Journal of Volcanology and Geothermal Research 324 (2016) 134–143

Contents lists available at ScienceDirect

Journal of Volcanology and Geothermal Research

journal homepage: www.elsevier.com/locate/jvolgeores

Automatic detection of volcano-seismic events by modeling state and event duration in hidden Markov models

Sohail Masood Bhatti [a,1], Muhammad Salman Khan [a,2], Jorge Wuth [a], Fernando Huenupan [b], Millaray Curilem [b], Luis Franco [c], Nestor Becerra Yoma [a,*]

[a] Department of Electrical Engineering, Universidad de Chile, Santiago, Chile
[b] Department of Electrical Engineering, Universidad de La Frontera, Temuco, Chile
[c] Observatorio Vulcanológico de los Andes Sur, Chile

ARTICLE INFO

Article history:
Received 26 November 2015
Received in revised form 1 April 2016
Accepted 23 May 2016
Available online 28 May 2016

Keywords:
Volcano monitoring
Duration model
Hidden Markov models
Signal processing
Pattern recognition

ABSTRACT

In this paper we propose an automatic volcano event detection system based on Hidden Markov Model (HMM) with state and event duration models. Since different volcanic events have different durations, therefore the state and whole event durations learnt from the training data are enforced on the corresponding state and event duration models within the HMM. Seismic signals from the Llaima volcano are used to train the system. Two types of events are employed in this study, Long Period (LP) and Volcano-Tectonic (VT). Experiments show that the standard HMMs can detect the volcano events with high accuracy but generates false positives. The results presented in this paper show that the incorporation of duration modeling can lead to reductions in false positive rate in event detection as high as 31% with a true positive accuracy equal to 94%. Further evaluation of the false positives indicate that the false alarms generated by the system were mostly potential events based on the signal-to-noise ratio criteria recommended by a volcano expert.

1. Introduction

Active volcanoes are continuously monitored to observe the underlying volcano-seismic activities. Each volcano-seismic activity or event is associated to a source process and gives an insight into the current state of the volcano that could potentially be used as precursors of an eruption. Automatically detecting (or classifying) these events has gained importance because of the growing need to monitor a high number of active volcanoes, in addition to providing an efficient, consistent and reliable machine-based framework.

Hidden Markov Models (HMMs) can be a powerful and efficient tool that has therefore also been used within multiple applications (Rabiner, 1989). Among others, multiple HMM based volcano-seismic event and earthquake detection and classification techniques have been proposed in recent years (Ohrnberger, 2001; Alasonati et al., 2006; Gutiérrez et al., 2006, 2009; Benítez et al., 2007; Beyreuther et al., 2008, 2012; Beyreuther and Wassermann, 2008, 2011; Ibáñez et al., 2009; Bicego et al., 2013). From the volcanology point of view, HMM is a very convenient method because it can offer a mathematical framework to model signals that are composed of non-stationary events. Gutiérrez et al. (2006), and later Ibáñez et al. (2009), presented an HMM based classification system for volcano events recorded at Stromboli and Etna volcanoes. Data was used from field surveys of years 1997 and 1999. The authors conclude that the recognition system is very efficient but acknowledged that for an improved performance reliable labeling by an expert is vital. Benítez et al. (2007) proposed a classification system developed based on the seismic events recorded during the 1994–1995 and 1995–1996 surveys at the Deception Island Volcano, Antarctica. The events included volcano-tectonic earthquakes, long period events, and volcanic tremor. The authors claimed that the system provides a classification accuracy of 90% and is suitable for real-time operation. The system was also tested on another data set, mainly with LP events, observed during the 2001–2002 field survey, and an accuracy of around 95% was achieved. In Beyreuther et al. (2008), HMMs were employed to detect and classify volcano-tectonic or tectonic earthquakes in continuous seismic data. HMMs were built for events and noise with the HTK software. In the implementation two classifiers were running in parallel, one containing all models of events and noise, and another one with only noise models. Valuable information for the seismological analyst to evaluate the detections was provided by using confidence measure. Beyreuther and Wassermann (2008) proposed the use of Discrete Hidden Markov Modeling (DHMM) for the detection of small to medium

* Corresponding author.
E-mail addresses: sohail.masood@cs.uol.edu.pk (S.M. Bhatti), salmankhan@uetpeshawar.edu.pk (M.S. Khan), jwuths@gmail.com (J. Wuth), fernando.huenupan@ufrontera.cl (F. Huenupan), millaray.curilem@ufrontera.cl (M. Curilem), luis.franco@sernageomin.cl (L. Franco), nbecerra@ing.uchile.cl (N.B. Yoma).
[1] Sohail Masood is now with the Department of Computer Science, University of Lahore, Gujrat Campus, Pakistan.
[2] Salman Khan is now with the Department of Electrical Engineering, University of Engineering and Technology Peshawar, Pakistan.

http://dx.doi.org/10.1016/j.jvolgeores.2016.05.015
0377-0273/

size earthquakes. The seismic signals were recorded with three stations of the Bavarian Earthquake Service. The performance of their algorithm was compared with a recursive LTA/STA detector, within a continuous one-month period. The detection rate was 81% and 90% for the proposed scheme and the LTA/STA respectively, in 69 earthquakes. A drawback when using DHMMs, compared to continuous HMMs, is that time series of real-valued feature vectors have to be converted to discrete valued time series via a vector quantization step as DHMMs are only capable to evaluate discrete symbol sequences. Also, continuous HMMs are more flexible than DHMM. Beyreuther and Wassermann (2011) proposed to use Hidden semi-Markov Models (HSMMs) applied to earthquake detection and classification. HSMMs extend the double stochastic HMMs by integrating a more realistic duration of the target waveforms. When using ordinary HMMs the probability of the duration for a single part in the HMM (called state) is an exponentially decaying function in time which is an unrealistic representation for the duration of earthquake classes or speech units (Oura et al., 2008). HSMMs use the more realistic Gaussians as state duration probability distributions. State transition probabilities and distributions are estimated jointly and automatically by an Expectation Maximization (EM) algorithm. Weighted Finite State Transducers (WFSTs) were built for the classification as the standard Viterbi algorithm cannot be employed because it relies strongly on the intrinsic HMM design. Detection and classification process using this approach was extremely slow (1/2 h CPU time for 1 h data). To avoid this problem, Beyreuther et al. (2012) employed an HMM based system detection with state clustering where states containing similar Gaussians are tied together according to a given metric. By doing so, a more efficient HMM detection was achieved when compared with HSMM. This methodology was used for detection of anthropogenically induced earthquakes and earthquake classification at Mt. Merapi volcano, Indonesia. This approach also incorporates minimum state duration similarly to the HSMM technique. However, the maximum state duration is not restricted. The HMM topology is limited to self-transitions and transitions to the next state. State transition probabilities and distributions are estimated by an EM algorithm as in the previous paper. Studying volcano San Cristóbal, Nicaragua, Gutiérrez et al. (2009) proposed an HMM based automatic volcano event detection and classification system. Data of over 600 h from field survey February to March 2006 was used and the events therein were manually labeled by an expert. They report a classification accuracy of around 80%. Bicego et al. (2013) proposed a new HMM based classification technique by enhancing the HMMs with a generative embedding scheme. The generative embedding uses the models to map signals into a vector space called the generative embedding space. In such a space, any discriminative vector-based classifier (e.g. kernel-based SVMs) can be applied. Experiments were performed on pre-triggered signals recorded at Galeras Volcano in Colombia, and indicated that the proposed approach generally performs better than the standard HMM scheme.

