Bulletin of the Seismological Society of America, Vol. 108, No. 1, pp. 238–247, February 2018, doi: 10.1785/0120170115

Ⓔ

An Earthquake Early Warning Algorithm Based on the *P*-Wave Energy Released in the t_S–t_P Interval

by Armando Cuéllar, Gerardo Suárez, and J. M. Espinosa-Aranda

Abstract The t_S–t_P earthquake early warning algorithm measures the energy of the *P*-wave coda on the vertical component in the t_S–t_P period to make a magnitude threshold estimation. The algorithm is based on two parameters: the logarithm of the peak ground acceleration, $\max(a(t_S–t_P))$ and the logarithmic cumulative acceleration $\mathrm{SA}(t_S–t_P)$. The model is built using a learning algorithm that iteratively parameterizes the linear fit of $\max(a(t_S–t_P))$ and $\mathrm{SA}(t_S–t_P)$ to M_w in segments. Training datasets were based on 324 accelerograms from 101 earthquakes ($4.8 \le M_w \le 8.1$) in the Mexican subduction zone from 1985 to 2013. The algorithm is supervised to avoid outliers in the data. The process results in a family of linear equations parameterizing the observations to magnitude $m(t_S–t_P)$ calibrated to the observed M_w. The algorithm was successfully tested using a dataset of 28 earthquakes in the Mexican subduction zone, from 2014 to 2017. The performance of t_S–t_P algorithm was tested as a warning tool using 89 earthquakes in the Mexican subduction zone from 1985 to 2017, that met the criterion of having at least two stations within 70 km from the epicenter. The results show that 79 were correctly screened. The magnitude of six events was overestimated and four were underestimated. These earthquakes had an unfavorable station distribution. The M_w 6 South Napa, California, earthquake of 24 August 2014 was used also as a test case. The two closest stations identified it as $M_w \ge 5.8$ within 2 s after the arrival of the *P* phase. This resulted in a lead time of 10 s in Berkeley and 12 s in San Francisco, prior to the arrival of the *S* waves. Thus, the t_S–t_P algorithm proves to be a reliable tool for seismic early warning where hypocenters are close to the target cities.

Electronic Supplement: Earthquake catalog and table showing performance evaluation of the t_S–t_P algorithm.

Introduction

Since its inception in 1991, the Mexican Seismic Alert System (SASMEX) was designed to cover the southeastern Guerrero seismic gap. At the time, the gap was considered potentially capable of generating a great $M_w > 8$ earthquake (McCann *et al.*, 1979; Singh *et al.*, 1981). The original goal of the system was to warn the population of Mexico City of an impending earthquake, similar to the destructive 19 September 1985 earthquake (Espinosa-Armanda *et al.*, 1995). Although the 1985 earthquake took place about 350 km from Mexico City, it produced a large loss of life and unprecedented material damage (e.g., Rosenblueth, 1986; Esteva, 1988). The soft and highly saturated soils, remnant of the pre-Columbian lake on which the city was built, were responsible for the large accelerations and the long duration of the strong motion of the ground. Therefore, Mexico City presented a unique opportunity to develop a seismic early warning system. If earthquakes are correctly identified and their magnitude calibrated, the distance of ~350–400 km from the subduction zone provides an enviable warning time of ~60 s before the strong shaking begins. No other city in the world offers this advantage to implement a seismic early warning system.

SASMEX has grown to 97 dedicated strong-motion stations covering the subduction zone, with an average spacing of ~25km (Fig. 1). Additional seismic stations are located inland to monitor the in-slab earthquakes that occur within the subducted Cocos plate (Cuéllar *et al.*, 2014). The original algorithm used by SASMEX to issue early seismic warnings to Mexico City is based on the growth of seismic energy in twice the observed t_S–t_P time. The $2(t_S–t_P)$ algorithm was

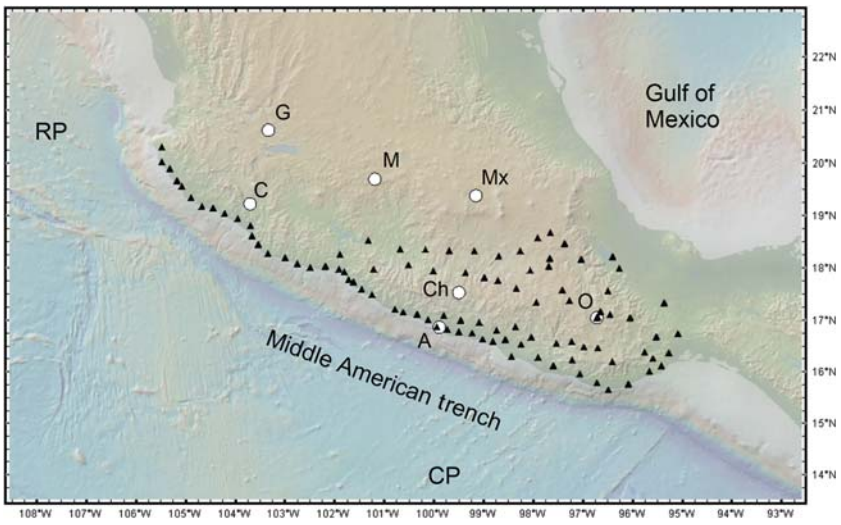

Figure 1. Current distribution of Mexican Seismic Alert System (SASMEX) strong-motion stations (triangles) and cities where the seismic early warning system is now operational (white circles). G, Guadalajara; M, Morelia; Mx, Mexico City; Ch, Chilpancingo; A, Acapulco; O, Oaxaca; RP, Rivera plate; CP, Cocos plate.

thoroughly tested and discussed by Cuéllar *et al.* (2017). Based on the energy released, it estimates an empirical magnitude range related to m_b.

In recent years, other cities were incorporated to SASMEX (Fig. 1). Some of these cities lie close to the subduction zone or to the regions where in-slab earthquakes take place beneath the continent. For this reason, many of these locations do not share Mexico City's advantage of having an almost minute-long warning time; thus, there is a need to develop a faster algorithm to detect earthquakes and to determine a magnitude threshold.

Here, we present an algorithm based on the seismic energy released in the interval t_S–t_P. The algorithm initiates from the detection of the *P* wave and ends with the arrival of the *S* wave. The acceleration of the *P* wave and its coda are measured on the vertical component. We propose a magnitude estimation model $m_{(t_S-t_P)}$ based on two parameters: the logarithm of the maximum quadratic acceleration $\max(a(t_S-t_P))$ and the logarithmic cumulative quadratic acceleration $\mathrm{SA}(t_S-t_P)$ of the *P* wave. The algorithm is calibrated to M_w using a training dataset of 101 earthquakes occurring from 1985 to 2013 in the Mexican subduction zone. Only accelerograms of earthquakes $M_w > 4.7$, located at epicentral distances < 70 km, are used.

The magnitude estimator $m_{(t_S-t_P)}$ is based on a family of logarithmic linear regression models fit in a least-squares sense, iteratively calculated using a machine learning strategy. In contrast to the two-tiered alerting protocol prescribed by the authorities in Mexico City (Cuéllar *et al.*, 2017), here we propose a minimum magnitude cutoff of $m_{(t_S-t_P)} \geq 5.8$ to activate the public earthquake early warning system. The t_S–t_P algorithm neither calculates the location of the earthquake nor estimates potential shaking of the ground. The algorithm waits for the confirmation of two neighboring stations that the magnitude threshold is exceeded to issue an alert.

