

Multitimescale Template Matching: Discovering Eruption Precursors across Diverse Volcanic Settings

Alberto Ardid, David Dempsey, Josh Corry, Owen Garrett, Oliver D. Lamb, Shane Cronin

Author's accepted manuscript (peer-reviewed, before publisher typesetting)

Published as: Ardid, Alberto, Dempsey, David, Corry, Josh, Garrett, Owen, Lamb, Oliver D., Cronin, Shane (2024). Multitimescale Template Matching: Discovering Eruption Precursors across Diverse Volcanic Settings. Seismological Research Letters, 95(5), 2611-2621. <https://doi.org/10.1785/0220240012>

Accepted manuscript. The version of record, published by the Seismological Society of America in Seismological Research Letters, is at <https://doi.org/10.1785/0220240012>. Not for commercial redistribution.

Open copy, summary and machine-readable versions:

<https://ddempsey.github.io/papers/2024-ardid-multitimescale-template-matching-discovering-eruption/>

Multitimescale Template Matching: Discovering Eruption Precursors across Diverse Volcanic Settings

Alberto Ardid^{*1}, David Dempsey¹, Josh Corry¹, Owen Garrett¹, Oliver D. Lamb², and Shane
Cronin³

1. University of Canterbury, Christchurch, New Zealand, <https://orcid.org/0000-0001-8040-8193>
(AA); <https://orcid.org/0000-0003-2135-5129> (OG); 2. GNS Science|Te Pū Ao, Wairakei Re-
search Centre, Taupo, New Zealand, <https://orcid.org/0000-0002-2254-4258> (ODL); 3. Uni-
versity of Auckland, Auckland, New Zealand, <https://orcid.org/0000-0001-7499-603X> (SC)

*Corresponding author: aardids@gmail.com

Abstract

Volcanic eruptions pose significant risks, demanding precise monitoring for timely hazard mitigation. However, interpreting noisy seismic data for eruptive precursors remains challenging. This study introduces a novel methodology that extends an earlier time-series feature engineering approach to include template matching against prior eruptions. We aim to identify subtle signals within seismic data to enhance our understanding of volcanic activity and future hazards. To do this, we analyze the continuous seismic record at a volcano and identify the time-series elements that regularly precede eruptions and the timescales over which these are observable. We conduct tests across various time lengths, ranging from 1 to 60 days. For Copahue (Chile/Argentina), Pavlof (Alaska), Bezymianny (Russia), and Whakaari (New Zealand) volcanoes, we confirm statistically significant eruption precursors. In particular, a feature named change quantiles (0.2–0.8), which is related to the conditional dynamics of surface acceleration at the volcano, emerges as a key indicator of future eruptions over 14-day timescales. This research offers new methods for real-time seismovolcanic monitoring, minimizing the effects of unknown, spurious noise, and discerning recurrent patterns through template matching. By providing deeper insights into pre-eruptive behavior, it may lead to more effective hazard reduction strategies, enhancing public safety around active volcanoes.

26 Introduction

27 An important aspect of volcanology is to monitor for impending volcanic eruptions and their po-
28 tential hazards. However, this is not easy because seismic data can be contaminated by signals of
29 unknown or spurious origin, or known to be unrelated to the volcanic system. Therefore, there is
30 a need for new methods that can detect and interpret the subtle changes in seismic activity that
31 occur at different timescales before an eruption.

32 Eruption precursors are defined here as observable signals that consistently precede volcanic
33 eruptions. They can be identified in seismic data (e.g., Chastin and Main, 2003; Pesicek et al., 2018;
34 Roman and Cashman, 2018), gas measurements (e.g., Aiuppa et al., 2007; de Moore et al., 2016;
35 Burton et al., 2021), and satellite observations (e.g., Oppenheimer et al., 1993; Biggs et al., 2022;
36 Girona et al., 2021). Eruption precursors can evolve over timescales ranging from years (Caudron
37 et al., 2019, 2022; Girona et al., 2021) to weeks (Pesicek et al., 2018; Roman and Cashman, 2018;
38 Ardid et al., 2022; Rey-Devesa, Benítez, et al., 2023; Rey-Devesa, Prudencio, et al., 2023) to hours
39 (Larsen et al., 2009; Dempsey et al., 2020; Cabrera et al., 2024).

40 Not every signal before an eruption should be classified as a precursor. Ardid et al. (2022) intro-
41 duced specific criteria to define an eruption precursor, emphasizing (1) that the signal consistently
42 recurs before multiple eruptions, and (2) it should be differentiable from periods of noneruptive
43 unrest or repose, recurring less frequently during these periods. The second criterion, focusing on
44 distinctiveness, is important for linking the pre-eruptive signal with the impending eruption.

45 One approach for exploring precursors over longer timescales is template matching. Ardid et
46 al. (2022) adapted this approach by combining the 2-day moving median time-series feature of the
47 displacement seismic amplitude ratio (DSAR) with a 28-day template recorded prior to the 2019
48 eruption at Whakaari volcano; the template appeared to recur with some similarity prior to eruptions
49 at Ruapehu and Tongariro volcanoes, also in New Zealand. The precursor was interpreted to reflect
50 a weeks-long consolidation of a low permeability mineralogical seal over an active hydrothermal
51 system, which contributed to the conditions conducive of phreatic eruptions. The same precursor
52 was used in a template matching investigation of fluid release events in volcanoes with crater lakes
53 (Ardid et al., 2023). However, to date, there has not been a systematic investigation of the timescales
54 of pre-eruptive unrest using template matching.

55 In this study, we apply template matching over multiple timescales ranging from 24 hr to 2
56 months to identify eruption precursors by comparing specific time-series elements that regularly
57 preceded eruptions and the timescales over which these are observable. We tested our approach on
58 four volcanoes: Copahue (Chile/Argentina), Pavlof (Alaska, United States), Bezymianny (Russia),
59 and Whakaari (New Zealand). We chose these volcanoes because they have extensive, continuous
60 seismic data sets with five or more eruptions recorded at each volcano. Our goal is to establish
61 whether there are specific timescales or eruption templates that are more informative of future
62 eruptions, and to explore how these timescales vary between volcanoes.

63 Methodology

64 Data preprocessing

65 In this study, we obtained data from a single station at each of the four target volcanoes (Fig. S1,
66 available in the supplemental material to this article; see Table S1 for more details on the stations).
67 In choosing stations for each volcano, our criteria prioritized proximity to the vent and number of
68 eruptions contained in the record. The primary criterion was the number of recorded eruptions
69 because our methods perform better when there are larger data sets to train on. We processed these
70 data as follows.

71 First, we extract the vertical velocity component of the daily waveform data and remove the in-
72 strument response. Although the removal of theoretical instrument response aims to ensure accurate
73 and consistent seismic data analysis, it is important to acknowledge that this alone may not guard
74 against all data artifacts. These artifacts could persist in both the training and testing data sets,
75 leaving room for uncertainty regarding their complete elimination. Nevertheless, the algorithm's
76 performance, as observed, presumably includes the detrimental effects of these artifacts. This sug-
77 gests that despite potential limitations in data quality, evidence of forecasting skill is still apparent.
78 However, it remains a priority to explore new avenues for improving the quality of data used in
79 training sets.

