Forecasting volcanic eruptions and other material failure phenomena: An evaluation of the failure forecast method

Andrew F. Bell,[1] Mark Naylor,[1] Michael J. Heap,[2] and Ian G. Main[1]

Received 16 May 2011; revised 13 June 2011; accepted 27 June 2011; published 6 August 2011.

[1] Power-law accelerations in the mean rate of strain, earthquakes and other precursors have been widely reported prior to material failure phenomena, including volcanic eruptions, landslides and laboratory deformation experiments, as predicted by several theoretical models. The Failure Forecast Method (FFM), which linearizes the power-law trend, has been routinely used to forecast the failure time in retrospective analyses; however, its performance has never been formally evaluated. Here we use synthetic and real data, recorded in laboratory brittle creep experiments and at volcanoes, to show that the assumptions of the FFM are inconsistent with the error structure of the data, leading to biased and imprecise forecasts. We show that a Generalized Linear Model method provides higher-quality forecasts that converge more accurately to the eventual failure time, accounting for the appropriate error distributions. This approach should be employed in place of the FFM to provide reliable quantitative forecasts and estimate their associated uncertainties. **Citation:** Bell, A. F., M. Naylor, M. J. Heap, and I. G. Main (2011), Forecasting volcanic eruptions and other material failure phenomena: An evaluation of the failure forecast method, *Geophys. Res. Lett.*, *38*, L15304, doi:10.1029/2011GL048155.

1. Introduction

[2] Accurate forecasts of the timing of natural hazard events, and reliable quantification of the associated forecast uncertainties, will enable more efficient application of hazard mitigation measures. Many material failure phenomena, such as volcanic eruptions [*Voight*, 1988; *Voight and Cornelius*, 1991; *Cornelius and Voight*, 1994; *Kilburn and Voight*, 1998; *Kilburn*, 2003; *Smith et al.*, 2007], landslides [*Kilburn and Petley*, 2003; *Petley et al.*, 2005], and failure of samples in the laboratory [*Lavallée et al.*, 2008; *Heap et al.*, 2009; *Smith et al.*, 2009], are preceded by clear accelerating rates of strain and seismicity. It has been widely suggested that these precursory signals could be the basis for forecasting the time of failure. *Voight* [1988] proposed a relation between the acceleration in a geophysical precursor Ω (such as strain or number of earthquakes) and its rate for conditions of constant stress and temperature:

$$\frac{d^2\Omega}{dt^2} = K\left(\frac{d\Omega}{dt}\right)^{\alpha} \qquad (1)$$

where α and K are constants, and α observed to take values between 1 and 2 [*Kilburn*, 2003]. In the general case that $\alpha \neq 1$, solutions to Voight's relation involving positive acceleration take the form of a power-law increase in the rate of precursory signals with time [*Voight*, 1988; *Kilburn and Voight*, 1998; *Main*, 1999]:

$$\frac{d\Omega}{dt} = k\left(t_f - t\right)^{-p} \qquad (2)$$

where t_f is the failure time, $p = 1/(\alpha - 1)$ is a power-law exponent and k is a multiplicative amplitude term. This solution involves a singularity at a finite time, corresponding to an (unphysical) instantaneously infinite rate, and interpreted as the likely time of bulk sample failure or eruption onset. Equation (2) can be solved by linearizing the problem in the form:

$$\left(\frac{d\Omega}{dt}\right)^{\frac{-1}{p}} = k^{\frac{-1}{p}}\left(t_f - t\right) \qquad (3)$$

and using standard least squares regression to determine the failure time [*Voight*, 1988; *Voight and Cornelius*, 1991]. This application of Voight's relation is known as the Failure Forecast Method (FFM) [*Cornelius and Voight*, 1994]. Commonly $p \approx 1$ [*Kilburn*, 2003], in which case the solution is a straightforward regression of inverse rate against time. However, although the FFM has widely been applied (retrospectively) to volcanic [*Voight*, 1988; *Kilburn and Voight*, 1998; *Kilburn*, 2003; *Smith et al.*, 2007], landslide [*Fukuzono*, 1985; *Kilburn and Petley*, 2003; *Petley et al.*, 2005] and laboratory data [*Lavallée et al.*, 2008; *Smith et al.*, 2009], the validity of its assumptions and its forecasting performance have never been formally evaluated.

