Supervised and Unsupervised Identification of Concept Drifts in Data Streams of Seismic-Volcanic Signals

Paola Alexandra Castro-Cabrera[1(✉)], Mauricio Orozco-Alzate[1], Cesar Germán Castellanos-Domínguez[1], Fernando Huenupán[2], and Luis Enrique Franco[3]

[1] Universidad Nacional de Colombia - Sede Manizales, Manizales, Colombia
{pacastroc,morozcoa,cgcastellanosd}@unal.edu.co
[2] Universidad de la Frontera (UFRO), Temuco, Chile
[3] Observatorio Vulcanológico de los Andes del Sur, Temuco, Chile

Abstract. The volcanic activity analysis by means of seismic signals is a scenario typically treated by studies in the Artificial Intelligence area under the assumption of invariant probability distribution over time. The literature in geophysics, on the other hand, qualitatively claims that the volcanic phenomenon evolves over long periods of time. This article shows, by three methods, one supervised and two unsupervised, the existence of significant changes in the intrinsic components of the data (*concept drifts*) generated within the volcanic phenomenon. Here it is also shown how the performance of a learning model is considerably affected in a classification task, when concept drifts are not treated in the analysis of a volcanic environment.

Keywords: Data stream · Change detection · Concept drift
Artificial Intelligence · Seismic signal · Volcanology

1 Introduction

The Signal Processing and Artificial Intelligence techniques can rarely be applied in a direct way to data from real-world applications. In dynamic environments, the properties of the generated data are usually time-varying; situation that should be handled with assumptions in order to fit the reality to theoretical frameworks. The most common assumption is that historical data come from the same distribution when actually they tend to change over time, condition known as *concept drift* [5]. Many sensor-based applications generate an increasing volume of data that must be continuously stored and analyzed. In such *data stream* environments, data arrive continuously, and concept drifts cause that patterns and relations in data evolve over time. Then, predictive models designed for either classification or regression tasks in those environments, may become obsolete due to changes in underlying physical processes or in the environment itself [11].

G. R. Simari et al. (Eds.): IBERAMIA 2018, LNAI 11238, pp. 193–205, 2018.
https://doi.org/10.1007/978-3-030-03928-8_16

Volcanic monitoring and the evaluation of the associated risks are issues of continuous study due to the vital importance of risk mitigation related to volcanic unrest. Nowadays, most of the active volcanoes around the world are monitored through different techniques such as geodesy, geochemistry, magnetometry, and specially, through seismological measurements. Seismic signals are the result of particular processes inside the volcano and, therefore, can be used to understand the volcanic phenomenon [1]. Consequently, based on the geophysical literature [1,9,15], the volcanic phenomenon and especially its monitoring by seismology, may be defined as a data stream and changing environment; however, no reference in the engineering area was found where such drifts are quantified or at least identified.

During the last two decades, Machine Learning studies have provided tools and techniques to face the challenge of the volcanic monitoring, by facilitating a time-consuming and repetitive task such as, for instance, the identification of categories and label assignment of registered seismic events. Applications of methods such as Hidden Markov Models, Artificial Neural Networks, Support Vector Machines, among others (see details in [10]), have shown satisfactory results in the state-of-the-art. However, these have been achieved with basic experimental configurations, that is, with small samples sizes taken in a short period of volcanic activity, and without taking into account the chronological order of the examples. Under this arrangement, it is difficult to identify drifts in the volcanic dynamics and, consequently, the classification performance might deteriorate once the learning model is deployed on-line.

This paper is aimed to demonstrate the existence of concept drifts in the volcanic phenomenon, exhibited in its seismic recordings. We maintained the chronological order of the examples to simulate an on-line situation. We use a straightforward supervised method that evaluates the performance of a classifier by its accuracy over time (called DDM). We also tested this hypothesis without the use of class labels through a semi-parametric method of log-likelihood (called SPLL). Additionally, we present a simple but direct method to identify time periods in which there is a change of a so-called *context* (components of an environment with a common dynamics), based on a sum of distances generated with the k nearest neighbor (k-NN) rule. As an illustrative case, this work analyzes the alterations of the dynamics of Villarrica volcano (Chile) over time from its seismic recordings.