This paper studies the Llaima volcano, one of the biggest by volume in Chile and one of the most active volcanoes in South America. The seismo-volcanic events considered in this study are: Long Period (LP), consisting in transient, volumetric signals; and, Volcano-Tectonic (VT), corresponding to ordinary earthquakes in the brittle rock within a volcanic edifice or in the crust beneath it. As will be shown later, duration distribution depends on the volcano events. As a consequence, duration models could provide useful complementary information to increase the accuracy in the volcano event detection problem. In this paper we therefore propose an HMM based volcano event detection system enhanced with duration modeling. This approach allows introducing minimum and maximum duration of HMM states. Furthermore, minimum and maximum duration constraints of the whole volcano event are also incorporated. Moreover, to reduce false events that may be detected at a frequency higher than true volcano event rate, a penalization is applied to every volcano activity detected. Finally, to reduce the computational load, the Viterbi algorithm, which is a very efficient scheme for HMM decoding (i.e. to find the optimal sequence of model and states for a given recorded signal), was adapted to incorporate the state and event duration constraints, and the penalization to new detected events. It is worth highlighting that the approach and analysis presented in this paper have not been applied to the problem of volcano event detection in the specialized literature.

2. Introduction to hidden Markov models

An HMM is composed of a sequence of states that can model non-stationary signals. HMMs are a finite state machine defined by three set of parameters: a) transition probabilities between states; b) observation probabilities in each state; and, c) initial probabilities for each state. In this research, volcano events are modeled using tree states to represent their dynamics. Fig. 1 shows an example of an HMM with 3 states S_1, S_2 and S_3. The first state models the beginning of the event signal, whilst the second and third states model the middle and end of the event, respectively.

2.1. Transition probabilities

As explained in more detail later, the recorded signal is divided in short-term windows, and within each window a set of features are estimated (i.e. feature extraction). The set of features estimated for each window is denominated frame. Each frame t is represented by a feature vector denoted as $\boldsymbol{O}_t$. Given an HMM, the current state may change from one frame to the following one, and the probability for changing to state j from state i is given by the transition probability a_{ij}. In this way, transitions from one state to another can be allowed by setting the corresponding transition probabilities to a positive number. In contrast, if a transition probability is set to zero, the corresponding transition is forbidden. The allowed state transitions are drawn as arcs in the model. In the HMM of Fig. 1, given a state i, only state transitions to the same state i or the state $i + 1$ are possible. This topology is known as left-to-right without state skip transition.

2.2. Observation probabilities

Given an HMM, to each state S_i corresponds an observation probability, $\Pr(\boldsymbol{O}_t|S_i,\lambda)$. In this research, the observation probability is modeled with a probability density function represented by a Gaussian Mixture Model (GMM) composed of G Gaussians. The observation probability is defined as:

$$Pr(\boldsymbol{O}_t|S_i,\lambda) = \sum_{g=1}^{G} \boldsymbol{\varphi}_{g,i,\lambda} \times \aleph\left(\boldsymbol{O}_t, \boldsymbol{\mu}_{\boldsymbol{g},\boldsymbol{i},\boldsymbol{\lambda}}, \sum_{g,i,\lambda}\right) \quad (1)$$

where λ denotes the HMM (LP, VT, or noise), G is the number of Gaussians per state, $\aleph(\cdot;\boldsymbol{\mu},\boldsymbol{\Sigma})$ is a multivariate Gaussian with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$:

$$\aleph(\boldsymbol{O};\boldsymbol{\mu},\boldsymbol{\Sigma}) = \frac{1}{\sqrt{(2\pi)^n|\Sigma|}} e^{-\frac{1}{2}(\boldsymbol{O}-\boldsymbol{\mu})^T\Sigma^{-1}(\boldsymbol{O}-\boldsymbol{\mu})} \quad (2)$$

where n is the dimensionality of $\boldsymbol{O}$, and $\boldsymbol{\varphi}_{g,i,\lambda}$ are the weights of the Gaussians.

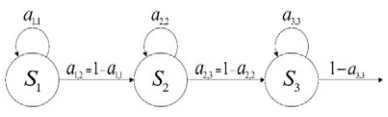

Fig. 1. Three state left-to-right HMM without state skip transition to model LP and VT events.

2.3. Initial probabilities

Another parameter that defines an HMM is the vector of initial probabilities for each state $\boldsymbol{\pi} = (\pi_2, \pi_{2,\ldots}\ \pi_{N_s})$, where N_S is the number of states within the HMM. Vector $\boldsymbol{\pi}$ represents the probability distribution on the initial state, i.e. the probability of the state allocated to the first frame within a given HMM. In this paper, the first frame within an HMM is assigned to the first state in the HMM. For the HMM in Fig. 1, $\boldsymbol{\pi} = (1,0,0)$.

2.4. Feature extraction

In the feature extraction process, the goal is to reduce the data dimensionality by converting the sampled waveform into a sequence of parameter vectors with less redundant information. Accordingly, a recorded continuous seismic signal of length equal to T frames, is represented by a sequence of observation vectors $\boldsymbol{O}_1, \boldsymbol{O}_2, \ldots \boldsymbol{O}_T$ composed of representative features of its main temporal and spectral characteristics. The feature extraction process is carried out by arranging the signal into 200–300 sample frames, usually overlapped, by employing a Hamming window. For each frame, different features are extracted to compose the observation vector of the frame. Cepstral coefficients (Beyreuther et al., 2012) and linear prediction filter coefficients (LPC) (Esposito et al., 2013) are commonly used in feature extraction. LPC are features that give compressed information about the shape of the spectral envelope of a signal. They can be used to predict a signal sample through a linear combination of previous samples. Cepstral coefficients are defined as the DCT (Discrete Cosine Transform) coefficients obtained from the log magnitude over the fast Fourier transform (FFT) of the signal. This feature vector is related to the statistical distribution of the spectral characteristics of the signal (Ibáñez et al., 2009).

Delta (Δ) and delta–delta (ΔΔ), also known as differential and acceleration coefficients, are also usually considered to include temporal evolution of the static features described above, e.g. Cepstral and LPC coefficients. This time evolution is established on the basis of the first and second derivatives of the corresponding static features. Δ and ΔΔ for the static features in frame t are defined as follows:

$$\Delta(t) = \frac{static\_features(t+1) - static\_features(t-1)}{2} \tag{3}$$

$$\Delta\Delta(t) = \frac{\Delta(t+1) - \Delta(t-1)}{2} \tag{4}$$

where $static\_features(t)$ denotes the corresponding static feature vector in frame t.

2.5. Training and decoding

After the feature extraction, the HMM parameters for each event and noise are estimated from a set of training recorded signals. This training procedure allows us to obtain the parameters that define a representative model for each defined class of event. For the training process it is necessary to determine the seismic event classes that will be addressed. Regarding the type of the volcanic seismic events considered in this work, it refers to commonly events often observed in most active volcanoes (Chouet and Matoza, 2013). These events are associated either to the dynamics of fluids into the volcano (long period type, LP) or with the breaking of brittle material (volcanotectonic type, VT). For the training and testing processes, the database stores events chosen by the human volcano experts from the continuously recorded data. The volcano specialists discriminate the events of interest from other types of volcanic signals (not included in this research) and from non-volcanic signals (like tectonic and environmental noise). LP and VT events of Llaima volcano were targeted in this study. Features such as the shape of the wave, envelope, event duration, time difference between primary (P) and secondary waves (S), frequency content and verification at various stations with consistent times of arrival, are the criteria usually employed by the human experts, to discriminate between LP and VT, and other volcanic and non-volcanic events. Then, we select examples of seismic events (training data) that can be identified as belonging to LP and VT classes. The training data set must be large enough to be statistically representative of the addressed classes. These trained HMMs are employed in the decoding process. In this context, given a signal recorded by a seismic station, decoding corresponds to allocate every frame in one of the models and states that compose an HMM network (see subsection 2.2). This network is defined by the interconnection of all the HMMs that represent the targeted volcano events and background noise. This HMM network is also determined by the realistic sequence of volcano events and noise. The decoding procedure (see Viterbi algorithm in Section 3.5) is both a detection and classification procedure because it delivers the most likely sequence of volcano events and noise. The decoding algorithm used in this research is the Viterbi algorithm (Viterbi, 1967; Rabiner, 1989; Rabiner and Juang, 1993).

