Stochastic Modeling of Past Volcanic Crises

Gordon Woo*

RMS, London, United Kingdom

The statistical foundation of disaster risk analysis is past experience. From a scientific perspective, history is just one realization of what might have happened, given the randomness and chaotic dynamics of nature. Stochastic analysis of the past is an exploratory exercise in counterfactual history, considering alternative possible scenarios. In particular, the dynamic perturbations that might have transitioned a volcano from an unrest to an eruptive state need to be considered. The stochastic modeling of past volcanic crises leads to estimates of eruption probability that can illuminate historical volcanic crisis decisions. It can also inform future economic risk management decisions in regions where there has been some volcanic unrest, but no actual eruption for at least hundreds of years. Furthermore, the availability of a library of past eruption probabilities would provide benchmark support for estimates of eruption probability in future volcanic crises.

Keywords: volcanic, crisis, stochastic, modeling, counterfactual, Montserrat, Ontake

INTRODUCTION

Probabilistic methods are increasingly appreciated as valuable tools in applied volcanology, especially in interfacing between the geohazards posed by volcanoes, and the georisks to which society is exposed. One of the most important decisions to be made in respect of a volcanic region involves the timing of the evacuation of people at risk, and also the timing of their return home. Where the size of the population at risk and the local transport infrastructure allow for their movement away from danger in a matter of hours, civil authorities can leave a crucial decision until volcanologists are quite sure that a dangerous eruption will actually occur. However, one or both of these criteria may fail for a given volcano, and crucial public safety decisions may have to be made under considerable uncertainty. In such challenging circumstances, the expected safety benefits of any early evacuation or any delay in return home should be assessed as outweighing the expected displacement and disruption costs (Marzocchi and Woo, 2007; Woo, 2011).

In order to weigh as rationally as possible the expected safety benefits and disruption costs, the probability of an imminent dangerous eruption, within a relevant practical time period such as the next 3 months, needs to be estimated, along with the proportion of the population at risk who would owe their lives to a timely evacuation or delay in return home. The former needs to synthesize all the monitoring data that are available. The latter involves identifying the alternative geographical footprints of pyroclastic currents and lahars, and assessing how likely each of these possibilities are. This requires state-of-the-art stochastic volcanic flow modeling.

Volcanology is an observational science, and the maximum knowledge and information needs to be extracted from the record of the past. Even for volcanologists who specialize in studying the volcanoes of one country, the past is not as familiar as might be supposed. A principal objective of this paper is to motivate more intensive research into individual past volcanic crises, to understand

not just what happened, but what else could or might have happened. Stochastic modeling is traditionally considered as an exercise for the future, but there is merit also in stochastic modeling for the past. Although the focus is on past volcanic crises, this paper is not a review of past crises, but rather presents a new perspective on the treatment of past crises, advocating a stochastic analytical approach. A number of examples illustrate this approach which can be extended to other volcanoes.

For her PhD dissertation, Hincks (2007) developed a quantitative framework for probabilistic volcano hazard analysis, and instigated a retrospective hazard analysis of the volcanic crisis at La Soufrière, Guadeloupe in 1976, one of the most controversial because of the mass evacuation called. Bayesian techniques are well-suited to stochastic modeling of an uncertain past (Balke and Pearl, 1994), where there is sufficient public risk mitigation interest to justify the effort to gain this extra knowledge. A Bayesian Belief Network methodology was used to quantify the eruption probability over the duration of the Guadeloupe crisis, taking evidential account of all the monitoring data (Hincks et al., 2014). The network methodology is particularly flexible for multivariate data synthesis.

Alternatively, there are event tree methods for estimating the likelihood of an eruption. The event tree typically includes branches for the origin of the eruption, the style of eruption, location, magma composition, and eruption size. The Bayesian Event Tree method developed by Marzocchi et al. (2010) is implemented in readily available and usable software BET_EF. The USGS Volcano Disaster Assistance Program (Newhall and Pallister, 2015) makes use of event trees, which were first introduced by USGS staff volcanologists Newhall and Hoblitt (2002). Another approach, which has been applied to Colima, Piton de la Fournaise and Merapi, is to embed the classic deterministic failure forecast method within a Bayesian Framework (Boué et al., 2016). A limitation of this method, acknowledged by the authors, is the condition for an effective forecast that the pre-eruptive sequence is long enough for civil protection to organize and manage an evacuation.

The search for the underlying causes of a disaster is an exercise in stochastic forensics. The aim is to explore and discover how likely the alternatives are. Outside volcanology, an important international application was to the probabilistic assessment of the final resting place of the Malaysian Airlines MH370 passenger jet that went missing in March 2014. An entire book (Davey et al., 2016), summarizing a major costly research effort, has charted the elaborate use of sophisticated Bayesian methods for weaving together in a coherent manner the multiple thin strands of evidence from satellite, avionic and oceanographic sources. It is well-known that evidence can be used to update the prior probability that a future event has a particular characteristic. Less apparent is that evidence can be used to update the prior probability that a *past* event had a particular characteristic.

An analysis of the issuance of volcanic alert levels during volcanic crises (Winson et al., 2014) shows that almost 80% of eruptions can be broadly characterized as not being anticipated by a change of alert levels. Inadequate monitoring is often a factor, and alert levels may not be intended for use in forecasting, but clearly much remains to be learned about estimating the

likelihood of an imminent eruption from such monitoring data as exist. For most past volcanic crises, little is actually known about these likelihoods. In the minds of decision-makers, these are typically vague implicit figures or their verbal equivalents: minimal, tiny, very small, small, moderate etc.