2 Concept Drift

In non-stationary environments, the distribution that generates the data can change over time, yielding a phenomenon known as *concept drift*. This scenario requires a different treatment to the traditional one when facing learning tasks. The formal definition of concept drift between times t_0 and t_1 is [5]:

$$\exists \mathbf{x} : p\left(\mathbf{x}, y\right)_{t_0} \neq p\left(\mathbf{x}, y\right)_{t_1} \tag{1}$$

where $p\left(\mathbf{x}, y\right)_{t_0}$ and $p\left(\mathbf{x}, y\right)_{t_1}$ are the joint distributions at times t_0 and t_1, respectively; $\mathbf{x}$ is a vector of input data (an observation or example), and y is the target variable, that for our case is the vector of class labels. In general terms, there is a concept drift if any component of this relation is altered, that is: (1) if the class prior probability $p(y)$ changes, or (2) if the class conditional probability $p(\mathbf{x}|y)$ changes. In consequence, the prediction is affected because the posterior probability of classes $p(y|\mathbf{x})$ may change. In this case, the drift is considered as *real*. Other kinds of concept drifts can be found in [5].

In recent years, the study of change detection in data streams has become more extensive because of its potential applications in real scenarios such as network traffic control, market analysis, fraud detection, medical condition monitoring, among others [5]. This is because most of this contexts may be environments where new sources of data generation arise, and where the new yielded data are detected and react to changes.

3 Methods

3.1 The Drift Detection Method (DDM)

We use the general framework (instead of the experimental one) proposed in [4], denominated drift detection method (DDM). The aim of the proposed method is to detect new "*contexts*", by understanding them as sequences of examples with a stationary distribution. According to DDM, a change in the data distribution occurs when the error increases until reaching a warning level at observation n_i, and a drift level at observation n_j, where $i < j$, in a sequence of n observations. For each point of the sequence, the error rate (E_i) is the probability of misclassification, with standard deviation defined as $s_i = \sqrt{E_i\,(1 - E_i)\,/i}$. Machine Learning theory assumes that, if the data distribution is stationary, the error of a learning model will reduce when the number of observations grows.

3.2 The Semi-parametric Log-Likelihood Detector (SPLL)

The semi-parametric log-likelihood detector (SPLL), proposed in [7], is a method that comes from joining the benefits of two log-likelihood frameworks, namely, the Kullback-Leibler (K-L) criterion and the Hotelling's t^2 test, but overcoming the weaknesses of both criteria and with a greater computational simplicity.

SPLL assumes that all the data come from the same distribution, generated from a Gaussian mixture $p_1(\mathbf{x})$, corresponding to the number of classes or categories (denoted as c), with the same covariance matrix. Two data windows are defined, W_1 and W_2. In the first one, the parameters of the Gaussian mixture are estimated, while in W_2 the change detection criterion is derived using an upper bound of the log-likelihood, which is one standard deviation of the mean of the criterion value in case the two distributions are the same. The criterion is calculated as follows:

$$SPLL\,(W_1, W_2) = \frac{1}{M_2} \sum_{\mathbf{x} \in W_2} \left(\mathbf{x} - \mu_{i*}\right)^{\mathrm{T}} \Sigma^{-1} \left(\mathbf{x} - \mu_{i*}\right) \qquad (2)$$

where $i* = \arg\min_{i=1}^{c} \left\{ (\mathbf{x} - \mu_i)^{\mathrm{T}}\, \Sigma^{-1} \,(\mathbf{x} - \mu_i) \right\}$ is the index of the component with the smallest squared Mahalanobis distance between $\mathbf{x}$ and its center, and M_2 as the amount of observations in W_2.

3.3 Sum of the Distances to the Nearest Neighbor

We propose to represent the differences between observation sets (or batches) through distances by using the k-NN rule with its simplest case: $k = 1$. We do not developed the k-NN rule as far as the label assignment but only until the storing of the distance measured between each observation of the test set and its nearest neighbor, using the Euclidean distance (d_{L_2}) as metric. Within a more formal definition, let $\mathbf{x}_i$ be one observation from the training set and $\mathbf{x}'_j$ one observation from the test set, with $i = 1, ..., N$ and $j = 1, ..., M$.