80 We then postprocess four data time series (Fig. S2): (1) real-time seismic amplitude measure
81 (RSAM; Endo and Murray, 1991), a 10-min average of the absolute velocity band-pass filtered
82 between 2 and 5 Hz; (2) medium frequency (MF); and (3) high frequency (HF), which are the

83 same as RSAM, but band-pass filtered at 4.5–8 and 8–16 Hz frequency bands, respectively; and (4)
 84 displacement seismic amplitude ratio (DSAR; Caudron et al., 2019), computed as the ratio between
 85 MF and HF bands when recalculated for absolute vertical displacement (Dempsey et al., 2020). The
 86 4.5 Hz threshold between RSAM and MF reflects an assumption that tremor predominantly radiates
 87 energy at frequencies below 4.5 Hz (Caudron et al., 2019), which allows us distinguish magma or
 88 fluid source processes from other attenuation related effects, for example, permeability changes
 89 because of shallow sealing. The DSAR parameter tracks transitions between frequency bands over
 90 time, hence capturing evolving spectral characteristics. Time-varying parameters are then used to
 91 compute feature time series, which have the potential to amplify underlying patterns in the seismic
 92 signal. To isolate a continuous tremor signal, we minimize signatures from regional earthquakes
 93 by applying an outlier detection algorithm to the raw waveform data before computing RSAM,
 94 MF, and HF (Dempsey et al., 2022). This algorithm identifies and removes outliers by excising a
 95 2-min envelope from the raw velocity trace whenever its value exceeds three standard deviations
 96 above the mean in the 10-min window. Although effective in removing signal contamination from
 97 most short-duration events, contamination from teleseismic waves of large-magnitude, long-duration
 98 subduction earthquakes can occur. To address this, a two-window (20 min) moving minimum is
 99 applied (Dempsey et al., 2022).

100 The outlier detection algorithm is intended to filter out regional earthquakes, and its calibration
 101 was based on data specifically gathered from regional earthquakes (Dempsey et al., 2022). However,
 102 it filters out all earthquakes without discriminating, including volcano tectonic (VT)—this is because
 103 our research interest is in online real-time algorithms that can operate without human assistance
 104 (e.g., classifying VT earthquakes)—future work could include adding automated discrimination of
 105 VTs from regional earthquakes (Ibáñez et al., 2008; Titos et al., 2018; Falcin et al., 2021).

106 Once calculated, the four data time series are normalized to unit lognormal distributions (Fig.
 107 S2). This normalization, known as z-score normalization, improves classification modeling and has
 108 been demonstrated to enhance the discrimination of precursors (Dempsey et al., 2022).

109 **Time-series feature calculation**

110 Each of the four data streams—RSAM, MF, HF, and DSAR—is further processed by extracting a
 111 large set of time-series features (Rouet-Leduc et al., 2019; Christ et al., 2018; Dempsey et al., 2020;

Ren et al., 2020). Each time-series feature that is created is a functional transformation of the data stream.

Time-series features are constructed from an initial time series by performing calculations on rolling windows. At each index in the initial time series, a backward-looking window of fixed length is constructed, bracketing a fixed subset of the data. For this window, a feature value is calculated. Examples of features include the maximum value of the data set, the median, or the number of peaks. The window is then shifted forward to the next data index, overlapping with the previous window, and a new feature value is calculated. The output is a feature time series with the same resolution as the original data, but now transformed according to the particular feature. In this study, we have calculated 700 time-series features using the Python package “tsfresh,” which computes various statistical moments, distribution properties such as energy or entropy, temporal features such as trend, and frequency domain features such as Fourier coefficients.

The time-series features are computed for each normalized 10-min data point, using a backward-looking 48-hr window (Dempsey et al., 2020, 2022; Ardid et al., 2022). The selection of a 48-hr window is deliberate, being able to capture short-term variations in the volcano state while also being long enough for attempts to filter potential noise sources via the feature calculation. Although Dempsey et al. (2020) investigated feature window lengths ranging from 12 hr to 5 days at Whakaari, here, we explore much longer timescales by incorporating a final step of template matching.

Template matching

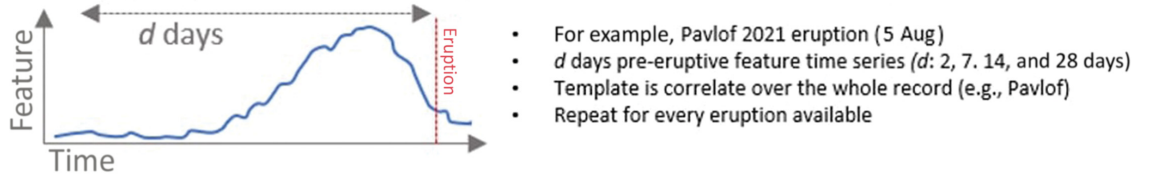
In this step, we seek a reference signal or pattern, referred to as the template. Here, a template is always defined as fixed-length subset of a time-series feature with the template endpoint being an eruption. This allows us to test the hypothesis that the template is reflective of a physical process or volcano state preceding an eruption. Ideally, the template is also interpretable within some suitable conceptual framework. Having defined a template, we then test whether it reappears frequently within the rest of the feature time series. Because there is likely to be natural variability in the underlying process or unknown, spurious signals not associated with the volcanic system, we expect the match between the template and rest of the feature time series will never be perfect.

Central to the application of template matching is the measurement of similarity between the template and a different part of the feature time series. Here, we quantify similarity using the

correlation coefficient, which is a numerical measure of the similarity between two signals' shapes once their magnitudes have been normalized. An alternative approach is dynamic time warping (Laperre et al., 2020; Li et al., 2020), which allows the template signal to be deformed on its time axis to improve the match. For a given feature, template eruption, and template length, we calculate the time series from the cross correlation of the template over the feature times series (Fig. 1). A value of 1 indicates a perfect match, which always occurs when the template is correlated with itself. A value of 0 indicates not correlation, whereas a -1 indicates perfect anticorrelation.

Template matching: Cross correlation on the dsar median

(a) **Template:** d days pre-eruptive feature time series computed from data-streams (RSAM, MF, HF, and DSAR)



(b) **Cross correlation:** of the template over the over the record

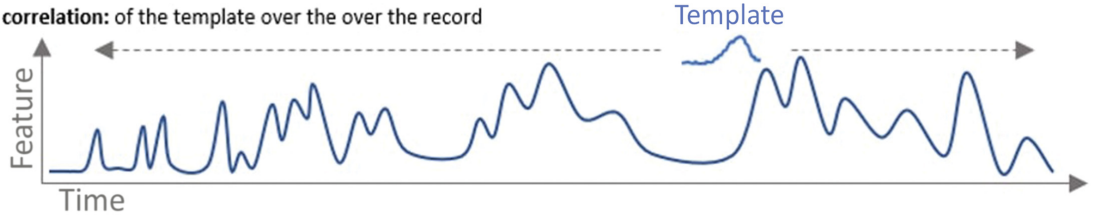


Figure 1: Scheme of template matching by using a template derived from the d days pre-eruptive features from an eruption (all eruptions are tested as template). (a) Example of template, which is a short time series, considered pre-eruptive. We explore multiple template lengths ranging from 1 to 60 days. (b) Running correlation of the template over the whole feature time series. The template matching process is based on comparing the template with the long time series. In this way, when it is repeated over time, we can say that the underlying physical process in the volcano is repeated, and that would be the future element to control. The color version of this figure is available only in the electronic edition.

Template matching may yield suboptimal results in the presence of a low signal-to-noise ratio. Distinguishing a precursor from the background signal is particularly challenging in volcanic tremor analysis (Steinke et al., 2023). This is why we attempt template matching for each of the different feature time series. A second challenge is that we neither know how many precursors may be present, nor what their durations are. Therefore, we have examined template lengths of 1, 2, 7, 14, 21, 28,

and 60 days, for each eruption at each volcano. For Copahue, with six eruptions (as templates), four data streams, 700 features per data stream, and seven template lengths, this results in 117,600 templates for testing.

Template selection by ranking p-values

To assess the performance of the candidate templates based on their differentiability (Ardid et al., 2022), we computed p-values for each correlation coefficient time series. The p-value was calculated from a two-sample Kolmogorov–Smirnov test, which tests whether correlation coefficient (CC) values prior to the nontemplate eruptions are significantly different from the background distribution of CC values.