Fig. 2 shows the proposed system architecture. Representative features from the training data are extracted and the HMMs are trained. The Gaussian distributions for the observation probabilities and the state and event duration parameters are also evaluated from the training data. In the decoding process, features are extracted from a given test signal and the optimal alignment is estimated with the Viterbi algorithm. The optimal alignment corresponds to the most likely assignment of frames to the states within the HMMs. The different components of the system are detailed as follows.

2.6. Events and noise model definitions

As mentioned above, LP and VT events of Llaima volcano were addressed in this study. Based on the dynamics of these events they are segmented into three sequence of frames, i.e. beginning, middle, and end. Each segment is allocated to one state within the LP and VT event HMMs, while only one state is used to model the noise in this paper. This choice is justified because we are not adopting any consideration about the stationarity of the background noise. Moreover, the GMM observation probability is adequate to model the diversity of the noise process. It is worth highlighting that the proposed method is not limited to the use of a fixed number of states per event.

Fig. 1 shows the three state left-to-right HMM without state skip transition adopted for LP and VT events. The HMM topology enforces the volcano events to go through all the three HMM states, starting in state 1 and ending in state 3. Observe that jumping from state 1 to state 3 is forbidden. The absence of volcano-seismic activity is represented by background noise. In this paper, the background noise is modeled by making use of only one state whose observation probability is modeled with a GMM to capture the diversity of the corrupting process. Fig. 3 presents the single-state HMM to model noise. As was mentioned above, each HMM λ (i.e. VT, LP, and noise) is defined by the initial probabilities of each state, the transition probabilities, and the observation probability mean vector and covariance matrix. In this paper we consider that each time an event takes place, the first frame of the event is allocated to the first state of its corresponding model. By doing so, the initial state probabilities of LP and VT HMMs are $\boldsymbol{\pi} = (1,0,0)$, as explained above. Notice that the noise HMM is composed of only one state. In this paper, we incorporate state duration constraints to the transition probabilities of LP and VT HMMs. In the case of the background noise HMM, no duration restrictions were applied. This is due to the fact that the training and testing data were generated by taking 20 min signals (10 min after and before an event identified by the volcano experts). In the case of the decoding process, the system outputs the information concerning the detected events (i.e. one or more) for each 20 min signal. As a result, the duration statistics for noise are not representative of the time separation between contiguous

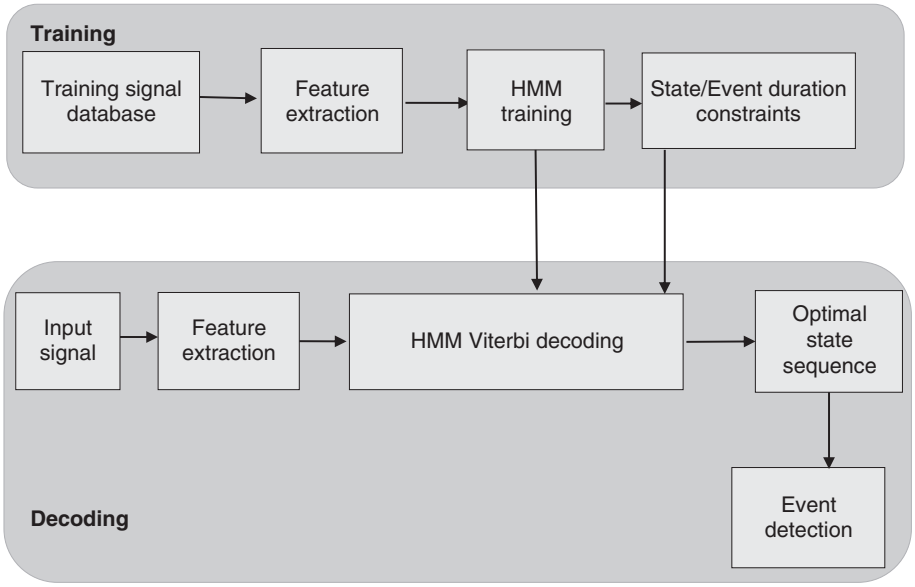

Fig. 2. Proposed system architecture.

volcano events. Actually, in a practical scenario, an event may start at any time (e.g. a few seconds or a few hours after a previous event). Consequently, this noise statistic model for state duration would impose very loose restrictions.

The observation probabilities are represented by GMMs. Each GMM is composed of eight Gaussians in this paper. The covariance matrix is considered diagonal, which is the result of the fact that the features are considered uncorrelated. The detailed parameter estimation process is given in Section 3.4.

2.7. Events and noise model network

Each volcano event is associated to an HMM and all the defined HMMs, including noise, compose the whole HMM network that represents the possible volcano event sequence as shown in Fig. 4. The transition probability a_{i_A,k_B} corresponds to the probability of going from state i in an HMM (A), given a frame, to state k in the same or different HMM (B) in the following frame. Transition probability from noise to the first state of LP and VT events, $a_{1_{Noise},1_{LP}}$ and $a_{1_{Noise},1_{VT}}$, respectively, are considered as $a_{1_{Noise},1_{LP}} = a_{1_{Noise},1_{VT}} = \frac{1-a_{1_{Noise},1_{Noise}}}{2}$. Transitions from the last state of LP and VT events to the single-state of noise are defined by the transition probabilities $a_{3_{LP},1_{Noise}}$ and $a_{3_{VT},1_{Noise}}$, respectively. These probabilities are evaluated as $a_{3_{LP},1_{Noise}} = 1 - a_{3_{LP},3_{LP}}$ and $a_{3_{VT},1_{Noise}} = 1 - a_{3_{VT},3_{VT}}$. According to the proposed HMM network, any volcano event has to be preceded and followed by an absence of activity, i.e. background noise. It is not possible to observe two consecutive volcano events without noise frames between them. The transition from one state to the next one is defined in such a way that the next state from noise is always the first state of LP or VT event HMM. Similarly, after the first state of LP or VT events, the following frames are allocated in the first or second state of the corresponding HMM according to the left-to-right without state skip transition topology. The transition from the third state within LP or VT HMM can be to the same state or to the noise state/model. For instance, after a period of no seismic activity, the volcano can present an LP event, including beginning, middle and end stages. After that, a new period without volcano activity takes place. Then, a new event can start. A possible sequence of states allowed by the HMM network, in Fig. 4, given a sequence of nine frames could be, for example, as follows:

(Noise model, state 1, frame 1), (Noise model, state 1, frame 2), (LP model, state 1, frame 3), (LP model, state 2, frame 4), (LP model, state 2, frame 5), (LP model, state 3, frame 6),(LP model, state 3, frame 7), (Noise model, state 1, frame 8), (VT model, state 1, frame 9), and so on.

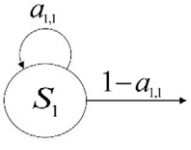

Fig. 3. One state HMM to model noise.

3. The proposed method

LP events are transient, volumetric signals consisting of a brief broadband onset, followed by a coda of decaying harmonic oscillations containing pronounced spectral peaks that are independent of azimuth and distance to the source (Chouet and Matoza, 2013; Zobin, 2012). The frequency energy for LP is typically located between 0.5–5 Hz (Chouet and Matoza, 2013). The worldwide observations of LP show a wide variability in temporal durations, which also depends on the specific volcano. The mean and standard deviation for the duration of LP events in the recorded signal database used in this paper is 32.9 ± 10.7 s.

VT earthquakes are ordinary earthquakes in the brittle rock within a volcanic edifice or in the crust beneath it. They are characterized by sharp, mostly impulsive onsets of P- and S-waves, with typically broad spectra extending up to 15 Hz (Lahr et al., 1994). VT has close similarity with tectonic earthquakes and is the result of abrupt frictional slip on opposing rock surfaces with a double-couple source mechanism (Chouet and Matoza, 2013). The mean and standard deviation for the duration of VT events in the recorded signal database used in this paper is 25.7 ± 14.5 s.

