Article:

Salvage, RO and Neuberg, JW (2016) Using a cross correlation technique to refine the accuracy of the Failure Forecast Method: Application to Soufrière Hills volcano, Montserrat. Journal of Volcanology and Geothermal Research, 324. pp. 118-133. ISSN 0377-0273

https://doi.org/10.1016/j.jvolgeores.2016.05.011

Using a cross correlation technique to refine the accuracy of the Failure Forecast Method: Application to Soufrière Hills volcano, Montserrat

R. O. Salvage[a,*], J. W. Neuberg[a]

[a] *Institute of Geophysics and Tectonics, School of Earth and Environment, University of Leeds, Leeds, LS2 9JT, United Kingdom*

Abstract

Prior to many volcanic eruptions, an acceleration in seismicity has been observed, suggesting the potential for this as a forecasting tool. The Failure Forecast Method (FFM) relates an accelerating precursor to the timing of failure by an empirical power law, with failure being defined in this context as the onset of an eruption. Previous applications of the FFM have used a wide variety of accelerating time series, often generating questionable forecasts with large misfits between data and the forecast, as well as the generation of a number of different forecasts from the same data series. Here, we show an alternative approach applying the FFM in combination with a cross correlation technique which identifies seismicity from a single active source mechanism and location at depth. Isolating a single system at depth avoids additional uncertainties introduced by averaging data over a number of different accelerating phenomena, and consequently reduces the mis-

*Corresponding Author

Email addresses: `beckysalvage@gmail.com` (R. O. Salvage), `j.neuberg@leeds.ac.uk` (J. W. Neuberg)

fit between the data and the forecast. Similar seismic waveforms were identified in the precursory accelerating seismicity to dome collapses at Soufrière Hills volcano, Montserrat in June 1997, July 2003 and February 2010. These events were specifically chosen since they represent a spectrum of collapse scenarios at this volcano. The cross correlation technique generates a five-fold increase in the number of seismic events which could be identified from continuous seismic data rather than using triggered data, thus providing a more holistic understanding of the ongoing seismicity at the time. The use of similar seismicity as a forecasting tool for collapses in 1997 and 2003 greatly improved the forecasted timing of the dome collapse, as well as improving the confidence in the forecast, thereby outperforming the classical application of the FFM. We suggest that focusing on a single active seismic system at depth allows a more accurate forecast of some of the major dome collapses from the ongoing eruption at Soufrière Hills volcano, and provides a simple addition to the well used methodology of the FFM.

Keywords: Volcano-seismology, Failure Forecast Method, low frequency, multiplets, Eruption forecasting, Soufrière Hills volcano

1. Introduction

Volcanic eruptions are often preceded by accelerating geophysical signals (McNutt, 2002), associated with the movement of magma or other fluid towards the surface. Of these precursors, seismicity is at the forefront of forecasting volcanic unrest since it is routinely observed and the change from background level can be observed in real time (Chouet et al., 1994; Cornelius and Voight, 1994; Kilburn, 2003; Ortiz et al., 2003). Since forecasting of volcanic eruptions relies upon the ability to forecast the timing of magma reaching the surface, low frequency seismicity, with a spectral range of $0.2 - 5$ Hz (Lahr et al., 1994), may potentially act as a forecasting tool since one of the interpretations of its low frequency content is its association with the movement of magmatic fluid at depth (Chouet et al., 1994; Neuberg et al., 2000).

The relationship between an accelerating geophysical precursor and the timing of failure of the system was first considered for landslides (Fukuzono, 1985) but has since been adapted for the forecasting of volcanic eruptions (Voight, 1988, 1989). The Material Failure Law or the Failure Forecast Method (FFM) as it is referred to in volcanology (Cornelius and Voight, 1995), is an empirical power-law relationship based on first principles associated with failing materials, which relates the acceleration of a precursor ($d^2\Omega/dt^2$) to the rate of that precursor ($d\Omega/dt$) at constant stress and temperature (Voight, 1988) by:

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

where K and α are empirical constants. Ω can represent a number of different geophysical precursors, for example low frequency seismic event rate (Hammer and Neuberg, 2009), event rate of all recorded seismicity (Kilburn and Voight, 1998), or the amplitude of the seismic events (Ortiz et al., 2003). The parameter

α is thought to range between 1 and 2 in volcanic environments (Voight, 1988; Voight and Cornelius, 1991), or may even evolve from 1 towards 2 as seismicity proceeds (Kilburn, 2003). An infinite $d\Omega/dt$ suggests an uncontrolled rate of change or a singularity and is associated with an impending eruption. The inverse form of $d\Omega/dt$ is linear if α = 2, and therefore the timing of failure is determined when a linear regression of inverse rate against time intersects the x-axis (Voight, 1988). It is important to note that this forecasted timing of failure is associated with the potential for an eruption due to accelerated magma flux at depth, and may not necessarily result in one, since a direct pathway of magma to the surface might be buffered.

The ongoing eruption of Soufrière Hills Volcano (SHV), Montserrat began in July 1995 with a series of phreatic explosions associated with vent openings around the crater (Young et al., 1998). Since November 1995, the andesitic volcano has undergone a repeated cycle of dome growth and collapse, with the collapse phases resulting in pyroclastic flows, lahars and ash fall events (Sparks and Young, 2002; Wadge et al., 2014). The first major dome collapse at SHV occurred in June 1997, killing 19 people (Loughlin et al., 2002), and generated pulsatory block and ash flows due to the collapse of 5×10^6 m^3 of material, removing the top 100 m of dome material (Voight et al., 1999). A number of other major dome collapses have occurred since 1995 including: 26 December 1997 (Voight et al., 2002); 29 July 2001, which lowered the dome height by over 150 m (Matthews et al., 2002); 12 July 2003, during which 210 million m^3 of material was displaced

(Herd et al., 2005); 20 May 2006 (Loughlin et al., 2010); and the most recent event on 11 February 2010 when 50 million m^3 of material was displaced (Stinton et al., 2014). As at many other volcanoes worldwide (e.g. Galeras, Colombia; Redoubt, Alaska; Mt Pinatubo, Philippines), an increase in the number of low frequency seismic events has been identified in hindsight prior to dome collapse events at SHV, in particular the event of June 1997 (Cruz and Chouet, 1997; Miller et al., 1998; Stephens and Chouet, 2001; Kilburn, 2003; Hammer and Neuberg, 2009).

White et al. (1998) first noted that low frequency earthquakes appear to occur in swarms of similar waveforms lasting from days to weeks, prior to and during unrest and extrusion periods at SHV. A swarm was originally described as a sequence of temporally close seismic events occurring within 15 km of a volcano (Benoit and McNutt, 1996). We take the narrower definition of Voight et al. (1999) who determine a swarm as when there are more than 10 events within an hour. Often, similarity between repeating events exists within these swarms, which suggests that the source location and source mechanism are identical (Geller and Mueller, 1980; Caplan-Auerbach and Petersen, 2005; Petersen, 2007). Events which are statistically similar to one another are known as multiplets, and can be grouped together into a family. Many authors have shown that it is possible to classify multiplets into a number of families of highly similar waveforms using a cross correlation technique, which therefore isolates and focusses on a single system at depth. Stephens and Chouet (2001) investigated a 23 hour swarm of low frequency seismic events prior to the eruption of Redoubt volcano, Alaska in December 1989, finding that the events could be sorted into 3 distinct families which evolved with time, of which the majority of events were correlated (cross correlation coefficient ≥ 0.68) with just one distinctive family. Later analysis of the 2009 Redoubt eruption by Buurman et al. (2013) also suggested the presence of multiplets, in particular prior to explosion events. Petersen (2007) suggested that a dominant family of multiplets exists within the low frequency seismic swarms at Shishaldin volcano, Alaska, although the dominant family is different within each swarm studied between 2002 and 2004. Thelen et al. (2011) suggest that the occurrence of multiplets at Mount St. Helens, Washington and Bezymianny volcano, Russia are related at least in part to the viscosity of the magma, and are therefore more prominent during dome building eruption events. Highly correlated high frequency events were observed at Mt. Unzen, Japan during significant endogenous growth of a lava dome between 1993 and 1994 and were classified into over 100 families (Umakoshi et al., 2003). Families of similar low frequency events have also been identified and studied at SHV in relation to tilt cycles by Voight et al. (1999) and Green and Neuberg (2006), who identified 9 multiplet families containing more than 45 similar events each over a time period of 6 days in June 1997, although not all of the families were active during each of the seismic swarms. In addition, Ottemöller (2008) found that 7100 hybrid events generated in the days prior to a large scale dome collapse at SHV in July 2003 all belonged to the same multiplet family.

In this paper we investigate accelerating event rates of precursory low frequency seismicity in the days prior to a number of large dome collapses at SHV on 25 June 1997; 12 July 2003; and 11 February 2010, and its use as a forecasting tool. These collapses were chosen since they represent a wide spectrum of collapses at SHV: the first; the largest; and the latest collapse. We use a cross correlation technique in order to further classify the seismicity based on waveform similarities as well as frequency content, as has been previously adopted to find similar seismic events at SHV. Green and Neuberg (2006) have already distinguished families of seismicity prior to the dome collapse in June 1997, however do not use these families in forecasting in any manner. Hammer and Neuberg (2009) used a single family from Green and Neuberg (2006) as a forecasting tool, however fail to detail the family used or whether any other family of seismicity produced a successful forecast. Here, we identify whether low frequency seismic families can be identified at another station at SHV compared to Green and Neuberg (2006), and produce forecasts using each of the seismic families identified, rather than picking just one. In addition, we identify similar seismic families and use these in hindsight analysis for forecasting dome collapses in July 2003 and February 2010 at SHV. This allows analysis of the wider application of the FFM at SHV, and whether it can be used in all circumstances as a forecasting tool. In Section 2 we describe the methodology of the cross correlation technique used to identify multiplets, with Section 3 showing the identified multiplet families for June 1997, July 2003 and February 2010 respectively. In Section 4 we apply the FFM to each of the families of events, showing that a more accurate forecast is generated when using a cross correlation technique to focus on one single system rather than simply using any low frequency seismicity. The implications of these results are discussed in Section 5.

2. The Cross Correlation Technique

2.1. Data selection

Since 1995 a continuous network of seismometers has been deployed on Montserrat to ensure the constant and consistent monitoring of the volcano by the Montserrat Volcano Observatory (MVO), originally installed by the USGS Volcano Disaster Assistance Program (Aspinall et al., 1998). At the time of the dome collapse event in June 1997, five three-component seismometers (Guralp CMG-40T with a 30 second corner frequency) and three vertical component Integra LA100/F 1Hz instruments were deployed, with data being digitized at 75 Hz. In March 2005, station MBLG was upgraded to a three-component broadband instrument (Guralp CMG-40T) with data digitized at 100 Hz. Station MBLG was chosen for analysis because of its close proximity to the dome which allowed a good signal-to-noise ratio and the availability of triggered and continuous data for the entire period under investigation.

MVO uses a STA/LTA (short-term average to long-term average) ratio triggering algorithm to identify individual events from the continuous incoming seismic record and place them into a catalogue of triggered events. It consists of two sliding windows, one investigating the short term amplitudes and is therefore very sensitive to incoming seismic signals, and one investigating the long term amplitude of the signal, which can provide information about the temporal amplitude of the noise at the site of the seismometer (Withers et al., 1998). Since the trigger is based on the ratio between these two windows, the algorithm is better able to record weak seismicity, compared to a simple amplitude only trigger mechanism (Trnkoczy, 2002). If a trigger (when a critical threshold of this ratio is exceeded) is found on three or more seismic stations simultaneously then an event is registered within the event count catalogue. To ensure the entirety of the earthquake signal is captured 2 seconds is added to the beginning of the trigger, and 10 seconds after the event once again drops below the triggering threshold. Using a Short-Term Averaging window length of 0.333 seconds, and a Long-Term Averaging window length of 60 seconds, with a trigger and detrigger ratio value of 4 and 2 respectively, a total of 1817 triggered events were placed within the catalogue from the 22 - 25 June 1997; 520 events placed in the catalogue from 8 - 13 July 2003; and 452 events identified from 8 - 12 February 2010. The instrument response and digitizer gains were removed from these seismograms to give velocity calibrated traces. STA/LTA analysis only identifies seismic events from the continuous seismic record, and does not classify them in any form. Traces were filtered with a band pass Butterworth two pole filter, with a low frequency cut-off of 0.5 Hz to reduce the influence of oceanic noise and a high frequency cut-off of 5 Hz, so as to concentrate entirely on the low frequency content of the waveform associated with magma movement.