Then, we computed adjusted p-values using the false discovery rate test (Benjamini and Hochberg, 1995). We then rank these adjusted templates by p-value and retained all those with a value less than 0.05. The p-value analysis helped us discern families of features with superior performance in discriminating pre-eruptive data.

Template selection by receiver operating characteristic (ROC) curves

We also computed ROC curves for each of the correlation coefficient time series suggested in the previous section to be significant (Fig. 2). ROC curves gauge the discriminability of classification models because they attempt to separate pre-eruptive and noneruptive signals (Ardid et al., 2024). These curves plot the true positive rate (sensitivity) against the false positive rate (1 specificity) for varying threshold values of CC. The area under the ROC curve (AUC) is a single measure of model performance across all thresholds, with values nearing 1 indicating optimal performance.

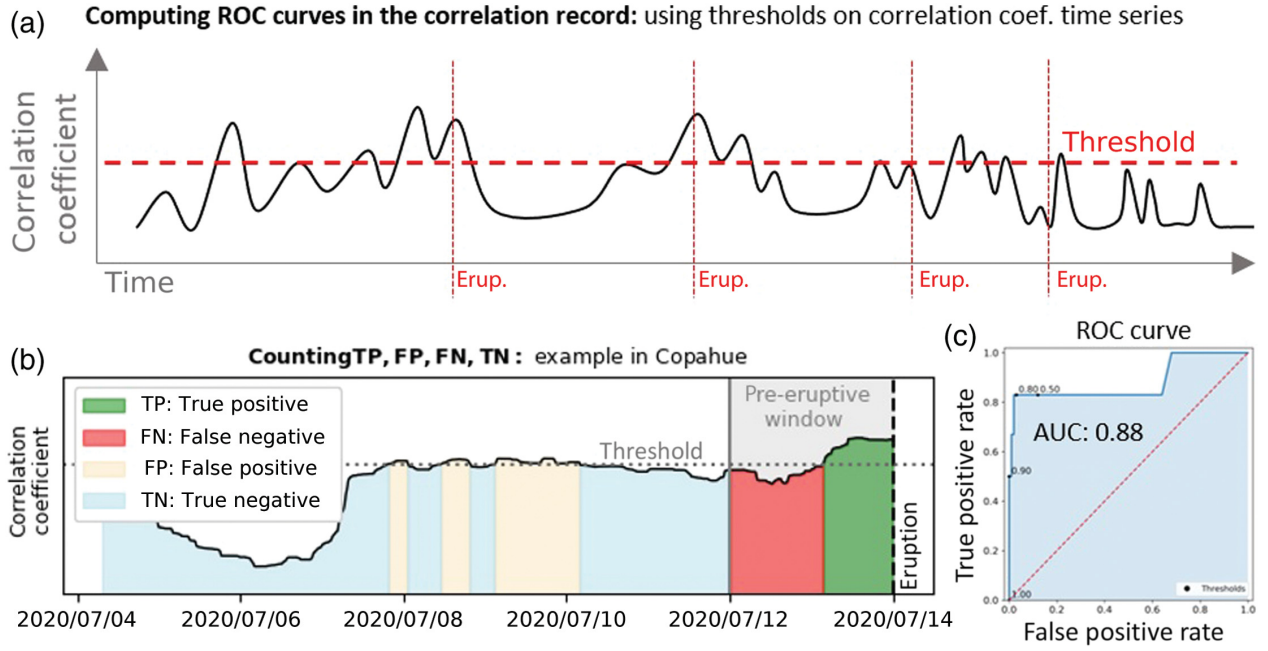


Figure 2: Scheme of performance evaluation of the template matching using receiver operating characteristic (ROC) curves. (a) The resulting correlation coefficient time series (template matching output) for a template with an arbitrary. (b) Illustration of how the correlation coefficient time series, is used to determine true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN) using a threshold and a pre-eruption eruption window. TP represents correctly predicted eruptions when the model output surpasses the threshold within the pre-eruptive window. FN indicates missed eruptions when the output fails to reach the threshold during this window. FP signifies incorrectly classified eruptions when the threshold is surpassed outside the pre-eruptive window, whereas TN denotes correctly classified noneruptive data. The choice of threshold greatly influences the counts of TP, FP, FN, and TN. (c) A hundred threshold values between 0 and 1 are used to count TP and FP rates for computing receiver operator curves (ROC) and their area under the curve (AUC). AUC is used as metric to evaluate the model performance when classifying as a whole (both pre-eruptive and noneruptive data) across all possible thresholds. The color version of this figure is available only in the electronic edition.

The construction of an ROC curve for a correlation coefficient time series (Fig. 2) involves (1) selecting a threshold value of CC (e.g., 0.5); (2) scoring a true positive (TP) or true negative (TN) in the time window before the eruption for the period defined by the template timescale, labeling “1”

for TP or “0” for TN; (3) scoring a false positive (FP) when the correlation coefficient surpasses the threshold outside the pre-eruptive data; (4) scoring a false negative (FN) for all other mismatches; (5) calculating the true positive rate (TPR) and false positive rate (FPR) as $TPR = TP/(TP + FN)$ and $FPR = FP/(FP + TN)$; (6) repeating steps (1–5) using 100 evenly spaced thresholds between 0 and 1 to form the ROC curve; and finally (7) computing AUC as the integral of the ROC curve. A higher AUC signifies a template that is better at discriminating other eruptions at the same volcano while maintaining a low FP rate.

Results and Discussion

Ranking template differentiability using p-values

P-values are calculated across all data streams, features, volcanoes, eruption templates, and timescales; this resulted in 431,200 p-values (following the Template selection by ranking p-values section). After ranking, the features are ~ 5000 potential features. Given the very large number of statistical tests, we provide a list for each volcano comprising the templates and their respective performance (see the supplemental material).

Notable contributors we found feature families such as Fourier coefficients and autoregression coefficients. Figure 3 shows the time series of two features of those families before eruptions at the four volcanoes. The graph visually demonstrates the ascending trend in these feature values prior to eruptions.

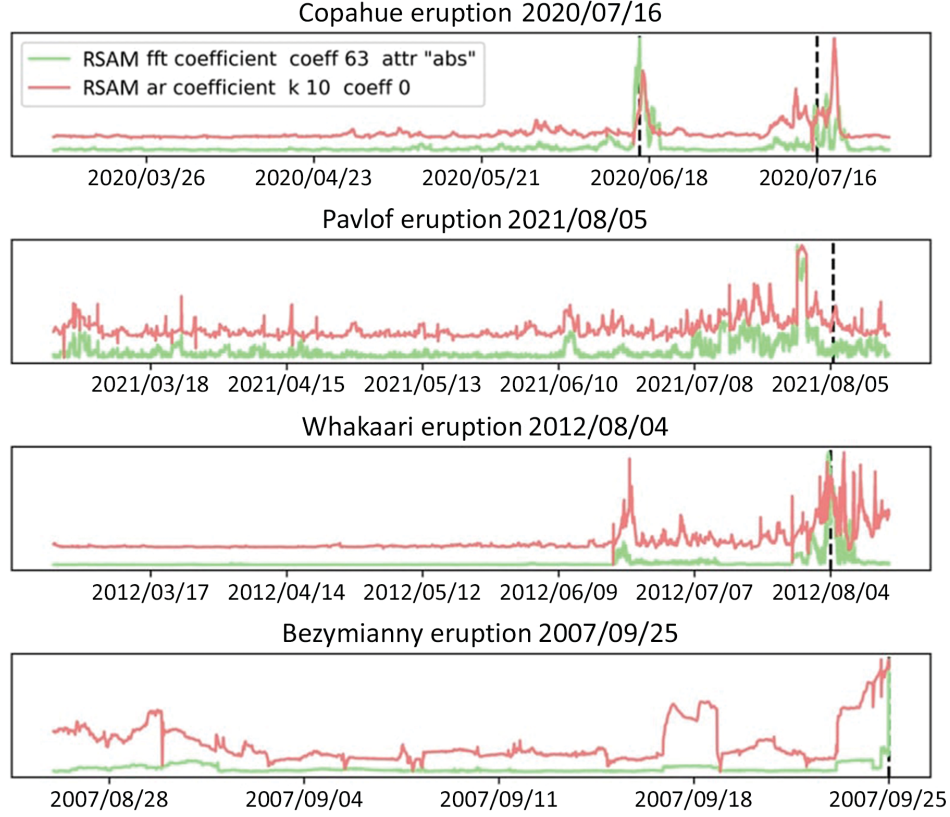


Figure 3: Time-series representation of two feature values before eruptions at the labeled volcanoes (titles), identified as significant in distinguishing between noneruptive and eruptive conditions based on p-values and AUC metrics. The graph illustrates the increasing trend in these feature values preceding eruptions, in accordance with the patterns captured by the p-values and AUC metrics. The color version of this figure is available only in the electronic edition.

