JGR Solid Earth

RESEARCH ARTICLE
10.1029/2020JB021294

Indicators of Volcanic Eruptions Revealed by Global M4+ Earthquakes

J. D. Pesicek[1], S. E. Ogburn[1], and S. G. Prejean[2]

[1]U.S. Geological Survey, Volcano Disaster Assistance Program, Vancouver, WA, USA, [2]U.S. Geological Survey, Volcano Disaster Assistance Program, Anchorage, AK, USA

Key Points:

- Eruption forecasting and detection benefits from analyses of global earthquake data, including seismicity unassociated with eruptions
- Analysis of global magnitude four and greater (M4+) earthquakes near volcanoes reveals patterns in seismic activity precursory to eruptions
- Globally, 1% of M4+ earthquakes within 30 km of volcanoes are followed by eruption within 2 weeks

Supporting Information:

- Supporting Information S1

Correspondence to:
J. D. Pesicek,
jpesicek@usgs.gov

Citation:
Pesicek, J. D., Ogburn, S. E., & Prejean, S. G. (2021). Indicators of volcanic eruptions revealed by global M4+ earthquakes. *Journal of Geophysical Research: Solid Earth*, *126*, e2020JB021294. https://doi.org/10.1029/2020JB021294

Received 4 NOV 2020
Accepted 6 FEB 2021

Abstract Determining whether seismicity near volcanoes is due primarily to tectonic or magmatic processes is a challenging but critical endeavor for volcanic eruption forecasting and detection, especially at poorly monitored volcanoes. Global statistics on the occurrence and timing of earthquakes near volcanoes both within and outside of eruptive periods reveal patterns in eruptive seismicity that may improve our ability to discern magmatically driven seismicity from purely tectonic seismicity. In this paper, we catalog magnitude four and greater (M4+) earthquakes near volcanoes globally and compute statistics on their occurrence with respect to various eruptive and volcanic attributes, evaluating their utility as diagnostic indicators of eruptions. Using a 2-week time window and a 30 km radius around the volcanoes, we find that 11% of eruptions are preceded by at least one M4+ earthquake, but only 1% of such earthquakes is followed by eruption. However, earthquakes located 5–15 km from the volcano, those with normal faulting mechanisms and/or large nondouble-couple components, and those occurring as groups are more commonly associated with eruptions, providing significant forecasting utility in some cases. Similarly, certain volcanoes are more likely to exhibit such precursors, such as those with long repose periods. We illustrate the use of these data in eruption forecasting scenarios, including rapid identification of analogous earthquake sequences at other volcanoes. When integrated within the context of multiparametric, multidisciplinary probabilistic assessments of volcanic activity, global earthquake statistics can improve eruption forecasts, and our work provides a model for use on other rapidly expanding global volcanological databases.

Plain Language Summary Magnitude four and larger (M4+) earthquakes in volcanic regions are only rarely followed rapidly by volcanic eruptions. In fact, most such earthquakes are caused by normal tectonic activity in the regions hosting the volcanoes. However, quickly determining the cause of any such earthquake is difficult. In this paper, we compute statistics of M4+ earthquakes near volcanoes globally in order to reveal patterns in the earthquake data that might be useful for improving volcano eruption forecasting and detection. Using a 2-week time window and a 30 km radius around the volcanoes, we find that 11% of eruptions are preceded by at least one M4+ earthquake, but only 1% of such earthquakes is followed by eruption. However, certain kinds of M4+ earthquakes are more likely to be directly caused by magma movement and certain kinds of volcanoes are more likely to experience these earthquakes prior to erupting than others. Our results provide significant forecasting utility in some cases and provide a model for future statistical analyses of other global volcano data sets, which promise to further improve volcano eruption forecasting and detection efforts.

1. Introduction

Increases in seismicity near volcanoes are frequently used to forecast eruptions worldwide. However, not all eruptions are preceded by such increases and some eruptions occur without any detectable precursory changes at all, complicating forecasting efforts. Depending on the characteristics of the volcano and its monitoring network, seismic precursors to explosive eruptions can occur anywhere on the spectrum from too subtle to detect (e.g., Fee et al., 2017) to multiple order of magnitude increases in seismic rate and energy release (e.g., Ruppert et al., 2011). In general, eruptions at frequently active volcanoes are less likely to be preceded by observable precursory changes in seismicity, whereas long-dormant volcanoes typically exhibit more pronounced seismicity changes prior to large eruptions (Cameron et al., 2018; Pesicek et al., 2018; White & McCausland, 2016, 2019). The historical and scientific literature document many cases of precursory behavior prior to explosive eruptions, including some of the largest and deadliest eruptions in history

(Abe, 1992; Harlow et al., 1996; McNutt, 1996; Pliny the Younger, 1835; Stewart, 1820; Verbeek, 1884; Voight, 1990; Yokoyama, 2001; Zobin, 2001).

An ultimate goal and grand challenge for volcano scientists is to provide precise and accurate predictions of future eruptive behavior to decision-makers and to the public. Ideally, predictions could be made using physicochemical models that incorporate observed seismic, geodetic, geochemical or other monitoring data in real time (National Academies of Sciences, Engineering, and Medicine, 2017). Unfortunately, most current deterministic modeling efforts are far from accomplishing this at any practicable level, even at extremely well monitored volcanoes (Poland & Anderson, 2020; Sparks, 2003). Thus, eruptive activity is usually "forecast" rather than predicted and modern forecasting efforts generally rely on probabilistic methods to estimate the likelihood of future volcanic events based on past behavior, commonly employing Bayesian event trees (Marzocchi et al., 2004, 2008; Newhall & Hoblitt, 2002; Newhall & Pallister, 2015).

Eruption forecasts at most of the world's volcanoes use a combination of pattern matching between evolving unrest and documented historic unrest coupled with process interpretation of monitoring data within the context of a conceptual model (Marzocchi & Bebbington, 2012; Sandri et al., 2004; Swanson et al., 1983; Voight et al., 1998). For any particular volcano, the pattern matching uses spatially and temporally limited data sets of analogous past behavior, either from the individual volcano in question or from some combination of analogous volcanoes drawn from various sources (Marzocchi & Bebbington, 2012; Ogburn et al., 2015; Tierz et al., 2019; Wright et al., 2018). Future gains in the field will come from expansion of volcano "big data," to which data mining and statistical techniques can be better applied (Poland & Anderson, 2020, and references therein).

One of the major impediments to using probabilistic approaches for eruption forecasting is that accounts documenting volcanic unrest, such as anomalous seismicity, tend to be "eruption-centric," meaning that the preponderance of published studies present eruptions and their precursory activity. Analysis of the anomalous behavior precursory to eruptions informs us about the processes that lead into eruption and informs our conceptual models of magmatic intrusion and eruption (e.g., Newhall & Pallister, 2015; Roman & Cashman, 2018; Syahbana et al., 2019; White & McCausland, 2019; Wright et al., 2018). However, eruption-centric studies only inform us about unrest that is followed by a known eruption. In the parlance of predictive analytics, these cases represent "True Positives," where unrest (e.g., seismic anomaly) is followed by eruption; or "False Negatives" (Type II errors), where eruptions are not preceded by unrest. Unfortunately, notable seismicity increases (or other changes in monitoring or observational data) not followed by an eruption, referred to as "False Positives" or Type I errors, are underreported in comparison to those preceding eruption (Moran et al., 2011) with few exceptions (e.g., Benoit & McNutt, 1996). This situation results from many innately human factors, including general public and observatory concern about eruptions, publication pressures of scientists, and limited resources of observatory staff. To go beyond the eruption-centric focus on True Positives and more fully assess the predictive value of seismic phenomena requires more complete statistics, including analysis of False Positives and False Negatives (Table 1).

In this study, we investigate the relation between seismicity and volcanic eruptions globally and assess the reliability of earthquakes as predictive indicators of eruptions at various types of volcanoes. In order to obtain truly global statistics, we focus on earthquakes large enough to be detected and cataloged by global monitoring agencies, generally M4+ earthquakes. Although far less common near volcanoes than smaller magnitude events, M4+ earthquakes near volcanoes are alarming to local residents and authorities. In many cases, the potential threat to the public is quite real. Multiple physical mechanisms involving stress and pore-pressure changes explain how ascending magma can trigger such earthquakes. It is also true however, that volcanoes are located in tectonically active areas, and earthquakes near volcanoes are frequently unrelated to magmatic processes directly. In rare circumstances, large tectonic earthquakes have also been invoked as triggering mechanisms for eruptions (Hill et al., 2002; Linde & Sacks, 1998; Walter, 2007). However, the most convincing cases of earthquake-triggered eruptions involved earthquakes of M7 and greater; even in these cases, large tectonic earthquakes are estimated to have triggered only small percentages of eruptions globally in the short term (Manga & Brodsky, 2006; Nishimura, 2017; Sawi & Manga, 2018). Thus, preeruptive earthquakes near a volcano are more likely triggered by magma intrusion than the inverse, and eruption forecasting must rely on more observations than detection of a solitary large earthquake and its aftershocks. Unfortunately, in poorly monitored regions without other corroborating evidence, it is often

Table 1
Abbreviations or Variables Used in Text

Abbrevation	Name	Notes
TP	True Positive	(1) Earthquake(s) followed by eruption within *t* (preeruptive)
		(2) Earthquake(s) during eruption (syneruptive)
FN	False Negative	Eruption lacking TP earthquake(s), also known as Type II error
FP	False Positive	Earthquake(s) lacking eruption within *t*, also known as Type I error
TPR	True Positive Rate = TP/(TP + FN)	Also known as recall or sensitivity
PPV	Positive Predictive Value = TP/(TP + FP)	Also known as precision
M	Magnitude of earthquake	Moment magnitude (Mw), except where otherwise stated
t	Preeruptive time window, or forecast window	2 weeks, except where otherwise stated
R	Radius (km)	Initially 50, then 30, except where otherwise stated
D	Depth (km)	30, except where otherwise stated

Note. Variables are indicated by italics. The values of variables used in the analysis are given in the "Notes" column.

difficult to quickly and confidently establish a causative link between any particular earthquake and magmatic processes at a nearby volcano.

Importantly, we do not expect that M4+ earthquakes by themselves will be sufficient for providing accurate eruption forecasts. In most cases, we expect their occurrence to better inform multidisciplinary probabilistic forecasts (e.g., Wright et al., 2018; Syahbana et al., 2019). Successful eruption forecasting is often based on a combination of changes in rates and characteristics of overall seismicity, including frequency changes of small earthquakes and shifts in earthquake characteristics in the context of changes in degassing, deformation, and other observables (e.g., Buurman et al., 2013; Chouet et al., 1994; Murray, 1992; Power & Lalla, 2010; Power et al., 1994, 1995; Swanson et al., 1983; White & McCausland, 2016, 2019). This work is instead analogous to that of Biggs et al. (2014), and Reath et al. (2019, 2020) who quantified the predictive strength of satellite-detected indicators of unrest at various types of volcanoes and assessed their utility for forecasting. However, anomalous seismicity, owing to nearly ubiquitous seismic networks, is often the earliest observed precursor to eruption, and for many remote and submarine volcanoes, global seismic networks comprise the only monitoring tool available. In this regard, these seismic networks provide the only truly global data set available for the ergodic, pattern-matching exercises that are needed to improve future probabilistic eruption forecasts. Thus, in this paper, we investigate the utility of publicly available earthquake data (Table 2) as a potential eruption-forecasting tool. We search for correlations between eruptive data from the Global Volcanism Program (GVP; Global Volcanism Program, 2013) and M4+ earthquakes in the proximity of the volcanoes that might help distinguish between tectonic and magmatically induced seismicity and we provide a framework for the use of these data in eruption forecasting scenarios, including rapid evaluation of analogous seismic sequences at other volcanoes. These methods are already in use at the U.S. Geological Survey (USGS) by the Volcano Disaster Assistance Program (VDAP) and we expect to apply them the next time there is an earthquake of concern near an active volcano, especially when other monitoring data are limited.

2. Data Selection and Processing

Data used in this study come from a variety of publicly available sources (Table 2). Characteristics of volcanoes (location, composition, morphology, etc.) and eruptions (dates, VEI, repose interval, etc.) from the GVP (GVP, 2013) are used to investigate differences between eruptions, though some categories are combined for simplicity. Although there are over 1,400 volcanoes in the GVP Holocene database, we focus on the 870 of these with confirmed Holocene eruptions, 439 of which have erupted since 1900 (see Supporting Information for other GVP data details). Parametric earthquake data used in this study come primarily

Table 2
Data Sources

Abbreviation	Name	Address	Reference
GVP	Global Volcanism Program	http://volcano.si.edu	GVP (2013)
ISC	International Seismological Centre	http://www.isc.ac.uk	ISC Bulletin (2017)
ComCat	Advanced National Seismic System (ANSS) comprehensive earthquake catalog	https://earthquake.usgs.gov/earthquakes/search/	
GCMT	Global Centroid Moment Tensor	http://www.globalcmt.org	Ekström et al. (2012)
JMA	Japan Meteorological Agency	http://www.data.jma.go.jp	
SSN	Mexican National Seismological Service	http://www2.ssn.unam.mx:8080/catalogo	Pérez-Campos et al. (2018)
SIL	Icelandic Meteorological Office's South Iceland Lowland network	http://hraun.vedur.is/ja/viku	Bodvarsson et al. (1996)
IGN	Spanish Instituto Geografico Nacional	http://www.ign.es/web/ign/portal/sis-catalogo-terremotos	
INGV	Italian Instituto Nazionale di Geofisica e Vulcanologia	http://cnt.rm.ingv.it/en	
GNS	New Zealand Geological and Nuclear Sciences	https://www.geonet.org.nz/data/types/eq_catalogue	
AVO	Alaska Volcano Observatory	https://doi.org/10.3133/sir20195037	Power et al. (2019)

from the International Seismological Centre (ISC) Bulletin (ISC Bulletin, 2017), which provides all reported hypocenters and Moment Tensors by any agency without imposing any filtering or selection criteria. Additional hypocenter data come from the country-specific monitoring agencies listed in Table 2.

For the 870 volcanoes, we combine all available hypocenter data to produce a composite catalog of seismicity for each volcano. Our ultimate goal is to produce a catalog of historical earthquakes for each volcano and rapidly query these catalogs for seismic sequences that are analogous to observed seismicity at restless volcanoes. Toward this goal, we prioritize data quantity in the creation of our volcano catalogs over data quality. That is, we do not initially exclude poorer quality data or nonuniform data. Instead, we assess the effects of seismic catalog quality, nonuniformity, and other potential data issues using sensitivity tests where appropriate, as discussed later and further detailed in the Supporting Information. The under-monitored status typical of most volcanoes and the short history of seismic observations compared to eruptive cycles necessitates this type of approach.

Because individual earthquakes may be recorded by multiple agencies and several included sources also contribute (subsets of) data to the ISC, we remove duplicate events using criteria similar to those used by the USGS's Advanced National Seismic System (ANSS) for creation of the historic ANSS composite earthquake catalog (Table 2). Events with origin time differences of <15 s and whose locations are within 100 km of each other are tagged as duplicate solutions. For duplicates, the reviewed ISC hypocenter is given preference if available, followed by those from country-specific monitoring agencies. To avoid subduction zone earthquakes and other events unrelated to crustal volcanic processes, all earthquakes with depths greater than 30 km are uniformly removed from all catalogs.

When multiple magnitude estimates are available for a given event, the magnitude is chosen hierarchically, with moment magnitude (Mw) being preferred. When Mw is not available, body (mb) and surface (Ms) wave magnitudes are converted to Mw using the relations provided by Lolli et al. (2014). Commonly, only local (ml) and/or duration (Md) magnitudes are available. In such cases, ml is preferred over Md as ml is generally coincident with Mw for events considered in this study (~M4–6). For many events, no details of the magnitude type are provided, and, in these cases, the magnitude is assumed to be ml. We acknowledge that some of the assigned magnitudes using this approach may still not be strictly comparable to Mw. For example, the relations provided by Lolli et al. (2014) are only directly applicable to magnitudes computed by the ISC, while Md requires local calibration and is generally not well suited for larger magnitude events. Previous studies have avoided similar issues by limiting their analyses to uniformly produced catalogs (e.g., Garza-Giron et al., 2018; Pesicek et al., 2018; Shuler et al., 2013a, 2013b; Vidale et al., 2006). Still others apply similar but more limited magnitude conversion strategies (e.g., Holtkamp & Brudzinski, 2011;

Holtkamp et al., 2011). In our preference for data quantity over quality, we retain all available events and show that the effects of nonuniformly computed magnitudes on our results are minimal (Figure S1; Supporting Information).

After preprocessing the earthquake data, individual earthquakes are linked to volcanoes using specific spatiotemporal criteria. For each earthquake, we find all volcanoes within a specific radius (initially 50 km) around the earthquake and link that earthquake to a preferred volcano according to the following criteria. If the earthquake occurs within an eruptive period at a nearby volcano, it is assigned to that eruption and volcano. Otherwise, the earthquake is assigned to the eruption that occurs closest in time following the event, and its associated volcano. Lag times from the earthquake to the assigned eruption are computed for each earthquake. If no eruption has yet occurred at any volcano within the search radius following the event, the event is assigned to the nearest volcano.

After creation of the earthquake catalogs for each volcano, we remove $M < 4$ earthquakes. Applying such a cutoff within the seismic continuum, although without physical basis, makes an otherwise intractable global analysis tractable. Global detection and location of all M4+ earthquakes nominally began in 1996 (ISC Bulletin, 2014). Although this time span is small compared to the repose periods of most volcanoes, its global scope should provide ergodicity. However, this threshold is applied to the exclusion of events whose magnitude uncertainty might span the threshold (see Supporting Information), and to the exclusion of the important magmatic processes that may dominate below the M4 level. Although our current analysis is limited by this cutoff, the choice of threshold is generally flexible and later (Section 9), we discuss the broader use of these methods.