Fig. 5 shows the normalized histograms of the LP and VT event durations. In Fig. 5 it can be seen that, on average, the LP events have longer time duration compared to VT signals. Also, the maximum duration of VT events is higher than the LP ones. Because the duration statistical distributions of the LP and VT volcano-seismic activities are different, they could provide useful complementary information to detect them. Similarly, Fig. 6 shows an example of volcano-seismic activity: its time domain signal; the label by a volcano expert; and, the spectrogram. We can see that there are three potential events but only one was labeled by the expert as a proper volcano-seismic signal, while the other two events were not labeled because, among other reasons, they present a

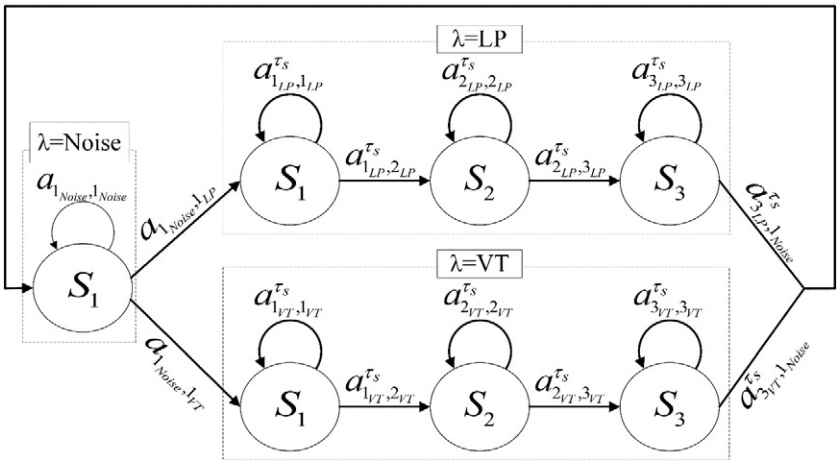

Fig. 4. HMM network for volcano events detection.

too short duration. These events could have been labeled as volcano-seismic activity if duration constraints were not enforced, indicating that the volcano experts possibly consider a priori information of the real volcano event duration in their analysis. This suggests that the incorporation of duration restrictions can provide useful information to improve the accuracy of a system for seismic activity detection.

3.1. State duration constraints

Duration probability distribution of states is not included in the ordinary HMM technique. According to the definition of HMMs, the probability of staying τ frames in state i_A can be estimated from the state transition probabilities as ${a_{i_A,i_A}}^{\tau-1}\cdot(1-a_{i_A,i_A})$, that results in a geometric distribution of state duration. In this paper, we incorporate state duration constraints to the transition probabilities in order to model the information contained in the time durations of LP and VT events. Duration constraints are incorporated in each state of the LP and VT HMMs (Fig. 4) according to Yoma et al. (2001). The transition probabilities a_{i_A,i_B} or a_{i_A,k_B}, as explained in Section 3.4.1, are constrained based on the minimum and maximum state duration τ_s:

$$a_{i_A,i_A}^{\tau_s}=\begin{cases}1, & if\ \tau_s<t_{min\_s_i}\\ 0, & if\ \tau_s\geq t_{max\_s_i}\\ a_{i_A,i_A}, & otherwise\end{cases} \tag{5}$$

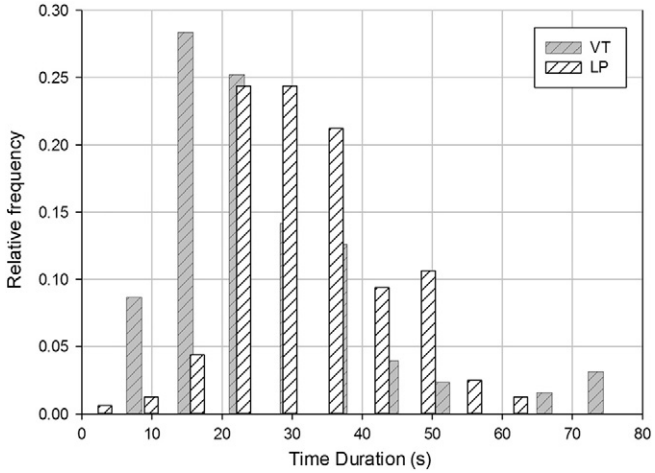

Fig. 5. Normalized histograms of the VT and LP event durations.

$$a_{i_A,k_B}^{\tau_s}=\begin{cases}0, & if\ \tau_s<t_{min\_s_i}\\ 1, & if\ \tau_s\geq t_{max\_s_i}\\ a_{i_A,k_B}, & otherwise\end{cases} \tag{6}$$

where:

in (6), if $[i_A=1\vee 2]\ =>k_B\ =i_A+1$, where $A=B=(LP\vee VT.)$
in (6), if $[i_A=3]\wedge[A=LP\vee VT]=>[B=noise\wedge k_B=1]\vee[B=A\wedge k_B=3]$

Here, $t_{\min\_s_i}$ and $t_{\max\_s_i}$ are the minimum and maximum duration thresholds for state i, respectively, and are computed as $t_{\min\_s_i}=tol\_\min\_s\times\ \min\_s_i(\tau_s)$ and $t_{\max\_s_i}=tol\_\max\_s\times\max\_s_i(\tau_s)$, where $tol\_\min\_s\ <1$, $tol\_\max\_s>1$, and $\min\_s_i(\tau_s)$ and $\max\_s_i(\tau_s)$ are the minimum and maximum durations for state i obtained from the training data. The constants $tol\_\min\_s$ and $tol\_\max\_s$ introduce a tolerance factor to the minimum and maximum duration for every state.

3.2. Event duration penalty

In combination with the state duration restrictions, as explained above, a duration constraint on the whole volcano event, *EDP* (event duration penalty) is proposed. EDP is also applied by making use of a truncated Gamma probability density function for each event. The following expression is applicable for both LP and VT events separately:

$$EDP=\begin{cases}-\infty, & if\ \tau_e<t_{\min\_e}\\ -\infty, & if\ \tau_e\geq t_{\max\_e}\\ \log\left(k.e-\alpha.\tau_e.{\tau_e}^{\rho-1}\right) & otherwise\end{cases} \tag{7}$$

where $\tau_e=0,1,2,\ldots$ is the duration of events LP or VT in number of frames; $t_{\min\_e}$ and $t_{\max\_e}$ are the minimum and maximum duration thresholds for events LP or VT, respectively, and are estimated as $t_{\min\_e}=tol\_\min\_e\times\ \min\_e(\tau_e)$ and $t_{\max\_e}=tol\_\max\_e\times\max\_e(\tau_e)$, where $tol\_\min\_e\ <1$, $tol\_\max\_e>1$, and $\min\_e(\tau_e)$ and $\max\_e(\tau_e)$ are the minimum and maximum durations for the event obtained from the training data. The constants $tol\_\min\_e$ and $tol\_\max\_e$ introduce a tolerance factor to the minimum and maximum duration for each event; and $\alpha>0$, $\rho>0$ and k is a normalizing term. The parameters α and ρ are estimated by

