

Support vector machines applied to fast determination of the geographical coordinates of earthquakes. The case of El Rosal seismological station, Bogotá - Colombia

Luis H. Ochoa ^a, Luis F. Niño ^b & Carlos A. Vargas ^a

^a Facultad de Ciencias, Universidad Nacional de Colombia, Bogotá, Colombia. lhochoag@unal.edu.co, cavargasj@unal.edu.co

^b Facultad de Ingeniería, Universidad Nacional de Colombia, Bogotá, Colombia. lfniinov@unal.edu.co

Received: October 9th, de 2018. Received in revised form: April 11th, 2019. Accepted: April 22th, 2019.

Abstract

The objective of this research was to determine the latitude and longitude of seismic events using support vector machines (SVMs) and seismic records from El Rosal station, which is located 40 kilometers northwest of Bogotá, Colombia. A total of 504 SVM models were tested to determine latitude and 504 models for longitude, with various combinations of the complexity factor and kernel function exponent, applied to earthquakes of 2, 2.5, 3 and 3.5 ML in time windows of 15, 10 and 5 seconds. The best results showed errors of 40 kilometers for latitude and 30 kilometers for longitude when identifying where tremors occurred. These outcomes might be improved by applying additional descriptors during the SVMs' training stages, such descriptors can be related to Fourier frequency spectra, predominant period and wavelet transform coefficients.

Keywords: earthquake early warning; rapid response; earthquake geographic coordinates; latitude, longitude; seismic event; Bogotá - Colombia; support vector machine (SVM); seismología.

Máquinas de vectores de soporte aplicadas a la determinación rápida de coordenadas geográficas de terremotos. Caso de la estación sismológica El Rosal, Bogotá - Colombia

Resumen

El objetivo de esta investigación fue determinar la latitud y la longitud de eventos sísmicos utilizando algoritmos de máquinas de vectores de soporte (MVS) y registros sísmicos de la estación "El Rosal", ubicada a 40 kilómetros al noroeste de Bogotá. Un total de 504 modelos de MVS fueron probados para determinar la latitud y 504 modelos para la longitud, con varias combinaciones de factor de complejidad y exponente de la función kernel, aplicados a terremotos de 2, 2.5, 3 y 3.5 ML en ventanas de tiempo de 15, 10 y 5 segundos. Los mejores resultados mostraron errores de 40 kilómetros para la latitud y 30 kilómetros para la longitud, con respecto al lugar donde se generaron los terremotos. Estos resultados podrían mejorarse mediante la aplicación de descriptores adicionales durante las etapas de entrenamiento de las MVS, dichos descriptores pueden estar relacionados con los espectros de frecuencia de Fourier, el período predominante y los coeficientes de transformada de la ondícula.

Palabras clave: alerta temprana de terremoto; respuesta rápida; coordenadas geográficas; latitud; longitud; evento sísmico; Bogotá - Colombia; máquina de vector de soporte (MVS); sismología.

1. Introduction

This study is part of a line of research which proposes that earthquake hypocentral parameters can be calculated by

applying artificial intelligence methods, in order to develop an early warning system for the city of Bogotá. Almost a third of Colombia's population live on the Bogotá savanna and surrounding areas, as it is the main economic center of

How to cite: Ochoa, L.H., Niño, L.F. and Vargas C.A., Support vector machines applied to fast determination of the geographical coordinates of earthquakes. The case of El Rosal seismological station, Bogotá - Colombia. DYNA, 86(209), pp. 230-237, April - June, 2019.

the country producing almost 40% of the gross domestic product [1]. For this reason a seismic early warning system for Bogotá is very important, and geographic latitude and longitude are some of the main parameters for this system. The task of an early warning system is to estimate earthquake hypocentral parameters in a short period of time with the best possible accuracy [2]. Contrary to prediction, seismic early warning systems emit an alert a few seconds after the event initiates, and from few seconds to a few tens of seconds before the stronger shaking movement starts.

Many early warning systems employ dense seismological networks to determine magnitude and localization using at least three stations with a very good accuracy [3-7], however, the density of seismological stations around Bogota is not high enough and they are too far away to conform to the travel time requirement for seismic events localization. A solution to this problem is to use seismological data from previous events recorded at one single station to calculate the earthquake hypocentral parameters [8].

The common way to calculate hypocentral parameters consists of applying velocity models for different earth rock layers and processing travel time signals of P and S waves recorded by seismic stations [9]. In recent years, alternative approaches based on machine learning techniques have been developed, most of them using genetic algorithms (GA) and fuzzy logic (FL). The FL approaches allow for the efficient exploration of the search area [10], while the GA are mainly used to determine the X, Y, Z coordinates of earthquake hypocenters [11]. Automatic computation algorithms in a single broadband three-component station have principally been developed for P and S waves onset detection, estimating earthquake location and the apparent surface speed [12-14] or seismic moment estimation [2,4,5,15-17]. Supervised machine learning techniques based on kernel methods have become a powerful tool for mathematicians, scientists and engineers, providing solutions in areas such as signal processing and pattern recognition.