3. Earthquake and Eruption Timing and Categorization

In this study, cases where we have linked M4+ earthquakes to eruptions using specific spatiotemporal parameters are considered to be True Positive (TP) earthquakes; these are the earthquakes traditionally considered as eruption precursors and are used in eruption-centric studies as the basis for forming conceptual models of preeruptive processes. In order to define an earthquake as a TP (Table 1), we must choose a temporal window prior to eruption beyond which an earthquake is no longer considered to be preeruptive. The choice of this window length is nontrivial as run-up times for well-documented eruptions vary by several orders of magnitude (Passarelli & Brodsky, 2012); thus, the definition of this window length will inevitably exclude some precursory events at some volcanoes and include some background seismicity at others. In Figure 1, we illustrate the distribution of observed lag times of M4+ earthquakes in relation to eruptions. We limit the plot to a maximum radius of 50 km from the GVP defined volcano coordinates and a maximum precursory window of 6 years White and McCausland (2016) showed that precursory TP seismicity can occur out to 40 km or more from the volcano in some cases, and up to 6 years prior to the eruption, in agreement with the maximum run-up time documented by Passarelli and Brodsky (2012). For these years, increasing numbers of events occur in the weeks and months prior to eruption, with a definitive peak in seismicity immediately preceding eruption (Figure 1a).

In order to minimize the inclusion of noneruptive events in our set of TPs, we choose a rather short 2-week preeruptive time window ($t = 14$ days) for analysis, corresponding to the peak in global M4+ rates prior to eruption (Figure 1a). A 2-week time frame is a common operational forecasting time window (e.g., Wright et al., 2018) and in prior work, this choice optimized the identification of preeruptive seismic-rate anomalies (Pesicek et al., 2018). Later, we explore the effects of time-window length on the results and the potential run-up times for preeruptive seismicity (Section 6). Using the 2-week, preeruptive window, we divide the earthquakes into four sets for comparison: (1) all earthquakes assigned to a volcano (volcanic); (2) those earthquakes not occurring during an eruption nor within the preeruptive time window (noneruptive); (3) those earthquakes occurring within the time window ($t = 14$ days) prior to the eruption (preeruptive); and (4) those earthquakes occurring within defined eruption times (syneruptive; Figure 1b). Sets (2)–(4) are nonoverlapping subsets of (1). Although uncertainty in the timing of these events can affect the categorization of specific earthquakes, in the Supporting Information we show that GVP eruption time uncertainties (generally much larger than earthquake origin time uncertainties) do not greatly affect the results

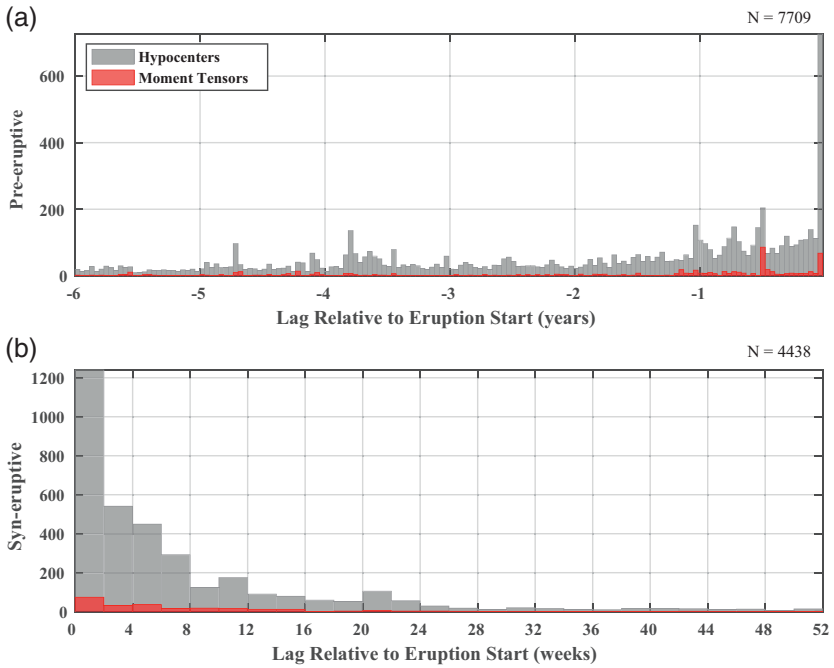


Figure 1. Histogram (2-week bins) of 7,709 M4+ earthquakes (gray) and available Moment Tensor (MT) solutions (blue) occurring within 50 km of the volcanoes and with depths <30 km. Lag times are relative to the eruption start times for (a) the 6 years prior to the eruptions (preeruptive), and (b) those occurring during the eruptions (syneruptive). Event counts peak in both plots in the 2-week bin closest to the eruption start time. Note that we categorize all 4,438 events occurring between eruption start and end times as syneruptive, even for eruptions that last longer than 1 year. However, only syneruptive events with lags <1 year are shown in (b).

(Figures S2 and S3). Thus, in the following sections, we assess differences in the statistical features of these sets and interpret robust features in terms of volcanic and eruptive processes.

4. True Positives: M4+ Earthquakes Preceding Eruptions

4.1. Epicenters

For the four defined sets of earthquakes, we first assess their locations relative to their assigned volcano. Figure 2 shows the distribution of earthquake-to-volcano map distances and focal depths for the volcanic, noneruptive, preeruptive, and syneruptive groups, plus a fifth set composed of an identically processed representative sample of global seismicity for comparison, which contains all events housed at the ISC for the years 2010–2015, regardless of their association with a volcano. In Figure 2, the shape of the distribution of M4+ hypocenters as a function of distance from volcano (left) and depth (right) appears quite different in eruptive subsets (d), (e), compared to the parent volcanic set (b). To formally quantify the statistical significance of this difference, we apply a Monte Carlo bootstrap resampling test. For the population of N earthquakes in a given subset, we create 1,000 random samples, each of size N, drawn from the larger parent catalog of volcanic earthquakes and show the 10th and 90th percentiles of variations in these distributions to illustrate the level of random variability expected solely due to subsampling of the data. When variations in the distributions lie well outside the ranges defined by the 10th and 90th percentiles, we interpret these as statistically significant variations.

The earthquake-to-volcano distance distributions of the preeruptive and syneruptive events show differences outside those expected due to random variability (Figure 2). The global volcanic and noneruptive sets generally show increasing event counts with increasing distance from the volcano, as would be expected for uniformly distributed seismicity where larger distances correspond to larger areas sampled. However, the distribution of preeruptive M4+ earthquakes shows that the peak occurs distal to the eruptive center, from 5 to 15 km from the vent, and that the number of preeruptive events drops significantly for distances >30 km.

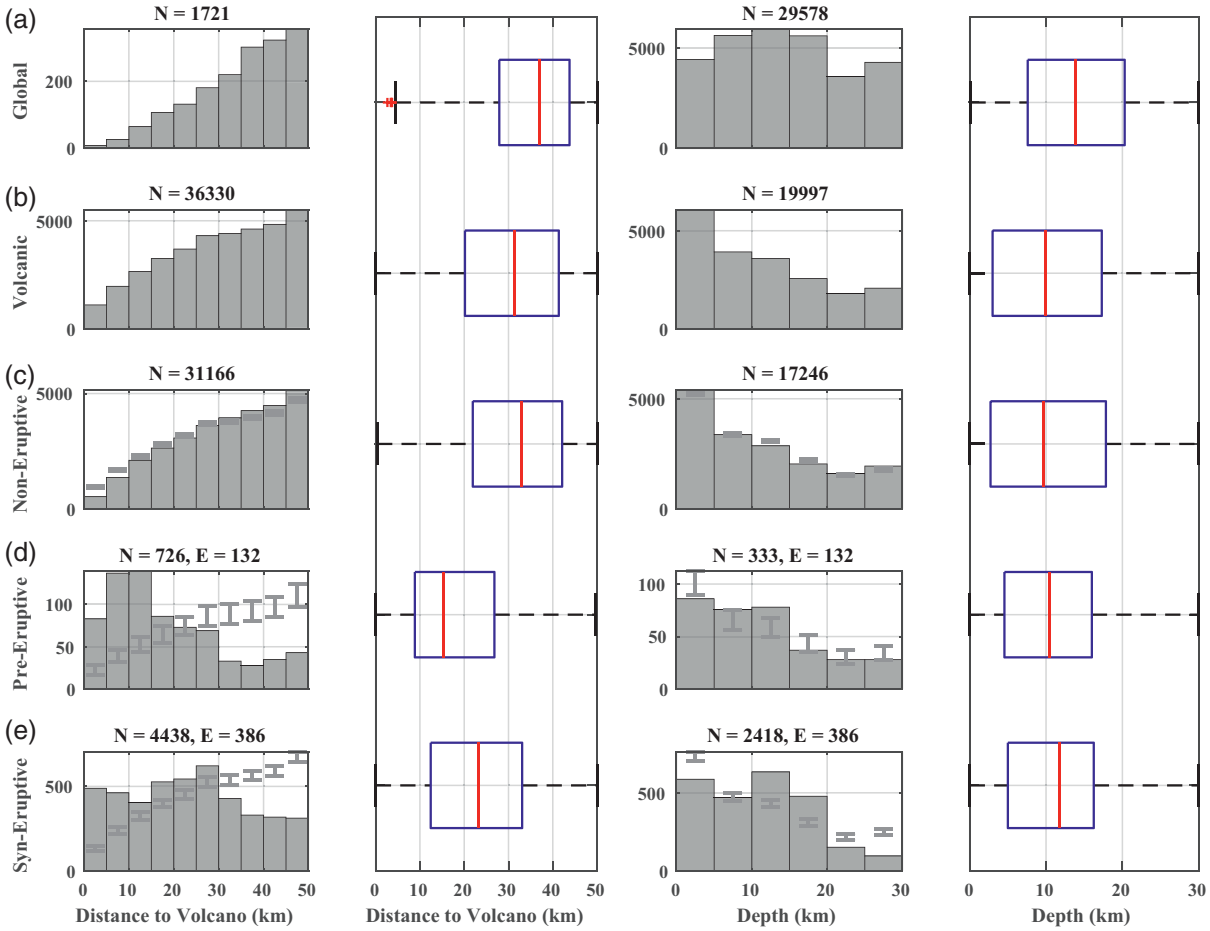


Figure 2. Distribution and statistical properties of M4+ hypocenters within 50 km of active volcanoes for five sets of earthquakes, as described in text: (a) a global reference set of ~70,000 events for the years 2010–2015, 1,721 of which are within 50 km of the volcanoes (b) all events assigned to a volcano (volcanic), (c) those events not occurring during an eruption or within the (t = 14 days) reference preeruptive time window (noneruptive), (d) those events occurring within the t preeruptive time window (preeruptive), and (e) those events occurring within defined eruption times (syneruptive). (c)–(e) are nonoverlapping subsets of (b). Dark gray bars in (c)–(e) show the 10th and 90th percentiles of variations expected due to random subsampling of (b) the volcanic distribution, as determined from the bootstrap analysis. Boxplots illustrate the medians (red), the 25th and 75th percentiles (blue), and data range (black whiskers) excluding outliers (red +) beyond 2.7σ (~99.3%), if any. Earthquakes with fixed depths are removed from the depth (below sea level) plots (right) for all sets, accounting for the smaller numbers (N) in (b)–(e). For the global reference set (a), depths are shown for all global events, not just those within 50 kms of the volcanoes, accounting for its larger N. The number of eruptions (e) represented in the preeruptive and syneruptive sets is also listed.

The syneruptive earthquakes are more widely distributed, with peaks closest to the volcano (0–5 km) and at 25–30 km. Neither of these plots shows eruptive M4+ seismicity to be mainly centered directly beneath the GVP volcano locations. Although we cannot explicitly assess location uncertainty, normally distributed location errors are not expected to produce such patterns and known biases in teleseismic hypocenters (due to unmodeled subducting slabs in the reference velocity models (e.g., Syracuse & Abers, 2009)) cannot explain global patterns for all volcano types whose hypocenters are frequently constrained with local seismic data.

These epicentral observations agree well with existing models of precursory seismic energy release being dominated by seismicity on nearby crustal faults (Coulon et al., 2017; White & McCausland, 2016, 2019), at least for M4+ earthquakes. The results support interpretations that precursory seismicity can reflect pressurization occurring over a broad region around the volcano, possibly triggering seismicity on nearby faults that are already stressed and poised to fail. For such triggered M4+ earthquakes, the results provide constraints on the distal extent of magmatic influence on seismicity in volcanic regions and emphasize the need to consider seismicity out to at least 30 km from the active vent for eruption forecasting purposes, as described by White and McCausland (2016). However, in some cases, seismicity out to five or more km from the GVP location may still be proximal to an eruption site. Thus, some distal M4+ seismicity in Figure 2 may

be related to conduit opening processes whereby shear faulting and tensile fracturing above intruding magma directly induces seismicity along favorably oriented faults. Unfortunately, in this work, we are not able to assess the effects of variable volcanic and magmatic system size and geometry, as our results reference a single GVP location for each volcano and a constant radius.

4.2. Focal Depths

Assessing variations in the earthquake depths is more difficult, due to higher uncertainties inherent in the determination of focal depth. In fact, many of the cataloged hypocenters are fixed at 0-km or 10-km depth by locating agencies (Bondár & Storchak, 2011). After removing these fixed-depth solutions from all sets, we see that the distribution of volcanic focal depths differs considerably from that of the global reference set (Figure 2). The global set has more uniformly distributed depths with a peak at 10–15 km while the volcanic seismicity is generally shallower. Depths of the preeruptive subset have fewer of the shallowest events but are otherwise similar to the volcanic set. The syneruptive subset is depleted in the shallowest and deepest events and enriched in events with depths between 10 and 20 km.

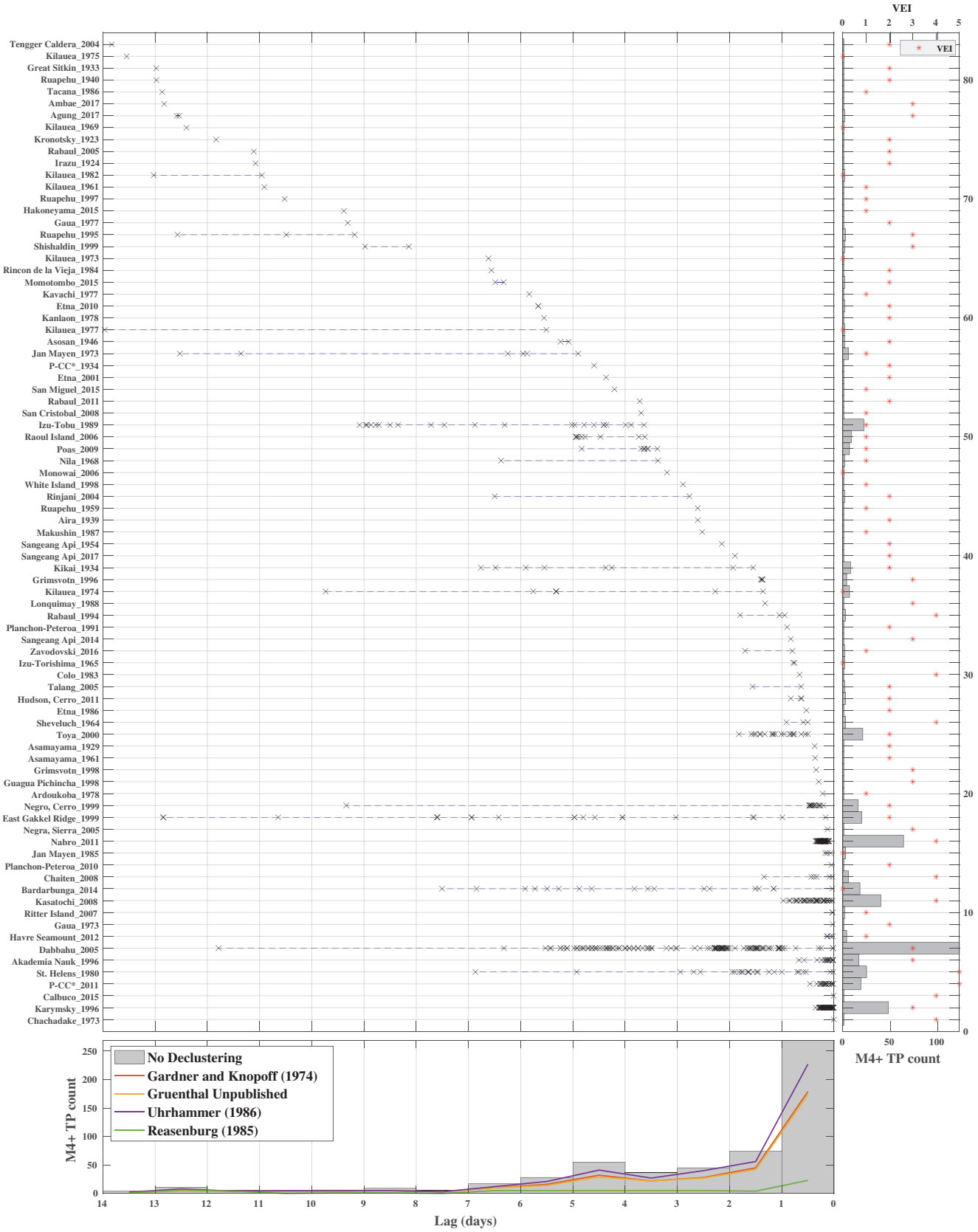
The depletion of shallow M4+ syneruptive seismicity may reflect the relaxation of shallow stresses once the pathway for magma and gas expulsion has been established (e.g., Scandone & Malone, 1985; Thelen et al., 2010). Although this appears true for M4+ earthquakes, we do not necessarily expect this to apply for smaller magnitude events, which often continue throughout syneruptive periods. The relative enrichment of deeper events in the syneruptive set may be explained by stress relaxation processes associated with magma reservoir depletion (e.g., Buurman et al., 2013; Moran, 1994; Mori et al., 1996; Power & Lalla, 2010; Power et al., 1994), and/or magma recharge events (e.g., Linde et al., 2016; Moran, 1994).

