Systematic mining and reanalysis of large volcano-seismic waveform datasets

Robin S. Matoza[1], Peter M. Shearer[1], Paul G. Okubo[2], Bernard Chouet[3], Phil Dawson[3], and Seth Moran[4]*

[1]*Institute of Geophysics and Planetary Physics, Scripps Institution of Oceanography, UC San Diego*
[2]*US Geological Survey, Hawaiian Volcano Observatory*
[3]*US Geological Survey, Menlo Park*
[4]*US Geological Survey, Cascades Volcano Observatory*

Summary

The analysis and interpretation of seismicity from mantle depths to the surface plays a key role in understanding how volcanoes work. We are developing and applying methods for the systematic reanalysis of waveforms from volcano-seismic networks, including high-precision earthquake relocation, spectral event classification, and source mechanism and stress drop estimates. Our datasets include 23 years of tectonic, volcano-tectonic, and long-period (LP, 0.5–5 Hz) seismicity from the ~50-station permanent network of the USGS Hawaiian Volcano Observatory (HVO), and millions of LP seismic events recorded on ~35 stations during the 2004–2008 eruption of Mount St. Helens, WA.

Introduction

Volcanoes generate seismicity by a full spectrum of fluid and solid processes. *Long-period seismicity* (LP, 0.5–5 Hz) is a class of volcano-seismic signal that is of particular interest. LP seismicity includes individual transient LP events and more continuous volcanic tremor. These signals are used routinely by volcano monitoring scientists to forecast volcanic eruptions despite an incomplete understanding of their origin (e.g., *Chouet et al.*, 1994; *McNutt*, 1996). LP events often consist of a brief broadband onset, followed by a coda of decaying harmonic oscillations containing pronounced spectral peaks that are independent of azimuth and distance to the source. This is commonly interpreted as a broadband, time-localized pressure excitation mechanism (or trigger mechanism), followed by the volumetric response of a fluid-filled resonator (e.g., *Chouet and Matoza,* 2013).

LP seismicity has been recorded for decades in the summit region of Kilauea Volcano, Hawaii, and is postulated as linked with the magma transport and shallow hydrothermal systems. The 2004–2008 eruption of Mount St. Helens produced millions of LP events, only a small fraction of which have been analyzed or modeled in detail. Many of the LP events during this eruption occurred with such precise regularity that they were termed "drumbeats" (*Moran et al.,* 2008), a phenomenon that has been observed at several other volcanoes (e.g., *Lees et al.,* 2008; *Power and Lalla,* 2014; *Firstov and Shakirova,* 2014).

Method

We have been developing a modular set of codes for the mining and analysis of large waveform datasets from volcanoes, based on previous work analyzing large event-based regional seismic datasets from Southern California (*Shearer*, 1997; *Shearer et al.*, 2005; *Lin et al.*, 2007). The processing steps include automated network-based template-matching detection to build an event database, analyses of the evolutionary dynamics of event metrics, waveform cross-correlation and high-precision relocations, and full-waveform inversions using representative stacks of similar events. Our combined cluster analysis and relative relocation method is described in more detail by *Matoza et al.* (2013, 2014).

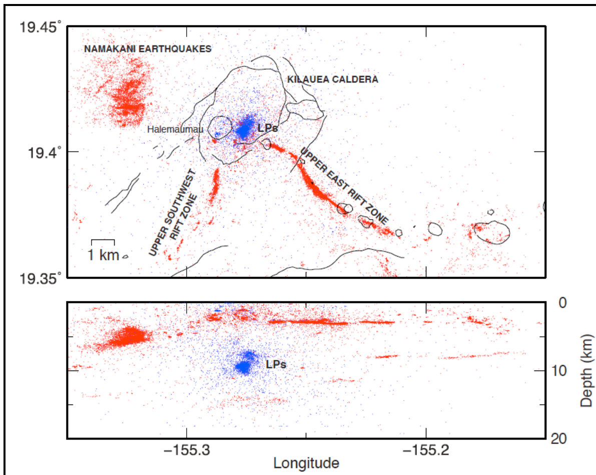

Figure 1: High-precision relocations of seismicity in the summit region of Kilauea Volcano, Hawaii from 1986 to 2009 (*Matoza et al.* 2013, 2014). Blue dots are events we automatically classify as LP and red dots are non-LP. We produce a dramatic sharpening of earthquake locations along faults and magmatic features compared to standard earthquake catalog locations.