Their implementation is quite simple and can be performed by applying mathematical functions that combine

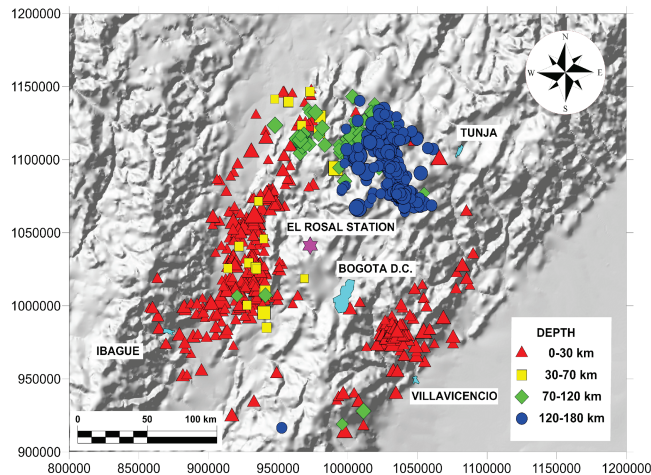


Figure 1. Location of El Rosal seismological station and earthquake distribution around Bogotá. Coordinates Gauss Kruger - Bogotá Origin Source: The Authors.

input variables as combinations of themselves, obtaining a new, enhanced space with more dimensions that facilitates the separation of classes. The methodology proposed in this study consists of applying support vector machines (SVMs) along with kernel functions in order to determine geographic latitude and longitude with the minimal processing of data acquired at the station, similar to the methodology applied in the fast determination of earthquake magnitude [18], epicenter distance [19] and depth [20] using a single seismological station.

2. Data set used and methods applied

In this research, the data set was collected at El Rosal seismological station, located northwest of Bogotá as Fig. 1 shows. This station is part of the Colombian Seismic Network managed by Servicio Geológico Colombiano - SGC (Colombian Geological Service).

The El Rosal station employs a Guralp CMG - T3E007 sensor in three components and a Nanometrics RD3-HRD24 digitizer, which provides simultaneous sampling of three channels with 24-bit resolution [21]. The data corresponds to three component raw waveforms recorded directly at this station and a seismic catalogue with 2164 selected events, from between January 1st 1998 and October 27th 2008; all of them located less than 120 kilometers from the station. The Colombian seismic network consists of 42 stations, with an average distance of 162 kilometers between them, which record and transmit seismic data in real time for the entire country.

2.1. Data pre-processing

Before starting to process the SVMs, waveform files from El Rosal station were converted to the American standard code for information interchange (ASCII) format, using a Seisan package tool; earthquakes with magnitudes lower than 2.0 ML were discarded and so the following processes were applied to the remaining 863 events. Since selected seismic records present varying levels of noise, it was necessary to filter it out with both high and low frequency filters. Low frequencies correspond to instrumental noise that can be easily eliminated using high-pass filters with a cut-off frequency of 0.075 Hz [22], while high frequencies were removed with low-pass filters with a cut-off frequency of 150 Hz.

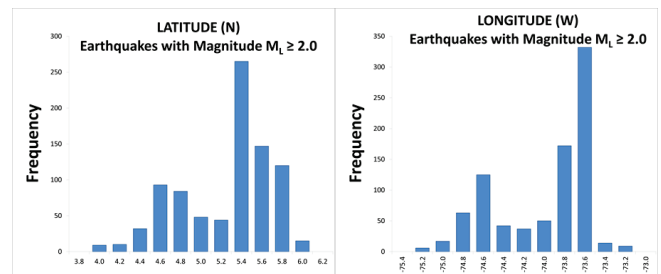


Figure 2. Statistical distribution of earthquake geographic coordinates recorded at El Rosal station. Source: The Authors.

The statistical distribution of geographic latitude and longitude is presented in Fig. 2, where the main distribution of the whole data set is observed. These histograms show bimodal patterns in both cases, suggesting different active seismogenic zones north-south and east-west of El Rosal station that produce the regular seismic behavior of the area.

2.2. Descriptor - Input data set of SVMs

In the first stage, parameters that have been previously used by other authors for earthquake magnitude estimation were calculated and employed as input variables or descriptors for the SVMs in this study. In this sense, the relationship between maximum P wave amplitudes and local earthquake magnitudes was considered [23], where a linear regression was performed for each of the station’s three components. Three parameters were taken from these linear regressions which correspond to slope M, independent term B and correlation coefficient R. The maximum amplitude values Mx obtained for each time window were also used as descriptors; therefore, each event had 12 descriptors related to this concept.

Second, 9 descriptors employed for epicenter distance estimation were added and adjust a linear regression of an exponential function in time *t* by applying the expression “ $Bt \exp(-At)$ ”; this expression belongs to the envelope of the seismic record in logarithmic scale [2] which is also determined by a linear regression and its respective correlation coefficient R, for each component in the seismic station. The correlation coefficient R along with the parameters A and B were also calculated; where B represents the slope of the initial part of the P waves and A is a parameter related to amplitude variations in time.

Finally, parameters for the determination of back-azimuth were used to include information about the source location of seismic events in the model. Maximum eigenvalues of a two-dimensional covariance matrix were employed as input, calculated as described in [12] and [24]. A windowing scheme with one second time windows was performed to obtain consecutive values for which a linear regression was calculated, also determining the slope M, independent term B, regression correlation factor R, and arithmetic mean of eigenvalues PP.

Table 1. Summary for the best epicenter distance models in each combination.

Related to Magnitude.	Related to Epicenter Distance.	Related to Back-Azimuth.
Slope (M)		Slope (M)
Independent Term (B)	Parameter (A)	Independent Term (B)
Correlation Coefficient (R)	Parameter (B)	Correlation Coefficient (R)
Maximum Amplitude (Mx)	Correlation Coefficient (R)	Maximum Amplitude (Mx)
12 Descriptors (4 in each component of the station).	9 Descriptors (3 in each component of the station).	Eigenvalues (P)
		4 Descriptors (all components work at the same time).