4.3. Aftershock and Outlier Sequences

We next assess properties of the frequency-magnitude distributions of volcanic earthquakes. Based on the sharp dropoff in earthquakes in Figure 2 beyond 30 km from the volcano, we now limit the analysis to events within a 30 km radius in order to further exclude noneruptive events. Figure 3 shows the 83 TP eruptions where at least one M4+ earthquake preceded the eruption within 2 weeks, along with the timing of M4+ events within those 2 weeks for each eruption. The vast majority of these earthquakes (~95%) occur within 7 days of the eruption, with increasing numbers closer to eruption. In fact, the median lag for the 83 first occurring TP M4+ earthquakes is 4.2 days, indicating the general rapidity of eruption following such earthquakes.

Figure 3 also shows that several individual eruptions have relatively high numbers of M4+ preeruptive earthquakes (e.g., Dabbahu volcano, 2005). This raises the possibility that one or a few particularly vigorous seismic sequences or eruptions are dominating the global statistics, potentially biasing the observed features. To investigate this possibility, we first consider whether the high numbers of preeruptive M4+ earthquakes for some eruptions result from aftershocks of a few particularly large earthquakes. Declustering the catalogs to remove aftershocks (see Supporting Information) reduces the numbers of included earthquakes (Figure 3) as expected but does not remove the relative temporal and spatial patterns seen in Figures 1–3 and others discussed later. It is still possible however that one or more eruptions with extremely vigorous M4+ seismicity are biasing the overall preeruptive and syneruptive group statistics. To examine this, we identify and remove outlier eruptions with extreme numbers of M4+ earthquakes and reexamine the statistical features of the remaining eruptions. The results (Figure S4) show that these outlier eruptions have a notable effect on the distribution of syneruptive earthquake locations, but not the preeruptive locations. The peak in syneruptive seismicity at 25–30-km distance (Figure 2) is not observed in the filtered data set, as many of outliers are related to dike intrusions distal to the GVP volcano location (see Supplementary Information). However, the syneruptive locations remain enriched in earthquakes located proximal to the volcanoes compared to the noneruptive group.

We also investigate the aftershock productivity of the M4+ earthquakes. Anomalous *b*-values for swarms of events are often reported at volcanoes and are typically interpreted as being caused by high pore pressures, low applied stresses, or heterogeneous materials (e.g., McNutt & Roman, 2015). Studies of temporal and/or spatial variations in *b*-values typically involve seismic activity well below the M4 level at densely monitored


VEI
0 1 2 3 4 5
VEI
Tengger Caldera_2004
Kilauea_1975
Great Sitkin_1933
Ruapehu_1940
Tacana_1986
Ambae_2017
Agung_2017
Kilauea_1969
Kronotsky_1923
Rabaul_2005
Irazu_1924
Kilauea_1982
Kilauea_1961
Ruapehu_1997
Hakoneyama_2015
Gaua_1977
Ruapehu_1995
Shishaldin_1999
Kilauea_1973
Rincon de la Vieja_1984
Momotombo_2015
Kavachi_1977
Etna_2010
Kanlaon_1978
Kilauea_1977
Asosan_1946
Jan Mayen_1973
P-CC*_1934
Etna_2001
San Miguel_2015
Rabaul_2011
San Cristobal_2008
Izu-Tobu_1989
Raoul Island_2006
Poas_2009
Nila_1968
Monowai_2006
White Island_1998
Rinjani_2004
Ruapehu_1959
Aira_1939
Makushin_1987
Sangeang Api_1954
Sangeang Api_2017
Kikai_1934
Grimsvotn_1996
Kilauea_1974
Lonquimay_1988
Rabaul_1994
Planchon-Peteroa_1991
Sangeang Api_2014
Zavodovski_2016
Izu-Torishima_1965
Colo_1983
Talang_2005
Hudson, Cerro_2011
Etna_1986
Sheveluch_1964
Toya_2000
Asamayama_1929
Asamayama_1961
Grimsvotn_1998
Guagua Pichincha_1998
Ardoukoba_1978
Negro, Cerro_1999
East Gakkel Ridge_1999
Negra, Sierra_2005
Nabro_2011
Jan Mayen_1985
Planchon-Peteroa_2010
Chaiten_2008
Bardarbunga_2014
Kasatochi_2008
Ritter Island_2007
Gaua_1973
Havre Seamount_2012
Dabbahu_2005
Akademia Nauk_1996
St. Helens_1980
P-CC*_2011
Calbuco_2015
Karymsky_1996
Chachadake_1973
0 50 100
M4+ TP count
No Declustering
Gardner and Knopoff (1974)
Gruenthal Unpublished
Uhrhammer (1986)
Reasenburg (1985)
M4+ TP count
Lag (days)
14 13 12 11 10 9 8 7 6 5 4 3 2 1 0

volcanoes (e.g., Farrell et al., 2009). In contrast, our global analysis of M4+ events is not capable of deciphering such detailed changes but may instead inform us about the deviations in aftershock productivity near volcanoes, above the M4 level. In general, we find that aftershock productivity for M4+ events in volcanic areas to be broadly similar to nonvolcanic areas, in agreement with other recent findings (Garza-Giron et al., 2018), and overall the *b*-values for the M4+ events do not differ significantly in preeruptive or syneruptive periods compared to others (Figure S5). Thus, we find no deviations in the frequency-magnitude distributions that might help to distinguish M4+ eruptive seismicity from tectonic seismicity.

4.4. Moment Tensors

Next, we explore aspects of the available source mechanisms in volcanic areas. Magmatic activity is often invoked to explain earthquake Moment Tensors (MTs) that deviate from pure double-couple (DC) shear slip (e.g., Miller et al., 1998; Shuler et al., 2013a) and are thought to result from a variety of mechanisms, including shear failure on volcanic ring faults, tensile faulting, and fluid transport (Julian et al., 1998; Shuler et al., 2013b). However, some percentage of the non-DC component for any single event is often attributed to data noise and/or modeling inaccuracies. As a result, it is often difficult to attribute non-DC components in volcanic areas solely to source process, for both the volumetric and Compensated Linear Vector Dipole (CLVD) components (e.g., Panza & Saraò, 2000; Pesicek et al., 2012). Furthermore, individual reporting agencies use different methods to compute their solutions and attempt to limit such effects in various ways. For example, the largest source of MT solutions comes from the Global Centroid Moment Tensor (GCMT) project (Table 2), whose solutions are constrained to be deviatoric. Unfortunately, this constraint limits our ability to directly investigate the volumetric components that are commonly attributed to magmatic processes (e.g., Julian et al., 1998). However, there is a known tradeoff between the volumetric component and vertical CLVD components for GCMTs determined by surface wave modeling and it is impossible to constrain the relative contributions of each (Shuler et al., 2013b). Thus, although we cannot expect to isolate the volumetric component, investigating the total non-DC component of the MTs likely includes many volcanic volumetric source processes in addition to CLVD source processes.

For the MT analysis, we limit the volcanic catalog to those MTs within 30 km of the volcanoes in order to better distinguish volcanic from nonvolcanic events. This cutoff is based on the sharp falloff observed in the preeruptive events in Figure 2 at 30 km, but also on the analysis of Shuler et al. (2013a), who found that ~90% of the GCMTs they linked to volcanic unrest fell within 30 km of their respective volcanoes. For the MTs, we investigate several aspects of the solutions and assess the non-DC components in two ways, (1) by overall non-DC percentage in the solution, which is regularly reported by the USGS for their rapid MT solutions (Benz, 2017) and (2) by the parameter$\in$, which is commonly reported in studies of non-DC MTs (e.g., Julian et al., 1998). The parameter $\in$ provides the relative proportions of DC vs. deviatoric components in terms of the eigenvectors of the MT. Pure DC events have $\in = 0$ while pure CLVD events have $\in = +/-0.5$. Positive (negative) $\in$ corresponds to extensional (compressional) polarity of the major dipole of the CLVD component. Finally, we also assess the plunges of the principal axes of the solutions and categorize the MTs by faulting type, using the method of Kagan (2005; as implemented by Álvarez-Gómez, 2014). Figures 4 and 5 illustrate the variations in these values for all MTs in each of the four volcanic populations and the global population along with the results of the bootstrap test to validate the significance of the variations.

The results show that MTs computed in volcanic areas tend to have larger non-DC components and higher numbers of strike-slip and normal faulting mechanisms compared to the overall global population. The number of preeruptive MTs is small ($N = 60$), but the set is statistically distinguishable from the larger volcanic population. The preeruptive events are more commonly normal faulting (Figure 5) and have non-DC components that are statistically significant at the extrema of the parameter ranges (Figure 4). Syneruptive MTs are more numerous ($N = 213$), are dominated by normal faulting mechanisms, and contain higher

Figure 3. M4+ earthquakes (black x) for all 83 True Positive (TP) eruptions, defined as those eruptions that were preceded by at least one M4+ earthquake within 30 km of the volcano and within the $t = 14$ days prior to the eruption. Eruptions are sorted by minimum lag time and labeled by their GVP name (except for P-CC*, which abbreviates Puyehue-Cordon Caulle) and year of eruption. The histograms show the overall count of M4+ earthquakes per eruption and red dots show eruption VEI (right) and counts of M4+ earthquakes per day within the 14-days window (bottom). Colored lines indicate counts of M4+ earthquakes per day for four different declustering methods as discussed in the Supporting Information. VEI, volcanic explosivity index.

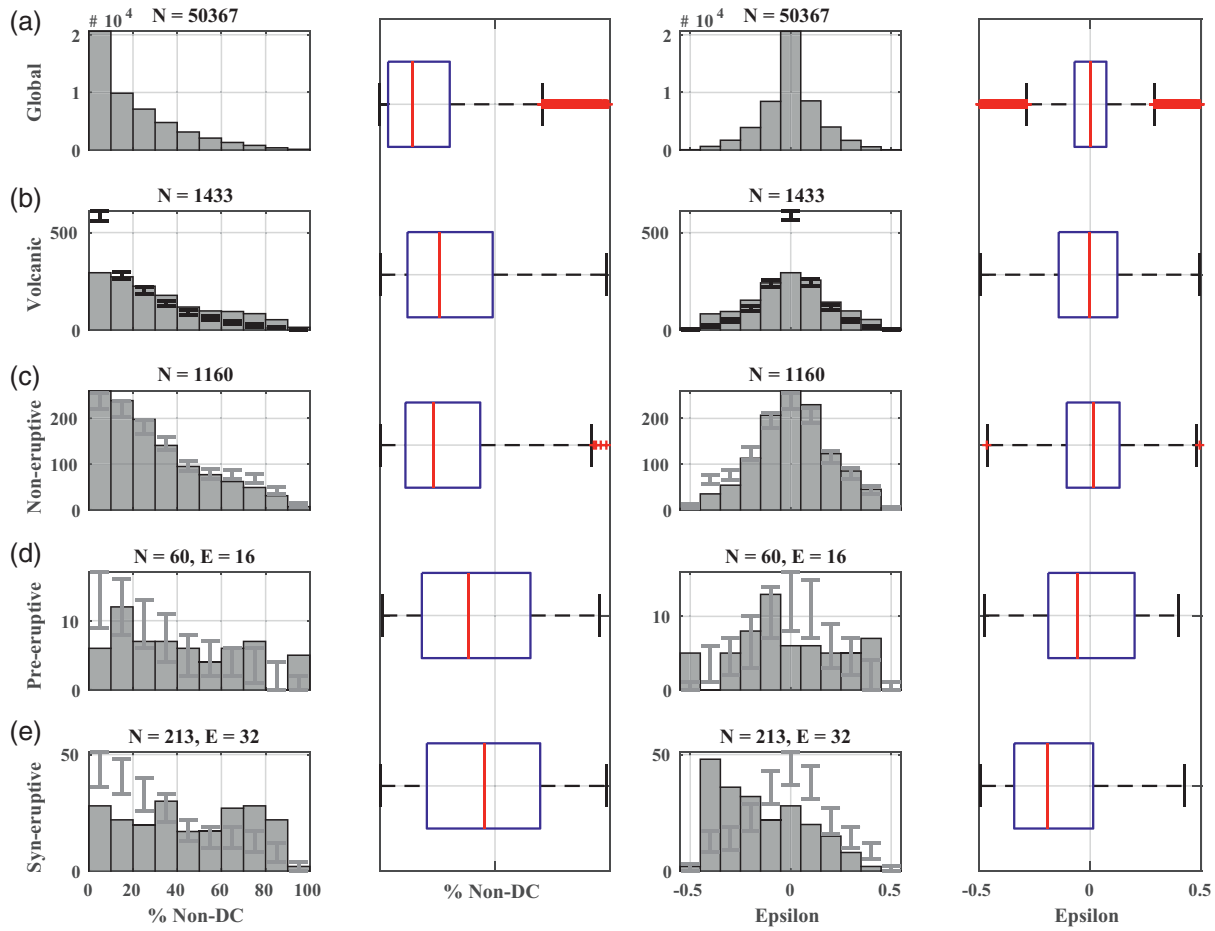


Figure 4. Distribution and statistical properties of the nondouble-couple (non-DC) components of the MT solutions (depths ≤ 30 km). Overall percent of nondouble-couple component (left) and the parameter ∈ (right) are computed for (a) all global crustal MTs, (b) the set within 30 km of active volcanoes, (c) the noneruptive set, (d) the preeruptive (t = 14 days) set, and (e), the syneruptive set. (c)–(e) are nonoverlapping subsets of (b). Black (gray) bars show the 10th and 90th percentiles of values expected due to random subsampling of the global (volcanic) catalog, as determined from a bootstrap analysis. See Figure 2 for additional details. MT, Moment Tensor.

non-DC components than others over a broader range. Together, the results suggest that MTs with large (>60%;∈ ≤ -0.3) non-DC components are the most likely to be linked to eruptive processes. Finally, we note that these observations do not change significantly when we limit the MT catalog to those uniformly computed by the GCMT (Figures S6 and S7).

For the non-DC components, we compare our results to those of Shuler et al. (2013a), who performed a detailed analysis of vertical CLVD type events globally near volcanoes. They recomputed GCMT solutions for 313 volcanic earthquakes with large non-DC components and assessed their timing in relation to volcanic unrest. They suggest that vertical-T (∈ > 0.20; T-axis plunge > 60°) CLVD earthquakes generally occur before the start of volcanic unrest while vertical-P (∈ < −0.20; P-axis plunge > 60°) CLVD earthquakes occur after the start of volcanic unrest. Though the unrest time windows examined by Shuler et al. (2013a) differ from our uniformly defined preeruptive time windows, we find general agreement with their results in that our syneruptive set contains large negative ∈ values (Figure 4) and large P-axis plunges (Figure 5). These events generally correspond to their vertical-P events, which tend to occur after their defined unrest/eruption start times.

