
Nick S. Roberts1, Andrew F. Bell1, and Ian G. Main1

1School of Geosciences, University of Edinburgh, Edinburgh, UK

Abstract The Gutenberg-Richter $b$ value is commonly used in volcanic eruption forecasting to infer material or mechanical properties from earthquake distributions. Such studies typically analyze discrete time windows or phases, but the choice of such windows is subjective and can introduce significant bias. Here we minimize this sample bias by iteratively sampling catalogs with randomly chosen windows and then stack the resulting probability density functions for the estimated $b$ value to determine a net probability density function. We examine data from the El Hierro seismic catalog during a period of unrest in 2011–2013 and demonstrate clear multimodal behavior. Individual modes are relatively stable in time, but the most probable $b$ value intermittently switches between modes, one of which is similar to that of tectonic seismicity. Multimodality is primarily associated with intermittent activation and cessation of activity in different parts of the volcanic system rather than with respect to any systematic inferred underlying process.

1. Introduction

The $b$ value of the Gutenberg-Richter relation, $\log(N) = a - bM$ [Gutenberg and Richter, 1954], describes the relative proportions of large- and small-magnitude earthquakes in a catalog. Theoretical and experimental studies suggest that $b$ values are influenced by a variety of factors, including stress [Scholz, 1968], mechanical properties [Schorlemmer et al., 2005], thermal gradient [Warren and Latham, 1970], pore fluid pressure [Sammonds et al., 1992; Raleigh et al., 1976], and material damage. The $b$ value for tectonic earthquakes, using best practice and large regional or global data sets, is commonly reported as $b = 1$ [Frolich and Davis, 1993]. In contrast the reported $b$ values from published studies of earthquake populations associated with volcanic unrest are commonly reported as being significantly higher than this [Roberts et al., 2015]. At volcanoes, the temporal evolution of the $b$ value has been used to infer changes in the physical processes controlling the approach to eruption, including material failure, and have been proposed as a potential forecasting tool. However, existing methods typically calculate $b$ within either a series of independent finite-time windows or overlapping fixed-width moving windows. In order to achieve the necessary fine-scale resolution of $b$ value changes, these studies often use small subcatalog sizes and assume a single value of the completeness magnitude [Ibanez et al., 2012; Lopez et al., 2012; Marti et al., 2013b]. Consequently, the errors in the $b$ values are likely to be large, correlated, and underestimated, and potentially give rise to biased results [Roberts et al., 2015]. At this stage of modeling we do not consider the effect of uncertainties in the individual magnitudes, so our error estimates are all underestimated to some extent.

Here, for the first time, we show the full probability distribution of the $b$ value for a volcanic earthquake catalog as it evolves with time. This result is achieved by combining two new methods. First, a stochastic windowing technique is used to recover fine-scale resolution of $b$ value changes, while avoiding arbitrary choices of window edges and small subcatalog sizes. Second, a more realistic uncertainty estimate for the $b$ value is determined for each subcatalog by joint estimation with the completeness magnitude. This process is repeated many times, and the resulting $b$ value samples and their errors combined to construct the full $b$ value probability distribution. A key benefit of this approach is that it is able to resolve different $b$ values associated with contemporaneous processes in the case where some generate high rates of events for short durations and others low rates for longer durations, characteristics that are typical for many volcanic processes.

2. Methods

2.1. Iterative Sampling Methodology

Figure 1 shows the inferred $b$ value and its error distribution obtained from finite samples of a synthetic catalog with a parent Gutenberg-Richter distribution above $M_c = 1.0$, with $b = 1$ for the first and last 5000 events...
and $b = 2$ for the middle 5000. Our new method clearly recovers the underlying $b$ values within the uncertainties involved. The method is as follows. First, the catalog is divided into subcatalogs of random size between a predefined minimum and maximum number of events of 50 and 1000 events, respectively, as justified in the “Choice of parameters” section in the supporting information. Three iterations are illustrated in Figure 1a. For each subcatalog $M_c$ is selected using the workflow in Figure 9 of Roberts et al. (2015). The $\tilde{b}$ value with associated error $\sigma_{\tilde{b}}$ is calculated for the complete part of the subcatalog using equations S2 and S3 (supporting information). The error multiplication factor $R(\tilde{b}, N)$ from Figure 11 of Roberts et al. (2015) is then used to determine the total error $R\sigma_{\tilde{b}}$, i.e., including the uncertainty introduced in estimating $M_c$. The average time