From a volcano hazard perspective, stochastic modeling of past episodes of volcanic unrest would provide valuable additional insight, especially for volcanoes having a sparse eruption history. Moran et al. (2011) have observed that episodes not leading to eruption may be under-reported, or not reported at all, by local observatories. They have further noted that such a bias can result in incorrect estimation of probabilities that a given episode of unrest will lead to an eruption. Effort taken to assess eruption probabilities for the historical past results in valuable benchmarks being available for future volcanic unrest.

There is an inherent human outcome bias in reviewing past experience. As Kahneman (2011) has pointed out, decisions tend to be judged according to the outcome. If the losses are minimal, such as for an unrest period that did not result in an eruption, or for an eruption that caused few if any casualties, the file on an event may be closed without the perceived need nor even any justification for a risk management review. Yet there are many lessons to be learned in risk awareness and catastrophe risk management from asking searching counterfactual questions about historical events - not just extreme events, but also those that might be classified as near misses. From a scientist's perspective, history was not inevitable, but is just one possible realization of what might have happened. Indeed, in his discussion of the logic of counterfactuals in causal inference, Pearl (2009) has emphasized that counterfactuals carry as clear an empirical message as any scientific laws, and indeed are fundamental to them.

COUNTERFACTUAL ANALYSIS OF VOLCANIC UNREST

In his explanatory review of Bayesian network analysis, Pearl (2000) has conceived the following basic illustrative example of a two-state system, which is a paradigm for the analysis of volcanic unrest and the predictive role of precursory activity.

In this example, the output system response Y (e.g., eruption) is a binary function of another binary precursory event indicator X (e.g., presence of volcanic tremor), and a set of other variables U that may influence $Y : Y = f(X, U)$. U include the forcing factors, such as internal pressure, that could cause a dynamic perturbation to a volcanic system, and thence a state transition to eruptive activity.

Since X and Y are either 0 or 1, for any given U, the relationship between X and Y [which might be denoted as $Y = Y(X)$] must be one of only four binary functions:

$$f_0 : \{Y(0) = 0 \ \& \ Y(1) = 0\}; \quad f_2 : \{Y(0) = 1 \ \& \ Y(1) = 0\}$$
$$f_1 : \{Y(0) = 0 \ \& \ Y(1) = 1\}; \quad f_3 : \{Y(0) = 1 \ \& \ Y(1) = 1\};$$

As U varies, the only effect is to switch the (X, Y) relationship among these four functions. The probability $P(U)$ thus induces a probability function over the possible response pairs $\{Y(0), Y(1)\}$.

Volcanologists are interested in the particular subset U_E of U values for which $\{Y(0) = 0 \quad \& \quad Y(1) = 1\}$, i.e., where the observation of X happens to be a true accurate and reliable predictor of Y.

In order to make better decisions on the interpretation of precursory observations, it is desirable to make progress in identifying which combinations of variables U comprise U_E, and in estimating $P(U_E)$, the probability of being in U_E. The eruption probability is thus associated in a tangible way with the measure of a physical parameter set. These parameters are not directly observable, but may be modeled. U_E can be expressed in terms of ranges of dynamic model variables.

Pearl (2009) has credited the development of Bayesian networks as a stepping stone toward a more profound transition, from reasoning about beliefs to reasoning about causal and counterfactual relationships. The probability that volcanic unrest might lead to an eruption is more than just a degree of belief, it reflects a deep counterfactual logic. Consistent with scientific knowledge of the state of the volcano during the unrest, key dynamic forcing variables could have been in the range leading to an eruption. The likelihood that they are in this range provides a physical measure of the counterfactual eruption probability.

Case Study: Unrest at Soufrière Hills, Montserrat

As a practical case study, consider the Soufrière Hills volcano in Montserrat. Christopher et al. (2015) have interpreted the observations since 1995 as being consistent with a vertically extensive crustal magmatic mush beneath the volcano. They have categorized an unrest period as being one where layers of melt and magmatic fluids connect and move upward, while the igneous mush collapses downward. The melts layers amalgamate to form magma chambers. Decompression of melts and magmatic fluids results in crystallization, and increases in magma and fluid pressure. These increases during periods of unrest may not be sufficient to cause an eruption. But once an eruption is instigated, instability increases due to a positive dynamical feedback effect.

With this geophysical and geochemical understanding, U can be defined in a simplified manner as the magma and fluid pressures inside the volcano. Correspondingly, U_E would be a sub-domain of the space of these pressure variables which is sufficiently jointly extreme as to be associated with an eruptive state. During a period of unrest, the probability of an eruption then can be estimated from the probability that the magma and fluid pressures exceed the critical threshold. This regime is associated with major destabilization of the trans-crustal layered igneous system.

In volcano catalog of eruptions, it has not been routine to include periods of unrest. These are times when human or scientific observations have been made of some external signs of activity, other than eruptive activity itself. Observations of unrest might include volcanic tremor, gas emissions, and inflation of the flanks of the volcano. This incompleteness of information reflects the traditional perception that such data are of scientific interest, but rather inessential for hazard estimation, which is primarily dependent on the geological and historical time series of the actual eruption events themselves. However, for volcanoes with a low frequency of eruption, occasional periods of unrest may be important indicators of failed eruptions which should be taken into account in volcano hazard assessment.

Soufrière Hills in Montserrat is a prime example of a volcano which, until 1995, had not erupted since around 1630, yet had given rise to three periods of significant unrest over the previous century. These might be interpreted as failed attempts at eruption. In the authoritative Smithsonian Institution catalog of volcanoes of the world, the second edition of which was published shortly before the eruption in 1995 (Simkin and Siebert, 1994), the 1630 eruption is included as the sole entry for Montserrat—but none of the unrest history is listed. The first recorded European landings on Montserrat occurred in 1628 and 1631 with no reports of settlers, so it is not possible to undertake a retrospective study of the 1630 eruption akin to that by Sandri et al. (2009) for the 1631 eruption of Vesuvius.