2.2. Event Classification

Similarities between waveform shape can be quantified using the cross correlation function:

$$r_{xy}(i, i-l) = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_{i-l} - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_{i-l} - \bar{y})^2}} \quad (2)$$

where *r* is the cross correlation coefficient, *x* and *y* represent the two traces in the correlation, and therefore x_i is the *i*th sample of the signal *x* and y_{i-l} is the *i*-*l*th sample of the signal *y*. The overbar represents the mean value of the signal and *l* is the lag between the two signals. Identical waveforms will result in a cross correlation function of 1 or -1 dependent upon the polarity of the signal. r_{xy} is a measure of similarity in waveform shape only, since events are normalised prior to calculation. Consequently, the cross correlation function gives no information on the amplitude ratios of the events. Waveforms which are similar, and therefore from the same source location and generated by the same source mechanism, represent a single active system of seismicity.

Since the triggered events varied in duration and the cross correlation technique requires events of the same length, a cross correlation window of 10 seconds was chosen. This length was chosen since it allowed the entirety of the waveform to be present within the window, but ensured that only one event was captured per window. Previous attempts at cross correlating events at SHV by Green and Neuberg (2006) used an 8 second cross correlation window, however upon inspection of the events this was not sufficient to include the majority of each coda at this station. In order to determine if any similar events were present on any particular day, each 10 second event was cross correlated with every other triggered event from the same day. The maximum cross correlation coefficient of each waveform with every other waveform was determined and placed into a cross correlation matrix (an example of which can be seen in Figure 1 for the 24 June 1997). In theory, identical events will give a cross correlation coefficient (r) of 1 (e.g. the autocorrelation of events, as seen along the diagonal of Figure 1), and events which have no correlation will have a cross correlation coefficient of 0 (r=0). Events were then classified as being significantly similar to one another if the maximum cross correlation coefficient was above 0.7, and are shown on a colour spectrum in Figure 1. Green and Neuberg (2006) and Thelen et al. (2011) show a clear justification of using 0.7 as a correlation threshold since it is significantly above the upper limit for random correlation between waveforms and noise. Therefore it is assumed that some events with a correlation coefficient of less than 0.7 are a consequence of random noise being correlated. Visual inspection of the stacked waveforms also confirmed that 0.7 was an appropriate choice and captured the majority of similar events with limited scatter of the waveforms once aligned. Figure 1 suggests distinct time periods when similar seismic events were active (coloured areas are separated by distinctly white areas), and that a highly correlated swarm of events occurred on this day (brighter and more concentrated colours).

Two different techniques, both utilizing the entire catalogue of triggered events

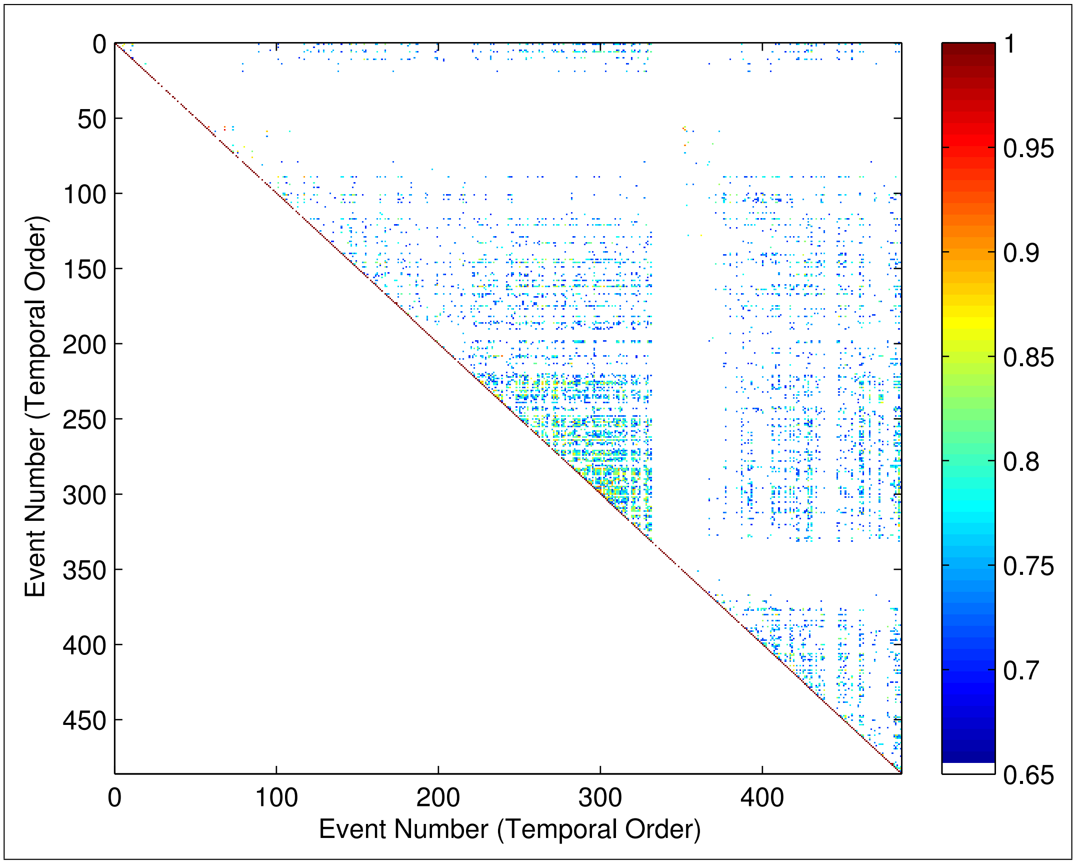


Figure 1: An example of a maximum cross correlation similarity matrix from station MBLG on 24th June 1997. A total of 486 triggered events were found within the 24 hour period and are represented from 1 to 486 along the x and y axis. Each row of the matrix therefore represents one triggered event compared to every other triggered event on that day. Only events with a cross correlation coefficient above 0.7 are shown on the colour spectrum and are deemed to be similar. The autocorrelation of each triggered event with itself (cross correlation coefficient equal to 1) is represented on the diagonal.

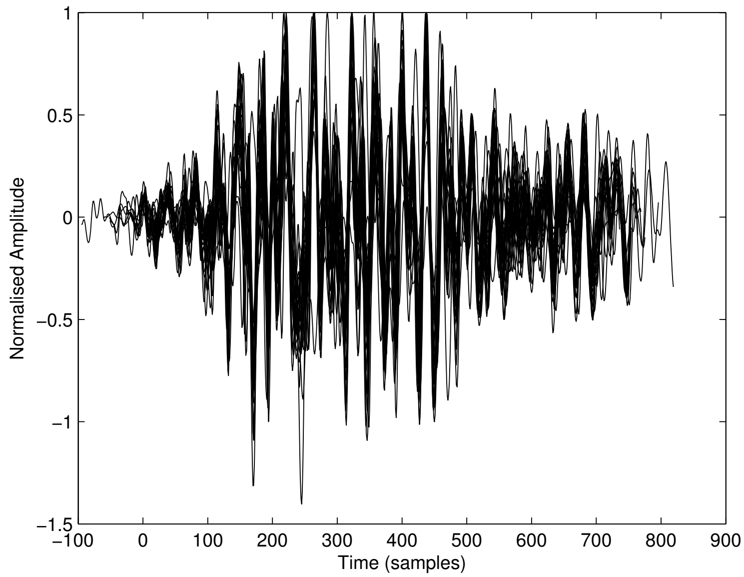
during periods of interest, were used to isolate multiplets and collate them into families. Following Petersen (2007), a dominant event for each day was identified as the event correlated with the highest number of other events from that day. The mean correlation value of each event with every other event was determined from the cross correlation matrix (Figure 1) and the event with the highest mean was taken to be the dominant event. The second technique followed Green and Neuberg (2006) where each triggered event in turn was correlated with every other event. Events with a cross correlation coefficient above 0.7 were subsequently grouped together, labelled as a multiplet family and removed from the time series. This procedure was repeated across the entire investigated time period until all events had been classified into a number of different families. This has the advantage of finding all families of multiplets which may be present in the continuous data, rather than simply the dominant one, as well as finding families which may be infrequent in their repetition but still important. This procedure also allows for the identification of evolving waveforms, either by migration of their source location or change in the source process. Families which contained fewer than 10 similar individual triggered waveforms were eliminated from further analysis. To avoid selection bias, the events within a single family (i.e. all had a minimum cross correlation coefficient of 0.7 with one another) were stacked, and the average waveform taken (Figure 2). This average waveform is hereafter referred to as the Master Event of each family, and is used as a statistical representation of this family in terms of waveform shape.

The master events were then cross correlated with the continuous seismic record at MBLG using a sliding window technique and multiplets identified when the cross correlation coefficient was greater than 0.7 between the master event and the continuous seismogram. The sliding window separation of 0.01 seconds allows the maximum number of multiplets to be identified, in particular those which are too small or are overlapping in the continuous seismic record to be identified by the triggered acquisition system at MVO (STA/LTA algorithm). The similar events were then grouped into a multiplet family.

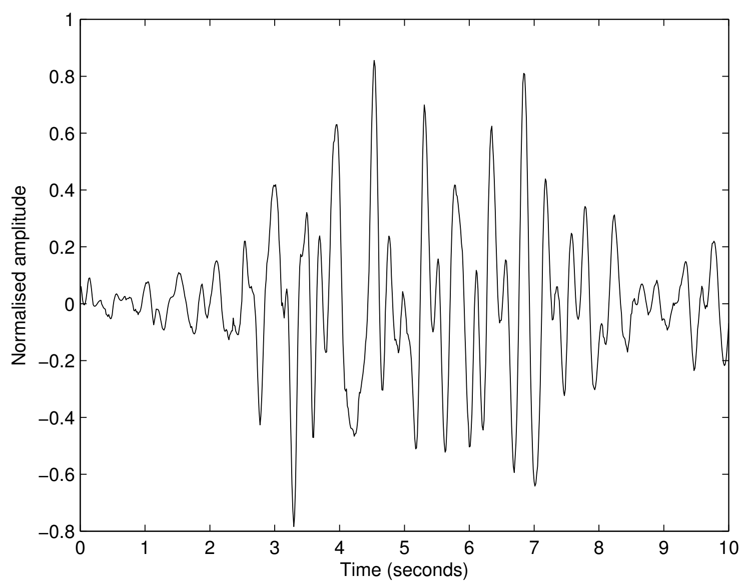
3. Similarity of Events

3.1. 22 - 25 June 1997

The total number of multiplets identified using the cross correlation procedure was 7653 from 22 to 25 June 1997, in comparison to only 1435 events identified using the triggered algorithm at MVO over the same time period. This methodology therefore represents a five fold increase in the number of events which can be identified and used in further analysis. The dominant multiplet family identified using the technique by Petersen (2007) contained a total of 878 multiplets (Table 1). 10 multiplet families containing over 250 multiplets each were identified using the technique of Green and Neuberg (2006), although the dominant master and Master event 001 have a cross correlation coefficient of 0.93, suggesting that they belong to the same family. This emulates the conclusions of Green and Neuberg (2006), who also identified 10 waveform families during the same time period at



(a) Stack of events highly correlated with identified dominant master event on 24 June at 11:18-53s from station MBLG. A total of 26 triggered events are included in the stack. Each event has been aligned at the peak correlation coefficient.



(b) Master Event : average waveform from stack

Figure 2: Stack of highly correlated waveforms and the resulting average dominant master event from station MBLG identified on 24 June 1997

Family	22nd	23rd	24th	25th	Total
Dominant	0	133	376	369	878
Master001	0	121	336	222	679
Master010	0	256	315	32	603
Master014	2	136	302	137	577
Master100	4	71	280	193	548
Master106	0	42	169	45	256
Master121	0	100	514	400	1014
Master136	3	131	483	542	1159
Master141	0	47	276	173	496
Master210	0	20	170	349	539
Master291	0	39	390	475	904

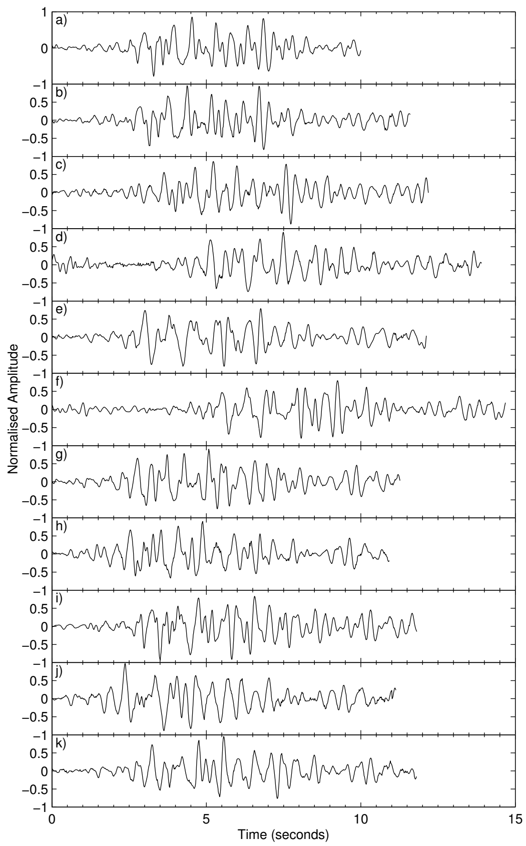
Table 1: Number of events within each family sorted into days from the 22 - 25 June 1997.

station MBGA, however they did not use these families as a forecasting tool. As these families of events were identified at a different station it is not possible to compare directly the results of the two studies.