![Figure 1.](image)

The synthetic catalog has a parent Gutenberg-Richter distribution above $M_c = 1.0$, with $b = 1$ for the first and last 5000 events and $b = 2$ for the middle 5000. (a) Each event is shown as a black dot. The red, blue, and green lines show the boundaries of each catalog in three successive iterations. (b) Cloud of data points produced by 100 iterations. (c) The $b$ value probability density by event number. The dashed grey line shows the underlying known $b$ value. The dotted black line marks $b_\nu$. The average time
and event number for the subcatalog is then calculated, allowing the $b$ value to be plotted as a function of either. Figure 1b shows the cloud of data points produced by 100 iterations. The advantage of plotting by event number is it normalizes the event rate making it easier to see variations when the event rate is high. This process is then iterated a desired number of times, each time using different random subcatalog sizes. Following enough iterations, a cloud of data points spanning the whole catalog is created.

For this method to be used as a real-time analysis tool, it is important that we are able to identify $b$ value variations in the most recent events. To make sure that the most recent events are always sampled, the catalog is divided into subcatalogs by starting with the youngest event then working back in time. This means that the oldest events may be undersampled if the number of remaining events is less than the minimum sample size.

2.2. Converting to a Probability Density Function

Each data point for a given randomly generated window has a $b$ value, an error, $\sigma_b$, and a multiplication factor $R$. For consistency with equation (53—supporting information), we use the parametrized version of the normal distribution to calculate the relative probability for any given $b$ value, $b_i$.

$$P(b_i) = \frac{1}{R\sigma_b\sqrt{2\pi}} e^{-\frac{(b_i - \bar{b})^2}{2\sigma_b^2}} \quad (1)$$

$P(b_i)$ is calculated over a $b$ value range from 0 to 4, at a resolution of 0.01 units.

2.3. Monte Carlo Simulation of Error Structure Using Moving Windows

In principle, stacking enough individual probability density functions (PDFs) at randomly sampled times should reveal the true structure of the overall PDF. However, from equation (1) $b$ values with smaller error contribute more to a stacked curve. As the size of the error is
proportional to the $b$ value (equation S3), higher $b$ values will also have higher errors (Figure 1). Therefore, care must be taken not to stack too many data points so that brief periods of high $b$ will not be hidden due to over smoothing.

To stack the PDFs, we take a fixed number of data points and sum the individual PDFs for all $b$ and then normalize the summed curve so that $\sum_{i=0}^{n} P(b_i) = 1$. This allows us to compare the peak probability for all the stacked PDFs. We take the time stamp for a given window to be, respectively, the average event number and average time and do this for all times and event numbers. Figure 2 shows sensitivity testing for 10, 20, 50, 100, and 200 data points in each window. We chose 50 data points as the ideal trade-off for resolution and smoothness. The results are visualized as a contour plot in probability-time space rather than as error bars at discrete points (both shown in Figure 3).

3. El Hierro Catalog

After testing and benchmarking on synthetic catalogs (Figures 1 and 2 are further discussed in the “Synthetic catalog testing” section in the supporting information), we applied the new method to real seismicity data from El Hierro volcano, Canary Islands [Ibanez et al., 2012; Marti et al., 2013b; Becerril et al., 2013; López et al., 2012; García et al., 2014] to generate a temporally evolving $b$ value probability density function (PDF), denoted $P(b)$. We analyze the Instituto Geográfico Nacional (IGN) earthquake catalog for El Hierro between July 2011 and December 2013—a period containing over 20,000 earthquakes associated with magma emplacement and a submarine eruption that began on 10 October 2011 [Ibanez et al., 2012; López...
et al., 2012]. An epicentral and hypocentral map of these events is shown in supporting information Figure S5, and the magnitudes and daily events rates are plotted in supporting information Figure S6. Some 4100 individual \( b \) values were generated over 100 iterations; these were then converted first to Gaussian PDFs for individual windows and then stacked to determine the net \( P(b) \) using a 50-point moving window.