The distance $D_{1\text{-}NN_j}$ defined for each $\mathbf{x}'_j$ is $D_{1\text{-}NN_j} = \min \left\{ d_{L_2} \left(\mathbf{x}'_j, \mathbf{x}_i \right) \right\}$, and finally, the batch distance is computed as: $D_{Batch} = \sum_{j=1}^{M} D_{1\text{-}NN_j}$.

In summary, for a batch formed by a sliding window of test data, if we add those distances from each observation $\mathbf{x}'_j$ to the one responsible of assigning the label, we can estimate how far the test points are from the training set.

4 Experimental Setup

4.1 Data Set Description

The data used for the experiments are composed by seismic signals from the Villarrica volcano, Chile. This volcano has a seismological network with eight stations which are monitored by *Observatorio Volcanológico de Los Andes del Sur* (OVDAS), located in Temuco and belonging to the *Servicio Nacional de Geología y Minería* of Chile.

The OVDAS distinguishes 9 types of events, whose class labels are assigned by trained analysts, guided by, among others, the visual definition of parameters on the shape and spectral content of the signal, namely: Volcano-Tectonic (VT), Long Period (LP), Tremor (TR), Hybrid (HB), Ice Quakes (IC), Very Long Period (VLP), "Tornillos" (TO), Volcano Distal (DVT) and Tectonics-Seismic events (non-volcanic earthquakes); see a typical description of them in [9]. Three types of seismic events predominate in the seismic activity of Villarrica volcano: LP, TR and VT events, which are considered the most characteristic ones. Therefore, we considered these three classes of events for the experiments. The employed dataset includes seismic signals recorded during 7.3 years (from January 2010 to April 2017), with a total of $317,648$ labeled events, distributed into the three classes as follows: LP $-$ $251,123$; TR $-$ $55,271$ and VT $-$ $2,081$ records. Each class was divided into two groups, one used for training and the other one for testing, depending on the time stamps of the seismic events, as explained in Sect. 4.3.

4.2 Data Processing

Pre-processing. In this work, only records from the vertical axis (Z) of the reference station (triaxial broadband seismometer) assigned by the OVDAS were considered due to its better signal/noise ratio (SNR), compared to the registers from the other two horizontal axes (N, E). Additionally, a data cleansing was carried out, excluding those registers with a low SNR or because the reference station was out of service, remaining a total of 260,365 records. Next, the records were filtered with a Butterworth band-pass filter of order 10 between 0.5 and 15 Hz, because most of the meaningful energy of the events is in this frequency band. After this pre-processing, the number of labeled events per class is: LP − 222,441; TR − 37,735 and VT − 189 records.

Feature Extraction. In this stage several parameters were computed which, according to the literature, have shown to be beneficial for classification tasks. They can be grouped into three families according to the information that they represent:

1. Morphological and geophysical features: those that provide information about the nature of the events; some of them are used by experts in manual classification: signal impulsiveness, event duration and number of zero crossings (dominant frequency of the signal within the trigger time window) [13].
2. Spectral features: which are obtained by converting the seismic register (considered as a time series) into the frequency domain using the Fourier Transform. They are: dominant frequency, mean of the 5 largest frequency peaks and pitch [2]. Other ones were obtained from the spectrogram: frequency range defined by the spectral contour, frequency of spectral contour centroid and frequency of the energy maximum in the spectrogram [13].
3. Transformed feature spaces: These features are obtained through transformations of either the signal waveform or the characterized signal to a domain different from the original one. The features obtained by transformations are: Wavelet energy from level 1 to 7 [2], 5 linear prediction coefficients (LPCs) [3] and 13 cepstral coefficients [6] (see equations of the features in the cited references).

After dividing the whole set into training and test sets, the extracted features were normalized by subtracting the mean (μ) and dividing by the standard deviation ($\mathbf{s}$) of the training set. Next, μ and $\mathbf{s}$ were used to normalize the test set. The characterized dataset was formed storing data in a matrix of 260,365 rows (characterized events) and 34 columns (features).

