AGU PUBLICATIONS

Geophysical Research Letters

RESEARCH LETTER
10.1002/2014GL061564

Key Points:
- The log-average period is better correlated with magnitude than other proxies
- The log-average period provides better magnitude prediction than other proxies
- The real-time log-average period is expected to enhance the performances of EEWS

Supporting Information:
- Readme
- Figure S1
- Figure S2

Correspondence to:
A. Ziv,
zivalon@tau.ac.il

Citation:
Ziv, A. (2014), New frequency-based real-time magnitude proxy for earthquake early warning, *Geophys. Res. Lett.*, *41*, doi:10.1002/2014GL061564.

Received 20 AUG 2014
Accepted 24 SEP 2014
Accepted article online 28 SEP 2014

New frequency-based real-time magnitude proxy for earthquake early warning

A. Ziv[1]

[1] Department of Geosciences, Tel-Aviv University, Tel-Aviv, Israel

Abstract The two frequency-based magnitude proxies currently employed by earthquake early warning systems in California are the predominant and the characteristic periods. These proxies, obtained using simple expressions that are valid for noise-free monochromatic signals, yield erroneous result. The log-average period, τ_{log}, introduced in this study, is calculated directly from the actual velocity spectrum of the first few seconds of the seismic record. Using data from South California and Japan consisting of 440 earthquakes whose magnitudes range between 3 and 7.3, it is demonstrated that τ_{log} is better correlated with the catalog magnitude than the predominant period and provides better magnitude assessment than the characteristic period for small magnitudes (M <4). The average prediction error is reduced with increasing the input interval up to 6 s. It appears that a single linear scaling describes the relation between $\log(\tau_{log})$ and the catalog magnitude for the entire magnitude range studied here.

1. Introduction

As short-term prediction of individual earthquakes is not feasible, considerable efforts are directed toward improving existing algorithms for earthquake early warning systems (EEWS) and developing new ones. These real-time algorithms utilize the early part of the seismic signals to assess the magnitude and/or the peak ground motion and use that information to issue warning to specific sites. Determining these parameters based on the first few seconds of the seismic record in real time is a challenging task. As both the amplitude and the frequency content of the early ground motion are related to the earthquake size, both may be used as real-time magnitude proxies. Yet while the distance dependence of the ground motion amplitude is strong, that of the *P* wave frequency content is moderate for distances relevant to EEWS [*Wu and Kanamori*, 2005b; *Zollo et al.*, 2010; *Colombeli et al.*, 2012]. Thus, frequency-based real-time magnitude proxies need not be corrected for hypocentral distances, and their implementation in EEWS is more straightforward than that of their amplitude-based counterparts.

The idea of using the frequency content of the initial ground motion as a real-time magnitude proxy is attributed to Nakamura, who took a simplifying approach of inferring the predominant period from the square root of the ratio between the integrated velocity squared and the integrated acceleration squared as [*Nakamura*, 1988]:

$$\tau_p = 2\pi\sqrt{V_n/D_n}, \tag{1a}$$

with:

$$V_n = \alpha V_{n-1} + v_n^2 \qquad \text{and} \qquad D_n = \alpha D_{n-1} + (\dot{v}_n)^2, \tag{1b}$$

where n is a time index, v is the recorded particle velocity, V is the smoothed particle velocity squared, the dot signifies time derivative, D is the smoothed acceleration and α is a smoothing coefficient, whose value is slightly less than unity. *Allen and Kanamori* [2003] defined the parameter τ_p^{max} as the maximum of τ_p within 2 or 4 s after the onset of the *P* wave. Owing to the erratic behavior of τ_p during low SNR intervals [see *Olson and Allen*, 2006, Figure 2], it is useful to disregard the earliest part of the signal. While *Olson and Allen* [2005] disregarded the first 0.05 s, here it is found that a slightly better result is obtained if the first 0.5 s is disregarded.