Example 1: Kilauea Volcano, Hawaii

We begin with event-triggered waveform data from the HVO permanent seismic network for 49,030 events occurring in the Kilauea summit region from January 1986 to March 2009. We also begin with catalog locations, a one-dimensional (1D) velocity model, and phase-pick data available for a subset of the events from HVO. We must first isolate LP events from the rest of the seismicity (tectonic and volcano-tectonic events) occurring in the summit region (Figure 1). The event classification scheme used at HVO is subjective and has not been consistently or systematically applied through time. We employ an automated method that can be applied in a standardized way to all event data, using the Frequency Index (FI) spectral ratio metric introduced by *Buurman and West* (2010) (Figure 1). We estimate 215,437 *P*-wave spectra, considering all events on all stations, and use the station-averaged FI to consistently classify LP and non-LP seismicity. We compute high-precision relative relocations for 5,327 LP events (43% of all classified LP events) using waveform cross correlation and cluster analysis with 6.4 million event pairs, combined with the source-specific station term method. The majority of intermediate depth (5-15 km) LPs collapse to a compact volume, with remarkable source location stability over 23 years indicating a source process controlled by geological or conduit structure (Figure 1) (*Matoza et al.*, 2014).

Example 2: Mount St. Helens, WA

An impressive seismic waveform dataset capturing rich and fascinating phenomena was collected during the 2004–2008 eruption of Mount St. Helens (e.g., *Moran et al.,* 2008; *Thelen et al.,* 2008; *Waite et al.,* 2008, *Matoza et al.,* 2009). However, the vast majority of these small volcanic earthquakes (LPs) did not make it into published catalogs and were primarily studied using single-channel analyses. The lack of an event catalog is the primary obstacle to performing large-scale systematic analyses such as those performed for Southern California and Hawaii.

We build an event catalog using a template matching approach, scanning an initial "seed" event through the continuous waveform data to produce correlation-coefficient functions on each station (Figure 2). We then stack the correlation coefficient functions for all stations across the network, resulting in considerable signal-to-noise ratio (SNR) gain and clearer detection thresholds. We define events to occur when the stacked correlation coefficient exceeds a given threshold, and we can choose not to allow the peaks in the correlation coefficient (i.e., events) to be too close together (Figure 2). Next, we re-extract the waveforms associated with each trigger, and

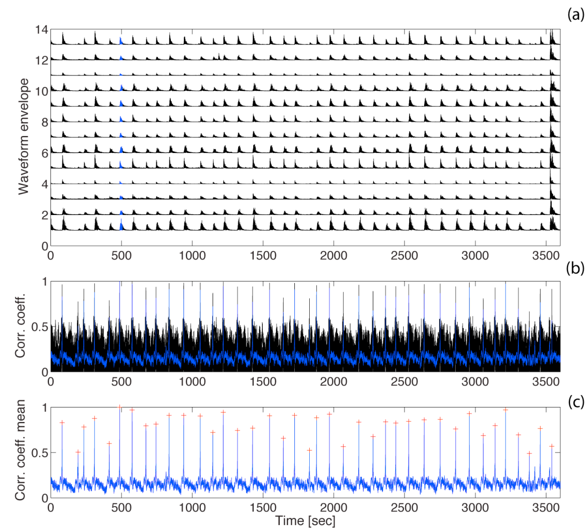

Figure 2: Illustration of the template matching approach using a 1-hour data example from 13 broadband stations of the temporary YB network (see *Waite et al*., 2008). (a) Normalized envelope functions of the 13 stations, with an arbitrarily and manually chosen starting or "seed" event shown in blue. (b) Running (1-sample increment) correlation coefficient with the chosen template. Individual stations are shown in black and the stack across the network is shown in blue. Note the considerable signal-to-noise ratio (SNR) gain for the stack. (c) The stacked correlation coefficient function is used to define event triggers. Red plus symbols are the event picks. In this example, an event is defined when the network-mean correlation coefficient exceeds 0.45 and events cannot be closer than 5 s apart in time.

stack them together to form a master event with higher signal-to-noise ratio. Finally, we scan the master event through all the data to build up an event catalog, and construct an event-based waveform database. Our database is built using a custom binary format that stores all event, waveform, and phase-pick data for a given event in a single file.

We apply our template matching algorithm to continuous waveform data from 35 stations (a mix of broadband and short-period stations) around Mount St. Helens that are publicly available at the IRIS DMC, from 1 October to 31 December 2005, identifying 94,365 events. We then pair each event with every other event within 5 days and perform *P*-wave cross-correlations for 409 million total pairs to estimate differential arrival times. We use this cross-correlation information to relocate 40,150 events (42% of original events) using 186 million "good" differential time pairs via the method described by *Matoza et al.,* (2013, 2014). The results (Figure 3) show seismicity collapsing to a tiny compact source volume, which is consistent with previous studies at this and other volcanoes.