Feature Selection. In order to find a subset of relevant features that provide greater discrimination in the classification task, we chose a traditional filter-type method called Relief, which uses a criterion independent from the learning algorithm, and provides a feature ranking according to their estimated relevance.

The Relief algorithm defines the quality of a feature (as strongly relevant) according to how well it distinguishes between two examples of different classes and, based on this criterion, it assigns the weight for each feature. However, the original version of Relief is limited because it cannot deal with incomplete data and only works for two-class problems. Its extended version, called ReliefF, is able to solve the aforementioned problems and other additional ones. The ReliefF algorithm searches for a nearest neighbor for each class and finds the feature weights penalizing the ones that give different values to neighbors of the same class, and rewards the features that give different values to neighbors of different classes (see more details in [12]).

Applying the ReliefF algorithm, the initial set of 34 features was reduced to 4, namely: mean of the 5 largest frequency peaks, event duration and wavelet energy at the levels 7 and 2.

4.3 Experiments

Three different methods were applied to demonstrate that the volcanic phenomenon is a changing environment where there are concept drifts, which are observable in the evolution of the seismic data over time.

Definition of Experimental Periods. Authors in geophysics have indicated that alterations in the records of volcanic signals, in relation to other indicators such as geochemical or geodesic parameters (among others), contribute to determine a possible instability period in the dynamics of a volcano, or are even precursors of eruptions [1,15]. In the case of volcanic seismology, the periods of variations in the internal dynamics of a volcano could generate underlying changes, which are evident in indicators such as changes in the number of recorded events, changes in the signal signatures and their intrinsic components (usually the frequency ones), among others [1]. Since the aim of this work is to identify a concept drift in volcanic phenomenon, it was necessary to acquire a priori knowledge about the dynamics of volcanic activity through the records of seismic signals in order to recognize such periods in time series processed with change detection algorithms.

Since the last major eruptions (1984–1985 and 1991 [16]), Villarrica volcano has established a “background” seismic activity, which, in several occasions and for short periods, has increased its levels, but without necessarily ending in an eruptive crisis (the most recent eruption was on March 3, 2015). Such a background activity is permanently studied by the OVDAS. The seismic activity is usually measured by volcanological observatories in terms of Real-Time Seismic Amplitude Measurements (RSAM), Spectral Seismic Amplitude Measurement (SSAM), Reduced Desplacement (RD), Local Magnitude (ML), and number/energy of the events [1]. Since the volcanic activity is dominated by fluid activity (LP and TR), a baseline level measured with the RD could be considered when it is below $10\,\mathrm{cm}^2$ and RSAM varies between 8 and 20 units [14].

According to the analysis carried out by the OVDAS experts about the seismic activity recorded in the Villarrica volcano from 2010 onwards, periods of volcanic stability and instability were defined in coherence with a study of records obtained in the geochemical area. These definitions were based on the congruence of either low or high levels of parameters such as number of daily events recorded per class, RSAM, SSAM and RD −in seismology− and levels of SO_2 −in geochemistry. Thus, the defined periods were:

- Stability: January 1, 2013 to December 31, 2013 (hereinafter referred to as the "Stability Period" frame (*SP frame*)), with $13,857$ examples, from which $9,762$ are LP, $4,065$ are TR and 30 are VT.
- Instability: June 1, 2014 to April 21, 2017 (hereinafter referred to as the "Instability Period" frame (*IP frame*)), with $213,301$ examples, from which $195,941$ are LP, $17,297$ are TR and 63 are VT.