Ranking template differentiability using ROC curves

In this evaluation directed at ranking template performance in discerning between noneruptive and eruptive states using ROC curves, we also provide a short list for each volcano containing the templates and their corresponding performance based on the AUC metric (refer to the supplemental material). Figure 4 illustrates examples of various ROC curves for different feature time series on the explored volcanoes. These ROC curves, along with their associated AUC values, provide insights into which parameters (features) harbor more differentiable information for distinguishing between noneruptive and eruptive conditions.

The ROC curves depicted in the figure demonstrate the efficacy of the model in binary classi-

fication, as indicated by the AUC. A higher AUC value, closer to 1, signifies strong classification ability, effectively distinguishing between positive and FP outcomes. Specifically, in this analysis, the ROC curves evaluate the time-series correlation coefficient resulting from template matching. A notable AUC for individual templates indicates a distinctive pre-eruptive pattern unique to eruptions, effectively separating them from noneruptive periods within the signal.

Insights into informative pre-eruptive timescales and frequency bands

From the 5% lowest p-values templates (see the Ranking template differentiability using p-values section), we examined which template timescales had the highest differentiability among the four volcanoes (Fig. 5a). For all volcanoes, we found 14-day time windows were discriminating of pre-eruptive behavior, although this timescale was most dominant for Copahue and Bezymianny. For both Pavlof and Whakaari, the 1-day timescale was important whereas, for Pavlof, 28-days was also prevalent. This implies that Pavlof may exhibit precursors over the course of a month, contrasting with the other volcano studied in which precursors are usually observed on shorter timescales (Fig. 5a).

The volcanoes also exhibit some dissimilarity in the relative importance of different data streams in the pre-eruptive period (Fig. 5b). For example, at Copahue, the RSAM band (2–5 Hz) dominates the significant templates, whereas for Pavlof, the four bands make a more balanced contribution, although MF carries the most weight. Both Bezymianny and Whakaari favor the RSAM band with minimal contribution from DSAR.

The occurrence of differentiable template time lengths and data streams provides initial insights into the methodology’s capability to differentiate between various time lengths of templates, highlighting those with higher discriminative potential for precursor identification.

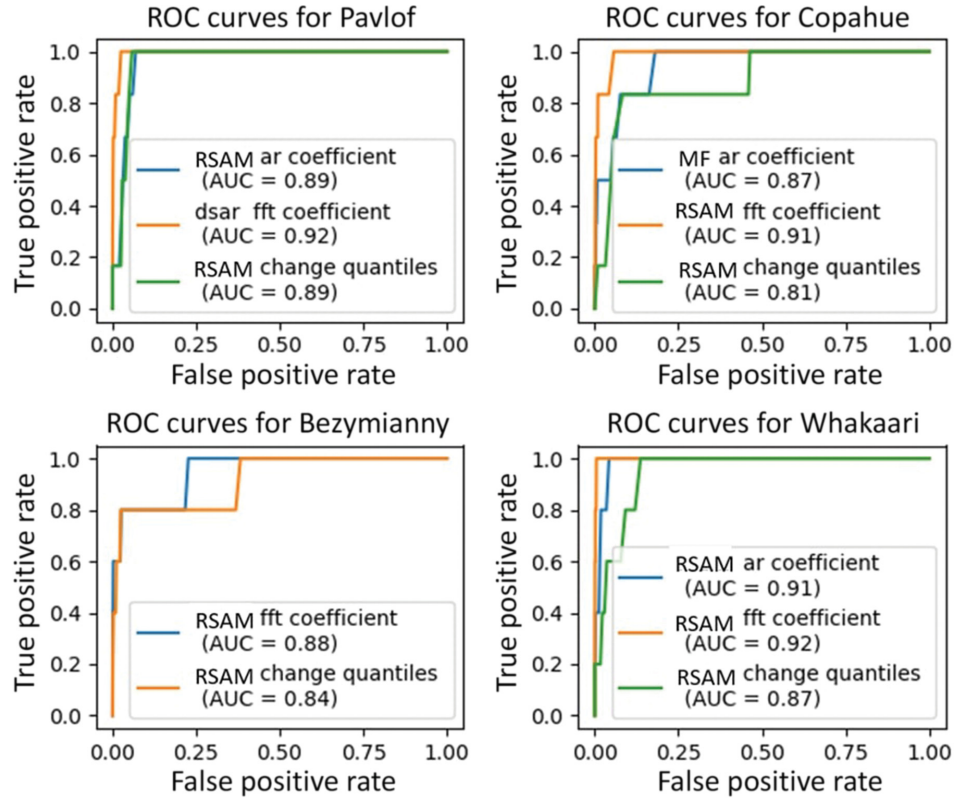


Figure 4: Example of ROC curves for three relevant feature time-series template in Pavlof, Copahue, Bezymianny, and Whakaari volcanoes. These features have been identified to be relatively good (compared to other features) for distinguishing between noneruptive and eruptive conditions using p-values and AUC metrics. The color version of this figure is available only in the electronic edition.

The divergent distributions of template data streams and time lengths across the four volcanoes could be attributed to various factors, including the volcano type, geological setting, content of fluids, magma composition, and depth of the magma reservoir. For example, Copahue is a basaltic-andesitic volcano that lies within a complex tectonic setting including major tectonic faults as well as caldera structures (Farías et al., 2023). The volcano features a shallow hydrothermal system above a magma chamber at ~ 7 km depth (Lundgren et al., 2017). Eruptive activity, dominated by phreatomagmatic events, are often preceded by changes in the seismic activity on timescales of days to weeks (Farías et al., 2023). Similarly, Whakaari features an established shallow hydrothermal system with eruptive activity dominated by phreatomagmatic events (Kilgour et al., 2021). Eruptive products from the volcano have indicated relatively shallow magma emplacement (< 1 km) before eruptive activity (Kilgour et al., 2021). In contrast, there are no indications of a hydrothermal

system at either Bezymianny or Pavlof. The former volcano typically exhibits effusive lava dome building behavior, interrupted by episodes of strong explosive eruptions (Girina, 2013). The explosive eruptions may be preceded by increased seismic activity up to 40 days beforehand (Koulakov et al., 2021). Seismicity also indicates a complex subsurface system with multiple magma chambers at 2–8 km depth (Koulakov et al., 2021) and interactions with other nearby volcanoes (Journeau et al., 2022). Finally, Pavlof is a basaltic–andesitic volcano that exhibits Strombolian explosion activity, which can rapidly escalate to intense lava fountaining within 1 day (Fee et al., 2017). Notably, shallow (<50 km) seismic events are relatively rare at the volcano (McNutt, 1987).

Apart from the discrepancies in template time length distribution, the results illustrate that the template matching approach and filtering metrics recover differentiable features, enabling a relative comparison of template time lengths that are assumed to be associated with the duration of potentially triggering eruptive processes. This methodology not only provides insights into volcanic activity but could also be extended to other geothermal phenomena. For instance, similar techniques could potentially be applied to hydrothermal eruptions in geysers (Ardid et al., 2019), in which understanding the timing and triggers of eruptions is crucial. Furthermore, this approach might be beneficial in studying geothermal fields (Ardid et al., 2021), offering a valuable tool for monitoring these systems based on the duration of observed thermal and seismic activities.

