Edinburgh Research Explorer

Statistical analysis of daily seismic event rate as a precursor to volcanic eruptions

OPEN ACCESS

Statistical analysis of daily seismic event rate as a precursor to volcanic eruptions

Sebastien F. M. Chastin and Ian G. Main

Edinburgh University, School of GeoSciences, Scotland, United Kingdom

[1]   We analyse time series of daily seismic event rate for the Kilauea, Hawaii, volcano between 1959 and 2000. Individual eruptions are not always preceded by an increase in event rate, and many increases in event rate do not lead to eruption. However, a mean field accelerating behaviour does emerge 10–15 days before eruption in data stacked in phase with the eruption time. In phase space the pre-eruptive dynamics is well defined by *Voight*'s [1988] equation, but so is that of the seismicity in the period between eruptions. We conclude that the underlying dynamics of the 'background' seismicity is similar to that of magma eruption. We use Bayesian methods to compare different time-to-failure models that have been suggested for precursors. Only a short-term forecast can be achieved, using a linear fit to inverse rate.   INDEX TERMS: 7223 Seismology: Seismic hazard assessment and prediction; 7280 Seismology: Volcano seismology (8419); 7299 Seismology: General or miscellaneous; 8419 Volcanology: Eruption monitoring (7280). **Citation:** Chastin, S. F. M., and I. G. Main, Statistical analysis of daily seismic event rate as a precursor to volcanic eruptions, *Geophys. Res. Lett.*, *30*(13), 1671, doi:10.1029/2003GL016900, 2003.

1.   Introduction

[2]   Precursory signals are often observed - in real time or in retrospect - to occur before individual volcanic eruptions. Earthquake swarms and ground deformation are amongst the most frequently cited precursors [*Scarpa and Tilling*, 1996], but no formal universal quantitative rules have so far been developed to forecast eruptions. In practice predictions are commonly based on a empirical assessments of patterns in a variety of precursory signals, and the expertise of observatory staff.

[3]   *Klein* [1984] first proposed a quantitative heuristic method for estimating eruption probability. He introduced a statistical test based on the comparison of the distribution of the precursory signal amplitude before an eruption with those during the repose time. More recently *Voight* [1988] suggested a forecasting model based on dynamics of material failure to describe accelerating earthquake event rate and ground deformation. The rationale is that the growth of magmatic pathways is driven by rock failure in the volcanic edifice under sustained, near-constant fluid pressure in the magma chamber. Therefore the rate of magma ascent is limited by the rate of fracture growth, recorded as seismicity.

[4]   Here we test Voight's hypothesis using daily earthquake event rates and eruption times from Kilauea, Hawaii.

We first seek the pre-eruptive dynamics in earthquake frequency time series and then test for detectability of any precursory signal against a null hypothesis of a random process. We use Bayesian techniques to compare different time to failure models and formally assess their forecasting power and confidence levels.

2.   Data

[5]   Kilauea is an active hotspot-type shield volcano, with generally effusive. Its activity is very well documented and there have been a comparatively large number of eruptions in recent times. This makes it a good candidate for statistical analysis. Earthquakes (know as Volcano Tectonic events) occurring within an 8.7×14.6 km area around the Kilauea caldera, as described by *Klein* [1984], with no restriction on depth, are counted daily. We do not introduce further apriori classification of events and we view the seismicity as a time series of acoustic emissions resulting from the general behaviour of the volcanic edifice. The daily seismic event rates and eruption times are plotted in Figure 1, covering the period from 10th January 1959 to 27th August 2000.

[6]   There were 35 eruptions in this area between 1959 and 2000, defined as the first day on which magma reached the surface of the volcano *Klein* [1984]. The HVO defines 55 different episodes during the long 1983 eruption [*Wolfe*, 1988]. These features are shown in Figure 1.

3.   Pre-Eruptive Seismicity

[7]   It is common knowledge among volcano seismologist that volcanic earthquake occur in swarms not necessarily correlated with eruptions. *Benoit and McNutt* [1996] described the log-normal distribution of swarm duration and showed an increase in the mean duration for swarms that lead to eruption. McNutt [in *Scarpa and Tilling*, 1996] gave a cinematic description of the evolution of the seismicity before eruptions. There are numerous detailed case studies of individual swarms, but outside of the work by *Voight* [1988], *Cornelius and Voight* [1994] and *Kilburn and Voight* [1998], little attention has been directed toward the generic temporal dynamics of these swarms.