Experiment 1. In Machine Learning, concept drifts are often managed either with weighted examples according to their age or with sliding time windows; in this case, we use the last option, with windows W of fixed size. The experimental procedure (see Sect. 3.1) was developed as in [4] and is summarized as follows:

1. As learning model, we use the k-NN classifier, where the optimal number of neighbors was defined heuristically, testing with $k = \{1, 3, 5, 7, 9, 11\}$ and keeping the W size fixed.
2. The examples were chronologically ordered to emulate the data stream context where new examples arrive to be processed consecutively. The observations within the *IP frame* were arranged as the training set, and those belonging to the *SP frame* as test set.
3. The learning model is generated from training data, and the validation is carried out n times with W_n sliding windows that move over the *IP frame* with an overlap of $\alpha\%$ and a window size M, where, $n = \{1, 2, 3, ..., N\}$, with N equals to the number of times that the window W of size M fits within the *IP frame*. For this experiment, we iteratively tested with $\alpha = \{20, 50, 80\}$ and $M = \{100, 200, 300, ..., N_{SP}\}$, where N_{SP} is the maximum number of examples belonging to the *SP frame* that fit in a single window.

Experiment 2. This test demonstrates how the seismic data of volcanic origin suffer changes in their distribution when they evolve over time, moving from a period of stable activity to a period of volcanic crisis that ended up with an eruption. This test is carried out through the SPLL criterion proposed in [7].

The procedure to implement this experiment was the following:

1. Given two (2) sliding data windows $W1$ and $W2$, a clustering is carried out on $W1$ into cl clusters applying the k-Means algorithm, with $k = 3$ as suggested in [8].
2. A weighted intra-cluster covariance matrix S is calculated.

3. The squared Mahalanobis distance is computed for each example in $W2$ with respect to the cluster centroids.
4. Calculate $SPLL(W1, W2)$, by using Eq. 2, as the mean of the minimum distances computed previously.
5. Swap windows $W1$ and $W2$ and repeat from step 1 to 4: $SPLL(W2, W1)$.
6. Take the maximum of the two values calculated in steps 4 and 5, and find out its chi-squared distribution to define the p-value.

The following considerations were taken into account for executing the experiment:

– Windows $W1$ and $W2$ were defined under the assumption that each one comes from two different probability mass functions. For this purpose, $W2$ was located at the beginning of the *IP frame*, and slided with an overlap $\alpha\%$. $W1$ was set to a fixed single window, centered within the *SP frame*. We ran experiments with $\alpha = \{20, 50, 80\}$.
– The size of the windows is a sensitive parameter in the performance of any change detection algorithm [4,7], so we tested exhaustively with different sizes. Proportional sizes were established for $W1$ and $W2$ windows as suggested in [7] and [8], respectively: $|W1| = |W2|$ and $|W1| = 2 * |W2|$.
– We iteratively tried with different window sizes, $M = \{100, 200, 300, ..., N_{SP}\}$, where N_{SP} is the maximum number of examples belonging to the *SP frame* that fit in a single window. We kept $cl = 3$ for the number of clusters, as recommended in [8].

Experiment 3. With this experiment we intend to show the temporal changes that can occur in the data of a recent period with respect to a previous one, making use of a distance (see Sect. 3.3). Given a window of previously unseen data, $W2$, which arrives in chronological order (test set), a sum of the distances is done from each of them to the prototype responsible for the label assignment in a 1-NN rule. This will serve as an estimation of how much the test data differ (proportional to how far they are) from the prototypes belonging to $W1$ (P_{W1}), where $W1$ is a data window assumed coming from the same distribution, and $i = \{1, 2, 3, ..., |W1|\}$.

For the execution of this experiment, the following considerations were taken into account:

1. As $W1$, a fixed window consisting of the data located in the *SP frame* was defined. $W2$ is a sliding window that moves across the *IP frame*, according to an overlap of $\alpha\%$. Here we tested with $\alpha = \{20, 50, 80\}$.
2. Calculate the distance from each point of $W2$ to its nearest neighbor, using the Euclidean distance as a metric: $D_{1\text{-}NN} = \min\left(d_{L_2}(X_{W2}, P_{W1})\right)$.
3. A representative value for each sliding window $W2$ is obtained from the sum of all the distances calculated for each (X_{W2}).

5 Results and Discussion

In this section, we evaluate the dynamics of the seismic volcanic activity in three different ways in order to determine whether or not there is a concept drift in this environment and, therefore, consider the need for a change detection stage in a classification task. It is not the aim of this article to compare the results for defining the best method for change detection, nor distinguishing between gradual or abrupt changes. The objective is just to obtain evidence of the presence of factors that generate changes in the volcano dynamics.

