Bull Volcanol
DOI 10.1007/s00445-012-0633-x

REVIEW ARTICLE

Probabilistic eruption forecasting at short and long time scales

Warner Marzocchi · Mark S. Bebbington

Received: 4 March 2012 / Accepted: 16 June 2012

Abstract Any effective volcanic risk mitigation strategy requires a scientific assessment of the future evolution of a volcanic system and its eruptive behavior. Some consider the onus should be on volcanologists to provide simple but emphatic deterministic forecasts. This traditional way of thinking, however, does not deal with the implications of inherent uncertainties, both aleatoric and epistemic, that are inevitably present in observations, monitoring data, and interpretation of any natural system. In contrast to deterministic predictions, probabilistic eruption forecasting attempts to quantify these inherent uncertainties utilizing all available information to the extent that it can be relied upon and is informative. As with many other natural hazards, probabilistic eruption forecasting is becoming established as the primary scientific basis for planning rational risk mitigation actions: at short-term (hours to weeks or months), it allows decision-makers to prioritize actions in a crisis; and at long-term (years to decades), it is the basic component for land use and emergency planning. Probabilistic eruption forecasting consists of estimating the probability of an eruption event and where it sits in a complex multidimensional time–space–magnitude framework. In this review, we discuss the key developments and features of models that have been used to address the problem.

Keywords Probabilistic eruption forecasting · Precursors · Stochastic modeling · Time-space-magnitude forecasts

Editorial responsibility: S. Self

W. Marzocchi (✉)
Istituto Nazionale di Geofisica e Vulcanologia,
Via di Vigna Murata 605, 00143 Rome, Italy
e-mail: warner.marzocchi@ingv.it

M. S. Bebbington
Volcanic Risk Solutions, Massey University,
Private Bag 11222, Palmerston North 4442, New Zealand
e-mail: m.bebbington@massey.ac.nz

Introduction

Volcanic eruptions can have a huge impact on society, in terms of human losses, environmental consequences, and economic costs. Moreover, the ever-increasing exposure and vulnerability of society are going to continuously amplify the overall risk from such events, leading to greater demand for effective risk mitigation measures. The primary component for planning sound risk mitigation actions at different time scales is to forecast when, where, and how big future eruptions will be. This scientific goal is complicated by the very high number of degrees of freedom, often nonlinearly coupled, that characterize the physical processes underlying a volcanic eruption. This inherent complexity and the large uncertainty in the knowledge of these processes lead to the practical impossibility of predicting deterministically, or even with a small uncertainty, the onset time, location, and size of the impending eruption (e.g., Marzocchi et al. 1997), with rare exceptions (e.g., the dilatometer-based prediction of the 1991 Hekla eruption, Linde et al. 1993). In practice, volcanologists can track the evolution of a volcanic system only in a probabilistic way, for instance, identifying different scenarios, each one with their own probability of occurrence.

For many years, the use of probabilistic models in volcanology was almost totally limited to long-term assessment (e.g., Wickman 1966a) with rare investigation of short-term forecasts (e.g., Decker 1986). While scientists accept the long-term unpredictability of a volcanic system, they appear much more reluctant to live with the fact that uncertainties may play a major role also in short-term assessment. De facto, the deterministic approach is still more or less explicitly adopted by many volcanologists, who consider monitoring data reliable enough to predict the evolution of a phase of unrest with a reasonable certainty. The reluctance to abandon this approach is underlain by both a training for determinism in classical physics and the more pragmatic reason that society and decision-makers usually press scientists to provide advice without the uncertainties that the decision-makers often do not know how to deal with. It is notable that this inability of decision-makers to manage scientific uncertainty in short-term forecasts is probably the weakest point of the whole decision-making process, not only in volcanology (Marzocchi and Woo 2009), but in seismology (Jordan et al. 2011) and with weather-related hazards. Conversely, at longer time scales, the use of probabilistic seismic hazard analysis is routine in critical facility design, building codes, and planning processes.

From a more scientific point of view, the deterministic and probabilistic approaches are often seen as two opposite and irreconcilable aspects of how we describe nature. We disagree with this dichotomic view; any probabilistic model that works well must incorporate the most relevant monitoring anomalies and/or pertinent physics. In this view, the only difference between deterministic and probabilistic approaches is that the latter incorporates the uncertainties, rather than ignoring them (cf., Jeffreys 1961). The lack of deterministic predictions should not be seen as a scientific failure but as a rational approach to the nature of the problem. Probabilistic forecasts are commonly used in several fields to justify mitigation actions, and even in our private life, we take daily decisions without being sure about consequences and the future possibilities. This means that, in order to make sound decisions to reduce volcanic risk, decision-makers have to be trained with the concept of decision-making under uncertainty (e.g., Marzocchi and Woo 2007, 2009; Woo 2008), and scientists have to do their best to estimate probabilities.

In the following sections, we describe in detail the current state of the art concerning short- and long-term eruption forecasting. For the purpose of this paper, probabilistic eruption forecasting is a generic term that embraces any kind of probabilistic assessment related to the occurrence of an eruption that can be expressed as the probability of occurrence of an eruption onset in a given time–space–size interval. The time interval of interest may vary from hours/days to decades depending on the type of mitigation actions envisaged by the forecast; here, we use "short-term" to indicate a forecast in a time horizon of hours/weeks or months, typically of interest in managing evolving episodes of volcanic unrest; and "long-term" for time windows of years to decades that are required for land use and evacuation planning. This distinction is a consequence of more than just the desired usage. Conceptually, any forecasting model can be applied to a wide range of forecasting time intervals; nonetheless, the kind of information used for short- and long-term forecasts are basically different: short-term forecasting is mostly driven by monitoring information, while long-term forecasts are primarily based on the past activity of the volcano (Decker 1986).

Quantitative probabilistic assessment

The term "probability" has diffused throughout the scientific literature, but its meaning is still very controversial and has been hotly debated for centuries (e.g., Gillies 2000). Generally speaking, we can identify two broad classes of probability definition: the *frequentist* (*objective* probability) and the *degree of belief* (*subjective* probability). In the objective interpretation, the probability represents the expected long-term frequency of the event that we are considering. In the subjective interpretation, the probability is no longer an expected frequency, but it represents the degree of belief about the occurrence of the *next* event. This distinction has many important philosophical and practical aspects that we do not discuss in this paper (see Gillies 2000; and Marzocchi and Zechar 2011 for a discussion in the context of earthquake forecasting). Here, we simply note that both views have important merits and that different types of available information may suggest the use of one rather than the other.

An estimate of objective probability can be obtained through a stochastic model (e.g., Cox and Lewis 1966) or empirical analysis. Stochastic modeling is widely used in long-term eruption forecasting (Bebbington 2009), where the primary information comes from historical and geological catalogs of eruption onsets and, in some cases, sizes. Eruptions may be considered as the outcome of stochastic point processes. This allows volcanologists to benefit from a well-developed theory of point-process models that has been developed for reliability analysis, such as, for example, aircraft component failures (Proschan 1963). The empirical approach simply calculates the frequency of past events, assuming that *exactly* the same frequency holds also for the future. Statistically, empirical analysis is "nonparametric," implying the absence of a model, and faithfully reproduces the random variation from the catalog in forecasts. Unless the catalog is a large one, unlikely in volcanology, the resulting forecasts tend not to be smooth and can be biased.

Subjective probability is estimated in a different way. The best procedure (Cooke 1991; Gillies 2000) is through the formalization of the degree of belief of a group of scientists (intersubjective probability). A description of the merits of this definition is beyond the scope of the present paper. Nonetheless, we emphasize that the degree of belief of a group of experts usually tends to be much more coherent than the degree of belief of one single researcher. Moreover, a group of experts evaluate the epistemic uncertainty from multiple perspectives, which increases the likelihood of considering a fuller range of information, improving the epistemic evaluation. Methods of eliciting a degree of belief from a group of experts is an active field of research. The most accepted procedure is probably the Delphi method (e.g., Cooke 1991) that has been used in volcanology (e.g., Aspinall 2006; Neri et al. 2008; Selva et al. 2010a). The Delphi method relies on a structured panel of experts where information is fed back in summary form, allowing the panel to discuss and revise assessments several times; opinions are usually kept anonymous in order to leave any researcher completely free from external conditions. Sometimes, the experts' opinions are weighted in order to give greater weight to the opinion of the "best" experts but, at the same time, to down-weight experts that are clearly overopinionated (see Cooke 1991). Despite the widespread use of this approach, it is worth mentioning that important initiatives in other fields, such as the Senior Seismic Hazard Assessment Committee (SSHAC; Budnitz et al. 1997), the Uniform California Earthquake Rupture Forecast (Field et al. 2007), and the Intergovernmental Panel on Climate Change (Solomon et al. 2007) adopt different ad hoc schemes to include expert opinions in their forecasts.

In a broad sense, the distinction between objective and subjective probability is of interest for scientists and practitioners because these probabilities correspond to two different kinds of uncertainty. Objective probability is expected to describe the inherent unpredictability of the system or the so-called aleatory uncertainty. Subjective probability is more closely linked to the epistemic uncertainty due to the imperfect knowledge of the system under study. The distinction between aleatory and epistemic uncertainties is not straightforward, and it may hide some conceptual misunderstanding. Nonetheless, we argue that this distinction is helpful in some way to scientists, because it marks a separation between the intrinsic randomness (or variability) of the system (aleatory uncertainty) that is irreducible and the epistemic uncertainty that may be reduced as new information, models, or data become available.

In probabilistic seismic hazard assessment, as well as in many other kinds of hazard assessment, scientists incorporate both kinds of uncertainty through the concept of the logic tree (e.g., Budnitz et al. 1997) . A logic tree is a branching structure where the different branches represent all relevant sources of epistemic uncertainty for hazard assessment. A single hazard curve, corresponding to an individual branch of the logic tree, quantifies the aleatory aspects of the corresponding model, while the spread of hazard curves for different values of the ground-motion parameter of interest is determined by epistemic uncertainty. Therefore, the distribution corresponding to the full suite of hazard curves captures both aleatory and epistemic uncertainties (Budnitz et al. 1997; Bommer and Scherbaum 2008).

In volcanic hazard assessment and eruption forecasting, aleatory and epistemic uncertainties can be incorporated in a structured form, although it is still not the norm. One approach that has been used is the event tree concept (Newhall and Hoblitt 2002) in a Bayesian framework (Marzocchi et al. 2004, 2008, 2010; Neri et al. 2008; Marti et al. 2008; Sobradelo and Marti 2010). Basically, the event tree is a graphical tree representation of events in which individual branches are alternative steps from a general prior event, state, or condition, through increasingly specific subsequent events (intermediate outcomes) to final outcomes. In this way, the scheme shows all relevant possible outcomes of volcanic unrest at progressively higher degrees of detail. The probability of each outcome is obtained by combining the probabilities at each node of the tree through classical probability theory (Fig. 1). With respect to the logic tree, the event tree does not directly incorporate aleatory and epistemic uncertainties; both kinds of uncertainty are included in the event tree through the use of Bayesian inference in calculating the probability at each node.

In the event tree, the probability at each node is described by a beta distribution, with density function

$$f_{\Theta}(\theta) = \frac{1}{B(\alpha, \beta)} \theta^{\alpha-1} (1-\theta)^{\beta-1}, \quad 0 < \theta < 1, \ \alpha > 0, \ \beta > 0, \tag{1}$$

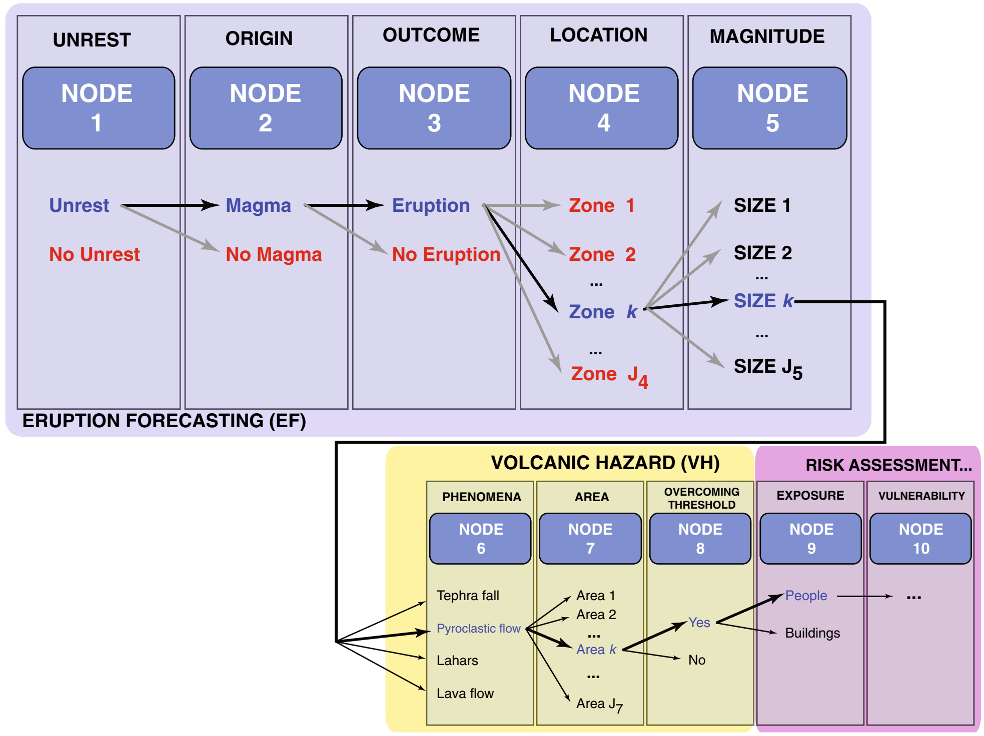

Fig. 1 Event tree scheme for eruption forecasting (up to *node 5*), volcanic hazard (up to *node 8*), and volcanic risk assessment (up to *node 10*). Each node is characterized by a probability conditioned by the events at the previous node; for instance, *node 3* is characterized by the conditional probability of eruption given the presence of magmatic unrest (*node 2*). The probability of eruption at one specific location and of one specific size is given by the combination of the probabilities of the first 5 nodes

where $B(\alpha, \beta)$ denotes the beta function

$$B(\alpha, \beta) = \int_0^1 \theta^{\alpha-1}(1-\theta)^{\beta-1}\, d\theta. \tag{2}$$

The expected value for this beta distribution is

$$E(\Theta) = \frac{\alpha}{\alpha+\beta}, \tag{3}$$

and the variance is

$$V(\Theta) = \frac{E(\Theta)(1-E(\Theta))}{(\alpha+\beta+1)}. \tag{4}$$

The beta distribution is often chosen because it is defined on the range [0, 1] and is the conjugate prior distribution in the binomial model (Gelman et al. 1995), which makes easier the merging of expert opinion/model output with observational data. This choice is subjective, and other distributions can be used, such as the Gauss distribution for the logistic transformation of the probability. In practice, the differences arising from the use of various distributions are not usually significant; for this reason, the beta distribution is the most commonly used in practical applications (e.g., Gelman et al. 1995). The parameters of the beta distribution can be set through expert opinion and updated using the data (if any) of past eruptions (see Marzocchi et al. 2008; Neri et al. 2008).

Bayesian inference, as applied to the event tree, brings two key advantages. First, the probability at each node is described by a distribution instead of a single value; this allows scientists to quantify formally the uncertainty related to the probability estimation and, in due course, to the hazard assessment. Second, the Bayesian inference provides a framework for merging all the relevant available information such as theoretical models, prior beliefs, monitoring measures, and past data (Fig. 1).

The choice of describing the probability as a distribution instead of a single value has many similarities with the approach taken by the SSHAC but has been criticized on philosophical grounds. Basically, critics interpret the probability as the only measure of uncertainty, and in this view, it does not make sense to assign an uncertainty to an uncertainty (Bedford and Cooke 2001). Here, we do not tackle this important controversy, but we note that the use of a probability distribution instead of a single value has the undoubted advantage of providing a measure of the reliability of any probabilistic assessment, especially in systems where the epistemic component is significant to the outcome.

Short-term forecast methods

Short-term forecasting is driven by the information provided by monitoring anomalies, i.e., by the occurrence of one or more concomitant monitoring signals outside the background range. During a phase of unrest, the long-term probability and the past frequency of eruptions become less important than what is being observed from the monitoring of physical quantities which are assumed to be related to the possibility of an imminent eruption. In a nutshell, we may describe the short-term eruption forecasting problem as the quantification of observed anomalies, also known as *precursors*, in terms of a probabilistic assessment. Despite its wide use, the term precursor has never been clearly defined; here, we use this term in a broad sense, indicating one or more anomalies that may anticipate an eruption with a certain probability in a specific time interval.

As in seismology (Jordan 2006), volcanic precursors were often interpreted in a deterministic sense in the past; more formally, the observation of a precursor was translated into a probability through the implicit use of the Heaviside function: the probability is zero if the precursor is not observed and jumps to one when the precursor is detected. Now, volcanologists acknowledge that eruptions may be anticipated by several different precursors, and some eruptions may occur without clear premonitory signals. Conversely, many identified precursors may occur without anticipating an eruption (e.g., Moran et al. 2011). This lack of a one-to-one correlation between precursors and eruptions emphasizes the need for probabilistic assessment (UNDRO 1985; Decker 1986; Sparks 2003).