The presence of large non-DC components in many, but not all of the preeruptive and syneruptive earthquakes likely reflects several influencing factors. For the preeruptive events, the relative dearth of large non-DC components is well explained by precursory seismicity triggering nearby faults that were previously tectonically loaded. These events might be expected to be dominantly shear and comprise the bulk

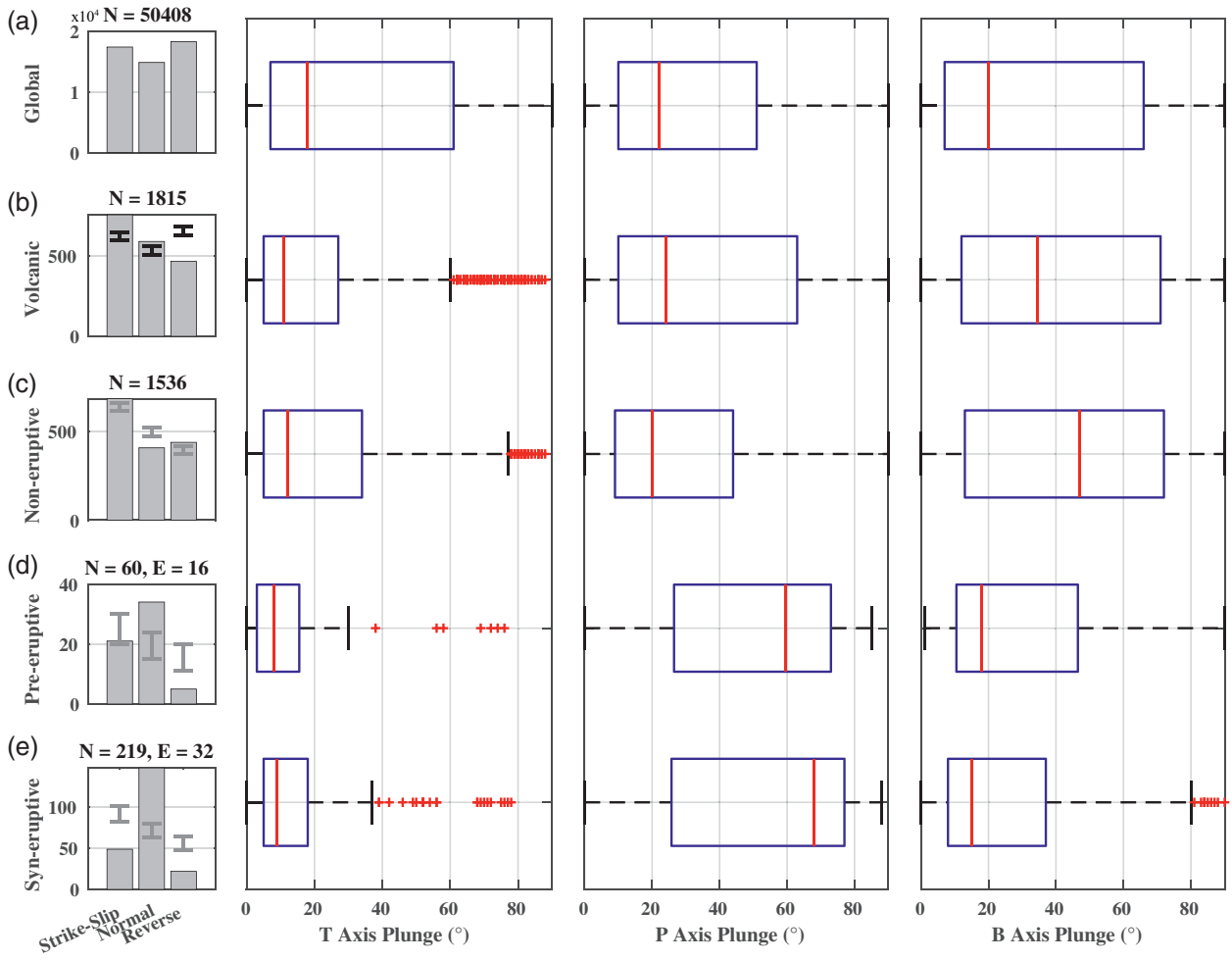


Figure 5. Fault type, distribution, and statistical properties of the plunges of the principal axes of the Moment Tensor (MT) solutions for (a) all crustal (depth < 30 km) MT solutions, (b) the subset within 30 km of the active volcanoes, (c) to the subset of preeruptive (t = 14 days) events, and (d) the subset of syneruptive TPs. Black (gray) bars show the 10th and 90th percentiles of values expected due to random subsampling of the global (volcanic) catalog, similar to Figures 2 and 5. See Figure 2 for additional details.

of moment release in the days to weeks preceding eruptions. Earthquakes closest to the eruption onset and immediately following eruption are more likely to directly reflect the processes of magma migration and eruption and may exhibit larger volumetric or CLVD source processes. Those occurring later in the eruption might reflect magma evacuation and collapse processes with complex sources that result in higher non-DC components, such as the ring faulting favored by Shuler et al. (2013b). The dominance of normal faulting events associated with eruptions may reflect several factors. Volcanoes commonly exist in local extensional or transtensional stress regimes, even where located in large scale compressive subduction zones. As a result, intruding magma may preferentially trigger preexisting normal faults. Vertical diffusion of fluids (aqueous, gaseous, and magmatic) is enhanced in areas of local extension where the least-compressive principal stress is in the horizontal plane. Fracture meshes involving vertical and subvertical faulting are likely common in volcanic areas (Hill, 1977; Shelly et al., 2015). In addition, normal faulting may be directly induced above inflating dikes (e.g., Belachew et al., 2013; Passarelli et al., 2015). Finally, preferential normal faulting may also be explained by motion along dipping faults beneath the volcanic edifice in response to stress adjustments around perturbed magmatic storage zones (e.g., Ekström, 1994; Saunders, 2000).

5. False Negatives: Eruptions Lacking M4+ Earthquake Precursors

A primary challenge in assessing the relation between earthquakes and eruptions comes not from finding when earthquakes preceded eruptions (TPs), but rather from determining which eruptions were not preceded by earthquakes. These False Negative (FN) cases are critical for evaluating the True Positive Rate (TPR; Table 1) of any predictive indicator, as well as the performance of our conceptual models of eruption. To find the FNs and compute the TPR, we must determine when earthquakes *could have been detected* in association with eruptions but weren't. In other words, we must find those eruptions that had a sufficient magnitude of completeness (M_c) to confirm a lack of earthquakes. The M_c of the reviewed ISC data is known to vary regionally (Woessner & Wiemer, 2005) and it may not be appropriate to assume that volcanoes in high-latitude regions, where instrument coverage is sparse, are as well monitored as those in more densely instrumented areas. Thus, for this work, we individually estimate the M_c of each eruption using all magnitude earthquakes for each volcano.

We seek M_c estimates that accurately reflect the state of monitoring of each eruption, where conditions range from monitoring with a dense local seismic network to having only data from global networks. To this end, we adopted a fast, reliable, robust, and commonly used catalog-based technique that relies on the maximum curvature of the frequency-magnitude distribution, which is assumed to approximate a power-law (Mignan, 2012; Wiemer & Wyss, 2000). Furthermore, we choose a relatively high minimum event (N_{min}) threshold (200 events) to ensure accurate estimates of M_c (Woessner & Wiemer, 2005). Finally, as maximum curvature methods are known to underestimate M_c in some cases, we apply a correction factor of 0.2 to our M_c threshold (e.g., Llenos & Michael, 2019; Woessner & Wiemer, 2005) and only use eruptions where the M_c is estimated to be 3.8 or less.

We estimate the M_c of each eruption using all earthquakes within a cylindrical volume around the volcano specified by a radius (R) and maximum depth (D) occurring in the 6 years preceding the eruption. The relatively large time window over which we compute M_c further ensures conservative M_c estimates and later (Section 6) allows us to assess variations in the t window up to the 6-year maximum run-up time (Section 3; Figure 1). We first attempt to estimate M_c for an eruption using those events located closest to the volcano ($R = 30$ km, $D = 30$ km). In cases where the N_{min} criterion is not met, we progressively expand R up to 50 km until N_{min} events are found. In cases where N_{min} is still not met, we instead use only reviewed ISC solutions to estimate M_c. In these cases, we use $D = 70$ km and a maximum radius (R) of up to 200 km from the volcano. If, after following this procedure for a given eruption, we still have not met the N_{min} criterion, we assume that the M_c is too large, and the eruption is excluded from consideration as a FN.

Using this strategy, we attempt to alleviate issues related to the composite character of the earthquake data while retaining as many eruptions for analysis as possible. However, our conservative approach means we are likely underestimating the true eruption FN count. However, we err on the side of excluding some sufficiently monitored eruptions rather than including insufficiently monitored eruptions, in order to ensure that any lack of precursory earthquakes ("evidence of absence") is not due to insufficient monitoring ("absence of evidence"). For $t = 14$ days, we find 645 FN eruptions where seismic monitoring was sufficient to detect all M4+ earthquakes based on our analysis of M_c, but where none were recorded prior to the eruption (Figure 6), including several well monitored VEI four eruptions (e.g., Budi-Santoso et al., 2013; Johnson et al., 2010; Power et al., 1995; Sigmundsson et al., 2010) and many others where no local earthquake data are available (e.g., those with $M_c \sim 3.5$–3.8 from 2000 on).

6. Eruption True Positive Rate

Having now categorized the eruptions as either TP or FN, we compute the TPR (Table 1). The results (Figure 7) illustrate the populations of eruptions where one or more M4+ earthquakes were detected within 14 days of eruption (TP) or where they could have been detected based on our analysis of M_c (Figure 6) but weren't (FN). The overall TPR for these eruptions indicates that 11% of sufficiently monitored eruptions globally were preceded by at least one M4+ earthquake. In addition, the results show significant variations in TPR based on different volcano and eruption attributes derived from the GVP data. Variations in the TPR for these subsets range from <10% to 100%, although many of the subset variations cannot be distinguished from random subsampling of the data, as illustrated by the bootstrap results (Figure 7).

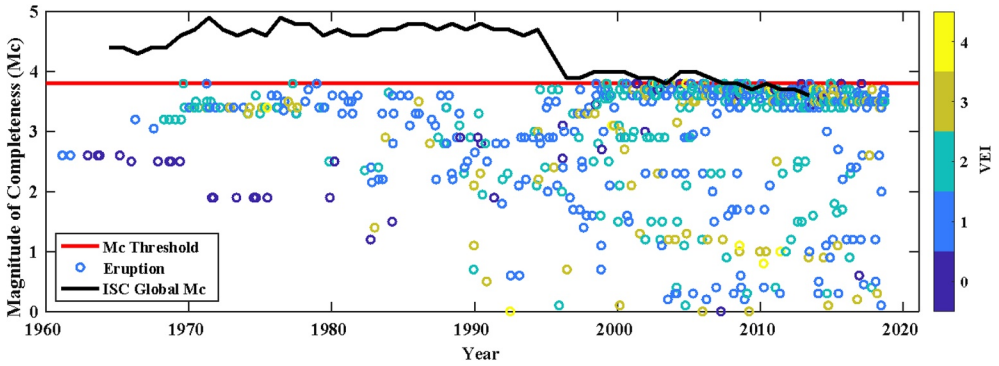


Figure 6. Plot showing computed magnitude of completeness (M_c) vs. eruption date for 645 False Negative (FN) eruptions. The black line shows the average M_c from the reviewed global ISC catalog (ISC Bulletin, 2017) while the red line shows the $M_c = 3.8$ threshold used to assign FN eruptions. The M_c is estimated over the 6-years preeruptive period for each eruption.

Variations in preeruptive TPR due to volcano or eruption attributes may inform us about where and under what circumstances we are more likely to observe large magnitude earthquake eruption precursors. However, care must be taken in interpreting these variations, as the groups defined by the GVP categories comprise overlapping subsets. Furthermore, the variations themselves depend critically on the choice of t. It is important to note that the TPR value necessarily increases with larger t windows, reflecting the recategorization of more events from noneruptive to preeruptive. For longer t windows, TPR values may capture more truly precursory earthquakes, but are also more affected by tectonic seismicity rates. Shorter windows reflect seismicity temporally closer to eruptions whose source processes are more likely to be magmatic but may not capture the complete seismic run-up sequence. Thus, although TPR always increases with t, the variations in this increase may provide insight into both eruption run-up time variations and background M4+ seismicity rates in different volcanic environments (Figure 8). Overall, the TPR vs. t curves resemble asymptotic functions whose y-axis (TPR) asymptote reflects the return to the background, long-term rate of M4+ events, while the different shapes of the curves away from the asymptote (toward $t = 0$) reflects their differences in run-up times before eruption.

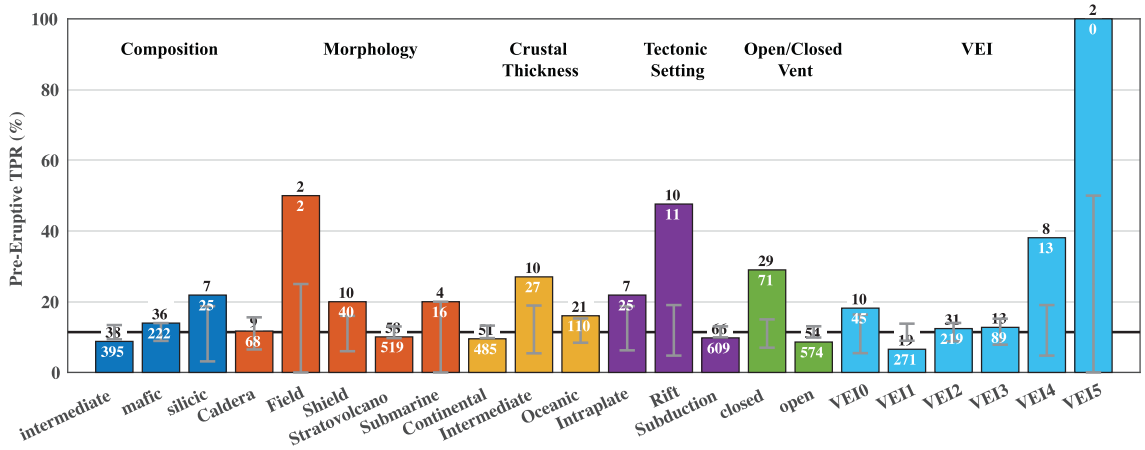


Figure 7. True Positive Rate (TPR; Table 1) comparing all eruptions that have at least one M4+ event within $t = 14$ days prior to eruption (TPs) to those eruptions (FNs) where none were recorded, but monitoring was sufficient to record them ($M_c \leq 3.8$). The TP (FN) count for each subcategory is shown at the top of each bar in black (white). Only subcategories with at least one TP are shown. Bars are colored by attribute and include chemistry (blue), morphology (red), crustal type (yellow), tectonic setting (purple), vent openness (green) at the time of the eruption, and VEI (cyan). Volcanic vents are defined as "open" if the eruption occurred within 15 years of the previous eruption. The overall TPR is 12% and is shown in black. The gray vertical bars show the 10th and 90th percentiles of the results of the bootstrap test for each subcategory, as discussed in the text. VEI, volcanic explosivity index.

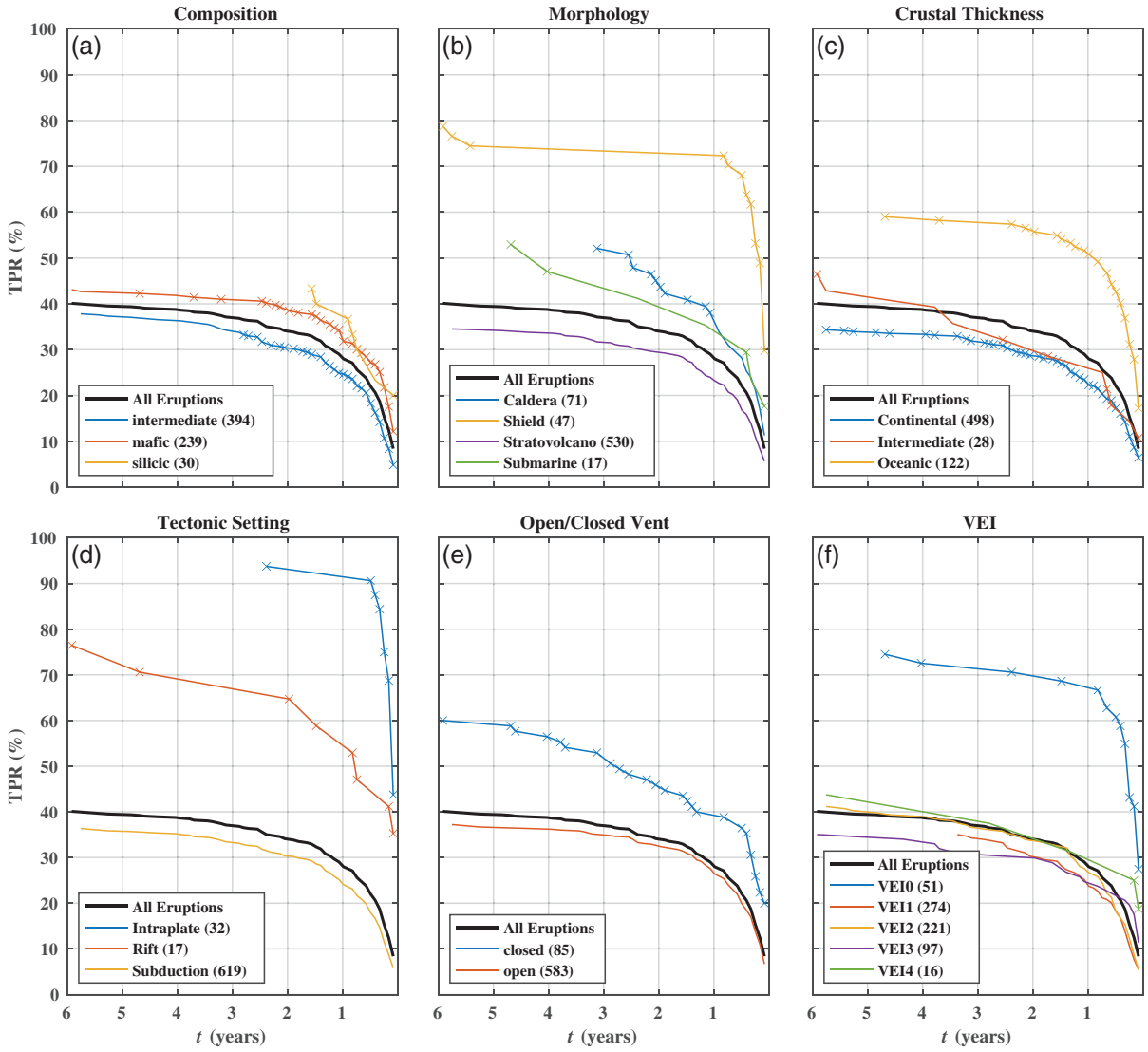


Figure 8. True Positive Rate (TPR; Table 1) as a function of preeruptive window size (t) and volcano attribute for the 668 eruptions with a sufficient M_c over 6 years prior to the eruption. The TPR is computed every 30 days and plotted when an increase in TPR occurs. t windows where the TPR falls outside of the 10th and 90th percentiles of the bootstrap tests are shown as "x" on top of the solid lines, indicating statistically significant deviations from the overall TPR (black). The number of eruptions for each attribute is given in the legend and those with fewer than 10 eruptions are omitted. Note that eruptions where one or more TP earthquakes were detected, but where the eruption M_c is >3.8 are excluded here (cf. Figure 7), in order to keep the number of eruptions constant with respect to t. See Figure 7 for additional details.

In many cases, a particular attribute is not statistically distinguishable from the overall trend, and this gives us insight into the dominance of that attribute in the overall eruption population and TPR. For example, the TPR vs. t curve for eruptions at subduction volcanoes is lower but statistically indistinguishable from the overall TPR curve for all t, illustrating how the total TPR is dominated by this largest of all groups (Figure 8d). While the numbers of eruptions at intraplate and rift volcanoes are much lower than for subduction zones, their higher TPR is statistically significant. For these curves, the lower smoothness is a function of both seismicity rate and eruption sample size, as larger groups of eruptions in highly seismogenic regions have the smoothest TPR curves (e.g., subduction zones). Smoother curves also imply broader transitions (higher variability) from background rate to run-up time for these groups and thus more difficulty distinguishing the two. In contrast to subduction zones, the TPR for intraplate eruptions climbs more quickly with increasing t and no new TP earthquakes are identified for $t > 3$ years, suggesting both shorter run-up times and reduced tectonic seismicity rates, and thus higher TPR. In between

these two are the rift type eruptions, which are fewest in number but still significantly higher in TPR than subduction eruptions.