In order to validate the method used in the first experiment, we employed $13,857$ chronologically ordered observations to train the model, which belong to a period of "quietness" in the volcanic activity, and tested with $4,625$ different test sets with 100 observations ($M = 100$ was the chosen window size because it showed the best resolution in the time), which resulted in sliding the time window with an overlap of $\alpha = 50\%$, that is, forgetting 50 observations for each new test set and including 50 new ones, also chronologically ordered. After trying with different values for k in the k-NN rule, we saw that the performance of the classifier did not show much difference. Results for $k = 1$ are shown below.

Figure 1 shows the deterioration of the classifier performance when the data stream grows and facing an atypical period that produces data generated by a different probability density function. The graph illustrates the predictive error curve pointing out episodes of importance in volcano activity. It is observed how the misclassification shows a significant increase centered between January and May of 2015, period in which the last eruptive crisis took place, which led to the eruption of March 3 of the same year. The error rate of the months of "quietness" of the volcanic activity rose from a maximum of 4% to 27% during the months that the critical phase occurred. Between June and August 2014 there was an important rise of the LP type seismicity (these annotations are documented in reports generated by the OVDAS found in [14]); this period requires a further examination. The last statements are based on the statistical theory which indicates that, if the class distribution is stationary, then the error of the decision model will reduce while the examples increase [4]. Therefore, a significant rise of the error rate of the learning algorithm performance suggests a change in the class distribution and that, consequently, the decision model used has become obsolete. Then, this experiment shows the importance of adding a detection change stage and making a concept drift handling within a classification task.

The subsequent two experiments show that such a change in the volcanic seismology dynamics can be dealt through unsupervised ways. According to the method proposed by Kuncheva (Sect. 3.2), a too large value of the SPLL criterion will indicate a change; however, she justifies that setting up a threshold for defining a change or no change is an aside problem not analyzed in her work. The author points out that such a threshold may be specific for each kind of data, and can be tuned according to the desired level of false and true positives.

According to the latest, if we look at the values of the blue line (and more specifically the red line that represents an average envelope of the blue one), we can notice that Fig. 2 shows changes in the region of the graph between

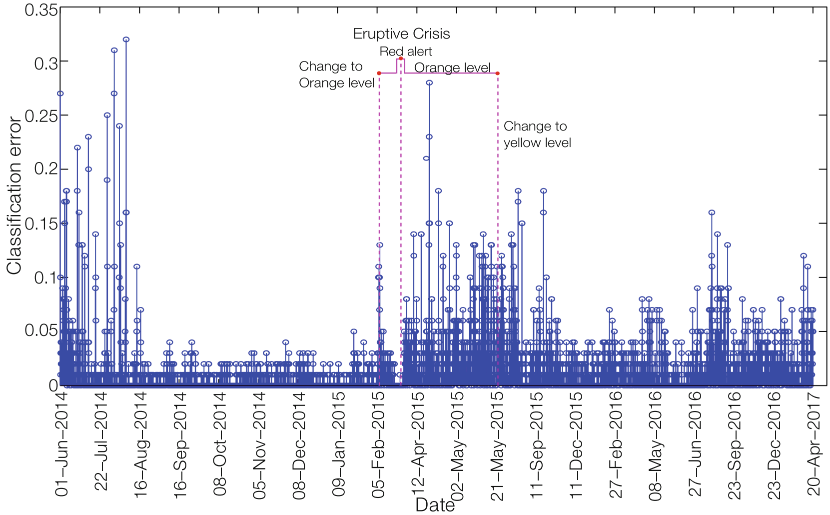

Fig. 1. Deterioration of the classifier accuracy going through different periods of volcano activity. Level of alerts are also indicated (from highest to lowest: red, orange, yellow, green [14])