The basic principle is that the observation of a precursor should increase the estimated probability of eruption. Despite the simplicity of this principle, we argue that the probabilistic assessment may be complex. Ideally, if a very large database of past monitored eruptions, and noneruptions, is available, we can estimate the probability related to each precursor as its frequency calculated from the database. However, at the level of individual volcanoes, such databases are very rare and usually small. Often, volcanologists face cases where no past monitoring data are available. The WOVOdat exercise led by Christopher Newhall (www.wovodat.org) aims at building a database of monitored episodes of unrest at a worldwide scale. This database may serve as the primary resource for a new field of volcano epistemology: during volcanic crises, it can be used to make queries along the lines of where else have X, Y, and Z been observed and what happened subsequently. Since different types of volcanos have different precursors, a critical source of epistemic uncertainty is the selection of the analogs used to subset the database; a poor choice of analog volcanoes can easily lead to a biased forecast. Notwithstanding the huge potential impact of WOVOdat, the ever growing monitoring capabilities at volcanoes will allow scientists to, at best, calculate the frequency of the most basic precursors, as the most elaborate and innovative monitoring measures will be available only for shorter time periods and fewer volcanoes. For this reason, the calculation of frequencies is not usually a viable way to make use of these innovative and new monitoring measures. More integrated probabilistic approaches that merge past data, conceptual models, and expert opinions will be required.

In the following, we present a brief overview of the most common seismic and ground deformation precursors, without pretending to a complete review of this very large and rapidly evolving field of research. There are several other precursors that may be important to forecast volcanic eruptions, such as geochemical signals, but their interpretation is usually strongly case-dependent and requires a volcano-specific conceptual model. Often, precursors and anomalies are identified after the eruption, focusing on only those processes that appeared to lead to eruption. This backward process is often misleading, as the false positives are not observed; in other words, there is no guarantee that a precursor identified a posteriori will show up again for future eruptions or that a different precursor may not serve as a functional prior for future events. For these reasons, we emphasize precursors that are commonly monitored and/or applied in a prospective way to forecast eruptions and are therefore of practical interest. After that, we describe probabilistic models that produce forecasts relying on the observation of precursors.

Seismic precursors

Seismic activity is considered the most important signal in forecasting eruptions. Its use is motivated by the physical basis of the signals, easy interpretation of data in real time, and the widespread availability of seismic networks to monitor volcanoes. Magma movement increases the stress of the surrounding rocks which is released mostly through seismic activity. Relative to tectonic seismic activity, the seismicity in volcanic areas presents a much wider variability, particularly in the spectral components of the seismic waves (e.g., Neuberg 2011). This variability reflects the large range of elastic behavior and rheology in the rising magma and surrounding rocks. Volcanoes are characterized by different magmas (e.g., basaltic and andesitic volcanoes) and mechanical conditions (e.g., open and closed conduits); all these features lead to a large variety of possible seismic precursors, with open conduit effusive volcanoes showing the most subtle signals, while closed conduit volcanoes with more silicic magmas usually present much clearer evidence of precursory seismic signals.

Upward migration of the seismic activity as the magma rises is not as common as might be expected; a few remarkable cases were reported for Tolbachik (Tokarev 1978), the Piton de la Fournaise (Battaglia et al. 2005) and Long Valley Caldera (Shelly and Hill 2009). Notably, the lead time to an eruption may be very short with respect to the upward migration of the seismicity, such as in the case of Chaiten (Castro and Dingwell 2009). This indicates that magma can migrate very quickly towards the surface once it finds a mechanical weakness. Lateral migrations are probably more interesting for practical purposes; they have been observed before the Pinatubo 1991 eruption and in other volcanoes (White and Power 2001; Pozgay et al. 2005). This precursor is particularly important because the lead time to eruption is usually much longer than for upward migration leaving more time for possible mitigation actions, and identifying lateral migration requires only epicentral estimation, which has greater precision than depth estimation.

The presence of fluids/magma leads to a wide variability in the frequency content of the seismic waves. Typically, earthquakes that involve fluids in some manner have a spectral content with more energy at lower frequencies. These events are called low-frequency or long-period events (LPEs hereafter), the presence of which is usually an indicator of pressurization in the system (e.g., Chouet 1996; Neuberg 2000). Notwithstanding the clear importance of fluids, the exact origin of these LPEs is still controversial. More important, from a practical point of view, is that it is not yet clear how to distinguish the presence of "cold" fluids from that of magma when analyzing LPEs. This implies a considerable uncertainty about how to interpret the presence of LPEs as a possible precursor of an impending magmatic eruption (Chouet 1996; Thomas and Neuberg 2012). For this reason, derivatives of LPEs such as the rate acceleration (Traversa et al. 2011), the link with the amplitude (Hammer and Neuberg 2009), and variation of the spectral content of seismic signals (Neuberg 2000; Bryan and Sherburn 2003) have been suggested as potential precursors.

One of the better known seismic precursors is the acceleration of the seismic rate and/or energy that has anticipated some past eruptions. The rationale behind this seismic acceleration is based on the extension and coalescence of fractures inside the volcano that can create a preferred pathway for magma ascent (Voight 1988, 1989; Cornelius and Voight 1996), even though other possible explanations exist (Lavallée et al. 2008). Despite such an acceleration having been retrospectively observed before recent large eruptions at Pinatubo in 1991 (Cornelius and Voight 1996; Smith and Kilburn 2010) and at Soufriere Hills in 1995 (Kilburn and Voight 1998), there is still concern about detection of acceleration prior to an eruption (Bell et al. 2011). Even in retrospective studies, the lead time to eruption seems very short, from a few hours to a few days (Smith and Kilburn 2010; Kilburn 2003; Chastin and Main 2003).

Probably, the most "definite" seismic precursor is volcanic tremor. The origin of tremor can be due to a coalescence of LPEs (Neuberg 2000) or by processes not related to earthquakes, such as fluid flow through rough-walled conduits (e.g., Benoit et al. 2003; Jellinek and Bercovici 2011). Notwithstanding its importance, tremor is not often very useful in practical forecasting applications because it shows up only immediately before and during an eruption.

Seismicity has also been used to detect stress changes before the reactivation of a volcano, using variations in the seismic noise (Brenguier et al. 2008), shear waveform splitting (Miller and Savage 2001; Gerst and Savage 2004), and fault plane solutions (Roman et al. 2006). Discrepancies in the time evolution of a seismic swarm relative to models used for tectonic earthquakes such as the epidemic type aftershock sequence (Ogata 1988, 1998) have been used to track the fluid/magma motion under the crust (Hainzl and Ogata 2005; Lombardi et al. 2006). Seismic clusters as potential precursors have also been studied using the pattern recognition approach (Mulargia et al 1991, 1992; Sandri et al 2005; Jaquet et al. 2006; Novelo-Casanova and Valdes-Gonzales 2008).

Finally, seismic activity has also been coupled to other observables in order to identify possible precursors. Thelen et al. (2010) found that the cumulative seismic moment and the repose time of the volcano may be indicative of impending eruptions at Mt. St. Helens. Similarly, Passarelli and Brodsky (2012) found a correlation between the duration of seismic unrest before an eruption and the repose time of the volcano.

Ground deformation precursors

The pressurization of the volcanic system induces a ground deformation that can be detected at the surface using different techniques/instruments such as electronic distance measurements (EDM), tiltmeters, dilatometers, global positioning system (GPS), and interferometry synthetic aperture radar (InSAR). EDM, tiltmeters, and GPS provide the deformation at the site of the instrument, and data can be collected in real time; InSAR provides a much broader spatial view, but the frequency at which data are collected is limited by the satellite's orbit (e.g., Dzurisin 2003). The observation of ground deformation on the surface is not necessarily linked to an impending eruption (e.g., Fournier et al. 2010) since many other processes, such as pressurization of the hydrothermal system, aborted intrusions, and tectonics may induce comparable signals (e.g., Bonafede 1991; Dzurisin 2003; Todesco et al. 2010).

Usually, volcanologists interpret surface ground deformations by means of physical modeling; assuming a physical model responsible for the deformation, the parameters of the model (e.g., overpressure, depth of the source, and rheology of the medium) are set to best reproduce the observations on the surface. This approach has been followed for many years using an isotropic point source model in a homogeneous elastic half-space (Mogi 1958), but in recent years, it has been demonstrated that the adoption of more general source models (Davis 1986), embedded in a realistically layered medium, can produce significantly different interpretations (e.g., Battaglia et al. 2003; Amoruso and Crescentini 2009, 2011; Trasatti et al. 2011).

Notwithstanding the clear and undoubted importance of ground deformation, the nonuniqueness of the inverse problem allows for multiple alternative interpretations of ground deformation. The most recent trend is to use other measurements, like the gravimetric field, to better constrain the deformation model (e.g., Battaglia and Hill 2009; Amoruso et al. 2008) and thus estimate the volume of newly intruded magma. For practical purposes, this kind of modeling is usually very time-consuming, preventing its use in tracking the evolution of a rapidly evolving unrest. For this reason, volcanologists often rely on simple empirical signatures of ground deformation, the most striking example being that of the paroxysmal phase of the Mt. St. Helens eruption in 1980. That eruption actually began in March 1980 in the summit region, without observed deformation precursors. On May 18, the paroxysmal phase, including a Plinian eruption, occurred. This phase was triggered by a flank collapse facilitated by a significant cumulative deformation of hundreds of meters produced over a few months by a cryptodome (Voight et al. 1981). Although the paroxysmal phase of the eruption on May 18 did not show any short-term precursors, the ground deformation clearly brought the volcano to a critical instability (Voight et al. 1981), and thus the cumulative ground deformation has to be considered an important precursor, even though the timing of the eruption is not estimable since the collapse may occur immediately or months later.

Many papers report some deformation before eruptions (e.g., Swanson et al. 1983; Voight et al. 1999; Bonaccorso et al. 2002; Iguchi et al. 2008), but it is not easy to identify a common pattern from them (cf., Fournier et al. 2010). In some cases, like Sakurajima (Iguchi et al. 2008), deformation has been remarkably successful as a precursor, with a lead time of minutes to a few hours. However, deformation is also often associated with nonerupting volcanoes (Fournier et al. 2010). As with seismicity, acceleration in ground deformation is usually considered an important precursor. Historical chronicles of the last eruption at Campi Flegrei (Guidoboni and Ciuccarelli 2011) highlighted a strongly localized ground deformation acceleration a few hours before the eruption, but the pattern was made more complicated by a strong inversion of the ground deformation signal, with a considerable deflation immediately before the eruption onset.

Precursors and probability

How to convert the detection of precursors into the probability of an eruption is probably one of the most difficult and less explored issues. The most obvious approach is the frequentist approach, i.e., calculating from the data of the past history of the volcano the hit rate:

$$\gamma = \frac{m}{N} \tag{5}$$

and the false alarm rate

$$\delta = \frac{M - m}{M}, \tag{6}$$

where N is the total number of eruptions, M is the number of times in which the precursor has been observed, and m is the number of eruptions anticipated by the precursor in an arbitrary time window τ. When the time window τ is long, another important parameter is the fraction of time in which the precursors indicate the possibility of an eruption. This approach has been applied to some frequently erupting volcanoes by using simple seismic patterns (Traversa et al. 2011, at Ubinas Volcano in Perù; Novelo-Casanova and Valdes-Gonzales 2008, at Popocatepetl Volcano in Mexico; Grasso and Ziapalin 2004, at Piton de la Fournaise; and Mulargia et al. 1991, 1992, at Etna). Retrospective testing has shown a large variability in these parameters, with γ from 60 to 100%, and δ from 20 to 60%. At Soufriere Hills Volcano, Jaquet et al. (2006) used precursors in a probabilistic framework through the use of stochastic simulation on the basis of potential evolution scenarios.

Unfortunately, as noted above, such unrest databases do not yet exist for most high-risk volcanoes.

The lack of databases and the constant improvement in high quality monitoring procedures makes the use of expert opinion unavoidable. As a matter of fact, volcanologists already rely strongly on subjective interpretation of precursors, usually in terms of alert levels (see, for instance, the alert levels at Vesuvius: http://www.protezionecivile.gov.it/jcms/en/view_pde.wp;jsessionid=9D8024967566483BC206C6859D2994E3?contentId=PDE12771#livelli_allerta). Nonetheless, we propose that a more structured and transparent probabilistic procedure to forecast volcanic eruptions may have significant advantages. Probably the most important advantage is that the use of probabilities facilitates the establishment of transparent and quantitative decision-making protocols (Marzocchi and Woo 2009). These protocols have to prepared before the crisis and take the form of quantitative rules that explain how different probability values are translated into mitigation actions. Such protocols can justify, even a posteriori, each step of the decision-making process; they are formidable educational and communication tools for both society and scientists. During an emergency, volcanologists can be under a great deal of pressure which may affect their perspective. A protocol established through an inclusive process in a period of repose may be much more effective than hasty evaluation under pressure.

To date, for high-risk volcanoes, several efforts have been devoted to translating the observation of one or more precursors into a probabilistic assessment using expert opinion. Notable are the event tree (Newhall and Hoblitt 2002; Marzocchi et al. 2004, 2008), and the Bayesian belief network (BBN, Aspinall et al. 2003, 2006). Both techniques have a similar structure and represent a general quantitative framework where all relevant monitoring observations are embedded into a probabilistic scheme through expert opinion, conceptual models, and possibly data of past monitored phases of unrest.

The BBN is a graphical representation of the relevant observations (nodes) and causal links among the nodes. Associated with each node, there is a set of conditional probabilities that describe the relationship between the states of the variable at the node with the states of the other variables at the connected nodes. These conditional probabilities are then combined through Bayes Theorem in order to get the probability of any specific event in which we may be interested. Noteworthy is the similarity with the event tree philosophy described above. The BBN developed by Aspinall et al. (2003) for Galeras is fed directly with conditional probabilities obtained by frequency of observations and/or expert opinion. In Montserrat, the monitoring observations were transformed into conditional probabilities through repeated expert elicitation sessions (Aspinall 2006). This approach may not be feasible during the evolution of a rapidly evolving crisis (e.g., the case of the Ruaumoko exercise, Lindsay et al. 2010) since it may take time to convene experts and to reach a coherent and shared assessment.

The Bayesian event tree (Marzocchi et al. 2008) and its subsequent applications (Sandri et al. 2009, 2012; Lindsay et al. 2010; Selva et al. 2012) follow a different strategy: expert opinion is elicited regarding which, and what level of, monitoring anomalies best characterize a phase of unrest (node 1), a magmatic displacement (node 2), and an imminent eruption (node 3; see Fig. 1). A monitoring anomaly is usually related to the observation of a single monitoring signal, but it can be also associated with the simultaneous observation of two or more signals of different nature. In this way, during a period of unrest, monitoring information allows us to calculate in almost real-time the time evolution of the eruption probability. For example, if monitoring at time τ identifies a number $Z(\tau)$ of monitoring anomalies at one specific node, this number is translated into a probability using a suitable transfer function. For nodes 2 and 3, the simplest learning curve is used (Marzocchi et al. 2008, see also Fig. 2)

$$E(\Theta) = 1 - a \exp(-b\, Z(\tau)) \tag{7}$$

where $E(\Theta)$ is the average of the beta distribution (see Eq. 3), a and b are parameters that are estimated from past monitored phases of unrest (if any) through Bayesian inference (Marzocchi et al. 2008). This functional form has some interesting features. Firstly, this relationship is monotonically increasing so that the larger the $Z(\tau)$, the larger the probability. Secondly, it implies that the largest increase in probability mean occurs when one of the monitoring variables shows some degree of anomaly; as more monitored variables become anomalous, the probability mean continues to rise but more slowly. It should also be noted that including more monitoring measures does not decrease the probability mean, even if they are all unobserved. This procedure implies that each monitoring anomaly (associated with a single or multiple monitoring signals) yields an absolute amount of information about the state of the preeruptive process that does not depend on how many other anomalies are observed or not. Marzocchi et al. (2008) set the variance of the beta distribution (see Eq. 4) at its largest value, accounting for the very rough knowledge of the preeruptive processes, but this value can be modified if more