[3] For least-squares linear regression to be valid, the residual error between the data and the mean rate must follow a Gaussian distribution. This is in general a valid assumption for measurement errors in the primary strain rates. In contrast earthquake occurrence is a point process whose rate uncertainties (in the absence of earthquake triggering) are well-described by a Poisson distribution [*Greenhough and Main*, 2008]. In both cases it is extremely unlikely that the resulting scatter in the transformed rate is Gaussian. In this paper we investigate methods for applying Voight's relation. We first use simulated data where, unlike real data, the parent distributions are known *a priori*, to show that the FFM gives biased and imprecise forecasts of failure times. We then describe a Generalized Linear Model (GLM) method for applying Voight's relation, which returns higher quality (more accurate and precise) forecasts. Finally,

[1]School of GeoSciences, University of Edinburgh, Edinburgh, UK.
[2]Laboratoire de Géophysique Expérimentale, Institut de Physique de Globe de Strasbourg, UMR 7516 CNRS, Université de Strasbourg/EOST, Strasbourg, France.

0094-8276/11/2011GL048155

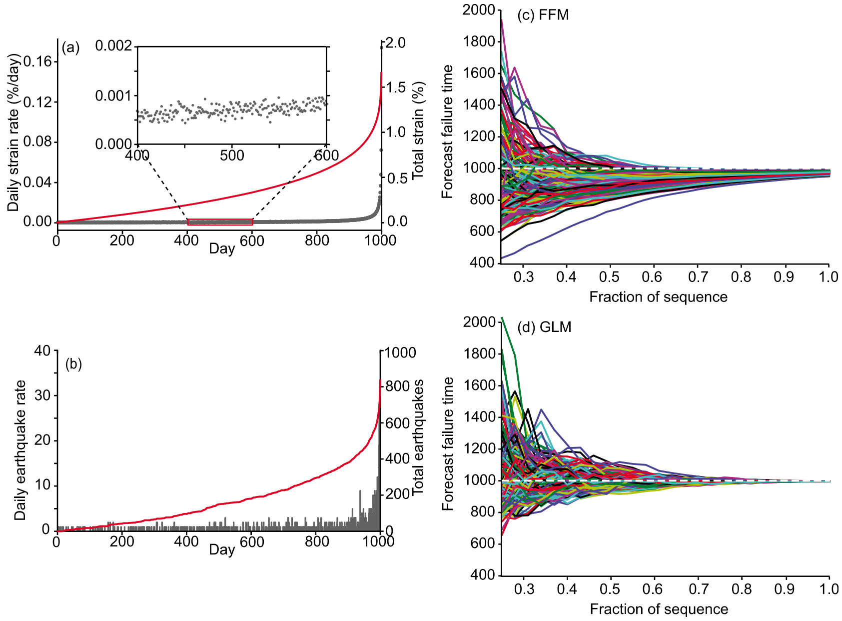

Figure 1. (a and b) Simulated sequences with mean rate given by equation (2). Grey points or bars are daily rates; red line is cumulative strain or number of earthquakes. Figure 1a shows strain, with Gaussian noise with standard deviation 0.0001. Inset shows detail of sequence for 200 day subset. Figure 1b shows earthquakes. In both instances $p = 0.8$, the failure time $t_f = 1000$. (c and d) Forecast failure time as a function of the fraction of total duration for 200 simulated power-law accelerations in the mean strain rate as for Figure 1a. Figure 1c shows FFM and Figure 1d shows GLM methods. Each colored line represents the evolution of the forecast for a single synthetic sequence. Dashed white line is the true failure time of 1000 days.

we compare the performance of the two methods on real data from laboratory experiments and volcanoes.

2. Comparison of the FFM and GLM Method

[4] We evaluate the FFM by applying it to synthetic strain and earthquake sequences where the mean rate evolves according to equation (2), each generated with identical parameters and with a failure time of 1000 days. In the case of strain-rate data we add a small amount of Gaussian noise to simulate the effect of measurement errors (Figure 1a). We simulate earthquake data as a Poisson process with mean rate and variance, λ, evolving as equation (2) (Figure 1b). Figure 1c illustrates how the failure time predicted using the FFM evolves for each of the 200 simulated strain rate sequences each modelled as in the example in Figure 1a. There is a degree of convergence as more data is fitted, but the convergence is slow, and the method systematically underestimates the failure time in all of the simulations in the last third of the time period, up to and including the actual failure time. Thus the method is biased and imprecise, even for retrospective analysis.