However there were three distinct episodes of unrest in the 1890s, 1930s and 1960s. Counterfactually, each of these unrest periods might have led to an eruption. The 1930s unrest was marked by sporadic bursts of seismic activity, which led up to a M6.2 earthquake on 10 November 1935 (Powell, 1938). As it turned out, seismicity levels and gas emission rates decayed quite rapidly afterwards, suggesting that this major offshore earthquake had relieved some dangerous crustal stresses. But this sizeable earthquake might potentially have triggered an eruption through opening rock fractures (Kennedy, 2017). Remarkably, the spatial pattern of earthquake epicenters has some striking similarities with that at the start of the 1995 eruption. For the third of the unrest periods prior to 1995, Shepherd et al. (1971) noted that sharp increases in seismic and solfataric activity occurred in 1966, and these events indicated the abnormally high risk of an eruption in the near future. Magma had intruded into the upper crust beneath the volcano, and migrated upwards.

Probabilistic volcano hazard assessment should not be based solely on actual past eruptions, which may be a very sparse dataset, but should take proper counterfactual account of the periods of unrest, which constitute a vital element of the knowledge of the activity status of a volcano. For each of the three periods of unrest, the eruption probability $P(U_E)$ may be interpreted intuitively as the probability that the magma and fluid pressures exceeded a critical threshold.

Scenarios may be regarded as counterfactual histories of the future (Weber, 1996). Accordingly, counterfactual thought experiments are insightful for risk assessment. As a counterfactual thought experiment in volcanology, suppose that there was an ensemble of dynamical computational models of the Soufrière Hills volcano, each with its own alternative possible parameterization for a given period of unrest. Such an ensemble might be constructed using simple internal representations of the volcano dynamics and geometry. Approximate Bayesian computation methods, that do not require knowledge of an explicit likelihood function, could then be applied to estimate volcano risk. The concept of an ensemble of computational models is already realized in practice in another major branch of natural hazards, namely meteorology, where ensembles are used

to inform probabilistic weather forecasts, such as for hurricane risk.

Within the context of this counterfactual thought experiment, the following question can be posed for each of the three periods of unrest: what proportion of the dynamical volcano models would show that the magma and fluid pressures exceeded the critical threshold for an eruption? Even if this proportion were as low as one-fifth for each unrest, there would have been an even chance of some eruption in the century from 1890 to 1990. This last risk statement is quite robust against varying these proportions, and representing them as plausibility ranges of experts (e.g., [0.1, 0.3]) rather than specific point figures. Indeed, the 1930s unrest period on its own had a good chance of generating an eruption when the M6.2 earthquake occurred on 10 November 1935.

Part of the purpose of stochastic modeling of the past is to redress the common human behavioral trait of outcome bias. To ponder what would have happened if things had turned for the worse is called a downward counterfactual. By contrast, an upward counterfactual considers what would have happened if things had been better. Psychologists of counterfactual thinking (Roese, 1997) observe that upward counterfactual thoughts are much more common than downward ones. The fact that no eruption resulted from any of the 1890s, 1930s, and 1960s volcanic crises on Montserrat affected the perception of volcanic hazard on the island. Indeed, the elapse time since the last previous eruption in 1630 was sufficiently long for the volcano to be presumed by the public to be dormant. And any downward counterfactual thought about volcano hazard would have been blown away by Hurricane Hugo, which struck the island on 17 September 1989, rendering 90% of the population homeless.

But there are important practical differences in risk implication between the following two hazard statements: one factual, the other counterfactual. Both might have been made in 1990, several years before seismic activity on Montserrat started.

[A]   There has been no eruption for more than 350 years.
[B]   There is an even chance of an eruption having occurred in the past 100 years.

Statement [A] might be invoked to satisfy a simple historical exclusionary criterion, whereas statement [B] is worded for application to a medium-term quantitative risk assessment for a dormant volcano. Among the stakeholders who would find practical value in the latter statement are insurers, and thereby residential, commercial and industrial property owners as well as builders. Losses from the 1995 eruption took insurers by surprise. A typical return period for portfolio insurance risk management is 250 years, which is shorter than the 350+ year time interval back to the last eruption. However, statement [B] indicates that the Soufrière Hills volcano presented a risk at a return period of about 200 years. As such it would have been used for accumulation insurance risk management to protect against extreme portfolio losses. Beyond volcano risk, counterfactual disaster risk analysis has been shown to be useful to insurers for all kinds of perils (Woo, 2016). Increasingly,

as the underlying scientific and engineering basis for property catastrophe insurance has expanded, geoscientists have become more motivated to understand and address geohazard and georisk issues useful to insurers.

COUNTERFACTUAL CASUALTY FIGURES

The protection of people from being killed or injured in a volcanic eruption is a cornerstone of public policy in every volcanic region. Insights into societal vulnerability and protection over time are found in an analysis of the global historical fatalities record, which has been undertaken by Auker et al. (2013). Further investigation of this record might extend to accounting explicitly for the uncertainty in the underlying data, and to carrying out a stochastic simulation of the past human losses from volcanic eruptions.

Official statistics on deaths from natural disasters are prone to human errors of mis-reporting and mis-recording, some of which may be intentional rather than accidental. Statistics may be underestimated for fear of public revolt, or overestimated to encourage donations. The uncertainty in keeping track of deaths from a volcanic eruption is compounded by the obliterating impact of volcanic flows on the integrity and identity of human bodies. Secondary indirect impacts, such as famine, may be extensive and global for a great eruption like Tambora in 1815. Less well-known, even in Indonesia, than its devastating impact in Europe is the starvation in China, which exacted a very high death toll (D'Arcy Wood, 2014).