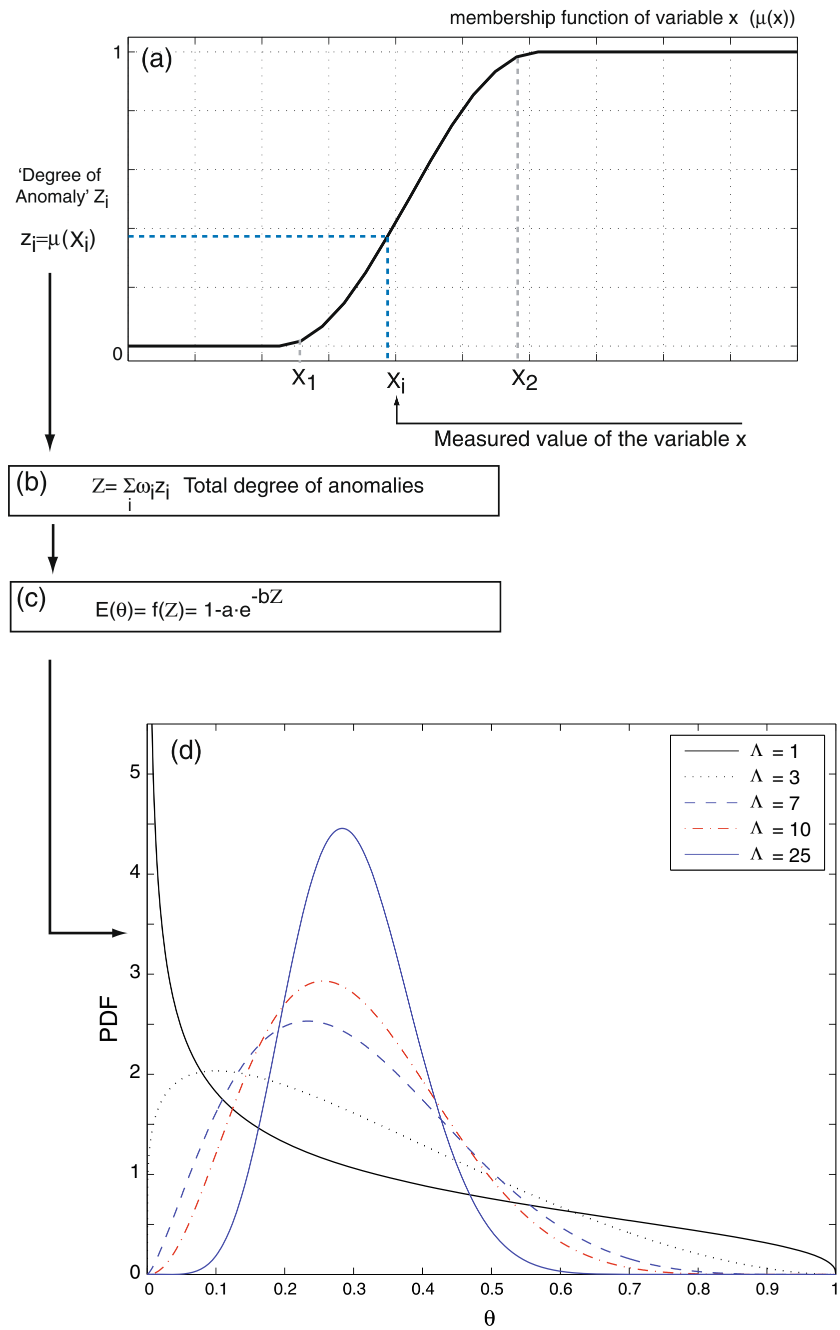

Fig. 2 An example of how monitoring measures are transformed into a probabilistic assessment for node 3 of the event tree (Fig. 1). How one monitoring measure x_i is translated in a degree of anomaly z_i according to a selected membership function $\mu(\cdot)$ (**a**). A measure below x_1 is considered background, above x_2 is anomalous, and in between it has a certain degree of anomaly. After collecting the degree of anomaly for all parameters considered, we combine them using a weighted average (ω_i is the weight of the i-th parameter) in order to obtain the total degree of anomaly (**b**). The parameters, weights, and thresholds may be selected by a panel of experts through elicitations. Then the total degree of anomaly is transformed into an average probability using a predefined function (see Eq. 3); for example, for node 3 Eq. 7 is used (**c**). Finally, the probability distribution is obtained by imposing the average as calculated in **c**, and an equivalent number of data Λ (Eq. 38) that mimics the reliability attached to the probability estimate (**d**). Usually, Λ is small when a substantial disagreement among scientists exist, and Λ is large when most of experts agree on the selected parameters and thresholds. The *plot* shows the distribution using different values of Λ

realistic conceptual models are available. The advantage of this procedure is that experts are queried directly about what they know best (monitoring anomalies) instead of being asked to provide probabilities, which are usually more challenging to assess directly. (For this reason, many expert elicitation sessions start with a short probability course illustrating the most common fallacies in assigning probabilities.) From a practical point of view, the method is able to provide probabilistic assessment in almost real-time since all rules to identify anomalies have been decided during a period of repose.

Long-term forecast methods

The probabilistic attitude to eruption forecasting was first adopted for long-term assessment by Wickman (1966a). The most extensive application has been that connected with the proposed high-level radioactive waste repository at Yucca Mountain, summarized by Crowe et al. (1998) (see also Connor et al. 2000). The risk associated with a possible eruption of Vesuvius has also received considerable attention (e.g., Scandone et al. 1993; Marzocchi et al. 2004), and the hazard from many other volcanoes has been estimated using a variety of approaches that we will discuss later. In contrast to short-term forecasting, where the focus is on the treatment of the monitoring observations, long-term hazard assessment is far more concerned with the quality and quantity of the available past eruption data. Nonetheless, more data are not always better, as inhomogeneity must need complicate the model, and different eruption catalogs can lead to different eruption forecasts (Wang and Bebbington 2011). The eruptive history of Mt. Etna is probably the most extensively examined, providing a detailed dataset that can be used for investigations of many types. Hence, the many studies of long-term volcanism at Mt. Etna have been more often concerned with determining "what makes it tick," rather than primarily with hazard assessment.

The shift from short-term to long-term eruption forecasting is determined by whether or not there are anomalous monitoring observations. In the absence of the latter, probabilistic forecasts have to be made of the basis of the past eruption history of the volcano.

Given this dependence on the past behavior, we are confronted with a question: Has the observed history of the volcano been characterized by activity at statistically different levels (or rates) for different intervals of that history? If not, then we can use a stationary model; otherwise, we need a nonstationary model, which may involve trend(s) or level changes in activity. As the models for the former are simpler, we will begin with them. First, we require some notation.

Let us suppose that our observed history consists of eruption onsets at times $0 = t_0 < t_1 < t_2 < \ldots < t_n$, where T is the elapsed time from t_0 to the present. We will assume that there is no information about activity before t_0, in keeping with the idea that absence of evidence is not evidence of absence. The volume of the ith eruption is denoted by v_i and the interonset times defined as $r_i = t_i - t_{i-1}$ for $i = 1, \ldots, n$. The latter are often referred to as repose times, although strictly that should mean the time between the end of an eruption and the onset of the subsequent eruption. The incomplete current repose $T - t_n$ is denoted r^*. For further discussion on the issue of treating eruptions as point events (in time), see Bebbington (2008) and Garcia-Aristizabal et al. (2012); a model incorporating eruption durations was proposed by Bebbington (2007).

Volcanological data are broadly of two types: historical (or observed) and geological. Historical records are usually short, often no longer than a few centuries, apart from at a handful of well-recorded volcanoes, and subject to incomplete observation, particularly in the earlier part of the records. Geological data on the other hand can go back a millennia. Such data are typically sourced from distal tephra records or dating of proximal volcanic products. In either case, the data needs to be age-interpolated consistent with stratigraphy, and multiple records may need to be merged. Also, the geologic record often incompletely preserves evidence of smaller eruptions, and burial of older deposits is common. Turner et al. (2008a) introduced the use of monotone spline functions and Monte Carlo simulation for sediment cores to fit hazard models to geological data. Other methods have been suggested by Cronin et al. (2001), Mendoza-Rosas and De la Cruz-Reyna (2008, 2010), and by Bebbington and Cronin (2011). Turner et al. (2009) explored the issue of merging multiple geological records, and Burt et al. (2001) have examined the problem of missing observations in the depositional record.

There are many types of models that can be fit to eruption onset data. Some decisions such as, for example, use of volume information, may be decided by the availability or reliability of data. Others, such as stationarity, can be determined by a simple statistical test. This leaves the question of which distribution or stochastic process from among the possible alternatives best describes the data. As different models can have different numbers of parameters, we need a means of compensating for the effect of additional parameters and thus avoiding overfitting. If models are fit using maximum likelihood techniques, this can be done using the Akaike information criterion (AIC) (Akaike 1977):

$$\mathrm{AIC} = -2p + 2 \log L, \tag{8}$$

where p is the number of parameters and $\log L$ the log likelihood. Larger AIC indicate better models. Note that we cannot, in general, use the corrected AIC (Hurvich and Tsai 1989), as there is no proof of its validity for point-process models (Claeskens and Hjort 2008), unlike for linear regression and autoregressive models. Linear regression is not usually applicable for model fitting, except for survival analysis techniques such as the Weibull plot (see, e.g., Bebbington and Lai 1996b). It cannot be used on the empirical survival curve (although least squares can), as the points on the survival curve are not independent. If the models are nested, then the difference in AIC is the basis of a formal likelihood ratio test (e.g., Bebbington and Marzocchi, 2011). Given a sufficiently long record, an alternative procedure for comparing models is sequentially via the probability gain (Passarelli et al. 2010b).

Stationary methods

In some sense, stationary models constitute a maximum ignorance alternative, in that no information about the temporal evolution of the activity is sought or possessed. They are also more robust, as incorporating an incorrect nonstationary model will lead to greater bias. Hence, from a statistical point of view, tests are best constructed with stationarity as the null hypothesis, evidence being required to reject this.

If we have n onsets in a time window of length T, then our definition of stationarity that the observed history of the volcano has been characterized by activity at a constant level would imply that the n events are distributed randomly along the interval T. In statistical terms, the interonset times $r_1, \ldots, r_n, r^*$ should not show any trend with time. This hypothesis can be checked with various parametric and nonparametric tests such as the sign test and the Laplace trend test (Cox and Lewis 1966). Although they are stationary, clustering processes, such as in the extreme case of the Auckland Volcanic Field (Bebbington and Cronin 2011), may not be amenable to such analysis.

Some stationary data show a correlation between $r_1, \ldots, r_n, r^*$ and the eruptive volumes $v_0, v_1, \ldots, v_n$, in which case some form of volume dependence needs to be included in the model. Otherwise, provided that the reposes are not autocorrelated (i.e., there is no correlation between successive repose lengths), we have a renewal process. If the repose lengths are consistent with an exponential distribution, this is the special case of a Poisson process.

Long-term average (Poisson process) behavior

The reason that the Poisson process is of particular importance is derived from the exponential distribution of the repose lengths. Suppose that the most recent onset occurred at some time u and that a time s has elapsed since. If $N(u, s)$ is the random variable corresponding to the number of onsets in the time interval (u, s), then the conditional distribution that the current repose extends at least a further time t is

$$\begin{aligned}
\Pr[N(u, s+t) = 0 | N(u, s) = 0] &= \frac{\Pr[N(u, s+t) = 0]}{\Pr[N(u, s) = 0]} \\
&= \frac{\exp[-\lambda(s+t)]}{\exp[-\lambda s]} \\
&= \exp(-\lambda t) \\
&= \Pr[N(s, s+t) = 0]\,.
\end{aligned} \tag{9}$$

In other words, the distribution of the remainder of the repose length is independent of the elapsed repose length. This is the memoryless property and implies that we know nothing about the temporal structure of the process. The parameter $\lambda = n/T$ is the average rate of onsets, and an equivalent characterization is to say that the probability of an onset in a short time interval of length Δ is approximately $\lambda\Delta$, or that the number of onsets in a time interval of length S has a Poisson distribution with a mean of λS.

Note that the memoryless nature of the process works two ways. If, particularly with geological data, we are unsure about the exact onset times, then assuming a Poisson process is robust to that uncertainty. The model is also parsimoniously parameterized and can be a nested model within any point-process formulation. These properties make it an obvious baseline model; other models can then be evaluated on their performance relative to the Poisson process by using measures such as the probability gain (e.g., Passarelli et al. 2010a).

There have been a number of approaches to perturbing the rate λ in the Poisson process, driven by the fact that the mean and standard deviation of repose lengths are not usually equal, as is implied by the exponential model. Ho (1990) suggested a Bayesian approach, in which λ has a gamma prior distribution. Solow (2001) augmented this to an empirical Bayes formulation by suggesting that an informative prior could be constructed from the eruption records of a group of similar volcanoes. However, the gamma prior results in the number of eruptions in a time interval having a negative binomial distribution and hence forces overdispersal on the data. This was addressed by Rodado et al. (2011), using the Conway–Maxwell–Poisson distribution, which allows both underdispersal and overdispersal, as a base. Bebbington and Lai (1998) took a different line by incorporating serial dependence in a generalized negative binomial distribution, applied to eruptions from Mt. Sangay (Ecuador).

For the Poisson process, the parameter λ is the rate at which new onsets occur. In this case, the rate is constant, and the result is called a (time)-homogeneous process. All other processes are time-inhomogeneous, where the constant rate λ is replaced by a function of time, $\lambda(t)$. This has the interpretation that the probability of an onset in the short time interval $(t, t+\Delta)$ is approximately $\lambda(t)\Delta$. The function $\lambda(t)$ is usually termed the "hazard rate" but can also be called the point-process intensity.

The hazard rate $\lambda(t)$ uniquely defines the forecast behavior, in that the probability of at least one eruption in a time interval (s_0, s_1) is as follows:

$$\Pr[N(s_0, s_1) > 0] = 1 - \exp\left[-\int_{s_0}^{s_1} \lambda(t)\mathrm{d}t\right]. \tag{10}$$

All desired forecast quantities can be obtained from Eq. 10. There are two basic families of time-inhomogeneous models, depending on whether $\lambda(t)$ varies stochastically as a function of the past history, or whether it is externally forced.

Previous-eruption dependence

As the basic character of the Poisson process implies that $\lambda(t)$ is constant, in order to weaken this constraint, we need to examine the tenets of the Poisson process. One of these tenets is that in a short enough time interval at most one event can occur, which is automatically satisfied by the definition of what makes an eruption. A second tenet is that events occur independently. There are many ways in which this dependency can be expressed, but the simplest is to assume that the previous eruption influences the timing of the next eruption onset. This is intuitively attractive for volcanoes, being the realization of a magma chamber model. If the size of the previous event is taken into account, we have a time-predictable model; otherwise, we have a renewal model.

Renewal models A renewal process is characterized by the intervals between events being independent and identically distributed with a distribution function F. Thus

$$\lambda(t) = \frac{f(t-s)}{1-F(t-s)}, \quad t > s, \tag{11}$$

where the most recent event occurred at time $s < t$, and $f = F'$ is the renewal density. As $t - s$ is the elapsed repose time, Eq. 11 can be rewritten:

$$\lambda(r) = \frac{f(r)}{1-F(r)}. \tag{12}$$

Thus only the elapsed time since the last eruption may control the time to the next eruption. Previous eruptions exert an influence only through their contribution to the parameter estimates in f. The special case in which λ of Eq. 11 does not depend on r is the exponential distribution (Poisson process) described above.

A variety of tests exist to detect whether an observed process deviates significantly from a renewal process. These reduce to testing for (a) a constant average rate of events, (b) independence between successive repose times, and (c) goodness-of-fit to the hypothesized renewal density f. Of particular interest in the latter is whether an exponential distribution also fits. This corresponds to testing the Poisson null hypothesis versus a more general renewal process where the time elapsed since the last eruption is important. For more details, see Reyment (1969), Bebbington and Lai (1996a), and Bebbington (2010).

Given a density $f(r)$ and observed interonset times r_i, $i = 1, \ldots, n$, the parameters can be estimated by maximum likelihood. That is, the values are chosen, either algebraically or numerically, to maximize the likelihood:

$$L(r_1, \ldots, r_n, r^*) = \left[1 - F(r^*)\right] \prod_{i=1}^{n} f(r_i), \tag{13}$$

where r^* is the elapsed time since the most recent eruption, and $1 - F(r)$ is the survival function for the reposes, i.e., the probability that a repose lasts longer than r. The likelihood Eq. 13 can be augmented by a term for a first incompletely observed repose, but as the start of the observation period is often impossible to establish at typical time scales, we usually consider observation to have started at the first recorded onset. Occasionally, other methods for estimating the parameters are used, such as the method of moments or least squares.