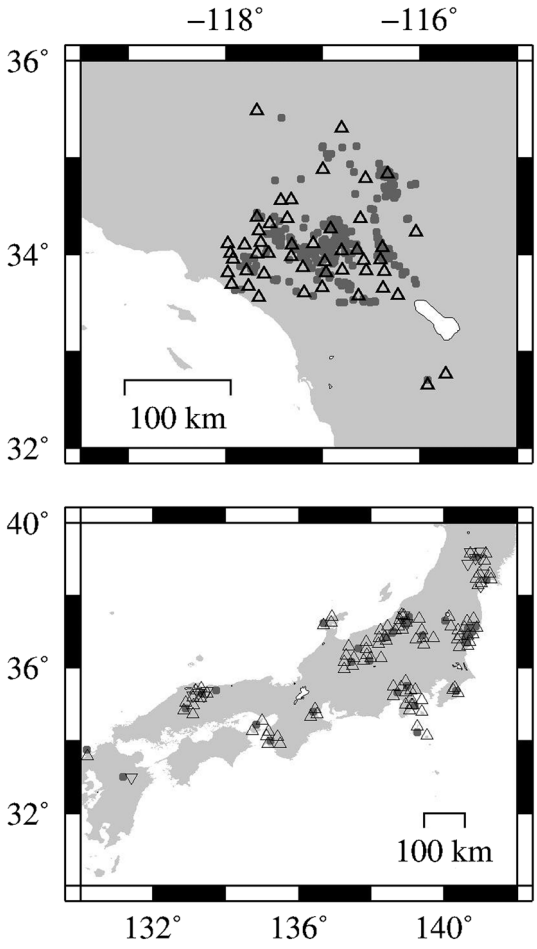

Figure 1. Location maps showing earthquake (gray circles) and station (triangles) distributions in (top) South California and (bottom) Japan. Borehole and surface instruments are indicated by down-pointing and up-pointing triangles, respectively.

Wu and Kanamori [2005a] proposed the characteristic period:

$$\tau_c = 2\pi \sqrt{\frac{\int_{t_0}^{t_0+\Delta t} u^2 dt}{\int_{t_0}^{t_0+\Delta t} v^2 dt}}, \quad (2)$$

where u is the ground displacement. *Shieh et al.* [2008] showed that τ_p is more sensitive to the presignal noise level than τ_c, and *Wolfe* [2006] showed that both proxies are nonlinearly dependent on the spectral amplitude and the frequency in a manner that overweights the highest frequencies in the spectrum.

In this study a new frequency-based real-time magnitude proxy is introduced: the log-average period, $\tau_{\log}$. It is shown that $\tau_{\log}$ is better correlated with the catalog magnitude than τ_c and $\tau_p^{\max}$ and therefore enhances the performance of EEWS.

2. Data

A composite data set is used that consists of 400 earthquakes from South California, whose magnitudes range between 3 and 5.7, and 40 earthquakes from Japan, whose magnitudes range between 4 and 7.3 (Figure 1). Only vertical components triggered seismograms are utilized. The total number of seismograms is about 1130, with 950 seismograms recorded by the Southern California Seismic Network (SCSN) broadband instruments, 160 accelerograms recorded by the K-NET strong motion sensors and 14 accelerograms recorded by the KiK-NET strong motion borehole sensors. The two data selection criteria that were used are that the SNR be greater than 10 and that the hypocentral distances be less than 30 km. These criteria guarantee that the data are of high quality and are (nearly) unaffected by anelastic attenuation. The sampling rate of the broadband records is equal to 40 Hz and that of the strong motion records is equal to 100 or 200 Hz. Acceleration records were integrated once and both integrated accelerograms and velocity seismograms were high-pass filtered at 0.075 Hz, in order to remove long-period trends and nonzero baseline. First P arrivals of $M \geq 5$ were picked manually, and those of smaller magnitude were picked automatically using a standard short-term averaging/long-term averaging algorithm.

3. The Log-Average Period

At the time τ_p was first introduced by *Nakamura* [1988], computing the actual spectrum of a few seconds long seismograms was much more time consuming than it is today; thus, it was not suitable for the EEWS algorithm that he designed, where the predominant periods are recalculated at each sample. Nowadays, such calculations are so fast that they may be readily incorporated in Nakamura's algorithm, and other algorithms adopting Nakamura's approach.