We are currently working on applying full waveform inversion methods (e.g., *Chouet et al.*, 2005; *Chouet and Dawson*, 2011) on representative stacks of events from the major event families. So far, these results point to dominantly volumetric source-mechanisms for LP events and smaller "subevents" at Mount St. Helens (*Matoza and Chouet*, 2010). A wide range of observed seismic phenomena during the 2004–2008 eruption of Mount St. Helens, including subevents, LP events, larger (M_d > 2) events, and phreatic explosions, can be explained within the framework of the interaction of shallow hydrothermal and magmatic systems.

Conclusions

Automated spectral identification and relocation of LP seismicity near the summit region of Kilauea Volcano shows that most intermediate depth (5–15 km) LP events occur within a compact volume that has remained at a fixed location for over 23 years, indicating a structural control. The integrated analyses of seismicity at Mount St. Helens point to the repetitive action of a non-destructive source process within a compact volume, interpreted as the cyclic collapse, resonance, and pressurization of sub-horizontal cracks within a perched shallow hydrothermal system.

References

Buurman, H., and M. E. West, 2010, Seismic precursors to volcanic explosions during the 2006 eruption of Augustine Volcano, in The 2006 eruption of Augustine Volcano, Alaska, edited by J. A. Power, M. L. Coombs, and J. T. Freymueller, chap. 2, pp. 41–57, U.S. Geological Survey Prof. Pap. **1769**.

Chouet, B., and P. Dawson, 2011, Shallow conduit system at Kilauea Volcano, Hawaii, revealed by seismic signals associated with degassing bursts. Journal of Geophysical Research: Solid Earth, **116,** 1978–2012.

Chouet, B.A., and R.S. Matoza, 2013, A multi-decadal view of seismic methods for detecting precursors of magma movement and eruption, J. Volcanol. Geotherm. Res., **252**, 108-175.

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

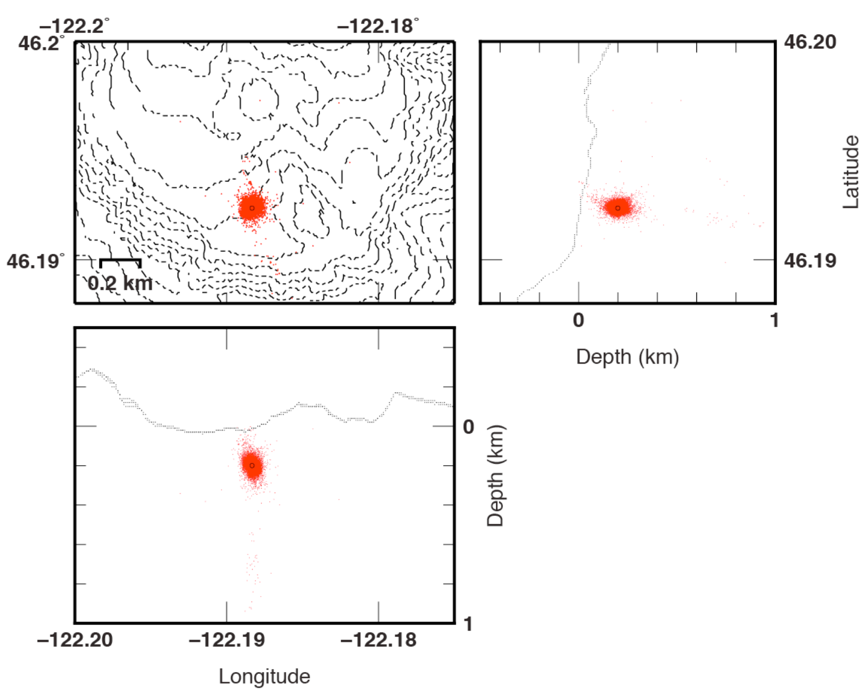

Figure 3: High-precision locations for 40,150 LP events at Mount St. Helens between 1 October and 31 December 2005 using waveform cross-correlation and a combined hierarchical cluster analysis and relocation method (*Matoza et al*., 2013; 2014). We identified 94,365 events during this time period using template matching, then performed *P*-wave cross-correlations for 409 million pairs of these events, where event pairs cannot be more than 5-days apart in time. The final 40,150 relocated events (42% of original events) are obtained using 186 million "good" differential times from 35 stations available during this time period.