[5] The Generalized Linear Model (GLM) method is a generalization of least-squares linear regression which can account for (1) a non-Gaussian distribution of errors from the mean (e.g. Poisson) and (2) a functional relation (the "link function", e.g. power-law) between the mean of the distribution and a basic linear model [*Nelder and Wedderburn*, 1972]. GLMs are commonly applied using maximum-likelihood, iteratively re-weighted means. In the case of a Gaussian error distribution and identity link function, the GLM method equates to least-squares linear regression. GLM routines are available in common computing packages (e.g. Python, Matlab and R).

[6] Equation (2) can be re-arranged to express the mean rate as a power-law function of a linear model, without changing the error distribution:

$$\frac{d\Omega}{dt} = (a + bt)^{-p} \qquad (4)$$

where $a = (k_p)^{\frac{-1}{p}} t_f$ and $b = -(k_p)^{\frac{-1}{p}}$. Consequently, it is possible to determine a and b using a GLM, specifying a power-law link function and a Gaussian or Poisson error structure for strain rates and seismicity rates, respectively. The forecast failure time is then given by $t_f = -a/b$. Figure 1d shows how the expected time of failure from the GLM method changes as the precursory sequence evolves for each of 200 simulated strain rate sequences. In contrast to the FFM, the forecast error represented by the scatter in the inferred failure time decreases much more quickly as the sequences proceed and the forecasts converge on the true failure time. The GLM clearly outperforms the FFM in the final third of the sequences.

[7] To illustrate this further Figure 2 compares probability density functions for the error in the forecast time of failure for 1000 synthetic strain and earthquake sequences as for Figures 1a and 1b. To fit to the earthquake data, where the Poisson process results in many days without any events, rates are determined for 10 evenly spaced time bins (as in previous application of the FFM to volcanic earthquakes [e.g., *Voight*, 1988; *Kilburn*, 2003]). Forecasts are made after 750, 850 and 950 days. For both FFM and GLM method, the variance of the forecast errors decreases with

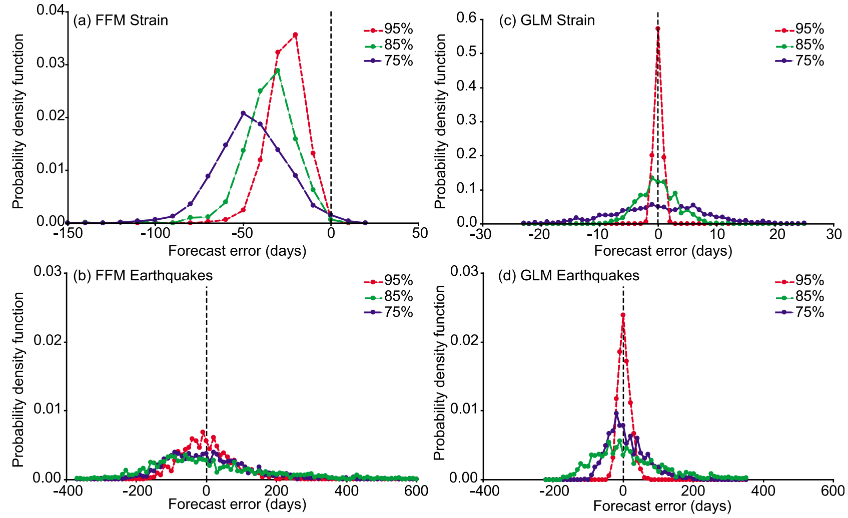

Figure 2. Probability density functions for the forecast error (in days, predicted-actual) at 750, 850 and 950 days for 1000 simulated sequences with failure time of 1000 days. (a) FFM on strain rates with Gaussian noise (see Figure 1c). (b) FFM on earthquake rates. (c) GLM on strain rates with Gaussian noise (see Figure 1d). (d) GLM on earthquake rates. Note the significant change in horizontal and vertical scales between Figures 2a and 2c.