4. Results

The results are shown as a function of both event number (Figure 3a) and time (Figure 3b). We use both plots because it is possible to see the fine structure better on the event number plot during times of high event rate. The periods with red colors in Figure 3 have the highest probability density \( P(b) \). The \( b \) value with peak probability, \( \tilde{b}_{P} \), is marked with a black dotted line to show how the maximum likelihood \( b \) value varies through time.

The results show \( \tilde{b}_{P} \) typically being between 1.0 and 1.5 and that it is best constrained when \( \tilde{b}_{P} \approx 1 \). Prior to the eruption \( \tilde{b}_{P} \) can be as high as 3.25 with several periods of \( \tilde{b}_{P} > 2.0 \). However, the most striking aspect of Figure 3 is the rapid switching between alternately higher and lower \( \tilde{b}_{P} \). The \( \tilde{b}_{P} \) value does not increase or decrease smoothly through the catalog; it jumps between values and then remains relatively constant until it changes again. This behavior is contrary to experimental observations, where \( \tilde{b} \) values evolve more gradually with time [Henderson et al., 1994; Main, 1996] and more reminiscent of a “mode-switching” process [Ben-Zion et al., 1999] observed in some models for tectonic earthquakes. Post-eruption, the \( \tilde{b} \) values stabilize at \( \tilde{b} = 1 \) for several months until a period from July 2012 to April 2013 where the \( \tilde{b} \) value is very poorly constrained, with \( P(\tilde{b}) < 0.020 \). The net \( P(\tilde{b}) \) is clearly multimodal, e.g., with persistent multiple local peaks in the range 0.6–2.5, in the time period October 2012 to December 2012.

Figure 3b also compares \( \tilde{b}_{P} \) with inferred \( \tilde{b} \) values obtained by conventional finite-time windows, using phase boundaries based on previously reported changes in hypocenter location [Roberts et al., 2015]. The uncertainty structure is much richer and more informative for the new method. While the previously inferred \( \tilde{b} \) values and uncertainties are not unreasonable using the conventional windowing method, they do not capture the multimodal behavior and mode switching in \( \tilde{b}_{P} \). There are also significant periods when \( \tilde{b}_{P} \) does not agree with \( \tilde{b} \) within its uncertainty, shown as two standard deviations, or approximately 95% confidence. In September 2012 and March 2013 \( \tilde{b}_{P} \) is systematically above this two standard deviation window, and through October–December 2012 and in May 2013, it is systematically below it.

Our method allows for the analysis of the full PDF, rather than assuming a Gaussian distribution [Aki, 1965], giving a much greater insight into the error structure and its potential multimodal character. The time or event number associated with each stacked window is taken to be the average event number or time, so it is possible to examine snapshots of the net PDF for any given event or time. Figure 4 shows three examples at different times corresponding to three significant events: (i) where \( \tilde{b}_{P} \) is very high (>3), (ii) where \( \tilde{b}_{P} \approx 1 \), and (iii) where \( \tilde{b}_{P} \) is multimodal. These examples are indicated in Figures 3 and 5. Event (i) has the highest \( \tilde{b}_{P} \) in the whole catalog where \( \tilde{b}_{P} = 3.25 \), with a broad, low-amplitude peak value \( P(\tilde{b}_{P}) \) of only ~0.006. There is a subsidiary peak at \( \tilde{b} \approx 1.3 \) corresponding to the baseline \( b \) value both before and after the period of very high \( \tilde{b}_{P} \). The window when \( \tilde{b}_{P} = 3.25 \) is only a few hundred events wide. This would be very hard to identify by eye in a time series because it occurs during a period with very high event rates.