Reanalyzing large volcano-seismic waveform datasets

Chouet, B., P. Dawson, and A. Arciniega-Ceballos, 2005, Source mechanism of Vulcanian degassing at Popocatepetl Volcano, Mexico, determined from waveform inversions of very long period signals. J. Geophys. Res., **110**, B070301.

Firstov, P.P., and A.A. Shakirova, 2014, Seismicity observed during the precursory process and the actual eruption of Kizimen Volcano, Kamchatka in 2009–2013. Journal of Volcanology and Seismology, **8**, 203-217.

Lees, J. M., J.B. Johnson, M. Ruiz, L. Troncoso, and M. Welsh, 2008, Reventador Volcano 2005: Eruptive activity inferred from seismo-acoustic observation, Journal of Volcanology and Geothermal Research, **176**, 179-190.

Lin, G., P. M. Shearer, and E. Hauksson, 2007, Applying a three-dimensional velocity model, waveform cross correlation, and cluster analysis to locate southern California seismicity from 1981 to 2005, J. Geophys. Res., **112**, B12309.

Matoza, R.S., and B.A. Chouet, 2010, Subevents of long-period seismicity: implications for hydrothermal dynamics during the 2004–2008 eruption of Mount St. Helens, Journal of Geophysical Research, **115**, B12206.

Matoza, R.S., M.A. Garces, B.A. Chouet, L. D'Auria, M.A.H. Hedlin, C. De Groot-Hedlin, and G.P. Waite, 2009, The source of infrasound associated with long-period events at Mount St. Helens. J. Geophys. Res., **114**.

Matoza, R.S., P.M. Shearer, G. Lin, C.J. Wolfe, and P.G. Okubo, 2013, Systematic relocation of seismicity on Hawaii Island from 1992 to 2009 using waveform cross correlation and cluster analysis, J. Geophys. Res. Solid Earth, **118**, 2275-2288.

Matoza, R.S., P.M. Shearer, and P.G. Okubo, 2014, High-precision relocation of long-period events beneath the summit region of Kilauea Volcano, Hawai`i, from 1986 to 2009, Geophys. Res. Lett., 3413-3421, **41**.

McNutt, S. R., 1996, Seismic monitoring and eruption forecasting of volcanoes: a review of the state-of-the-art and case histories. In Monitoring and mitigation of volcano hazards (pp. 99-146). Springer Berlin Heidelberg.

Moran, S.C., S.D. Malone, A.I. Qamar, W. Thelen, A.K. Wright, and J. Caplan-Auerback, 2008, 2004-2005 seismicity associated with the renewed dome-building eruption of Mount St. Helens. In A Volcano Rekindled: The Renewed Eruption of Mount St. Helens, 2004-2006, editors D. R. Sherrod, W. E. Scott, and P. H. Stauffer. U.S. Geological Survey Professional Paper **1750**.

Power, J.A., and D.J. Lalla, 2010, Seismic observations of Augustine Volcano, 1970–2007. In: Power, J.A., Coombs, M.L., Freymueller, J.T. (Eds.), The 2006 Eruption of Augustine Volcano, Alaska: U. S. Geological Survey Prof. Pap. 1769, pp. 3–40 (Ch. 1).

Shearer, P.M., 1997, Improving local earthquake locations using the L1-norm and waveform cross-correlation: Application to the Whittier Narrows, California, aftershock sequence, J. Geophys. Res., **102**, 8269-8283.

Shearer, P., E. Hauksson, and G. Lin, 2005, Southern California hypocenter relocation with waveform cross-correlation, part 2: Results using source-specific station terms and cluster analysis, Bull. Seismol. Soc. Am., **95**, 904–915.

Thelen, W. A., R.S. Crosson, and K.C. Creager, 2008, Absolute and relative locations of earthquakes at Mount St. Helens, Washington using continuous data: Implications for magmatic processes. In A Volcano Rekindled: The Renewed Eruption of Mount St. Helens, 2004-2006, editors D. R. Sherrod, W. E. Scott, and P. H. Stauffer. U.S. Geological Survey Professional Paper **1750**.

Waite, G.P., B.A. Chouet, and P.B. Dawson, 2008, Eruption dynamics at Mount St. Helens imaged from broadband seismic waveforms: Interaction of the shallow magmatic and hydrothermal systems. J. Geophys. Res., **113**.

Acknowledgments

We thank the staff of the USGS Hawaiian Volcano Observatory for maintaining the seismic network and enabling studies of long-term seismicity variations. Mount St. Helens datasets were collected by the Cascades Volcano Observatory and G. P. Waite, and downloaded via IRIS. This work was partially supported by NSF award EAR-1045035.