Source: The Authors.

This last process works with all components of the station at the same time, thus four descriptors were added to this process as input.

In sum, the SVMs of this study employ 25 time signal descriptors as input (Table 1); 12 of them related to works on magnitude calculation, 9 were associated with epicenter distance estimations and the last 4 were used for back-azimuth determination. These descriptors were calculated for 5, 10 and 15 second signals of the 863 selected events.

2.3. The SVM models

The SVMs are a group of supervised learning algorithms related to classification and regression problems. When a sample is used in training, it can be separated into classes and so train a SVM to predict the classes of a new sample; an SVM represents the points of a sample in a space, separating classes within these points into the widest possible spaces. When new samples are projected onto this model, they can be classified into any class in function of proximity of the points. The SVM models applied in this study are based on the complexity factor C and the selected kernel function. The complexity factor regulates the accuracy of the model; this factor can support the proper training of the model (generalization), or else, it can reach a point of overfitting. A proper generalization allows the model to accurately classify several samples that are different from those employed during the training stage; moreover, overfitting occurs when the model can only classify correctly the sample used in training.

The Kernel function projects a data set on a space of specific characteristics and uses algorithms related to linear algebra, geometry and statistics to identify linear patterns in the dataset. Any solution using kernel methods comprises two phases; the first phase consists of a module that maps the projected data; the second phase consists of an algorithm designed to detect linear patterns in the space where this data is projected [25]. The kernel function applied in this study was a polynomial type, using Equation 1.

$$K(x, y) = [(x, y) + 1]^E \tag{1}$$

Where E is a parameter representing the kernel exponent and K represents the kernel function depending on variables *x* and *y*. The kernel exponent E and the complexity factor C must be provided by the users in each model.

In this study, the models were trained with the refined data set for each time window using the Waikato Environment for Knowledge Analysis WEKA 3.6 and the 25 descriptors explained above (Descriptor - Input data set of SVMs). These algorithms have strong statistical support and can easily be implemented at the station by electronic processing cards. In order to choose the E and C parameters, correlation factors and minimum absolute error obtained by a 10-fold cross-validation process were compared for geographic latitude and longitude. These processes were carried out by testing multiple combinations of E and C for selected earthquake magnitudes and time signals. The correlation coefficient calculated for each partition corresponds to the Pearson’s Coefficient, which measures the

linear relationship between two variables independently of their scales. This coefficient uses values between -1 and 1; a value of zero means that a linear relationship between two variables could not be found. A relationship with a positive value means that two variables change in a similar way, i.e., high values of one variable correspond to high values of the other and vice versa. The closer this value is to one, the greater certainty that two variables have a linear relationship.

3. Results

Using the 25 descriptors and earthquake magnitudes for each seismic event, a group of 24 datasets was evaluated (Tables 2A and 2B). Each dataset corresponds to combinations of 4 minimum magnitude filters (2.0, 2.5, 3.0 and 3.5 ML) and 3 signal length filters (5, 10 and 15 seconds), evaluating combinations of 7 values for the kernel exponent ($E = 1.5, 2, 4, 5, 10, 20$ and 50) and 6 values for the complexity factor ($C = 1, 3, 5, 10, 20$ and 50); testing 1008 models of SVMs, which correspond to 504 models for latitude and 504 models for longitude in order to find the combination of parameters with the best correlation factor. Tables 3A and 3B show a statistical summary for the best models of latitude and longitude in each combination of time signals and magnitudes.

Tables 2A and 3A demonstrate that the best combinations of parameters for earthquake latitude determination are 10 for E and 2 for C using a cut-off magnitude of ≥ 2.5 ML and a time signal of 5 seconds; with these combinations, the best correlation factor is 0.32, reaching 0.36 degrees of standard deviation, which correspond to approximately 40 kilometers of error. Fig. 3 shows the cross-plot with the relationship between real latitude (X axis) and latitude calculated by the model (Y axis), where a normal statistical pattern can be observed in the distribution of residual values (histogram). The dashed blue line represents the linear behavior of the predicted data, corresponding to the locus where the prediction is equal to real values; this plot also confirms a standard deviation of 0.35 degrees in latitude determination. Employing the same methodology for longitude estimation, the best combinations are 2 for E and 5 for C , using a cut-off magnitude of ≥ 3.0 ML and a time signal of 5 seconds (Tables 2B and 3B). The standard deviation is 0.31 degrees, equivalent to approximately 30 kilometers, also confirmed in the cross-plot in Fig. 4.

4. Conclusions and recommendations

The SVMs used for this research provide part of the required information to develop an early warning system for seismicity that may affect the city of Bogotá. These algorithms showed an error of 40 and 30 kilometers in the determination of the latitude and longitude of earthquakes respectively.

The accuracies in this study were lower than those obtained by other authors [26-29], who achieved an accuracy of 7 kilometers on average. However, it is important to note that this study was carried out with information from a single seismological station and additional descriptors such as the predominant period. Fourier frequency spectra and wavelet

transform coefficients should be considered in order to extract additional features allowing for better correlation and better estimation of the geographic coordinates where tremors occurred.

It is recommended that the models be complemented with earthquake data that was not considered in this research, particularly events since October 27th 2008 to the present, providing a greater data set and improving the accuracy of these SVMs.

5. Acknowledgments

The authors would like to thank Servicio Geológico Colombiano (SGC) for providing the data set used in this study and Universidad Nacional de Colombia for supporting our efforts to achieve a fast and reliable early warning system for Bogotá D.C. - Colombia.