3.1.   Individual Sequences and Correlations

[8]   Examining Figure 1 we see that some individual eruptions are preceded by accelerating seismicity, but many are not. Many clear phases of accelerating seismicity do not terminate in eruptions during the repose time. The background seismicity in non-eruptive periods has a mean level of 5 events per day. Only 19 out of the 35 eruptions have an increased seismicity, on the day of eruption, above this background level. Event rates sometimes far exceed the
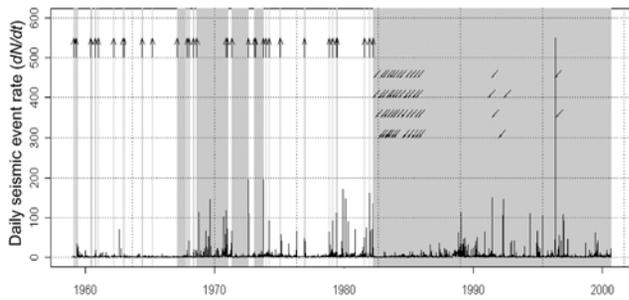
Figure 1. Time series of daily seismic event rate at Kilauea 10th January 1959 untill 27th August 2000. Eruptions are marked by upward arrows the eruption period as grey area. Eruptive episodes within the 1983 eruption are marked as slanted downward arrows.

eruption day rate within a 100-day window prior to eruption. This is true in all cases where the eruption rate is below 50 events per days. There is no systematic critical threshold value of event rate, before eruption. Moreover we have found no correlation between the terminal event rate before eruption and the size of eruption, measured as the volume of lava erupted, or inter-eruption time or eruption duration. Thus a simple threshold value cannot be used reliably as a precursor for this data set.

3.2. Mean Field Behaviour

[9] Due to the presence of significant "background noise", individual sequences are not suitable for deriving a generic dynamic behaviour of pre-eruptive seismicity. To increase the signal to noise ratio, we extract a mean field behaviour by superposing the sequences in a stack with the eruption date as common final point. In other words we add together many independent realisation of the same phenomena. This is a technique used routinely in statistical physics and introduced as the superimposed epoch analysis by *Grasso* in *Collombet et al.* [2002].

[10] The stacked sequence, displayed in Figure 2a, shows a clear increase in seismic rate dN/dt $10-15$ days prior to the eruption date, apparently following a power law acceleration according to Voight's prediction for quasi-static loading

$$\frac{dN}{dt} \propto (t_e - t)^{-\nu} \qquad (1)$$

where t_e is the eruption date and ν is a positive scaling exponent. The derivation of (1) assumes that, in the last few days before eruption, the relationship between eruptive mechanism and seismicity is identical for each individual case, and that the load is approximately constant. The filling rate of the magma chamber at Kilauea is slow [*Swanson*, 1972], and the mean inter-eruptive time ($\sim$241 days) is large compared to the $10-15$ days period we consider. Therefore, it is a reasonable assumption.

[11] The stacked sequence is not dominated by a single large pre-eruptive series, but there is a possibility that the power law increase occurred by chance. To test this formally we generated sets of surrogate data [*Theiler et al.*, 1992], which are independent random events with the same statistical distribution and Fourier amplitude spectra as the original pre-eruptive sequences, but with random phases.

The time correlation is therefore deliberately destroyed and surrogate data are series of independently random events. The surrogate stack, in Figure 2c does not show the precursory power law increase of the seismic rate, proving that the pre-eruptive dynamics is a robust generic process.

4. Pre-Eruptive Dynamics and Detectability

[12] Power law acceleration of the seismic event rate has been observed before eruption in individual cases by *Voight* [1988], *Cornelius and Voight* [1994], and *Kilburn and Voight* [1998]. It can be described by the general empirical equation

$$\frac{d^2N}{dt^2} - A\left(\frac{dN}{dt}\right)^{-\alpha} = 0 \qquad (2)$$

relating the event rate dN/dt to the acceleration d^2N/dt^2. The constant α is a positive power law exponent, and A is a constant, which depends on initial conditions including the background level.

[13] We obtained estimates of α and A from simple linear regression, by taking the log of this expression, for the stacked pre-eruptive sequence. The estimation was performed for pre-eruptive window ranging from 100 days to 4 days. We compared these estimates to those obtained for: the surrogate data stack (Figure 2c), a stack of sequences during non-eruptive repose (Figure 2d), a stack of 35 sequences preceding episodes of the 1983 eruption (Figure 2b) to ascertain whether these parameters are significantly different in pre-eruptive period, and hence if the pre-eruptive dynamics is detectable in this data set.