Graphical example of template matching and feature interpretability

Among the most relevant precursor features identified by our study is a family called “change quantiles.” Figure 6 illustrates how one example of this feature, termed “rsam change quantiles (0.8 to 1.0) absolute mean,” is constructed. First, a quantile mask is applied to select only the highest 20% of RSAM values in a 2-day window. Second, for the selected values, a difference is computed between adjacent 10-min values. Third, the mean of the absolute value of these differences is calculated.

Although this feature appears to be statistically significant, its physical significance is not yet clear. Nevertheless, we propose one physical explanation for the signal, primarily to illustrate how explainable features of machine learning could be used to advance new hypotheses for later testing. In this way, machine learning can complement robust observational science, which upholds the role of data collection and analysis to test new hypotheses.

Our proposed explanation of this feature is as follows. RSAM, as a measure of average seismic

velocity, is a proxy for the seismic energy density that is radiated by some generating source. For a volcano experiencing unrest prior to phreatic eruptions, this generating source may be related to shallow degassing of magma (i.e., Jaupart and Vergnolle, 1988; Journeau et al., 2022) or other movement of fluids in the subsurface (i.e., Ferrick et al., 1982; Yamada et al., 2021). Then, by applying a 0.8 to 1.0 quantile mask, we are selecting for further study the highest energy part of the signal and, by implication, the highest energy source processes occurring during the 2-day feature window. Another way to think about this is that the quantile mask has binarized the time series according to a state variable of the volcano: in this case, high and low radiated energy.

The second step in the feature calculation associates a characteristic of the signal with mask-selected state. In this case, we are computing 10-min absolute differences in the RSAM taking their average value during high-energy periods. This quantity will be high when there is a lot of short-term volatility in how much energy is being radiated from the source process. Attributing a physical meaning to this volatility is difficult and likely depends on the nature of that source. For instance, it could be associated with rapid fluctuations in degassing if the tremor signal has a flux origin (Girona et al., 2019). Alternatively, it may relate to a complex sequence of crack opening and closure in a fracture network subject to varying fluid pressure (Konstantinou and Schlindwein, 2003; Kennedy et al., 2020; Li et al., 2022).

Figure 7 shows a 14-day template of the change quantiles feature that was extracted prior to the first Copahue eruption (Fig. 7a). It is characterized by a 13-day quiet period followed by a rapid increase in the feature value in the day before the eruption. This template has then been overlaid on the feature time series prior to the other five recorded eruptions, highlighting the time of maximum cross correlation. Qualitatively, the match is reasonable for all but the fourth eruption. In contrast, the raw RSAM time series only exhibits anomalous values prior to the first two eruptions.

Although this template recurs before the majority of eruptions (Fig. 7b,c,e,f), its absence prior to the fourth eruption is notable (Fig. 7d). This discrepancy highlights that recurrent templates may not reliably recur prior to every eruption at a volcano. This could be due to natural variation in the shape or strength of pre-eruptive signals. Another explanation for the discrepancy is that this fourth eruption was reported by local residents as a minor ash plume without further alerts from official agencies. Because all other eruptions were reported by agencies, this event may have been smaller than other eruptions and therefore had dissimilar precursors, if any.

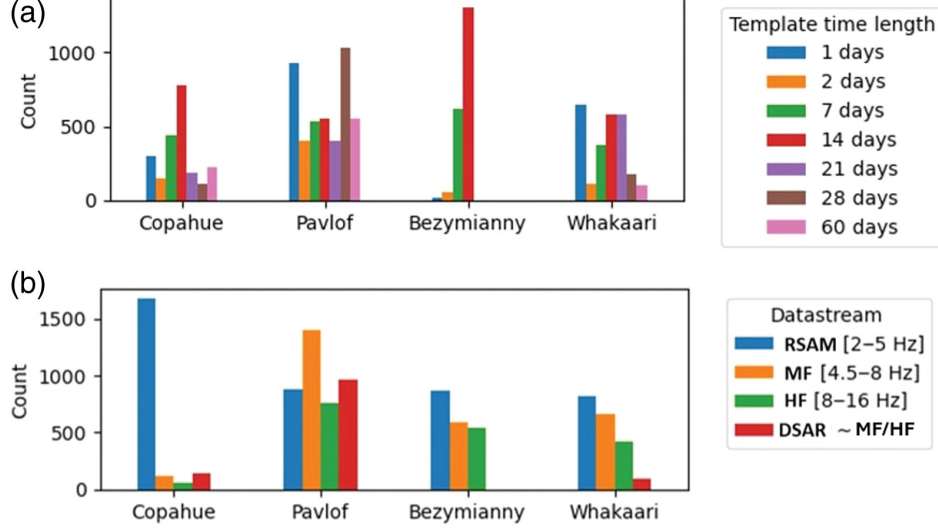


Figure 5: (a) Distribution of occurrences of different template time lengths after the filtering feature time series on distinguishing between noneruptive and eruptive conditions using p-values and AUC metrics, for Copahue, Pavlof, Bezymianny, and Whakaari volcanoes. The time lengths are divided into seven categories: 1, 2, 7, 14, 21, 28, and 60 days. (b) Distribution of occurrences of different data streams after template selection using p-values and ROC curves for distinguishing between pre-eruptive differentiable and nondifferentiable patterns: for Copahue, Pavlof, Bezymianny, and Whakaari volcanoes. The data streams correspond to seismic amplitude measurement real-time seismic amplitude measure (RSAM), medium frequency (MF), and high frequency (HF), and for the ratio seismic measurement displacement seismic amplitude ratio (DSAR). The plot suggests which frequency bands are more informative when looking for precursory activity. The color version of this figure is available only in the electronic edition.

Unfortunately, not all the features used in this analysis are interpretable. For instance, the autoregression coefficient (Fig. 4), which describes a time series as a linear combination of its past values (up to n lagged observations), lacks physical interpretability. The case is similar for the Fourier coefficients (Fig. 4), which highlight harmonics of greater intensity preceding eruptions. However, assigning a physical variable that correlates with these coefficients proves to be challenging. Nevertheless, the value of time-series features lies more in their ability to capture subtle patterns in the signal prior to eruptions. These patterns can be used with statistical tools to forecast future system behaviors, for example, by training machine learning models (e.g., Dempsey et al., 2020,

2022; Ardid et al., 2024).

Although the analysis of seismic parameters provides insights into volcanic dynamics, it is important to acknowledge the inherent limitations. The direct association of seismic parameters with specific physical phenomena can be speculative due to the complexity of volcanic systems. Therefore, although discussing the interpretability of features like the “rsam change quantiles (0.2–0.8) mean,” it is relevant to recognize the speculative nature of attributing specific physical processes solely based on seismic data. The multifaceted influences outlined highlight the interplay of various factors shaping volcanic activity. Consequently, although certain features may capture intrinsic state variations and extrinsic behavioral shifts within volcanic systems, the broader context underscores the need for cautious interpretation, and a multidisciplinary approach in understanding and monitoring volcanic activity.

We recognize that comprehending a monitoring parameter enhances our ability to make informed decisions, but the primary focus of a monitoring parameter should be its capability for identifying precursors. Whether or not we fully understand it is of secondary importance.

Conclusions

Our study proposes a methodology for detecting eruption precursors across various timescales at different volcanoes, enhancing our understanding of volcanic activity and hazard mitigation strategies. Through multitimescale template matching, we identified statistically significant eruption precursors at Copahue, Pavlof, Bezymianny, and Whakaari volcanoes, shedding light on the diverse temporal dynamics preceding eruptions.