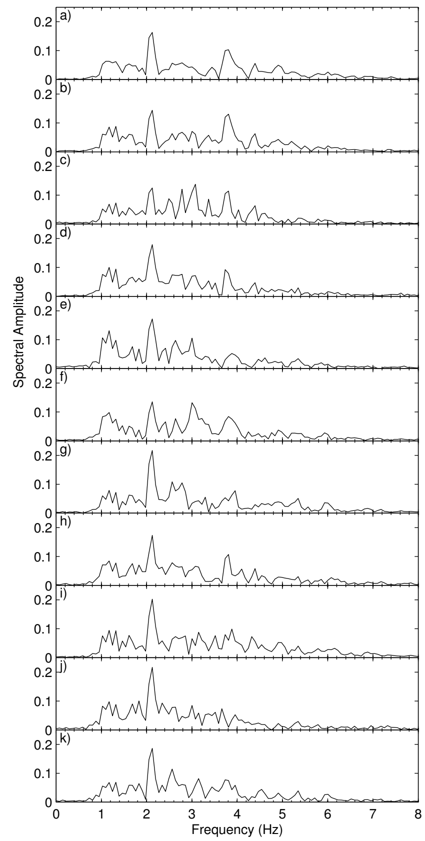
Clear differences between each of the families can be found in terms of the onset timing of events and the waveform characteristics (Figure 3). In particular significant differences are noted in the expression of the waveform coda. As expected for low frequency events, master event 210 (Figure 3a(i)) clearly decays in a harmonic manner, however this is not the case for every master event (e.g. master event 100, Figure 3a(e)). However, very little difference is seen in the amplitude spectra (Figure 3b). All of the low frequency master events have a dominant spectral peak at 2.1Hz, suggesting a fundamental similarity in their resonance behaviour. A secondary spectral peak can be seen at 3.8 Hz for some master events. Variations in the amount of energy distributed from 1-5 Hz varies amongst master event waveforms, although not significantly. No significant phase shifts were detected from the cross correlation analysis.

Some differences are evident in the duration and timings of swarms of each master event in relation to the dome collapse on the 25 June (Figure 4). In particular, only three master events appear within the swarms identified on the 22 June, and these are very short lived. The beginning and ending of each swarm varies only slightly throughout the rest of the sequence from the 23 - 25 June. Since MBLG was destroyed by volcanic activity associated with the dome collapse at 16:55 UTC (Luckett, 2005) no swarms are able to be identified past this (vertical line in Figure 4). Each of the multiplet families are persistent across each of the six swarms which occur from the 23 June onwards, suggesting that sources at the same locations are being reactivated by the same process during this time, as was also concluded by Green and Neuberg (2006). Besides master events 010, 014 and 106, all other master events appear within swarms which are active right up to the dome collapse (Figure 4).

Unlike Stephens and Chouet (2001), Umakoshi et al. (2003) and Petersen (2007) the waveforms within the multiplet families observed at SHV in 1997 do not appear to significantly evolve with time, since clustering of events is not seen along the diagonal in Figure 1 (Caplan-Auerbach and Petersen, 2005) and each swarm appears to contain similar waveforms with limited evolution in the cross



(a) Dominant master and 10 other identified master event waveform signatures identified in June 1997 at station MBLG



(b) Amplitude spectra of the waveforms of the dominant master and the 10 other identified master events in June 1997 at station MBLG

Figure 3: Comparison of all master events identified from 22 to 25 June 1997 at station MBLG, by both techniques. The waveform and amplitude spectra labelled (a) is the dominant waveform identified from the cross correlation coefficient matrix (Figure 1). Waveforms and corresponding amplitude spectrum labelled (b) to (k) represent the master events 001, 010, 014, 100, 106, 121, 136, 141, 210 and 291 respectively.

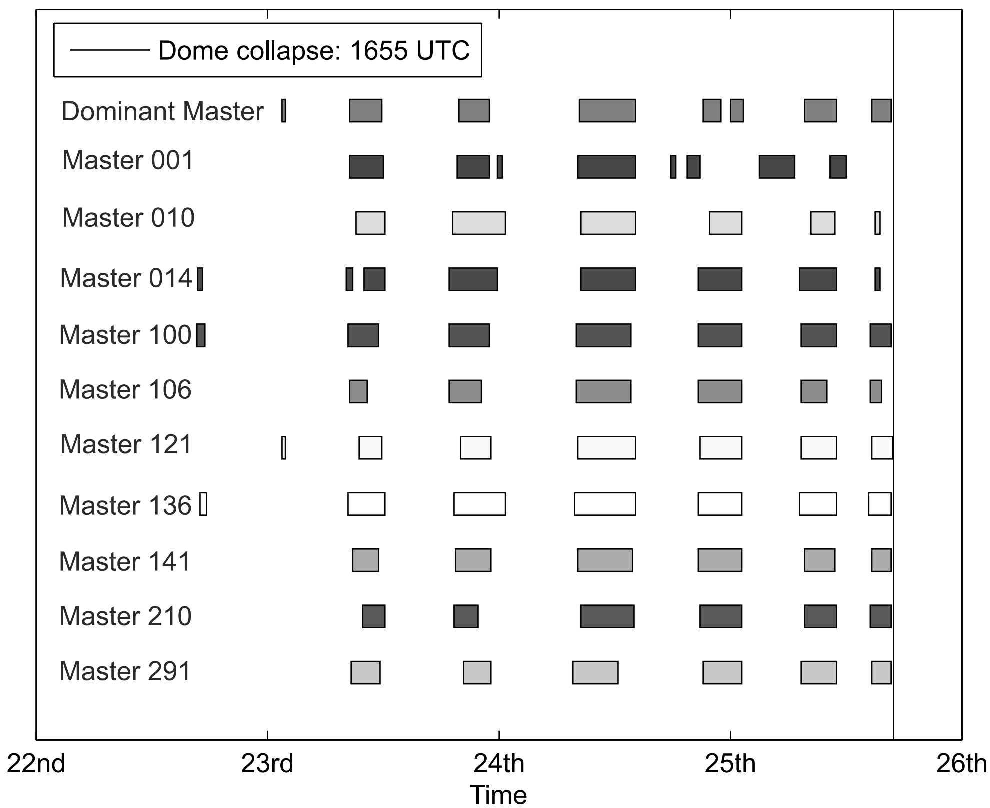


Figure 4: Comparison of the timing and duration of swarms related to each of the master events identified. The timing of the dome collapse is represented by the vertical line on the 25th June. The y axis is only an indication of each of the families present separated in space for the purpose of clarity on the plot and does not represent time or dominance, each master event is simply drawn below the last so that all can be compared. Each coloured rectangular box represents the times when the master event was active during the 22 - 25 June analysis period.

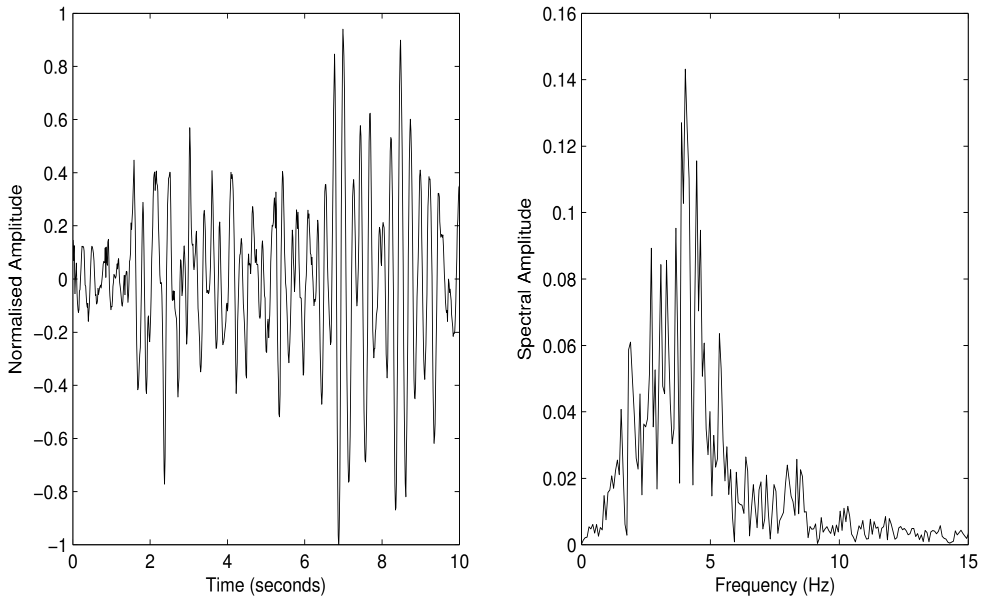
correlation coefficient with time. This suggests the waveforms are stable and persistent and therefore the trigger location and source process must also be. The cross correlation coefficient can vary by up to 0.25 within each swarm, however the difference between the maximum and minimum mean cross correlation coefficient for each swarm as determined by the 11 master events only varies by ≤ 0.06, suggesting limited evolution in the waveforms.

3.2. 8 - 12 July 2003

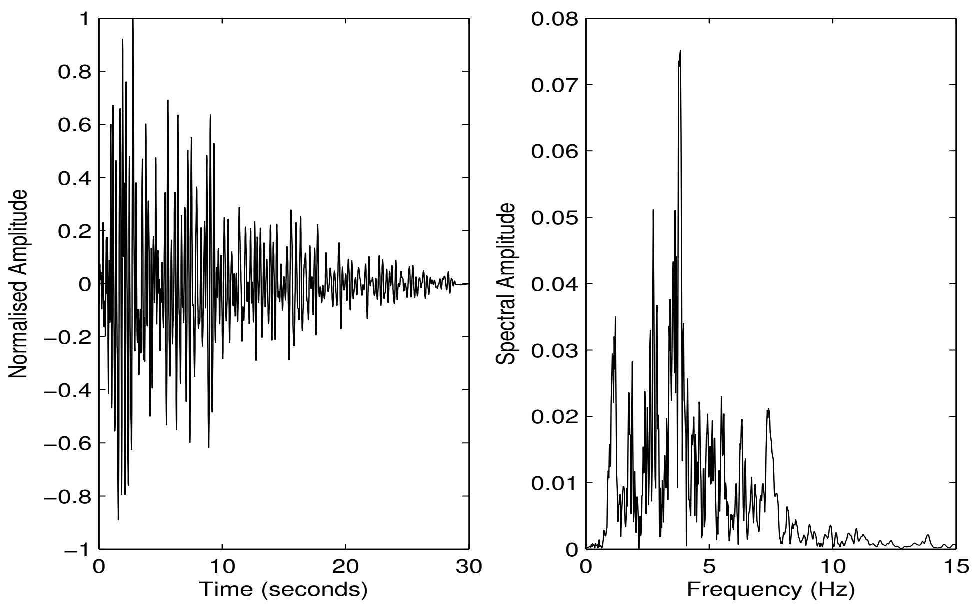
A total of 520 events were identified from the continuous seismic record at station MBLG from 8 to 12 July 2003 using the MVO STA/LTA algorithm. In comparison, the total number of multiplets identified for the same time period was 2241 events, representing a four fold increase in the number of seismic events identified. However, unlike the multiplets identified in June 1997, only one family of events could be identified. A total of 79 events from this single family were stacked to create an average master event, shown in Figure 5(a), all of which had a cross correlation coefficient of above 0.7 to maintain a high signal to noise ratio and consequently pick the most similar events for use in the forecast. Ottemöller (2008) also identified one single dominant family of events during a similar time period from 00:00 on 9 July to 12:00 on 12 July, however suggests that a total of 7100 events could be identified. The large difference in the number of identified events from the cross correlation technique is put down to the fact that Ottemöller (2008) used a much lower cross correlation coefficient to identify events ranging from 0.6 to 0.66, whereas the event identified in this study consistently used a cross correlation coefficient threshold of 0.7.