[14] For 100-day pre-eruptive windows, estimates in all four cases belong to the same generic set defined by the likelihood surface of estimate in parameter spaces (Figure 3). When we consider windows $\leq$12 days prior to eruption these results change (Figure 3). The pre-eruptive sequence is a subset of the inter-eruptive dynamics, but is clearly distinguishable from the surrogate stack in terms of its variability within the parameter space. This indicates that the pre-eruptive dynamics in not solely a random process and that the stacking is robust. However the power law exponent characterising the event rate acceleration does not change with the onset of volcanic eruption. We note that the exponent for the 1983 episodes stack is lower and parame-
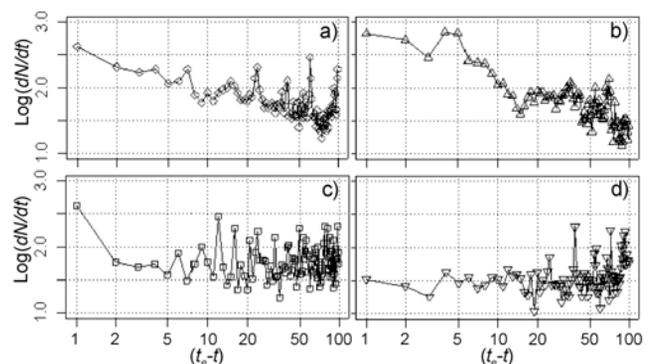
Figure 2. Log-log plot of seismic rate stacked sequences versus $(t_e - t)$. (a) Pre-eruptive stack, (b) Pre-episode stack, (c) Surrogate pre-eruptive stack, (d) Repose non-eruptive stack.
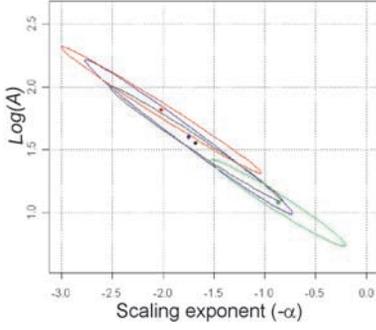
Figure 3. Likelihood contour and best estimates (as points) of the dynamics parameters: α the scaling exponent and log(A) background rate, for pre eruptive stack (black), surrogate data (red), repose stack (blue), pre-episode (green).

ters distinguishable from the pre-eruptive set. This implies that lava breaks after the first eruption tend to have less marked non-linearity in their dynamics.

5. Time to Failure Models

[15] Time to failure analysis routinely used with acoustic emission in material science has been applied in different forms to earthquake [*Bufe and Varnes*, 1993; *Sornette and Sammis*, 1995] *Voight* [1988] and eruption forecasting.

[16] We use the standard Newton-Raphson method to fit the following time to failure models to the pre-eruptive mean sequence:

[17] - Exponential rate growth [*Lockner*, 1991]

$$\frac{dN}{dt} = B + AExp[(t_e - t)] \tag{3}$$

[18] - Power law model equation (1) [*Voight*, 1988].

[19] - Inverse rate model [*Voight*, 1988].

$$\left(\frac{dN}{dt}\right)^{-1} = at + b \tag{4}$$

[20] This is a linear version of (1), near $t_e = -b/a$

[21] - Log-periodic model [*Sornette and Sammis*, 1995]

$$\frac{dN}{dt} = A(t_e - t)^{-\nu}\left[1 + C.\cos\left(2\pi\left(\frac{\log(t_e - t)}{\log(\lambda)}\right) + \psi\right)\right] \tag{5}$$

[22] - A random distribution with constant mean μ and standard deviation σ as null hypothesis

$$\frac{dN}{dt} = \aleph(\mu, \sigma) \tag{6}$$

[23] We repeated the procedure for pre-eruptive windows ranging from 100 days to the minimum number of data needed to retain 1 degree of freedom in the regression. Figure 4 shows the best fit to the various models in a 100 days window prior to the eruption time.

5.1. Model Ranking

[24] To get a reliable estimate of the eruption date we must choose a forecasting model with the smallest number of free parameters but which retains an accurate description of the dynamics. Consequently we rank the competing

models using a Bayesian Information Criterion (BIC) [*Leonard and Hsu*, 1999]. BIC is a standard goodness-of-fit measure that introduces a penalty for the number of free parameters involved. Hence, it gives a measure of the level of useful information obtained from a model.

[25] Table 1 summarises the results. For a 100-days window, BIC is highest overall for the random distribution (6), so the null hypothesis cannot be rejected. The random sequence is objectively the most appropriate model for these data, but provides no forecast for t_e. As the size of the window is reduced to $10-15$ days the linear model (4) begins to outperform the null hypothesis (6), and all the other models. This implies there is real predictive power in the linear model for the stacked data sequence a week or two prior to the eruption.

