Journal of Volcanology and Geothermal Research 286 (2014) 15–23

Contents lists available at ScienceDirect

Journal of Volcanology and Geothermal Research

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

Long-term forecasting of eruption hazards: A hierarchical approach to merge analogous eruptive histories

Tom Sheldrake

School of Earth Sciences, University of Bristol, Wills Memorial Building, Queen's Road, Bristol BS8 1RJ, United Kingdom

ARTICLE INFO

Article history:
Received 10 January 2014
Accepted 21 August 2014
Available online 28 August 2014

Keywords:
Hierarchical
Bayesian
Volcanic hazard
Exchangeability
Forecasting
Power-law
Frequency–magnitude

ABSTRACT

Estimating the hazard associated with a volcanic eruption requires an understanding of previous eruptive episodes to forecast future events. This involves calculating how destructive a future eruption is likely to be by estimating the magnitude of eruptive activity and likelihood of various hazardous phenomena. Importantly though, eruptive histories for individual volcanoes can suffer from a lack of observations and thus might not be representative for all future eruption scenarios. Consequently, a methodology is developed to combine eruptive histories from multiple volcanoes into an event tree framework to inform forecasts at individual volcanoes. It is based on a hierarchical Bayesian approach in which model parameters are derived for a group of volcanoes and then updated on an individual basis. However, eruptive histories are not simply aggregated and the model allows for possible heterogeneities in eruptive regimes. Continuous probability distributions are employed to capture the relative uncertainties of both global and individual records and posterior distributions for eruption magnitudes and hazardous phenomena are computed using Markov chain Monte Carlo techniques. The model is designed to initially include no subjective judgement of probabilities but is developed so that information from other analyses can be incorporated. While this article uses the hierarchical Bayesian approach specifically for event forecasting, the methodology has the potential to be used in a wide range of problems regarding hazard assessment and for the purposes of causal inference and data reduction.

1. Introduction

Volcanoes pose significant risks for human populations living in proximal areas. Attempts to forecast volcanic eruptions occur at a variety of timescales from hours to decades and can involve a multitude of different datasets, such as eruptive histories or time series of monitoring observables (Marzocchi and Bebbington, 2012). However, within the field of volcanic hazard assessment a common issue with interpreting and analysing datasets is a lack of observations at individual volcanoes. This could be due to resource constraints restricting the ability to record eruptive activity or perhaps simply that the eruptive history of a volcano is limited. The result is that hazard assessments performed for volcanoes where data are scarce are likely to be extremely uninformative and potentially inaccurate. Consequently, this has resulted in the employment of subjective judgement using techniques such as expert elicitation (Aspinall and Cooke, 2013) rather than the results from empirical analyses when little data exists.

Where data are scarce an alternative empirical approach is to aggregate datasets from groups of analogous volcanoes (Rodado et al., 2011). However, this involves some subjectivity as to how to combine these data, and assumptions that observations are independent and identically distributed. However, it is tricky to make this assumption as no two volcanoes are precise analogues of each other and thus assessments will be susceptible to heterogeneities. Nevertheless, artificially increasing the size of datasets by incorporating observations from multiple volcanoes has promise. It has the potential to improve the certainty of assessments while also ensuring that so-called "black swan" events are accounted for (Woo, 2011). For example, if a volcanic record contains only low magnitude eruptions, it might not be correct to assume that all future eruption magnitudes will be of similar nature and thus a full range of eruption magnitudes must be considered. Consequently, the occurrence rate for unobserved magnitudes must be made based upon observations from other records. Further reasoning for the concept of combining observations from multiple volcanoes is supported by the continuing development of global databases of monitoring data and eruption records (Siebert and Simkin, 2002; Crosweller et al., 2012), which are improving the ease with which such studies could be performed.

A common approach for eruption forecasting is the use of event tree structures that contain a series of interconnected nodes (Newhall and Hoblitt, 2002; Marzocchi et al., 2004; Sparks and Aspinall, 2004; Neri et al., 2008; Sobradelo and Martí, 2010). As one progresses through the event tree, along a series of unique branches, the nodes relate to more specific events until a final outcome is reached. To calculate the probabilities at each node Bayesian methods have been developed

E-mail address: tom.sheldrake@bristol.ac.uk.

http://dx.doi.org/10.1016/j.jvolgeores.2014.08.021
0377-0273/

that allow various data sources, including both empirical observations and subjective judgement, to be integrated together (Marzocchi et al., 2008; Marzocchi et al., 2010; Sobradelo et al., 2013). The purpose of this article is to couple a robust statistical approach based on a Bayesian method to share data between volcanoes and to account for both group-level similarities and individual-volcano heterogeneities with an event tree approach to assess long-term volcanic hazard.