Quite apart from the numerical vagueness of the volcanic fatalities record, the number of deaths from a volcanic eruption is subject to a substantial degree of dynamic volatility, which should be recognized in volcano risk management. A correlation is already known to exist (Ambraseys and Bilham, 2011) between earthquake fatalities in a country, and the level of corruption, as gauged by the Transparency International index. Inevitably, given the common elements of geohazards and georisks, it should be expected that politics exert an influence on the death toll from volcanic activity.

Consider the issuance of an effective public hazard warning. Commensurate with scientific progress in understanding volcanoes, the lives saved through volcano warnings and evacuations around the world should increase progressively. A counterfactual statistical analysis of the lives-saved metric would be an interesting complement to the statistical analysis of casualties. Such analysis would make use of state-of-the-art methods in the modeling of volcanic flows (e.g., Bayarri et al., 2015). The human behavioral dimension to the volcano casualty toll is not recognized by analyzing the raw death tolls. Clearly, there are numerous socio-economic factors that affect the death toll from volcanic eruptions. In particular, there is an extraneous political dimension to hazard warnings which needs to be appreciated. This applies to hurricanes as well as volcanic eruptions. Indeed, Dyson (2007) likened the death toll in New Orleans from Hurricane Katrina in 2005 to that in Pompeii from Vesuvius in the year 79, with the underprivileged being least likely to evacuate.

For the greatest volcanic eruptions like Tambora 1815 and Krakatau 1883, a very high death toll can only be averted through maximally effective enforced evacuation procedures. The third highest volcanic death toll since 1600, that from the eruption of Mt. Pelée, Martinique, on 8 May 1902, might have been averted if the safety of citizens in St. Pierre had taken precedence over an imminent election on 11 May. The coincidence of the eruption with the week of the election was clearly bad luck. Counterfactually, there was a high chance that this would not have happened. In October of the same year, the Santa Maria volcano in Guatemala erupted, killing more than 8,000 people—the seventh highest death toll since 1600. At the time, the government had been preoccupied with a festival generating propaganda for the president.

Just as with the spread of pandemic disease, ongoing military conflict erodes the capability of a national government to take effective safety decisions in controlling a crisis. At the time of the catastrophic eruption and lahar at Nevado del Ruiz on 13 November 1985, the Colombian government and army were preoccupied by the guerrilla war from giving due attention to the mudflow that ultimately killed almost 25,000 of its citizens. Counterfactually, had the government been able to focus attention on the natural hazard adversary, rather than fighting on two fronts, many lives could have been saved.

Counterfactual thinking works both ways; more lives might have been lost or saved in an eruption than actually were registered. Human populations are in constant flux. The death toll will depend crucially on the endangered population at the time of eruption. Sometimes this will be lower than on average. The eruption of Mt. St. Helens on 18 May 1980 occurred at 8.32 am on a Sunday morning. At that time, there were few employees of the Weyerhauser timber company working. With millions of dollars in timber revenue at stake, local government had been pressurized by loggers to make the restricted areas smaller than recommended by the USGS. Counterfactually, the death toll could have been more than an order of magnitude larger than 57 if the eruption had occurred the following day, when thousands of Weyerhauser employees would have been at work on Monday morning. The chance of the triggering earthquake occurring early on a Sunday morning was only a few percent.

INTERNATIONAL LIBRARY OF ERUPTION PROBABILITIES

Quantitative disaster risk analysis is founded on scenarios for the future. The more uncertain the future is, the greater the number and broader diversity of disaster scenarios that need to be developed. But even with highly elaborate scenario analyses, events may yet occur that come as a partial or even complete surprise. This reflects a cognitive bias limiting human imagination beyond what has already been experienced. Steps to correcting this bias can be made by considering not just what has happened, but also what almost happened, or what might have happened before. This requires a research agenda in counterfactual historical disaster analysis, especially periods of unrest of apparently dormant volcanoes.

As a cause for human surprise at events, Weber (1996) cites the failure of divergent thinking: due weight needs to be given to the variety of pasts that might have occurred. He believes that counterfactual reasoning in the past and in the future have to be treated as essentially identical operations in logical terms. Just as there is stochastic modeling of the future, there should be stochastic modeling of the past. Yet there is a longstanding anthropocentric tendency to treat the historical past as fixed and determined. From an objective scientific perspective what happened is just one of numerous realizations of a complex stochastic process. The study of natural disasters is an observational rather than experimental discipline. To fill gaps in knowledge of rare extreme events, historical disasters can be used much more extensively as an available test laboratory for scenario discovery (Woo et al., 2017). Whilst seismological progress is being made in the development of earthquake likelihood models (e.g., Zechar et al., 2013), volcanic eruption likelihood models need an expanded database.

Past episodes of volcanic unrest constitute a valuable but under-investigated and under-utilized database for improving the reliability of estimates of eruption probability. Caldera eruptions are notoriously difficult to forecast reliably. The Santorini unrest of 2011–2012 gave rise to major concerns over public safety as well as Greek economic loss of tourism. A Bayesian Belief Network analysis, combining multiple strands of scientific and observational evidence, was developed (Aspinall and Woo, 2014) to support UK government response planning. One of the practical conclusions was that, ideally, such hazard and risk assessments should be elaborated in detail and critiqued well before crisis-level unrest develops—not initiated and implemented within a few hours just when a situation looks ominous. In particular, careful analysis of all information is required to determine and represent parameter uncertainties comprehensively and dependably.