5.2. Forecast Confidence

[26] Decision-makers need to know the confidence level that can be attributed to forecasts. To complete our comparison we calculate Bayesian Inference level for the time to failure models. We computed the posteriori density for each model using Bayes theorem [*Leonard and Hsu*, 1999]. The only prior information we have is that t_e is larger or equal to the last day in the time series (day 100). We keep the prior assessment as uninformative as possible, by assigning a uniform distribution to t_e between bounds $t_e \in [100, t_{max}]$. To avoid problems with the tail we have chosen to restrict t_{max} to 300, that is $t_e + 200$ days (close to the mean inter eruption period). From the posteriory probability density function, the most likely time of eruption can be confined to a window set by the 95% Bayesian interval in t_e.

[27] The results are also summarised in Table 1. The linear model, equation (4), gives the narrowest uncertainty on t_e for the 15-day window [between $100-105$ days at the 95% confidence level]. All the other models return Bayesian intervals at this confidence level that are much larger than the 15-day timescale of the pre-eruptive sequence.

6. Discussion

[28] Kilauea and the Hawaii volcanic system are remote from any plate tectonic boundary. The seismicity must be driven by other phenomena: rather than plate boundary stresses.
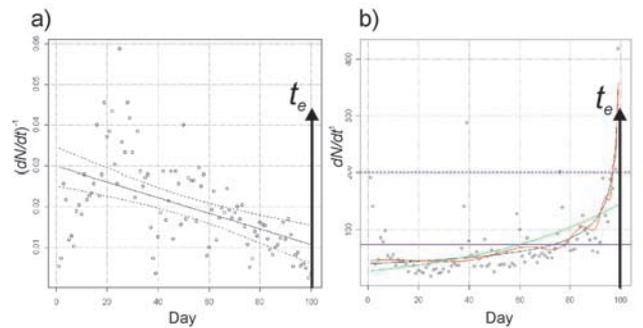
Figure 4. Fit of Time to Failure Model a) inverse rate linear model dashed line represent the 95% confidence interval b) power law in black, exponential in green, log periodic in red, random distribution with constant mean in blue (with standard deviation in blue dashed line).

Table 1. BIC Scores for a 100 and 15 Days Pre-Eruptive Window, 95% Confidence Interval on t_e, Bayesian Confidence Level in a 15 Days Window, and Number of Free Parameters for Time to Failure Models

Statistics	BIC 100 days window	BIC 15 days window	95 % Interval of confidence	Bayesian Inference Level (%)	Number of free parameters
Models					
Linear (4)	653	178	$100-105$	~95	3
Exponential (3)	1065	169	$100-184$	44	4
Power Law (1)	1029	151	$100-147$	78	4
Log Periodic (5)	1020	146	$100-185$	54	8
Random (6)	1114	160	NA	NA	2

[29] In the seismicity time series analysed here only ~15% of observed individual sequences of accelerating earthquake counts correspond to reported lava reaching the surface. Eruptions are direct and clear indication of magmatic dike intrusion. It is tempting to associate the remaining 85% with intrusions not reaching the surface. This would be consistent with geological observations. According to *Gudmundsson* [2002], over 90% of basaltic intrusions in the form of dikes are arrested before reaching the surface. If indeed these accelerating swarms are the seismic response to injection of magma in the volcanic rock edifice, small and large intrusions will result in power law (scale invariant) accelerations of the seismic event rate, with an identical exponent. Consequently we cannot discriminate between the seismic signature of intrusions and eruptions.

[30] On the other hand it is difficult to find a generic and consistent behaviour for seismicity associated with dike propagation [*Rubin et al.*, 1998] even when the presence of an intrusion is confirmed by lava breaking the surface. It is not proven that intrusion plays an active part in generating seismic events and fractures. On the contrary, there are indications that intrusions are passive [*Cervelli et al.*, 2002], and propagate following fracture networks opening in the rock mass under other driving forces and the interaction between crack populations. In this respect it is not surprising that the pre-eruptive seismicity dynamics cannot be distinguished from the background seismicity.

[31] The lack of firm evidences that dike intrusions actively alter the seismicity and propagate generating VT events, should change the way eruption mechanism is perceived. The idea that swarms are due to the impeding eruption forcing its way through the volcanic edifice seems unlikely. The reversed causality relation is more probable. Effusive eruptions should be viewed as passive events resulting from the opening of a percolation path to the surface under more or less static load, as hypothesised by *Voight* [1988].