Description of the t_S–t_P Algorithm

Detection of the *P* and *S* Waves

The detection process of the *P* and *S* phases uses the three orthogonal components of the strong-motion data, running in parallel to the average square input (Allen, 1978) and vertical-to-horizontal (Nakamura, 1996) algorithms. The real-time detection procedure was described in detail for the $2(t_S-t_P)$ algorithm (Cuéllar *et al.*, 2017). The phase detection process takes into account empirical criteria that correct for site effects and for the characteristics of the seismic instrumentation (Cuéllar *et al.*, 2017). In cases where the seismic station detects abnormally high background noise, the algorithm recalibrates its operational parameters. In the case of persistent levels of high noise, the seismic station is temporarily taken out of the detection process until the background seismic noise returns to the predefined level.

Estimation of Parameters $\max(a(t_S-t_P))$ and $\mathrm{SA}(t_S-t_P)$

The algorithm uses the vertical component of three orthogonal component accelerograms. The energy in the t_S–t_P time window is composed essentially of the *P* wave and its coda. The parameter $\max(a(t_S-t_P))$ is the logarithm of the maximum value of the quadratic acceleration of the vertical channel within the t_S–t_P interval (equation 1). The parameter $\mathrm{SA}(t_S-t_P)$ is the sum of the logarithmic cumulative quadratic acceleration of the vertical channel observed within the same interval (equation 2). The algorithm terminates with the detection of the *S* wave and proceeds to estimate the corresponding seismic magnitude threshold.

$$\max(a(t_S-t_P)) = \log_{10}[\max(a(i)^2)]_{i=t_P}^{t_S} \tag{1}$$

$$\mathrm{SA}(t_S-t_P) = \log_{10}\sum_{i=t_P}^{t_S} a(i)^2, \tag{2}$$

in which $a(i)$ are the acceleration values on the vertical channel of the triaxial accelerometer and i is the sample number starting from the *P* wave and ending with the detection of the *S* wave.

Correlation between M_w and $\max(a(t_S-t_P))$ and $\mathrm{SA}(t_S-t_P)$

To explore the correlation between the magnitude M_w of earthquakes in the Mexican subduction zone and parameters

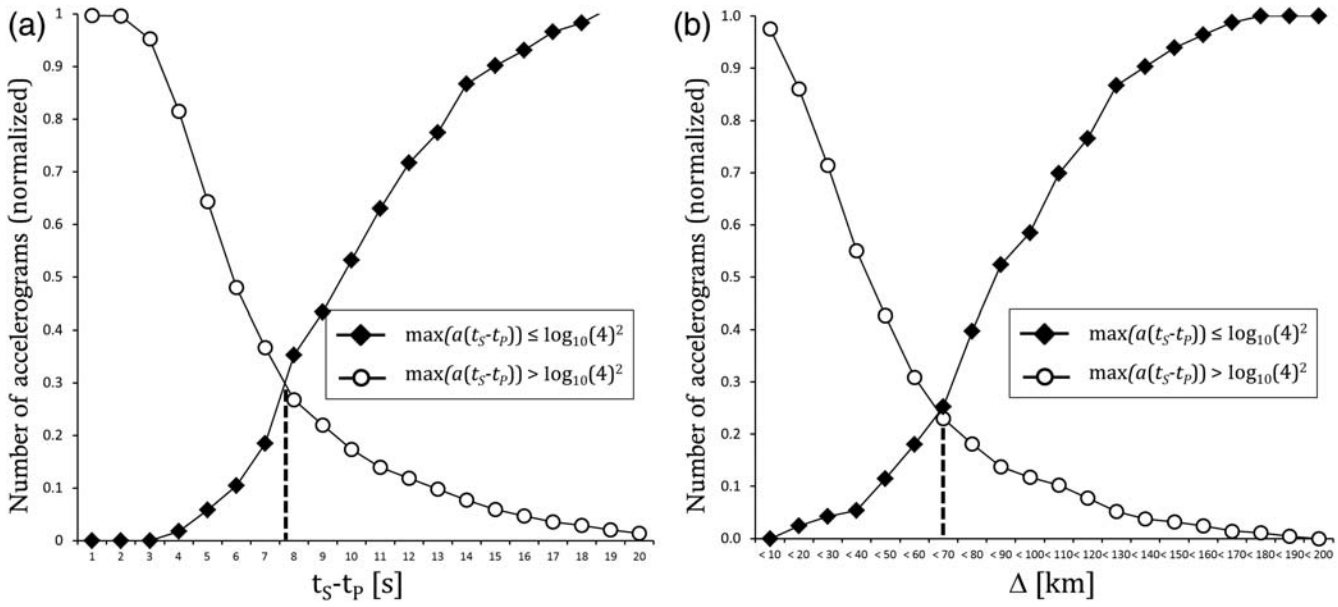

Figure 2. Cumulative distribution of the 1114 acceleration records since 1985–2017 $4.7 > M_w \leq 8.1$ grouped into two categories: $\max(a(t_S–t_P)) \leq \log_{10}(4^2)$ (open circles) and $\max(a(t_S–t_P)) > \log_{10}(4^2)$ (closed diamonds) plotted with respect to (a) the $t_S–t_P$ time and (b) the epicentral distance Δ.

$\max(a(t_S–t_P))$ and $SA(t_S–t_P)$, data were collected from the following catalogs: the Mexican Seismological Service (SSN), the Global Centroid Moment Tensor (CMT) (Dziewonski *et al.*, 1981, 1999; Ekström *et al.*, 2012), the strong-motion data recorded by the Instituto de Ingeniería, Universidad Nacional Autónoma de México (Pérez-Yañez *et al.*, 2014), and the SASMEX network, operated by the Centro de Instrumentación y Registro Sísmico. A selection criteria were applied as follows:

- A1: earthquakes reported in the Mexican subduction zone in the SSN catalog with focal depths <40 km;
- A2: earthquakes of magnitude $M_w > 4.8$, as reported by the CMT catalog; and
- A3: accelerograms with clear recordings of both *P* and *S* waves.

A total of 1114 acceleration records were selected. In some cases, the accelerograms show very low accelerations that are close to the self-noise of the instruments. The resolution of the majority of the early strong-motion SASMEX instruments was 10 or 12 bits. In those cases, the noise level of the seismic acceleration recordings is at least ±1.0 or ±0.5 cm/s^2, respectively. To be well above the noise level, only acceleration records with values >4 cm/s^2 in the $t_S–t_P$ time window were utilized. This value of acceleration corresponds to $\max(a(t_S–t_P)) > \log_{10}(4^2)$. More recently, the instruments have higher dynamic ranges using 16 or 24 bit digitizers; the threshold criteria would be also acceptable in these cases.

The strong-motion records were grouped in two: one where $\max(a(t_S–t_P)) > \log_{10}(4^2)$ and the other with $\max(a(t_S–t_P)) \leq \log_{10}(4^2)$. The cumulative number of these two groups of accelerograms was plotted relative to the epicentral distance and to the $t_S–t_P$ time (Fig. 2). In the case of records with large $t_S–t_P$ times, the cumulative distribution for $\max(a(t_S–t_P)) > \log_{10}(4^2)$ decreases. In contrast, $\max(a(t_S–t_P) \leq \log_{10}(4^2)$ increases as a function of time (Fig. 2). We propose using the intersection of these curves to empirically define a criterion of when the acceleration records are above the noise level in the $t_S–t_P$ time. The intersection of these curves corresponds to epicentral distances of ~70 km and $t_S–t_P \sim 7$ s.