Risk communication is especially challenging for caldera regions where a significant proportion of the local population may be unaware of the volcanic risk. Following four strong magnitude 6 shocks in May 1980, USGS detected dome-like swelling in the middle of Long Valley Caldera, California. In response to these signs of renewed volcanic unrest, USGS intensified their monitoring of the Long Valley area. Measurements showed that the center of the caldera had risen almost a foot since the summer of 1979, after decades of stability. Modeling of the process of new magma rising beneath the caldera could inform real-time approximate estimates of eruption probability during the period of active deformation. USGS has pointed out that unrest can temporarily increase the odds of an eruption, depending on the nature, intensity, and location of the unrest. Tracking the hazard variation over time would be particularly useful for consistent and informed decision-making on public warnings.

This would apply as well to another caldera, Campi Flegrei, Italy. Hazard mapping there is actively being researched (Bevilacqua et al., 2017). A retrospective study of the 1538 Monte Nuovo eruption (Di Vito et al., 1987) provides historical insight into the sequence of events leading up to this notable previous eruption. Twenty shocks were felt at Pozzuoli and Naples on

the day and in the night before the eruption. Frequent local earthquakes had in fact been felt in the preceding 2 years. As with the 1631 eruption of Vesuvius referenced below, it would be insightful if some approximate estimate of eruption probability were assigned to this historical eruption in the Naples region. Specifically, such an exercise would provide a useful quantitative hazard benchmark for current and future unrest at Campi Flegrei.

Dealing with unrest at any volcano situated in a densely populated region is of course much more than a scientific issue, it is a complex risk management challenge. In this georisk context, volcanologists should heed the wisdom of the inventor and philosopher of modern business management, Peter Drucker (Lavinsky, 2017): *"If you can't measure it, you can't improve it."* To make improvements in volcano safety, additional monitoring is undertaken, and the data should be converted to risk measures. To measure volcanic risk, a library of probability estimates for notable past episodes of volcanic unrest needs to be developed to help calibrate and validate the quantitative procedures for hazard estimation which may be adopted in future volcanic crises. Many public and private stakeholders would benefit from a research program in stochastic modeling of past volcanic crises. Such stakeholders include civil authorities, infrastructure planners, and insurers.

Implementation within WOVOdat

The World Organization of Volcano Observatories (WOVO) oversees a comprehensive global database on volcanic unrest, WOVOdat, aimed at understanding pre-eruptive processes and improving eruption forecasts. WOVOdat is presently hosted at the Earth Observatory of Singapore (www.wovodat.org). A lack of standardization in data formats and database architectures has frustrated comparative studies of volcanic unrest, and searches for analog to any current unrest. WOVOdat translates and compiles this data into common formats with the goal of making them freely web-accessible, for reference during volcanic crises, comparative studies, and basic research on pre-eruptive processes. WOVOdat is intended for reference during volcanic crises, comparative studies, basic research on pre-eruption processes, teaching, and outreach (Newhall et al., 2017).

Another novel application of WOVOdat is to provide data for a numerical eruption simulator. Airline pilots train and prepare for dangerous situations by spending time on a numerical flight simulator. Observatory volcanologists could also train and prepare for dangerous situations by spending time on a numerical eruption simulator. The International Air Transport Association (IATA), which is a major stakeholder in volcano risk mitigation, would appreciate the purpose and value of such training. The stochastic modeling of past volcanic crises would provide much needed input for such a simulator.

The eruption of Eyjafjallajökull, Iceland, in April 2010, caused massive aviation disruption in northern Europe. Analysis of the seismic energy released around its big neighbor Katla provided evidence of a rising intrusive magma body at Godabunga on the western flank of the volcano, which had erupted in October 1918. Some preliminary analysis of the likelihood of a triggered eruption of Katla was undertaken at the time to advise the UK

government, but a more elaborate and extensive retrospective study would be very useful as a comparative benchmark for future hazard levels.

As observed in the WOVOdat Developer Team (2014), WOVOdat is to volcanology as epidemiological databases are to medicine, providing valuable tools for research and crisis response. Scientists needing to forecast the outcome of a fresh volcanic crisis will be able to search for analog, find the past outcomes, and estimate probabilities of how the fresh unrest will evolve. In order to benchmark these probabilities for current periods of unrest, it will be instructive to estimate eruption probabilities for the more significant periods of unrest in the past. Estimating the likelihood of eruption for past periods of significant unrest is a preparedness exercise for testing probabilistic volcano hazard analysis:

- It can help to validate and check hazard analysis procedures. The methodology for hazard analysis can be tested by benchmarking against past experience.
- It can improve the analytical tools needed and streamline the computational procedures to be followed in the future. Approximate Bayesian computation is one of these other procedures.
- It can illuminate past volcanic crisis evacuation decisions. The skill level may be judged using an expected loss consequence metric, as in meteorology. This involves assigning losses to the four outcomes: [evacuation, no eruption]; [evacuation, eruption]; [no evacuation, no eruption] and [no evacuation, eruption]. The biggest loss occurs when an eruption occurs but there is no evacuation, and a proportion of the prevailing population at risk become casualties. Accordingly, a low eruption probability in this eventuality would incur a high loss penalty. By contrast, zero loss occurs when no eruption occurs, and there is no evacuation.
- It can inform future probabilistic volcanic hazard assessments. The more that can be learned about past probability estimates, the more confidence can be placed in future probability estimates.

Comparing Current and Past Eruption Probabilities

In respect of the last point, suppose that in some future unrest at a specific volcano, the eruption probability is estimated to be *P*. Within the context of population safety, the threshold size of an eruption is one that would threaten the local population, and the time window would typically be a few months. A useful benchmark and sense check of this estimate would be a quantitative comparison with other previous periods of unrest. If eruption probabilities have been calculated for past unrests, this comparison can made more easily and transparently, and with greater confidence. In the absence of such probabilities, a comparison with the past would have to be made on a rather ambiguous, ill-defined, and weakly constrained multi-variate ranking basis; e.g., the tremor now is stronger than in previous unrest X, but the CO_2 flux is weaker and the tilt is smaller etc.. Such is the indeterminate ambiguity in multi-variate ranking,

that some plausible ranking process could then be found to support any eventual decision.