The rationale underling τ_p, τ_c, and the new parameter introduced here, $\tau_{\log}$, is exactly the same; because the frequency content of the initial ground motion is magnitude dependent, parameters quantifying the average/dominant frequency content of the first few seconds of the seismic signal may be used as real-time

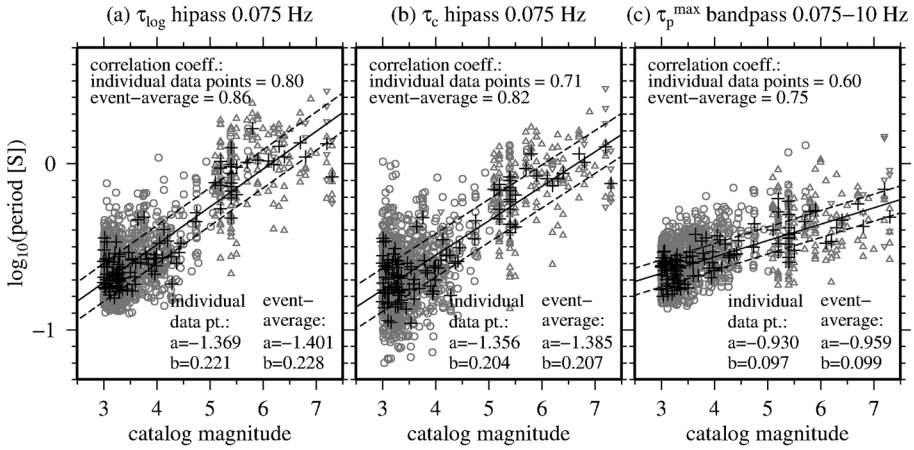

Figure 2. Period of the first 3 s after the *P* wave arrival as a function of catalog magnitude. (a) Log-average period, τ_{log}, of high-passed filtered seismograms at 0.075 Hz. (b) Characteristic period, τ_c, of high-passed filtered seismograms at 0.075 Hz. (c) Predominant period, τ_p^{max}, of band-passed seismograms between 0.075 and 10 Hz. Individual data points and event-average values are indicated by gray symbols and dark crosses, respectively. Circles indicate data recorded by the South California broadband seismometers. Strong motion data recorded by the KiK borehole and the K-NET strong motion networks are indicated by down-pointing and up-pointing triangles, respectively. Solid and dashed lines are the best fits of event-average values to equation (4) and their respective standard deviations. Correlation coefficients and fitting coefficients corresponding to individual data points and event-average are indicated on each panel.

magnitude proxies. While τ_p and τ_c are evaluated through application of simple expressions (equations (1) and (2)) that yield correct solutions only in cases where the signal is noise free and monochromatic [*Wolfe*, 2006], τ_{log} is calculated directly from the actual velocity spectra and therefore reflects the signal's true frequency content.

The log-average period, τ_{log}, is calculated from the spectrum of the early part of the velocity seismogram as follows: (a) a prespecified interval is extracted, starting at the time of the *P* wave arrival, (b) a Hann window is applied to reduce abrupt end effects, (c) the resulting signal is Fourier transformed to get the set of power spectrum coefficients $P_i(w_i)$, (d) the set of uniformly spaced P_i is resampled to get a set that is spaced every 0.1 log unit of frequency, between 0.1 and 10 Hz, and finally, (e) the log-average period, τ_{log}, is obtained through:

$$\log\left(\tau_{\log}\right) = \frac{\sum_i \left(P_i^*(w_i)\log(1/w_i)\right)}{\sum_i \left(P_i^*(w_i)\right)}, \tag{3}$$

where the asterisk indicates that the power spectrum is resampled according to step (d). It is emphasized that replacing P_i with P_i^* in the above expression, and/or multiplying the power spectrum coefficients by $1/w_i$ instead of $\log(1/w_i)$ would result in an average period that is biased toward the highest frequencies in the spectrum. Next it is shown that the switch from approximate quantities, such as τ_c and τ_p^{max}, to the exact quantity τ_{log} improves real-time magnitude estimates.

3.1. Fixed Interval

The magnitude proxies τ_{log}, τ_c and τ_p^{max} are calculated using 3 s long intervals starting at the time of the *P* wave arrival. The logarithm of the three periods as a function of catalog magnitude is shown in Figure 2. Empirical relations between each proxy and earthquake magnitude are obtained through fitting the three data sets with:

$$\log(\text{period}) = a + b \times \text{magnitude}, \tag{4}$$

where *a* and *b* are the fitting coefficients of the linear regression. Despite the large data scatter, the correlations between the log of periods and the magnitudes are statistically meaningful for all three magnitude proxies (correlation coefficients are reported on each panel in Figure 2). It is found that τ_{log} is better correlated with earthquake magnitude than τ_c, and that the correlation of the latter with the magnitude is better than that of τ_p^{max}.