time (as more data is used to make the forecast and the rate increases). The variances for the FFM are large and the mean forecast time is consistently earlier than the true failure time (Figures 2a and 2b). For the GLM method, the variances are much reduced and the mean error is close to zero (Figures 2c and 2d). For the strain rate data, the variance is very small, reflecting the large number of data points and relatively low levels of Gaussian noise. The variance increases as the standard deviation of the added noise is increased. The spread of predicted failure times is larger for the earthquake rate data, reflecting the relatively few data points and higher stochastic variability inherent in the Poisson process. In summary, Figures 1 and 2 demonstrate that the GLM produces higher quality forecasts when

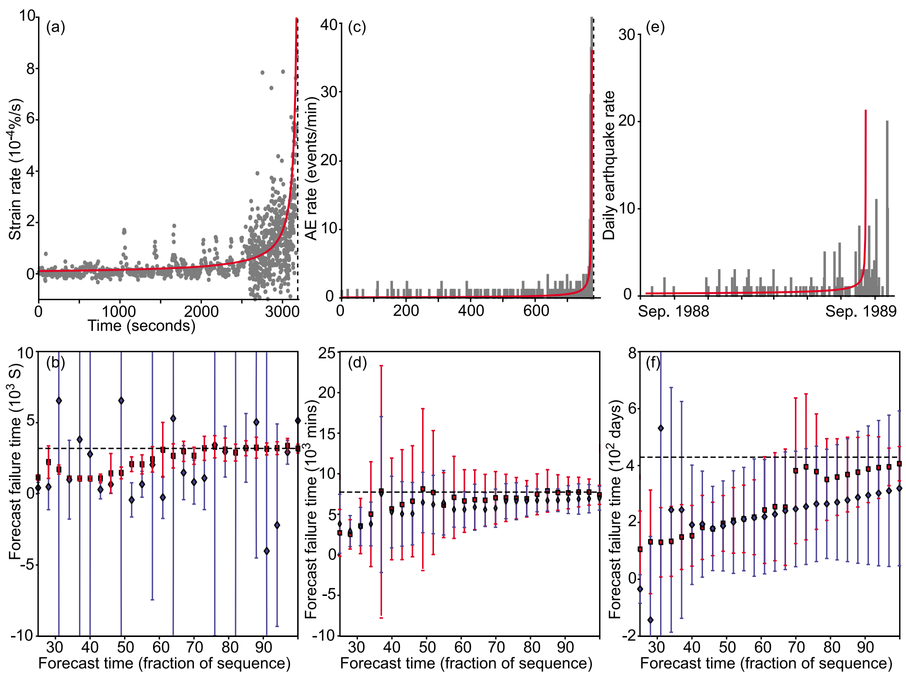

Figure 3. Application of the FFM and GLM methods to real data. (a, c, and e) Strain, AE or earthquake rates (grey points or bars) and example GLM models determined at 95% of true failure time (red line). (b, d, and f) temporal evolution of forecast failure times and standard errors using the FFM (blue diamonds) and GLM method (red squares). Black dashed line is true failure time. Data are for strain during a brittle creep experiment on Darley Dale Sandstone (Figures 3a and 3b), AE during a brittle creep experiment on Darley Dale Sandstone (Figures 3c and 3d), and earthquakes preceding the September 1989 flank eruption of Mt Etna (Figures 3e and 3f).

applied to the same data, being both more precise and less biased.

3. Application of the GLM Method to Real Data

[8] We first compare the forecasting performance of the FFM and GLM method to examples of laboratory brittle creep experiments performed on samples of Darley Dale Sandstone [*Heap et al.*, 2009]. The final stages of these experiments are characterised by accelerating strain and Acoustic Emission (AE) rates; theoretical models [e.g., *Main*, 2000] suggest that rates should evolve according to a power-law acceleration (equation (2)). We fit the models to the final third of the experimental data, defined by the period when acceleration to failure might be expected to have started. Figure 3a shows the strain rates and an example GLM fitted at 95% of the sequence duration for an experiment with a steady-rate strain rate of $10^{-7}s^{-1}$. Analysis of the entire sequence shows that it is best modelled with a p-value of 1.0 (or $\alpha = 2$) and we assume this value for application of the two forecast methods. Figure 3b shows the predicted failure times evolve with time during the sequence. By inspection, the GLM method clearly outperforms the FFM, being both more precise and less biased with respect to the actual failure time. Figure 3c shows the AE rates and an example GLM fitted at 95% of the sequence duration for a second experiment with a steady-rate strain rate of $10^{-8}s^{-1}$. We apply the two forecast methods, in this instance using a previously determined p-value of 0.9 [*Bell et al.*, 2011]. Again, the GLM outperforms the FFM (Figure 3d).