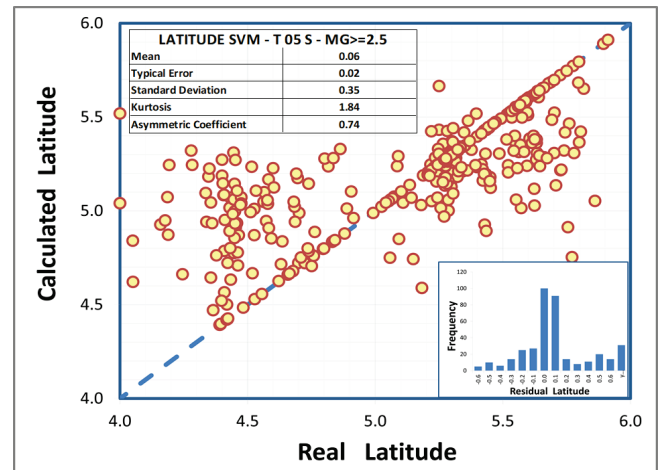


Figure 3. Correlation between real and calculated geographic latitude with SVM.

Source: The Authors.

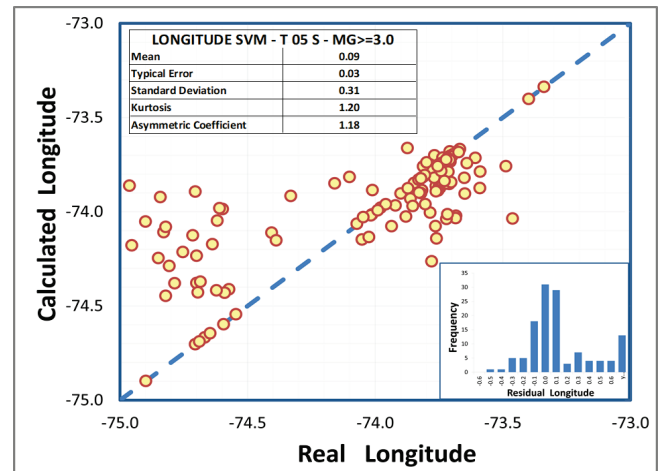


Figure 4. Correlation between real and calculated geographic longitude with SVM.

Source: The Authors.

Table 2A.

Cut-off Magnitude and Time Length combination (Geographic Latitude).