Based on these results, the accelerograms selected had to meet the following criteria:

- B1: acceleration values in the vertical channel $a > 4$ cm/s^2 during the $t_S–t_P$ time;
- B2: epicentral distances of <70 km to the strong-motion site; and
- B3: acceleration records with $t_S–t_P < 7$ s.

From the initial dataset of 1114 acceleration records, only 400 accelerograms recorded from 129 earthquakes ranging in magnitude from $4.8 \leq M_w \leq 8.1$ met these criteria (Fig. 3 and Ⓔ Table S1, available in the electronic supplement to this article). These 400 accelerograms are divided into two sets: the first 324 are used as training data to develop the algorithm, and the other 76 strong-motion records are used as test data.

$\text{Max}(a(t_S–t_P))$ and $SA(t_S–t_P)$ were computed for the selected dataset of 400 acceleration records and plotted against the reported magnitude M_w. The correlation between $\max(a(t_S–t_P))$ and M_w is $r = 0.40$. In the case of the

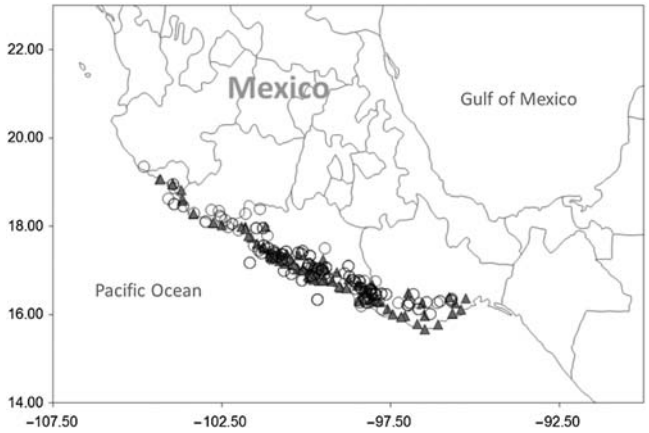

Figure 3. Map showing the 129 selected earthquakes in the Mexican subduction zone from 1985 to 2017 ($M_w > 4.9$; depth < 40 km) used as training data (open circles). The seismic field sensors (triangles) are from the networks of the Instituto de Ingeniería, Universidad Nacional Autónoma de México, and SASMEX.

cumulative acceleration SA(t_S–t_P), the correlation coefficient with magnitude is $r = 0.50$ (Fig. 4). The correlation between max($a(t_S$–$t_P)$) and SA(t_S–t_P) is $r = 0.88$, indicating that these two parameters are not completely independent.

Although the plot of max($a(t_S$–$t_P)$) versus M_w indicates that there is a linear correlation between these two parameters, there are several outliers, particularly in the case of large earthquakes $M_w > \sim 7$ (Fig. 4a). The maximum acceleration of the ground observed in the near field in the first few seconds does not reflect necessarily the resulting magnitude of the earthquake. A case in point is the 19 September 1985 earthquake (M_w 8.1) (Fig. 4a and Ⓔ Table S2). The low acceleration observed in station CALE was attributed to the fact that the nucleation of the 1985 earthquake was located immediately beneath that station (Mendez and Anderson, 1991).

The correlation between SA(t_S–t_P) versus M_w shows an asymptotic growth. As in the case of max($a(t_S$–$t_P)$), the cumulative acceleration shows the presence of outliers, corresponding to earthquakes $M_w > \sim 7$. In some cases, outliers are due to earthquakes where the directivity of the seismic energy is radiated away from the recording seismic stations. For example, in the case of the 1995 Copala earthquake (M_w7.3), the seismic energy radiated away from the closest stations (Anderson *et al.*, 1995). As a result, the values of SA(t_S–t_P) and max($a(t_S$–$t_P)$) at stations GR11, COPL, GR10, and GR09 were low at short epicentral distances for this large magnitude earthquake (Fig. 4 and Ⓔ Table S2).

Calibration of Magnitude Based on max($a(t_S$–$t_P)$) and SA(t_S–t_P)

The correlation of parameters max($a(t_S$–$t_P)$) and SA(t_S–t_P) with M_w suggests the possibility to build an algorithm relating the magnitude and the acceleration observed in the *P*-wave coda (Fig. 4). To this end, we use a training set of 324 accelerograms corresponding to 101 earthquakes occurring from 1985 to 2013 that meet the selection criteria described above (Ⓔ Table S1).

The relationships between magnitude and the parameters reflecting the energy released are estimated by a learning machine algorithm exemplified in Figure 5. The input data to the learning machine is vector $[\bar{x}_j, \bar{y}_j]$, in which $\bar{x}_j = (\mathrm{SA}(t_S\text{–}t_P)_j, \max(a(t_S\text{–}t_P))_j)$, formed from the 324 acceleration records, and $\bar{y}_j$ is the target data: $\bar{y}_j = M_{w_j}$. Before it is introduced into the learning machine process, the data are ordered so that $\mathrm{SA}(t_S\text{–}t_P)_j < \mathrm{SA}(t_S\text{–}t_P)_{j+1}$. The process solves the overdetermined set of equations $\log_{10} M_{w_j} = \alpha \mathrm{SA}(t_S\text{–}t_P)_j + \beta \max(a(t_S\text{–}t_P))_j$.

Considering that the cumulative energy released SA(t_S–t_P) behaves in an asymptotic manner, the process is

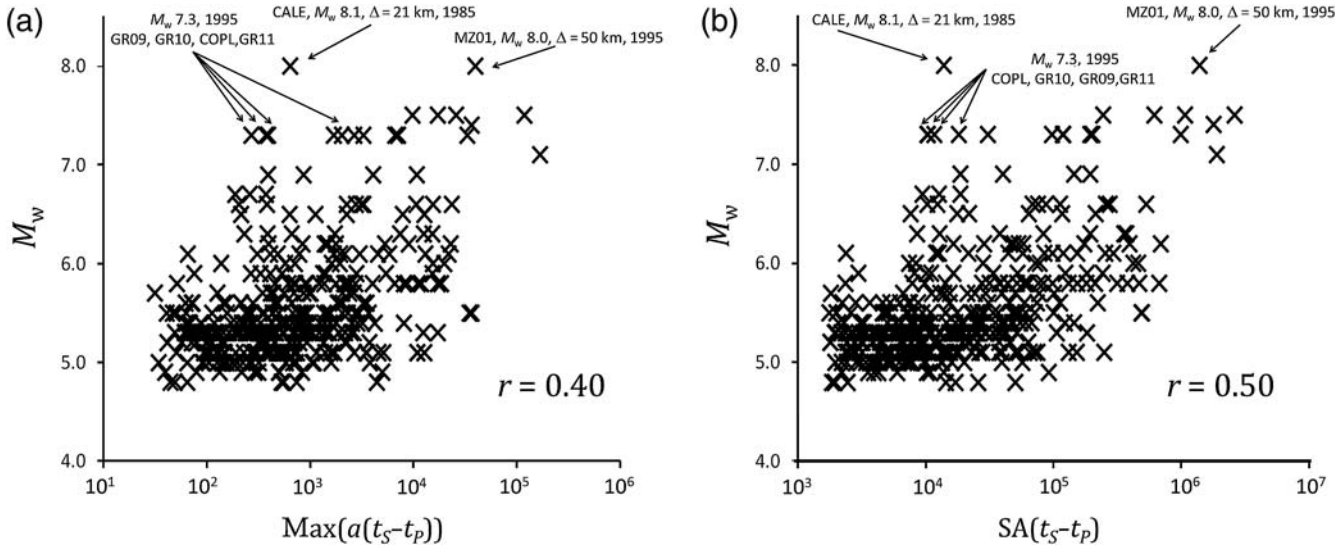

Figure 4. Correlation between parameters (a) max($a(t_S$–$t_P)$) and (b) SA(t_S–t_P) with M_w. The dataset shows the strong-motion records of 324 accelerograms recorded from 1985 to 2013.