The dominant waveform has an emergent onset and it is difficult to pick out significant seismic phases (Figure 5(a)). Unlike the low frequency seismicity identified in June 1997, the coda of the waveform does not decay in a smooth manner. It is evident that the peak energy is centred at approximately 4 Hz, in comparison to the June 1997 events which all had a dominant frequency of approximately 2.1 Hz, however it is distributed across 0 Hz to 5 Hz band. Contrary to Petersen (2007) who suggested that some multiplets can be active over a number of years, there are no similarities between the events identified in June 1997 and the dominant master event identified in July 2003, pointing to an evolving system over this time period.

Figure 6 suggests a clear evolution of the cross correlation coefficient with time. One possible explanation is that this may represent a slightly migrating source location at depth. The events with a relatively lower cross correlation coefficient on the 9 and 10 July occurred further away from the dominant master event location, than those on 11 July. Perhaps more significantly, it should be noted that the similar seismic waveforms appear to stop in the hours prior to the dome collapse (vertical line in Figure 6), although a large amount of data is missing from this time period so it is not possible to tell the exact timing of the change from very similar events to non-similar events. This is significant because it may represent a time delay function between events occurring at depth and those at the



(a) *Left*: Dominant master waveform identified by stacking similar events (cross correlation coefficient greater than 0.7) from 8 to 12 July 2003 at station MBLG. A total of 79 events were used in the stack to create the average Master event waveform. *Right*: Single sided amplitude spectrum of the dominant master waveform.



(b) *Left*: Dominant master waveform identified by stacking similar events (cross correlation coefficient greater than 0.7) on 11 February 2010. A total of 3 events were used in the stack to create the average Master event waveform. *Right*: Single sided amplitude spectrum of the dominant master waveform.

Figure 5: Dominant Master waveforms identified in precursory seismicity in July 2003 and February 2010; and their associated frequency spectrum.

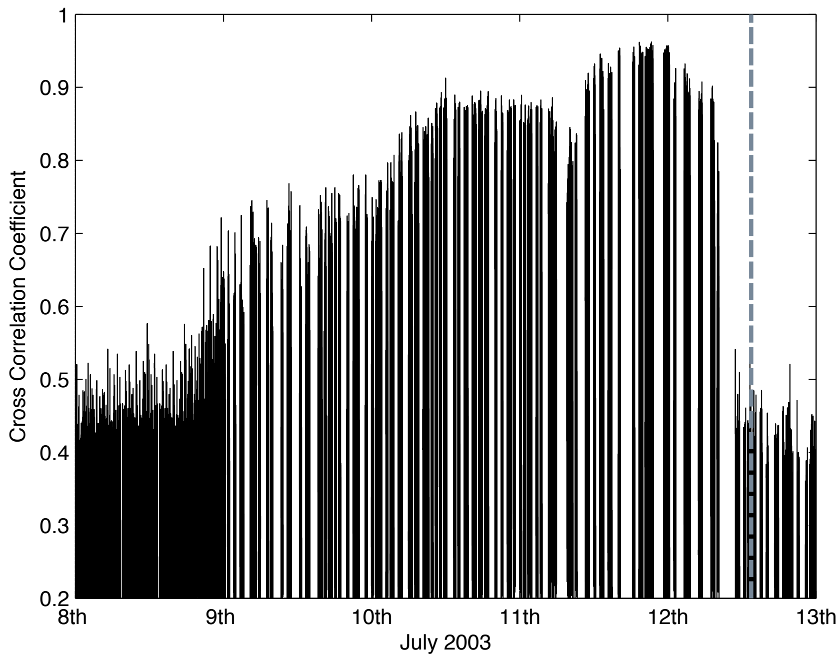


Figure 6: The evolution of the cross correlation coefficient with time: July 2003 with the Dominant Master Event, station MBLG. The dome collapse occurred at the time of the vertical line (13:30 on 12 July 2003). The gaps in the data represent gaps in the seismometer recordings rather than a dip in the cross correlation coefficient.

surface as first envisaged by Voight (1988).

3.3. 8 - 11 February 2010

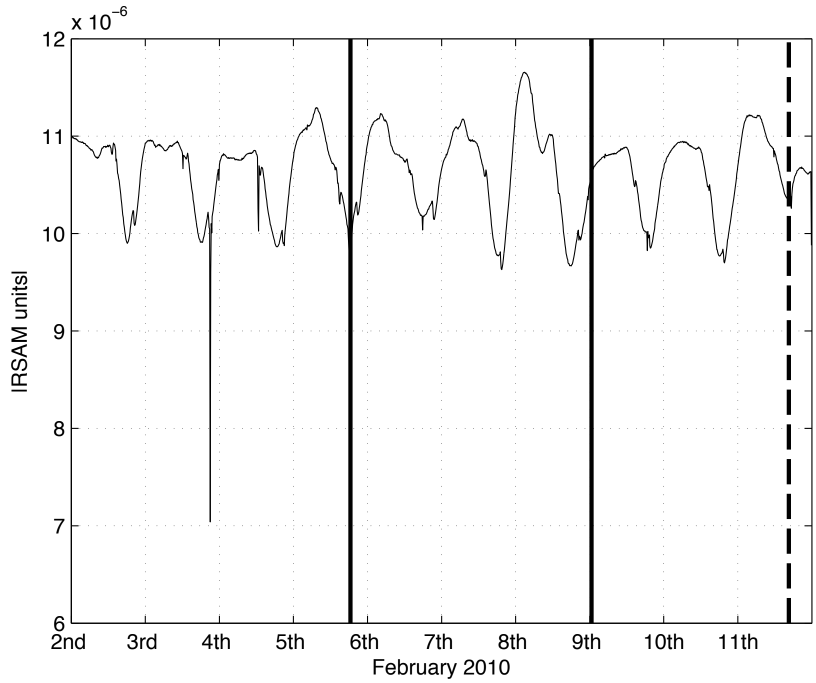


Figure 7: RSAM (10 minute averages) at Soufrière Hills volcano, 2 to 11 February 2010, station MBLG. RSAM is representative of average ground velocity in meters per second. The first two vertical lines (solid) represent two small Vulcanian explosions, on the 5 and 8 February. The final vertical line (dotted) represents the onset of the dome collapse and associated pyroclastic flows on 11 February.

Figure 7 suggests a clear cyclicity in seismicity in February 2010, which is unaffected by two small Vulcanian explosions in the days prior to the dome collapse. However, Stinton et al. (2014) reported very little precursory seismic activity before the dome collapse on 11 February 2010. Indeed, in order to identify individual seismic events from the continuous record, it was necessary to change the input STA/LTA parameters. In particular, the short term averaging window length changed from 0.333 seconds (in 1997 and 2003) to 2 seconds (the maximum period likely for a low frequency signal between the frequency of 0.5 and 5 Hz), allowing the concentration on temporally longer events. The long term averaging window was modified from 60 to 120 seconds, again to accommodate longer events, meaning that very short events were not identified as an event. The trigger ratio value, the value at which the algorithm begins to detect an event, was moved from 4 to 8 since large amounts of noise appeared to dominate most of the signal. A total of 452 events were identified using this method between the 8 and 12 February 2010.

However, of these 452 events identified in STA/LTA analysis, only 10 events were identified as true low frequency events, detected by filtering and manual inspection. Of these 10 events, 3 were very similar to one another with a cross correlation coefficient of over 0.7. These events were subsequently stacked and a Master waveform produced (Figure 5(b)). The master waveform is much longer than the previous master events identified in 1997 and 2003, lasting $\approx$30 seconds. The distribution of energy is also very different, with energy existing over the ranges of 0 to 10 Hz, suggesting a hybrid nature, with a higher proportion of energy concentrated above 2 Hz. However, on Montserrat, hybrid and low frequency earthquakes appear to occur on a continuum, with these two types bestowing the idealised end members (Neuberg et al., 2000). Therefore, although characteristically different to the master events identified prior to the 1997 and 2003 dome collapses, it is still assumed that the master event for 2010 is a low frequency type earthquake related to the movement of magma at depth (Chouet, 1988; Neuberg et al., 2000).

4. Forecasting Volcanic Dome Collapses at SHV

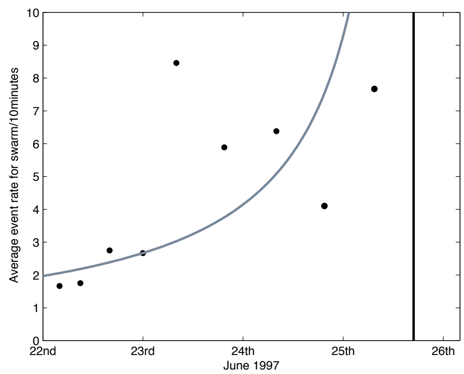
The FFM was first developed as a tool for identifying accelerating material creep and relating this to a slope failure; a cause and consequence of one single active system generating failure (Fukuzono, 1985). However, a volcanic system is inherently more complex, and accelerating magma ascent could be detected at several positions in the magma plumbing system with different phase delays and amplitudes. Therefore, in order for only one system to be analysed as input in the FFM it is necessary to focus only on one "family" of low frequency waveforms which originate from the same source mechanism and location (Geller and Mueller, 1980; Neuberg et al., 2000; Thelen et al., 2011). Classification by waveform similarity in addition to frequency content allows low frequency seismicity of a single source and depth to be exclusively analysed, meaning it is less likely that multiple sources become mixed in the forecast.

Hammer and Neuberg (2009) suggested that although individual swarm analysis was not suitable for the FFM since a deceleration phase is almost always evident, consecutive swarm analysis might be at SHV. Instead of taking the event rate every 10 minutes from the continuous data, the event rate per 10 minutes is averaged across the entire duration of the seismic swarm, suggesting an overall acceleration in the event rate from swarm to swarm. This paper therefore also focuses on multiple swarm analysis for forecasting the timing of dome collapses at SHV.

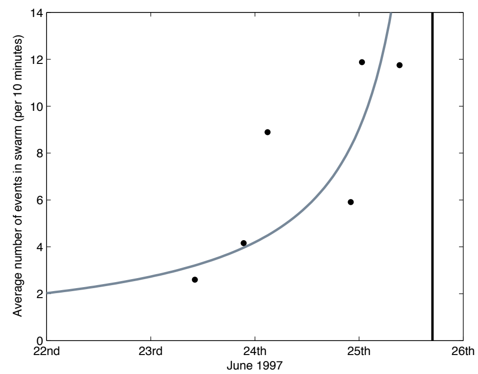
4.1. 22 - 25 June 1997

Using identified events from the STA/LTA algorithm, as used at MVO, Figure 8a shows an initial acceleration in the average number of events per 10 minutes across swarms identified in June 1997, although the trace of the least squares fitted curve suggests a slowing of the acceleration up to the point of dome collapse (vertical line). Figure 8b shows the acceleration of swarms which have been identified using the dominant master event. Further classification of the low frequency seismic events into families appears to tighten the least squares fit and lead to a more convincing accelerating pattern of average number of events within 10 minutes of each swarm. This is further verified in Figure 8c, 8d and 8e which all show an acceleration in the average event rate with time up until the dome collapse, using master events 121, 136 and 141 respectively.