Pearson's Coefficient for Each Combination of Magnitude and Time Signal																
		Magnitude $\geq 2.0 M_L - E$							Magnitude $\geq 2.5 M_L - E$							
Time Signal = 15 s - C		1.5	2	4	5	10	20	50	1	1.5	2	4	5	10	20	50
	1		0.426	0.440	0.474	0.480	0.471	0.493	0.457	1	0.310	0.336	0.379	0.388	0.392	0.393
3		0.446	0.467	0.478	0.469	0.478	0.479	0.410	3	0.363	0.380	0.393	0.390	0.373	0.358	0.299
5		0.461	0.472	0.471	0.465	0.481	0.466	0.392	5	0.378	0.389	0.391	0.391	0.371	0.319	0.305
10		0.470	0.478	0.463	0.463	0.474	0.432	0.376	10	0.392	0.397	0.389	0.376	0.361	0.271	0.306
20		.478	0.479	0.459	0.463	0.465	0.384	0.355	20	0.398	0.395	0.372	0.353	0.317	0.253	0.306
50		0.481	0.476	0.458	0.457	0.423	0.327	0.352	50	0.399	0.389	0.350	0.344	0.242	0.251	0.306
Time Signal = 10 s - C		1.5	2	4	5	10	20	50	1	1.5	2	4	5	10	20	50
	1		0.329	0.354	0.383	0.395	0.412	0.411	0.399	1	0.232	0.248	0.284	0.297	0.344	0.353
3		0.356	0.375	0.395	0.398	0.398	0.386	0.343	3	0.247	0.274	0.325	0.340	0.337	0.350	0.314
5		0.370	0.381	0.394	0.395	0.380	0.373	0.328	5	0.272	0.294	0.338	0.339	0.314	0.323	0.310
10		0.380	0.392	0.388	0.393	0.359	0.351	0.318	10	0.298	0.311	0.334	0.320	0.320	0.283	0.310
20		0.388	0.390	0.386	0.376	0.331	0.309	0.315	20	0.310	0.316	0.313	0.292	0.313	0.274	0.310
50		0.392	0.377	0.354	0.324	0.302	0.268	0.311	50	0.308	0.314	0.274	0.287	0.255	0.269	0.310
Time Signal = 05 s - C		1.5	2	4	5	10	20	50	1	1.5	2	4	5	10	20	50
	1		0.274	0.280	0.289	0.290	0.278	0.245	0.207	1	0.263	0.272	0.289	0.286	0.288	0.309
3		0.278	0.283	0.286	0.288	0.258	0.206	0.176	3	0.272	0.286	0.294	0.290	0.321	0.314	0.314
5		0.276	0.283	0.288	0.285	0.235	0.186	0.169	5	0.277	0.289	0.295	0.309	0.312	0.317	0.315
10		0.276	0.278	0.285	0.268	0.202	0.158	0.144	10	0.275	0.289	0.321	0.324	0.309	0.295	0.302
20		0.272	0.276	0.267	0.249	0.171	0.127	0.137	20	0.269	0.291	0.321	0.312	0.304	0.279	0.293
50		0.271	0.282	0.236	0.201	0.142	0.111	0.133	50	0.279	0.289	0.303	0.308	0.268	0.260	0.291
		Magnitude $\geq 3.0 M_L - E$							Magnitude $\geq 3.5 M_L - E$							
Time Signal = 15 s - C		1.5	2	4	5	10	20	50	1	1.5	2	4	5	10	20	50
	1		0.095	0.079	0.170	0.184	0.157	0.183	0.270	1	0.293	0.258	0.264	0.253	0.266	0.180
3		0.082	0.146	0.153	0.165	0.155	0.216	0.269	3	0.262	0.276	0.316	0.325	0.341	0.235	-0.08
5		0.113	0.190	0.183	0.184	0.150	0.220	0.269	5	0.261	0.270	0.336	0.347	0.375	0.235	-0.08
10		0.172	0.186	0.206	0.197	0.177	0.217	0.269	10	0.267	0.321	0.357	0.389	0.375	0.235	-0.08
20		0.174	0.200	0.216	0.167	0.190	0.217	0.269	20	0.317	0.327	0.406	0.409	0.375	0.235	-0.08
50		0.206	0.255	0.166	0.177	0.183	0.217	0.269	50	0.302	0.332	0.407	0.409	0.375	0.235	-0.08
Time Signal = 10 s - C		1.5	2	4	5	10	20	50	1	1.5	2	4	5	10	20	50
	1		0.048	0.047	0.222	0.282	0.298	0.240	0.257	1	0.229	0.215	0.250	0.261	0.314	0.232
3		0.063	0.151	0.296	0.415	0.253	0.225	0.272	3	0.242	0.467	0.332	0.347	0.388	0.262	-0.02
5		0.106	0.233	0.315	0.338	0.209	0.245	0.272	5	0.258	0.276	0.355	0.424	0.388	0.262	-0.02
10		0.200	0.255	0.336	0.296	0.200	0.256	0.272	10	0.261	0.302	0.470	0.464	0.388	0.262	-0.02
20		0.232	0.265	0.287	0.213	0.208	0.256	0.272	20	0.297	0.338	0.470	0.464	0.388	0.262	-0.02
50		0.252	0.340	0.190	0.194	0.225	0.256	0.272	50	0.319	0.438	0.470	0.464	0.388	0.262	-0.02
Time Signal = 05 s - C		1.5	2	4	5	10	20	50	1	1.5	2	4	5	10	20	50
	1		0.074	0.056	0.029	0.037	0.153	0.154	0.135	1	0.356	0.312	0.241	0.198	0.128	0.037
3		0.005	-0.01	0.054	0.147	0.193	0.115	0.130	3	0.296	0.263	0.215	0.207	0.152	0.092	-0.14
5		-0.03	-0.02	0.142	0.194	0.146	0.144	0.130	5	0.264	0.270	0.228	0.246	0.150	0.092	-0.14
10		-0.04	-0.02	0.197	0.219	0.070	0.165	0.130	10	0.305	0.268	0.230	0.154	0.150	0.092	-0.14
20		-0.05	0.034	0.199	0.150	0.123	0.171	0.130	20	0.297	0.239	0.128	0.135	0.150	0.092	-0.14
50		0.00	0.124	0.086	0.058	0.154	0.171	0.130	50	0.228	0.168	0.128	0.135	0.150	0.092	-0.14

Source: The Authors.

Table 2B.
Cut-off Magnitude and Time Length combination (Geographic Longitude).