[32] The linear method applied to inverse rate proposed by *Voight* [1988] and further developed by *Kilburn and Voight* [1998] succesfully 'hindcasts' the eruption date with a significant level of confidence when applied to the stacked pre-eruptive sequence. However in prospective mode we expect this confidence level to drop dramatically. Forecasts of eruption dates strongly depend on the choice of initial conditions and the width of temporal window. Moreover the model seems more generic of the seismicity on volcano hence it is likely to forecast eruptions for all accelerating avalanches. We therefore can expect a false alarm ratio of the order of 90%, if seismic event rate data is used on its own as a forecasting tool.

7. Conclusion

[33] Acceleration of earthquake rates is not observed systematically before all individual eruptions in the Kilauea catalogue. However a robust power law increase emerges 10 to 15 days prior to the eruption date. However its characteristic parameters are statistically indistinguishable from those of the background generic physical process. Neither a critical threshold level for the event rate, nor any change in the scaling exponent can be used as a reliable indicator of pre-eruptive periods. Seismic event rate data only provide a short-term forecasting capability using a linear model of the inverse rate. As a final caveat, these results are obtained for an effusive volcano and might not apply to explosive volcanoes. This will be the subject of further work.

References

Bufe, C. G., and D. J. Varnes, Predictive modelling of the seismic cycle of the greater San Francisco Bay region, *Bull. Seism. Soc. Am.*, *86*, 868–889, 1993.

Cervelli, P., P. Segall, F. Amelung, H. Garbeil, C. Meertens, S. Owen, A. Miklius, and M. Lisowski, The 12 September 1999 Upper East Rift Zone dike intrusion at Kilauea Volcano, Hawaii, *J. Geophys. Res.*, *107*, 2150, 2002.

Cornelius, R. R., and B. Voight, Seismological aspect of the 1989–1990 eruption at Redoubt volcano, Alaska: the material failure forecast method with RSAM and SSAM seismic data, *J. Volcanol. Geotherm. Res.*, *62*, 469–498, 1994.

Collombet, M., J. R. Grasso, and V. Ferrazzini, Seismicity rate and volcano dynamics at Piton de la Fournaise Volcano, (Indian Ocean), *Eos, Trans. AGU*, *83*(47), Fall Meet. Suppl., abstract V21A-1171, 2002.

Gudmundsson, A., Emplacement and arrest of sheets and dikes in central volcanoes, *J. Volcanol. Geotherm. Res.*, *116*(3-4), 279–298, 2002.

Kilburn, R. J., and B. Voight, Slow rock fracture as eruption precursor at Soufriere Hills volcano, Montserrat, *Geophys. Res. Lett.*, *25*(19), 3665–3668, 1998.

Klein, J., Eruptive Forecasting at Kilauea Volcano Hawaii, *J. Geophys. Res.*, *89*, 3059–3073, 1984.

Leonard, T., and J. S. J. Hsu, Bayesian methods, *Cambridge Univ. Press*, New York, 1999.

Lockner, D. A., J. D. Byerlee, V. Kuksenko, A. Ponomaev, and A. Sidorin, Quasi-static growth and shear fracture energy in granite, *Nature*, *350*, 39–42, 1991.

Rubin, A. M., D. Gillard, and J. L. Got, A reinterpretation of seismicity associated with the January 1983 dike intrusion at Kilauea Volcano, Hawaii, *J. Geophys. Res.*, *103*, 10,003–10,015, 1998.

Scarpa, R., and R. I. Tilling (Eds), Monitoring and mitigation of volcano hazards, 841 pp., Springer-verlag, Berlin, 1996.

Sornette, D., and C. G. Sammis, Complex critical exponents from renormalization group theory of earthquakes: implications for earthquake predictions, *J. Phys. I, France*, *5*, 607–619, 1995.

Swanson, D. A., Magma supply rate at Kilauea volcano, 1952–1971, *Science*, *175*, 169–170, 1972.

Theiler, J., S. Eubank, A. Longtin, B. Galdrikian, and J. D. Farmer, Testing for non-linearity in time series: The method of surrogate data, *Physica D*, *58*, 77, 1992.

Voight, B., A method for prediction of volcanic eruptions., *Nature*, *332*, 125–130, 1988.

Wolfe, E. W., (ed), The Pu'u 'O'o eruption of Kilauea Volcano: U.S., *Geological Survey Professional Paper*, *1463*, 251 p., 1988.

————————————

S. F. M. Chastin and I. Main, Edinburgh University, Geology and Geophysics Department, Grant Institute, Kings Building, West Mains Road, Edinburgh, EH9 3JW, Scotland, UK. (Sebastien.Chastin@glg.ed.ac.uk; Ian.Main@glg.ed.ac.uk)