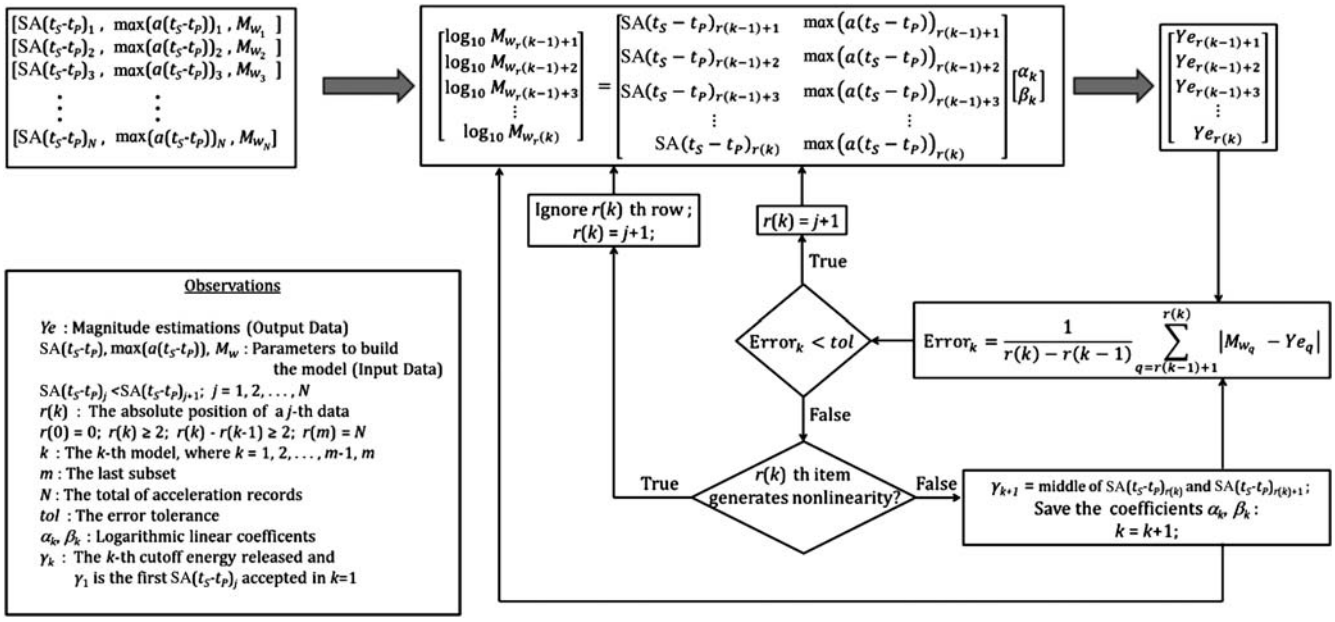

Figure 5. Flow chart showing the process followed by the learning procedure to determine the linear fit in k segments.

linearized in m sections of the curve, in which for each section k, the system of equations is solved with the $(r(k) - r(k-1))$ elements, in which $r(k)$ represents the position of the ordered input data and its value corresponds to $[\bar{x}_j, \bar{y}_j]$ (Fig. 5). The solution to the overdetermined set of equations determines the value of coefficients $[\alpha_k, \beta_k]$ and the error in the linear fit of the kth segment. Error_k is the mean absolute error $|\overline{M_{\text{w}k}} - \overline{Ye}_k|$; in which $\overline{M_{\text{w}k}}$ is the reported magnitude and $\overline{Ye}_k$ is the resulting magnitude of the inversion process in section k.

The process starts by building a solution to the first segment $k = 1$. When the error for that segment is $\text{Error}_k <$ tolerance, the process recycles, adds the next data point $[\bar{x}_{j+1}, \bar{y}_{j+1}]$, and again solves the overdetermined set of equations (Fig. 5). This iterative process continues until $\text{Error}_k >$ tolerance. After several tests, we settled for an error tolerance of 5% in the least-squares fit for each segment. At that point, the algorithm moves on to solve the next section $k + 1$ and sets the upper bound γ of section k, as $\gamma_k = \log_{10}(0.5(10^{\text{SA}(t_S - t_P)_j} + 10^{\text{SA}(t_S - t_P)_{j+1}}))$. Thus, the last data point in section k is the jth term $[\bar{x}_j, \bar{y}_j]$. Subsequently, the first data point in section $k + 1$ will be the $j + 1$th item $[\bar{x}_{j+1}, \bar{y}_{j+1}]$.

The correlations between $\text{SA}(t_S - t_P)$ and $\max(a(t_S - t_P))$ with M_{w} show that the accelerograms of large magnitude earthquakes in the near field often exhibit relatively low values of peak ground acceleration and cumulative energy (Fig. 4). These outliers introduce a nonlinear behavior and increase the value of Error_k. Thus, the inversion process is supervised at each iteration. If a data point $[\bar{x}_j, \bar{y}_j]$ is identified as a clear outlier in model k, the data point is manually eliminated by an active supervision of the learning algorithm (Fig. 5).

The resulting $t_S - t_P$ algorithm, based on the 324 training acceleration records, is defined by nine linear equations that empirically relate $\max(a(t_S - t_P))$ and $\text{SA}(t_S - t_P)$ with $m_{(t_S - t_P)}$ through coefficients α_k and β_k in the m segments of the process. As a result, these nine models are grouped in a family of linear equation parameterized by the γ-values:

$$m_{(t_S - t_P)} = \begin{cases} (10^{\text{SA}(t_S - t_P)})^{\alpha_1} \times (10^{\max(a(t_S - t_P))})^{\beta_1} & \text{if } \text{SA}(t_S - t_P) > \gamma_1 \\ (10^{\text{SA}(t_S - t_P)})^{\alpha_2} \times (10^{\max(a(t_S - t_P))})^{\beta_2} & \text{if } \text{SA}(t_S - t_P) > \gamma_2 \\ (10^{\text{SA}(t_S - t_P)})^{\alpha_3} \times (10^{\max(a(t_S - t_P))})^{\beta_3} & \text{if } \text{SA}(t_S - t_P) > \gamma_3 \\ (10^{\text{SA}(t_S - t_P)})^{\alpha_4} \times (10^{\max(a(t_S - t_P))})^{\beta_4} & \text{if } \text{SA}(t_S - t_P) > \gamma_4 \\ (10^{\text{SA}(t_S - t_P)})^{\alpha_5} \times (10^{\max(a(t_S - t_P))})^{\beta_5} & \text{if } \text{SA}(t_S - t_P) > \gamma_5, \\ (10^{\text{SA}(t_S - t_P)})^{\alpha_6} \times (10^{\max(a(t_S - t_P))})^{\beta_6} & \text{if } \text{SA}(t_S - t_P) > \gamma_6 \\ (10^{\text{SA}(t_S - t_P)})^{\alpha_7} \times (10^{\max(a(t_S - t_P))})^{\beta_7} & \text{if } \text{SA}(t_S - t_P) > \gamma_7 \\ (10^{\text{SA}(t_S - t_P)})^{\alpha_8} \times (10^{\max(a(t_S - t_P))})^{\beta_8} & \text{if } \text{SA}(t_S - t_P) > \gamma_8 \\ (10^{\text{SA}(t_S - t_P)})^{\alpha_9} \times (10^{\max(a(t_S - t_P))})^{\beta_9} & \text{if } \text{SA}(t_S - t_P) > \gamma_9 \end{cases} \tag{3}$$