Pearson's Coefficient for Each Combination of Magnitude and Time Signal																
		Magnitude $\geq 2.0 M_L - E$							Magnitude $\geq 2.5 M_L - E$							
Time Signal = 15 s - C		1.5	2	4	5	10	20	50	1	1.5	2	4	5	10	20	50
		1	0.503	0.528	0.572	0.585	0.629	0.635	0.594	1	0.521	0.544	0.592	0.604	0.652	0.640
	3	0.530	0.559	0.604	0.618	0.631	0.614	0.546	3	0.559	0.580	0.618	0.645	0.648	0.611	0.567
	5	0.544	0.568	0.615	0.625	0.631	0.591	0.516	5	0.567	0.584	0.640	0.656	0.633	0.600	0.552
	10	0.561	0.585	0.621	0.624	0.621	0.563	0.482	10	0.576	0.576	0.654	0.651	0.609	0.577	0.546
	20	0.573	0.599	0.620	0.622	0.596	0.521	0.450	20	0.567	0.595	0.643	0.637	0.575	0.549	0.545
	50	0.589	0.600	0.612	0.600	0.545	0.443	0.442	50	0.585	0.631	0.626	0.612	0.516	0.512	0.545
Time Signal = 10 s - C		1.5	2	4	5	10	20	50	1	1.5	2	4	5	10	20	50
		1	0.475	0.500	0.536	0.546	0.583	0.590	0.559	1	0.490	0.507	0.554	0.568	0.591	0.571
	3	0.507	0.524	0.560	0.578	0.597	0.577	0.516	3	0.522	0.537	0.580	0.589	0.586	0.552	0.463
	5	0.518	0.530	0.575	0.584	0.593	0.566	0.495	5	0.526	0.547	0.587	0.594	0.573	0.538	0.445
	10	0.525	0.544	0.580	0.584	0.582	0.535	0.478	10	0.541	0.553	0.587	0.595	0.548	0.493	0.445
	20	0.536	0.559	0.577	0.584	0.564	0.497	0.459	20	0.541	0.564	0.589	0.580	0.519	0.436	0.445
	50	0.554	0.564	0.568	0.564	0.525	0.452	0.452	50	0.551	0.564	0.566	0.545	0.466	0.390	0.445
Time Signal = 05 s - C		1.5	2	4	5	10	20	50	1	1.5	2	4	5	10	20	50
		1	0.416	0.429	0.448	0.455	0.466	0.459	0.366	1	0.360	0.398	0.439	0.442	0.477	0.446
	3	0.446	0.448	0.463	0.468	0.471	0.422	0.329	3	0.424	0.433	0.458	0.472	0.467	0.399	0.370
	5	0.449	0.454	0.470	0.475	0.464	0.407	0.304	5	0.430	0.435	0.473	0.476	0.444	0.386	0.358
	10	0.452	0.457	0.474	0.466	0.440	0.384	0.286	10	0.433	0.438	0.470	0.466	0.419	0.379	0.349
	20	0.453	0.464	0.458	0.454	0.408	0.361	0.267	20	0.431	0.458	0.449	0.442	0.386	0.356	0.342
	50	0.459	0.463	0.430	0.422	0.376	0.293	0.257	50	0.446	0.461	0.422	0.411	0.352	0.325	0.343
Time Signal = 15 s - C		Magnitude $\geq 3.0 M_L - E$							Magnitude $\geq 3.5 M_L - E$							
		1.5	2	4	5	10	20	50	1	1.5	2	4	5	10	20	50
	1	0.554	0.578	0.621	0.638	0.610	0.564	0.526	1	0.597	0.606	0.620	0.636	0.549	0.441	0.271
	3	0.603	0.615	0.644	0.638	0.604	0.537	0.526	3	0.609	0.628	0.583	0.523	0.427	0.428	0.271
	5	0.615	0.622	0.638	0.638	0.577	0.542	0.526	5	0.631	0.675	0.495	0.457	0.416	0.428	0.271
	10	0.610	0.621	0.638	0.621	0.557	0.543	0.526	10	0.673	0.599	0.394	0.360	0.416	0.428	0.271
	20	0.603	0.629	0.600	0.590	0.553	0.543	0.526	20	0.609	0.495	0.314	0.339	0.416	0.428	0.271
	50	0.601	0.582	0.575	0.560	0.558	0.543	0.526	50	0.485	0.325	0.308	0.339	0.416	0.428	0.271
Time Signal = 10 s - C		1.5	2	4	5	10	20	50	1	1.5	2	4	5	10	20	50
		1	0.005	0.523	0.559	0.572	0.616	0.610	0.481	1	0.559	0.582	0.595	0.594	0.586	0.593
	3	0.526	0.555	0.582	0.601	0.630	0.578	0.479	3	0.591	0.572	0.550	0.547	0.588	0.593	0.425
	5	0.545	0.571	0.590	0.609	0.62	0.551	0.479	5	0.568	0.531	0.27	0.535	0.588	0.593	0.425
	10	0.557	0.548	0.581	0.580	0.586	0.548	0.479	10	0.494	0.476	0.511	0.539	0.588	0.593	0.425
	20	0.535	0.539	0.546	0.564	0.522	0.548	0.479	20	0.425	0.448	0.516	0.539	0.588	0.593	0.425
	50	0.513	0.513	0.531	0.513	0.510	0.548	0.479	50	0.389	0.412	0.516	0.539	0.588	0.593	0.425
Time Signal = 05 s - C		1.5	2	4	5	10	20	50	1	1.5	2	4	5	10	20	50
		1	0.379	0.400	0.480	0.471	0.462	0.426	0.287	1	0.588	0.613	0.579	0.555	0.479	0.334
	3	0.477	0.519	0.476	0.485	0.469	0.421	0.286	3	0.665	0.637	0.542	0.515	0.424	0.279	-0.03
	5	0.519	0.536	0.499	0.496	0.444	0.399	0.286	5	0.681	0.638	0.526	0.526	0.418	0.279	-0.03
	10	0.533	0.512	0.475	0.449	0.446	0.389	0.286	10	0.690	0.597	0.521	0.478	0.418	0.279	-0.03
	20	0.530	0.535	0.395	0.399	0.427	0.391	0.286	20	0.632	0.543	0.494	0.486	0.418	0.279	-0.03
	50	0.545	0.415	0.342	0.361	0.417	0.391	0.286	50	0.547	0.512	0.495	0.486	0.418	0.279	-0.03

Source: The Authors.

Table 3A.
Summary for the best model of geographic latitude in each combination.

SUMMARY IN DETERMINATION OF EARTHQUAKE LATITUDE - RESIDUAL												
	Signal Time = 15 s				Signal Time = 10 s				Signal Time = 05 s			
	2.0	2.5	3.0	3.5	2.0	2.5	3.0	3.5	2.0	2.5	3.0	3.5
Minimum Magnitude	2.0	2.5	3.0	3.5	2.0	2.5	3.0	3.5	2.0	2.5	3.0	3.5
Mean	0.03	0.04	0.04	0.00	0.04	0.04	0.04	0.07	0.09	0.06	0.03	0.08
Typical Error	0.01	0.02	0.02	0.05	0.01	0.02	0.02	0.05	0.01	0.02	0.02	0.05
Standard Deviation	0.4	0.4	0.2	0.3	0.4	0.4	0.2	0.3	0.4	0.3	0.2	0.3
Kurtosis	1.29	2.10	14.84	0.84	1.60	3.00	7.47	0.03	0.57	1.84	7.82	0.28
Asymmetric Coefficient	0.59	0.76	3.48	-0.12	0.77	0.79	2.22	0.42	0.83	0.74	1.69	0.49
Count	863	379	126	33	863	379	126	33	864	377	126	33
Level of Confidence (95%)	0.025	0.038	0.039	0.101	0.024	0.036	0.043	0.100	0.027	0.035	0.040	0.102

Source: The Authors.