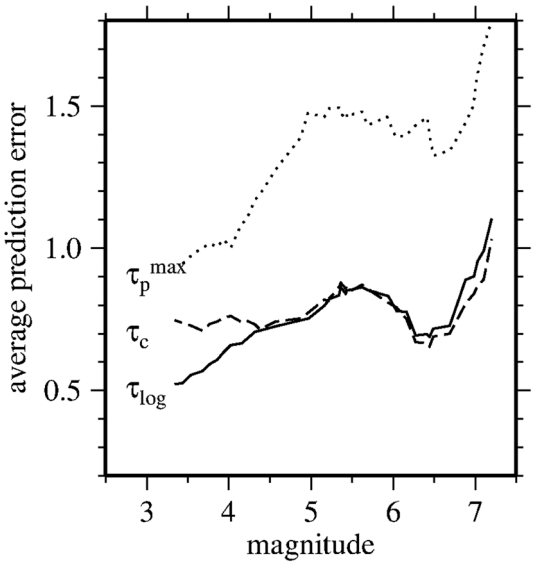

Figure 3. Average prediction errors as a function of magnitude for τ_{log} (solid line), τ_c (dashed line), and τ_p^{max} (dotted line). Periods are calculated using 3 s long intervals.

Previous studies show that while the correlation of individual τ_p^{max} or τ_c with catalog magnitude is rather weak, that of event-average is much stronger [*Allen and Kanamori*, 2003; *Brown et al.*, 2011; *Lancieri et al.*, 2011; *Sadeh et al.*, 2013]. Out of the 440 earthquakes analyzed in this study, 113 were recorded by four or more stations. Indeed, it is found that the correlation between event-average magnitude proxies and earthquake magnitude is significantly better than that of individual proxies. Again here, event-average τ_{log} are better correlated with earthquake magnitude than event-average τ_c and τ_p^{max}.

To further assess the performances of the three magnitude proxies, it is useful to examine the difference between observed and predicted magnitudes, with the latter being estimated through application of equation (4). Absolute prediction errors as a function of catalog magnitudes are shown in Figure S1, and their running averages are compared in Figure 3. On average, the magnitude assessments obtained using τ_{log} and τ_c are more accurate than those obtained using τ_p^{max}. In addition, τ_{log} yields better estimates than τ_c for earthquakes with magnitude less than about 4. It is therefore inferred that use of τ_{log} is expected to reduce the rate of false alarms, and thus enhance the overall performances of both regional (i.e., network-based) and onsite (i.e., station-based) EEWS.

3.2. Prediction Error Versus Interval Length

The increase in the prediction error with increasing magnitude for the $M>6$ earthquakes (Figure 3) is not surprising, as the rupture of these earthquakes must have lasted more than 3 s. To see if the worsening in the magnitude prediction for $M>6$ is due to the input data being too short compared to the rupture duration, prediction errors were analyzed using different interval lengths (Figures S2 and 4). Indeed, the correlation between $\log(\tau_{log})$ and the catalog magnitude increases steadily upon increasing the interval of the velocity