A large number of renewal distributions have been used, as noted in Table 1, including the exponential

$$f(r) = \lambda \exp(-\lambda r), \quad \lambda > 0 \tag{14}$$

which is exactly the Poisson process, the Weibull

$$f(r) = \alpha(\beta r)^{\alpha-1} \exp\left[-(\beta r)^{\alpha}\right], \quad \alpha, \beta > 0, \tag{15}$$

Table 1 Renewal models for volcanic hazard

Distribution	Trend[a] in $\lambda(r)$	Volcano	References
Linear	I	Hekla, Katla	Thorlaksson (1967)
Exponential[b]	C	Kilauea, Mauna Loa	Klein (1982)
		Yucca Mountain	Crowe et al. (1982)
		Villarrica, Llaima, Tupungatito	Munoz (1983)
		Etna	Mulargia et al. (1985)
		Aggregate	De la Cruz-Reyna (1991)
		Azores	Caniaux (2005)
Pareto	D	Colima	Medina Martinez (1983)
Lognormal	U	Various	Bebbington and Lai (1996a)
		Katla	Eliasson et al. (2006)
Weibull	I or D	Various	Bebbington and Lai (1996a)
		Ruapehu, Ngauruhoe	Bebbington and Lai (1996b)
Power law	D	Aggregate	Pyle (1998)
Gamma	I or D	Citlaltepetl	De la Cruz-Reyna and Carrasco-Nunez (2002)
Log-logistic	U	Soufriere Hills	Connor et al. (2003)
		Various	Watt et al. (2007)[c]
Mixture of exponentials	V[d]	Colima	De la Cruz-Reyna (1993)
		Colima, Popocatepetl	Mendoza-Rosas and De la Cruz-Reyna (2009)
		Villarrica, Llaima	Dzierma and Wehrmann (2010)[c,e], Wehrmann and Dzierma (2011)[c,e]
Mixture of Weibulls	V[f]	Taranaki	Turner et al. (2008a, 2009)
Inverse Gaussian[g]	U	Miyakejima	Garcia-Aristizabal et al. (2012)[c,e,h]

[a]How the estimated hazard $\lambda(r)$ changes as the current repose length r increases. *C* Constant, *I* Increasing, *D* Decreasing, *U* Unimodal, *V* Various
[b]Poisson process
[c]Also used the Weibull density
[d]The two component version is decreasing
[e]Also used the exponential and log-logistic densities
[f]The two component version has eight possible shapes (Jiang and Murthy 1998)
[g]Also known as the Brownian passage time model
[h]Also used the log-normal and gamma densities

which includes the exponential as a special case, as does the gamma

$$f(r) = \beta^\alpha r^{\alpha-1} \exp(-\beta r)/\Gamma(\alpha), \quad \alpha, \beta > 0, \tag{16}$$

which is the sum of α independent exponential random variates if α is an integer. While Settle and McGetchin (1980) fitted a Gaussian density to the interonset times of eruptions at Stromboli, the nonzero probability of a negative repose means that it cannot be used for forecasting purposes. Other distributions tried so far include the Pareto

$$f(r) = a(1 + bu)^{-a/b-1}, \quad a, b > 0, \tag{17}$$

power law

$$f(r \mid r > c) = \frac{r^{-1-1/b}}{b c^{-1/b}}, \quad r > c > 0, b > 0, \tag{18}$$

log-normal

$$f(r) = \frac{1}{r\sigma\sqrt{2\pi}} \exp\left[-\frac{(\ln r - \mu)^2}{2\sigma^2}\right], \quad \sigma > 0, \tag{19}$$

log-logistic

$$f(r) = \frac{\eta\gamma r^{\gamma-1}}{(1 + \eta r^\gamma)^2}, \quad \eta, \gamma > 0, \tag{20}$$

and inverse Gaussian

$$f(r) = \sqrt{\frac{\mu}{2\pi\alpha^2 r^3}} \exp\left[-\frac{(r-\mu)^2}{2\alpha^2 \mu r}\right], \quad \mu, \alpha > 0. \tag{21}$$

The basic distributions above can also be used to construct mixture distributions

$$f(r) = \sum_i \pi_i f_i(r), \quad \sum_i \pi_i = 1, \tag{22}$$

where the components f_i need not be from the same distribution. The idea of a mixture renewal model is that the distribution of the next repose is chosen randomly according to the probabilities $\{\pi_i\}$. Thus, for example, a two component model might represent the possibility of the previous eruption closing the conduit (cf., Marzocchi and Zaccarelli 2006) or the depletion of the upper magma plumbing system, requiring a recharge from the mantle on a longer time scale (cf., Turner et al. 2008b). As each additional component adds $p + 1$ parameters, where p is the number of parameters in the additional density, mixture models can easily have a large number of parameters and, hence, a complex structure. This allows them to fit data much more closely, and the danger of overparameterization must be guarded against by comparing their performance with that of simpler (ideally nested) models

Table 2 Nyamuragira eruption data from Wadge and Burt (2011)

Onset date (year)	DRE volume (10^6 m^3)
1901.5	59
1904.364	21
1905.556	54
1912.923	62
1938.079	170
1948.164	52
1951.877	19
1954.142	53
1956.877	1
1957.992	5
1958.600	116
1967.310	71
1971.227	62
1976.978	43
1980.082	62
1981.984	118
1984.148	64
1986.540	51
1987.997	6
1989.318	93
1991.721	198
1994.507	46
1996.918	45
1998.795	62
2000.074	43
2001.101	127
2002.564	54
2004.351	62
2006.907	42
2010.005	50

using AIC or similar methods. The underlying conundrum is known as the bias-variance tradeoff. Given a sufficiently complex model, the observed data can be fitted arbitrarily closely. However, assuming some degree of aleatory uncertainty, the model will conform too closely to the observed sample, resulting in a high degree of variance in the forecasts made from it. If the model is not complex enough, the forecasts will exhibit little variance but have large bias relative to the underlying system.

In general, a renewal density can be characterized by the shape of the associated hazard rate $\lambda(r)$. If eruptions tend to cluster in time, then the hazard rate will be greatest immediately after the previous event, and so $\lambda(r)$ will be decreasing with r. Both increasing and unimodal (first increasing and then decreasing) hazard rates imply that eruptions will tend to be more regular in time than random (Poisson process) behavior. The decrease beyond the peak of a unimodal hazard rate may signal a change in the behavior of the volcano. For example, the inverse Gaussian hazard rate tends to a constant, or Poisson, behavior at long intervals, which is the open–closed conduit model proposed by Marzocchi and Zaccarelli (2006). On the other hand, the Weibull and gamma distributions can be of increasing or decreasing character and also include the Poisson process as a special case, which means they can be used to identify clustering or periodicity in the volcanic record. The mixture of Weibulls renewal model is capable of a complex variety of shapes, with both the Weibull renewal and Poisson processes nested within it, and thus can be used for more detailed investigation and modeling. Examples of renewal densities fitted to the onset data for Nyamuragira 1901–2010 (Wadge and Burt 2011), reproduced here as Table 2, are shown in Fig. 3 along with the derived hazard rate functions. We see that the mixture of Weibulls density is the only one that reproduces the observed features of the data and suggests a strong mode of eruption onsets separated by 2–3 years but with significant likelihood of longer reposes. The estimated hazard decreases if the mode is passed without an eruption occurring, before increasing again.

Time-predictable models It has been observed that the onset process of volcanic eruptions can exhibit more structure than is consistent with a simple Poisson

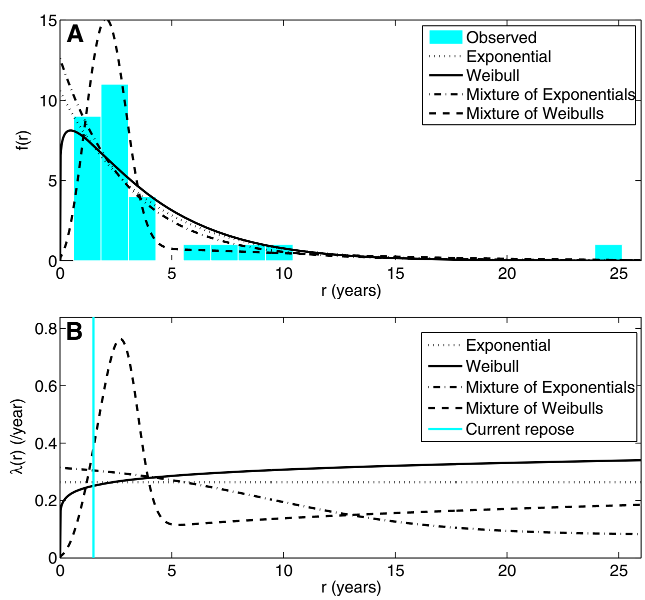

Fig. 3 Renewal distributions (**a**) and renewal model hazard rate (**b**) for the Nyamuragira onset data in Table 2, using various example distributions. Models are fitted by maximizing the likelihood Eq. 13. The model densities (*curves*) in **a** should reproduce the observed data (*histogram*), and give the likelihood of a given repose length under the model. The hazard rate *curves* in **b** show how the likelihood of an eruption changes with time elapsed since the last onset for each model. The *vertical line* shows the current repose length

process or a renewal process. Martin and Rose (1981) observed that the intervals between eruptions of Fuego Volcano (Guatemala) appeared to be proportional to the volume of the preceding eruption. Hill et al. (1998) successfully forecast the 1999 eruption of Cerro Negro Volcano in Nicaragua using time–volume relationships that were supported by petrogenetic models. This is most simply explained by assuming that the rate of magma input is constant, and an eruption occurs when a certain magma level is reached. Such behavior is termed "time-predictable."

If a central volcano can be thought of as a magma reservoir fed from below at a constant rate (Wadge 1982), erupting when a critical level is reached, the time-predictable model is appropriate. However, the critical level may vary due to factors such as the state of the conduit, presence or absence of a dome, tectonic influences, and geochemical and hydrological conditions (De la Cruz-Reyna 1991). Hence, there is a less than perfect correlation between the previous eruption volumes and the subsequent repose or, alternatively, the reposes have a probability distribution which is correlated with the size of the previous eruption.

This correlation was first described by Burt et al. (1994) who performed a regression analysis of $\{r_{i+1}\}$ on $\{v_i\}$. Sandri et al. (2005) generalized the test for time predictability to a regression analysis of $\{\log r_{i+1}\}$ on $\{\log v_i\}$, so that an estimated slope of b significantly different to zero implies a time-predictable relation $r_{i+1} \propto v_i^b$. A similar analysis of $\{\log v_i\}$ on $\{\log r_i\}$ can be conducted to see if there is any relation between the repose length and the subsequent eruption size. This is the size-predictable model, which will be discussed later. Examples of the analysis for the Nyamuragira data in Table 2 are shown in Fig. 4. The sequence appears to be time-predictable but not size-predictable, in agreement with the analysis of Wadge and Burt (2011) and Burt et al. (1994).

The logarithmic transformation in the repose-size regression analysis is also advisable to reduce the high leverage of the tail points due to both the repose and volume distributions usually being highly skewed. Additionally, the transformation leads directly to a stochastic model, as the residuals in a linear regression are assumed to be normally distributed. Hence, the underlying regression model in the Sandri et al. (2005) analysis is as follows:

$$\log r = a + b \log v + \epsilon \tag{23}$$

where ϵ is the normally distributed error. This implies that r has a log-normal distribution with density Eq. 19, where σ^2 is the variance of ϵ and

$$\mu = a + b \log v. \tag{24}$$

In this form, it was proposed as the *generalized time predictable model (GTPM)* by Marzocchi and Zaccarelli (2006). The density Eq. 19 with the link function Eq. 24 can be plugged into the renewal model, and the parameters a, b, and σ estimated via the usual

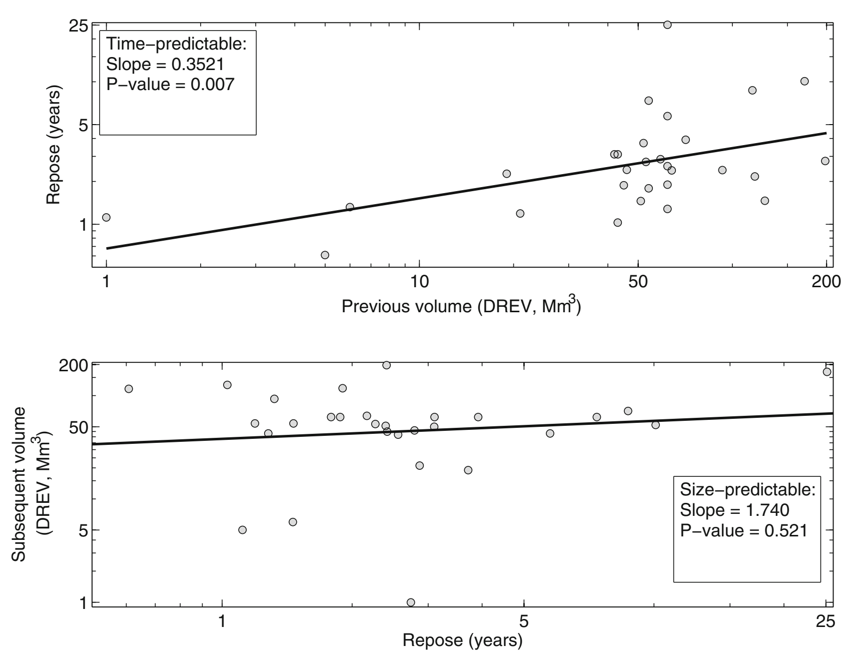

Fig. 4 Time (**a**) and size predictability (**b**) of the Nyamuragira eruption record from Wadge and Burt (2011). The *lines* are fitted by linear regression, and the *P* value gives the likelihood of the observed slope or greater under a null hypothesis of no trend. The time-predictable model in **a** assesses whether there is a correlation between the repose length and the previous eruption size, i.e., whether larger eruptions lead to longer reposes. The size-predictable model in **b** assesses similarly whether longer reposes are more likely to be terminated by larger eruptions

maximum likelihood procedures. Passarelli et al. (2010a) put the result within a Bayesian hierarchical model to allow the incorporation of information from additional covariates, such as eruption duration.

Markov processes Another way to introduce dependence on the previous eruption is to model the behavior via a continuous-time Markov chain, first proposed by Wickman (1966b). This entails defining a number of eruptive and repose states, with prescribed permitted transitions between them. The sojourn time in each state has an exponential distribution, and the time between onsets is the passage time through the eruptive and repose states, which will have a compound exponential distribution, i.e, be the sum of a number of exponential random variables of differing means. Hence, there are similarities to a mixture of gammas distribution model. The difference is that the different temporal distributions are explicitly identified in the Markov chain model, and hence the record of the volcano has to be matched to the states (e.g., Carta et al. 1981) from geologic data, although hidden Markov model techniques to automate the process have been suggested by Bebbington (2007).

Nonstationary methods

Renewal processes, including the homogeneous Poisson process and the time-predictable model are stationary in time in that the distribution of the number of events in an interval depends only on the length of the interval, not its location. In other words, events occur at the same average rate at all times, which is the third tenet of the Poisson process. More general models can incorporate a trend in the occurrence rate with time.

Guttorp and Thompson (1991) outlined a nonparametric method of estimating the occurrence rate over time using a kernel smoothing approach, allowing for incomplete observation. Time series methods can then be used to predict the future hazard rate and, hence, forecast the next eruption. However, the uncertainties rapidly grow beyond those of a parametric approach.

Nonhomogeneous Poisson processes

In a nonhomogeneous Poisson process, the number of events in the interval $(0, t]$ has a Poisson distribution with mean $\mu(t) = \int_0^t \lambda(s)ds$. Ho (1991) used an example of a nonhomogeneous Poisson process known as the *power law process*, with

$$\lambda(t) = \frac{\beta}{\theta}\left(\frac{t}{\theta}\right)^{\beta-1}. \tag{25}$$

This includes the homogeneous Poisson process as a special ($\beta = 1$) case, while if $\beta \neq 1$, the process is nonstationary. Note that despite the similarity in formulation, this is a very different model from the Weibull renewal process described above. The renewal process is stationary, with hazard depending on the time elapsed since the previous eruption. Here, the power law process is nonstationary, with hazard depending on the time from some fixed origin, which is not reset following an eruption. The hazard rate Eq. 25 is monotonic and, hence, can model either an increase or decrease in volcanic activity, but not both. There are a number of tests for whether the data are consistent with a power law process (Bebbington and Lai 1996a).

The parameter estimates in Eq. 25 are sensitive to the position of the time origin (Bebbington and Lai 1996a), and estimates of $\beta > 2$, which are quite feasible due to selection of the time origin and short length of records relative to the life span of the volcano, indicate a constantly accelerating, or convex, hazard rate. Neither of these properties is particularly desirable from a physical viewpoint, and hence the power law process is not suitable to model entire volcanic histories. Salvi et al. (2006) identified that the Mt. Etna flank eruption sequence was stationary before 1980 and constructed a composite model using the power law process for the post-1980 section. This characterization was formalized by Smethurst et al. (2009) using piecewise linear hazard rates, with a constant level of 0.11 onsets/year until 1964, after which the fitted hazard rate rises at 0.016/year per year.

Jaquet et al. (2000) (see also Jaquet and Carniel 2001) proposed a temporal occurrence model using the Cox process. This is a doubly stochastic process where the occurrence rate $\lambda(t)$ is itself a stochastic process, which has the advantage of allowing for correlation in time. The disadvantage is that the fitting and inference procedures are very complex, and the standard cautions regarding overparameterization, particularly on small or sparse datasets, apply.

There is a framework in which the inhomogeneous Poisson process and the renewal process models can be unified. Given an increasing function $\Psi(t)$ and a renewal distribution F, the *trend renewal process* (TRP) is defined by the values $r_i^\# = \Psi(t_{i+1}) - \Psi(t_i)$ being independent and identically distributed random variables with distribution function F. This preserves the independence property of the reposes, which is both statistically and physically attractive. If $\Psi(t) = \int_0^t \psi(s)ds$ for some function $\psi(s) \geq 0$ such that $\Psi(T) = T$, then $\psi(t)$ is a time scaling, taking values less (greater) than one where the onset rate is lower (higher) than average.

