Volcanic Eruption Forecasting Using Shannon Entropy: Multiple Cases of Study

Pablo Rey-Devesa1,2, Carmen Benítez3, Janire Prudencio1,2, Ligdamis Gutiérrez1,2, Guillermo Cortés-Moreno1,2, Manuel Titos3, Ivan Koulakov4,5, Luciano Zuccarello6, and Jesús M. Ibáñez1,2,7

[1]Department of Theoretical Physics and Cosmos. Science Faculty. Avd. Fuentenueva s/n. University of Granada. 18071. Granada. Spain.

[2]Andalusian Institute of Geophysics. Campus de Cartuja. University of Granada. C/Profesor Clavera 12. 18071. Granada. Spain.

[3]Department of Signal Theory, Telematics and Communication. University of Granada. Informatics and Telecommunication School. 18071. Granada. Spain.

[4]Trofimuk Institute of Petroleum Geology and Geophysics SB RAS, Prospekt Koptyuga, 3, 630090 Novosibirsk, Russia

[5]Institute of the Earth's Crust SB RAS, Lermontova 128, Irkutsk, Russia

[6]Istituto Nazionale di Geofisica e Vulcanologia, Sezione di Pisa, via Cesare Battisti, 53, 56125, Pisa, Italy.

[7]Istituto Nazionale di Geofisica e Vulcanologia, Sezione di Catania, Piazza Roma, Osservatorio Etneo, Catania, Italy.

Corresponding author: Jesús M. Ibáñez (jibanez@ugr.es)

1. General Table of Feature parameters

A key point to identify and clarify seismic signals is how to represent observations; that is, the determination of a set of meaningful features that relate to measurements made on observations. Such representations are typically obtained by extracting parameters (features) from the data and using them in a new frame of reference to perform a classification of isolated seismic events. In this work initially we extracted features from both the waveform (statistical, shape descriptors, etc.) and the spectrum according their different natures. Ultimately, we choose a subset of 26 features (Table S1).

Table S1. Features selected.

FEATURE	FORMULA
1. Energy	$E = \sum_{i=1}^{n} S[i]^2; E_i = S[i]^2$
2. Frequency index	$FI = log_{10}\left(\frac{E_{high\ frequencies}}{E_{low\ frequencies}}\right)$
3. Attack ratio	$max_i\left(\frac{S[i] - S[i-1]}{n}\right)$
4. Decay ratio	$min_i\left(\frac{S[i] - S[i+1]}{n}\right)$
5. Mean	$\mu_s = \frac{1}{n}\sum_i S[i]$
6. Standard deviation	$\sigma_s = \sqrt{\frac{1}{(n-1)}\sum_i (S[i] - \mu_s)^2}$
7. Skewness	$\frac{1}{n}\sum_i \left(\frac{S[i] - \mu_s}{\sigma_s}\right)^3$
8. Kurtosis	$\frac{1}{n}\sum_i \left(\frac{S[i] - \mu_s}{\sigma_s}\right)^4$
9. i of central energy	$\bar{i} = \frac{1}{E}\sum_i E_i \cdot i$
10. RMS bandwidth	$B_i = \sqrt{\frac{1}{E}\sum_i i^2 \cdot E_i - \bar{i}^2}$
11. Mean skewness	$\sqrt{\frac{\sum_i (i - \bar{i})^3 E_i}{E \cdot B_i^3}}$
12. Mean kurtosis	$\sqrt{\frac{\sum_i (i - \bar{i})^4 E_i}{E \cdot B_i^4}}$
13. Entropy	$\sum_i P_f(TFD(n,f))\, log_2\left(P_f(TFD(n,f))\right)$
14. Brightness	$\frac{\sum_i f \cdot TFD(n,f)}{\sum_i TFD(n,f)}$
15. Shannon entropy	$-\sum_i P(S_i) log_2\left(P(S_i)\right)$
16. Rényi entropy	$\frac{1}{1-a} log_2\left(\sum_i P(S_i)^a\right)$
17. LPC (5 coefficients)	$S[n] = \left(\sum_i a_k \cdot S[n-k]\right) + err[n]$
18. Cepstral coefficients (5 coefficients)	$Re\left(FFT^{-1}\left(log\left(abs\left(FFT(S_i)\right)\right)\right)\right)$

Note: The symbols used in this table are explained in Esmaili et al. (2004), Alvarez et al., (2011), and Malfante et al. (2018a,b). RMS, Root Mean Square; LPC, linear predictive coefficients. Parameters 1 to 4 are phenomenological features; parameters 5 to 16 are statistical features; parameters 17 and 18 (10 coefficients) are signal domain transform parameters.