To date, some retrospective studies have been undertaken on estimating eruption probabilities for actual volcanic disasters. Sandri et al. (2009) researched the primary literature sources, including Latin documents, to perform a Bayesian event tree analysis for the 1631 eruption of Vesuvius, which killed between 3,000 and 6,000 inhabitants. The original historical documents (Guidoboni, 2008) were especially needed for the reconstruction of the chronology of events leading up to the eruption. Sandri et al. (2009) point out that the historical information collected likely represents a subset of all the phenomena that occurred in the 1631 pre-eruptive phase. Hence, their results are a lower limit on the eruption forecasting ability of the software of BET_EF.

The 1631 pre-eruptive phase showed a diverse range of signals including seismic activity, smoke emissions and ground uplift. These precursors were able to increase the absolute probability of eruption up to 10% about a month before the beginning of the eruption; this probability increased to more than 30%, 7–10 days before the onset of the event, when a distinct unrest phase was established. This is an important finding, since the latter probability level counterfactually should have been sufficient to justify efforts at evacuation, thus mitigating the loss of life. The fact that these probabilities have already been estimated for the 1631 eruption will provide a useful benchmark supporting decision-making during a future volcanic crisis at Vesuvius. In particular, recognizing the hundreds of thousands who may not be officially listed as living on the slopes of Vesuvius, protracted population evacuation times may require that the evacuation process is instigated well before the eruption probability attains a high value, as in 1631.

Case Study: Mt. Ontake, Japan

To appreciate how the estimation of comparative eruption probabilities can assist in operational volcano hazard forecasting, consider Mt. Ontake in Japan. There was an eruption in late March 2007, which was preceded by a substantial amount of precursory activity. In December 2006, there was inflation of the volcanic edifice, and an increase in shallow seismicity below the summit. During January 2007, numerous earthquakes were detected, and low-frequency tremor was observed. It is clear that, in early 2007, there was a substantial chance of an eruption occurring, although this was not quantified. An aspiration for WOVOdat is that estimates of eruption probability such as this would be assigned, perhaps facilitated by the elicitation of structured expert judgement (Aspinall and Cooke, 1998).

Fortunately, there were no casualties in the eruption of March 2007. By contrast, at least 58 hikers were killed when Ontake erupted again at 11.52 a.m. on Saturday 27 September 2014. This eruption was preceded by little precursory activity. There was some unusual seismic activity detected on 10-11 September 2014, but no volcanic tremor or crustal deformation was detected. The main reason the JMA chose not to raise the volcano warning level at Ontake in September 2014 was that the number of low-frequency earthquakes was much fewer than observed with the March 2007 eruption (Yamaoka et al., 2016). At the time, the base warning level 1 corresponded to "Normal." This is

now clarified as "Potential for Increased Activity." In January 2017, legal proceedings against the JMA were instigated by some bereaved families arguing that the JMA should have raised the alert level (Bretton and Aspinall, 2017).

An interesting question that arises in the aftermath of this worst Japanese volcanic disaster since 1926 is what eruption probability level might have warranted a more assertive JMA response than issuing volcano observation information to nearby local governments. According to the Japan Times (2014), the problem was that nobody was very interested in such information. A translation is needed from volcanic unrest information to individual risk mitigation action. This translation is provided by economic cost-benefit analysis (Woo, 2014). Implicit in the mind of any decision-maker is the trade-off between the cost to the public of volcano access restriction, and the safety benefit of avoiding fatalities. Access is important: the need to listen to public access demands was articulated by the head of research at Japan's National Research Institute for Earth Science and Disaster Prevention (Cyranoski, 2014).

Suppose access to the crater on Saturday, 27 September 2014, had been denied through raising the volcano warning level from the level of 1 (Normal) to 2 (do not Approach the Crater). This was the best time of the year to see the autumnal colors on the sacred mountain. The loss of enjoyment to hikers would have had a significant economic cost that might have been found by allowing access only on payment of a substantial entry fee to cover the expense of emergency rescue services. The previous several weekends had been rainy, so to deter entry this fee would have had to be of the order of 100 US dollars, which is the approximate cost of a day trip at a substitute resort location. According to Holtz (2013), the Japanese value of a statistical life for cost-benefit analysis is as high as $10 million. This is consistent with the equivalent figure in USA (Butry et al., 2007). The ratio of these amounts is then $100/10,000,000 = 1/100,000$.

Writing the daily chance of an eruption as P_D, and the likelihood that a hiker on the mountain would be killed by such an eruption as H, then a change in alert level might have been warranted if $P_D > 1/(100,000 * H)$. Given the spatial footprint of the eruption, about one-fifth of the 300 mountain hikers were killed. Furthermore, the afternoon time window of hiker exposure to an eruption on 27 September is about one-fifth of a 24 h day. Accordingly, $H \approx 1/25$, so the criterion becomes $P_D > 1/4000$. Expressed in terms of a reference benchmark time window of 3 months, this is equivalent to a threshold eruption probability of about 2.3%. In qualitative language, this threshold corresponds to the hazard being low.

This shows that the decision to consider changing the volcano warning level in September 2014 was quite finely balanced. Given that at least 58 died, the decision had to have been close. If the actual hazard level were perceived to be well below the threshold, the criterion for changing the warning level would have been inadequate. Comparison with the large number of low-frequency earthquakes preceding the March 2007 eruption was a key factor in the decision not to raise the warning level. Clearly, additional decision support would have been valuable.