In general, subduction zones are highly tectonically active areas with time variable criticality. The earthquake cycle of loading, rupture, and reloading may occur over a faster time scale in subduction zones than in intraplate and rift settings. Interestingly, remote dynamic earthquake triggering occurs more rarely in seismically and volcanically active subduction zones in Japan and Alaska as compared to some intraplate and extensional tectonic settings (Harrington & Brodsky, 2006; Prejean & Hill, 2018), suggesting that the subduction zone environments are poised at criticality a smaller percentage of the time. Thus, higher stress changes may be required to trigger M4+ earthquakes in subduction zones than in other environments. Furthermore, eruptions in rift and intraplate settings are often related to mafic dike intrusions whereby magma injects into the surrounding country rock (e.g., Belachew et al., 2013; Neal et al., 2018; Passarelli et al., 2015) rather than ascending through well established and preferred pathways to central vents, which may result in large magnitude induced earthquakes. Finally, mid-ocean ridge volcanoes often locate near associated transform faults, which tend to have smaller maximum magnitudes and shorter recurrence intervals than other tectonic environments (e.g., Boettcher & Jordan, 2004), which would also tend to increase the TPR.

Eruptions related to mafic dike intrusions in oceanic crust may also explain other volcano and eruption attributes that produce higher TPRs—for example, the relatively high TPR for VEI 0 eruptions. For VEI ≥ 1 eruptions, there is a correlation between VEI and TPR for small t (Figure 7), with higher TPRs for larger VEIs. This may be expected since more explosive eruptions are more likely to have larger and more numerous seismic precursors as more work is required to move larger volumes of magma from depth to the surface, perturbing larger regions of the tectonically stressed crust. In addition, rapidly ascending magma can lead to rapid fluid exsolution, which can overpressurize the crust, further inducing seismicity. However, VEI 0 eruptions are fundamentally different from VEI ≥ 1 eruptions, as the former are entirely effusive and often occur at basaltic calderas and shield systems in oceanic crust (such as Kilauea, Hawaii, e.g.). As such, these eruptions often also involve rapid movement of large volumes of magma through the subsurface, often as propagating dikes.

Although eruptions with the attributes common to dike intrusions (intraplate, rift, oceanic, VEI 0, shield) tend to have higher TPRs (and shorter run-up times) than eruptions in other settings, other attributes produce high TPRs as well. For example, silicic volcanoes have higher TPR than both mafic (most t) and intermediate (all t) volcanoes (Figures 7 and 8a). Many of the more evolved silicic volcanoes are calderas, which also have relatively high TPRs for larger t (Figure 8b). In addition, background M4+ seismicity rates near these volcanoes appear lower; no additional M4+ earthquakes (and thus TPs) are added beyond $t > \sim 3$ years. In these cases, large, highly evolved magma storage regions frequently erupt diverse magmas from various locations with ephemeral connections to the magma sources. These types of eruptions may also require more work to open pathways for magma expulsion. Alternatively, overpressurized crust and the frequent inflation and deflation of many restless calderas, resulting from geothermal activity, and/or magmatic gas exsolution, may facilitate more frequent, smaller earthquakes. Finally, many M4+ earthquakes observed in these environments are associated with syneruptive processes (e.g., caldera collapses) rather than run-up processes and are not represented in Figures 7 and 8.

A final attribute correlating with a higher TPR is long repose periods preceding eruptions. To illustrate the relation between repose and TPR, we classify volcanic vents as either "open" or "closed" system at the time of eruption based upon whether the previous eruption occurred less than or greater than 15 years before, respectively (see also Cameron et al., 2018; Pesicek et al., 2018). Eruptions at closed-system volcanoes have higher TPRs and are more commonly preceded by M4+ earthquakes than those at open-system volcanoes (Figures 7 and 8e). This increase in TPR for closed systems is insensitive to small changes to the 15-years criterion ($10 \leq t \leq 20$ years; Table S1) and may be underestimated, considering how the GVP defines eruption start and stop times (see Supplementary Information). These findings are consistent with prior work showing that larger eruptions from long-dormant systems tend to have more obvious seismic precursors than open, frequently active systems (e.g., Cameron et al., 2018; Pesicek et al., 2018; White & McCausland, 2019).

7. Use in Eruption Forecasting

In the previous sections, we focused on bulk characteristics of earthquakes that precede eruption onsets (e.g., their locations, magnitudes, MTs) and then considered how frequently different types of eruptions at different volcanoes experience such precursors. In this section, we shift focus to uses of global earthquake data where no association with an eruption is presumed, and where the ultimate goal is assessing the utility of these earthquakes for eruption forecasting. This necessarily expands our analysis to include seismicity at all volcanoes considered active, even those without monitored eruptions. That is, we now explore the occurrence of False Positive (FP) earthquakes that are NOT followed by eruption, which are rarely considered in eruption-focused literature. When combined with other geologic and geophysical observables, such as ground deformation, degassing, and fumarole activity, significant improvements to probabilistic eruption forecasts can be achieved, even for M4+ earthquakes.

7.1. M4+ Earthquake Positive Predictive Value

During a seismic crisis or immediately following a large-magnitude earthquake near a remote volcano lacking local monitoring, we might want to know the probability of an eruption following the earthquake within a specific time period based on the global database of similar earthquakes. In order to compute such a probability, we must know the prevalence of FPs. This allows us to compute the Positive Predictive Value (PPV) of the seismic event for forecasting within the chosen time window, without regard to source process. The PPV (Table 1) depends on the number of TPs and FPs but not FNs, and thus does not explicitly depend on the estimation of M_c (Section 5). In order to calculate the PPV, we specifically define a FP as a t-day window in which at least one M4+ earthquake occurred without a subsequent eruption. In this sense, we are not counting individual earthquakes *per se*, but time windows with earthquakes, such that t windows with multiple M4+ earthquakes that are not followed by an eruption are only counted once.

For all M4+ earthquakes and $t = 14$ days, we calculate an overall PPV of 1%, illustrating that the vast majority of such earthquakes are not rapidly followed by eruptions at nearby active volcanoes. Additional sensitivity tests on this value show little variation in the result. For example, the PPV only increases by 0.1% if we limit the volcano population to those with historic eruptions, rather than using the larger Holocene population. Furthermore, variations due to specific GVP attributes of the volcano cause changes in the overall PPV value that are <1%, most of which cannot be distinguished from random subsampling of the data (Figure S10). Similarly, applying the suite of aftershock removal methods (Section 4.2; Supporting Information) changes the overall PPV by <1%. Finally, we note that smaller radii (R) correlate with larger PPVs, but only minimally (e.g., for $R = 10$ km, PPV = 2%). Thus, although M4+ earthquakes commonly precede eruptions (Figures 1, 3, 7, and 8) the occurrence of a single M4+ earthquake by itself is a poor predictor of subsequent eruption occurrence. However, the occurrence of distinctive types or sequences of earthquakes can be more diagnostic. For example, preeruptive M4+ events are usually normal faulting events (Figure 5) with high percentages of non-DC components (Figure 4) and many eruptions are preceded by multiple M4+ earthquakes over short time periods (Figure 3). Computing the PPV for certain types or specific groups of M4+ may be more informative for forecasting purposes.

Figure 9a illustrates the relation between the number of M4+ earthquakes that preceded the eruptions (Figure 3), the number of similar sized sequences that did not precede eruptions (the FPs), and the resulting PPV. In this case we use $t = 7$ days, which contains ~95% of the events shown in Figure 3. Although FPs are far more common (low PPV) when considering small numbers of M4+ earthquakes, the ratio of TPs to FPs (and thus PPV) generally increases with increasing minimum sequence size. For example, in cases where 25 or more M4+ earthquakes occurred within a 7-days window (32 cases), an eruption followed within 7 days 19% of the time (6 TPs; 26 FPs). When at least 50 M4+ earthquakes occurred within 7 days (7 cases), an eruption has followed twice (PPV = 29%). For the one case where >100 M4+ earthquakes occurred within 7 days near a volcano (Dabbahu volcano, 2005; Figure 3), an eruption also followed. Although the total number of identified cases decreases with increasing sequence size, the increasing PPV illustrates both the increasing likelihood that the earthquakes are magmatic (beyond what would be expected for purely tectonic sequences) and the increasing likelihood that unrest leads to eruption with higher numbers of

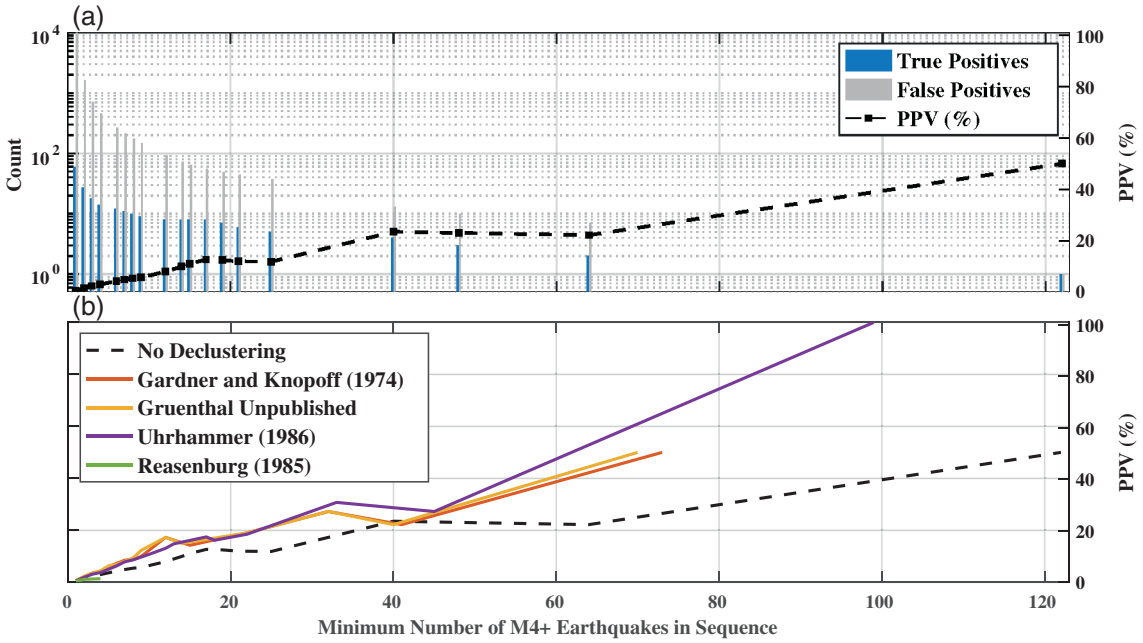


Figure 9. Positive Predictive Value (PPV) vs. the count of M4+ earthquakes for $t = 7$ days, which contains ~95% of the events shown in Figure 3. (a) Histograms of True Positive (blue) and False Positives (gray) are used to compute PPV (black) for eruption TP counts shown at right in Figure 3. (b) Comparison of PPV computed using four different declustering methods discussed in the Supporting Information.

similar-sized earthquakes. Thus, although a single M4+ earthquake is not predictive of an eruption, increasing numbers of such events provide an increasing probability of eruption in the days following.

In the computation of PPV for Figure 9a, we have not attempted to remove aftershocks from the analysis. This is partly due to the difficulties of declustering (Section 4.2; Supporting Information), but also because during a volcanic seismic crisis, it is usually not easy to rapidly and confidently differentiate and remove aftershocks from sustained high seismicity rates due to magma intrusion. Moreover, when we apply these declustering methods, the slope of the PPV trend steepens, suggesting an even stronger relation between numbers of M4+ events and eruptions (Figure 9b).

7.2. Application

Taken together, these results further illustrate the frequent association of seismic rate increases prior to many large eruptions, observable even at the M4+ level. The results support existing observatory practices, as increasing numbers and energy of earthquakes leading into eruption has been a cornerstone of eruption forecasting for decades (e.g., Tokarev, 1971). The occurrence of so many M4+ earthquakes in the few days prior to eruption emphasizes the need to evaluate seismicity quickly to determine if the volcano is showing other signs of unrest (such as increases in fumarolic activity, deformation, degassing, etc.). Observatories need to determine if the earthquakes are decaying like an aftershock sequence (thus more likely tectonic), or if they are continuing to occur in increasing numbers, implying continued pressurization, and evolution of seismicity toward more fluid-driven events (long-period events, tremor, etc.) (e.g., White & McCausland, 2019). Next, we describe two case studies illustrating these concepts in forecasting scenarios at Kasatochi (Alaska) and Rinjani (Indonesia) volcanoes.

7.2.1. Kasatochi Volcano, Alaska, United States

On August 7, 2008, a VEI four eruption began at Kasatochi volcano in the Aleutian Islands of Alaska. In the days prior to the eruption, a vigorous seismic swarm including >12 M4+ earthquakes (Figure 3) was recorded and located using regional stations (Ruppert et al., 2011). Because the volcano was not locally monitored, distinguishing whether these events had tectonic or magmatic origins was initially difficult. However, increasing numbers and magnitudes of earthquakes with time, and ultimately, the detection of

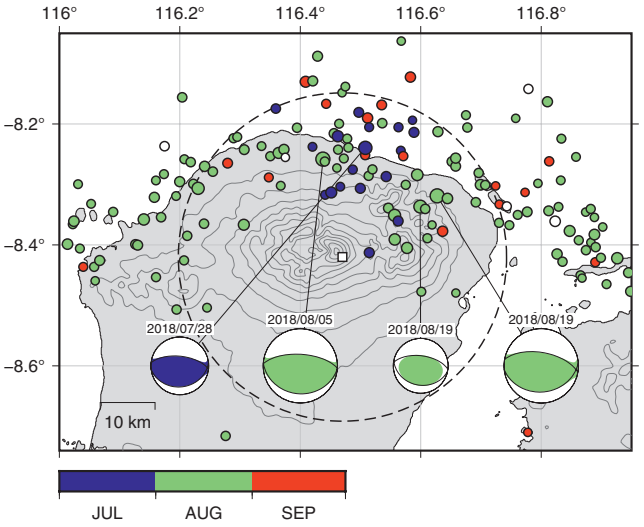


Figure 10. M4+ seismicity near Rinjani Volcano (white square), Lombok Island, Indonesia in 2018. Earthquake data from the USGS ComCat (Table 2). Moment Tensors are shown for M6+ events, as discussed in the text. Dashed circle shows the 30 km radius used to associate earthquakes with volcanoes.

volcanic tremor made it clear that the activity was magmatic. The eruption was successfully forecast (Cameron et al., 2018), including forecast of high explosivity based on high earthquake magnitudes, reaching M5.6.

7.2.2. Rinjani Volcano, Lombok, Indonesia

A more recent example comes from a vigorous seismic sequence near Rinjani volcano in 2018. In this case, the methods and data discussed herein were compared to USGS rapid solutions provided by the National Earthquake Information Center (Benz, 2017) and the results were used directly to frame eruption likelihood estimates during the sequence. On July 28, 2018, an M6.4 earthquake struck ~20 km north of Rinjani volcano at a depth of 14 km. It was followed by an M6.9 earthquake on August 5, ~10 km to the west at 34-km depth. Then, 13 days later an M6.3 occurred (16-km depth) ~20-km to the northeast of Rinjani, which was followed 10 h later by another M6.9 event at 21-km depth. All four M6+ events had MTs with characteristics of back-arc thrusting, with strikes roughly parallel to the subduction trench (Figure 10). Non-DC components for the four events ranged from 7% to 28%. Prior to 2018, only a handful of teleseismically recorded earthquakes had been reported within 30 km of Rinjani volcano, the most recent of which was in 2005. Rinjani volcano is frequently active, last erupting in 2016 and was the site of a VEI seven eruption in 1257 A.D. Currently, over one million people live within 30 km of the volcano and the spate of M6+ events caused over 500 fatalities and widespread destruction on Lombok Island (Supendi et al., 2020). At Rinjani volcano, these earthquakes caused remobilization of ash and landslides, resulting in one fatality and necessitating the rescue of 690 climbers (GVP, 2018). Although no other signs of volcanic unrest were observed, the proximity of these large events to Rinjani, the lack of similar historical seismicity in the area, and the potential swarm-like nature of the events all raised concerns that they could be signaling awakening of the volcano.