2. Study of General evolution of the Seismic Features for Mt. Etna Volcano.

We studied the temporal evolution of the selected seismic features to determine which presented significant pre-eruptive variation in comparison with a non-eruptive period. For this purpose, we selected the data for a lava fountain eruption at Mt. Etna on November 11th, 2013. This choice was based on the fact that among our case study volcanoes, Mt. Etna has the densest seismic network and the seismic records for this event are of high quality over different distances. Figure S1 shows the evolution of the 26 seismic features at station EBEM, divided into six subgroups for 2 days before and 1 day after the eruption. Some features were found to be invariant or to exhibit random changes prior to the eruption (e.g., skewness and mean skewness, the i of central energy, brightness, the LPC coefficients, and the cepstral coefficients). Among the cepstral coefficients, only CEP1 showed pre-eruptive change; however, this is likely because CEP1 is associated with energy. Energy (Figure S1a), the frequency index (Figure S1a), Shannon entropy (Figure S1b), and kurtosis (Figure S1c) all showed clear changes in their temporal evolution prior to the eruption, particularly in the 12 h before the eruption, and represented the best candidates for short-term volcanic eruption prediction. The other features that also showed variation prior to the eruption were at least partially linked to energy, kurtosis, frequency index, or Shannon entropy; as such, they are not discussed further.

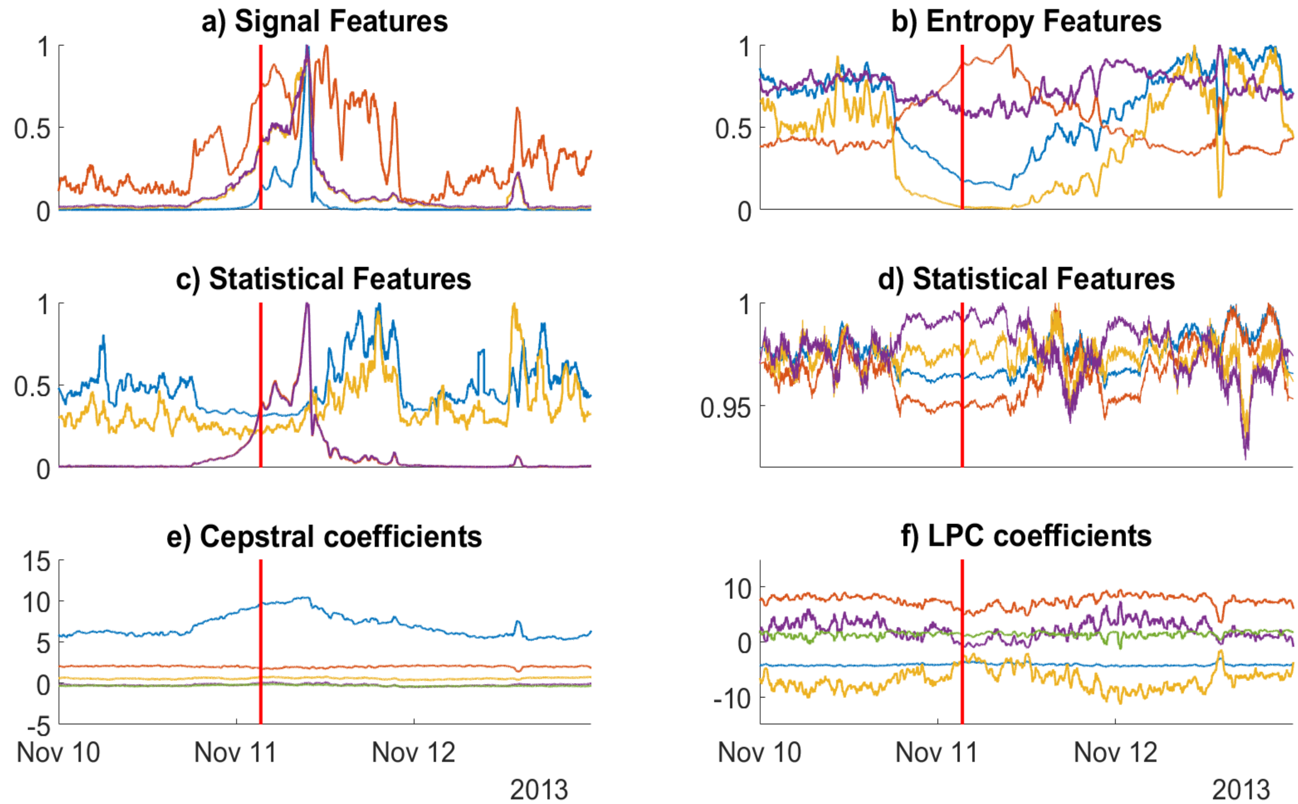

Figure S1. Time evolution of 26 seismic features at station EBEM for 2 days before and 1 day after a lava fountain eruption of Mt. Etna on November 11th, 2013. Features are divided into six subgroups: (**a**) energy and associated features (blue is energy; orange is the frequency index; yellow is the rate of decay, and purple is the rate of attack); (**b**) entropy features (blue is entropy; orange is Rényi entropy; yellow is Shannon entropy, and purple is the brightness); (**c**) statistical features (blue is kurtosis; orange in the mean; yellow is the skewness, and purple is the standard deviation); (**d**) statistical features (blue is mean skewness; orange is mean kurtosis; yellow is the i of central energy, and purple is the root mean square [RMS] bandwidth); (**e**) cepstral coefficients; and (**f**) linear predictive coefficients (LPC) coefficients. In (e) and (f), blue is the 1st, orange the 2nd, yellow the 3rd, purple the 4th, and green is the 5th coefficient. The time of the eruption is represented by the vertical red line.