The model can be fitted to the observed data by a small modification of the likelihood Eq. 13,

$$L = \{1 - F[\Psi(T) - \Psi(t_n)]\} \prod_{i=1}^{n} f[\Psi(t_i) - \Psi(t_{i-1})]\psi(t_i). \tag{26}$$

We also see that the TRP is a stationary renewal process when $\psi(t) = 1$, while if $f = F'$ is the exponential density Eq. 14, then the TRP is a nonhomogeneous Poisson process with hazard rate $\lambda\psi(t)$. Hence, Bebbington (2010) used the Weibull density Eq. 15, which includes increasing and decreasing hazard rates, and trend functions including the power law Eq. 25, wax-and-wane, and cyclic terms, concluding that observed clustering is better explained via nonstationarity than between-event clustering. The speculated control of eruption recurrence patterns at Colima by magma differentiation processes (Luhr and Carmichael 1990) can also be modeled using this technique (Bebbington 2010).

While methods such as the AIC can identify which of multiple candidates is the model best fitting the data, they cannot assess whether or not the model is a good fit in absolute terms. To determine this, various goodness-of-fit tests, like the Kolmogorov–Smirnov one sample test and the chi-square test, can be applied. In particular, nonhomogeneous Poisson processes can be assessed for their absolute fit to the observed data by using the point-process compensator (Ogata 1988). This involves the scaling

$$\tau_i = \int_0^{t_i} \lambda(t)\mathrm{d}t \tag{27}$$

to produce the residual process times $\tau_1, \ldots, \tau_n$. If the model is a satisfactory fit to the observed record, then the residual process will be a Poisson process of unit rate. Standard tests then apply (e.g., Bebbington and Harte 2001).

Regimes

It has been suggested (Wickman 1966a; Wadge 1982; Marzocchi and Zaccarelli 2006) that volcanoes may exhibit different, but constant, levels of activity, rather than a trend with time. These *regimes* may represent changes in the eruption mechanism, the mechanism for transport of magma to the surface, or the eruptive style. If in each regime, the volcano exhibits stationarity, transitions between regimes are themselves random, and the regimes are recurrent, then the model will have no long-term trend.

Proposed means of identifying changes in regime include a running mean of time between onsets (Klein 1982), the theory of change-point problems (Mulargia et al. 1987), the cumulative count of eruptions in a statistical control chart (Ho 1992), and rank order statistics for the size of event (Pyle 1998), although none of them provide a means of forecasting future changes in regime.

A mixture model can be seen as a trivial example of a regime model, but the regime of each repose is being chosen randomly, independent of the previous and subsequent regimes, and hence, there is no temporal structure to the regimes. This was addressed by Cronin et al. (2001) in a hierarchical model where onsets occurred in "episodes", the number of eruptions in each episode being random, with different renewal distributions corresponding to interepisode and intraepisode reposes. The piecewise-constant model fitted to Mt. Etna flank eruptions by Smethurst et al. (2009) is also a regime model, although the regimes have a trend in this case.

More general structural dependence in the regimes was introduced by Bebbington (2007) using a hidden Markov model. The unobserved state represents the regime, which controls a renewal process, whose parameters are different in each regime. The number of hidden states required can be determined via AIC, which was shown to be more consistent than other alternative procedures. Regimes can be statistically identified via the Viterbi algorithm, which finds the most likely path through the hidden states, and as the regime operates as a Markov chain, switching (or not) after eruptions, future eruption probabilities are readily forecast. Bebbington (2007) also formulated an extension to a bivariate renewal model controlled by a hidden Markov chain, allowing for repose-size dependence.

Volume dependence

The time-predictable model has the repose length depending on the volume of only the most recent eruption. It is possible to generalize this, as postulated by De la Cruz-Reyna (1991) in his general load-and-discharge model.

The *volume-history (dependent) model* (Bebbington 2008) has

$$\lambda(t) = \exp\{\alpha + \nu[\rho t - V(t)]\}, \tag{28}$$

where $V(t) = \sum_{k:t_k<t} v_k$ is the cumulative volume erupted prior to time t. The parameter α incorporates the unknown state of the volcano at time 0. The

parameter estimates must be found by maximizing the point-process log likelihood

$$\log L = \sum_{i=1}^{n} \log \lambda(t_i) - \int_0^T \lambda(t)\mathrm{d}t, \qquad (29)$$

which rewards (first term) a large value of $\lambda(t)$ at an observed onset time, and penalizes (second term) large values of $\lambda(t)$ elsewhere. Bebbington (2008) fitted the model Eq. 28 to flank eruptions of Mt. Etna, identifying possible long-term quasicyclic behavior, and to Mauna Loa, finding a long-term decrease in activity. An example of the hazard rate estimated from the volume-history model for the Nyamuragira data in Table 2 is shown in Fig. 5, along with the estimate from the GTPM defined above, cf. Eq. 24. We see that the two models interpret the time evolution of the system in very different ways. According to the AIC criterion Eq. 8, the GTPM is the best fitting of the models considered for Nyamuragira in this paper. It can also be shown that the residual process Eq. 27 in the GTPM is stationary, although the untransformed onset times are not. Moreover, both the untransformed and transformed repose times are autocorrelated, which appears to be picking up the variation in stress-field controls identified by Wadge and Burt (2011).

Excitation process

Motivated by an observed "flare-up," with at least 30 of the 49 Auckland Volcanic Field events in the 250 ky record occurring within a 20-ky period, Bebbington and Cronin (2011) proposed a model with

$$\lambda(t) = \mu + \frac{\nu}{\sigma}\sqrt{\frac{2}{\pi}} \sum_{j:t_j<t} \exp\left[-\frac{(t-t_j)^2}{2\sigma^2}\right]. \qquad (30)$$

Thus, each eruption adds to the likelihood of a subsequent eruption, but this addition decays over time in a sigmoid fashion. This is an extreme example of a clustering process and can be fitted to the data by maximizing the point-process log likelihood Eq. 29. The estimated additional contribution from each previous event in the fitted model has a half-life on the order of 3,000 years but starts at a level four times that of the background rate μ.

Externally forced and coupled models

Volcanic eruptions do not occur in isolation. There is always a driver, be it only magma input from the mantle. The modeling techniques above examine the eruption time–volume recurrence data as a closed

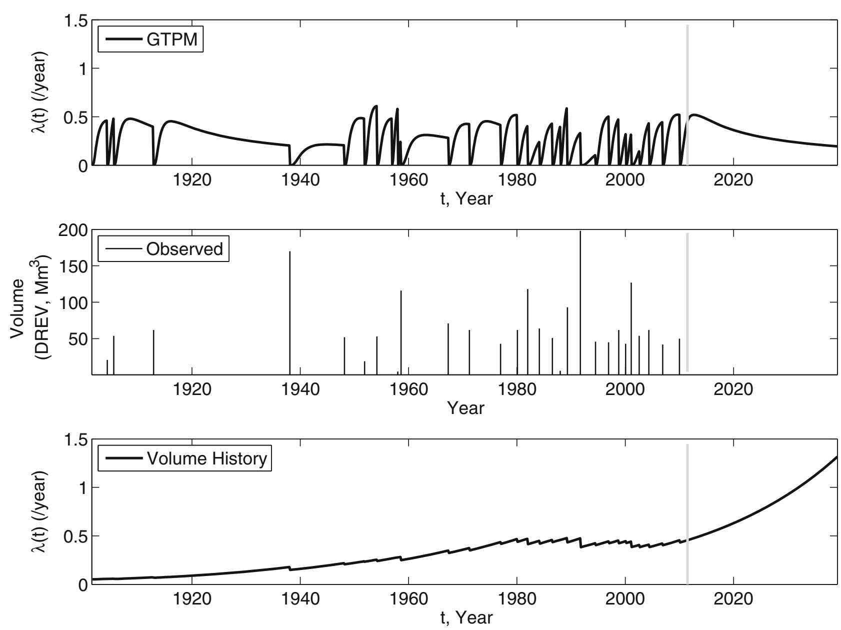

Fig. 5 Hazard rate estimates for Nyamuragira based on time and volume data. The *middle panel* shows the observed (time–volume) data. The GTPM hazard rate in the *top panel* and the volume-history hazard rate in the *bottom panel* are fitted by maximizing the point-process log likelihood Eq. 29. This rewards the model for having a large hazard rate at the time an eruption occurs, and penalizes it for having large hazard rates at other times. The volume–history model explains the increasing rate of eruptions as being due to a gradual increase in cumulative magma storage, while the GTPM explains it as being due to a correlation between individual eruption volumes and subsequent reposes

system including any magmatic drivers. However, other processes may perturb the observed eruption record.

Earthquake dependence

Static stress changes decay as $1/L^3$, and so changes from earthquakes resolved at dike locations tend to be small, with the exception of earthquakes within a few fault-lengths of the dike. However, physical mechanisms that explain how small changes in static stress, or larger but transient dynamic stress changes, can have observable effects on a volcanic system also imply the possibility of delayed triggering.

Many accounts (see Manga and Brodsky 2006; Bebbington and Marzocchi 2011, for reviews) have noted a causal effect on volcanic eruptions from large, not too distant, earthquakes, but none provided a model able to provide forecasts. Bebbington and Marzocchi, (2011) instead took a point-process modeling viewpoint, multiplying the hazard rate λ in some of the models above (the Poisson process, Weibull renewal process, GTPM, and volume-history model) by a time-, distance-, and magnitude-dependent triggering term. This was applied to volcanic eruptions (VEI 2+) and large earthquakes (Mw 7+) in the Indonesian arc since 1900. The method weighs both positive and negative (i.e., absence of eruptions following an earthquake) evidence of triggering. Of 35 volcanoes with at least three eruptions in the study region, seven (Marapi, Talang, Krakatau, Slamet, Ebulobo, Lewotobi, and Ruang) showed statistical evidence of triggering over varying temporal and spatial scales, but only after the eruption volume history of the volcano was accounted for, indicating that the earthquake triggering was only significant in cases where an eruption was already imminent due to the buildup of magma. However, those volcanoes with strong time-predictable behavior were not susceptible to earthquake triggering. Earthquake triggering appeared to be independent of the number of eruptions and their size.

Coupled volcanic sources

An extension (Bebbington 2008) to the volume-dependent eruption rate model (28) can be used to investigate the relation between summit and flank eruptions (cf., Takada 1997), the relation between neighboring volcanoes (cf., Klein 1982; Bebbington and Lai 1996b), or between different styles of eruption (cf., Turner et al. 2008b). The central idea is that eruptions of one class can extract magma from, or add magma to, the stored volume associated with other classes. In order to make the model at all tractable, these "transfer rates" are assumed to be constant proportions of the erupted volume, and thus we are seeking to measure an average tendency, a procedure akin to the regression fit Eq. 24 underlying the GTPM.

We suppose that the stored magma volume is divided among the classes, with the stored volume associated with class i evolving as

$$U_i(t) = U_i(0) + \rho_i t - \sum_j \theta_{ji} V^{(i)}(t) + \sum_{j \neq i} \theta_{ij} V^{(j)}(t) \qquad (31)$$

for $i = 1, \ldots, m$, where $V^{(j)}(t)$ is the cumulative volume erupted from class j, and θ_{ij} the proportion of erupted volume transferred from class j to class i ($\theta_{ii} = 1$), which may be positive (i.e., an eruption from class j transfers additional stored magma to class i) or negative (an eruption of class j partially drains the reservoir for class i). The ρ_i are the rates of magma input. Developing this model results in a point process with hazard rate

$$\lambda_i(t) = \exp\left[\alpha_i + \nu_i \left(\rho_i t - \sum_j \theta_{ji} V^{(i)}(t) + \sum_{j \neq i} \theta_{ij} V^{(j)}(t)\right)\right], \qquad (32)$$

for each class i, which can be fitted to data using a sum of point-process likelihoods Eq. 29.

Bebbington (2008) found significant evidence that summit eruptions of Mt. Etna are dependent on preceding flank eruptions, with both flank and summit eruptions being triggered by the other type. The model also found a marginally significant relationship between eruptions of Mauna Loa and Kilauea, consistent with the hypothesized invasion of the latter's plumbing system by magma from the former.

Forecasting eruption location and energetics

Long-term eruption location

Forecasting the location of the impending eruption in the case of central-vent volcanoes may appear trivial, even though there a spatial element may arise in the form of flank eruptions that are potentially dangerous and important (e.g., Salvi et al. 2006; Selva et al. 2012). On the other hand, spatial variability is probably more important than temporal variability in calderas or volcanic fields due to the infrequent nature of eruptions. This problem is often oversimplified by adopting a single scenario (i.e., selecting one source location) to assess the volcanic hazard; the use of a single scenario, even the most likely one, can lead to a gross under-estimation of the real hazard (Selva et al. 2010b). To date, spatial forecasting has been developed for long-term models (e.g., Connor and Hill 1995; Conway et al. 1998; Martin et al. 2004; Magill et al. 2005; Connor and Connor 2009; Bebbington and Cronin 2011; Selva et al. 2012), but very few efforts have been made to use the localization of monitoring anomalies for short-term spatial forecasts. In some cases, this problem may have a trivial solution; for example, a strong superficial seismic sequence may indicate where magma is going to erupt, but this pattern might only be observed shortly before the eruption, limiting drastically the scope for risk mitigation actions (e.g., Tokarev 1978).

Monogenetic volcanic fields

In monogenetic volcanic fields, where cones correspond to single eruptions or eruption sequences, the hazard is spatiotemporal. In this case, we extend the hazard rate notation to encompass the spatial dimension, so that $\lambda(t, x)\Delta t\, b(\Delta x)$ is approximately the probability of having an eruption in the time-space window $(t, t + \Delta t) \times (y : \|y - x\| < \Delta x)$, where $\|\cdot\|$ is the appropriate Euclidean norm and $b(\Delta x)$ denotes the measure (area in our case) of a circle of radius Δx. Often, however, the spatiotemporal hazard is calculated as the product of independent spatial and temporal terms $\lambda(t, x) = \lambda(t)\eta(x)$. The independence can be assumed, in the absence of sufficient data to test the hypothesis, or tested (e.g., Bebbington and Cronin 2011).

Models for spatiotemporal hazard are generally non-parametric or kernel-type. The *spatiotemporal nearest neighbor estimate* (Connor and Hill 1995) is calculated by inversely weighting the spatial density of volcanic centers according to the elapsed time since the eruption. More technically, suppose we have n centers, that the formation of the ith center occurred at time t_i, and that u_i is the area of a circle with radius equal to the distance between the point x of interest and the ith center. If $j = 1, \ldots, m$ indexes the mth nearest neighbors to the point x using the distance metric $u_i(t - t_i)$, then the estimated intensity at the point x is

$$\lambda(x, t) = \frac{m}{\sum_{j=1}^{m} u_j(t - t_j)}, \tag{33}$$

where m is the number of nearest neighbors used in the estimation, which is usually the number of centers, although Condit and Connor (1996) proposed a method for determining an optimal m. A *kernel estimate* of the spatial intensity is given by the following:

$$\eta(x) = \frac{1}{e_h} \sum_{i=1}^{n} \frac{\kappa_i}{h^2}, \tag{34}$$

where e_h is an edge correction,

$$\kappa_i = \frac{1}{h\sqrt{2\pi}} \exp\left[-(\mathrm{d}_i/h)^2\right] \tag{35}$$

is the Gaussian kernel (Conway et al. 1998), and d_i is the distance from x to the ith center. The parameter h is a smoothing constant, or "bandwidth." A small value of h concentrates the probability close to existing centers, while a large value distributes it more uniformly. This has far greater influence on the estimated hazard than the form of the kernel. Estimating the best value of h is generally done via some means of cross validation (e.g., Duong 2007). The kernel in Eq. 34 is isotropic, i.e., radially symmetric. An anisotropic kernel

$$\eta(x) = \frac{1}{2\pi\sqrt{H}} \sum_{i=1}^{n} \exp\left[-0.5(x - x_i)^T H^{-1}(x - x_i)\right] \tag{36}$$

has been suggested by Connor and Connor (2009) and implemented by Kiyosugi et al. (2010) for the Abu Monogenetic Volcanic Group and by Bebbington and Cronin (2011) for the Auckland Volcanic Field. In the latter case, the orientation of the fitted kernel appeared to be in accordance with the underlying tectonics, but in the former, it was not. The anisotropic estimate is sensitive to the criterion used to determine the kernel. Examples of both isotropic and anisotropic kernels fitted to vent locations from the Auckland Volcanic Field are shown in Fig. 6. The former is more robust to individual vent locations but provides little insight into possible directional controls on the spatial hazard.