$$\alpha=\frac{E(\tau_e)}{Var(\tau_e)} \tag{8}$$

$$\rho=\frac{E^2(\tau_e)}{Var(\tau_e)} \tag{9}$$

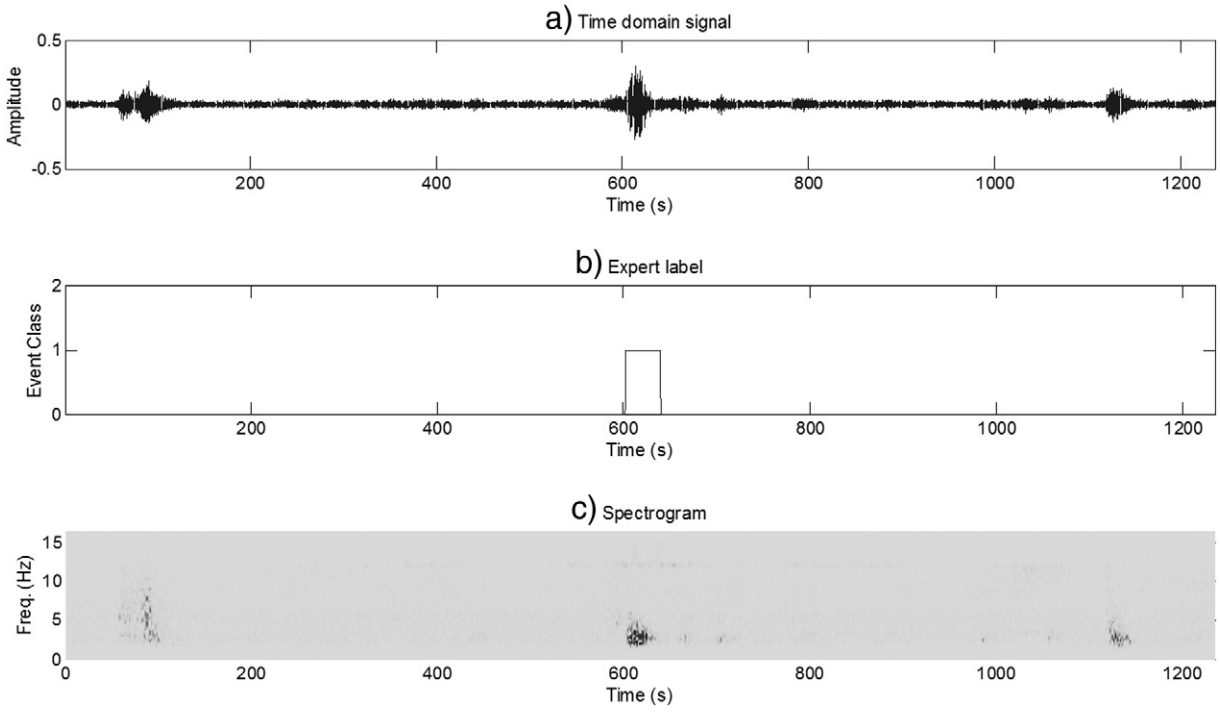

Fig. 6. Example of a seismic signal labeled by a volcano expert: (a) time domain signal of a volcano event; (b) labeling according to the volcano expert, where label 0 is noise, label 1 is LP, and label 2 is VT; and, (c) the corresponding spectrogram of the signal.

where the mean duration, $E(\tau_e)$, the variance, $Var(\tau_e)$, and the min_e and max_e durations are estimated for every event occurrence obtained from the training data labeled by volcano experts. EDP is applied as an additive term in the log-likelihood domain when an observation vector $\boldsymbol{O}_t$ is in the last state of the LP or VT HMM and $\boldsymbol{O}_{t+1}$ jumped to the noise state/model (see the Viterbi algorithm in Section 3.5). As a result, those alignments where detected volcano events presents likely durations according to (7) will have higher log-likelihoods than those alignments where event durations are unlikely. Consequently, the Viterbi decoding procedure will tend to deliver optimal alignments where volcano-seismic signals are likely according to (7).

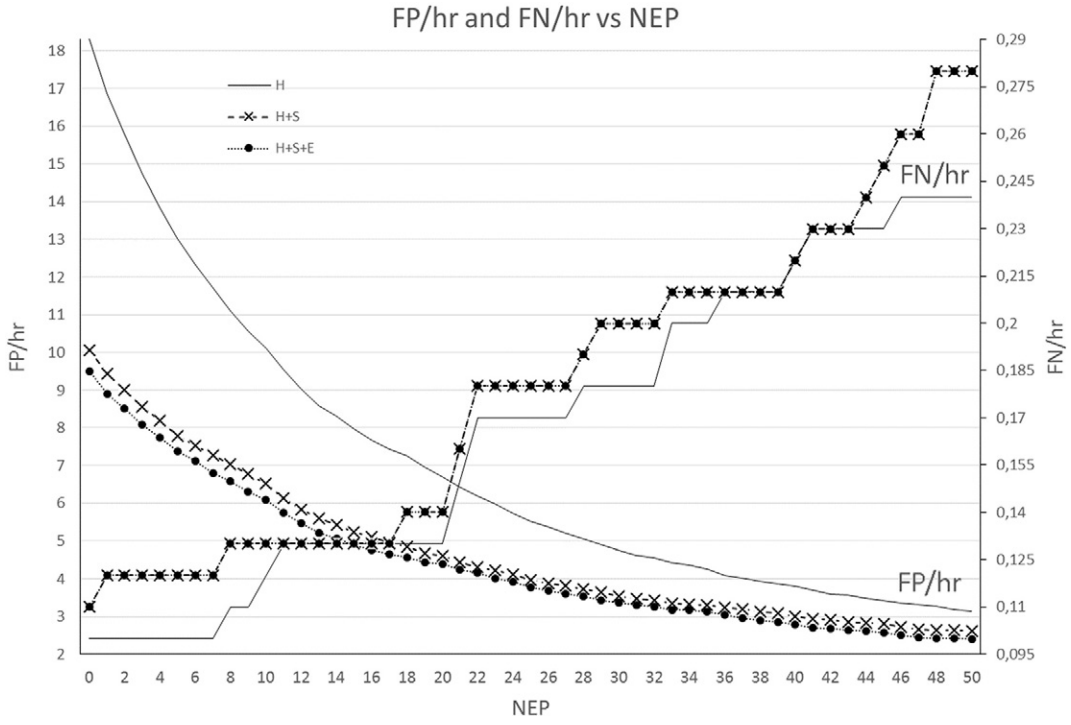

Fig. 7. False negatives rate (FN/h) and false positives rate (FP/h) vs new event penalty (NEP).

3.3. New event penalization

As suggested by Fig. 6, false events may be detected at a rate higher than true volcano event occurrences. In order to reduce the false event detection rate in the proposed HMM network shown in Fig. 4, a penalization is applied in the transition from single-state noise model to the first state of LP or VT events. The new event penalization, *NEP*, is a scalar constant and is also applied in the log-likelihood domain. *NEP* is applied as an additive term in the log-likelihood domain when an observation vector $\boldsymbol{O}_t$ is in the noise state/model and $\boldsymbol{O}_{t+1}$ jumped to the first state of the LP or VT HMM (see the Viterbi algorithm in Section 3.5). As a result, those alignments with a higher number of events will have lower log-likelihoods than those alignments with just a few detected volcano-seismic signal. Consequently, the Viterbi decoding procedure will tend to deliver optimal alignments with just a few volcano events.

3.4. Training procedure

All the signals, both for training and testing, were labeled by an expert from The Southern Andes Volcano Observatory (OVDAS), the state agency that continuously monitors over forty active Chilean volcanoes including Llaima. Each signal was divided in 300-sample windows overlapped by 50%. Feature vectors were estimated for each window. For training all the LP and VT event occurrences were uniformly divided in three non-overlapping segments of feature vectors that were allocated to state 1, 2, and 3 sequentially. In contrast, all the feature vectors for noise were assigned to the single-state noise HMM. The statistics required to estimate all the HMMs parameters are computed with the frames in each state. The training method used to build the HMMs for each class is the expectation maximization algorithm, EM, to estimate the optimal parameters of the Gaussian mixture models that represent the observation probabilities (Huang et al., 1990). The Gaussian distributions and the state and event duration parameters were computed from the training data.

3.4.1. Ordinary transition probabilities

The ordinary HMM state transition probabilities are estimated according to Yoma et al. (2005):

$$a_{i_A,i_A} = 1 - \frac{1}{E_i[\tau_s]} \tag{10}$$

$$a_{i_A,k_B} = \begin{cases} 1 - a_{i_A,i_A}, & \text{if } i = 1,2 \text{ for } A = LP \vee VT, \\ & k = i+1 \text{ for } B = A, \\ & \text{and if } i = 3 \text{ for } A = LP \vee VT, \\ & k = 1 \text{ for } B = noise \\ \dfrac{1 - a_{i_A,i_A}}{2}, & \text{if } i = 1 \text{ for } A = noise, \\ & k = 1 \text{ for } B = LP \vee VT \end{cases} \tag{11}$$

where $E_i[\tau_s]$ is the expected value of the duration of the i-th state in a given model in the training signals.