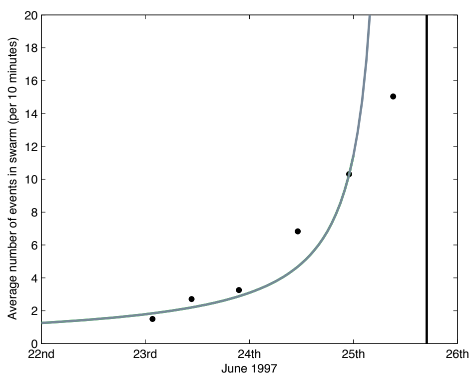
Figure 9 represents the application of the FFM to each of the accelerating event rates identified in Figure 8. α = 2 allowed graphical extrapolation to the forecasted timing of collapse as a simple linear regression. Inverse event rate trends were identified if at least three consecutive swarms formed an inverse trend. This was to try and eliminate spurious trends since a single decrease in the swarm event rate may be due to external factors, for example such as an increase in noise which obscures the number of events determined. Table 2 shows the results from



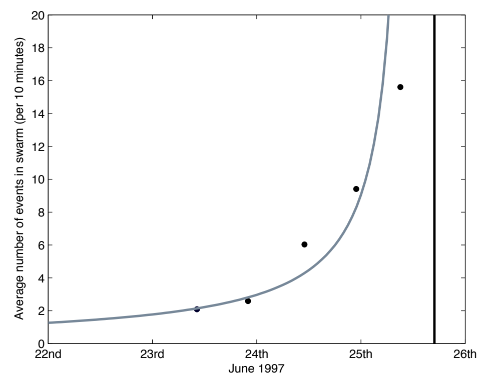
(a) All triggered low frequency seismicity



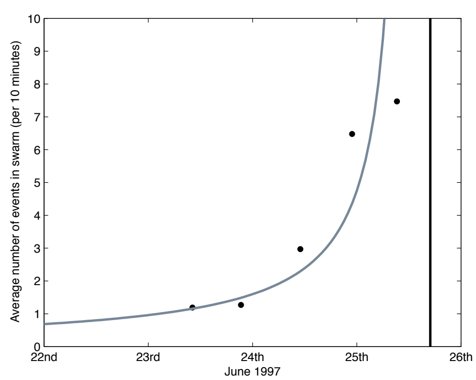
(b) Dominant master event



(c) Master event 121

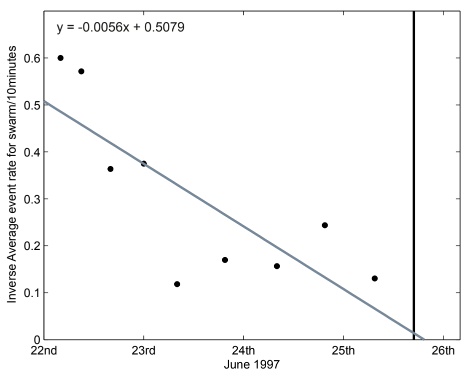


(d) Master event 136

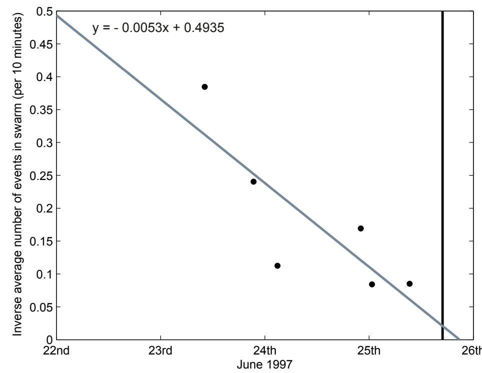


(e) Master event 141

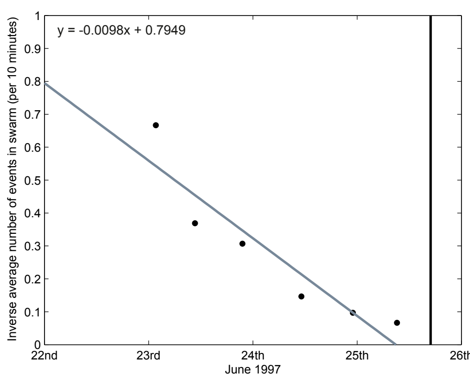
Figure 8: The average event rate per 10 minutes within swarms from 22 - 25 June 1997 at station MBLG. Each data point represents the average event rate for each individual swarm. The vertical line represents the known timing of dome collapse on the 25 June 1997 at 16:55 UTC. The acceleration of these events is depicted with the curve.



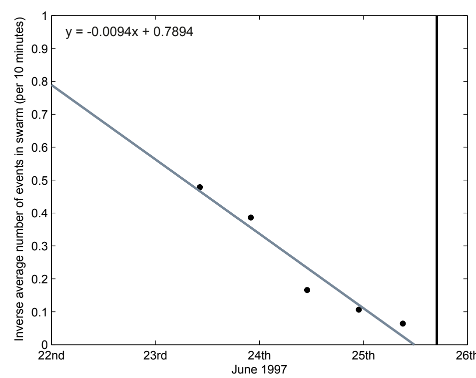
(a) All triggered low frequency seismicity



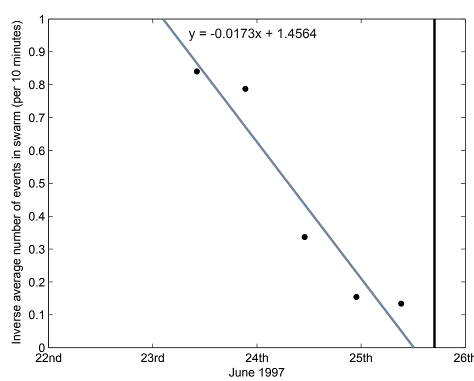
(b) Dominant master event



(c) Master event 121



(d) Master event 136



(e) Master event 141

Figure 9: Application of the FFM: the inverse average event rate per 10 minutes within swarms from 22 - 25 June 1997 at station MBLG. Each data point represents the inverse average event rate for each individual swarm. The vertical line represents the known timing of dome collapse on the 25 June 1997 at 16:55 UTC. The graphical representation of the FFM is depicted by the linear regression (it is assumed that $\alpha = 2$) and the forecasted timing of failure can be read off the x-axis at the point where the linear regression crosses it.

Event	Known Timing (HH:MM)	Forecasted Timing (HH:MM)	Difference (HH:MM)	Forecasted early/late	R^2
Triggered Low frequency	88:55	91:24	02:29	Late	0.63
Dominant Master	88:55	92:50	03:55	Late	0.69
Master Event 001	88:55	95:12	06:17	Late	0.73
Master Event 010	88:55	n/a	n/a	n/a	n/a
Master Event 014	88:55	77:56	10:59	Early	0.60
Master Event 100	88:55	82:29	06:26	Early	0.86
Master Event 106	88:55	85:58	02:57	Early	0.59
Master Event 121	88:55	80:54	08:01	Early	0.87
Master Event 136	88:55	83:43	05:12	Early	0.94
Master Event 141	88:55	84:10	04:45	Early	0.92
Master Event 210	88:55	82:48	06:07	Early	0.84
Master Event 291	88:55	75:55	13:00	Early	0.83

Table 2: Timings of forecasted failure. Timings (known, forecasted and difference) are depicted in hours and minutes. The R^2 value, as defined in the text, ranges from 0 to 1 and is the a scale of how well the model (FFM) can explain the data (event rate). Besides master events 014 and 106, each of the forecasts which were forecasted early using the FFM and a separate master event have a higher R^2 value than those who were forecasted late, or when using all low frequency seismicity. Only master event 010 was unable to be used in analysis using the FFM since no acceleration in the event rate per swarm was identified. The R^2 value for all triggered low frequency seismicity suggests that the FFM is inappropriate to describe the inverse event rates seen, since the linear regression does not fit well to the data.

the same analysis with all identified master events. When using triggered event data, although the timing of the forecasted dome collapse is within two hours of the known failure time, the fit of the linear regression (the FFM) to the data is poor. The application of the FFM to a single family of events in June 1997 has already been published in Hammer and Neuberg (2009) however, they fail to identify the family they used in analysis or the accuracy of the forecast in terms of fit to that data in a quantitative manner. Here, we show that any of the families of similar seismicity in June 1997 can be used as a forecasting tool with the FFM, with the exception of Master Event 010, which did not show an accelerating trend in the number of seismic events.

The forecasted timing of the dome collapse was never greater than 13 hours away from the known timing of collapse when using the cross correlation technique first to identify similar events, and in most cases the collapse was forecasted early (Table 2). Despite increasing the difference between the known and forecasted failure times, further classification of multiplets into families consistently allows for a better fit of the linear regression to the data, as can be seen from the high R^2 values.

After Barrett (1974), R^2 is defined as:

$$R^2 = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad (3)$$

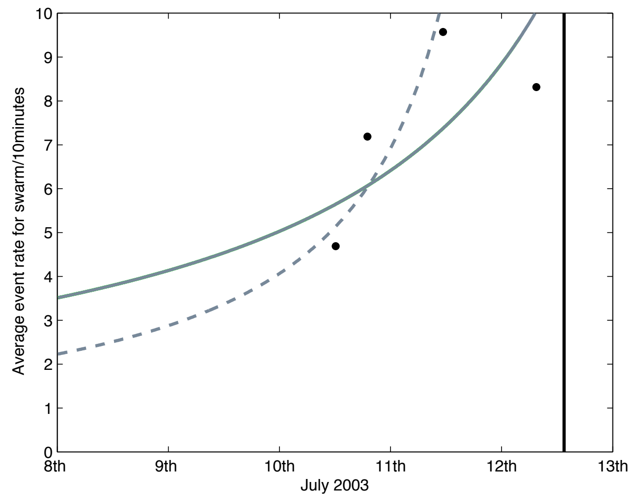
where y_i represents the observed parameter at position i (i.e. the inverse event rate at a given time), $\hat{y}_i$ represents the predicted parameter of y at i (i.e. the FFM linear regression at this time), and $\bar{y}$ represents the mean value of all of the y values. R^2 or the coefficient of determination is the proportion of variability which can be explained by the model, and ranges from a minimum value of 0 which suggests that the model does not explain any part of the data, up to a maximum of 1 which suggests the model perfectly describes the data. In our case in particular, the R^2 value shows how well future outcomes can be predicted by the model (the FFM), and therefore the closer the value is to one, the more confidence we can have in the forecast. R^2 values of less than 0.65 are considered to represent a poor relationship between the observed data and the fitted FFM model.

Since seven out of eleven master events identified had R^2 values of greater than 0.7, it can be assumed that the FFM is appropriate for this data set. The R^2 value for the dominant master is slightly lower at 0.69, however this still demonstrates a good fit between the model and the observations. Significantly, the R^2 value for using the FFM with all triggered low frequency result gives a value of 0.63, which is deemed to not be a good fit. This is confirmed by the wide discrepancy between the observed event rates and the theoretical application of the FFM (Figure 9a).

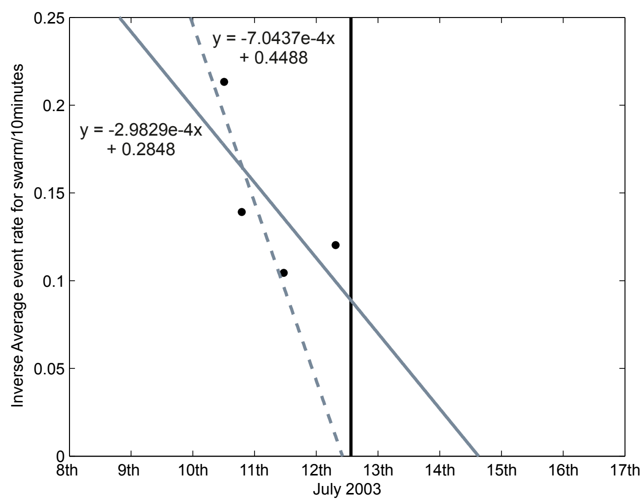
4.2. 8 - 12 July 2003

Seismicity prior to the dome collapse in July 2003 did not take the form of well defined swarms, as observed in June 1997. Instead, seismicity was pulsatory which was further highlighted by data gaps for MBLG. In addition, Figure 6 suggests that the cross correlation coefficient is constantly evolving, with highly similar events occurring throughout the precursory sequence and no evidence of an acceleration in the event rate. However "swarms" of events were identified in the triggered incoming seismicity as recorded by MVO in near real time (STA/LTA parameters unknown) and therefore the timings of these swarms were used with the FFM as a forecasting tool.

A clear acceleration can be seen in the swarms which occurred from the 10 to 12 July 2003 (Figure 10(a)), followed by a slight deceleration in the event rate for the final swarm before the dome collapse (vertical line). Application of the FFM to all swarms identified (i.e. if the forecast was made on 12 July 2003 after the last swarm had ended) then the forecasted timing of collapse would be on 14 July 2003 at approximately 15:00 h. However, the confidence in the forecast would be low, with an R^2 value of 0.51. If only the first three swarms are used (i.e. only those swarms which exhibit an acceleration) then a forecast is made for 12 July at approximately 10:12 (dotted line in Figure 10(b)), just over 3 hours before the known timing of failure at 13:30 on 12 July. A greater amount of confidence can also be placed on this forecast, since the R^2 value is 0.82, suggesting a significant relationship between the observed data points (event rate in each swarm) and the model (FFM). This is a significant improvement upon a forecast using all low frequency seismicity identified during this period, which forecasted a failure time at 17:50 h on 11 July, with an R^2 value of 0.54.