[9] We now compare the forecasting performance of the FFM and GLM methods on real volcanic data using the sequence of volcano-tectonic earthquakes preceding the 1989 eruption at Mt Etna [*Vinciguerra*, 2002]. Figure 3e shows the earthquake rates and an example GLM fitted at 95% of the sequence duration. We apply the two forecast methods using a previously determined p-value of 0.6 [*Bell et al.*, 2011]. In this example both methods consistently under-predict the time of the eruption, though the GLM method predicts a failure time much closer to the true failure time, and with higher precision, than the FFM (Figure 3f). The FFM routinely forecasts a failure time earlier than the time at which the forecast was made.

4. Discussion

[10] Our results clearly show that the FFM is biased and imprecise, even for retrospective data analysis. The transformation of the rate data means that the error distribution assumed in the least-squares linear regression component of the FFM is not valid, nor are any metrics derived from the regression lines. The GLM method provides better quality forecasts (more precise and less biased) of the failure time and its uncertainty, and with modern statistical packages is equally easy to apply. Here we focus on the method of forecasting when the value of p is known *a priori*. If this is not the case, it is possible to fit many GLMs using different trial values of p, and use a Monte-Carlo approach to delineate the optimal forecast failure time and its confidence limits.

[11] We consider a Gaussian or Poissonian distribution to be realistic first approximations for the error distribution of strain or earthquake rates, respectively, about the mean trend predicted by Voight's relation. It is possible that additional physical processes may result in fluctuations in strain or earthquake rate about the mean rate that differ to those expected for a Gaussian or Poissonian distribution (as is the effect of earthquake triggering in tectonic seismicity). At volcanoes and in the laboratory, strain, earthquake and AE event rate error distributions are currently not well characterized. These distributions, and their implications for the application of Voight's relation, should be an important area for future work. Nevertheless, here we have demonstrated that the GLM method assuming a Gaussian or Poissonian distribution produces a more stable estimation of the failure time than the FFM on examples using real data.

5. Conclusions

[12] The FFM for forecasting on the basis of Voight's relation is imprecise and systematically under-estimates failure times based on accelerating mean rates of strain or earthquakes. This poor performance is not due to a failing of Voight's relation, but because the assumed error structure after linearization is not valid. On synthetic data the GLM method, assuming an error structure can be approximated by a Gaussian or Poissonian distribution, provides substantially higher quality (less biased and more precise) forecasts and a reliable measure of their uncertainty. The GLM method also out performs the FFM on real example datasets from deformation experiments in the laboratory and preceding volcanic eruptions. Further work is necessary to quantify the forecasting power of the GLM method under truly prospective conditions and where the Gaussian and Poissonian error distributions approximations are not met.

[13] **Acknowledgments.** This work was funded by NERC standard award Ne/H02297x/1 (AB), NERC studentship NER/S/A2005/13553 (MJH) and a Scottish Government and Royal Society of Edinburgh fellowship (MN). We gratefully acknowledge Neil Hughes, Steve Boon, John Bowles, Patrick Baud, Phil Meredith and two anonymous reviewers.

[14] The Editor thanks an anonymous reviewer for their assistance in evaluating this paper.

References

Bell, A. F., J. Greenhough, M. J. Heap, and I. G. Main (2011), Challenges for forecasting based on accelerating rates of earthquakes at volcanoes and laboratory analogues, *Geophys. J. Int.*, *185*, 718–723, doi:10.1111/j.1365-246X.2011.04982.x.

Cornelius, R. R., and B. Voight (1994), Seismological aspects of the 1989–1990 eruption at Redoubt Volcano, Alaska: The materials failure forecast method (Ffm) with RSAM and SSAM seismic data, *J. Volcanol. Geotherm. Res.*, *62*(1–4), 469–498, doi:10.1016/0377-0273(94)90048-5.