Had the 3-month eruption probability in late March 2007 been previously estimated, this would have provided an additional

informative quantitative benchmark to inform the judgement of decision-makers in September 2014, when the level of seismic activity was much lower. A sense of relative risk is helpful as a guide. A 3-month eruption probability in late March 2007 below 10% would have given authorities greater confidence in not raising the warning level in September 2014.

A higher figure might still have been compatible with not raising the warning level, according to the following reasoning. Maintaining the warning level at "Normal" would correspond implicitly to assuming that hikers were prepared to pay a higher personal cost to access the crater on 27 September 2014 to enjoy the Autumnal foliage. This socio-economic dimension of volcano risk merits investigation prior to any developing crisis. Sociological surveys could be undertaken to elicit public preferences relating to volcano access.

This suggestion applies to all active volcanoes, in USA, New Zealand etc., popular with hikers, tourists and other visitors. In situations, such as Ontake on 27 September 2014, where the hazard level is low, but not negligible, an access fee could be charged commensurate with the elicited public preferences. The access fee acts as a "nudge" (Thaler and Sunstein, 2008),

discouraging those who would be quite happy to spend their time in ways other than mountaineering. The access fee explicitly captures the value to the individual of spending time on the volcano. In a cost-benefit analysis, this value to the individual is weighed against the small chance of becoming a casualty in a sudden eruption.

AUTHOR CONTRIBUTORS

The author confirms being the sole contributor of this work and approved it for publication.

FUNDING

ACKNOWLEDGMENTS

REFERENCES

Ambraseys, N., and Bilham, R. (2011). Corruption kills. *Nature* 469, 153–155. doi: 10.1038/469153a

Aspinall, W. P., and Cooke, R. M. (1998). "Expert judgement and the Montserrat volcano eruption," in *PSAM4*, eds A. Mosleh and R. A. Bari (New York, NY: Springer), 2223–2118.

Aspinall, W. P., and Woo, G. (2014). Santorini unrest 2011-2012: an immediate Bayesian belief network analysis of eruption scenario probabilities for urgent decision support under uncertainty. *J. Appl. Volcanol.* 3:12. doi: 10.1186/s13617-014-0012-8

Auker, M. R., Sparks, R. S. J., Siebert, L., Crosweller, H. S., and Ewert, J. (2013). A statistical analysis of the global historical record. *J. Appl. Volcanol.* 2:2. doi: 10.1186/2191-5040-2-2

Balke, A., and Pearl, J. (1994). "Counterfactual probabilities: computational methods, bounds, and applications," in *Uncertainty in Artificial Intelligence*, eds R. Lopez de Mantara and D. Poole (San Francisco, CA: Morgan Kaufmann), 46–54.

Bayarri, M. J., Berger, J. O., Calder, E. S., Patra, A. K., Pitman, E. B., Spiller, E. T., et al. (2015). Probabilistic quantification of hazards: a methodology using small ensembles of physics-based simulations and statistical surrogates. *Int. J. Uncert. Quant.* 5, 297–325. doi: 10.1615/Int.J.UncertaintyQuantification.2015011451

Bevilacqua, A., Neri, A., Bisson, M., Esposito Ongaro, T., Flandoli, F., Isaia, R., et al. (2017). The effects of vent location, event scale, and time forecasts on pyroclastic density current hazard maps at Campi Flegrei caldera (Italy). *Front. Earth Sci.* 5:72. doi: 10.3389/feart.2017.00072

Boué, A., Lesage, P., Cortés, G., Valette, B., Reyes-Dávila, G., Arámbula-Mendoza, R., et al. (2016). Performance of the material Failure Forecast Method in real-time situations: a Bayesian approach applied on effusive and explosive eruptions. *J. Volc. Geotherm. Res.* 327, 622–633. doi: 10.1016/j.jvolgeores.2016.10.002

Bretton, R., and Aspinall, W. P. (2017). Risk assessments face legal scrutiny. *Nature* 350:188. doi: 10.1038/550188b

Butry, D. T., Brown, M. H., and Fuller, S. K. (2007). *Benefit-Cost Analysis of Residential Fire- Sprinkler Systems*. NIST Office of Applied Economics Report 7451.

Christopher, T. E., Blundy, J., Cashman, K., Cole, P., Edmonds, M., Smith, P. J., et al. (2015). Crustal-scale degassing due to magma system destabilization and magma- gas decoupling at Soufrière Hills Volcano, Montserrat. *Geochem. Geophy. Geosyst.* 16, 2797–2811. doi: 10.1002/2015GC005791

Cyranoski, D. (2014). Why japan missed volcano's warning signs. *Nature News: Explainer.* London.

D'Arcy Wood, G. (2014). *Tambora*. Princeton, NJ: Princeton University Press.

Davey, S., Gordon, N., Holland, I., Rutten, M., and Williams, J. (2016). *Bayesian Methods in the Search for MH370*. London: SpringerOpen.

Di Vito, M., Lirer, L., Mastrolorenzo, G., and Rolandi, G. (1987). The 1538 Monte Nuovo eruption (Campi Flegrei, Italy). *Bull. Volcanol.* 49, 608–615. doi: 10.1007/BF01079966

Dyson, M. E. (2007). *Come Hell or High Water: Hurricane Katrina and the Color of Disaster*. New York, NY: Basic Civitas books.

Guidoboni (2008). Vesuvius: a historical approach to the 1631 Eruption – "cold data" from the analysis of three contemporary treatises. *J. Volcano Geother. Res.* 178, 347–358. doi: 10.1016/j.jvolgeores.2008.09.020

Hincks, T. K. (2007). *Probabilistic Volcanic Hazard and Risk Assessment*. Ph.D. thesis, University of Bristol.