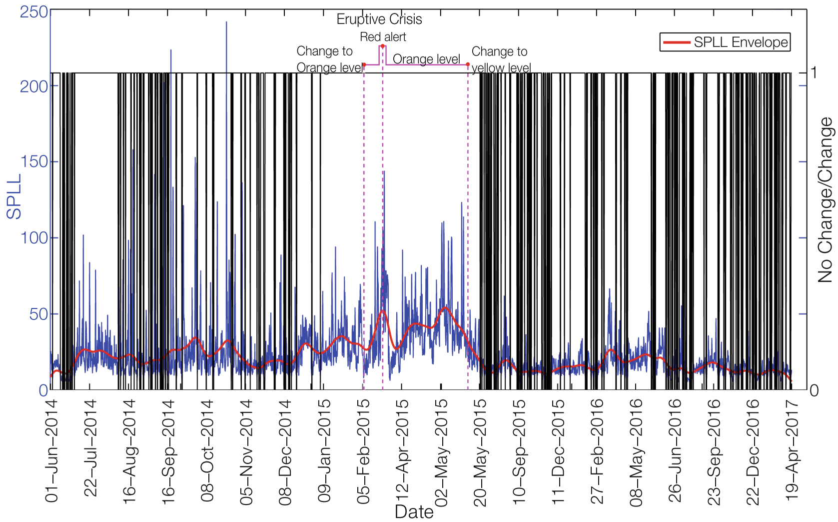

Fig. 2. Definition of a changing environment according to the SPLL criterion on the dynamics of the Villarrica volcano. (Color figure online)

January and May 2015, which coincides with those shown in Fig. 1. The black line indicates the change/no-change criterion. Since we did not make a fine-tuning of the threshold, an alternating variation of such a criterion will be understood as a no change or a small and little significant change.

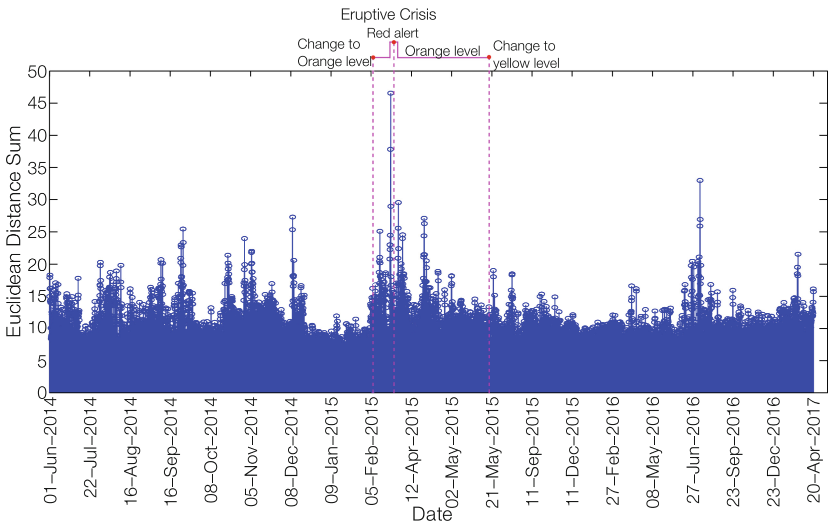

Fig. 3. Identification of significant changes in volcanic activity through the sum of 1-NN distances.

Regarding the third experiment, as shown in Fig. 3, the sum of the distances had the higher values in the region between February and May 2015, in particular, one of them coincides with the eruptive crisis. This is explained as a movement of the points of such period to a region of lower density, which would increase the distance (it can be understood as a mismatch between the compared points) between the events that represents the "future" dynamics of the volcano with respect to the prototypes of the past that within the 1-NN rule would assign the label. If there are changes, then it is presumed that the assignment of the label will be less reliable, and that a decrease in the confidence can be indirectly measured through the distance from the prototype that assigns the label to the test point.

6 Conclusion

Several techniques for the automatic analysis of seismic-volcanic signals have been proposed but they are often limited because of having been designed and trained under assumptions of unchanging data distribution over time. Within a volcanic environment, it was demonstrated how a learning model can increase its error rate up to 23% when the volcano faces a different dynamics, as it happens in phases of eruptive crisis. Additionally, it was demonstrated by two unsupervised methods (one of them very simple and straightforward, proposed by us) that, by setting thresholds, concept drifts can be identified in certain periods that cause changes in contexts. As future work, it is convenient to make a detailed analysis of the deterioration of classification performance by class in order to determine the contributions of each class. Also, the possibilities are open for verifying the improvement of the performance of an automatic system for identifying seismic-volcanic events. This would be possible by adding a stage of change detection which feedbacks the learning model.