2. Hierarchical Bayesian model

Hierarchical modelling is a method used to break down a complex problem into layers, where each layer can be modelled using a probability distribution. The Bayesian approach developed by Marzocchi et al. (2008, 2010) specifies a probability model for the data given specific parameters. A hierarchical approach will then specify a model for these parameters given hyper-parameters; and perhaps a model for the hyper-parameters given other, higher-level hyper-parameters. The number of layers depends upon how the problem is modelled and how individuals (i.e. volcanoes) in the model are grouped together. Ultimately there is no restriction on the number of layers, although as the number increases the model often becomes more abstract and potentially difficult to interpret. Here, the number of levels is restricted to two; firstly a probability distribution for each event (i.e. model parameters); and secondly parameters on these model parameters (i.e. model hyper-parameters). By introducing these hierarchies it allows multiple volcanoes to be modelled together by assuming that the parameters of the prior distribution for an event are constant across the group. This means that without observing the data (i.e. no of eruptions for each respective VEI) the probability of each event (i.e. VEI eruption) is constant across the group. The posterior distributions for these model hyper-parameters are thus dependent on data from all the volcanoes in the model. The group-level prior is then updated individually for each volcano based upon its respective eruptive record. For this structure to exist, an important starting assumption of exchangeability must be made. A set of random variables are exchangeable (and thus exhibit symmetry) if their joint distribution is invariant to permutations of their labels. Therefore, the parameters of the model (e.g. probability of a VEI magnitude at a particular volcano) are identically distributed if nothing but the data (e.g. number of eruptions) can be used to distinguish between the volcanoes. Consequently, the basis upon which analogue volcanic records are chosen to be exchangeable is important for the validity of the model.

The model that is developed here expands the approach of Marzocchi et al. (2008, 2010), which aims to calculate the probabilities for N mutually exclusive events and employs Bayesian inference to update prior beliefs using appropriate observations. In this previous approach each mutually exclusive event (j) relates to one of a number of states (J) that an individual volcano can take. For example, j may be a specific VEI magnitude from a range of VEI eruption magnitudes. The prior beliefs are parameterised using a Dirichlet distribution and updated using a Multinomial likelihood, which is a conjugate distribution to the Dirichlet. The Dirichlet prior is defined by J alpha parameters ($_j$) and the likelihood the number observations for each state (x_j) where n is the total number of recorded events for a volcano (e.g. total number of eruptions recorded). These same priors and likelihood functions are used here with the addition of an index parameter i representing each individual volcano (Eqs. (1) and (2)).

$$\theta_{i,j} \sim Dir\left(\alpha_j\right) \tag{1}$$

$$x_{i,j} \sim Multi\left(\theta_{i,j}, n_i\right) \tag{2}$$

Choosing the parameters of the model (i.e. the α parameters) will determine the informativeness of the likelihood function and the overall performance of the model. Where no relevant observations or understanding exists the prior can be chosen to be noninformative, for example a uniform distribution (e.g. Sandri et al., 2012), which reflects no prior knowledge. Alternatively the prior can be chosen to be informative and based on previous observations and the understanding of the volcanic processes. The benefit of a noninformative prior is that it allows the data to fully control the posterior distribution without the introduction of biases associated with subjective judgement of model parameters. However, due to the sparse nature of eruption records and the various recording biases associated with historical and geological records (Coles and Sparks, 2006), a noninformative approach will likely lead to inaccurate results that are not representative of the true understanding of the volcanic processes. Consequently, model parameters are commonly defined using results from other analytical models or using methods such as expert elicitation (e.g. Sobradelo and Martí, 2010).

In the model developed by Marzocchi et al. (2008, 2010) the variance in the model, or confidence in the prior distribution is characterised by an “equivalent number of data” parameter. The value of this parameter requires a subjective judgement of the informativeness of the prior distribution. Alternatively, using the HASSET tool this is done by weighting the informativeness of the prior distribution in calculating the posterior distribution relative to the data or likelihood function (Sobradelo et al., 2013).

One approach to estimate values for the model parameters is to use observations from a group of volcanoes. However, there are many issues when using analogue datasets to calculate point values for the model parameters based on the average rate of an event and measures of the data informativeness or variance. For example, using a group of analogous volcanoes to define point estimates for the parameters of the prior distribution will artificially increase the precision of the posterior distributions for each volcano as the eruptive record would be used twice, once in the calculation of the prior distribution and secondly in updating the prior distribution (Gelman et al., 2004).

The following hierarchical approach is designed with the introduction of an additional level to the model by defining hyperpriors that are rearranged in terms of moments (mean average and equivalent number of data) of the prior distribution. The introduction of hyperpriors allows a noninformative prior distribution, where the parameters can be updated using the relevant data from analogue volcanoes, without the need for subjective judgement of the informativeness of the prior distribution. Consequently, this allows the prior distribution to be extremely flexible and range from being noninformative to extremely informative depending upon the size and similarity of eruptive records in the group dataset. Thus, hyperpriors ensure that the α parameters are not defined by point estimates and allow uncertainty in the prior distribution to be quantified.

2.1. Hyperpriors

The first hyperprior (Eq.(3)) is defined for each alpha parameter or state (j) of the Dirichlet distribution (e.g. a specific VEI magnitude). They are chosen to be diffuse with low precision so that the data can speak for itself. To do this a logit transformation is used to convert the parameter space of the mean average to an unbounded continuous scale ($-\infty,\ \infty$). This allows the hyperpriors to be parameterised using a diffuse normal distribution with low precision. The final transformation of the hyperprior is the multiplication of the logit of the mean by ($J-1$). This ensures that the hyperprior can be defined by a normal distribution centred on zero ($\mu_j = 0$). This means that with no data in the model the alpha parameters each equal one, reflecting a noninformative uniform distribution across all the possible outcomes (i.e. all the VEI magnitudes are equally likely).

The second hyperprior is a uniform distribution the strength of the prior and degree of applicability of the analogue set (Eq. (4)). The hyperprior is parameterised using α_0, which is the sum of all the alpha parameters. This reflects the equivalent number of data