in which

$$\begin{aligned} &\alpha_1 = 23.570 \times 10^2; \quad \beta_1 = -29.573 \times 10^3; \quad \gamma_1 = 3.230; \\ &\alpha_2 = 23.556 \times 10^2; \quad \beta_2 = -56.052 \times 10^3; \quad \gamma_2 = 3.380; \\ &\alpha_3 = 21.214 \times 10^2; \quad \beta_3 = -33.404 \times 10^3; \quad \gamma_3 = 3.690; \\ &\alpha_4 = 21.214 \times 10^2; \quad \beta_4 = -44.066 \times 10^3; \quad \gamma_4 = 4.016; \\ &\alpha_5 = 18.176 \times 10^2; \quad \beta_5 = -8.438 \times 10^3; \quad \gamma_5 = 4.149; \\ &\alpha_6 = 21.492 \times 10^2; \quad \beta_6 = -73.543 \times 10^3; \quad \gamma_6 = 4.278; \\ &\alpha_7 = 22.792 \times 10^2; \quad \beta_7 = -102.763 \times 10^3; \quad \gamma_7 = 4.637; \\ &\alpha_8 = 10.767 \times 10^2; \quad \beta_8 = -47.098 \times 10^3; \quad \gamma_8 = 5.394; \\ &\alpha_9 = 18.747 \times 10^2; \quad \beta_9 = -62.704 \times 10^3; \quad \gamma_9 = 5.609. \end{aligned}$$

Based on expressions (4) and (5), described below, the mean absolute error and its standard deviation are estimated:

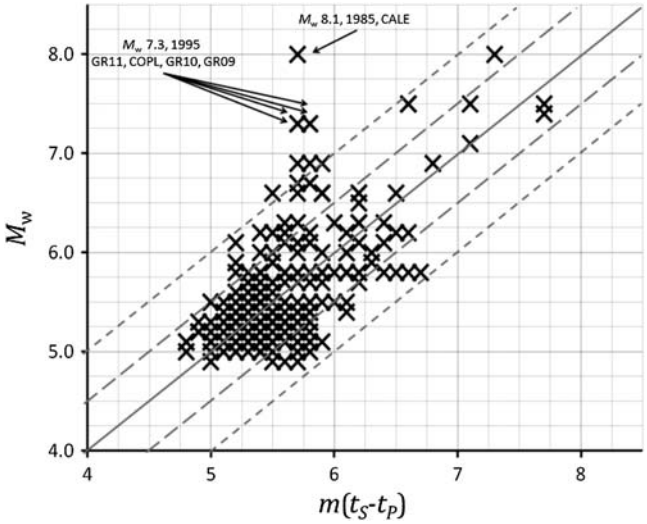

Figure 6. The estimated magnitude $m(t_S–t_P)$ plotted versus the observed M_w of 324 strong-motion records of 101 earthquakes $M_w < 4.8$ recorded from 1985 to 2013 in the Mexican subduction zone.

$$m_{\text{error}} = \frac{1}{N}\sum_{j=1}^{N} |M_{w_j} - m(t_S–t_P)_j| \tag{4}$$

$$\sigma_{\text{error}} = \left(\frac{1}{N}\sum_{j=1}^{N}(m(t_S–t_P)_j - m_{\text{error}})^2\right)^{1/2}. \tag{5}$$

The mean absolute error of the estimated magnitude of the 324 accelerograms corresponding to the training set was $m_{\text{error}} = 0.32$, with a standard deviation of $\sigma_{\text{error}} = 0.30$. This mean absolute error in magnitude is within the tolerance accepted in the algorithm training process. Assuming an error tolerance in magnitude of ±0.5 and ±1.0 as considered in Japan (Hoshiba *et al.*, 2008), the algorithm accurately predicts the magnitude with a success rate of 85% and 96%, respectively. The results of the proposed algorithm, comparing the estimated magnitude $m_{(t_S–t_P)}$ versus the catalog M_w is shown in Figure 6.

Magnitude Estimation $m(t_S–t_P)$ on Independent Test Data from 2014 to 2017

To test the resulting algorithm, a test dataset composed of 76 acceleration records of 28 earthquakes occurring in the Mexican subduction zone from 2014 to 2017 was selected. The magnitude of these 28 earthquakes ranges from $4.8 < M_w < 7.3$. Because of the short time span of the test data available, there are only three earthquakes with $M_w > 6$.

Based on expressions (4) and (5), the mean absolute error of the estimated magnitude was $m_{\text{error}} = 0.5$, with a standard deviation $\sigma_{\text{error}} = 0.38$. Thus, the resulting algorithm predicted the magnitude $m_{(t_S–t_P)}$ of the test earthquakes within the prescribed error tolerance of ±0.5. Considering an error tolerance of ±0.5 and ±1.0 in the magnitude estimation, the success rate was 70% and 93%, respectively. The magnitude estimated by the $t_S–t_P$ algorithm of the test data is compared with the catalog magnitude in Figure 7 and Ⓔ Table S2.

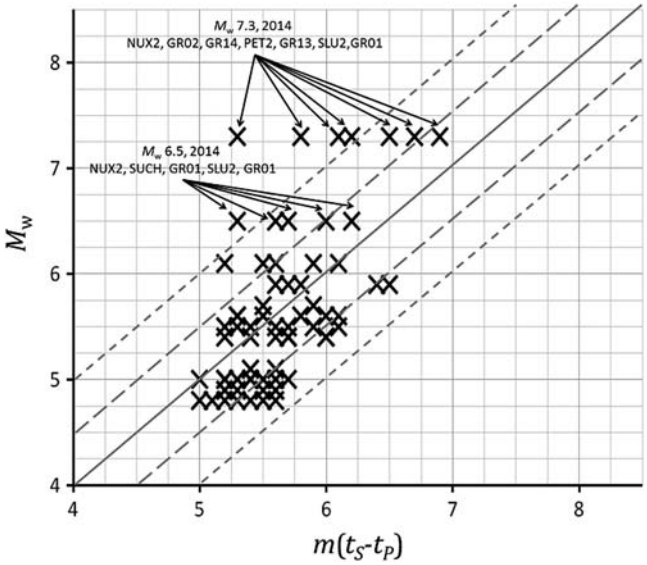

Figure 7. The estimated magnitude $m(t_S–t_P)$ plotted versus the observed M_w for the test data composed of 76 strong-motion records of 28 earthquakes $M_w > 4.8$ recorded from 2014 to 2017 in the Mexican subduction zone.

In the case of the two largest earthquakes in the test dataset, 18 April 2014 (M_w 7.3) and 8 May 2014 (M_w 6.5), the two strong-motion stations closest to the epicenter estimated the magnitude within an error of ±0.5. The largest errors in magnitude estimation are in stations that are far from the epicenter. However, these other stations, although they are not at an ideal epicentral distance for the algorithm, lie within the established distance range. Thus, they were included for completeness and transparency in the tests.