3.4.2. Training of observation probabilities

The number of Gaussians, G in (1), was made equal to eight. All the Gaussian mixture parameters were trained by using the EM algorithm with the frames allocated to the corresponding state (McLachlan and Peel, 2000).

3.4.3. State/event duration parameters estimation

Minimum and maximum state duration, required by (5) and (6), and minimum, maximum, mean and standard deviation of event duration, required by (7), are estimated by using the frames allocated to all the states and event occurrences in the training data. In the case of the single-state noise model, state duration statistics were not estimated.

3.5. Decoding process

To detect events in a testing signal, first the feature set is calculated for each frame. Then, the optimal alignment is estimated with the Viterbi algorithm (Huang et al., 1990), described below. The proposed Viterbi algorithm incorporates transition probabilities that are functions of state and event duration, and penalization NEP defined above.

Step 1. Initialization:
When frame $t = 1$
When state = 1 in λ_N,

$$\delta(state = 1, \lambda = \lambda_N, t = 1) = \log\left[Pr(\boldsymbol{O}_0 \mid state = 1, \lambda = \lambda_N)\right]$$

$$\omega(state = 1, \lambda = \lambda_N, t = 1) = 0$$

For every other state in the HMM network (Fig. 4),

$$\delta(state, \lambda, t = 1) = -\infty$$

Step 2. Recursive calculation of the ML (maximum likelihood) state sequences:
for frame t =2, . . ., T, where T is the total number of frames
for all the states and all the models in the HMM network (Fig. 4):
case 1: if state=1 and $\lambda = \lambda_N$,

$$\delta(state, \lambda, t) = Pr(\boldsymbol{O}_t \mid state, \lambda) + \max\left[\delta(state, \lambda, t-1) + \log\left(a^{\tau_s}_{state_\lambda, state_\lambda}\right);\right.$$

$$\left.\delta(3, \lambda_{LP/VT}, t-1) + \log\left(a^{\tau_s}_{3_{\lambda_{LP/VT}}, state_{\lambda_N}}\right) + EDP\right]$$

$\omega(state, \lambda, t) =$ optimal state and model at $t-1$ in the previous optimization

case 2: if state=1 and $\lambda = \lambda_{LP}$ or λ_{VT},

$$\delta(state, \lambda, t) = \Pr(\boldsymbol{O}_t \mid state, \lambda) + \max\left[\delta(state, \lambda, t-1) + \log\left(a^{\tau_s}_{state_\lambda, state_\lambda}\right);\right.$$

$$\left.\delta(1, \lambda_N, t-1) + \log\left(a^{\tau_s}_{1_{\lambda_N}, state_\lambda}\right) - NEP\right]$$

$\omega(state, \lambda, t) =$ optimal state and model at $t-1$ in the previous optimization

case 3: if state=2 or 3 and $\lambda = \lambda_{LP}$ or λ_{VT},

$$\delta(state, \lambda, t) = \Pr(\boldsymbol{O}_t \mid state, \lambda) + \max\left[\delta(state, \lambda, t-1) + \log\left(a^{\tau_s}_{state_\lambda, state_\lambda}\right);\right.$$

$$\left.\delta(state - 1, \lambda, t-1) + \log\left(a^{\tau_s}_{state-1_\lambda, state_\lambda}\right)\right]$$

$\omega(state, \lambda, t) =$ optimal state and model at $t-1$ in the previous optimization

end state
end frame
Step 3. Finalization: retrieval of the most likely final state as per the following

$$\delta\left(state^*, \lambda^*, T\right) = \underset{state,\ \lambda}{\operatorname{argmax}}\left[\delta(state, \lambda, T)\right]$$

where $state^*$ and λ^* are the optimal state and model at time T

$$\omega^*\left(state^*, \lambda^*, T\right) = \omega\left(state^*, \lambda^*, T\right)$$

where ω^* is the optimal alignment
Step 4. Optimal alignment estimation:
for $t = T$-1 to $t = 2$

$$\omega^*\left(state^*, \lambda^*, t\right) = \omega\left(state^*, \lambda^*, t\right)$$

where $state^*$ and λ^* are given by $\omega^*(state, \lambda, t+1)$
end
Step 5. The optimal alignment is employed to determine the optimal sequence of states and events.

3.5.1. Viterbi algorithm with duration modeling

The Viterbi algorithm incorporating state and event duration is summarized above. The likelihood of a state in a model at frame t is represented by δ, and the corresponding optimal state and model at time $t-1$ is denoted by ω; λ_N = *noise HMM*, λ_{LP} = *LP HMM*, λ_{VT} = *VT HMM*, and $\lambda_{LP/VT}$ = *LP or VT HMMs* within the algorithm description.

4. Experimental setup and system description

As mentioned before, this paper studies the Llaima volcano. Llaima is a compound strato-volcano of basaltic to andesite-basaltic composition and is located in the Araucanía Region (38° 41′S–71° 44′W), on the western edge of the Andes. Llaima volcano has 9 seismic stations, which together with other instruments are used to monitor its activity. In the present study only one station i.e. Laguna Verde (LAV) is considered. The station is located at −38.700988°; −71.651116, 7 km from the crater. It is a broadband station, Guralp 6TD of 30s period and only the Z component is considered, since in most of the events it provides

a better signal-to-noise ratio. The seismo-volcanic events considered in this study are Long Period (LP) and Volcano-Tectonic (VT), recorded between 2011 and 2014. The sample rate of the data was 100 Hz.

The identification of each volcanic event, location, duration and classification was made manually by an OVDAS analyst. The data from the digital binary SUDS format was exported to the Matlab environment, and it was segmented by centering a given event by considering 10 min before its start and 10 min after its end. As a result, the length of each signal is 20 min. A total of 167 signals containing LP events and 133 signals containing VT events, which gives a total of 100 h of samples, were used in this study. Experiments are performed on two subsets of the database where the first half is referred to as subset1 and the second half as subset2. Subset1 contains the first 83 LP and 67 VT events. Subset2 contains the remaining 84 LP and 66 VT events of the database.

4.1. Feature extraction

Feature extraction procedure was performed as mentioned in Section 2. Considering that the sampling frequency is 100 Hz, the [0–50 Hz] bandwidth was divided in 16 uniform bandpass filters. The features considered are: five linear prediction filter coefficients (LPC) (Esposito et al., 2013); twenty Cepstral coefficients (Beyreuther et al., 2012) estimated with the 16 uniform bandpass filters mentioned above; and the energy in the 5th wavelet band (1.56–3.125 Hz) (Curilem et al., 2014) obtained with the wavelet transform using a Daubechies mother type five and five decomposition levels. This feature was computed as the ratio between the sum of the components of the 5th wavelet band over the sum of all the wavelet components (in all the bands). Finally, the first (Δ) and second ($\Delta\Delta$) derivatives of the static features, as defined in (3) and (4), are also estimated. Consequently, the total number of static features was equal to 26, which combined with Δ and $\Delta\Delta$ parameters provide a final feature vector of 78 coefficients.

4.2. Performance metrics

The performance of the proposed HMM based volcano event detection system is evaluated using "true positives per hour", "false negatives per hour", and "false positives per hour". The correct detection of a labeled event occurrence is considered as a "true positive" (TP). A labeled event occurrence that is not detected is counted as a "false negative" (FN), while the detection of an event that was not labeled by the human expert is termed as a "false positive" (FP).

5. Results and discussion

Three versions of the volcano event detection system based on Viterbi algorithm are evaluated: ordinary Viterbi algorithm with constant transition probabilities without EDP, termed as H; the Viterbi algorithm with state duration constraint according to (5), (6) without EDP, denoted as H + S; and, the Viterbi algorithm with state duration constraint as given in (5), (6) combined with EDP as defined in (7), indicated as H + S + E. Subset1 was used for training and subset2 for testing, and vice versa. Observe that both EDP and NEP are additive parameters in the log-likelihood domain in the Viterbi algorithm described above. All the three versions of the Viterbi algorithm were tuned by varying the penalty NEP from 0 to 50 with a step size equal to 1. Average results with subset1 and subset2 versus NEP are presented in Fig. 7.