Immediately following the second M6.9 earthquake, we compared our volcano earthquake catalogs to rapid earthquake solutions reported by the USGS National Earthquake Information Center (Benz, 2017) (sourced from the USGS ComCat; Table 2). As the results of any query of the global volcano earthquake catalogs depend critically on the parameters of the search and the exact form of the query, we explored a range of reasonable parameters to identify analogous historical sequences near other volcanoes. We first searched for other cases where four M6+ events occurred near a volcano ($R = 30$ km) over the same number of days ($t = 25$) as the Rinjani events, excluding syneruptive sequences. We found only one other matching case near Karkar volcano, Papua New Guinea in 1984. However, in the case of Karkar, the M6+ events occurred several years after the previous eruption in 1979 and the next eruption did not occur until 2012; thus, these earthquakes were post-eruptive, this analog is classified as a FP, and the PPV = 0. Next, we broadened the search to explore for other possible analogs, including cases where at least two M6.5+ events were recorded within $t = 25$ days of each other and within $R = 50$ km of a volcano. We found two additional cases, neither of which can be plausibly associated with eruptive activity, one at an unnamed volcano (GVP volcano number: 281030) offshore Taiwan in 1951 and the other at Izu-Tobu volcano in Japan in 1923. Lowering the magnitude threshold to M6+ identifies several additional earthquake sequences that could be potentially associated with magmatic systems (Mammoth Mountain, 1980; Asosan, 2016; Hakoneyama, 1923; Reventador, 1987), but none of these were followed by an eruption within $t = 6$ months of the events (PPV = 0). In addition, reverse faulting events like those at Rinjani are the least likely type to precede an eruption (Figure 5) and the relatively low level of non-DC component for these events was not indicative of eruption (Figure 4). Finally, and most importantly, there were no changes in other volcano monitoring parameters (D. Syahbana, pers. comm.). Taken together, these observations rapidly provided an assessment that suggested the seismicity would not be preeruptive. Later detailed analysis of the sequence using regional seismic data supports this assessment (Supendi et al., 2020). High precision relocations of the sequence revealed three separate regions of seismicity and support interpretations that the initial event triggered subsequent slip on back-arc faults east and west of the August 5 M6.9, closest to Rinjani. Although it remains possible that

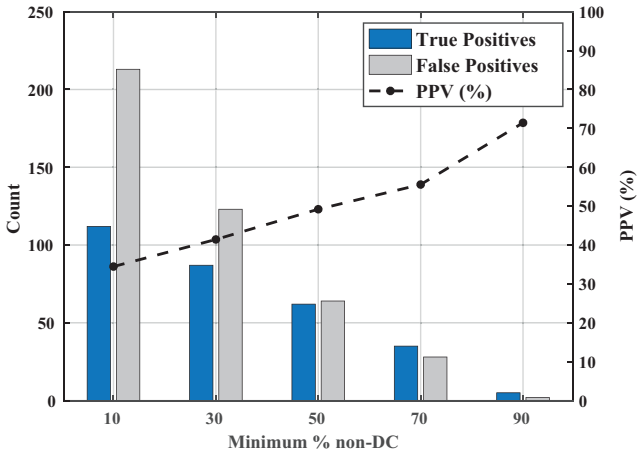


Figure 11. Positive Predictive Value (PPV; Table 1) vs. percentage of nondouble-couple (non-DC) component of the GCMT solutions. Histograms of True Positive (blue) and False Positive (gray) are used to compute PPV (black) for increasing minimum non-DC percentages. In contrast to previous figures, TPs are defined as events within the t = 7-days window or the syneruptive window.

magma intrusion was the ultimate trigger for seismicity on these faults, no other indications of unrest at the volcano have been reported and the alert level of the volcano remains unchanged.

8. Use in Eruption Detection

The previous section focused on the utility of M4+ earthquake occurrence for forecasting eruption onsets and excluded syneruptive events from the PPV calculations. We extend the analyses here in order to test the utility of M4+ earthquakes for eruption detection, which is also extremely valuable, especially for remote or submarine volcanoes.

8.1. Nondouble-Couple M4+ Earthquakes

In Section 4, we showed that preeruptive and syneruptive earthquakes have common features that help distinguish them from noneruptive earthquakes (Figures 2, 4, and 5) that may be useful for probabilistic eruption detection. For example, syneruptive earthquakes often have large non-DC components (Figure 4) that, similar to seismicity rates (Figure 9), may be useful as diagnostics indicators of eruption. Inclusion of syneruptive MTs greatly increases the number of MTs available for consideration, allowing better statistical comparisons between sequences. Figure 11 illustrates how we might use the full population of eruptive MTs rapidly following a large earthquake to inform forecasts. Using the statistics of the GCMT solutions for earthquakes within 30 km of active volcanoes and the eruption timing information, we formulate a PPV for the percentage of non-DC component occurring during the 7-days preeruptive window (t) or a syneruptive period as defined by the eruption start and stop times. The results show that earthquakes with very large non-DC components are much more likely to be related to an eruption than those with smaller ones. Earthquakes with >50% non-DC component are more likely than not (PPV > 50%) to be associated with an eruption. If an earthquake near a volcano has >90% non-DC component, the probability that the event is associated with an eruption is quite high (PPV = ~70%). This may be particularly useful when conditions prevent visual confirmation of eruption onsets, such as at night, during poor atmospheric conditions, or for submarine volcanoes.

8.2. Application: Mayotte Island, Comoros Islands, France

An example where these methods were applied directly for eruption detection comes from the recent seismicity near Mayotte volcano. On May 10, 2018, a seismic crisis began ~50 km east of Mayotte Island (Cesca et al., 2020). The sequence of events continued into early 2020 and included over 200 M4+ earthquakes, the largest of which was an M5.9 on May 15, 2018 whose MT suggests strike-slip faulting with negligible non-DC component (2%). A pause in activity of ~2 months began in July, with M4+ activity resuming again in September, farther west, closer to Mayotte Island (Figure 12). On November 11, 2018, a large monochromatic very low frequency (VLF) event lasting 20 min was recorded by stations around the world, piquing the interest of seismologists worldwide and receiving global media coverage. No seismicity had been located by the USGS in this area prior to May 2018. Recent analysis of available seismic and deformation data points to the explosive birth of a new submarine volcano east of Mayotte Island (Cesca et al., 2020). However, prior to acquisition of new bathymetric data in May 2019, confirmation of an eruption east of Mayotte was not possible, and detailed analysis of the timing of eruption with respect to the seismic sequence remains unclear.

Shortly after the 11 November VLF event, we searched our global volcano earthquake catalogs for prior seismic sequences with similar characteristics to the real-time earthquake data reported by the USGS National Earthquake Information Center (Benz, 2017). We estimated the M_c of the real-time catalog in the area to be 4.4 and we searched for analogous sequences where at least 104 M4.4+ events had occurred over 183 days, which corresponds to the time period between the first such event east of Mayotte and the November 11 globally recorded VLF. We found 11 matching earthquake sequences, six of which were positively associated

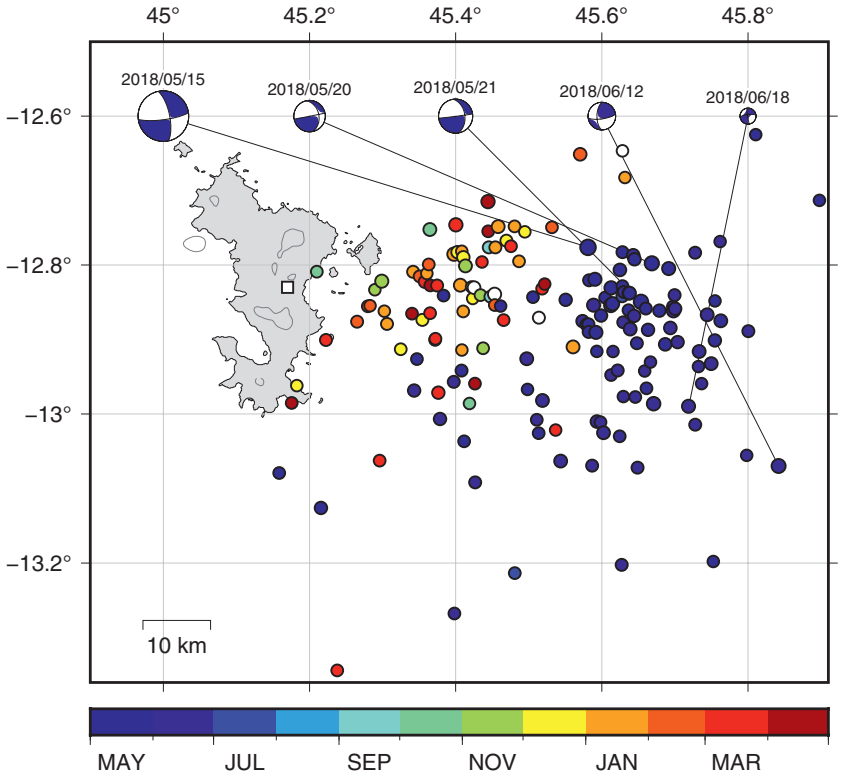


Figure 12. M4+ seismicity near Mayotte Island volcano (white square), Comoros Islands, France in 2018–2019. Earthquake data from the USGS ComCat (Table 2). All available Moment Tensors are shown, including for the largest event (M5.9) on 15 May.

with eruptions (St. Helens volcano, 1980; East Gakkel Ridge, 1999; Miyakejima volcano, 2000; Dabbahu volcano, 2005; Bardarbunga volcano, 2015; Kilauea volcano, 2018) and five that were not (PPV = 55%).

The sequence east of Mayotte also included 19 M5+ events occurring over the course of 1 year. Using these parameters as filters narrows the list of analogous sequences to two TP eruptions at Miyakejima volcano in 2000 and Bardarbunga volcano in 2015 (PPV = 100%). Thus, our analysis of global analog data for this level of sustained seismicity concurs with later detailed analyses that indicate eruptive diking at Mayotte was considerably more likely to be the cause of the earthquake swarms than purely tectonic activity. But was this seismicity preeruptive or syneruptive? One way to search the data is by the timing of the largest magnitude event. The PPV of a single M5.9+ event within 30 km of global volcanoes occurring within a syneruptive period is only 6% based on prior occurrences. Similarly, the likelihood of the M5.9 event being followed by an eruption before the large VLF event occurred (t = 181 days) is 9%. Thus, deciphering eruption timing based solely on the single largest magnitude event, even at the ~M6 level, is not particularly diagnostic. However, we have shown that rates of large magnitude events can be much more diagnostic (Figure 9), where at Mayotte there were 14 M5+ events in the first swarm over the course of ~60 days (Figure 12). We find eight other similar cases globally using these criteria, six of which are known to be syneruptive (PPV = 75%). Thus, the overall high association of eruptions with similar past sequences suggests that there is a high probability that an eruption was already underway by the start of the second swarm in September.

9. Discussion

9.1. Eruption Forecasting and Detection Using Global Earthquake Data

The global compilation of M4+ earthquakes near volcanoes has highlighted several associations between these events and eruptive activity and allowed us to determine the global prevalence of these eruption precursors. The results in turn allow us to update and improve our conceptual models of magma intrusion and

ascent, a goal in line with many prior eruption-centric studies. Few studies have fully quantified the prevalence of similar anomalies outside of eruptive periods (the FPs). In this study, we establish that globally, 1% of M4+ events within 30 km of active volcanoes are followed by eruption within 2 weeks. Thus, although common prior to eruptions, any single M4+ earthquake by itself is a poor predictor of eruptive behavior. Determining the prevalence of these and other precursors outside of eruptive periods is critical for properly weighing their importance within probabilistic eruption forecasts.

Although a single M4+ earthquake has low predictive value, the PPV is greatly affected by the individual characteristics of the events and their temporal occurrence with respect to other such anomalies. For example, we have shown that larger numbers of M4+ events occurring over the forecast window increase the PPV (Figure 9), as do larger percentages of non-DC component (Figure 11). Ultimately, the definition of the anomaly, the chosen forecast window, and other specifics of the data and query employed determine the value and relevance of the results. In forecasting scenarios, the users of such queries must properly evaluate both the characteristics of importance of the anomalous activity and the associated errors and uncertainties of the queried data. For example, the occurrence of a large, shallow, normal faulting event closely located to a restless volcano does not necessarily imply that we should exclude strike-slip and reverse mechanism events from analog searches of global data. In this case, the faulting type may be unique to the local stress conditions and less important than its magnitude and shallow location beneath the active vent. Similarly, excluding events deeper than the earthquake of interest does not account for uncertainties in focal depth, which can be 10 km or more in many cases. These and other similar issues illustrate the susceptibility of the query results to errors in the data and to artificial binning and filtering of phenomena that naturally occur over a continuum.

In this study, we have conducted a variety of tests to assess the robustness of the results with respect to various parameter and binning choices and other data quality issues (see Supplementary Information). Overall, we find the patterns and results discussed herein to be stable with respect to data quality and categorization. None of these tests revealed significant changes in the results and in general confirm the robustness of the statistical features revealed and thus their value to better inform conceptual models of precursory phenomena. Nevertheless, when detailed queries are used in a forecasting context to identify analog sequences or to assess the PPV of a particular anomaly, the results of the query depend critically on its exact formulation and the parameter choices. For simple queries with large populations of analogs, the parameters matter less. But for more complicated queries with fewer matching sequences, the specifics of the query and data can be critical. In these cases, it is important to fully explore the sensitivity of the results within the uncertainty of the data in order to identify appropriate analogs. In Sections 7 and 8, we illustrated two such applications using real-time data reported by the USGS near Rinjani and Mayotte volcanoes and explored the changes in the results due to simple variations in the query formulation.

The analysis presented here has focused on a single indicator of unrest, defined specifically to ensure ergodicity of the results. Although useful for this global analysis, there is no technical limit to applying this query process above or below the M4 cutoff, and we have done so in other cases where circumstances of ongoing unrest have dictated. For real-time comparisons, the M_c of the earthquake data at the volcano of interest is the ultimate limiter of the analog searches. For the example near Rinjani volcano (Section 7.2.2), the earthquakes of interest were well above global real-time completeness thresholds while for those near Mayotte volcano (Section 8.2), we estimated the M_c of the real-time data and adapted the data queries accordingly, ensuring that the results were global in scope. In contrast, using earthquake data below the nominal M4 completeness threshold of the global earthquake catalogs risks breaking the ergodic assumption, as the results will become geographically limited to more densely instrumented regions. Nevertheless, quantitative searches such as these, even when regionally limited, provide unbiased sets of analogs for further comparison. In the framework of a multidisciplinary probabilistic eruption forecast, the query results and their inherent limitations (geographic and otherwise) can be properly weighed in the context of all available data from other monitoring streams. Furthermore, in volcanic crises where local data are limited, these queries provide an important source of baseline information for forecasts of eruptive phenomena, even if only regional. So many of the world's active volcanoes lack the necessary monitoring data over the multiple eruptive cycles needed to build metrics for their specific eruptive behavior. Only by leveraging observations of analogous volcanic behavior at other volcanoes, along with local data when available, can we hope to

overcome the limited temporal span of observations at any given volcano and fully explore the range of possible eruption scenarios and their precursors.

9.2. Future Forecasting and Detection Using Improved, Multidisciplinary Data Sets

As volcano monitoring continues to improve and databases of monitoring data (e.g., WOVOdat; Newhall et al., 2017) expand, similar global analyses of smaller magnitude earthquakes will become more feasible, allowing us to test whether these observations may also apply to precursors occurring below our imposed M4 threshold. Previous studies using smaller magnitude earthquakes have documented clear seismic precursors prior to closed-system eruptions (e.g., Pesicek et al., 2018; White & McCausland, 2016). In other cases, we may not always expect similar results for smaller seismic anomalies. For example, we see no statistically significant distinction between the frequency-magnitude distributions of eruptive periods and noneruptive periods for the M4+ catalog (Figure S4). For smaller earthquakes however, we might expect differences in aftershock productivity, as have been observed in association with volcanic and geothermal activity (e.g., Farrell et al., 2009; McNutt & Roman, 2015), and hydraulic fracturing (e.g., Skoumal et al., 2015).

Importantly, the inclusion of smaller magnitude earthquakes and local seismic data (particularly in real-time) will allow for more robust analysis of other aspects of seismic unrest, including earthquake rates, locations, types, RSAM, reduced displacement, etc., as well as seismic metrics not yet commonly tracked, but which show potential forecasting value. For example, temporal changes in earthquake families (e.g., Buurman & West, 2010; Rodgers et al., 2013; Rowe et al., 2004) and seismic velocity (e.g., Bennington et al., 2018; Brenguier et al., 2008; Machacca-Puma et al., 2019) have been observed preceding some eruptions. Although studies of these changes inform our conceptual models of intrusion and eruption, their forecasting value is limited without more complete statistics on their overall occurrence in relation to eruptions, including analysis of their FNs and FPs. The lack of such statistics, even for the most basic monitoring metrics, remains an impediment to improving future multidisciplinary probabilistic eruption forecasts.

To facilitate such analyses, more observatories would need to submit data from local monitoring networks to global databases such as WOVOdat and the ISC. However, even if this were accomplished globally, significant challenges exist in the standardization of seismic unrest indicators in order for such methods to be applied. In general, seismic metrics tracked at one volcano are not easily comparable to other volcanoes unless the metrics are standardized. This is well illustrated by the earthquake catalogs analyzed herein, where many data issues had to be addressed in the creation of the hypocenter catalogs for each volcano (Section 2). The development of standardized metrics and their potential utility for forecasting remains an outstanding issue for the wider application of these statistical techniques. Although this problem appears tractable for seismic data (e.g., Pérez et al., 2020; this study), it may be less so for other types of equally important monitoring data, including deformation, gas geochemistry, and remote sensing data, among others. As more of these data become broadly available, similar global analyses become feasible, both individually (e.g., Biggs et al., 2014; this study) and within the context of multiparametric statistical analyses of unrest (e.g., Reath et al., 2020). Multidisciplinary approaches that synthesize data from multiple sources have the potential to outperform those that rely on only one data stream.

Finally, future improvements in probabilistic eruption forecasting also require more detailed eruption chronologies. The analysis in this paper focuses on forecasting eruption onsets, largely because eruption start and end dates are the most easily obtainable eruption chronology information available, despite their limitations (see Supplementary Information). Current efforts are underway by our group and others to establish more detailed eruption chronologies globally, which will facilitate a more event-based statistical analysis of monitoring data and allow for forecasting changes in eruptive size or style after an eruption has begun. For example, with more detailed information on explosion sizes and dates during eruptions, the analysis in this paper could be extended to intraeruptive explosions or changes in syneruptive behavior.