Performance Evaluation of the $t_S–t_P$ Algorithm on Mexican Subduction Earthquakes $M_w > 4.7$ from 1985 to 2017

Iglesias *et al.* (2007) and Suárez *et al.* (2009) pointed out the difficulties of the $2(t_S–t_P)$ algorithm of correctly identifying the magnitude in the narrow bins prescribed by the Mexico City authorities to issue preventive alerts ($5.5 < M_w < 6$). To avoid this problem of estimating magnitude within very narrow bins, we propose a single earthquake early warning criterion: $M_w \geq 5.8$.

The activation criterion traditionally used by SASMEX to minimize spurious signals (Espinosa-Aranda *et al.*, 1995) requires that two nearby sensors confirm that the magnitude estimate on both is above the prescribed threshold to issue an alert. To test the alerting robustness of the $t_S–t_P$ algorithm using a dataset with a broader range of magnitudes, we use the 89 earthquakes from 1985 to 2017 that meet the criterion of having accelerograms within epicentral distances of

< 70 km. (Ⓔ Tables S1 and S2). Therefore, this test evaluates the robustness of the proposed algorithm of discriminating earthquakes above or below the M_w 5.8 threshold, not only the magnitude estimation based on individual accelerograms.

Out of these 89 earthquakes, in 79 cases the seismic alert based on the $t_S–t_P$ algorithm would have been correctly broadcast. In four cases, even though the magnitude of the earthquakes reported in the seismic catalogs was $M_w > 5.8$, the magnitude $m_{(t_S–t_P)}$ was underestimated and no alert would have been issued by the $t_S–t_P$ algorithm. In the case of six events, the algorithm overestimated the magnitude. This means that 89% of the earthquakes would have been correctly classified based on the pre-established warning criteria.

Earthquakes that are larger than the prescribed magnitude threshold, but for which no alert is issued, represent the possibility that an alert may not be issued for a damaging earthquake. Thus, it is worth reviewing the four earthquakes $M_w > 5.8$, for which the magnitude $m_{(t_S–t_P)}$ was underestimated. In the case of the 31 May 1990 earthquake (M_w 5.8), one out of four accelerograms (LLAV) estimated the magnitude as $m_{(t_S–t_P)}$ 6.4. The other three records yield an estimate of $m_{(t_S–t_P)} < 5.8$. The second earthquake is the M_w 6.0 earthquake on 1 January 2004. In this case, the only two stations available estimated $m_{(t_S–t_P)} < 5.8$ (Ⓔ Table S2). Both strong-motion stations lie at the limit of the design criteria of the algorithm, with $t_S–t_P$ times of 5.75 and 6.80 s, respectively. The third case is the 14 June 2004 M_w 5.9 earthquake. It was recorded by five acceleration stations. Station PNTP estimated $m_{(t_S–t_P)}$ 6.3. The next closest station OX04 with $t_S–t_P = 3.01$ s, estimated $m_{(t_S–t_P)}$ 5.5 (Ⓔ Table S2). Finally, the 2 April 2012 M_w 6.0 earthquake, in this case two stations selected estimated $m_{(t_S–t_P)}$ 5.7.

It is worth pointing out that these earthquakes occurred prior to the installation of the new SASMEX stations and, therefore, had a very unfavorable distribution of strong-motion records. Furthermore, they correspond to cases where the reported magnitude is very close to the prescribed threshold of M_w 5.8.

It is also interesting to evaluate the performance of the $t_S–t_P$ algorithm on the eight earthquakes $M_w > 7$ in the dataset (Ⓔ Table S2). Three of these eight earthquakes, the 14 September 1995 Copala (M_w 7.3), the 20 March 2012 (M_w 7.5), and the 18 April 2014 (M_w 7.3), had at least two strong-motion records. Although the true magnitude was underestimated, the $t_S–t_P$ algorithm identified them as large magnitude events and would have issued a correct early warning in all cases (Ⓔ Table S2).

In the case of the other five earthquakes $M_w > 7$, only one accelerogram was available. Thus, we could not evaluate them using the criteria of two confirming stations. Nevertheless, except for the 1985 M_w 8.1 Michoacán earthquake with magnitude estimation $m_{(t_S–t_P)}$ 5.7, the accelerograms of the other four earthquakes were identified as having a magnitude $M_w > 5.8$ and would have had an alert. As mentioned before, in the case of the 1985 Michoacán earthquake, the low acceleration observed at station CALE, located immediately above the hypocenter of the 1985 earthquake, is the result of the nucleation process. From these results, it may be concluded that, even under unfavorable distribution of strong-motion stations, the $t_S–t_P$ algorithm is a reliable tool to issue alerts in a short time span based only on the *P*-wave coda. It is clear that with better seismic coverage, the $t_S–t_P$ algorithm would have a better performance than the one reported here.

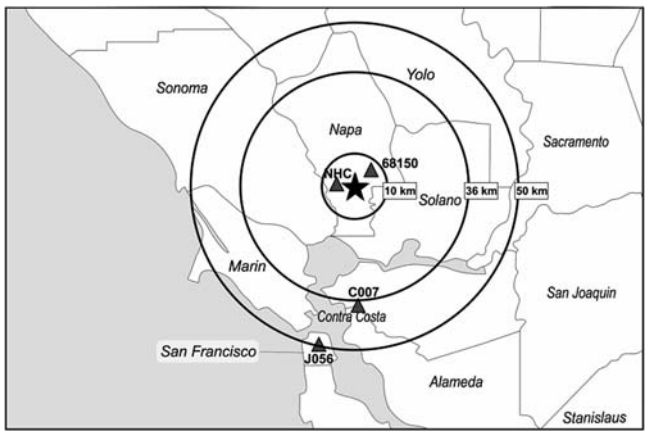

Figure 8. Location of the Napa earthquake of 24 August 2014 (M_w 6) and of the four closest Center for Engineering Strong Motion Data (CESMD) seismic stations. NHC and station 68150 are located at distances of less than 10 km from the epicenter. The stations C007 and JO56, in Berkeley and San Francisco, respectively, are used to determine the arrival of the *S* waves at these locations.

Performance of M_w 6 South Napa Earthquake

The South Napa earthquake is an ideal example of a shallow crustal earthquake that occurs close to important urban areas where the $t_S–t_P$ algorithm can be tested. Several instruments from the Center for Engineering Strong Motion Data (CESMD) recorded the South Napa earthquake of 24 August 2014 (M_w 6) at relatively close distances. The NHC-Huichica Creek and the 68150-Napa College strong-motion stations are the closest to the epicenter and were the first to record the earthquake. They are located 3.9 and 6.4 km away from the epicenter, respectively (Fig. 8).