(a) The average event rate per 10 minutes within swarms from 10 to 13 July 2003, station MBLG. The acceleration of all swarms is depicted with the solid curve and only the accelerating swarms by the dotted curve.



(b) Application of the FFM: the inverse average event rate per 10 minutes within swarms from 10 to 13 July 2003, station MBLG. The graphical representation of the FFM is depicted by the linear regression (it is assumed that $\alpha = 2$) and the forecasted timing of failure can be read off the x-axis at the point where the linear regression crosses it. The solid regression includes data from all swarms; the dotted regression is only the swarms which were accelerating.

Figure 10: Acceleration of swarms of seismicity and application of the FFM: 10 to 13 July 2003. Each data point therefore represents the inverse average event rate for each individual swarm. The vertical line represents the known timing of dome collapse on the 12 July 2003 at 13:30.

4.3. 8 - 11 February 2010

Despite changing the STA/LTA parameters in order to identify seismic events during the precursory period of 8 to 11 February 2010, no additional seismicity from the continuous seismic data was identified using the cross correlation technique. This not only suggests that there was very little precursory seismicity to the dome collapse, in terms of event counts or identified from accelerations in RSAM which simply appeared cyclic up to the collapse (Figure 7), but also suggests that at this time events which were detected were not similar. This is in stark contrast to both the dome collapses of 1997 and 2003 which were dominated by similar seismic events, and which showed an acceleration in seismicity prior to the collapse.

The fact that this collapse was not preceded by low frequency seismicity suggests that the collapse originated in processes unrelated to the movement of magmatic fluid at depth in the days before the collapse or that this movement was aseismic in nature. Stinton et al. (2014) suggest that the collapse occurred due to over-steepening of the dome and talus which led to a gravitational collapse of the material. In the 4 months prior to the collapse, intensive extrusive and explosive activity had been observed, and since the collapse occurred in a piecemeal fashion over a number of hours, gravitational instability of a large dome is thought to have been a primary driving factor in collapse.

5. Discussion

5.1. Characteristics of Seismicity Observed

The seismicity prior to the dome collapse of June 1997 exhibited many features which have been commonly observed prior to eruptive events at SHV, and many other volcanoes worldwide. Not only is an increase in the number of seismic events observed in the days prior to collapse, but seismicity occurs in well defined swarms. The identification of multiplets within these swarms in June 1997 and July 2003 echo the conclusions of Green and Neuberg (2006) and Petersen (2007) that a stable source process and location must be present in generating these events, and therefore the seismic energy must travel along similar ray paths to the seismometer. Further evidence for this comes from the fact that there is little clustering of similar events close to the diagonal of the cross correlation matrix (e.g. Figure 1). Slight clustering on the 24 June of events 250 to 350 is perhaps evident but not deemed significant. This suggests that very highly correlated events can be found over sustained periods of time (hours) and indicates very little change in the source conditions over this period (Caplan-Auerbach and Petersen, 2005). Since the multiplets are repeated in swarms over a number of days, the source mechanism must be non-destructive and the trigger mechanism must be able to recharge quickly, since successive similar events occur in the continuous seismic record within seconds of one another. The identification of eleven families of multiplets each with their own waveform characteristics in June 1997 (Figure 3) suggests that a number of source mechanisms and/or source locations must been active at SHV. This reflects the diversity of sources and physical processes which act simultaneously at SHV. Unlike Green and Neuberg (2006), we did not find that any one multiplet family was more active than any others, suggesting that each of the source mechanisms were sustained for the duration of the precursory sequence. The identification of only one single active family in July 2003 suggests a change in source dynamics such that only one source was active during this time.

Using the quarter wavelength hypothesis of Geller and Mueller (1980), the repeating multiplets in June 1997 must occur within a maximum source distance of ≈ 300m, assuming a dominant frequency of 2.1 Hz and an average P wave velocity of 2500 ms^{-1} in the dome region as described by the current MVO velocity model for SHV. Paulatto et al. (2010) however have suggested that the upper 2.5 km of the dome at SHV could have a P wave velocity as low as 1510 ms^{-1}, therefore transforming the source volume to ≈ 178 m. Furthermore, since this hypothesis is only really applicable for general seismic body waves, Neuberg et al. (2006) have suggested that for low frequency events to be deemed similar, the source location within a heterogeneous volcanic environment may vary by as little as one tenth of the wavelength. For our results this suggests that the each of the low frequency families in June 1997 could be located within a source volume of ≈ 120 m (using the MVO velocity model) or ≈ 72 m (using the model of Paulatto et al. (2010)). This is in good agreement with De Angelis and Henton (2011) who located multiplet events from the same time period (22-25 June 1997) and suggest they consistently occur within a narrow depth range of 100 m to 300 m below sea level (approximately 1200 m to 1400 m beneath the volcano summit), within an equally narrow longitude and latitude. Rowe et al. (2004) also relocated almost 4000 similar seismic events from July 1995 to February 1996 and found that the source volume was approximately $1km^3$, with source dimensions of 10m to 100m in diameter. Using the dominant frequency of 4.03 Hz for the similar seismicity in July 2003, the quarter wavelength hypothesis states that the events must have occurred within a maximum distance of 155 m of one another. This is is good agreement with Ottemöller (2008) who found similar hybrid seismic events from swarms over the same time period were mostly located between 1500 and 1700 m below the dome summit, within a radius of ≤ 150m.

Comparison of precursory seismicity before three large dome collapses at SHV suggests that precursory conditions are not stable or constant. This is best illustrated by the fact that 11 families of similar seismicity were identified in June 1997, only one family of similar seismicity was identified in July 2003 and no families of similar seismicity were identified in February 2010. Stinton et al. (2014) has suggested that the lack of seismicity in February 2010 may be due to changes in the underlying processes of collapse, and that in 2010 gravity and over-steepening of the dome played a much larger part in the collapse event than the internal dynamics of moving magma. Forecasting dome collapse events at SHV cannot therefore solely rely upon seismicity, but must include analysis of dome size and shape in an equal capacity.

5.2. The Similarity of Master Events in June 1997

The similarity in frequency content between each of the master events identified in June 1997 (Figure 3(b)) suggests that the master events themselves could be similar to one another. If only a shift in phase separates them from one another it could be indicative of a migrating source location at depth or a slight change in the active source process. In order to see if the master events were similar, we adopted the method of Thelen et al. (2011), who cross correlate all of the master events against one another, but use a higher cross correlation coefficient as a threshold for similarity (≥ 0.8). This is a results of a trade off between maintaining an acceptable cross correlation coefficient (≥ 0.7) for each of the individual events classified into each family, and reducing bias from the combination of multiplets across several days. Table 3 shows that with a cross correlation coefficient of 0.8, the multiplets can be further organised into a number of subfamilies, for example master events 100, 106 and 210 can all be combined into a single multiplet family with a cross correlation coefficient greater than the threshold. A lower threshold allows each of the multiplets to be placed into fewer families, however the visual similarity between the waveforms is lost. For example, with a cross correlation coefficient of 0.65 the multiplets can be sorted into just two families, however there is a much larger scatter in the waveform similarity.

Using a smaller subset of multiplets (i.e. by combining families which have a cross correlation coefficient ≥ 0.8) also provides an accurate forecast of the dome collapse event in June 1997. For example, combining master events 100, 106 and

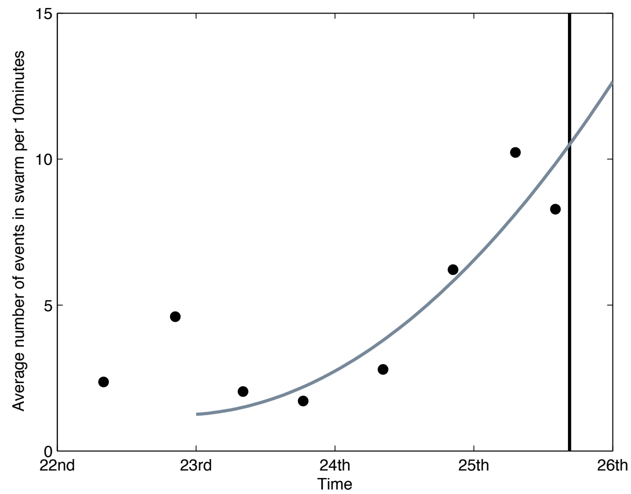
Master Event	001	010	014	100	106	121	136	141	210	291
001	1.000	0.778	0.567	0.590	0.669	0.727	0.699	0.700	0.453	0.682
010		1.000	0.694	0.389	0.462	0.590	0.736	0.359	0.468	0.599
014			1.000	0.776	0.614	0.699	**0.828**	0.536	0.734	0.666
100				1.000	**0.845**	0.783	0.656	0.794	**0.811**	0.707
106					1.000	0.574	0.384	0.754	0.592	0.463
121						1.000	0.789	0.762	0.732	**0.827**
136							1.000	0.573	0.776	0.798
141								1.000	0.668	0.701
210									1.000	0.794
291										1.000

Table 3: Cross Correlation Coefficients of each identified master event with every other master event. Values highlighted in bold represent those whose cross correlation coefficient exceeds the threshold of 0.8, and are therefore deemed to be similar. In this case, master events 100, 106 and 210 were stacked and an average waveform taken to provide a new master event. The same was done for master events 014 and 136, and 121 and 291.

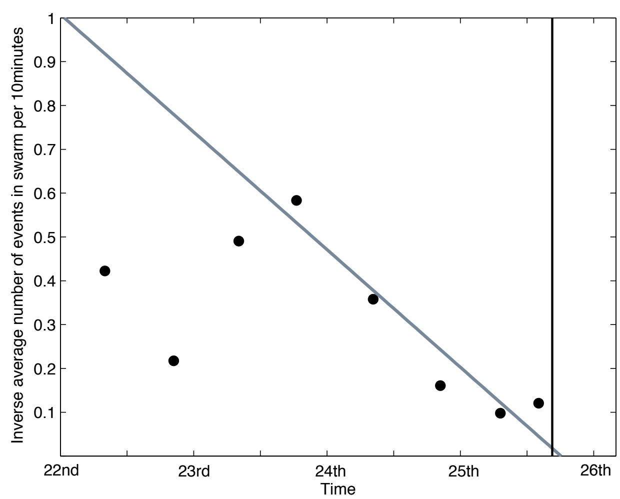
210 by stacking them at their maximum point of correlation and then taking an average of this stack, does not appear to produce a distinct acceleration which is tending towards a singularity (Figure 11(a)). However, upon application of the FFM we find that the timing of the dome collapse is forecasted less than 2 hours away from the known timing of the dome collapse, with a high degree of certainty between the linear regression and the inverse event rate ($R^2 = 0.9054$) (Figure 11), once a clear regression has been identified.

5.3. The Generation of Multiplets

It would be a remarkable coincidence if all of the successful forecasts of volcanic eruptions using the FFM, whether in hindsight or real-time were chance occurrences, suggesting that some real link between the activity at depth and the



(a) The average event rate per 10 minutes within swarms from 22 - 25 June. The acceleration of these events is depicted with the curve.



(b) Application of the FFM: the inverse average event rate per 10 minutes within swarms from 22 - 25 June. The graphical representation of the FFM is depicted by the linear regression (it is assumed that $\alpha = 2$) and the forecasted timing of failure can be read off the x-axis at the point where the linear regression crosses it.

Figure 11: The acceleration of swarms and application of the FFM to seismicity found in swarms which are a result of a new master event as a results of the combination of master events 100, 106 and 210. Each dot therefore represents the average event rate for each individual swarm. The vertical line represents the known timing of dome collapse on the 25 June 1997 at 16:55 UTC. Since swarms 1-3 did not produce a significant ongoing negative linear regression they were omitted from the application of the FFM.

surface must be plausible. A number of models have been proposed to explain the occurrence of low frequency seismicity in volcanic settings. The model of Iverson et al. (2006) suggests that the generation of low frequency seismicity occurs as a magmatic plug moves incrementally upwards within a conduit due to the movement of buoyant magmatic fluid behind the plug. In this instance it would expected for seismicity to migrate with the movement of the plug, and therefore become shallower with time. This is not observed on Montserrat, where seismicity consistently occurs at ≈1500 m depth below the dome summit (Aspinall et al., 1998; Rowe et al., 2004; Ottemöller, 2008; De Angelis and Henton, 2011). In addition, families of similar seismic waveforms must occur within a small spatial extent in order to maintain their similarity, and single families have been observed being sustained over a number of hours and days. These observations at Soufrière Hills volcano are inconsistent with this model, since the same family of earthquakes would not be sustained over this period of time without major changes to the similarity of the waveforms. The evolution of cross correlation coefficients, such as that which were observed in July 2003, would require only a very small migration of the seismicity which would not be expected from the incremental movement of a volcanic plug, since it can still be classed as from the same family over the sustained period of time. This model also fails to explain the occurrence of a number of distinct families as observed in June 1997, which are repeatedly activated then deactivated.