There have also been more specific investigations of alignments in location of volcanic vents, in the hope that these can be related to geological features. In the *two-point azimuth method* (Lutz 1986), the azimuth between every two vents is measured, generating $n(n - 1)/2$ measurements for n vents. These are then binned in 10° intervals, and the result are tested for departure from randomness. A correction has to be made for noncircular fields and/or a nonhomogeneous density of points as these can produce a preferred orientation. This can be done via Monte Carlo simulation of random points from a spatial kernel density (Lutz and Gutmann 1995) to produce a reference distribution from which departures can be detected. In the *Hough transform* (Wadge and Cross 1988), a point (x, y) is converted into the normal parameterization (ρ, θ), where $\rho = x \cos\theta + y \sin\theta$. In this way, each point generates a curve in the (ρ, θ) plane as θ varies, and multiple (more than two curves) intersections of these define the

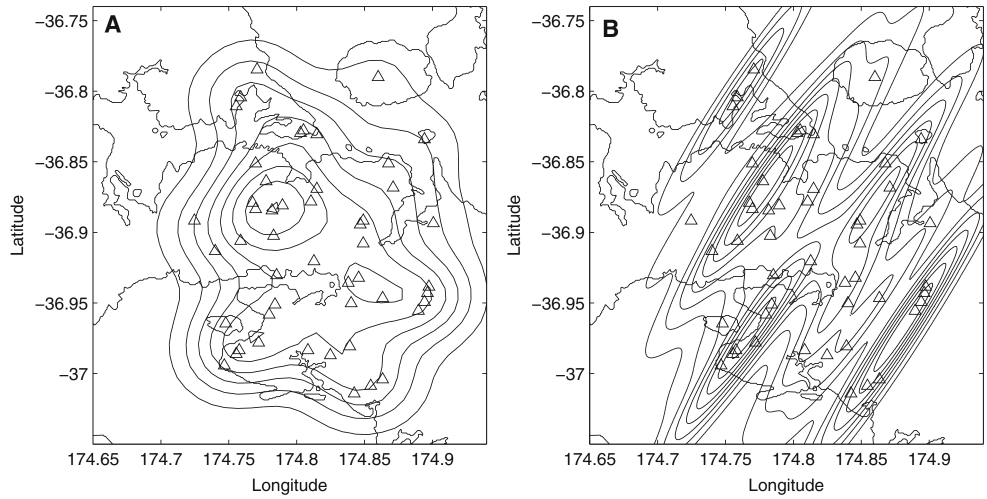

Fig. 6 Spatial hazard density for the Auckland Volcanic Field via kernel density estimation using a least squares cross-validation criterion. **a** Isotropic kernel from Eq. 34. **b** Anisotropic kernel from Eq. 36. *Contours* are at intervals of 0.0005

alignment of colinear points. Again, significance levels can be obtained via Monte Carlo techniques.

Martin et al. (2004) (see also Connor et al. 2000, and Weller 2004) embedded the spatiotemporal kernel into a Bayesian formulation to incorporate information from other geophysical data such as P-wave velocity perturbations, geothermal gradients, or gravity data. The spatiotemporal hazard is used as the prior, a likelihood function is generated by conditioning the geophysical data on the locations of volcanic events, and the result inverted via Bayes' Theorem to obtain the posterior estimate of the hazard.

A novel approach to spatiotemporal clustering was suggested by Magill et al. (2005) for the Auckland Volcanic Field. A statistic based on Ripley's K-function was used to determine a characteristic distance range for clustering, and the 49 centers were thus collapsed into 18 events using the eruption order. Separate stochastic models were then constructed for the spatial distribution within and between events. Unfortunately, in the absence of any age data, many neighboring centers were assigned consecutive places in the eruption order, introducing an element of circular reasoning. Using additional age determinations including inversion of tephra dispersal, Bebbington and Cronin (2011) found no dependence between the spatial and temporal aspects. However, the method remains valid and is a recommended starting point in any situation where a time hierarchy of events in space exists.

Polygenetic volcanoes

In general, spatial quantification of hazard on polygenetic volcanoes is of interest mainly if flank eruptions or collapses are possible or if multiple vents exist and provide distinct channels for lahars, lava, or pyroclastic flows. If the spatial dimension can be usefully categorized into a small number of possibilities, a simple Markov chain approach can be used (e.g., Cronin et al. 2001).

A number of studies have considered the relationship between summit and flank eruptions. Klein (1982) showed that the classification of the previous event at Kilauea provided no information about the time of the next eruption, but at Mauna Loa, reposes following flank eruptions were significantly longer than those following summit eruptions. While Kilauea summit eruptions tend to cluster in time due solely to the long summit sequence of 1924–1954, Mauna Loa displays no tendency for clustering or alternation of summit and flank eruptions. However, Bebbington (2008) found that summit and flank eruptions at Mt. Etna have a statistical tendency to trigger an event of the other type.

Salvi et al. (2006) considered the azimuth distribution of flank eruptions from Mt. Etna, detecting no significant difference in the pre- and post-1536 distributions, although Smethurst et al. (2009) did detect a change in the spatial pattern of flank eruptions following the formation of the South East Crater in 1971. Wadge et al. (1994) created a spatial density map for the opening of a flank vent and simulated the resulting lava hazard.

Calderas and rift zones

In general, the methods described above for monogenetic fields are applicable in varying degrees to calderas. In the case of rift zones, such as Taveuni, the spatial location can be reduced in dimension to the distance along the rift (Cronin et al. 2001). Using a Markov chain approach, Eliasson et al. (2006) divided the Katla caldera, Iceland, into three sectors to examine the spatial progression of volcanogenic flood events.

A different approach has been recently proposed by Selva et al. (2012) and applied to the spatial forecast of the next eruption at Campi Flegrei. The method relies on Bayesian inference to merge prior information and data of past activity. The possible vent openings are binned into a grid of n cells covering the whole caldera, on which is estimated a probability distribution of vent opening through a Dirichlet distribution which generalizes the beta distribution (see Eq. 1). The choice of this distribution implies that the cells represent a set of mutually exclusive and exhaustive possibilities, i.e., the eruption can occur only in one specific cell and the possibility of multiple simultaneous vents is neglected. The posterior Dirichlet distribution is obtained as

$$\begin{aligned}\Theta_{post} &= Di_n(\alpha_1 + y_1, \ldots, \alpha_n + y_n) \\ &= Di_n(E[\Theta_1](\Lambda + n - 1) \\ &\quad + y_1, \ldots, E[\Theta_n](\Lambda + n - 1) + y_n)\end{aligned} \tag{37}$$

where y_j is the number of times that a past eruption has occurred in the j-th cell and α_j, $j = 1, \ldots, n$ are the parameters of the prior Dirichlet distribution. The right-hand side describes the same distribution in terms of simpler and more intuitive parameters, i.e., the central values $E[\Theta_j](j = 1, ..., n)$ and the equivalent number of data Λ (Marzocchi et al. 2008)

$$\alpha_j = E[\Theta_j](\Lambda + n - 1). \tag{38}$$

Equation 38 shows that the parameter Λ is related to the parameters of the Dirichlet distribution, but it has a simpler practical interpretation. In fact, it quantifies the worth of the prior information in terms of equivalent number of data; specifically, $\Lambda = 1$ means that the weight of the prior information has the same importance of one single real datum. In practice, the averages $E[\Theta_j]$ represent the best-guess probabilities, while Λ is a measure of the dispersion around the average, quantifying our confidence regarding the best-guess probabilities (epistemic uncertainty; see Marti et al. 2008). The parameter Λ is a single value belonging to the entire prior model, not to the individual locations, and is a measure of the confidence in the prior distribution that is simpler and more intuitive than the usual variance. The larger Λ, the greater the confidence in the reliability of the prior model, so that the number of past data needed to significantly modify the prior must be larger. Conversely, if Λ is small, even a small number of past data can drastically modify the prior. In the case of maximum ignorance, Λ is set to 1 and $E[\Theta_j]$ to $1/n$, resulting in a uniform distribution.

In the case of Campi Flegrei, Selva et al. (2012) defined the prior as

$$E[\Theta_j] = \frac{W_j}{\sum\limits_{k=1}^{n} W_k}, \tag{39}$$

and $\Lambda = 1$, where W_j is a parameter that takes into account the mechanical weakness of the floor of the cell. It is high where faults and/or past vents are present and low for the intact caldera floor (more details can be found in Selva et al. 2012). This prior information has been incorporated in the posterior distribution using the locations of the most recently active vents (Eq. 37). As with the kernel estimates covered above under monogenic fields, this procedure implies that cells hosting a high number of past eruptions are more likely to be the location of future events.

Eruption energetics

Forecasting the size of an impending eruption is probably one of the most controversial scientific issues, with a huge impact on decision-making (Marzocchi et al. 2004). In the long-term perspective, the problem has been approached using the concept of "maximum expected event" (e.g., Rosi 1996; Mastrolorenzo et al. 2006), or "worst credible event." The former definition is unsatisfactory from a scientific point of view, as it is impossible to define a maximum expected event from many natural systems that are often characterized by power law distribution (e.g., Bak et al. 1987). Similarly, worst credible event depends on the meaning of "credible," The essence of probabilistic hazard analysis is that any size of eruption has a specific probability of occurrence which can be estimated (e.g., Bebbington et al. 2008; Orsi et al. 2009); the choice of what size of eruption becomes "not credible" thus imposes the choice of a probability threshold. This selection has many similarities with the choice of an "acceptable" risk; we believe that the choice of what is credible or acceptable (or of any other probability threshold) is not a scientific issue. The task of the scientist is to estimate the full range of possible events and their probability of occurrence and pass these on to the decision-makers for action (e.g., Marzocchi and Woo 2007, 2009; Sandri et al. 2012). To do otherwise is for the scientists to usurp the functions of the decision-maker.

For short-term forecasting, no monitoring parameter(s) that indicates the size of the impending eruption has been identified to date. A tutorial example was given by the Pinatubo eruptions in the early 1990s, where a similar preeruptive pattern was detected both before the June 1991 event and before a much smaller eruption which occurred the year following (Ramos et al. 1996).

VEI, magnitude, volume, and intensity

The size of volcanic eruptions can be summarized through different quantities such as the volcanic explosivity index (VEI; Newhall and Self 1982), intensity, magnitude, and volume. Here, we report the definition of these quantities as found in the *Encyclopedia of Volcanoes* (Sigurdsson et al. 2000) and *Volcanoes of the World* (Simkin and Siebert 1994).

- *Intensity* A measure of the rate at which magma is discharged during an eruption. It is measured in kilograms per second.
- *Magnitude* The total mass of material ejected during an eruption. It is measured in kilograms. Pyle (2000) proposed a magnitude scale analogous to that for earthquakes as magnitude $=$ $\log_{10}$ erupted mass $-$ 7.
- *Volume* The total volume of material ejected during an eruption. It is measured in cubic meters.
- *VEI* A crude measure of the potential impact of an eruption on the atmosphere, without consideration of the sulfer gas release, based, in order, on tephra volume, column height, “explosivity,” eruption type and duration. It is measured on a scale from 0 to 8.

Notwithstanding the more subjective nature of the VEI measure, it is the most widely used because it can, and has been, assessed for most known eruptions. This is difficult, where even possible, for the other measures. However, while VEI is commonly used as a magnitude scale for explosive eruptions, it does not measure well the magnitude of nonexplosive eruptions, as it is not a mass-based scale.

Statistical distributions

As for the size of many other natural events, such as earthquakes, floods, landslides, etc., the global catalog seems to show a power law distribution between the number of events and their magnitude/volume/intensity (Simkin and Siebert 1994).

$$\log(N_i) = a - b\,\log(S_i) \tag{40}$$

where N_i is the number of events with magnitude/volume/intensity S_i, and a and b are parameters. As for the Gutenberg–Richter in seismology, the distribution Eq. 40 becomes exponential when the VEI is taken into consideration, or we use the logarithm of the volume or intensity. While this distribution is widely accepted for global activity, significant doubts persist about what is the best distribution for each single volcano, mirroring the debate in seismology about characteristic magnitudes for individual fault segments. In some cases, past activity seems to fit well power law distributions, like at Campi Flegrei (Orsi et al. 2009), Vesuvio (Marzocchi et al. 2004) and Mt. Taranaki (Bebbington et al. 2008). However, in other cases like at Miyakejima Volcano, the volume distribution appears much more regular, being characterized by a preferred size (Garcia-Aristizabal et al. 2012). We note that the hypothesis of a power law distribution requires many data to be adequately tested, particularly in the tail describing the largest events, and is susceptible to bias from differential reporting rates in geological and historical records. Extreme value distributions for the eruption size have been inverted to estimate catalog completeness (Deligne et al. 2010; Furlan 2010).

Size-predictable models

Estimating hazard using an event size drawn independently from a power law (Bebbington et al. 2008) or extreme value distribution (Mendoza-Rosas and De la Cruz-Reyna 2008) is common, but very little research has appeared regarding the possible size of future events conditioned on the prior eruptive history. Burt et al. (1994), and Wadge and Burt (2011) formulated the size-predictable model as a regression analysis of the subsequent volume $\{v_i\}$ on the length of the repose $\{r_i\}$ but found it described the behavior of Nyamuragira less well than the time-predictable model (see also Fig. 4). A global analysis by Marzocchi and Zaccarelli (2006), generalizing the regression analysis to one of $\{\log v_i\}$ on $\{\log r_i\}$, reached the same conclusion. Bebbington (2007) introduced a coupling between the distributions of the repose and subsequent eruption duration via a hidden Markov model. Tested on flank eruptions from Mt. Etna, it likewise provided less information than the alternative eruption duration-subsequent repose coupling.

An open research direction is whether other ancillary data can be used to constrain the size-prediction problem. For example, Wadge and Burt (2011) tested the relationship over the last 110 years between the local stress field and 30 flank eruptions at Nyamuragira. Eruptions fed by dikes parallel to the East African Rift Valley were found to have longer durations (and larger volumes) than those fed by other dikes. The intrusion of a major dike during the 1977 volcano–tectonic event at neighboring Nyiragongo Volcano appeared to have changed the system dynamics, and since then, most eruptions have been of short-duration fed by dikes perpendicular to the Rift. A subsequent volcano–tectonic event in 2002 may have resulted in a further change.

Future research directions

The primary means for scientists to abate natural risks is through reducing uncertainties. Uncertainties cannot be completely eliminated because natural processes usually own an intrinsic unpredictability (aleatory uncertainty), but they can be reduced significantly through the development of more reliable and skilled forecasting models. A careful analysis of the past and recent literature on eruption forecasting highlights the need to make more efforts on all fronts. Nonetheless, some gaps appear more evident.

First, the most urgent (and conceptually simple) need is that of obtaining more and more reliable data to build and test theoretical and empirical models. While statistics is always called upon to make forecasts from less data than desired (or at times, even sufficient) in volcanology, the uneven coverage of, lack of homogeneity in, and inaccessibility of, the data is particularly evident and poses barriers to improvement in eruption forecasting science.

Secondly, and dependent on more data being assembled, there is an urgent need to test the power and applicability of various forecasting models on a wide range of volcanoes. At present, methods are developed based on one, or a few, volcanoes, and we lack any evidence as to whether they are portable even to geologically similar settings.

Finally, and from a more scientific point of view, the largest gap is in forecasting the size of future/impending eruptions. Current best practice is to use an independent power law or extreme value distribution, and the monitoring data do not seem to carry discernable information constraining the dimension of the impending eruption. We feel that this represents the most important future research direction in this field. A further promising avenue is in the investigation of how additional measurements, for example, from covariates of the eruptions themselves or from earthquakes (e.g., Bebbington and Marzocchi, 2011), tectonic models, GPS deformation etc., can be mined for information concerning future eruption times and sizes.

Conclusion

At the present state of knowledge, the use of probability is unavoidable in order to properly account for uncertainties in eruption forecasts. We are optimistic that precursors can be generalized and recast into a more suitable probabilistic forecast. Probabilistic forecasts can be more rationally used in societal decision making to select mitigation actions (Marzocchi and Woo 2007, 2009; Woo 2008) in a way that more generic or deterministic predictions cannot, as probabilistic forecasts can properly take uncertainties into account. This benefit is particularly important for managing the evolution of an unrest phase in high-risk volcanoes, where mitigation actions, such as evacuation, require some lead time before the eruption but also incur considerable costs and may result in unacceptable loss of life and property if incorrectly applied.

Acknowledgements The authors thank Geoff Wadge and Britt Hill for their painstaking and helpful reviews of the manuscript content and style, and editor Stephen Self for careful checking of, and additional suggestions on, the result. Commitments elsewhere forced Willy Aspinall to withdraw from the authoring team, after making a number of contributions to the paper organization.