The application of the $t_S–t_P$ algorithm on the NHC-Huichica Creek seismic record results in a magnitude $m_{(t_S–t_P)}$ 6.2 with a processing time of 2.02 s. The second seismic record in 68150-Napa College processed the algorithm parameters in 2.08 s and estimated a magnitude $m_{(t_S–t_P)}$ 6.5 (Ⓔ Table S3). Based on these two stations, which are the closest to the epicenter, an earthquake early warning would have been broadcast ~2 s after the arrival of the *P* wave (Fig. 9). Considering that the strong shaking coincides generally with the arrival of the *S* waves, the warning time in Berkeley would have been ~10 s, prior to the arrival of the *S* wave, as measured on the C007-Forest Lane strong-motion

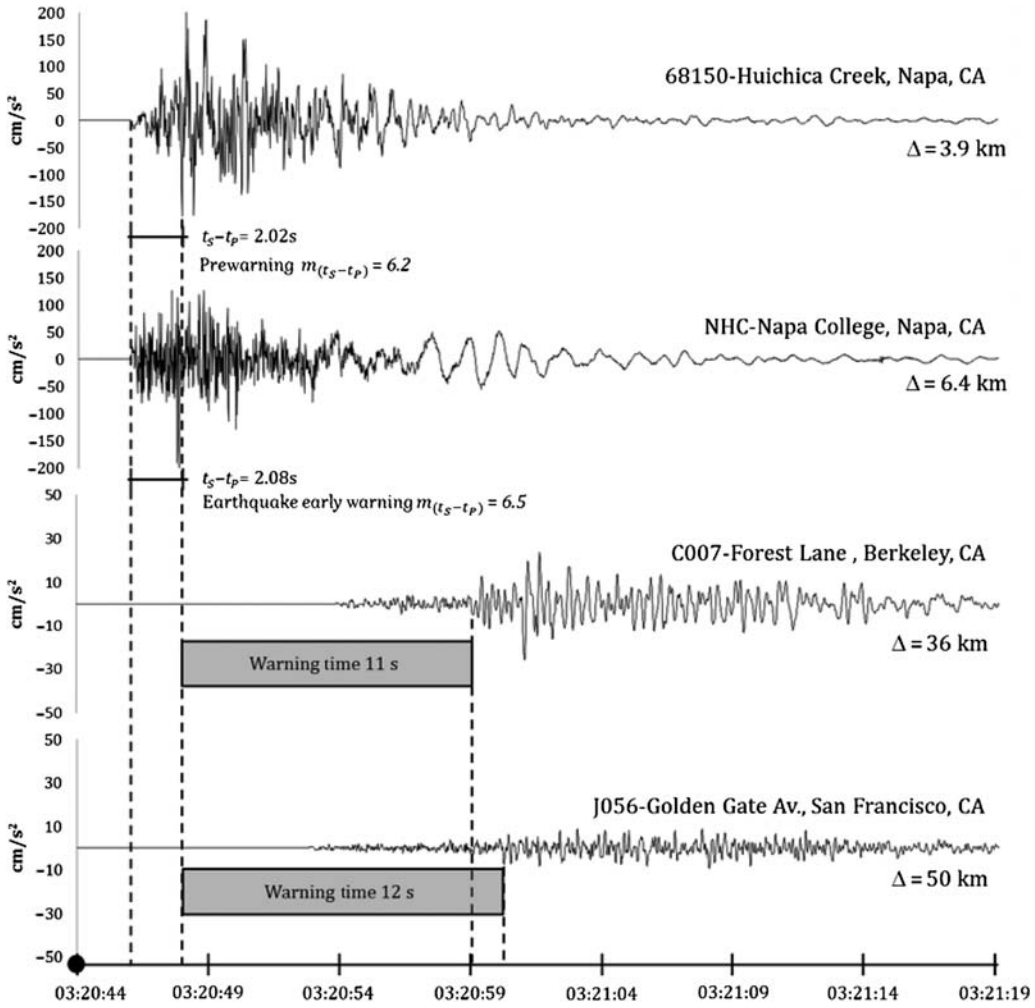

Figure 9. Graphical representation of the performance of t_S–t_P algorithm applied to the 24 August 2014 M_w 6 South Napa, California, earthquake.

record (Figs. 7 and 8). In the case of San Francisco, the warning time would have been 12 s prior to the arrival of strong shaking, based on the *S*-wave arrival at the J056-Golden Gate Avenue station (Fig. 8).

It is interesting to compare the warning times reported by ShakeAlert to those obtained by the t_S–t_P algorithm. The warning time reported for the Napa earthquake based on ShakeAlert was 5 s prior to the arrival of the *S* waves in Berkeley (Allen *et al.*, 2014; Grapenthin *et al.*, 2014a,b). As a result of this process, the t_S–t_P algorithm would almost double the early warning time. A similar result is obtained in San Francisco, where a warning time of 12 s would allow the implementation of basic civil protection procedures.

Summary and Conclusions

The recent expansion of SASMEX created the demand for possible early earthquake warnings for cities that lie close to the seismic zones. In these cases, the $2(t_S–t_P)$ algorithm, traditionally used by the Mexican seismic early warning system, would shorten the warning time. The purpose of deriving a fast algorithm is to rapidly issue early warnings to cities that lie at distances of a few tens of kilometers from the epicenter. Although the warning time may be of only a few seconds, it may help in establishing rapid and automated civil protection measures.

The proposed t_S–t_P algorithm reduces the time needed to estimate the magnitude in half, making use of two parameters measured from the vertical component: the logarithm of the maximum acceleration $\max(a(t_S–t_P))$ and the logarithmic cumulative quadratic acceleration $\mathrm{SA}(t_S–t_P)$. Based on a linear regression model that maps $\mathrm{SA}(t_S–t_P)$ and $\max(a(t_S–t_P))$ to M_w, a magnitude $m_{(t_S-t_P)}$ is estimated. The construction of the algorithm makes use of a supervised machine learning process that linearizes piecewise the relation between maximum and cumulative acceleration and the magnitude.

The algorithm was derived using a training dataset composed of 324 accelerograms that correspond to 101 earthquakes occurring between 1985 and 2013 (Ⓔ Table S1).

The mean absolute error is $m_{error} = 0.32$, within the accepted error tolerance of ± 0.5. In other words, considering an error tolerance of ± 0.5 and ± 1.0 in the magnitude estimation, the success rate was 85% and 96%, respectively.

The test dataset was based on 28 earthquakes occurring in the subduction zone from 2014 to 2017. The results of this test on 76 accelerograms results in the absolute mean error of 0.5. The test dataset had never been seen by the algorithm. Assuming error tolerances of ± 0.5 and ± 1.0 in the magnitude estimation, the success rate was 70% and 93%, respectively.

A test on the performance of t_S–t_P algorithm as a warning tool was conducted using 89 earthquakes in the Mexican subduction zone that met the criteria of having two strong-motion records of nearby stations to issue a seismic warning to avoid false alerts. The results show that 79 earthquakes, out of the 89 tested, were correctly classified as being above or below the magnitude threshold M_w 5.8. The magnitude of six events was overestimated and four were underestimated. This represents a success rate of 89%.

Admittedly, if the detection of the S phase is late, the value of $\max(a(t_S–t_P))$ would increase and, as a result, the magnitude would be overestimated. However, $\max(a(t_S–t_P))$ is measured on the vertical channel, where normally the amplitude of the S waves is smaller relative to the horizontal components, which are not incorporated in the algorithm. Nevertheless, using two stations to issue the alert would help avoid this situation, as shown in Ⓔ Table S2.

Clearly, now that the seismic coverage is substantially better along the subduction zone (Fig. 1), the magnitude estimations will tend to improve. It is important to underline that in the implementation of this algorithm in the earthquake early warning system, new relevant data will be systematically incorporated as part of the process of an effective learning machine.

The t_S–t_P algorithm was also tested on the 24 August 2014 Napa earthquake. The results for a hypothetical earthquake early warning for the Napa earthquake show that the t_S–t_P algorithm would have identified the event as having $M_w > 5.8$ at both stations within 2 s after the arrival of the P phase. This would have given a warning time of ~10 s in the city of Berkeley prior to the arrival of the S waves. In the case of San Francisco, the warning time would have been 12 s.

Early seismic warning algorithms face the challenge of reducing the processing time needed to identify earthquakes and discriminating a magnitude threshold with a high degree of confidence to make a decision on issuing a seismic alert. In this respect, the results of the t_S–t_P algorithm proposed here are encouraging for regions with large population centers close to seismic zones. Although the warning time would be only a few seconds, well-designed protocols could activate a civil protection response designed to mitigate damage and loss of life.