It can be seen in Fig. 7 that the number of FPs/h starts decreasing rapidly with the increase in NEP and then converges at high values of NEP for the H, the H + S, and the H + S + E configurations. However, the FPs/h curve corresponding to H is always dramatically higher than the FPs/h curves for H + S and H + S + E. The H + S + E gives the best results in terms of FPs/h as compared to the H and the H + S, while generally, the H + S also reduces FPs/h as compared to the H version of Viterbi. For instance, at NEP = 27, the H + S and the H + S + E lead to a reduction of 27% and 31% in FPs/h as compared with H. This is because the application of state and event duration constraints restricts detection of those events that have durations less than the minimum or greater than the maximum durations obtained from the training data. On the other hand, the number of FNs/h increases slowly with the increment in NEP for all the three versions. Also at NEP = 27, H + S and H + S + E lead to a slight increase of 6% in FNs/h when compared with the baseline H. Observe that the FN/h curve for H + S and H + S + E are identical. A NEP equal to 27 seems to reduce FP/h dramatically without a considerable increment in FN/h for the three methods. It is worth highlighting that the state duration constraints applied in the H + S method lead to the event minimum and maximum duration bounds, given by the sum of the corresponding minimum and maximum state durations, respectively. This event duration constraint is then refined in the H + S + E method leading to a reduction in FP/h equal to 5% when compared to the H + S method at NEP = 27.

Table 1
TP rates, FN rates and FP rates with the H, H + S, H + S + E Viterbi decoding for NEP = 27. Subset1 was used for training and subset2 was employed for testing.

	TPs per hour	FNs per hour	FPs per hour
H	2.8	0.2	4.16
H + S	2.8	0.2	3.38
H + S + E	2.8	0.2	3.06

Results with NEP = 27 for training on subset1 and testing on subset2, and vice versa, are shown in Tables 1 and 2, respectively. It can be seen from Tables 1 and 2 that there is a significant reduction in FPs with the H + S and H + S + E methods as compared to the baseline H configuration. In Table 1, where training is performed on subset1 and testing on subset2, the H, H + S, and H + S + E methods detect 2.8 events per hour corresponding to a detection accuracy of 93.3% each (140 out of 150 events detected correctly). The H method generates 4.16 false positives per hour (FPs/h). The H + S, that incorporates state duration model, reduces the false alarm rate to 3.38 corresponding to a reduction of 19%, while further including the event duration constraint, i.e. H + S + E, produces 3.06 false alarms per hour, a 26% reduction when compared with the H scheme. The three methods generate 0.2 false negatives per hour (FNs/h). In Table 2, where training is performed on subset2 and testing on subset1, the results basically follow the same trend as in Table 1. The H, H + S, and H + S + E methods achieve detection accuracies of 95.3%, 94.7%, and 94.7%, respectively. The H method generates 6.26 FPs/h. The H + S scheme achieves 4.24 FPs/h that corresponds to a reduction of 32% when compared with H. The H + S + E version of the Viterbi decoding delivers 4.16 FPs/h that corresponds to a reduction of 34% over the method H. The slight reduction of 0.6% in the true positive rate must be due to the fact that the duration statistics obtained from subset2 are not totally representative of subset1. However, the reduction in the FP rate in both Tables 1 and 2 is due to the fact that the state and event duration constraints discards the potential events that do not fulfill the duration criteria based on the statistics obtained from the training data. On the other hand, the H method generates 0.14 FNs/h. The H + S and H + S + E schemes achieve 0.16 FNs/h that corresponds to an increase of 14% when compared with H. This increment of FN/h is much lower than the reduction on FP/h.

Table 2
TP rates, FN rates and FP rates with the H, H + S, H + S + E Viterbi decoding for NEP = 27. Subset2 was used for training and subset1 was employed for testing.

	TPs per hour	FNs per hour	FPs per hour
H	2.86	0.14	6.26
H + S	2.84	0.16	4.24
H + S + E	2.84	0.16	4.16

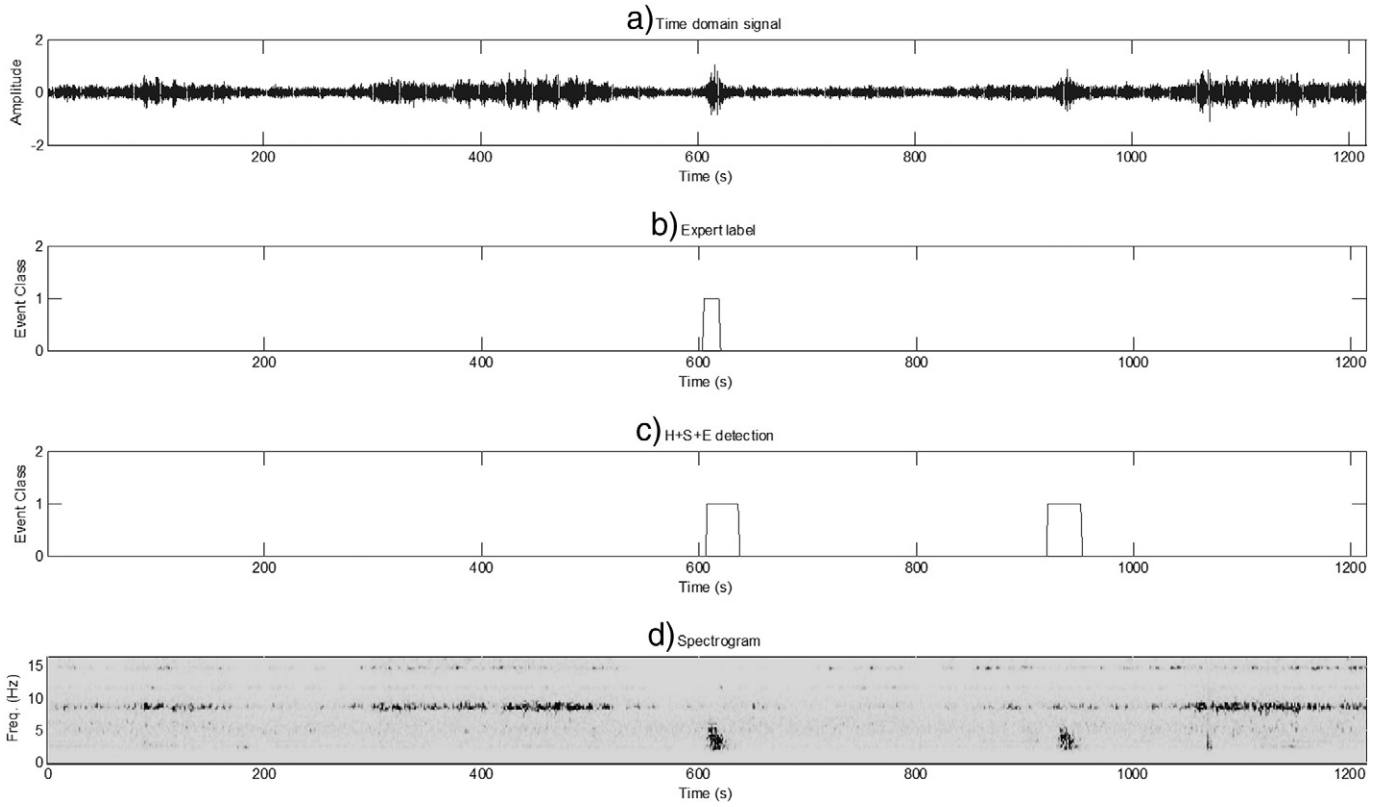

Fig. 8. Example of a signal with a FP: (a) time domain signal of a volcano event; (b) labeling according to the volcano expert, where label 0 is noise, label 1 is LP, and label 2 is VT; (c) detection delivered by the H + S + E system; and, (d) the corresponding spectrogram of the signal.