References

Akaike H (1977) On entropy maximization principle. In: Krishnaiah PR (ed) Applications of statistics. North-Holland, Amsterdam, pp 27–41

Amoruso A, Crescentini L (2009) Shape and volume change of pressurized ellipsoidal cavities from deformation and seismic data. J Geophys Res 114:B02210

Amoruso A, Crescentini L (2011) Modelling deformation due to a pressurized ellipsoidal cavity, with reference to the Campi Flegrei caldera, Italy. Geophys Res Lett 38:L01303

Amoruso A, Crescentini L, Berrino G (2008) Simultaneous inversion of deformation and gravity changes in a horizontally layered half-space: evidences for magma intrusion during the 1982–1984 unrest at Campi Flegrei caldera (Italy). Earth Plan Sci Lett 272:181–188

Aspinall W (2006) Structured elicitation of expert judgement for probabilistic hazard and risk assessment in volcanic eruptions. In: Mader HM, Coles SG, Connor CB, Connor LJ (eds) Statistics in volcanology. IAVCEI Publications. ISBN 978-1-86239-208-3, pp 15–30

Aspinall WP, Woo G, Voight B, Baxter PJ (2003) Evidence-based volcanology: application to eruption crises. J Volcanol Geotherm Res 128:273–285

Aspinall W, Carniel R, Jaquet O, Woo G, Hincks T (2006) Using hidden multi-state Markov models with multi-parameter volcanic data to provide empirical evidence for alert level decision-support. J Volcanol Geotherm Res 153:112–124

Bak P, Tang C, Wiesenfeld K (1987) Self-organized criticality: an explanation of the 1/f noise. Phys Rev Lett 59:381–384

Battaglia M, Hill DP (2009) Analytical modeling of gravity changes and crustal deformation at volcanoes: the Long Valley caldera, California, case study. Tectonophysics 471:45–57

Battaglia M, Roberts C, Segall P (2003) The mechanics of unrest at Long Valley caldera, California: 2. Constraining the

nature of the source using geodetic and micro-gravity data. J Volcanol Geotherm Res 127:219–245
Battaglia J, Ferrazzini V, Staudacher T, Aki K, Cheminee J (2005) Pre-eruptive migration of earthquakes at the Piton de la Fournaise volcano (Reunion Island). Geophys J Int 161:549–558
Bebbington MS (2007) Identifying volcanic regimes using hidden Markov models. Geophys J Int 171:921–942
Bebbington M (2008) Incorporating the eruptive history in a stochastic model for volcanic eruptions. J Volcanol Geotherm Res 175:325–333
Bebbington M (2009) Volcanic eruptions: stochastic models of occurrence patterns. In: Meyers B (ed) Encyclopedia of complexity and system science, vol 9. Springer, New York, pp 9831–9861
Bebbington M (2010) Trends and clustering in the onsets of volcanic eruptions. J Geophys Res 115:B01203
Bebbington M, Cronin SJ (2011) Spatio-temporal hazard estimation in the Auckland Volcanic Field, New Zealand, with a new event-order model. Bull Volcanol 73:55–72
Bebbington M, Harte D (2001) On the statistics of the linked stress release process. J Appl Probab 38A:176–187
Bebbington MS, Lai CD (1996a) On nonhomogeneous models for volcanic eruptions. Math Geol 28:585–600
Bebbington MS, Lai CD (1996b) Statistical analysis of New Zealand volcanic occurrence data. J Volcanol Geotherm Res 74:101–110
Bebbington MS, Lai CD (1998) A generalised negative binomial and applications. Commun Stat, Theory Methods 27:2515–2533
Bebbington MS, Marzocchi, W (2011) Stochastic models for earthquake triggering of volcanic eruptions. J Geophys Res 116:B05204
Bebbington M, Cronin S, Chapman I, Turner M (2008) Quantifying volcanic ash fall hazard to electricity infrastructure. J Volcanol Geotherm Res 177:1055–1062
Bedford T, Cooke R (2001) Probabilistic risk analysis: foundations and methods. Cambridge University Press, Cambridge
Bell A, Naylor M, Heap M, Main I (2011) Forecasting volcanic eruptions and other material failure phenomena: an evaluation of the failure forecast method. Geophys Res Lett 38:L15304
Benoit JP, McNutt SR, Barboza V (2003) Duration-amplitude distribution of volcanic tremor. J Geophys Res 108:2146
Bommer J, Scherbaum F (2008) The use and misuse of logic trees in probabilistic seismic hazard analysis. Earthq Spectra 24:9970–1009
Bonaccorso A, Aloisi M, Mattia M (2002) Dike emplacement forerunning the Etna July 2001 eruption modeled through continuous tilt and GPS data. Geophys Res Lett 29:1624
Bonafede M (1991) Hot fluid migration: an efficient source of ground deformation—application to the 1982–84 crisis at Phlegraean Fields, Italy. J Volcanol Geotherm Res 48:187–198
Brenguier F, Shapiro N, Campillo M, Ferrazzini V, Duputel Z, Coutant O, Nercessian A (2008) Towards forecasting volcanic eruptions using seismic noise. Nat Geosci 1:126–130
Bryan CJ, Sherburn S (2003) Eruption-induced modifications to volcanic seismicity at Ruapehu, New Zealand, and its implications for eruption forecasting. Bull Volcanol 65:30–42
Budnitz RJ, Apostolakis G, Boore DM, Cluff LS, Coppersmith KJ, Cornell CA, Morris PA (1997) Senior Seismic Hazard Analysis Committee; recommendations for probabilistic seismic hazard analysis: guidance on uncertainty and use of experts, vol 1–2. U.S. Nuclear Regulatory Commission, U.S. Dept. of Energy, Electric Power Research Institute; NUREG/CR-6372, UCRL-ID-122160
Burt ML, Wadge G, Curnow RN (2001) An objective method for mapping hazardous flow deposits from the stratigraphic record of stratovolcanoes: a case example from Montange Pelée. Bull Volcanol 63:98–111
Burt ML, Wadge G, Scott WA (1994) Simple stochastic modelling of the eruption history of a basaltic volcano: Nyamuragira, Zaire. Bull Volcanol 56:87–97
Caniaux G (2005) Statistical analysis of the volcanic eruption frequency in the Azores Islands: a contribution to risk assessment. Bull Soc Géol Fr 176:107–120
Carta S, Figari R, Sartoris G, Sassi R, Scandone R (1981) A statistical model for Vesuvius and its volcanological implications. Bull Volcanol 44:129–151
Castro JM, Dingwell DB (2009) Rapid ascent of rhyolitic magma at Chaitén volcano, Chile. Nature 461:780–784
Chastin SFM, Main IG (2003) Statistical analysis of daily seismic event rate as a precursor to volcanic eruptions. Geophys Res Lett 30:1671
Chouet B (1996) Long-period volcano seismicity: its source and use in eruption forecasting. Nature 380:309–316
Claeskens G, Hjort NL (2008) Model selection and model averaging. Cambridge University Press, Cambridge
Condit CD, Connor CB (1996) Recurrence rates of volcanism in basaltic volcanic fields: an example from the Springerville volcanic field, Arizona. Geol Soc Am Bull 108:1225–1241
Connor CB, Connor LJ (2009) Estimating spatial density with kernel methods. In: Connor CB, Chapman NA, Connor LJ (eds) Volcanic and tectonic hazard assessment for nuclear facilities. Cambridge University Press, Cambridge, UK, pp 346–368
Connor CB, Hill BE (1995) Three nonhomogeneous Poisson models for the probability of basaltic volcanism: application to the Yucca Mountain region, Nevada. J Geophys Res 100:10107–10125
Connor CB, Stamatakos JA, Ferrill DA, Hill BE, Ofoegbu G, Conway FM, Sagar B, Trapp JS (2000) Geologic factors controlling patterns of small-volume basaltic volcanism: application to a volcanic hazards assessment at Yucca Mountain, Nevada. J Geophys Res 105:417–432
Connor CB, Sparks RSJ, Mason RM, Bonadonna C, Young SR (2003) Exploring links between physical and probabilistic models of volcanic eruptions: the Soufriere Hills volcano, Montserrat. Geophys Res Lett 30:1701
Conway FM, Connor CB, Hill BE, Condit CD, Mullaney K, Hall CM (1998) Recurrence rates of basaltic volcanism in SP cluster, San Francisco volcanic field, Arizona. Geology 26:655–658
Cooke RM (1991) Experts in uncertainty: opinion and subjective probability in science. Oxford Univ Press, Oxford
Cornelius RR, Voight B (1996) Real-time seismic amplitude measurement (RSAM) and seismic spectral amplitude measurement (SSAM) analyses with the materials failure forecast method (FFM), June 1991 explosive eruption at Mount Pinatubo. In: Newhall CG, Punongbayan RS (eds) Fire and mud, eruptions and lahars of Mount Pinatubo Philippines, pp 249–267
Cox DR, Lewis PAW (1966) The statistical analysis of series of events. Methuen, London
Cronin S, Bebbington M, Lai CD (2001) A probabilistic assessment of eruption recurrence on Taveuni volcano, Fiji. Bull Volcanol 63:274–288
Crowe BM, Johnson ME, Beckman RJ (1982) Calculation of the probability of volcanic disruption of a high-level radioactive

waste repository within southern Nevada, USA. Radioact Waste Manage 3:167–190
Crowe BM, Wallmann P, Bowker LM (1998) Probabilistic modeling of volcanism data: Final volcanic hazard studies for the Yucca Mountain site. In: Perry FV et al (eds) Volcanism studies: final report for the Yucca Mountain project. Los Alamos National Laboratory Report LA-13478, Los Alamos National Laboratory, Los Alamos, NM, p 415
Davis PM (1986) Surface deformation due to inflation of an arbitrarily oriented triaxial ellipsoidal cavity in an elastic half-space, with reference to Kilauea Volcano, Hawaii. J Geophys Res 91:7429–7438
Decker RW (1986) Forecasting volcanic eruptions. Ann Rev Earth Planet Sci 14:267–291
De la Cruz-Reyna S (1991) Poisson-distributed patterns of explosive eruptive activity. Bull Volcanol 54:57–67
De la Cruz-Reyna S (1993) Random patterns of occurrence of explosive eruptions at Colima volcano, Mexico. J Volcanol Geotherm Res 55:51–68
De la Cruz-Reyna S, Carrasco-Nunez G (2002) Probabilistic hazard analysis of Citlaltepetl (Pico de Orizaba) volcano, eastern Mexican volcanic belt. J Volcanol Geotherm Res 113:307–318
Deligne NI, Coles SG, Sparks RSJ (2010) Recurrence rates of large explosive volcanic eruptions. J Geophys Res 115:B06203. doi:10.1029/2009JB006554
Duong T (2007) ks: kernel density estimations and kernel discriminant analysis for multivariate data in R. J Statist Softw 21(7):1–16
Dzierma Y, Wehrmann H (2010) Eruption time series statistically examined: Probabilities of future eruptions at Villarrica and Llaima Volcanoes, Southern Volcanic Zone, Chile. J Volcanol Geotherm Res 193:82–92
Dzurisin D (2003) A comprehensive approach to monitoring volcano deformation as a window on the eruption cycle. Rev Geophys 41:1001
Eliasson J, Larsen G, Gudmundsson MT, Sigmundsson F (2006) Probabilistic model for eruptions and associated flood events in the Katla caldera, Iceland. Comput Geosci 10:179–200
Field EH, Dawson TE, Felzer KR, Frankel AD, Gupta V, Jordan TH, Parsons T, Petersen MD, Stein RS, Weldon RJ, Wills CJ (2007) The uniform California earthquake rupture forecast, Version 2 (UCERF 2). USGS Open File Report 2007–1437
Fournier T J, Pritchard ME, Riddick SN (2010) Duration, magnitude, and frequency of subaerial volcano deformation events: new results from Latin America using InSAR and a global synthesis. Geochem Geophys Geosyst 11:Q01003
Furlan, C (2010) Extreme value methods for modelling historical series of large volcanic magnitudes. Statist Model 10:113–132
Garcia-Aristizabal A, Marzocchi W, Fujita E (2012) A Brownian model for recurrent volcanic eruptions: an application to Miyakejima volcano (Japan). Bull Volcanol 74:545–558
Gelman A, Carlin JB, Stern HS, Rubin DB (1995) Bayesian data analysis. CRC, Boca Raton, FL
Gerst A, Savage M (2004) Seismic anisotropy beneath Ruapehu Volcano: a possible eruption forecasting tool. Science 306:1543–1547
Gillies D (2000) Philosophical theories of probability. New York, Routledge
Grasso JR, Zaliapin I (2004) Predictability of volcano eruption: Lessons from a basaltic effusive volcano. Geophys Res Lett 31:L05602
Guidoboni E, Ciuccarelli C (2011) The Campi Flegrei caldera: historical revision and new data on seismic crises, bradyseisms, the Monte Nuovo eruption and ensuing earthquakes (twelfth century 1582 AD). Bull Volcanol 73:655–677
Guttorp P, Thompson ML (1991) Estimating second-order parameters of volcanicity from historical data. J Am Statist Assoc 86:578–583
Hainzl S, Ogata Y (2005) Detecting fluid signals in seismicity data through statistical earthquake modeling. J Geophys Res 110:B05S07
Hammer C, Neuberg J (2009) On the dynamical behavior of low-frequency earthquake swarms prior to a dome collapse of Soufrière Hill volcano, Montserrat. Geophys Res Lett 36:L06305
Hill BE, Connor CB, Jarzemba MS, La Femina PC, Navarro M, Strauch W (1998) 1995 eruptions of Cerro Negro volcano, Nicaragua, and risk assessment for future eruptions. Geol Soc Am Bull 110:1231–1241
Ho CH (1990) Bayesian analysis of volcanic eruptions. J Volcanol Geotherm Res 43:91–98
Ho CH (1991) Nonhomogeneous Poisson model for volcanic eruptions. Math Geol 23:167–173
Ho CH (1992) Statistical control chart for regime identification in volcanic time-series. Math Geol 24:775–787
Hurvich CM, Tsai CL (1989) Regression and time series model selection in small samples. Biometrika 76:297–307
Iguchi M, Yakiwara H, Tameguri T, Hendrasto M, Hirabayashi JI (2008) Mechanism of explosive eruption revealed by geophysical observations at the Sakurajima, Suwanosejima and Semeru volcanoes. J Volcanol Geotherm Res 178:1–9
Jaquet O, Carniel R (2001) Stochastic modelling at Stromboli: a volcano with remarkable memory. J Volcanol Geotherm Res 105:249–262
Jaquet O, Low S, Martinelli B, Dietrich V, Gilby D (2000) Estimation of volcanic hazards based on Cox stochastic processes. Phys Chem Earth (A) 25:571–579
Jaquet O, Carniel R, Sparks S, Thompson G, Namar R, Dicecca M (2006) DEVIN: a forecasting approach using stochastic methods applied to the Soufriere Hills Volcano. J Volcanol Geotherm Res 153:97–111
Jellinek AM, Bercovici D (2011) Seismic tremors and magma wagging during explosive volcanism. Nature 470:522–526. doi:10.1038/nature09828
Jeffreys H (1961) Theory of probability, 3rd edn. Oxford University Press, Oxford
Jiang R, Murthy DNP (1998) Mixture of Weibull distributions—parametric characterization of failure rate function. Appl Stoc Models Data Anal 14:47–65
Jordan TH (2006) Earthquake predictability, brick by brick. Seismol Res Lett 77:3–6
Jordan TH, Chen Y-T, Gasparini P, Madariaga R, Main I, Marzocchi W, Papadopoulos G, Sobolev G, Yamaoka K, Zschau J (2011) Operational earthquake forecasting: state of knowledge and guidelines for implementation. Ann Geophys 54:315–391. doi:10.4401/ag-5350
Kiyosugi K, Connor CB, Zhao D, Connor LJ, Tanaka K (2010) Relationships between volcano distribution, crustal structure, and P-wave tomography: an example from the Abu Monogenetic Volcano Group, SW Japan. Bull Volcanol 72:331–340
Klein FW (1982) Patterns of historical eruptions at Hawaiian volcanoes. J Volcanol Geotherm Res 12:1–35
Kilburn C (2003) Multiscale fracturing as a key to forecasting volcanic eruptions. J Volcanol Geotherm Res 125:271–289
Kilburn CRJ, Voight B (1998) Slow rock fracture as eruption precursor at Soufriere Hills volcano, Montserrat. Geophys Res Lett 25:3665–3668
Lavallée Y, Meredith P, Dingwell D, Hess K, Wassermann J, Cordonnier B, Gerik A, Kruhl J (2008) Seismogenic lavas and explosive eruption forecasting. Nature 453:507–510