Table 3B.
Summary for the best model of geographic longitude in each combination.

Summary in determination of earthquake longitude - Residual												
	Signal Time = 15 s				Signal Time = 10 s				Signal Time = 05 s			
	2.0	2.5	3.0	3.5	2.0	2.5	3.0	3.5	2.0	2.5	3.0	3.5
Minimum Magnitude	2.0	2.5	3.0	3.5	2.0	2.5	3.0	3.5	2.0	2.5	3.0	3.5
Mean	0.03	0.06	0.07	0.00	0.05	0.07	0.01	0.03	0.11	0.09	0.09	0.02
Typical Error	0.01	0.01	0.02	0.03	0.01	0.02	0.01	0.04	0.01	0.02	0.03	0.04
Standard Deviation	0.3	0.3	0.2	0.2	0.3	0.3	0.1	0.2	0.4	0.3	0.3	0.2
Kurtosis	4.34	4.53	1.19	5.19	4.81	8.12	17.28	6.16	1.11	3.55	1.20	5.44
Asymmetric Coefficient	0.99	1.63	1.11	0.64	1.56	2.39	3.22	2.12	1.09	1.69	1.18	1.15
Count	863	379	126	33	863	379	126	33	864	377	126	33
Level of Confidence (95%)	0.019	0.026	0.038	0.067	0.019	0.030	0.023	0.076	0.024	0.033	0.056	0.072

Source: The Authors.

References

- [1] Ojeda, A., Martinez, S., Bermudez, M. and Atakan, K., The new accelerograph network for Santa Fe de Bogota, Colombia. *Soil Dynamics and Earthquake engineering*, 22(9-12), pp. 791-797, October-December, 2002. DOI: 10.1016/S0267-7261(02)00100-8
- [2] Odaka, T. et al., A new method for quickly estimating epicentral distance and magnitude from a single seismic record. *Bulletin of the Seismological Society of America*, 93(1), pp. 526-532, February, 2003. DOI: 10.1785/0120020008
- [3] Anderson, J.G. and Chen, Q., Beginnings of earthquakes in the Mexican subduction zone on strong-motion accelerograms. *Bulletin of the Seismological Society of America*, 85(4), pp. 1107-1115, August, 1995. DOI: /85/4/1107/102623
- [4] Espinosa-Aranda, J.M. et al., Mexico city seismic alert system. *Seismol. Res. Lett.*, 66(6), pp. 42-53, November, 1995. DOI: 10.1785/gssrl.66.6.42
- [5] Wu, Y.M., Shin, T.C. and Tsai, Y.B., Quick and reliable determination of magnitude for seismic early warning. *Bulletin of the Seismological Society of America*, 88(5), pp. 1254-1259, October, 1998.
- [6] Wu, Y.M. and Teng, T.L., A virtual Sub-Network approach to earthquake early warning. *Bulletin of the Seismological Society of America*, 92(5), pp. 2008-2018, June, 2002.
- [7] Allen, R.M. and Kanamori, H., The potential for earthquake early warning in southern California. *Science*, Issue 300, pp. 685-848, May, 2003. DOI: 10.1126/science.1080912
- [8] Ochoa, L.H., Niño, L.F. and Vargas, C.A., Severity classification of a seismic event based on the Magnitude-Distance ratio using only one seismological station. *Earth Sciences Research Journal*, 18(2), pp. 115-122, December, 2014. DOI: 10.15446/esrj.v18n2.41083
- [9] Zhang, M., Tian, D. and Wen, L., A new method for earthquake depth determination: stacking multiple-station autocorrelograms. (A. Access, Ed.) *Geophysical Journal International*, (197), pp. 1107-1116, 2014. DOI: 10.1093/gji/ggu044
- [10] Lin, K-W. and Sanford, A., Improving regional earthquake locations using modified G Matrix and Fussy Logic. *Bulletin of the Seismological Society of America*, pp. 82-93, 2001.
- [11] Sambrige, M. and Gallagher, K., Earthquake hypocenter locations using genetic algorithms. *Bulletin of the Seismological Society of America*, [online]. 83(5), pp. 1467-1491, 1993. Available at: <http://rsees.anu.edu.au/~malcolm/papers/pdf/sg93.pdf>
- [12] Magotra, N., Ahmed, N. and Chael, E., Seismic event detection and source location using single station (three components) data. *Bulletin of the Seismological Society of America*, 77(3), pp. 958-971, June, 1987.
- [13] Roberts, R.G., Christofferson, A. and Cassidy, F., Real-time detection, phase identification and source location estimation using single station three component seismic data. *Geophysical Journal*,