In the post-eruption phase (Figure 4ii) \( \tilde{b} \) is very well constrained to \( \tilde{b}_{P} = 1.0 \), similar to that of tectonic seismicity. There is a single peak with a narrow symmetric base, indicating simple unimodal behavior. Figure 4iii shows the net PDF associated with event (iii) in December 2012. The PDF is clearly multimodal, with five notable peaks at \( \tilde{b} = 0.5, 0.7, 1.1, 1.7 \), and 2.6, although the peak at 0.5 has a very small cumulative probability (<1%) associated with it and hence may not be significant. The question arises, what might be the underlying cause of the four peaks that do make a significant contribution to the cumulative probability?

Closer inspection of the seismic catalog reveals that there are four clusters (A–D) of events that can be separated spatially (Figure 5) as well as temporally, with A starting first, B and C second at the same time but at
different depths, and then finally D (supporting information Figure S4). All clusters then ended at approximately the same time. Accordingly, we determined \( \hat{b} \) and its error bar for the four clusters separately. Clusters A–C contain >500 events in the incomplete catalog, comparable with an average sampled catalog size of 525. Accordingly, there is sufficient data to identify statistically distinct \( \hat{b} \) values if the underlying values are different. Cluster D contains 113 events, so although it will be much harder to sample, the minimum catalog size of 50 events should allow for it to be suitably represented in the PDF, given enough sampling iterations.

Clusters B and D both have complete catalog sizes significantly below the advised threshold [Roberts et al., 2015] of \( N_c \geq 200 \) (89 and 38, respectively), so the results should be treated with caution. However, the calculated \( \hat{b} \) values for all four clusters fit well within error to the four most significant peaks in the PDF (Figure 4). The mode-switching behavior can then be attributed to different parts of the volcanic system being activated

![Figure 4](image1.png)

**Figure 4.** Cross sections of the \( b \) value probability density function marked in Figure 3, using a 50 data point window. Colored error bars in (iii) represent the \( \hat{b} \) values calculated for the clusters in Figure 5.

![Figure 5](image2.png)

**Figure 5.** Epicentral and hypocentral locations of earthquakes above 0.0 \( M \) in Phase 6 (Table 1). Each color marks a separate cluster of events, A–D. All clusters ceased activity by 18 March 2013. Cluster A started on 13 March 2012. Clusters B and C started on 14 September 2012. Cluster D began on 31 December 2012. Events colored grey were judged not to be part of any of the clusters of events. The colored polylines show where the majority of events in phases 1–5, 7, and 8 are clustered. The black and grey ovals and lines in the cross section indicate the relative positions of the magma chambers described in Marti et al. [2013a, 2013b]. The direction of migration in Phase 2 is marked with an arrow. In map view, the solid black line shows the shoreline.
Table 1. Phases of Seismic Activity at El Hierro From July 2011 to December 2013 (Figure 5)