10. Conclusions

Global statistical analyses of eruption precursors and their variations due to volcanic and eruptive attributes can improve conceptual models of magmatic intrusion and eruption, greatly benefiting eruption forecasts. Statistics of global eruptive data reveal patterns that may not be obvious in more limited data sets where

errors and uncertainties have greater effects. In this study, we have presented a framework for such efforts by evaluating the forecasting value of M4+ earthquakes as proxies for volcanic unrest and by computing baseline statistics on their global occurrence with respect to eruptions. We used metrics borrowed from predictive analytics (Table 1) and applied them to global volcanic and earthquake data (Table 2) to reveal patterns in eruptive data and to quantify their prevalence both within and outside of eruptive periods. We have not only examined the characteristics of commonly studied preeruptive seismicity (TPs) on a global scale, but also those eruptions without such seismicity (FNs) and earthquakes not followed by eruptions (FPs). This examination of FNs and FPs is generally lacking in the volcanological literature. Finally, the use of predictive indicators (e.g., PPV) on monitoring anomalies and the unbiased groups of analogs that they provide promise to further improve future probabilistic eruption forecasts. When used during a crisis, the queries and parameters used (e.g., Table 1) can be dynamically updated as changes in monitoring data dictate, providing rapid quantitative assessments of similar historical volcanic activity, improving eruption forecasts and the decisions that rely on them.

Specific conclusions from our analysis include:

- Most precursory M4+ earthquakes (TPs) occur in the few days immediately preceding eruption (Figures 1 and 3), 5–15 km from the volcanoes (Figure 2), and most commonly have normal faulting mechanisms (Figure 5) (Section 4)
- Syneruptive M4+ earthquakes tend to locate over a broad spatial area, with events often occurring farther from the volcanic center and deeper than preeruptive seismicity (Figure 2). Syneruptive earthquakes also most commonly occur on normal faults (Figure 5), often with large non-DC components (Figure 4) (Section 4)
- We find that 83 out of 728 total eruptions (TPR = 11%) were preceded by a M4+ earthquake within 30 km of the volcano and within 2 weeks of the eruption (Figure 7; Section 6). The median lag from the first such earthquake to eruption is ~4 days (Figure 3; Section 4.3)
- We find 645 FN eruptions where no M4+ precursor was recorded, yet there was monitoring adequate to detect them (Figure 6; Section 5)
- Several types of volcanoes are correlated with higher occurrences of M4+ earthquake precursors (TPR up to 40%) including shield volcanoes and calderas, volcanoes on oceanic crust, in rift and intraplate settings, those with silicic and mafic lavas, and those with long repose times (Figures 7 and 8; Section 6)
- Effusive VEI 0 eruptions are frequently associated with M4+ earthquake precursors (Figures 7 and 8f). These eruptions commonly occur at shield volcanoes in intraplate, rift, and oceanic settings, attributes which also show higher predilection for M4+ earthquake precursors and where mafic diking is also common (Section 6)
- For explosive eruptions (VEI ≥ 1), the occurrence of M4+ precursors increases with increasing VEI (Figures 7 and 8f)
- Globally, 1% of M4+ earthquakes occurring within 30 km of active volcanoes are followed by an eruption within 2 weeks (PPV = 1%) (Section 7.1). However, larger numbers of M4+ earthquakes (Figure 9) and larger degrees of non-DC component in the MT (Figure 11) increase the PPV and can be diagnostic in some cases

Data Availability Statement

All data used in this study are available from the data sources listed in Table 2 (http://volcano.si.edu; http://www.isc.ac.uk; https://earthquake.usgs.gov/earthquakes/search/; https://www.globalcmt.org; http://www.data.jma.go.jp; http://www2.ssn.unam.mx:8080/catalogo; http://hraun.vedur.is/ja/viku; http://www.ign.es/web/ign/portal/sis-catalogo-terremotos; http://cnt.rm.ingv.it/en; https://www.geonet.org.nz/data/types/eq_catalogue; https://doi.org/10.3133/sir20195037). The primary source of parametric earthquake data used is the International Seismological Centre (ISC) Bulletin (ISC Bulletin, 2017).

Acknowledgments
We thank all those responsible for maintaining these databases, without which this study would not have been possible. In addition, we thank Kristín Jónsdóttir (Icelandic Met Office) for help accessing the Icelandic earthquake data, and Ivan Rodriguez Rasilla and Victor Hugo Espíndola (Servicio Sismológico Nacional) for help accessing the Mexican national earthquake catalog. This effort also benefited from numerous discussions with Christina Widiwijayanti and the other developers of WOVOdat at the Earth Observatory of Singapore. In addition, we thank USGS colleagues John Ewert, Heather Wright, Jake Lowenstern, Gavin Hayes, and Larry Mastin for critical reviews of various versions of this manuscript. Finally, we are very grateful to Laura Sandri and Luigi Passarelli for thoughtful and thorough reviews.
VDAP receives funding from the U.S. Agency for International Development's Bureau for Humanitarian Assistance and the U.S. Geological Survey's Volcano Hazards Program. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

References

Abe, K. (1992). Seismicity of the caldera-making eruption of Mount Katmai, Alaska in 1912. *Bulletin of the Seismological Society of America*, *82*(1), 175–191.

Álvarez-Gómez, J. A. (2014). FMC: A one-liner Python program to manage, classify and plot focal mechanisms. *Geophysical Research Abstracts*, *16*, 10887.
Belachew, M., Ebinger, C., & Coté, D. (2013). Source mechanisms of dike-induced earthquakes in the Dabbahu-Manda Hararo rift segment in Afar, Ethiopia: Implications for faulting above dikes. *Geophysical Journal International*, *192*(3), 907–917. https://doi.org/10.1093/gji/ggs076
Bennington, N., Haney, M., Thurber, C., & Zeng, X. (2018). Inferring magma dynamics at Veniamin of Volcano via application of ambient noise. *Geophysical Research Letters*, *45*, 11650–11658. https://doi.org/10.1029/2018GL079909
Benoit, J. P., & McNutt, S. R. (1996). *Global volcanic earthquake swarm database 1979-1989 (USGS Numbered Series No. 96–69)*. U.S. Geological Survey. Retrieved from https://pubs.er.usgs.gov/publication/ofr9669
Benz, H. (2017). Building a National Seismic Monitoring Center: NEIC from 2000 to the present. *Seismological Research Letters*, *88*(2B), 457–461. https://doi.org/10.1785/0220170034
Biggs, J., Ebmeier, S. K., Aspinall, W. P., Lu, Z., Pritchard, M. E., Sparks, R. S. J., & Mather, T. A. (2014). Global link between deformation and volcanic eruption quantified by satellite imagery. *Nature Communications*, *5*, 3471. https://doi.org/10.1038/ncomms4471
Bodvarsson, R., Rognvaldsson, S. T., Jakobsdottir, S. S., Slunga, R., & Stefansson, R. (1996). The SIL data acquisition and monitoring system. *Seismological Research Letters*, *67*(5), 35–46. https://doi.org/10.1785/gssrl.67.5.35
Boettcher, M. S., & Jordan, T. H. (2004). Earthquake scaling relations for mid-ocean ridge transform faults. *Journal of Geophysical Research*, *109*, B12302. https://doi.org/10.1029/2004JB003110
Bondár, I., & Storchak, D. (2011). Improved location procedures at the International Seismological Centre. *Geophysical Journal International*, *186*(3), 1220–1244. https://doi.org/10.1111/j.1365-246X.2011.05107.x
Brenguier, F., Shapiro, N. M., Campillo, M., Ferrazzini, V., Duputel, Z., Coutant, O., & Nercessian, A. (2008). Towards forecasting volcanic eruptions using seismic noise. *Nature Geoscience*, *1*(2), 126–130. https://doi.org/10.1038/ngeo104
Budi-Santoso, A., Lesage, P., Dwiyono, S., Sumarti, S., Subandriyo, S., Surono, S., et al. (2013). Analysis of the seismic activity associated with the 2010 eruption of Merapi Volcano, Java. *Journal of Volcanology and Geothermal Research*, *261*, 153–170. https://doi.org/10.1016/j.jvolgeores.2013.03.024
Buurman, H., & West, M. E. (2010). Seismic precursors to volcanic explosions during the 2006 eruption of Augustine Volcano. In *The 2006 eruption of Augustine volcano, Alaska* (Numbered Series No. 1769-2, Chap. 2, Vol. *1769–2*, p. 4157). U.S. Geological Survey. https://doi.org/10.3133/pp17692
Buurman, H., West, M. E., & Thompson, G. (2013). The seismicity of the 2009 Redoubt eruption. *Journal of Volcanology and Geothermal Research*, *259*, 16–30. https://doi.org/10.1016/j.jvolgeores.2012.04.024
Cameron, C. E., Prejean, S. G., Coombs, M. L., Wallace, K. L., Power, J. A., & Roman, D. C. (2018). Alaska Volcano observatory alert and forecasting timeliness: 1989–2017. *Frontiers in Earth Science*, *6*, 86. https://doi.org/10.3389/feart.2018.00086
Cesca, S., Letort, J., Razafindrakoto, H. N. T., Heimann, S., Rivalta, E., Isken, M. P., et al. (2020). Drainage of a deep magma reservoir near Mayotte inferred from seismicity and deformation. *Nature Geoscience*, *13*(1), 87–93. https://doi.org/10.1038/s41561-019-0505-5
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. https://doi.org/10.1016/0377-0273(94)90030-2
Coulon, C. A., Hsieh, P. A., White, R., Lowenstern, J. B., & Ingebritsen, S. E. (2017). Causes of distal volcano-tectonic seismicity inferred from hydrothermal modeling. *Journal of Volcanology and Geothermal Research*, *345*, 98–108. https://doi.org/10.1016/j.jvolgeores.2017.07.011
Duputel, Z., Ferrazzini, V., Brenguier, F., Shapiro, N., Campillo, M., & Nercessian, A. (2009). Real time monitoring of relative velocity changes using ambient seismic noise at the Piton de la Fournaise volcano (La Réunion) from January 2006 to June 2007. *Journal of Volcanology and Geothermal Research*, *184*(1), 164–173. https://doi.org/10.1016/j.jvolgeores.2008.11.024
Ekström, G. (1994). Anomalous earthquakes on volcano ring-fault structures. *Earth and Planetary Science Letters*, *128*(3), 707–712. https://doi.org/10.1016/0012-821X(94)90184-8
Ekström, G., Nettles, M., & Dziewoński, A. M. (2012). The global CMT project 2004–2010: Centroid-moment tensors for 13,017 earthquakes. *Physics of the Earth and Planetary Interiors*, *200–201*, 1–9. https://doi.org/10.1016/j.pepi.2012.04.002
Farrell, J., Husen, S., & Smith, R. B. (2009). Earthquake swarm and b-value characterization of the Yellowstone volcano-tectonic system. *Journal of Volcanology and Geothermal Research*, *188*(1–3), 260–276. https://doi.org/10.1016/j.jvolgeores.2009.08.008
Fee, D., Haney, M. M., Matoza, R. S., Eaton, A. R. V., Cervelli, P., Schneider, D. J., & Iezzi, A. M. (2017). Volcanic tremor and plume height hysteresis from Pavl of Volcano, Alaska. *Science*, *355*(6320), 45–48. https://doi.org/10.1126/science.aah6108
Gardner, J. K., & Knopoff, L. (1974). Is the sequence of earthquakes in Southern California, with aftershocks removed, Poissonian? *Bulletin of the Seismological Society of America*, *64*(5), 1363–1367.
Garza-Giron, R., Brodsky, E. E., & Prejean, S. G. (2018). Mainshock-aftershock clustering in volcanic regions. *Geophysical Research Letters*, *45*, 1370–1378. https://doi.org/10.1002/2017GL075738
Global Volcanism Program. (2013). In E. Venzke (Ed.), Volcanoes of the world, v. 4.6.7. Smithsonian Institution. https://dx.doi.org/10.5479/si.GVP.VOTW4-2013
Global Volcanism Program. (2018). Report on Rinjani (Indonesia). In S. K. Sennert (Ed.), *Weekly volcanic activity report, 25 July-31 July 2018*. Smithsonian Institution and US Geological Survey.
Harlow, D. H., Power, J. A., Laguerta, E. P., Ambubuyog, G., White, R. A., & Hoblitt, R. P. (1996). Precursory seismicity and forecasting of the June 15, 1991, eruption of Mount Pinatubo. In C. G. Newhall, & R. S. Punongbayan (Eds.), *Fire and mud: Eruptions and lahars of mount Pinatubo, Philippines* (pp. 285–305). Seattle: Philippine Institute of Volcanology and Seismology, Quezon City and University of Washington Press.
Harrington, R. M., & Brodsky, E. E. (2006). The absence of remotely triggered seismicity in Japan. *Bulletin of the Seismological Society of America*, *96*(3), 871–878. https://doi.org/10.1785/0120050076
Hill, D. P. (1977). A model for earthquake swarms. *Journal of Geophysical Research*, *82*(8), 1347. https://doi.org/10.1029/JB082i008p01347
Hill, D. P., Pollitz, F., & Newhall, C. (2002). Earthquake–volcano interactions. *Physics Today*, *55*(11), 41–47. https://doi.org/10.1063/1.1535006
Hill, D. P., & Prejean, S. (2005). Magmatic unrest beneath Mammoth Mountain, California. *Journal of Volcanology and Geothermal Research*, *146*(4), 257–283. https://doi.org/10.1016/j.jvolgeores.2005.03.002
Holtkamp, S. G., & Brudzinski, M. R. (2011). Earthquake swarms in circum-Pacific subduction zones. *Earth and Planetary Science Letters*, *305*(1–2), 215–225. https://doi.org/10.1016/j.epsl.2011.03.004
Holtkamp, S. G., Pritchard, M. E., & Lohman, R. B. (2011). Earthquake swarms in South America. *Geophysical Journal International*, *187*(1), 128–146. https://doi.org/10.1111/j.1365-246X.2011.05137.x

ISC. (2017). International Seismological Centre. Thatcham, UK:Internatl. Seismol. Cent. On-line Bulletin. Retrieved from https://www.isc.ac.uk

Johnson, J. H., Prejean, S., Savage, M. K., & Townend, J. (2010). Anisotropy, repeating earthquakes, and seismicity associated with the 2008 eruption of Okmok volcano, Alaska. *Journal of Geophysical Research*, *115*, B00B04. https://doi.org/10.1029/2009JB006991

Julian, B. R., Miller, A. D., & Foulger, G. R. (1998). Non-double-couple earthquakes 1. Theory. *Reviews of Geophysics*, *36*(4), 525–549. https://doi.org/10.1029/98RG00716

Kagan, Y. Y. (2005). Double-couple earthquake focal mechanism: Random rotation and display. *Geophysical Journal International*, *163*(3), 1065–1072. https://doi.org/10.1111/j.1365-246X.2005.02781.x

Knopoff, L., & Gardner, J. (1972). Higher seismic activity during local night on the raw worldwide earthquake catalogue. *Geophysical Journal of the Royal Astronomical Society*, *28*, 311–313.

Linde, A. T., Kamigaichi, O., Churei, M., Kanjo, K., & Sacks, S. (2016). Magma chamber recharging and tectonic influence on reservoirs: The 1986 eruption of Izu-Oshima. *Journal of Volcanology and Geothermal Research*, *311*, 72–78. https://doi.org/10.1016/j.jvolgeores.2016.01.001

Linde, A. T., & Sacks, I. S. (1998). Triggering of volcanic eruptions. *Nature*, *395*(6705), 888–890. https://doi.org/10.1038/27650

Llenos, A. L., & Michael, A. J. (2019). Ensembles of ETAS models provide optimal operational earthquake forecasting during swarms: Insights from the 2015 San Ramon, California Swarm. *Bulletin of the Seismological Society of America*, *109*(6), 2145–2158. https://doi.org/10.1785/0120190020

Lolli, B., Gasperini, P., & Vannucci, G. (2014). Empirical conversion between teleseismic magnitudes (mb and Ms) and moment magnitude (Mw) at the global, Euro-Mediterranean and Italian scale. *Geophysical Journal International*, *199*(2), 805–828. https://doi.org/10.1093/gji/ggu264

Machacca-Puma, R., Lesage, P., Larose, E., Lacroix, P., & Anccasi-Figueroa, R. M. (2019). Detection of pre-eruptive seismic velocity variations at an andesitic volcano using ambient noise correlation on 3-component stations: Ubinas volcano, Peru, 2014. *Journal of Volcanology and Geothermal Research*, *381*, 83–100. https://doi.org/10.1016/j.jvolgeores.2019.05.014

Manga, M., & Brodsky, E. (2006). Seismic triggering of eruptions in the far field: Volcanoes and geysers. *Annual Review of Earth and Planetary Sciences*, *34*(1), 263–291. https://doi.org/10.1146/annurev.earth.34.031405.125125

Marzocchi, W., & Bebbington, M. S. (2012). Probabilistic eruption forecasting at short and long time scales. *Bulletin of Volcanology*, *74*(8), 1777–1805. https://doi.org/10.1007/s00445-012-0633-x

Marzocchi, W., Sandri, L., Gasparini, P., Newhall, C., & Boschi, E. (2004). Quantifying probabilities of volcanic events: The example of volcanic hazard at Mount Vesuvius. *Journal of Geophysical Research*, *109*, B11201. https://doi.org/10.1029/2004JB003155

Marzocchi, W., Sandri, L., & Selva, J. (2008). BET_EF: A probabilistic tool for long-and short-term eruption forecasting. *Bulletin of Volcanology*, *70*(5), 623–632.