- 97(3), pp. 471-480, June, 1989. DOI: 10.1111/j.1365-246X.1989.tb00517.x
- [14] Saita, J. and Nakamura, Y., The early warning systems for mitigation of disasters caused by earthquakes and tsunamis. In: Zschau, J. and Koppers, A., eds. Early warning systems for natural disaster reduction. Springer-Verlag, Berlin, 2003, pp. 453-460. DOI: 10.1007/978-3-642-55903-7_58
- [15] Talandier, J., Reymond, D. and Oka, E.A., Use of variable mantle magnitude for the rapid one-station estimation of teleseismic moments. *Geophysical Research Letters*, 14(8), pp. 840-843, August, 1987. DOI: 10.1029/GL014i008p00840
- [16] Reymond, D., Hyvernaud, O. and Talandier, J., Automatic detection, location and quantification of earthquakes. *Pure and Applied Geophysics*, March, 135(3), pp. 361-382, 1991. DOI: 10.1007/BF00879470
- [17] Ochoa, L. H., Niño, L. F. and Vargas, C. A., 2017. Fast magnitude determination using a single seismological station record implementing machine learning techniques. *Sciences Direct, Geodesy and Geodynamic*, March, ppp. 1-8. DOI: 10.1016/j.geog.2017.03.010
- [18] Ochoa, L. H., Niño, L.F. and Vargas, C.A., Fast estimation of earthquake epicenter distance using a single seismological station with machine learning techniques. *DYNA*, 85(204), pp. 161-168, 2018. DOI: 10.15446/dyna.v85n204.68408
- [19] Ochoa, L.H., Niño, L.F. and Vargas, C.A., Fast determination of earthquake depth using seismic records of a single station, implementing machine learning techniques. *Ingeniería e Investigación*, 38(2), pp. 91-103, 2018. DOI: 10.15446/ing.investig.v38n2.68407
- [20] Bermúdez, M.L. and Rengifo, F., El Rosal: la estación sismológica del CTBTO en Colombia. Bogotá, Primer Simposio Colombiano de Sismología, 2002, 8 P.
- [21] Wu, Y.M. and Zhao, L., Magnitude estimation using the first three seconds P-wave amplitude in earthquake early warning. *Geophys. Res. Lett.*, 33(16), pp. L16312, August, 2006. DOI: 10.1029/2006GL026871
- [22] Wu, Y.M. and Kanamori, H., Rapid assessment of damage potential of earthquakes in Taiwan from beginning of P waves. *Bulleting of the Seismological Society of America*, 93(1), pp. 526-532, February, 2005. DOI: 10.1785/0120040193
- [23] Magotra, N., Ahmed, N. and Chael, E., Single-station seismic event detection and location. *IEEE Transactions on Geoscience and Remote Sensing*, 27(1), pp. 15-23, January, 1989. DOI: 10.1109/36.20270
- [24] Taylor, J.S. and Cristianini, N., *Kernel methods for pattern recognition*. First ed. Cambridge University Press, Cambridge, United Kingdom, 2004.
- [25] Hsiao, N-Ch, Wu, Y.M., Zhao, L., Chen, D., Huang, W-T., Kuo, K-H., Shin, T-Ch. and Leu, P-L., A new prototype system for earthquake early warning system in Taiwan. *Soil Dynamic and Earthquake Engineering*, 31(2), pp. 201-208, 2011. DOI: 10.1016/j.soildyn.2010.01.008
- [26] Chen, D-Y., Hsiao, N-Ch. and Wu, Y-M., The earthworm based earthquake alarm reporting system in Taiwan. *Bulletin of the Seismological Society of America*, 105(2A), pp. 568-570, April, 2015.
- [27] Sheen, D-H., A robust maximum-likelihood earthquake location method for early warning. *Bulletin of the Seismological Society of America*, 105(3), pp. 1301-1313, 2015. DOI: 10.1785/0120140188.
- [28] Zhang, H. et al., An earthquake early warning system in Fujian, China. *Bulletin of the Seismological Society of America*, 106(2), pp. 755-765, 2016. DOI: 10.1785/0120150143

L.H. Ochoa, is an associated professor at Universidad Nacional de Colombia, in the Sciences Faculty, Geosciences Department. Received his BSc. Eng in Civil Engineering in 1988, a MSc. degree in Geophysics in 2003, a MSc. degree in Geomatics in 2007, and a PhD. in System Engineering in 2017, all of them from the Universidad Nacional de Colombia, Bogota. He dedicates himself to research in intelligent systems, geophysics and geodesy. His main interests include earthquake early warning using intelligent systems.
ORCID: 0000-0002-3607-7339

L.F. Niño, is an associated professor at Universidad Nacional de Colombia, in the Engineering Faculty, Systems Engineering Department. Received his BSc. Eng in System Engineering (1990) and a MSc. degree in Mathematics in 1995 from Universidad Nacional de Colombia, Bogota - Colombia. A MSc. degree in Computer Science (1999) and a PhD. degree in Computer Science at The University of Memphis - United States of America. He dedicates himself to teaching and research in computer science, and his main academic interests include computational intelligence and its applications, particularly in the life science.
ORCID: 0000-0003-4703-0007

C.A. Vargas, is a full professor at Universidad Nacional de Colombia, in the Sciences Faculty, Geosciences Department. Received his BSc. in Geology (1993) from Universidad de Caldas - Colombia. A MSc. degree in Seismic Engineering and Structural Dynamic (2000) from Universidad Politécnica de Catalunya - Spain. A MSc. degree in Physics Instrumentation (2006) from Universidad Tecnológica de Pereira - Colombia. His PhD. in Seismic Engineering and Structural Dynamic (2003) from Universidad Politécnica de Catalunya - Spain. His postdoctoral research (2010) from University of Texas, Institute for Geophysics - United States of America. He dedicates himself to research in geodynamic and basin analysis. His main interest includes subduction systems and their response in sedimentary basins.
ORCID: 0000-0002-5027-9519