The more recent model of Bean et al. (2014) suggests that slow rupture failure within unconsolidated volcanic material in the edifice can induce low frequency seismicity, whose waveform characteristics are fundamentally dependent upon the wave propagation path. It it is envisaged that families and swarms of low frequency seismicity develop due to slow deformation at a number of points within the upper edifice where the stress is reduced which could be induced by gas influx, gravity or magma migration. On Montserrat, the seismicity systematically occurs at the same depth such that it would require the stress drop to be maintained at the same location over a number of years. In addition, no explanation is given for the clear acceleration in seismicity which has been observed prior to the dome collapse events studied.

The generation of low frequency seismicity has also been attributed to the brittle failure of the magma itself (Webb and Dingwell, 1990; Goto, 1999) through an increase in viscosity of the melt and/or high strain rates (Lavallée et al., 2008) as the melt enters a glass transition stage. In volcanic environments, conditions which may induce this glass transition stage may include: changes in crystal and/or bubble content of the magma (Goto, 1999); an increase in the ascent rate of magma (Neuberg et al., 2006); or through the introduction of a restriction in the conduit (Thomas and Neuberg, 2012). The brittle failure model allows for the acceleration in LF seismicity observed at Soufrière Hills prior to dome collapses, since accelerations in magma ascent has been shown to increase the strain rate within a volcanic conduit simulation, and therefore instigate the brittle failure of the melt (Neuberg et al., 2006). A much simpler way to induce an increased strain rate is to introduce a constriction within the conduit (Thomas and Neuberg, 2012). Such a constriction fits with observations at Soufrière Hills that LF seismicity consistently occurs at ≈ 1500 m below the dome summit (Aspinall et al., 1998; Rowe et al., 2004; Ottemöller, 2008; De Angelis and Henton, 2011), as well as the fact that multiple LF sources (i.e. families of similar seismic events) may be active at any given time, since a number of locations may exist where the strain rate threshold for brittle failure is overcome.

This model not only accounts for the acceleration in the number of multiplets that is observed and the stable source mechanism and its location, but can also account for the possible phase delay between the timing of failure at depth (i.e. when a magmatic pathway is created) and the surface expression of this failure (Voight, 1988). This could be represented by a time delay in the FFM, which forecasts the timing of failure before the known timing of the dome collapse since technically the FFM is forecasting the failure at depth related to the accelerating seismicity, rather than the surface manifestation that we observe and try to relate it to.

5.4. Forecasting Potential

The concept of forecasting using the FFM in hindsight analysis once the eruption has occurred is common (e.g. Cornelius and Voight (1994); Kilburn and Voight (1998); De la Cruz-Reyna and Reyes-Dávila (2001); Ortiz et al. (2003); Hammer and Neuberg (2009); Smith and Kilburn (2010)). It is much less common to employ and rely upon these tactics during developing unrest as huge responsibility is placed upon generating accurate forecasts which are often simply plagued with too many uncertainties. Using a cross correlation technique in conjunction with the FFM allows the isolation of single system at depth. Isolating a single system at depth avoids additional uncertainties introduced by averaging data over a number of different accelerating phenomena, and consequently reduces the misfit between the data and the forecast. On occasions when precursory seismicity could be identified prior to dome collapses at SHV in June 1997 and July 2003, use of similar seismicity and the FFM provided a more successful and more accurate forecast to the timing of collapse events than simply using the FFM in isolation with all incoming low frequency seismicity. Further investigation is required to determine whether these techniques are applicable to other volcanoes around the world, and whether it can be used to forecast volcanic phenomena other than dome collapses.

If this technique is to be used in real time forecasting, the identification of similar seismic events will also need to be undertaken in real time, i.e. a continuous search for similar events and defining of master events will have to become part of routine monitoring. Experience at Soufrière Hills volcano teaches us that some similar seismic events are consistent over a number of days (e.g. June 1997, July 2003) and therefore their early identification would allow simple manipulation of the incoming seismicity. In other circumstances, the master events for similar seismicity may be short-lived, which requires the recalculation of master events more often. There are also occasions when similar seismic events may not be apparent in the seismicity at all (e.g. Soufrière Hills, February 2010) and other forecasting methods will need to be relied upon.

Other parameters used in forecasting and the identification of similarity will also need to be undertaken in real time. In this study, we have found that the use of 0.7 as a correlation threshold is appropriate for identifying different families of similar seismicity. However, Ottemöller (2008) used a cross correlation threshold of between 0.6 and 0.66 for continuous data from 9 to 12 July 2003 at SHV, allowing the identification of far more similar seismic events. Upon visual inspection of our data, we found these cross correlation thresholds to be too low to identify similar seismicity without noise. The determination of a cross correlation threshold is extremely important since if it is placed too low then there is a risk of placing events which are not similar into the same family, and if it is too high there may be many similar events which are not detected due to poor signal to noise ratios. Identifying this threshold in real time is likely to be another difficult parameter, and further analysis of more swarms of similar seismicity is required to determine whether it is always appropriate to use 0.7 as a threshold at SHV.

As the FFM follows a least squares regression analysis when α is equal to 2, the residual error between the observed event rate and the mean event rate should follow a typical Gaussian distribution (Bell et al., 2011b). Greenhough and Main (2008) have suggested that since earthquake occurrence is a point process, the rate uncertainties are best described by a Poisson distribution. This was also suggested by Bell et al. (2011a) who determined that the daily earthquake rates at Mauna Loa preceding the 1984 eruption were consistent with a Poisson regime, within 95 per cent confidence limits. In this instance, a generalised linear model (GLM) where α = 1, rather than a least squares regression model (α = 2) may be more appropriate, since it can allow for a distribution of data that is non-Gaussian (Bell et al., 2011b). However, although using a GLM as a fitting tool does provide a higher R^2 value for each of the forecasts generated from similar seismicity (always > 0.8), suggesting a GLM is a better fit to the data than a least squares linear regression, the forecasted timing of the dome collapse in 1997 was consistently late (the best forecast was still over 30 hours away from the known timing of failure). Moreover, forecasting using a GLM for the July 2003 collapse generated a forecast over a week from the known timing of the collapse when using all of the available data. If only the accelerating swarms are used, the forecast is over 72 hours after the known timing of failure. This is in contrast to Bell et al. (2011b) who suggest that the GLM provides more accurate forecasts for the timings of eruptions than the FFM with a linear least squares regression.

Although the GLM solves the problem of needing to use a model which can account for the appropriate error structure, its use may not be applicable in a volcanic setting (Hammer and Ohrnberger, 2012). Forecasting volcanic eruptions using the FFM and rates of temporal seismicity requires that the system has a memory, and therefore that events which have occurred before can influence the outcome in the future. A poisson process is the exact opposite to this: it requires a memoryless system, in which events evolve independently. This would therefore make the use of the GLM with $\alpha = 1$ and FFM together invalid, since one of the overriding assumptions of the FFM is that previous geophysical observables form the basis of the forecast, and therefore suggesting that the system has a memory. Analysis of a number of dome building eruptions at Mt. St. Helens in 1985 and 1986, which may be comparable to Soufrière Hills volcano, suggested that an exponential model did not adequately explain the precursory trends in earthquake event rates (Bell et al., 2013), possibly implying that an exponential model is not appropriate for forecasting at andesitic dome building volcanoes, which may be one reason as to why the forecasts using a GLM in this instance were not successful.

6. Conclusions

Utilizing the cross correlation technique for the forecasting of large scale dome collapses at SHV appears to provide more consistent and more accurate forecasts than when using all precursory low frequency seismicity, since we can assume that we are forecasting using only one active system at depth. The potential for an improved forecast by isolating a single system was first proposed by Kilburn and Voight (1998) who suggested that two populations may be in effect at SHV in a qualitative manner from graphics alone (their Figure 2(c)). Here we show that this is indeed the case by a quantitative method of cross correlation for dome collapses in June 1997 and July 2003. The most recent dome collapse at SHV in February

2010 was not able to be forecast using the FFM since no precursory seismicity was detected. This suggests that not every dome collapse at SHV is forecastable, and suggests some events may come without warning. However, when precursory seismicity is detected, the cross correlation technique allows for a five-fold increase in the number of detectable events from the continuous record, and therefore provides a more holistic picture of the ongoing seismicity. The magnitude of the events and the possible superposition of events is insignificant for detection since the continuous seismic record is normalised in near real-time and the search criteria can be set to smaller increments within the cross correlation procedure to find closely spaced events.

In June 1997, 10 families of similar waveforms were detected, signalling a number of active sources at depth occurring at the same time. Using any one of these families of seismicity in conjunction with the FFM provided more accurate forecasts to the timing of the dome collapse, with a greater degree of confidence due to high R^2 values which suggest that the model (FFM) fits to the data well (Figure 9 and Table 2). In July 2003, only one family of events were identified, in agreement with Ottemöller (2008). Analysis of the cross correlation coefficients suggests a migration of the source with time (Figure 6). Significantly, and in contrast to the events of June 1997, the similar seismicity ceased hours before the dome collapse, perhaps an indication of a delay function between the seismicity at depth and the collapse at the surface as first envisaged by Voight (1988). Forecasting using only the accelerating swarms in July 2003 provided an accurate forecast for the timing of the dome collapse, but proves the difficulty of using the FFM in real time, as the last swarm in the sequence led to a forecast a number of weeks from the known timing of collapse. Despite clear cyclic activity in RSAM (Figure 7), no families of similar seismic events could be identified in the precursory seismicity of the February 2010 collapse. This echoes the conclusions of Stinton et al. (2014) who suggested that no acceleration in seismicity was observed prior to the collapse, and in fact, seismicity remained remarkably low. Further investigation is required to determine whether the cross correlation technique used in conjunction with the FFM is applicable for forecasting other volcanic phenomena at other volcanoes around the world.

The overwhelming question in forecasting volcanic eruptions however still remains: how can we tell the difference between accelerating seismicity which is precursory to an eruptive event, and that which does not appear to lead to a surface expression? This is fundamental for forecasting, since the generation of false alarms can lead to deteriorating confidence in the observatory making the forecasts. False alarms are an inherent part of forecasting volcanic eruptions; the forecasts are never going to be 100% correct, 100% of the time, primarily due to the incidental nature of nature itself. However, keeping false alarms to a minimum is essential. Forecasting is complicated by precursory activity often having more than one acceleration event, and in particular at SHV having cyclic acceleration events (e.g. Figure 7). The FFM does not account well for multiple accelerations within a system, which is why searching for an overall acceleration in the precursory activity (e.g. taking the average event rate of each swarm of activity, rather than per unit time) allows a more successful application. The use of the cross correlation technique in conjunction with the FFM appears to further enhance the success of the forecast since focus is on a single active seismic system at depth and therefore can be directly related to surface activity.

Acknowledgements

We would like to thank all of the current and past staff at the Montserrat Volcano Observatory who continue to monitor and maintain seismic stations around the volcano, without whom this data would not have been available. ROS was funded through a NERC studentship at the University of Leeds (NE/J50001X/1). JWN received funding from the European Union Framework Program 7 (Grant 282759, “VUELCO”) and from the NERC/ESRC project “Strengthening Resilience in Volcanic Areas” (STREVA). We would like to thank William Murphy for initial discussions. We would also like to thank two anonymous reviewers for helpful, constructive reviews that greatly improved the manuscript.

References

Aspinall, W., Miller, A., Lynch, L., Latchman, J., Stewart, R., White, R., Power, J., 1998. Soufrière Hills eruption, Montserrat, 1995-1997: volcanic earthquake locations and fault plane solutions. Geophysical Research Letters 25, 3397–3400.

Barrett, J. P., 1974. The coefficient of determinationsome limitations. The American Statistician 28 (1), 19–20.

Bean, C. J., De Barros, L., Lokmer, I., Metaxian, J. P., O Brien, G., Murphy, S., 2014. Long-period seismicity in the shallow volcanic edifice formed from slow-rupture earthquakes. Nature Geoscience 7 (1), 71–75.