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). Berlin, Heidelberg: Springer. Retrieved from https://link.springer.com/chapter/10.1007/978-3-642-80087-0_3

McNutt, S. R., & Roman, D. C. (2015). Chapter 59—Volcanic seismicity. In H. Sigurdsson (Ed.), *The encyclopedia of volcanoes* (2nd ed.,pp. 1011–1034). Amsterdam: Academic Press. https://doi.org/10.1016/B978-0-12-385938-9.00059-6

Mignan, A. (2012). *Estimating the magnitude of completeness for earthquake catalogs. Community online resource for statistical seismicity analysis.* Retrieved from https://e-citations.ethbib.ethz.ch/view/pub:100272

Miller, A. D., Foulger, G. R., & Julian, B. R. (1998). Non-double-couple earthquakes 2. Observations. *Reviews of Geophysics*, *36*(4), 551–568. https://doi.org/10.1029/98RG00717

Moran, S. C. (1994). Seismicity at Mount St. Helens, 1987–1992: Evidence for repressurization of an active magmatic system. *Journal of Geophysical Research*, *99*(B3), 4341–4354. https://doi.org/10.1029/93JB02993

Moran, S. C., Newhall, C., & Roman, D. C. (2011). Failed magmatic eruptions: Late-stage cessation of magma ascent. *Bulletin of Volcanology*, *73*(2), 115–122. https://doi.org/10.1007/s00445-010-0444-x

Mori, J., White, R. A., Harlow, D. H., Okubo, P., Power, J. A., Hoblitt, R. P., et al. (1996). Volcanic earthquakes following the 1991 climactic eruption of Mount Pinatubo: Strong seismicity during a waning eruption. In C. G. Newhall, & R. S. Punongbayan (Eds.), *Fire and mud: Eruptions and lahars of mount Pinatubo, Philippines* (pp. 339–350). Seattle: PHIVOLCS and University of Washington.

Murray, T. L. (1992). A system for acquiring, storing, and analyzing low-frequency time-series data in near-real time. In J. W. Ewert, & D. A. Swanson (Eds.), *Monitoring volcanoes: Techniques and strategies used by the staff of the Cascades volcano observatory, 1980-90: U.S.* (pp. 37–43). Geological Survey Bulletin 1966.

National Academies of Sciences, Engineering, and Medicine. (2017). *Volcanic eruptions and their repose, unrest, precursors, and timing*. Washington, DC: The National Academies Press. https://doi.org/10.17226/24650

NCEDC. (2016). *Northern California earthquake data center*. UC Berkeley Seismological Laboratory. Dataset. https://doi.org/10.7932/NCEDC

Neal, C. A., Brantley, S. R., Antolik, L., Babb, J., Burgess, M., Calles, K., et al. (2018). The 2018 rift eruption and summit collapse of Kīlauea Volcano. *Science*, *363*, 367–374. https://doi.org/10.1126/science.aav7046

Newhall, C., & Hoblitt, R. (2002). Constructing event trees for volcanic crises. *Bulletin of Volcanology*, *64*(1), 3–20.

Newhall, C. G., Costa, F., Ratdomopurbo, A., Venezky, D. Y., Widiwijayanti, C., Win, N. T. Z., et al. (2017). WOVOdat—An online, growing library of worldwide volcanic unrest. *Journal of Volcanology and Geothermal Research*, *345*, 184–199. https://doi.org/10.1016/j.jvolgeores.2017.08.003

Newhall, C. G., & Pallister, J. S. (2015). Using multiple data sets to populate probabilistic volcanic event trees. In *Volcanic Hazards, risks and disasters* (pp. 203–232). Elsevier.

Newhall, C. G., & Self, S. (1982). The volcanic explosivity index (VEI) an estimate of explosive magnitude for historical volcanism. *Journal of Geophysical Research*, *87*(C2), 1231–1238. https://doi.org/10.1029/JC087iC02p01231

Nishimura, T. (2017). Triggering of volcanic eruptions by large earthquakes. *Geophysical Research Letters*, *44*, 7750–7756. https://doi.org/10.1002/2017GL074579

Ogburn, S. E., Loughlin, S. C., & Calder, E. S. (2015). The association of lava dome growth with major explosive activity (VEI ≥ 4): DomeHaz, a global dataset. *Bulletin of Volcanology*, *77*(5), 40.

Panza, G. F., & Saraò, A. (2000). Monitoring volcanic and geothermal areas by full seismic moment tensor inversion: Are non-double-couple components always artefacts of modelling? *Geophysical Journal International*, *143*(2), 353–364. https://doi.org/10.1046/j.1365-246X.2000.01250.x

Passarelli, L., & Brodsky, E. E. (2012). The correlation between run-up and repose times of volcanic eruptions. *Geophysical Journal International*, *188*(3), 1025–1045. https://doi.org/10.1111/j.1365-246X.2011.05298.x

Passarelli, L., Rivalta, E., Cesca, S., & Aoki, Y. (2015). Stress changes, focal mechanisms, and earthquake scaling laws for the 2000 dike at Miyakejima (Japan). *Journal of Geophysical Research: Solid Earth*, *120*, 4130–4145. https://doi.org/10.1002/2014JB011504

Pérez, N., Benítez, D., Grijalva, F., Lara-Cueva, R., Ruiz, M., & Aguilar, J. (2020). ESeismic: Towards an Ecuadorian volcano seismic repository. *Journal of Volcanology and Geothermal Research*, *396*, 106855. https://doi.org/10.1016/j.jvolgeores.2020.106855

Pérez-Campos, X., Espíndola, V. H., Pérez, J., Estrada, J. A., Monroy, C. C., Bello, D., et al. (2018). The Mexican National Seismological Service: An overview. *Seismological Research Letters*, *89*(2A), 318–323. https://doi.org/10.1785/0220170186

Pesicek, J. D., Sileny, J., Prejean, S. G., & Thurber, C. H. (2012). Determination and uncertainty of moment tensors for microearthquakes at Okmok Volcano, Alaska. *Geophysical Journal International*, *190*(3), 1689–1709. https://doi.org/10.1111/j.1365-246X.2012.05574.x

Pesicek, J. D., Wellik, J. J. I., Prejean, S. G., & Ogburn, S. E. (2018). Prevalence of seismic rate anomalies preceding volcanic eruptions in Alaska. *Frontiers in Earth Science*, *6*, 100. https://doi.org/10.3389/feart.2018.00100

Pliny the Younger. (1835). *Select letters of Pliny the Younger*. Boston: Perkins, Marvin, & Co. Retrieved from the Library of Congress.

Poland, M. P., & Anderson, K. R. (2020). Partly cloudy with a chance of Lava Flows: Forecasting volcanic eruptions in the twenty-first century. *Journal of Geophysical Research: Solid Earth*, *125*, e2018JB016974. https://doi.org/10.1029/2018JB016974

Power, J., Jolly, A., Page, R., & McNutt, S. (1995). *Seismicity and forecasting of the 1992 Eruptions of Crater Peak Vent, Mt. Spurr, Alaska: An overview* (pp. 149–159): U.S. Geological Survey. https://doi.org/10.3133/b2139

Power, J. A., Friberg, P. A., Haney, M. M., Parker, T., Stihler, S. D., & Dixon, J. P. (2019). *A unified catalog of earthquake hypocenters and magnitudes at volcanoes in Alaska—1989 to 2018* (p. 17). U.S. Geological Survey Scientific Investigations Report 2019-5037. https://doi.org/10.3133/sir20195037

Power, J. A., Lahr, J. C., Page, R. A., Chouet, B. A., Stephens, C. D., Harlow, D. H., et al. (1994). Seismic evolution of the 1989–1990 eruption sequence of Redoubt Volcano, Alaska. *Journal of Volcanology and Geothermal Research*, *62*(1), 69–94. https://doi.org/10.1016/0377-0273(94)90029-9

Power, J. A., & Lalla, D. J. (2010). Seismic observations of Augustine Volcano, 1970-2007. In *The 2006 eruption of Augustine volcano, Alaska* (USGS Numbered Series No. 1769–1, Chap. 1, p. 340). U.S. Geological Survey. https://doi.org/10.3133/pp17691

Prejean, S. G., & Hill, D. P. (2018). The influence of tectonic environment on dynamic earthquake triggering: A review and case study on Alaskan volcanoes. *Tectonophysics*, *745*, 293–304. https://doi.org/10.1016/j.tecto.2018.08.007

Reasenberg, P. (1985). Second-order moment of central California seismicity, 1969–1982. *Journal of Geophysical Research*, *90*(B7), 5479–5495. https://doi.org/10.1029/JB090iB07p05479

Reath, K., Pritchard, M., Biggs, J., Andrews, B., Ebmeier, S. K., Bagnardi, M., et al. (2020). Using conceptual models to relate multiparameter satellite data to subsurface volcanic processes in Latin America. *Geochemistry, Geophysics, Geosystems*, *21*, e2019GC008494. https://doi.org/10.1029/2019GC008494

Reath, K., Pritchard, M., Poland, M., Delgado, F., Carn, S., Coppola, D., et al. (2019). Thermal, deformation, and degassing remote sensing time series (CE 2000–2017) at the 47 most active volcanoes in Latin America: Implications for volcanic systems. *Journal of Geophysical Research: Solid Earth*, *124*, 195–218. https://doi.org/10.1029/2018JB016199

Rodgers, M., Roman, D. C., Geirsson, H., LaFemina, P., Muñoz, A., Guzman, C., & Tenorio, V. (2013). Seismicity accompanying the 1999 eruptive episode at Telica Volcano, Nicaragua. *Journal of Volcanology and Geothermal Research*, *265*, 39–51. https://doi.org/10.1016/j.jvolgeores.2013.08.010

Roman, D. C., & Cashman, K. V. (2018). Top–down precursory volcanic seismicity: Implications for 'Stealth' magma ascent and long-term eruption forecasting. *Frontiers in Earth Science*, *6*, 124. https://doi.org/10.3389/feart.2018.00124

Ruppert, N. A., Prejean, S., & Hansen, R. A. (2011). Seismic swarm associated with the 2008 eruption of Kasatochi Volcano, Alaska: Earthquake locations and source parameters. *Journal of Geophysical Research*, *116*, B00B07. https://doi.org/10.1029/2010JB007435

Sandri, L., Marzocchi, W., & Zaccarelli, L. (2004). A new perspective in identifying the precursory patterns of eruptions. *Bulletin of Volcanology*, *66*(3), 263–275.

Sawi, T. M., & Manga, M. (2018). Revisiting short-term earthquake triggered volcanism. *Bulletin of Volcanology*, *80*(7), 57. https://doi.org/10.1007/s00445-018-1232-2

Scandone, R., & Malone, S. D. (1985). Magma supply, magma discharge and readjustment of the feeding system of mount St. Helens during 1980. *Journal of Volcanology and Geothermal Research*, *23*(3), 239–262. https://doi.org/10.1016/0377-0273(85)90036-8

Shelly, D. R., Taira, T., Prejean, S. G., Hill, D. P., & Dreger, D. S. (2015). Fluid-faulting interactions: Fracture-mesh and fault-valve behavior in the February 2014 Mammoth Mountain, California, earthquake swarm. *Geophysical Research Letters*, *42*, 5803–5812. https://doi.org/10.1002/2015GL064325

Shuler, A., Ekström, G., & Nettles, M. (2013a). Physical mechanisms for vertical-CLVD earthquakes at active volcanoes. *Journal of Geophysical Research: Solid Earth*, *118*, 1569–1586. https://doi.org/10.1002/jgrb.50131

Shuler, A., Nettles, M., & Ekström, G. (2013b). Global observation of vertical-CLVD earthquakes at active volcanoes. *Journal of Geophysical Research: Solid Earth*, *118*, 138–164. https://doi.org/10.1029/2012JB009721

Siebert, L., Simkin, T., & Kimberly, P. (2011). *Volcanoes of the world* (3rd ed., p. 568). University of California Press. Retrieved from https://www.ucpress.edu/book/9780520268777/volcanoes-of-the-world

Sigmundsson, F., Hreinsdóttir, S., Hooper, A., Árnadóttir, T., Pedersen, R., Roberts, M. J., et al. (2010). Intrusion triggering of the 2010 Eyjafjallajökull explosive eruption. *Nature*, *468*(7322), 426–430. https://doi.org/10.1038/nature09558

Skoumal, R. J., Brudzinski, M. R., & Currie, B. S. (2015). Earthquakes induced by hydraulic fracturing in Poland Township. *Ohio Bulletin of the Seismological Society of America*, *105*(1), 189–197. https://doi.org/10.1785/0120140168

Sparks, R. S. J. (2003). Forecasting volcanic eruptions. *Earth and Planetary Science Letters*, *210*(1–2), 1–15. https://doi.org/10.1016/S0012-821X(03)00124-9

Stewart, G. A. (1820). XIV. Description of a volcanic eruption in the Island of Sumbawa. *The Philosophical Magazine*, *56*(268), 96–100. https://doi.org/10.1080/14786442008652373

Supendi, P., Nugraha, A. D., Widiyantoro, S., Pesicek, J. D., Thurber, C. H., Abdullah, C. I., et al. (2020). Relocated aftershocks and background seismicity in eastern Indonesia shed light on the 2018 Lombok and Palu earthquake sequences. *Geophysical Journal International*, *221*(3), 1845–1855. https://doi.org/10.1093/gji/ggaa118

Swanson, D. A., Casadevall, T. J., Dzurisin, D., Malone, S. D., Newhall, C. G., & Weaver, C. S. (1983). Predicting eruptions at Mount St. Helens, June 1980 through December 1982. *Science*, *221*(4618), 1369–1376.

Syahbana, D. K., Kasbani, K., Suantika, G., Prambada, O., Andreas, A. S., Saing, U. B., et al. (2019). The 2017–19 activity at Mount Agung in Bali (Indonesia): Intense unrest, monitoring, crisis response, evacuation, and eruption. *Scientific Reports*, *9*(1), 8848. https://doi.org/10.1038/s41598-019-45295-9

Syracuse, E. M., & Abers, G. A. (2009). Systematic biases in subduction zone hypocenters. *Geophysical Research Letters*, *36*, L10303. https://doi.org/10.1029/2009GL037487

Thelen, W. A., Malone, S. D., & West, M. E. (2010). Repose time and cumulative moment magnitude: A new tool for forecasting eruptions? *Geophysical Research Letters*, *37*, L18301. https://doi.org/10.1029/2010GL044194

Tierz, P., Loughlin, S. C., & Calder, E. S. (2019). VOLCANS: An objective, structured and reproducible method for identifying sets of analogue volcanoes. *Bulletin of Volcanology*, *81*(12), 76.

Tokarev, P. I. (1971). Forecasting volcanic eruptions from seismic data. *Bulletin Volcanologique*, *35*(1), 243–250. https://doi.org/10.1007/BF02596821

Uhrhammer, R. (1986). Characteristics of Northern and Central California seismicity. *Earthquake Notes*, *57*(1), 21–29.

van Stiphout, T., Zhuang, J., & Marsan, D. (2012). *Seismicity declustering, community online resource for statistical seismicity analysis*. Retrieved from https://www.corssa.org. https://doi.org/10.5078/corssa-52382934

Verbeek, R. D. M. (1884). The Krakatoa eruption. *Nature*, *30*(757), 10–15. https://doi.org/10.1038/030010a0

Vidale, J. E., Boyle, K. L., & Shearer, P. M. (2006). Crustal earthquake bursts in California and Japan: Their patterns and relation to volcanoes. *Geophysical Research Letters*, *33*, L20313. https://doi.org/10.1029/2006GL027723

Voight, B. (1990). The 1985 Nevado del Ruiz volcano catastrophe: Anatomy and retrospection. *Journal of Volcanology and Geothermal Research*, *44*(3), 349–386. https://doi.org/10.1016/0377-0273(90)90027-D

Voight, B., Hoblitt, R. P., Clarke, A. B., Lockhart, A. B., Miller, A., Lynch, L., & McMahon, J. (1998). Remarkable cyclic ground deformation monitored in real-time on Montserrat, and its use in eruption forecasting. *Geophysical Research Letters*, *25*(18), 3405–3408. https://doi.org/10.1029/98GL01160

Walter, T. R. (2007). How a tectonic earthquake may wake up volcanoes: Stress transfer during the 1996 earthquake–eruption sequence at the Karymsky Volcanic Group, Kamchatka. *Earth and Planetary Science Letters*, *264*(3), 347–359. https://doi.org/10.1016/j.epsl.2007.09.006

White, R., & McCausland, W. (2016). Volcano-tectonic earthquakes: A new tool for estimating intrusive volumes and forecasting eruptions. *Journal of Volcanology and Geothermal Research*, *309*, 139–155. https://doi.org/10.1016/j.jvolgeores.2015.10.020

White, R. A., & McCausland, W. A. (2019). A process-based model of pre-eruption seismicity patterns and its use for eruption forecasting at dormant stratovolcanoes. *Journal of Volcanology and Geothermal Research*, *382*, 267–297. https://doi.org/10.1016/j.jvolgeores.2019.03.004

Wiemer, S., & Wyss, M. (2000). Minimum magnitude of completeness in earthquake Catalogs: Examples from Alaska, the Western United States, and Japan. *Bulletin of the Seismological Society of America*, *90*(4), 859–869. https://doi.org/10.1785/0119990114

Woessner, J., & Wiemer, S. (2005). Assessing the quality of earthquake catalogues: Estimating the magnitude of completeness and its uncertainty. *Bulletin of the Seismological Society of America*, *95*(2), 684–698. https://doi.org/10.1785/0120040007

Wright, H. M. N., Pallister, J. S., McCausland, W. A., Griswold, J. P., Andreastuti, S., Budianto, A., et al. (2018). Construction of probabilistic event trees for eruption forecasting at Sinabung volcano, Indonesia 2013–14. *Journal of Volcanology and Geothermal Research*. *382*, 233–252. https://doi.org/10.1016/j.jvolgeores.2018.02.003

Yokoyama, I. (2001). The largest magnitudes of earthquakes associated with some historical volcanic eruptions and their volcanological significance. *Annals of Geophysics*, *44*(5–6). https://doi.org/10.4401/ag-3553

Zobin, V. M. (2001). Seismic hazard of volcanic activity. *Journal of Volcanology and Geothermal Research*, *112*(1), 1–14. https://doi.org/10.1016/S0377-0273(01)00230-X