Our analysis underscores the importance of considering different timescales in volcano monitoring, as each volcano exhibits distinct patterns of pre-eruptive behavior. Although 14-day time windows emerge as discriminative for Copahue and Bezymianny, Pavlof shows importance at 1-day and 28-day scales, highlighting varied eruption precursors across volcanoes.

The “change quantiles” feature family, highlighted in our study, unveils subtle variations in seismic activity preceding volcanic eruptions. Specifically, the feature “rsam change quantiles (−0.8 to 1.0) absolute mean,” captures changes in seismic energy density before eruptions. This interpretation suggests a link between intense seismic energy and processes like magma degassing or fluid movement

underground. By isolating high-energy segments within the signal and examining their fluctuations, we gain insights into potential precursory patterns. Although interpreting patterns in features poses challenges, they offer valuable clues for forecasting volcanic activity and generating hypotheses for further investigation.

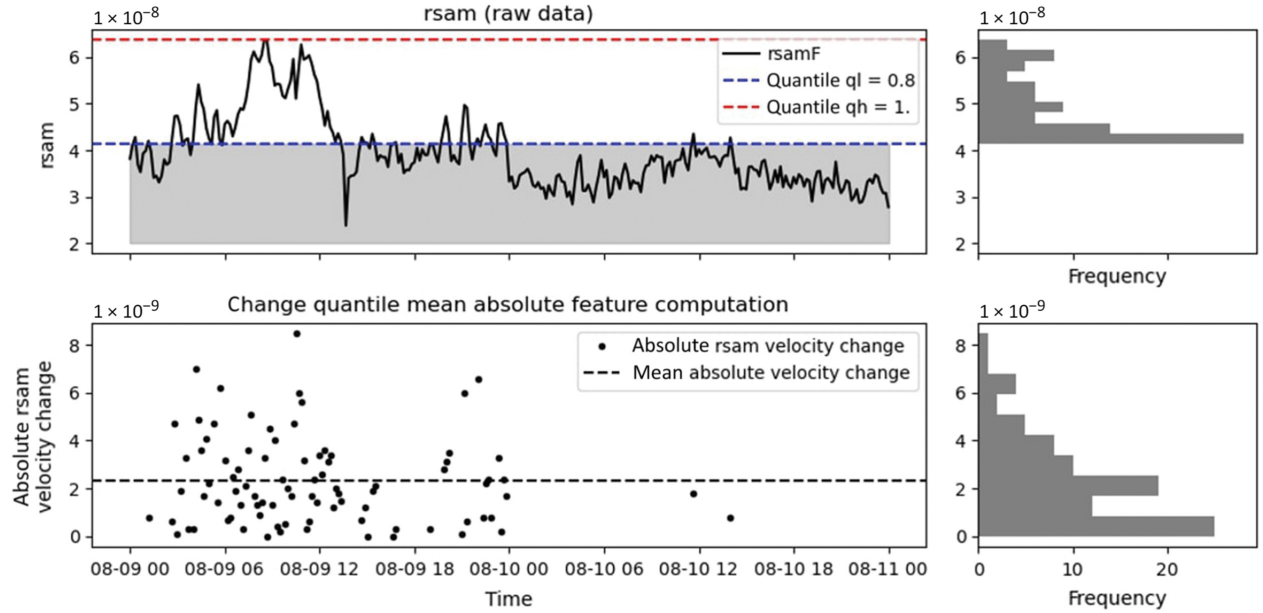


Figure 6: An example of how rsam change quantiles mean is calculated, showing the raw RSAM time series on a window basis (2 days, 288 samples), the distribution used to define quantiles (top right), the RSAM time series including only values in the body of the distribution. The plot at the bottom shows how the selected values are differenced to obtain a measure of how average velocity is changing. The mean of these differences gives the feature value for this 2-day window. The color version of this figure is available only in the electronic edition.

Moving forward, the integration of multitimescale template matching into real-time volcano monitoring systems holds promise for enhancing eruption forecasting capabilities. By leveraging machine learning approaches and putting these tools into operation, we can further refine our ability to predict volcanic eruptions and improve safety measures in volcanic regions. Continued research and application of these methodologies are essential for advancing volcanic monitoring and hazard mitigation efforts worldwide.

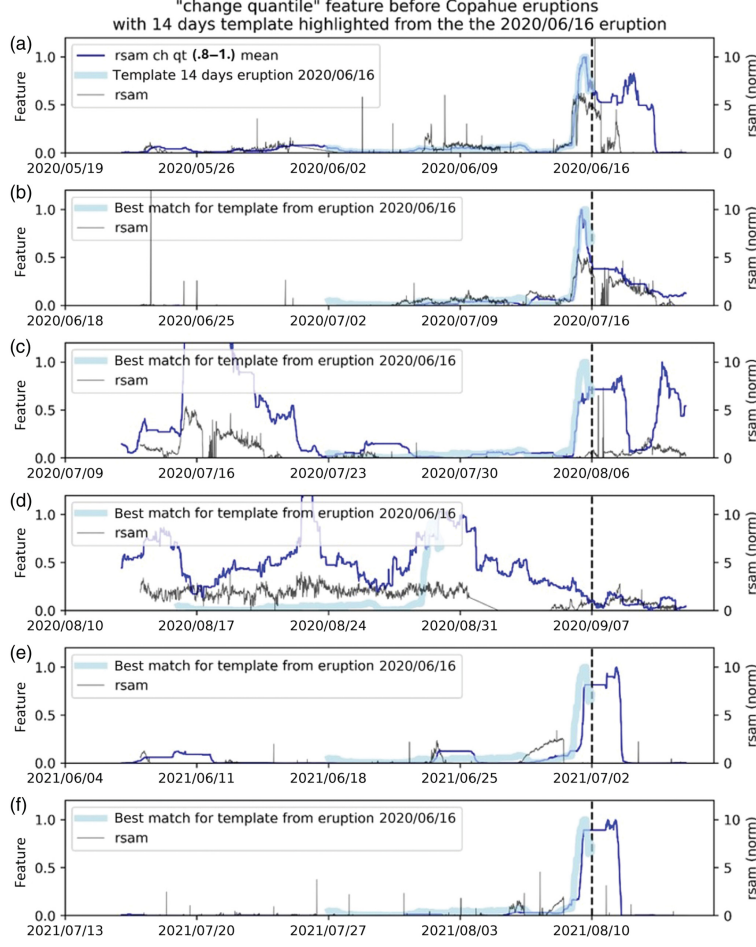


Figure 7: Temporal evolution of the “change quantiles (0.8–1.0) mean” feature in the RSAM band leading up to six eruptions in Copahue, along with the rsam data for comparison. The rsam values have been standardized using the z-score, meaning they have been adjusted by subtracting the historical mean and dividing by the standard deviation. (a) First eruption where the template is extracted. In cyan (light blue) the 14-day template extracted before the first eruption (notice how it matches the first eruption, as it is expected). The template from the first eruption is overlaid on the other five eruptions at the point of maximum coincidence in template matching, revealing a clear match for most eruptions (b,c,e,f) except for the fourth eruption (d). The color version of this figure is available only in the electronic edition.

Data and Resources

The codes required to replicate the results of this study are available at https://github.com/aardid/temp_matching_srl, released under the MIT license. Codes are provided in the supplemental material to compute the receiver operating characteristic (ROC) curves for all three types of foresters. Raw waveform data for New Zealand volcanoes can be downloaded from GeoNet (<https://www.geonet.org.nz/>); and for Alaskan volcanos from Incorporated Research Institutions for Seismology (IRIS) (<http://service.iris.edu>). Both are operable as clients through the International Federation of Digital Seismograph Networks (FDSN) webservice (<https://www.fdsn.org/networks/>). All websites were last accessed in May 2024.

Declaration of Competing Interests

The authors acknowledge that there are no conflicts of interest recorded.