Linde AT, Ágústsson K, Sacks IS, Stefánsson R (1993) Mechanism of the 1991 eruption of Hekla from continuous borehole strain monitoring. Nature 365:737–740
Lindsay J, Marzocchi W, Jolly G, Constantinescu R, Selva J, Sandri L (2010) Towards real-time eruption forecasting in the Auckland Volcanic Field: application of BET_EF during the New Zealand national disaster exercise 'Ruaumoko'. Bull Volcanol 72:185–204
Lombardi AM, Marzocchi W, Selva J (2006) Exploring the evolution of a volcanic seismic swarm: the case of the 2000 Izu Islands swarm. Geophys Res Lett 33:L07310
Luhr JF, Carmichael ISE (1990) Petrological modeling of cyclical eruptive activity at Volcan Colima, Mexico. J Volcanol Geotherm Res 42:235–260
Lutz TM (1986) An analysis of the orientation of large-scale crustal structures: a statistical approach based on areal distributions of pointlike features. J Geophys Res 91:421–434
Lutz TM, Gutmann JT (1995) An improved method for determining and characterizing alignments of point-like features and its implications for the Pinacate volcanic field, Sonora, Mexico. J Geophys Res 100:17659–17670
Magill CR, McAneney KJ, Smith IEM (2005) Probabilistic assessment of vent locations for the next Auckland volcanic field event. Math Geol 37:227–242
Manga M, Brodsky, E (2006) Seismic triggering of eruptions in the far field: volcanoes and geysers. Ann Rev Earth Planet Sci 34:263–291
Marti J, Aspinall WP, Sobradelo R, Felpeto A, Geyer A, Ortiz R, Baxter P, Cole P, Pacheco J, Blanco MJ, Lopez C (2008) A long-term volcanic hazard event tree for Teide-Pico Viejo stratovolcanoes (Tenerife, Canary Islands). J Volcanol Geoth Res 178:543–552
Martin AJ, Umeda K, Connor CB, Weller JN, Zhao D, Takahashi M (2004) Modeling long-term volcanic hazards through Bayesian inference: an example from the Tohoku volcanic arc, Japan. J Geophys Res 109:B10208
Martin DP, Rose WI (1981) Behavioral patterns of Fuego volcano, Guatemala. J Volcanol Geotherm Res 10:67–81
Marzocchi W, Woo G (2007) Probabilistic eruption forecasting and the call for an evacuation. Geophys Res Lett 34: L22310
Marzocchi W, Woo G (2009) Principles of volcanic risk metrics: theory and the case study of Mt. Vesuvius and Campi Flegrei (Italy). J Geophys Res 114:B03213
Marzocchi W, Zaccarelli L (2006) A quantitative model for the time-size distribution of eruptions. J Geophys Res 111:B04204
Marzocchi W, Zechar JD (2011) Earthquake forecasting and earthquake prediction: different approaches for obtaining the best model. Seismol Res Lett 82:442–448
Marzocchi W, Mulargia F, Gonzato G (1997) Detecting low-dimensional chaos in geophysical time series. J Geophys Res 102:3195–3209
Marzocchi W, Sandri L, Gasparini P, Newhall C, Boschi E (2004) Quantifying probabilities of volcanic events: the example of volcanic hazard at Mt Vesuvius. J Geophys Res 109: B11201
Marzocchi W, Sandri L, Selva J (2008) BET_EF: a probabilistic tool for long- and short-term eruption forecasting. Bull Volcanol 70:623–632
Marzocchi W, Sandri L, Selva J (2010) BET_VH: a probabilistic tool for long-term volcanic hazard assessment. Bull Volcanol 72:705–716
Mastrolorenzo G, Petrone P, Pappalardo L, Sheridan MF (2006) The Avellino 3780 yr BP catastrophe as worst-case scenario for a future eruption at Vesuvius. Proc Natl Acad Sci USA 103:4366–4370
Medina Martinez F (1983) Analysis of the eruptive history of the Volcan de Colima, Mexico (1560–1980). Geofis Int 22:157–178
Mendoza-Rosas AT, De la Cruz-Reyna S (2008) A statistical method liking geological and historical eruption time series for volcanic hazard estimations: application to active polygenetic volcanoes. J Volcanol Geotherm Res 176:277–290
Mendoza-Rosas AT, De la Cruz-Reyna S (2009) A mixture of exponentials distribution for simple and precise assessment of the volcanic hazard. Nat Hazards Earth Syst Sci 9:425–431
Mendoza-Rosas AT, De la Cruz-Reyna S (2010) Hazard estimates for El Chicon volcano, Chiapas, Mexico: a statistical approach for complex eruptive histories. Nat Hazards Earth Syst Sci 10:1159–1170
Miller V, Savage M (2001) Changes in seismic anisotropy after volcanic eruptions: evidence from Mount Ruapehu. Science 293:2231–2233
Mogi K (1958) Relations between the eruptions of various volcanoes and the deformations of the ground surface around them. Bull Earthq Res Inst Univ Tokyo 36:99–134
Moran SC, Newhall CG, Roman DC (2011) Failed magmatic eruptions: late-stage cessation of magma ascent. Bull Volcanol 73(2):115–122 doi:10.1007/s00445-010-0444-x
Mulargia F, Tinti S, Boschi E (1985) A statistical analysis of flank eruptions on Etna volcano. J Volcanol Geotherm Res 23:263–272
Mulargia F, Gasperini P, Tinti S (1987) Identifying regimes in eruptive activity: an application to Etna volcano. J Volcanol Geotherm Res 34:89–106
Mulargia F, Gasperini P, Marzocchi W (1991) Pattern recognition applied to volcanic activity: identification of the precursory patterns to Etna recent flank eruptions and periods of rest. J Volcanol Geotherm Res 45:187–196
Mulargia F, Marzocchi W, Gasperini P (1992) Statistical identification of physical patterns which accompany eruptive activity on Mount Etna, Sicily. J Volcanol Geotherm Res 53:289–296
Munoz M (1983) Eruption patterns of the Chilean volcanoes Villarrica, Llaima, and Tupungatito. Pure Appl Geophys 121:835–852
Neri A, Aspinall W, Cioni R, Bertagnini A, Baxter P, Zuccaro G, Andronico D, Barsotti S, Cole P, Esposti Ongaro T (2008) Developing an event tree for probabilistic hazard and risk assessment at Vesuvius. J Volcanol Geotherm Res 178:397–415
Neuberg J (2000) Characteristics and causes of shallow seismicity in andesite volcanoes. Philos Trans Roy Soc A 358:1533–1546
Neuberg JW (2011) Earthquakes, volcanogenic. In: Gupta HK (ed) Encyclopedia of solid earth geophysics, vol 1. Springer, pp 261–269
Newhall C, Hoblitt R (2002) Constructing event trees for volcanic crises. Bull Volcanol 64:3–20
Newhall CG, Self S (1982) The volcanic explosivity index (VEI): an estimate of the explosive magnitude for historical eruptions. J Geophys Res 87:1231–1238
Novelo-Casanova DA, Valdes-Gonzales C (2008) Seismic pattern recognition techniques to predict large eruptions at the Popocatépetl, Mexico, volcano. J Volcanol Geotherm Res 176:583–590
Ogata Y (1988) Statistical models for earthquake occurrences and residual analysis for point processes. J Am Statist Assoc 83:9–27

Ogata Y (1998) Space-time point-process models for earthquake occurrences. Ann Inst Stat Math 50:379–402
Orsi G, Di Vito MA, Selva J, Marzocchi W (2009) Long-term forecast of eruption style and size at Campi Flegrei caldera (Italy). Earth Plan Sci Lett 287:265–276
Passarelli L, Brodsky EE (2012) The correlation between run-up and repose times of volcanic eruptions. Geophys J Int 188(3):1025–1045 doi:10.1111/j.1365-246X.2011.05298.x
Passarelli L, Sandri L, Bonazzi A, Marzocchi W (2010a) Bayesian hierarchical time predictable model for eruption occurrence: an application to Kilauea Volcano. Geophys J Int 181:1525–1538
Passarelli L, Sanso B, Sandri L, Marzocchi W (2010b) Testing forecasts of a new Bayesian time-predictable model of eruption occurrence. J Volcanol Geotherm Res 198:57–75
Pozgay SH, White RA, Wiens DA, Shore PJ, Sauter AW, Kaipat JL (2005) Seismicity and tilt associated with the 2003 Anatahan eruption sequence. J Volcanol Geotherm Res 146:60–76
Proschan F (1963) Theoretical explanation of decreasing failure rate. Technometrics 5:375–383
Pyle DM (1998) Forecasting sizes and repose times of future extreme volcanic events. Geology 26:367–370
Pyle DM (2000) Sizes of volcanic eruptions. In: Sigurdsson H et al (eds) Encyclopedia of volcanoes. Academic Press, San Diego, pp 263–269
Ramos EG, Laguerta EP, Hamburger MW (1996) Seismicity and magmatic resurgence at Mount Pinatubo in 1992. In: Newhall CG, Punongbayan RS (eds) Fire and mud, eruptions and lahars of Mount Pinatubo, Philippines, pp 387–406
Reyment RA (1969) Statistical analysis of some volcanologic data regarded as series of point events. Pure Appl Geophys 74:57–77
Rodado A, Bebbington M, Noble A, Cronin S, Jolly G (2011) On selection of analogue volcanoes. Math Geosci 43:505–519
Roman D, Neuberg J, Luckett R (2006) Assessing the likelihood of volcanic eruption through analysis of volcanotectonic earthquake faultplane solutions. Earth Plan Sci Lett 248:229–237
Rosi M (1996) Quantitative reconstruction of recent volcanic activity: a contribution to fore- casting of future eruptions. In: Scarpa R, Tilling RI (eds) Monitoring and mitigation of volcano hazards. Springer, Berlin, pp 631–674
Salvi F, Scandone R, Palma C (2006) Statistical analysis of the historical activity of Mount Etna, aimed at the evaluation of volcanic hazard. J Volcanol Geotherm Res 154:159–168
Sandri L, Marzocchi W, Gasperini P (2005) Some insights on the occurrence of recent volcanic eruptions of Mount Etna volcano (Sicily, Italy). Geophys J Int 163:1203–1218
Sandri L, Guidoboni E, Marzocchi W, Selva J (2009) Bayesian Event Tree (BET) for eruption forecasting at Vesuvius, Italy: a retrospective forward application to 1631 eruption. Bull Volcanol 71:729–745
Sandri L, Jolly G, Lindsay J, Howe T, Marzocchi W (2012) Combining long- and short-term probabilistic volcanic hazard assessment with cost-benefit analysis to support decision making in a volcanic crisis from the Auckland Volcanic Field, New Zealand. Bull Volcanol 74(3):705–723 doi:10.1007/s00445-011-0556-y
Scandone R, Arganese G, Galdi F (1993) The evaluation of volcanic risk in the Vesuvian area. J Volcanol Geotherm Res 58:263–271
Selva J, Marzocchi W, Civetta L, Del Pezzo E, Papale P (2010a) Emergency preparedness: community-based short-term eruption forecasting at Campi Flegrei. EGU General Assembly 2010, held 2–7 May 2010 in Vienna, Austria, p 10318
Selva J, Costa A, Marzocchi W, Sandri L (2010b) BET_VH: exploring the influence of natural uncertainties on long-term hazard from tephra fallout at Campi Flegrei (Italy). Bull Volcanol 72:717–733
Selva J, Orsi G, Di Vito M, Marzocchi W, Sandri L (2012) Probability hazard map for future vent opening at the Campi Flegrei caldera, Italy. Bull Volcanol 74(2):497–510 doi:10.1007/s00445-011-0528-2
Settle M, McGetchin TR (1980) Statistical analysis of persistent explosive activity at Stromboli, 1971: Implications for eruption prediction. J Volcanol Geotherm Res 8:45–58
Shelly D, Hill DP (2009) Migrating swarms of brittle-failure earthquakes in the lower crust beneath Mammoth Mountain, California. Geophys Res Lett 38:L20307
Sigurdsson H (editor in chief) et al (2000) Encyclopedia of volcanoes. Academic Press, San Diego
Simkin T, Siebert L (1994) Volcanoes of the world. Geosciences, Tucson, Ariz
Smethurst L, James MR, Pinkerton H, Tawn JA (2009) A statistical analysis of eruptive activity on Mount Etna, Sicily. Geophys J Int 179:655–666
Smith R, Kilburn C (2010) Forecasting eruptions after long repose intervals from accelerating rates of rock fracture: the June 1991 eruption of Mount Pinatubo, Philippines. J Volcanol Geotherm Res 191:129–136
Sobradelo R, Marti J (2010) Bayesian event tree for long-term volcanic hazard assessment: application to Teide-Pico Viejo stratovolcanoes, Tenerife, Canary Islands. J Geophys Res 115:B05206
Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, Miller HL, eds. (2007) Climate change 2007: the physical science basis: contribution of Working Group I to the fourth assessment report of the intergovernmental panel on climate change. Cambridge and New York, Cambridge University Press
Solow AR (2001) An empirical Bayes analysis of volcanic eruptions. Math Geol 33:95–102
Sparks RSJ (2003) Forecasting volcanic eruptions. Earth Plan Sci Lett 210:1–15
Swanson DA, Casadevall TJ, Dzurisin D, Malone SD, Newhall CG, Weaver CS (1983) Predicting eruptions at Mount St Helens, June 1980 through December 1982. Science 221:1369–1376
Takada A (1997) Cyclic flank-vent and central-vent eruption patterns. Bull Volcanol 58:539–556
Thelen WA, Malone SD, West ME (2010) Repose time and cumulative moment magnitude: a new tool for forecasting eruptions? Geophys Res Lett 37:L18301
Thomas ME, Neuberg JW (2012) What makes a volcano tick-A first explanation of deep multiple seismic sources in ascending magma. Geology 40(4):351–354. doi:10.113/G32868.1
Thorlaksson JE (1967) A probability model of volcanoes and the probability of eruptions of Hekla and Katla. Bull Volcanol 31:97–106
Todesco M, Rinaldi AP, Bonafede M (2010) Modeling of unrest signals in heterogeneous hydrothermal systems. J Geophys Res 115:B09213
Tokarev PI (1978) Prediction and characteristics of the 1975 eruption of Tolbachik volcano, Kamchatka. Bull Volcanol 41:251–258
Trasatti E, Bonafede M, Ferrari C, Giunchi C, Berrino G (2011) On deformation sources in volcanic areas: modeling the Campi Flegrei (Italy) 1982–84 unrest. Earth Plan Sci Lett 306:175–185

Traversa P, Lengliné O, Macedo O, Metaxian J, Grasso J, Inza A, Taipe E (2011) Short term forecasting of explosions at Ubinas volcano, Perù. J Geophys Res 116:B11301

Turner M, Cronin S, Bebbington M, Platz T (2008a) Developing a probabilistic eruption forecast for dormant volcanos: a case study from Mt Taranaki, New Zealand. Bull Volcanol 70:507–515

Turner M, Cronin S, Smith I, Bebbington M, Stewart RB (2008b) Using titanomagnetite textures to elucidate volcanic eruption histories. Geology 36:31–34

Turner M, Bebbington M, Cronin S, Stewart RB (2009) Merging eruption datasets: building an integrated Holocene eruptive record of Mt Taranaki. Bull Volcanol 71:903–918

UNDRO (1985) Volcanic emergency management. Office of the United Nations Disaster Relief Co-Ordination, Geneva

Voight B (1988) A method for prediction of volcanic eruptions. Nature 332:125–130

Voight B (1989) A relation to describe rate-dependent material failure. Science 243:200–203

Voight B, Glicken H, Janda RJ, Douglass PM (1981) Catastrophic rockslide avalanche of May 18. In: Lippman PW, Mullineaux DR (eds) The 1980 eruptions of Mount St. Helens, Washington. US Geological Survey Professional Paper, 1250, pp 347–377

Voight B et al (1999) Magma flow instability and cyclic activity at Soufriére Hills volcano, Montserrat British West Indies. Nature 283:1138–1142

Wadge G (1982) Steady state volcanism: evidence from eruption histories of polygenetic volcanoes. J Geophys Res 87:4035–4049

Wadge G, Burt L (2011) Stress field control of eruption dynamics at a rift volcano: Nyamuragira, D.R.Congo. J Volcanol Geotherm Res 207:1–15

Wadge G, Cross A (1988) Quantitative methods for detecting aligned points: an application to the volcanic vents of the Michoacan-Guanajuato volcanic field, Mexico. Geology 16:815–818

Wadge G, Young PAV, McKendrick IJ (1994) Mapping lava flow hazards using computer simulation. J Geophys Res 99:489–504

Wang T, Bebbington M (2011) Robust estimation for the Weibull process applied to eruption records. In: Proceedings of the international union of geology and geodesy XXV general assembly, Melbourne, Australia, 28 June–7 July, IAVCEI JV02:920

Watt SFL, Mather TA, Pyle DM (2007) Vulcanian explosion cycles: patterns and predictability. Geology 35:839–842

Wehrmann H, Dzierma Y (2011) Applicabilty of statistical eruption analysis to the geological record of Villarrica and Llaima Volcanoes, Southern Volcanic Zone, Chile. J Volcanol Geotherm Res 200:99–115

Weller JN (2004) Bayesian inference in forecasting volcanic hazards: an example from Armenia. MS Thesis Paper 1298, University of South Florida

White RA, Power JA (2001) Distal volcano-tectonic earthquakes: diagnosis and use in eruption forecasting. Trans AGU (Eos) 82:47

Wickman FE (1966a) Repose-period patterns of volcanoes. I. Volcanic eruptions regarded as random phenomena. Arch Miner Geol 4:291–367

Wickman FE (1966b) Repose-period patterns of volcanoes. V. General discussion and a tentative stochastic model. Arch Mineral Geol 4:351–367

Woo G (2008) Probabilistic criteria for volcano evacuation decisions. Nat Hazards 45:87–97. doi:10.1007/s11069-007-9171-9