Bell, A., Greenhough, J., Heap, M., Main, I., 2011a. Challenges for forecasting based on accelerating rates of earthquakes at volcanoes and laboratory analogues. Geophysical Journal International.

Bell, A., Naylor, M., Heap, M., Main, I., 2011b. Forecasting volcanic eruptions and other material failure phenomena: An evaluation of the failure forecast method. Geophysical Research Letters 38, L15304.

Bell, A. F., Naylor, M., Main, I. G., 2013. The limits of predictability of volcanic eruptions from accelerating rates of earthquakes. Geophysical Journal International 194 (3), 1541–1553.

Benoit, J. P., McNutt, S. R., 1996. Global volcanic earthquake swarm database 1979-1989. US Department of the Interior, US Geological Survey.

Buurman, H., West, M. E., Thompson, G., 2013. The seismicity of the 2009 Redoubt eruption. Journal of Volcanology and Geothermal Research 259, 16–30.

Caplan-Auerbach, J., Petersen, T., 2005. Repeating coupled earthquakes at Shishaldin Volcano, Alaska. Journal of Volcanology and Geothermal Research 145 (1), 151–172.

Chouet, B., 1988. Resonance of a fluid-driven crack: Radiation properties and implications for the source of long-period events and harmonic tremor. Journal of Geophysical Research: Solid Earth (1978–2012) 93 (B5), 4375–4400.

Chouet, B. A., Page, R. A., Stephens, C. D., Lahr, J. C., Power, J. A., 1994. Precursory swarms of long-period events at Redoubt Volcano (1989–1990), Alaska: Their origin and use as a forecasting tool. Journal of Volcanology and Geothermal Research 62 (1), 95–135.

Cornelius, R., Voight, B., 1994. Seismological aspects of the 1989–1990 eruption at Redoubt Volcano, Alaska: the Materials Failure Forecast Method (FFM) with RSAM and SSAM seismic data. Journal of Volcanology and Geothermal Research 62 (1), 469–498.

Cornelius, R., Voight, B., 1995. Graphical and PC-software analysis of volcano eruption precursors according to the Materials Failure Forecast Method (FFM). Journal of Volcanology and Geothermal Research 64 (3-4), 295–320.

Cruz, F. G., Chouet, B. A., 1997. Long-period events, the most characteristic seismicity accompanying the emplacement and extrusion of a lava dome in Galeras Volcano, Colombia, in 1991. Journal of Volcanology and Geothermal Research 77 (1), 121–158.

De Angelis, S., Henton, S., 2011. On the feasibility of magma fracture within volcanic conduits: Constraints from earthquake data and empirical modelling of magma viscosity. Geophysical Research Letters 38 (19).

De la Cruz-Reyna, S., Reyes-Dávila, G. A., 2001. A model to describe precursory material-failure phenomena: applications to short-term forecasting at Colima volcano, Mexico. Bulletin of Volcanology 63 (5), 297–308.

Fukuzono, T., 1985. A new method for predicting the failure time of a slope. In: Proceedings of the 4th International Conference and Field Workshop in Landslides, Tokyo. pp. 145–150.

Geller, R., Mueller, C., 1980. Four similar earthquakes in central California. Geophysical Research Letters 7 (10), 821–824.

Goto, A., 1999. A new model for volcanic earthquake at Unzen Volcano: Melt rupture model. Geophysical Research Letters 26 (16), 2541–2544.

Green, D., Neuberg, J., 2006. Waveform classification of volcanic low-frequency earthquake swarms and its implication at Soufrière Hills Volcano, Montserrat. Journal of Volcanology and Geothermal Research 153 (1), 51–63.

Greenhough, J., Main, I., 2008. A poisson model for earthquake frequency uncertainties in seismic hazard analysis. Geophysical Research Letters 35 (19).

Hammer, C., Neuberg, J., 2009. On the dynamical behaviour of low-frequency earthquake swarms prior to a dome collapse of Soufrière Hill volcano, Montserrat. Geophysical Research Letters 36 (6), L06305.

Hammer, C., Ohrnberger, M., 2012. Forecasting seismo-volcanic activity by using the dynamical behavior of volcanic earthquake rates. Journal of Volcanology and Geothermal Research 229, 34–43.

Herd, R. A., Edmonds, M., Bass, V. A., 2005. Catastrophic lava dome failure at Soufriere Hills volcano, Montserrat, 12–13 July 2003. Journal of Volcanology and Geothermal Research 148 (3), 234–252.

Iverson, R., Dzurisin, D., Gardner, C., Gerlach, T., LaHusen, R., Lisowski, M., Major, J., Malone, S., Messerich, J., Moran, S., et al., 2006. Dynamics of seismogenic volcanic extrusion at Mount St Helens in 2004–05. Nature 444 (7118), 439–443.

Kilburn, C., 2003. Multiscale fracturing as a key to forecasting volcanic eruptions. Journal of Volcanology and Geothermal Research 125 (3-4), 271–289.

Kilburn, C. R., Voight, B., 1998. Slow rock fracture as eruption precursor at Soufriere Hills volcano, Montserrat. Geophysical Research Letters 25 (19), 3665–3668.

Lahr, J., Chouet, B., Stephens, C., Power, J., Page, R., 1994. Earthquake classification, location, and error analysis in a volcanic environment: Implications for the magmatic system of the 1989–1990 eruptions at Redoubt Volcano, Alaska. Journal of Volcanology and Geothermal Research 62 (1), 137–151.

Lavallée, Y., Meredith, P., Dingwell, D., Hess, K.-U., Wassermann, J., Cordonnier, B., Gerik, A., Kruhl, J., 2008. Seismogenic lavas and explosive eruption forecasting. Nature 453 (7194), 507–510.

Loughlin, S., Calder, E., Clarke, A., Cole, P., Luckett, R., Mangan, M., Pyle, D., Sparks, R., Voight, B., Watts, R., 2002. Pyroclastic flows and surges generated by the 25 June 1997 dome collapse, Soufrière Hills Volcano, Montserrat. Memoirs-Geological Society Of London 21, 191–210.

Loughlin, S., Luckett, R., Ryan, G., Christopher, T., Hards, V., De Angelis, S., Jones, L., Strutt, M., 2010. An overview of lava dome evolution, dome collapse and cyclicity at Soufrière Hills Volcano, Montserrat, 2005–2007. Geophysical Research Letters 37 (19).

Luckett, R., 2005. Seismic Data from the Montserrat Eruption at BGS. Tech. rep., British Geological Survey Open Report, OR/09/57.

Matthews, A. J., Barclay, J., Carn, S., Thompson, G., Alexander, J., Herd, R., Williams, C., 2002. Rainfall-induced volcanic activity on Montserrat. Geophysical Research Letters 29 (13), 22–1.

McNutt, S. R., 2002. Volcano seismology and monitoring for eruptions. International Geophysics Series 81 (A), 383–406.

Miller, A., Stewart, R., White, R., Luckett, R., Baptie, B., Aspinall, W., Latchman, J., Lynch, L., Voight, B., 1998. Seismicity associated with dome growth and collapse at the Soufriere Hills volcano, Montserrat. Geophysical Research Letters 25 (18), 3401–3404.

Neuberg, J., Luckett, R., Baptie, B., Olsen, K., 2000. Models of tremor and low-frequency earthquake swarms on Montserrat. Journal of Volcanology and Geothermal Research 101 (1-2), 83–104.

Neuberg, J., Tuffen, H., Collier, L., Green, D., Powell, T., Dingwell, D., 2006. The trigger mechanism of low-frequency earthquakes on Montserrat. Journal of Volcanology and Geothermal research 153 (1), 37–50.

Ortiz, R., Moreno, H., Garcıa, A., Fuentealba, G., Astiz, M., Peña, P., Sánchez, N., Tárraga, M., 2003. Villarrica volcano (Chile): characteristics of the volcanic tremor and forecasting of small explosions by means of a material failure method. Journal of Volcanology and Geothermal Research 128 (1), 247–259.

Ottemöller, L., 2008. Seismic hybrid swarm precursory to a major lava dome collapse: 9–12 July 2003, Soufriere Hills Volcano, Montserrat. Journal of Volcanology and Geothermal Research 177 (4), 903–910.

Paulatto, M., Minshull, T., Baptie, B., Dean, S., Hammond, J., Henstock, T., Kenedi, C., Kiddle, E., Malin, P., Peirce, C., et al., 2010. Upper crustal structure of an active volcano from refraction/reflection tomography, Montserrat, Lesser Antilles. Geophysical Journal International 180 (2), 685–696.

Petersen, T., 2007. Swarms of repeating long-period earthquakes at Shishaldin Volcano, Alaska, 2001–2004. Journal of Volcanology and Geothermal Research 166 (3), 177–192.

Rowe, C., Thurber, C., White, R., 2004. Dome growth behavior at Soufriere Hills Volcano, Montserrat, revealed by relocation of volcanic event swarms, 1995–1996. Journal of Volcanology and Geothermal Research 134 (3), 199–221.

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. Journal of Volcanology and Geothermal Research 191 (1), 129–136.

Sparks, R., Young, S., 2002. The eruption of Soufriere Hills Volcano, Montserrat (1995-1999): overview of scientific results. Geological Society, London, Memoirs 21 (1), 45–69.

Stephens, C., Chouet, B., 2001. Evolution of the December 14, 1989 precursory long-period event swarm at Redoubt Volcano, Alaska. Journal of Volcanology and Geothermal Research 109 (1), 133–148.

Stinton, A. J., Cole, P. D., Stewart, R. C., Odbert, H. M., Smith, P., 2014. The 11 February 2010 partial dome collapse at Soufrière Hills Volcano, Montserrat. Geological Society, London, Memoirs 39 (1), 133–152.

Thelen, W., Malone, S., West, M., 2011. Multiplets: Their behavior and utility at dacitic and andesitic volcanic centers. Journal of Geophysical Research: Solid Earth (1978–2012) 116 (B8).

Thomas, M. E., Neuberg, J., 2012. What makes a volcano tick –A first explanation of deep multiple seismic sources in ascending magma. Geology 40 (4), 351–354.

Trnkoczy, A., 2002. Understanding and parameter setting of STA/LTA trigger algorithm. IASPEI New Manual of Seismological Observatory Practice 2, 1–19.

Umakoshi, K., Shimizu, H., Matsuwo, N., 2003. Seismic activity associated with the endogenous growth of lava dome at Unzen Volcano, Japan. Tech. rep., Abstract V10.

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 (4888), 200–203.

Voight, B., Cornelius, R., 1991. Prospects for eruption prediction in near real-time. Nature 350 (6320), 695–698.

Voight, B., Komorowski, J., Norton, G., Belousov, A., Belousova, M., Boudon, G., Francis, P., Franz, W., Heinrich, P., Sparks, R., et al., 2002. The 26 December (Boxing Day) 1997 sector collapse and debris avalanche at Soufriere Hills volcano, Montserrat. Geological Society, London, Memoirs 21, 363–408.

Voight, B., Sparks, R., Miller, A., Stewart, R., Hoblitt, R., Clarke, A., Ewart, J., Aspinall, W., Baptie, B., Calder, E., et al., 1999. Magma flow instability and cyclic activity at Soufriere Hills volcano, Montserrat, British West Indies. Science 283 (5405), 1138–1142.

Wadge, G., Voight, B., Sparks, R., Cole, P., Loughlin, S., Robertson, R., 2014. An overview of the eruption of Soufriere Hills Volcano, Montserrat from 2000 to 2010. Geological Society, London, Memoirs 39 (1), 1–40.

Webb, S. L., Dingwell, D. B., 1990. Non-newtonian rheology of igneous melts at high stresses and strain rates: Experimental results for rhyolite, andesite, basalt, and nephelinite. Journal of Geophysical Research: Solid Earth (1978–2012) 95 (B10), 15695–15701.

White, R., Miller, A., Lynch, L., Power, J., 1998. Observations of hybrid seismic events at Soufriere Hills volcano, Montserrat: July 1995 to September 1996. Geophysical Research Letters 25 (19), 3657–3660.

Withers, M., Aster, R., Young, C., Beiriger, J., Harris, M., Moore, S., Trujillo, J., 1998. A comparison of select trigger algorithms for automated global seismic phase and event detection. Bulletin of the Seismological Society of America 88 (1), 95–106.

Young, S. R., Sparks, R. S. J., Aspinall, W. P., Lynch, L. L., Miller, A. D., Robertson, R. E., Shepherd, J. B., 1998. Overview of the eruption of Soufriere Hills volcano, Montserrat, 18 July 1995 to December 1997. Geophysical Research Letters 25 (18), 3389–3392.