Acknowledgments

The authors acknowledge GeoNet, New Zealand and Alaskan Volcano Observatory (AVO, United States) seismic monitoring system for supplying free access to their volcanic seismic data. The authors wish to thank NZ MBIE UOAX1913 (Transitioning Taranaki to a Volcanic Future) for support of A. A., D. D., and S. J. C. during this study. The authors wish to thank NZ MBIE E7774 (Adapting to climate change through stronger geothermal enterprises) for support of A. A., D. D., and S. J. C. during this study. The authors thank Ivan Melchor and “Instituto de Investigación en Paleobiología y Geología” (UNRN-CONICET), Argentina, for providing the seismological data of the Copahue volcano. O. L. was supported by the NZ MBIE through the Strategic Science Investment Fund (SSIF) Hazards and Risk Management program.

References

- [1] Aiuppa, A., R. Moretti, C. Federico, G. Giudice, S. Gurrieri, M. Liuzzo, P. Papale, H. Shinohara, M. Valenza, and M. L. I. Witt (2007). Forecasting Etna eruptions by real-time observation of volcanic gas composition, *Geology* 35, no. 12, 1115–1118.

- [2] Ardid, A., R. Archer, E. Bertrand, F. Sepulveda, P. Tarits, and D. Dempsey (2021). Heat transfer through the Wairakei-Tauhara geothermal system quantified by multi-channel data modeling, *Geophys. Res. Lett.* 48, no. 8, e2020GL092056, doi: 10.1029/2020GL092056.
- [3] Ardid, A., D. Dempsey, C. Caudron, and S. Cronin (2022). Seismic precursors to the Whakaari 2019 phreatic eruption are transferable to other eruptions and volcanoes, *Nat. Commun.* 13, no. 1, 1–9, doi: 10.1038/s41467-022-29681-y.
- [4] Ardid, A., D. Dempsey, C. Caudron, S. Cronin, B. Kennedy, T. Girona, S. Potter, O. Lamb, C. Miller, L. Cabrera, et al. (2024). Generalized eruption forecasting models using machine learning trained on seismic data from 24 volcanoes, *Res. Square*, doi: 10.21203/rs.3.rs-3483573/v1.
- [5] Ardid, A., D. Dempsey, C. Caudron, S. J. Cronin, C. A. Miller, I. Melchor, D. Syahbana, and B. Kennedy (2023). Using template matching to detect hidden fluid release episodes beneath crater lakes in Ruapehu, Copahue, and Kawah Ijen volcanoes, *J. Geophys. Res.* 128, no. 10, e2023JB026729, doi: 10.1029/2023JB026729.
- [6] Ardid, A., E. Vera, C. Kelly, M. Manga, C. Munoz-Saez, A. Maksymowicz, and F. CulaciatiOrtega- (2019). Geometry of geyser plumbing inferred from ground deformation, *J. Geophys. Res.* 124, no. 1, 1072–1083.
- [7] Benjamini, Y., and Y. Hochberg (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing, *J. Roy. Stat. Soc.* 57, no. 1, 289–300.
- [8] Biggs, J., N. Anantrasirichai, F. Albino, M. Lazecky, and Y. Maghsoudi (2022). Large-scale demonstration of machine learning for the detection of volcanic deformation in Sentinel-1 satellite imagery, *Bull. Volcanol.* 84, no. 12, 100, doi: 10.1007/s00445-022-01608-x.
- [9] Burton, M., C. Hayer, C. Miller, and B. Christenson (2021). Insights into the 9 December 2019 eruption of Whakaari/White Island from analysis of TROPOMI SO₂ imagery, *Sci. Adv.* 7, no. 25, eabg1218, doi: 10.1126/sciadv.abg1218.
- [10] Cabrera, L., A. Ardid, I. Melchor, S. Ruiz, B. Symmes-Lopetegui, J. Baez, F. Delgado, P. Martinez-Yáñez, D. Dempsey, and S. J. Cronin (2024). Eruption forecasting model for Copahue

volcano (southern Andes) using seismic data and machine learning: A joint interpretation with geodetic data (GNSS and InSAR), *Seismol. Res. Lett.* doi: 10.1785/0220240022.

[11] Caudron, C., Y. Aoki, T. Lecocq, R. De Plaen, J. Soubestre, A. Mordret, L. Seydoux, and T. Terakawa (2022). Hidden pressurized fluids prior to the 2014 phreatic eruption at Mt Ontake, *Nat. Commun.* 13, no. 1, 6145, doi: 10.1038/s41467-022-32252-w.

[12] Caudron, C., T. Girona, B. Taisne, Suparjan, H. Gunawan, Kristianto, and Kasbani (2019). Change in seismic attenuation as a long-term precursor of gas-driven eruptions, *Geology* 47, no. 7, 632–636.

[13] Chastin, S. F., and I. G. Main (2003). Statistical analysis of daily seismic event rate as a precursor to volcanic eruptions, *Geophys. Res. Lett.* 30, no. 13, doi: 10.1029/2003GL016900.

[14] Christ, M., N. Braun, J. Neuffer, and A. W. Kempa-Liehr (2018). Time series feature extraction on basis of scalable hypothesis, *Neurocomputing* 307, 72–77.

[15] de Moore, J. M., A. Aiuppa, J. Pacheco, G. Avard, C. Kern, M. Liuzzo, and T. P. Fischer (2016). Short-period volcanic gas precursors to phreatic eruptions: Insights from Poás volcano, Costa Rica, *Earth Planet. Sci. Lett.* 442, 218–227.

[16] Dempsey, D. E., S. J. Cronin, S. Mei, and A. W. Kempa-Liehr (2020). Automatic precursor recognition and real-time forecasting of sudden explosive volcanic eruptions at Whakaari, New Zealand, *Nat. Commun.* 11, 3562, doi: 10.1038/s41467-020-17375-2.

[17] Dempsey, D. E., A. W. Kempa-Liehr, A. Ardid, A. Li, S. Orenia, J. Singh, A. J. Tyler, and S. J. Cronin (2022). Evaluation of short-term probabilistic eruption forecasting at Whakaari, New Zealand, *Bull. Volcanol.* 84, no. 10, 1–26.

[18] Endo, E. T., and T. Murray (1991). Real-time seismic amplitude measurement (RSAM): A volcano monitoring and prediction tool, *Bull. Volcanol.* 53, 533–545, doi: 10.1007/BF00298154.

[19] Falcin, A., J. P. Métaixian, J. Mars, É. Stutzmann, J. C. Komorowski, R. Moretti, M. Malfante, F. Beauducel, J. M. Saurel, C. Dessert, et al. (2021). A machine-learning approach for automatic classification of volcanic seismicity at La Soufrière volcano, Guadeloupe, *J. Volcanol. Geotherm. Res.* 411, 107151, doi: 10.1016/j.jvolgeores.2020.107151.