Acknowledgment. This work is supported by *Programa Nacional de Fomento a la Formación de Investigadores, Doctorados Nacionales – COLCIENCIAS*. The authors would like to thank OVDAS and UFRO staff for their contributions.

References

1. Carniel, R.: Characterization of volcanic regimes and identification of significant transitions using geophysical data: a review. Bull. Volcanol. **76**(8), 848 (2014)
2. Curilem, M., et al.: Feature analysis for the classification of volcanic seismic events using support vector machines. In: Gelbukh, A., Espinoza, F.C., Galicia-Haro, S.N. (eds.) MICAI 2014. LNCS (LNAI), vol. 8857, pp. 160–171. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-13650-9_15
3. Esposito, A.M., D'Auria, L., Giudicepietro, F., Peluso, R., Martini, M.: 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 (2013)
4. Gama, J., Medas, P., Castillo, G., Rodrigues, P.: Learning with drift detection. In: Bazzan, A.L.C., Labidi, S. (eds.) SBIA 2004. LNCS (LNAI), vol. 3171, pp. 286–295. Springer, Heidelberg (2004). https://doi.org/10.1007/978-3-540-28645-5_29
5. Gama, J., Žliobaitė, I., Bifet, A., Pechenizkiy, M., Bouchachia, A.: A survey on concept drift adaptation. ACM Comput. Surv. (CSUR) **46**(4), 44 (2014)
6. Ibáñez, J.M., Benítez, C., Gutiérrez, L.A., Cortés, G., García-Yeguas, A., Alguacil, G.: The classification of seismo-volcanic signals using Hidden Markov Models as applied to the Stromboli and Etna volcanoes. J. Volcanol. Geoth. Res. **187**(3–4), 218–226 (2009)
7. Kuncheva, L.I.: Change detection in streaming multivariate data using likelihood detectors. IEEE Trans. Knowl. Data Eng. **25**(5), 1175–1180 (2013)
8. Kuncheva, L.I., Faithfull, W.J.: PCA feature extraction for change detection in multidimensional unlabeled data. IEEE Trans. Neural Netw. Learn. Syst. **25**(1), 69–80 (2014)
9. McNutt, S.R., Roman, D.C.: Volcanic seismicity. In: The Encyclopedia of Volcanoes (Second Edition), pp. 1011–1034. Elsevier (2015)
10. Orozco-Alzate, M., Acosta-Muñoz, C., Londoño-Bonilla, J.M.: The automated identification of volcanic earthquakes: concepts, applications and challenges. In: Earthquake Research and Analysis-Seismology, Seismotectonic and Earthquake Geology. InTech (2012)
11. Pears, R., Sakthithasan, S., Koh, Y.S.: Detecting concept change in dynamic data streams. Mach. Learn. **97**(3), 259–293 (2014)
12. Robnik-Šikonja, M., Kononenko, I.: Theoretical and empirical analysis of ReliefF and RReliefF. Mach. Learn. **53**(1–2), 23–69 (2003)
13. Ibs-von Seht, M.: Detection and identification of seismic signals recorded at Krakatau volcano (Indonesia) using artificial neural networks. J. Volcanol. Geoth. Res. **176**(4), 448–456 (2008)
14. SERNAGEOMIN, RNVV, OVDAS: Reportes de actividad volcánica (2015). http://sitiohistorico.sernageomin.cl/volcan.php?pagina=5&iId=22

15. Tárraga, M., Martí, J., Abella, R., Carniel, R., López, C.: Volcanic tremors: good indicators of change in plumbing systems during volcanic eruptions. J. Volcanol. Geoth. Res. **273**, 33–40 (2014)
16. Van Daele, M., et al.: The 600 yr eruptive history of Villarrica volcano (Chile) revealed by annually laminated lake sediments. GSA Bull. **126**(3–4), 481–498 (2014)