Fukuzono, T. (1985), A method to predict the time of slope failure caused by rainfall using the inverse number of velocity of surface displacement, *J. Jpn. Landslide Soc.*, *22*, 8–13.

Greenhough, J., and I. G. Main (2008), A Poisson model for earthquake frequency uncertainties in seismic hazard analysis, *Geophys. Res. Lett.*, *35*, L19313, doi:10.1029/2008GL035353.

Heap, M. J., P. Baud, P. G. Meredith, A. F. Bell, and I. G. Main (2009), Time-dependent brittle creep in Darley Dale sandstone, *J. Geophys. Res.*, *114*, B07203, doi:10.1029/2008JB006212.

Kilburn, C. R. J. (2003), Multiscale fracturing as a key to forecasting volcanic eruptions, *J. Volcanol. Geotherm. Res.*, *125*(3–4), 271–289, doi:10.1016/S0377-0273(03)00117-3.

Kilburn, C. R. J., and D. N. Petley (2003), Forecasting giant, catastrophic slope collapse: Lessons from Vajont, northern Italy, *Geomorphology*, *54*(1–2), 21–32, doi:10.1016/S0169-555X(03)00052-7.

Kilburn, C. R. J., and B. Voight (1998), Slow rock fracture as eruption precursor at Soufriere Hills volcano, Montserrat, *Geophys. Res. Lett.*, *25*(19), 3665–3668, doi:10.1029/98GL01609.

Lavallée, Y., P. G. Meredith, D. B. Dingwell, K. U. Hess, J. Wassermann, B. Cordonnier, A. Gerik, and J. H. Kruhl (2008), Seismogenic lavas and explosive eruption forecasting, *Nature*, *453*(7194), 507–510, doi:10.1038/nature06980.

Main, I. G. (1999), Applicability of time-to-failure analysis to accelerated strain before earthquakes and volcanic eruptions, *Geophys. J. Int.*, *139*(3), F1–F6, doi:10.1046/j.1365-246x.1999.00004.x.

Main, I. G. (2000), A damage mechanics model for power-law creep and earthquake aftershock and foreshock sequences, *Geophys. J. Int.*, *142*(1), 151–161, doi:10.1046/j.1365-246x.2000.00136.x.

Nelder, J. A., and R. W. M. Wedderburn (1972), Generalized linear models, *J. R. Stat. Soc., Ser. A*, *135*(3), 370–384.

Petley, D. N., T. Higuchi, D. J. Petley, M. H. Bulmer, and J. Carey (2005), Development of progressive landslide failure in cohesive materials, *Geology*, *33*(3), 201–204, doi:10.1130/G21147.1.

Smith, R., C. R. J. Kilburn, and P. R. Sammonds (2007), Rock fracture as a precursor to lava dome eruptions at Mount St Helens from June 1980 to October 1986, *Bull. Volcanol.*, *69*(6), 681–693, doi:10.1007/s00445-006-0102-5.

Smith, R., P. R. Sammonds, and C. R. J. Kilburn (2009), Fracturing of volcanic systems: Experimental insights into pre-eruptive conditions, *Earth Planet. Sci. Lett.*, *280*(1–4), 211–219, doi:10.1016/j.epsl.2009.01.032.

Vinciguerra, S. (2002), Damage mechanics preceding the September-October 1989 flank eruption at Mount Etna volcano inferred by seismic scaling exponents, *J. Volcanol. Geotherm. Res.*, *113*(3–4), 391–397, doi:10.1016/S0377-0273(01)00274-8.

Voight, B. (1988), A method for prediction of volcanic eruptions, *Nature*, *332*(6160), 125–130, doi:10.1038/332125a0.

Voight, B., and R. R. Cornelius (1991), Prospects for eruption prediction in near real-time, *Nature*, *350*(6320), 695–698, doi:10.1038/350695a0.

A. F. Bell, M. Naylor, and I. G. Main, School of GeoSciences, University of Edinburgh, Edinburgh EH9 3JW, UK. (a.bell@ed.ac.uk)

M. J. Heap, Laboratoire de Géophysique Expérimentale, Institut de Physique de Globe de Strasbourg, UMR 7516 CNRS, Université de Strasbourg/EOST, 5 rue René Descartes, F-67084 Strasbourg CEDEX, France.