Data and Resources

Accelerograms from the Mexican Seismic Alert System (SASMEX) stations were provided by the Centro de Instrumentación y Registro Sísmico (CIRES) with the authorization of the Institutopara la Seguridad en las Construcciones del Distrito Federal in Mexico City and the Coordinación Estatal de Protección Civil de Oaxaca, Oaxaca state. Strong-motion data for the Mexican subduction earthquakes are from the webpage of the Instituto de Ingeniería of the Universidad Nacional Autónoma de México (UNAM) https://aplicaciones.iingen.unam.mx/AcelerogramasRSM/ (last accessed September 2017). Data from the Global Centroid Moment Tensor Project are from http://www.globalcmt.org/CMTsearch.html (last accessed September 2017).

Acknowledgments

The authors express their gratitude to the institutions that provided the strong-motion data. The governments of Mexico City and of the state of Oaxaca allowed us use the waveform database of Mexican Seismic Alert System (SASMEX). Thanks are also due to the Instituto de Ingeniería of the National University in Mexico (UNAM) for accelerograms in Mexico, and to the Center for Engineering Strong Motion Data for the Napa earthquake records. G. S. acknowledges support of the Consejo Nacional de Ciencia y Tecnología (CONACYT) under Grant 082821. Special thanks are due to Guadalupe Rico, Sandra Ramos, Marisela Palomino, and Cecilia Hernández for their invaluable support in collecting and organizing the data presented in this article. The article was substantially improved by the valuable and constructive comments and corrections made by three anonymous reviewers and by Associate Editor D. Melgar.

References

Allen, R. V. (1978). Automatic earthquake recognition and timing from single traces, *Bull. Seismol. Soc. Am.* **68,** no. 5, 1521–1532.

Allen, R. V., D. Given, T. H. Heaton, and J. Vidale (2014). Successful ShakeAlert performance for the Napa quake, *2014 AGU Fall Meeting*, San Francisco, California, S44D-01.

Anderson, J., R. Quaas, S. K. Singh, J. M. Espinosa, A. Jimenez, J. Lermoand, and S. Alcocer (1995). The Copala, Guerrero, Mexico earthquake of September 14, 1995 ($M_w = 7.4$): A preliminary report, *Seismol. Res. Lett.* **66,** no. 6, 11–39.

Cuéllar, A., J. M. Espinosa-Aranda, G. Suárez, G. Ibarrola, A. Uribe, F. H. Rodríguez, and B. Frontana (2014). The Mexican Seismic Alert System (SASMEX): Its alert signals, broadcast results and performance during the M 7.4 Punta Maldonado earthquake of March 20th, 2012, in *Early Warning for Geological Disasters*, Springer, Berlin, Germany, 71–87.

Cuéllar, A., G. Suárez, and J. M. Espinosa-Aranda (2017). Performance evaluation of the earthquake detection and classification algorithm $2(t_S–t_P)$ of the Seismic Alert System of Mexico (SASMEX), *Bull. Seismol. Soc. Am.* **107,** no. 3, 1451, doi: 10.1785/0120150330.

Dziewonski, A. M., T. A. Chou, and J. H. Woodhouse (1981). Determination of earthquake source parameters from waveform data for studies of global and regional seismicity, *J. Geophys. Res.* **86,** 2825–2852, doi: 10.1029/JB086iB04p02825.

Dziewonski, A. M., G. Ekström, and N. N. Maternovskaya (1999). Centroid-moment tensor solutions for April–June, 1998, *Phys. Earth Planet. In.* **112,** no. 1, 11–19.

Ekström, G., M. Nettles, and A. M. Dziewonski (2012). The global CMT project 2004–2010: Centroid-moment tensors for 13,017

earthquakes, *Phys. Earth Planet. In.* **200/201,** 1–9, doi: 10.1016/j.pepi.2012.04.002.

Espinosa-Aranda, J. M., A. Jimenez, G. Ibarrola, F. Alcantara, A. Aguilar, M. Inostroza, and S. Maldonado (1995). Mexico City seismic alert system, *Seismol. Res. Lett.* **66,** no. 6, 42–53.

Esteva, L. (1988). The Mexico earthquake of September 19, 1985—Consequences, lessons, and impact on research and practice, *Earthq. Spectra* **4,** no. 3, 413–426.

Grapenthin, R., I. Johanson, and R. M. Allen (2014a). Operational real-time GPS-enhanced earthquake early warning, *J. Geophys. Res.* **119,** no. 10, 7944–7965.

Grapenthin, R., I. Johanson, and R. M. Allen (2014b). The 2014 M_w 6.0 Napa earthquake, California: Observations from real-time GPS-enhanced earthquake early warning, *Geophys. Res. Lett.* **41,** no. 23, 8269–8276.

Hoshiba, M., O. Kamigaichi, M. Saito, S. Y. Tsukada, and N. Hamada (2008). Earthquake early warning starts nationwide in Japan, *Eos Trans. AGU* **89,** no. 8, 73–74.

Iglesias, A., S. K. Singh, M. Ordaz, M. A. Santoyo, and J. Pacheco (2007). The seismic alert system for Mexico City: An evaluation of its performance and a strategy for its improvement, *Bull. Seismol. Soc. Am.* **97,** no. 5, 1718–1729.

McCann, W. R., S. P. Nishenko, L. R. Sykes, and J. Krause (1979). Seismic gaps and plate tectonics: Seismic potential for major boundaries, in *Earthquake Prediction and Seismicity Patterns*, Birkhäuser, Basel, Switzerland, 1082–1147.

Mendez, A. J., and J. G. Anderson (1991). The temporal and spatial evolution of the 19 September 1985 Michoacan earthquake as inferred from near-source ground-motion records, *Bull. Seismol. Soc. Am.* **81,** no. 3, 844–861.

Nakamura, Y. (1996). Real-time information systems for seismic hazards mitigation UrEDAS, *Q. Rep. Railway Tech. Res. Inst.* **37,** 112–127.

Perez-Yañez, C., L. Ramirez-Guzman, G. A. L. Ruiz, D. R. Delgado, A. Macias C Marco, G. H. Sandoval, N. L. Alcantara, and R. A. Quiroz (2014). Strong ground motion database system for the Mexican seismic network, *AGU Fall Meeting*, San Francisco, California, 14–19 December.

Rosenblueth, E. (1986). The 1985 earthquake: Causes and effects in Mexico City, *J. Am. Concr. Inst.* **8,** no. 5, 23–24.

Singh, S. K., L. Astiz, and J. Havskov (1981). Seismic gaps and recurrence periods of large earthquakes along the Mexican subduction zone: A reexamination, *Bull. Seismol. Soc. Am.* **71,** no. 3, 827–843.

Suárez, G., D. Novelo, and E. Mansilla (2009). Performance evaluation of the seismic alert system (SAS) in Mexico City: A seismological and a social perspective, *Seismol. Res. Lett.* **80,** no. 5, 707–716.

Centro de Instrumentación y Registro Sísmico (CIRES)
Anaxagoras 814
Colonia Narvarte
Mexico City, Mexico 03020
cuellara@cires.org.mx
espinosajm@cires.org.mx
(A.C., J.M.E.-A.)

Instituto de Geofísica
Universidad Nacional Autónoma de México
Mexico City, Mexico 04510
gerardo@geofisica.unam.mx
(G.S.)

Manuscript received 6 April 2017;
Published Online 19 December 2017