<table>
<thead>
<tr>
<th>Phase No.</th>
<th>Observation</th>
<th>Inferred Process</th>
<th>Time Range</th>
<th>Event No. Range</th>
<th>Modal b</th>
<th>95% b Range</th>
<th>±2σ From Ref. 18</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Spatially stationary events at 8–16 km depth.</td>
<td>Magma intrusion and accumulation in upper chamber [Marti et al., 2013b]</td>
<td>19 Jul 2011 to 20 Sep 2011 [Ibanez et al., 2012]</td>
<td>0–7,033</td>
<td>1.0, 2.3, 3.2</td>
<td>0.8–3.6</td>
<td>2.39 ± 0.39</td>
</tr>
<tr>
<td>2</td>
<td>Southward laterally migrating events at 10–18 km depth.</td>
<td>Sill migration along Moho discontinuity [López et al., 2014; Marti et al., 2013b]</td>
<td>20 Sep 2011 [Ibanez et al., 2012] to 18 Oct 2011 [Ibanez et al., 2012]</td>
<td>7,034–9,883</td>
<td>1.5</td>
<td>0.7–1.9</td>
<td>1.52 ± 0.27</td>
</tr>
<tr>
<td>3</td>
<td>Spatially stationary events at 15–25 km depth.</td>
<td>Deflation above lower chamber [Marti et al., 2013b]</td>
<td>18 Oct 2011 to 14 Dec 2011</td>
<td>9,884–11,706</td>
<td>1.0</td>
<td>0.8–1.4</td>
<td>1.11 ± 0.10</td>
</tr>
<tr>
<td>4</td>
<td>Spatially stationary events at 9–18 km depth.</td>
<td>Expansion above recharging upper chamber [Marti et al., 2013b]</td>
<td>14 Nov 2011 to 24 Jun 2012</td>
<td>11,707–12,503</td>
<td>1.0, 1.5</td>
<td>0.6–1.8</td>
<td>1.16 ± 0.16</td>
</tr>
<tr>
<td>5</td>
<td>Southwestward laterally migrating events at 17–24 km depth.</td>
<td>A magma intrusion [Marti et al., 2013b]</td>
<td>24 Jun 2012 to 13 Aug 2012</td>
<td>12,504–14,956</td>
<td>1.2, 1.5</td>
<td>0.8–2.5</td>
<td>1.38 ± 0.13</td>
</tr>
<tr>
<td>6</td>
<td>Four clusters of spatially stationary events (Figure 5) at varying depths</td>
<td>Deflation above both chambers</td>
<td>13 Aug 2012 to 18 Mar 2013</td>
<td>14,957–16,492</td>
<td>0.6, 0.7, 1.1, 1.7, 2.6</td>
<td>0.6–2.8</td>
<td>1.70 ± 0.27</td>
</tr>
<tr>
<td>6A</td>
<td>Spatially stationary events at 9–13 km depth.</td>
<td>Deflation above upper chamber</td>
<td>13 Aug 2012 to 18 Mar 2013</td>
<td>501 events</td>
<td>1.1</td>
<td>0.94–1.26</td>
<td>1.03 ± 0.10</td>
</tr>
<tr>
<td>6B</td>
<td>Spatially stationary events at 11–14 km depth.</td>
<td>Deflation above the upper chamber</td>
<td>14 Sep 2012 to 18 Mar 2013</td>
<td>645 events</td>
<td>0.7</td>
<td>0.69–0.94</td>
<td>0.89 ± 0.22</td>
</tr>
<tr>
<td>6C</td>
<td>Spatially stationary events at 17–24 km depth</td>
<td>Deflation above and to the south of the lower chamber</td>
<td>14 Sep 2012 to 18 Mar 2013</td>
<td>772 events</td>
<td>1.7</td>
<td>1.38–1.96</td>
<td>1.49 ± 0.21</td>
</tr>
<tr>
<td>6D</td>
<td>Spatially stationary events at 16–20 km depth</td>
<td>Deflation above and to the east of the lower chamber</td>
<td>31 Dec 2012 to 18 Mar 2013</td>
<td>113 events</td>
<td>2.6</td>
<td>1.96–3.12</td>
<td>2.71 ± 0.70</td>
</tr>
<tr>
<td>7</td>
<td>Westward laterally migrating events at 12–20 km depth</td>
<td>A magma intrusion</td>
<td>18 Mar 2013 to 4 Apr 2013</td>
<td>16,493–18,720</td>
<td>1.1, 2.0</td>
<td>0.8–2.5</td>
<td>1.07 ± 0.06</td>
</tr>
<tr>
<td>8</td>
<td>Spatially stationary events at 9–14 km depth</td>
<td>Deflation above upper chamber</td>
<td>4 Apr 2013 to 22 Dec 2013</td>
<td>18,721–20,385</td>
<td>0.8, 1.2, 1.5</td>
<td>0.6–1.9</td>
<td>1.26 ± 0.16</td>
</tr>
</tbody>
</table>