Hincks, T. K., Komorowski, J.-C., Sparks, R. S. J., and Aspinall, W. P. (2014). Retrospective analysis of uncertain eruption precursors at La Soufrière volcano, Guadeloupe, 1975-1977: volcanic hazard using a Bayesian Belief Network approach. *J. Appl. Volcanol.* 3:3. doi: 10.1186/2191-5040-3-3

Holtz, H. (2013). *Global Health Care*. Burlington, MA: Jones & Bartlett Learning.

Japan Times (2014). *The Mount Ontake Eruption*. Tokyo.

Kahneman, D. (2011). *Thinking, Fast and Slow*. London: Allen Lane.

Kennedy, B. (2017). What effects to earthquakes have on volcanoes? *Geology* 45, 765–766. doi: 10.1130/focus0820172.1

Lavinsky, D. (2017). *The Two Most Important Quotes in Business*. Available online at: www.growthink.com

Marzocchi, W., Sandra, L., and Selva, J. (2010). BET-VH: a probabilistic tool for long-term volcanic hazard assessment. *Bull. Volc.* 72, 705–716. doi: 10.1007/s00445-010-0357-8

Marzocchi, W., and Woo, G. (2007). Probabilistic eruption forecasting and the call for evacuation. *Geophys. Res. Lett.* 34:L22310. doi: 10.1029/2007GL031922

Moran, S. C., Newhall, C., and Moran, D. C. (2011). Failed magmatic eruptions: late-stage cessation of magma ascent. *Bull. Volc.* 73, 115–122. doi: 10.1007/s00445-010-0444-x

Newhall, C. G., Costa, F., Ratdomopurbo, A., Venezky, D. Y., Widiwijayanti, C., Win, N. T. Z., et al. (2017). WOVOdat – an online, growing library of worldwide volcanic unrest. *J. Volc. Geotherm. Res.* 345, 184–199. doi: 10.1016/j.jvolgeores.2017.08.003

Newhall, C. G., and Hoblitt, R. P. (2002). Constructing event trees for volcanic crises. *Bull. Volc.* 64, 3–20. doi: 10.1007/s004450100173

Newhall, C. G., and Pallister, J. S. (2015). "Using multiple data sets to populate volcanic event trees," in *Volcano Hazards, Risks and Disasters,* eds J. Shroder and P. Papale (Amsterdam: Elsevier).

Pearl, J. (2000). The logic of counterfactuals in causal inference. *J. Am. Stat. Assoc.* 95, 428–435.

Pearl, J. (2009). *Causality: Models, Reasoning and Inference.* Cambridge University Press.

Powell, C. F. (1938). The royal society expedition to Montserrat, B.W.I. final report. *Philos. Trans. R. Soc. A* 237, 1–34. doi: 10.1098/rsta.1938.0002

Roese, N. J. (1997). Counterfactual thinking. *Psychol. Bull.* 121, 131–148. doi: 10.1037/0033-2909.121.1.133

Sandri, L., Guidoboni, E., Marzocchi, W., and Selva, J. (2009). Bayesian event tree for eruption forecasting (BET_EF) at vesuvius, Italy: a retrospective forward application to the 1631 eruption. *Bull. Volcanol.* 71, 729–745. doi: 10.1007/s00445-008-0261-7

Shepherd, J. B., Tomblin, J. F., and Woo, D. A. (1971). Volcano-seismic crisis in Montserrat, West Indies, 1966-1967. *Bull. Volcanol.* 35, 143–162. doi: 10.1007/BF02596813

Simkin, T., and Siebert, L. (1994). *Volcanoes of the World, 2nd Edn.* Washington, DC: Smithsonian Institution, Geoscience Press.

Thaler, R. H., and Sunstein, C. R. (2008). *Nudge.* New Haven, CT: Yale University Press.

Weber, S. (1996). "*Counterfactuals, past and* future," in *Counterfactual Thought Experiments in Politics,* eds P. E. Tetlock and A. Belkin (Princeton, NJ: Princeton University Press), 268–290.

Winson, A. E. G., Costa, F., Newhall, C. G., and Woo, G. (2014). An analysis of the issuance of volcanic alert levels during volcanic crises. *J. Appl. Volcanol.* 3:14. doi: 10.1186/s13617-014-0014-6

Woo, G. (2011). *Calculating Catastrophe.* London: Imperial College Press.

Woo, G. (2014). "Cost-benefit analysis in volcano risk," in *Volcanic Hazards, Risks and Disasters,* eds J. Shroder and P. Papale (Amsterdam: Elsevier), 289–300.

Woo, G. (2016). *Counterfactual Disaster Risk Analysis. Variance Journal.* Arlington, VA: Casualty Actuarial Society.

Woo, G., Maynard, T., and Seria, J. (2017). *Reimagining History*: Counterfactual *Risk Analysis.* Lloyd's Emerging Risk Report. Lloyd's, London.

WOVOdat Developer Team (2014). *VOVOdat Manual 1.1.* Available online at: www.wovodat.org

Yamaoka, K., Geshi, N., Hashimoto, T., Ingebritsen, S. E., and Oikawa, T. (2016). Special issue 'The Phreatic Eruption of Mt. Ontake Volcano in 2014'. *Earth Planets Space* 68, 175–182. doi: 10.1186/s40623-016-0548-4

Zechar, J. D., Schorlemmer, D., Werner, M. J., Gerstenberger, M. C., Rhoades, D. A., and Jordan, T. H. (2013). Regional earthquake likelihood models 1: first-order results. *Bull. Seism. Soc. Am.* 103, 787–798. doi: 10.1785/0120 120186