- [20] Farías, C., J. Lazo, D. Basualto, M. Saavedra, F. Muñoz-Quiroz, L. Zúñiga-Urrea, R. Martínez-Bravo, I. Huentenao-Inostroza, and R. Sáez-Opazo (2023). One decade of b-value variations from volcano-tectonic seismicity as an early indicator of episodes of crisis in a volcano: the case of Copahue, southern Andes, *Front. Earth Sci.* 11, 1181177, doi: 10.3389/feart.2023.1181177.
- [21] Fee, D., M. M. Haney, R. S. Matoza, A. R. Van Eaton, P. Cervelli, D. J. Schneider, and A. M. Iezz (2017). Volcanic tremor and plume height hysteresis from Pavlof volcano, Alaska, *Science* 355, no. 6320, 45–48.
- [22] Ferrick, M. G., A. Qamar, and W. F. St. Lawrence (1982). Source mechanism of volcanic tremor, *J. Geophys. Res.* 87, no. B10, 8675–8683.
- [23] Girina, O. A. (2013). Chronology of Bezymianny volcano activity, 1956–2010, *J. Volcanol. Geotherm. Res.* 263, 22–41.
- [24] Girona, T., C. Caudron, and C. Huber (2019). Origin of shallow volcanic tremor: The dynamics of gas pockets trapped beneath thin permeable media, *J. Geophys. Res.* 124, 4831–4861, doi: 10.1029/2019JB017482.
- [25] Girona, T., V. Realmuto, and P. Lundgren (2021). Large-scale thermal unrest of volcanoes for years prior to eruption, *Nature Geosci.* 14, 238–241.
- [26] Ibáñez, J. M., E. Del Pezzo, C. Bengoa, A. Caselli, G. Badi, and J. Almendros (2008). Volcanic tremor and local earthquakes at Copahue volcanic complex, southern Andes, Argentina, *J. Volcanol. Geotherm. Res.* 174, no. 4, 284–294.
- [27] Jaupart, C., and S. Vergnolle (1988). Laboratory models of Hawaiian and Strombolian eruptions, *Nature* 331, no. 6151, 58–60.
- [28] Journeau, C., N. M. Shapiro, L. Seydoux, J. Soubestre, I. Y. Koulakov, A. V. Jakovlev, I. Abkadyrov, E. I. Gordeev, D. V. Chebrov, D. V. Droznin, et al. (2022). Seismic tremor reveals active trans-crustal magmatic system beneath Kamchatka volcanoes, *Sci. Adv.* 8, no. 5, eabj1571, doi: 10.1126/sciadv.abj1571.
- [29] Kennedy, B. M., A. Farquhar, R. Hilderman, M. C. Villeneuve, M. J. Heap, S. Mordensky, G. Kilgour, A. Jolly, B. Christenson, and T. Reuschlé (2020). Pressure controlled permeability in a

conduit filled with fractured hydrothermal breccia reconstructed from ballistics from Whakaari (White Island), New Zealand, *Geosciences* 10, no. 4, 138, doi: 10.3390/geosciences10040138.

[30] Kilgour, G., B. Kennedy, B. Scott, B. Christenson, A. Jolly, C. Asher, M. Rosenberg, and K. Saunders (2021). Whakaari/White Island: A review of New Zealand’s most active volcano, *New Zeal. J. Geol. Geophys.* 64, nos. 2/3, 273–295.

[31] Konstantinou, K. I., and V. Schlindwein (2003). Nature, wavefield properties and source mechanism of volcanic tremor: A review, *J. Volcanol. Geotherm. Res.* 119, nos. 1/4, 161–187.

[32] Koulakov, I., P. Plechov, R. Mania, T. R. Walter, S. Z. Smirnov, I. Abkadyrov, A. Jakovlev, V. Davydova, S. Senyukov, N. Bushenkova, et al. (2021). Anatomy of the Bezymianny volcano merely before an explosive eruption on 20.12. 2017, *Sci. Rep.* 11, no. 1, 1758, doi: 10.1029/2011GL046957.

[33] Laperre, B., J. Amaya, and G. Lapenta (2020). Dynamic time warping as a new evaluation for dst forecast with machine learning, *Front. Astron. Space Sci.* 7, 39, doi: 10.3389/fspas.2020.00039.

[34] Larsen, J., C. Neal, P. Webley, J. Freymueller, M. Haney, S. McNutt, D. Schneider, S. Prejan, J. Schaefer, and R. Wessels (2009). Eruption of Alaska volcano breaks historic pattern, *Eos Trans. AGU* 90, 173–174, doi: 10.1029/2009EO200001.

[35] Li, H., J. Liu, Z. Yang, R. W. Liu, K. Wu, and Y. Wan (2020). Adaptively constrained dynamic time warping for time series classification and clustering, *Inform. Sci.* 534, 97–116.

[36] Li, K. L., C. J. Bean, A. F. Bell, M. Ruiz, S. Hernandez, and J. Grannell (2022). Seismic tremor reveals slow fracture propagation prior to the 2018 eruption at Sierra Negra volcano, Galápagos, *Earth Planet. Sci. Lett.* 586, 117533, doi: 10.1016/j.epsl.2022.117533.

[37] Lundgren, P., M. Nikkhoo, S. V. Samsonov, P. Milillo, F. Gil-Cruz, and J. Lazo (2017). Source model for the Copahue volcano magma plumbing system constrained by InSAR surface deformation observations, *J. Geophys. Res.* 122, no. 7, 5729–5747.

[38] McNutt, S. R. (1987). Eruption characteristics and cycles at Pavlof volcano, Alaska, and their relation to regional earthquake activity, *J. Volcanol. Geotherm. Res.* 31, nos. 3/4, 239–267.

- [39] Oppenheimer, C., P. W. Francis, D. A. Rothery, R. W. Carlton, and L. S. Glaze (1993). Infrared image analysis of volcanic thermal features: Lascar volcano, Chile, 1984–1992, *J. Geophys. Res.* 98, no. B3, 4269–4286.
- [40] Pesicek, J. D., J. J. Wellik, S. G. Prejean, and S. E. Ogburn (2018). Prevalence of seismic rate anomalies preceding volcanic eruptions in Alaska, *Front. Earth Sci.* 6, 100, doi: 10.3389/feart.2018.00100.
- [41] Ren, C. X., A. Peltier, V. Ferrazzini, B. Rouet-Leduc, P. A. Johnson, and F. Brenguier (2020). Machine learning reveals the seismic signature of eruptive behavior at Piton de la Fournaise volcano, *Geophys. Res. Lett.* 47, no. 3, e2019GL085523, doi: 10.1029/2019GL085523.
- [42] Rey-Devesa, P., C. Benítez, J. Prudencio, L. Gutiérrez, G. Cortés-Moreno, M. Titos, I. Koulakov, L. Zuccarello, and J. M. Ibáñez (2023). Volcanic early warning using Shannon entropy: multiple cases of study, *J. Geophys. Res.* 128, e2023JB026684, doi: 10.1029/2023JB026684.
- [43] Rey-Devesa, P., J. Prudencio, C. Benítez, M. Bretón, I. Plasencia, Z. León, F. Ortigosa, L. Gutiérrez, R. Arámbula-Mendoza, and J. M. Ibáñez (2023). Tracking volcanic explosions using Shannon entropy at Volcán de Colima, *Sci. Rep.* 13, no. 1, 9807, doi: 10.1038/s41598-023-36964-x.
- [44] Roman, D. C., and K. V. Cashman (2018). Top-down precursory volcanic seismicity: Implications for “stealth” magma ascent and long-term eruption forecasting, *Front. Earth Sci.* 6, 124, doi: 10.3389/feart.2018.00124.
- [45] Rouet-Leduc, B., C. Hulbert, and P. Johnson (2019). Continuous chatter of the Cascadia subduction zone revealed by machine learning, *Nat. Geosci.* 12, no. 1, 75–79.
- [46] Steinke, B., A. D. Jolly, R. Carniel, D. E. Dempsey, and S. J. Cronin (2023). Identification of seismo-volcanic regimes at Whakaari/White Island (New Zealand) via systematic tuning of an unsupervised classifier, *J. Geophys. Res.* 128, no. 3, e2022JB026221, doi: 10.1029/2022JB026221.

- [47] Titos, M., A. Bueno, L. García, M. C. Benítez, and J. Ibañez (2018). Detection and classification of continuous volcano-seismic signals with recurrent neural networks, *IEEE Trans. Geosci. Remote Sens.* 57, no. 4, 1936–1948.
- [48] Yamada, T., A. K. Kurokawa, A. Terada, W. Kanda, H. Ueda, H. Aoyama, T. Ohkura, Y. Ogawa, and T. Tanada (2021). Locating hydrothermal fluid injection of the 2018 phreatic eruption at Kusatsu-Shirane volcano with volcanic tremor amplitude, *Earth Planets Space* 73, 1–15.