aDescription of the inferred process are given. The time and event ranges are stated. Phases 1–3 have been previously defined by Ibanez et al. [2012]. Modal b, and approximate 95% b range are all identified using Figure 3. The b values given in the final column are from Roberts et al. [2015] and shown in Figure 3b for comparison.
b95% range estimated from PDF producing multimodal errors in Figure 4iii.
at different times, each with an otherwise relatively stationary or slowly evolving $\tilde{b}$ value and location, rather than a systematic change in say the global stress state for the whole area of unrest that may have been the preferred interpretation using the current standard practice.

Comparison of the hypocenters and characteristics of the clusters with previously published volcanological models of the activity preceding and during the eruption [Ibanez et al., 2012; Marti et al., 2013b, 2013a; López et al., 2012, 2014] shows that Clusters C and D are likely to be associated with eastward lateral magma emplacement, originating from a magma chamber in the upper mantle at ~20 km depth. Clusters A and B at ~12 km depth occur just above the inferred location of the upper magma chamber at the Moho discontinuity, ~14 km depth. Table 1 shows the modal $\tilde{b}$P values for all eight phases between July 2011 and December 2013. The two most common modes are $\tilde{b}_P = 1.0$ and 1.5. Spatiotemporal clustering is associated with inferred changes in the magma volume-inducing events above the upper and lower chambers and lateral magma emplacement (laterally propagating clusters). There is no apparent systematic correlation between these inferred processes and $\tilde{b}_P$. Instead, $\tilde{b}_P$ changes systematically as different parts of the volcanic system are activated or deactivated, each having its own characteristic $b$ value. Changes in $\tilde{b}$ are purely spatial, and apparent changes with time are only observed because of systematically changing locations.

This ability to resolve multiple clusters of varying $\tilde{b}$ values occurring at overlapping time windows is the most important advance compared to current methodology. Conventional windowing effectively masks any multimodality and mode-switching behavior, in cases where the full net uncertainty structure is not actually Gaussian, for example, due to different parts of the volcanic system being activated at different times, as demonstrated here.

5. Conclusions

We have developed and applied a new iterative sampling method to the 2011–2013 El Hierro catalog. The method minimizes bias associated with finite sampling of time windows and reveals a complex net probability density function in the real volcanic data. We report high $\tilde{b}$ values of up to 3.25 before the main submarine eruption on 10 October 2011, followed by a relatively stable period when $\tilde{b} = 1$, i.e., similar to that of natural tectonic seismicity. From August 2012 to March 2013 we observe strongly multimodal behavior with four significant local peaks. Through further investigation into the catalog, we discover that these can be associated with spatially separate concurrent clusters of seismic activity (Figure 5) and that high $\tilde{b}$ values are not inherently linked with a specific volcanic process. Our results confirm that conventional windowing with a linear (Gaussian) error structure often provides a good first-order estimate of the $\tilde{b}$ value at a given time but lacks resolution and detail of our iterative sampling method and misses key intervals where the $\tilde{b}_P$ value is outside the estimated error.

Critically, we observe mode switching of $\tilde{b}_P$ as it jumps between otherwise stationary values. In one time period multimodality is associated with different component parts of the system being active at different times. This introduces a new possibility in interpreting $\tilde{b}$ values and may motivate an effort in physical modeling of volcanic processes to explain the mechanical bases for mode-switching behavior, as well as a reappraisal of the how $\tilde{b}$ values and their full uncertainty structure may be used in eruption forecasting.

Acknowledgments

Nick Roberts is a NERC-funded PhD student at the University of Edinburgh. We thank the Instituto Geográfico Nacional for making the seismic catalog for El Hierro available, and two anonymous reviewers whose insightful, thoughtful, and constructive feedback significantly improved the manuscript.

References