The results in Fig. 7, and Tables 1 and 2 highlight the advantage achieved due to the introduction of the state and event duration models. As a further investigation, the FPs that were generated by the proposed H + S + E system were analyzed by an expert. It was found that most of the FPs could contain some volcano-seismic activity, as shown in an example event in Fig. 8. In Fig. 8, the top row shows the time domain signal, the second row depicts the corresponding label, the third row displays the detection by the H + S + E method, and the fourth row shows the spectrogram. It can be observed that only one event was labeled by the expert whereas the second event detected by the system contains energy in similar spectral region as seen in the spectrogram. Therefore, based on the recommendation by experts, a signal-to-noise ratio (SNR) threshold was used to analyze the FPs to identify the potential volcano-seismic activity. The SNR is defined as $SNR = \frac{max(|event\,signal(t)|)}{mean(|noise(n)|)}$, where $event\_signal(t)$ is the entire event signal in time domain, $noise(n)$ corresponds to the 3 s signal of noise before the event starts, and |.| is the absolute value. If $SNR > 3$, the detected event can be considered as a potential volcano event. Consequently, the FPs generated by H, H + S, and H + S + E methods with NEP = 27 were evaluated based on the SNR criterion. Accordingly, it was found that 97% of the false alarms show a SNR greater than three, and hence they are potential events. This must be due to the fact that those FP events were detected in the station employed in this paper but not in the other stations that also monitor Llaima volcano. It is worth highlighting that, as a common accepted practice, human experts correlate the information from more than one station to validate real volcano events.

6. Conclusion

In this paper, an automatic volcano event detection system based on HMMs together with state and event duration models is proposed. The system was trained and tested with the volcano-seismic data of the Llaima volcano. Experiments with data from Llaima volcano with two types of events, i.e. long period and volcano-tectonic, show that the incorporation of state and event duration modeling can lead to reductions in false positive rate in event detection as high as 31% with a true positive accuracy equal to 94%. Further evaluation based on the signal-to-noise ratio, recommended by a volcano expert, shows that 97% of the false alarms generated by the proposed system were potential events. The detection and classification of events by making use of the information provided by several stations with duration modeling is proposed for future research.

Acknowledgments

The research reported in this paper was funded by the Chilean National Commission for Scientific and Technological Research (CONICYT), PIA, Anillo project ACT-1120 and FONDEF IDeA CA13I10273. Also many thanks to OVDAS, who provided the data and geological knowledge that supported the study and the analysis of the results.

References

Alasonati, P., Wassermann, J., Ohrnberger, M., 2006. Signal classification by wavelet-based hidden Markov models: application to seismic signals of volcanic origin. In: Mader, H.M., Coles, S.G., Connor, C.B., Connor, L.J. (Eds.), Statistics in VolcanologySpecial Publications of IAVCEI, 1. Geological Society, London, pp. 161–173.

Benítez, M.C., Ramírez, J., Segura, J.C., Ibáñez, J.M., Almendros, J., García-Yeguas, A., Cortés, G., 2007. Continuous HMMs-based seismic event classification at Deception Island, Antarctica. IEEE Trans. Geosci. Remote Sens. 45 (1), 138–146.

Beyreuther, M., Wassermann, J., 2008. Continuous earthquake detection and classification using discrete hidden Markov models. Geophys. J. Int. 175 (3), 1055–1066.

Beyreuther, M., Wassermann, J., 2011. Hidden semi-Markov model based earthquake classification system using weighted finite-state transducers. Nonlinear Process. Geophys. 18 (1), 81–89.

Beyreuther, M., Carniel, R., Wassermann, J., 2008. Continuous hidden Markov models: application to automatic earthquake detection and classification at Las Cañadas caldera, Tenerife. J. Volcanol. Geotherm. Res. 176 (4), 513–518.

Beyreuther, M., Hammer, C., Wassermann, J., Ohrnberger, M., Megies, T., 2012. Constructing a hidden Markov model based earthquake detector: application to induced seismicity. Geophys. J. Int. 189 (1), 602–610.
Bicego, M., Acosta-Munoz, C., Orozco-Alzate, M., 2013. Classification of seismic volcanic signals using hidden-Markov-model-based generative Embeddings. IEEE Trans. Geosci. Remote Sens. 51 (6), 3400–3409.
Chouet, B.A., Matoza, R.S., 2013. A multi-decadal view of seismic methods for detecting precursors of magma movement and eruption. J. Volcanol. Geotherm. Res. 252, 108–175.
Curilem, M., Vergara, J., Martin, C.S., Fuentealba, G., Cardona, C., Huenupan, F., Chacón, M., Khan, M.S., Hussein, W., Yoma, N.B., 2014. Pattern recognition applied to seismic signals of the Llaima volcano (Chile): an analysis of the events' features. J. Volcanol. Geotherm. Res. 282, 134–147.
Esposito, A.M., D'Auria, L., Giudicepietro, F., Peluso, R., Martini, M., 2013. Automatic recognition of landslides based on neural network analysis of seismic signals: an application to the monitoring of Stromboli volcano (southern Italy). Pure Appl. Geophys. 170 (11), 1821–1832.
Gutiérrez, L., Ramírez, J., Benítez, C., Ibañez, J., Almendros, J., García-Yeguas, A., 2006. HMM-based classification of seismic events recorded at Stromboli and Etna volcanoes. Proc. IEEE IGARSS, pp. 2765–2768.
Gutiérrez, L., Ibañez, J., Cortés, G., Ramírez, J., Benítez, C., Tenorio, V., Álvarez, I., 2009. Volcano-seismic signal detection and classification processing using hidden Markov models. Application to San Cristóbal volcano, Nicaragua. Proc. IEEE IGARSS 4, 522–525.
Huang, X.D., Ariki, Y., Jack, M.A., 1990. Hidden Markov Models for Speech Recognition. Vol. 2004. Edinburgh university press, Edinburgh.
Ibáñez, J.M., Benítez, C., Gutiérrez, L.A., Cortés, G., García-Yeguas, A., Alguacil, G., 2009. The classification of seismo-volcanic signals using hidden Markov models as applied to the Stromboli and Etna volcanoes. J. Volcanol. Geotherm. Res. 187 (3/4), 218–226.
Lahr, J.C., Chouet, B.A., Stephens, C.D., Power, J.A., Page, R.A., 1994. Earthquake classification, location and error analysis in a volcanic environment: implications for the magmatic system of the 1989-1990 eruptions at redoubt volcano, Alaska. J. Volcanol. Geotherm. Res. 62, 137–151.
McLachlan, G., Peel, D., 2000. Finite Mixture Models. John Wiley & Sons, Inc., Hoboken, NJ.
Ohrnberger, M., 2001. Continuous Automatic Classification of Seismic Signals of Volcanic Origin at Mt. Merapi, Java, Indonesia (Ph.D. dissertation) Univ. Potsdam, Postdam, Germany.
Oura, K., Zen, H., Nankaku, Y., Lee, A., Tokuda, K., 2008. A fully consistent hidden semi-Markov model-based speech recognition system. IEICE Trans. Inf. Syst. E91-D, 2693–2700.
Rabiner, L.R., 1989. A tutorial on hidden Markov models and selected applications in speech recognition. Proc. IEEE 77 (2), 257–286.
Rabiner, L., Juang, B., 1993. Fundamentals of Speech Recognition. Prentice Hall.
Viterbi, A.J., 1967. Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. IEEE Trans. Inf. Theory 13 (2), 260–269.
Yoma, N.B., McInnes, F.R., Jack, M.A., Stump, S.D., Ling, L.L., 2001. On including temporal constraints in Viterbi alignment for speech recognition in noise. IEEE Trans. Speech Audio Process. 9 (2), 179–182.
Yoma, N.B., Busso, C., Soto, I., 2005. Packet-loss modelling in IP networks with state-duration constraints. IEEE Proc.- Commun. 152 (1).
Zobin, V.M., 2012. Introduction to Volcanic Seismology. second ed. Elsevier.