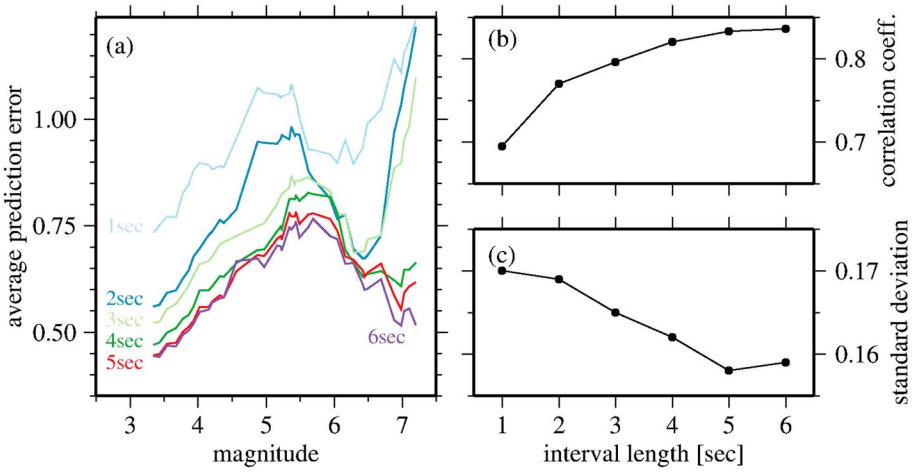

Figure 4. Diagrams summarizing the τ_{log}-vs-magnitude plots in Figure S2. (a) Average prediction error as a function of magnitude for interval lengths between 1 and 6 s. (b and c) Correlation coefficient and standard deviation as a function of interval length corresponding to best fitting $\log(\tau_{log})$ versus magnitude according to equation (4).

seismograms from 1 to 6 s. Consequently, the average magnitude discrepancy decreases dramatically with increasing the input time series interval up to 6 s (Figure 4a).

The notable increase in the average prediction error for $5 < M < 6$ observed for all intervals considered in Figure 4 is attributed mainly to a couple of dozen accelerograms that were recorded by the K-NET strong motion instruments. Visual inspection of these records reveals the presence of long period noise. Eliminating this noise by high passing the input record at a frequency higher than the 0.075 Hz used here is not a solution, as this will inevitably result in the wiping off of some of the signal. Because, in general, accelerograms recorded by the K-NET instruments are noisier than those recorded by the KiK borehole instruments, it is sensible to assign more weight to data recorded by the KiK borehole instruments.

Changes in scaling for moderate and large earthquakes were reported by previous researchers. *Allen and Kanamori* [2003] inferred a different τ_p^{max} versus magnitude relation for magnitudes between 3 to 5 and for those greater than 5. *Lancieri et al.* [2011] reported a change in the scaling of τ_c between $M_w < 6$ and $M_w > 6$ earthquakes. No such scaling break is apparent in the distribution of τ_{log} versus magnitude in Figures S2e and S2f, and it appears that a single linear scaling describes the relation between $\log(\tau_{log})$ and the catalog magnitude for the entire magnitude range studied here.

4. Summary and Conclusions

Because the initial ground motion caused by small earthquakes is richer in frequencies higher than that caused by large earthquakes, a parameter that quantifies the frequency content of the first few seconds of the seismic signal may be used as real-time magnitude proxy and be implemented for EEWS [*Nakamura*, 1988; *Allen and Kanamori*, 2003; *Wu and Kanamori*, 2005a]. The two frequency-based distance-independent real-time magnitude proxies that are currently in use or are under test are the predominant period, τ_p^{max} [*Nakamura*, 1988; *Allen and Kanamori*, 2003], and the characteristic period, τ_c [*Wu and Kanamori*, 2005a]. These proxies, however, are extremely sensitive to the presignal noise level and are nonlinearly dependent on the spectral amplitude and the frequency in a manner that overweight the highest frequencies in the spectrum [*Wolfe*, 2006]. Unlike the τ_p^{max} and the τ_c parameters, the log-average period ,τ_{log}, introduced in this study is obtained from the actual spectrum of the first few seconds after the *P* wave arrival.

The performance of τ_{log} is compared to that of τ_c and τ_p^{max}, using a composite data set that consists of broadband data from South California and strong motion data from Japan. It is shown that τ_{log} is better correlated with catalog magnitude than τ_c and that the correlation of the latter with catalog magnitude is better than that of τ_p^{max}. In addition, τ_{log} yields better magnitude estimates than τ_c for $M < 4$ and yields similar accuracy for larger earthquakes. Thus, replacing τ_c with τ_{log} is expected to improve EEWS performance, by reducing the fraction of false alarms. The correlation between $\log(\tau_{log})$ and the catalog magnitude increases steadily, and the average magnitude discrepancy decreases upon increasing the interval of the velocity seismograms from 1 to 6 s. It seems that a single linear scaling describes the relation between $\log(\tau_{log})$ and the catalog magnitude for magnitudes between 3 and 7.3, if the data interval is longer than 4 s.

Speeding-up real-time magnitude determination by a few seconds may be of great importance at short hypocentral distances. It is therefore instructive to seek ways to minimize the computation time, while keeping the magnitude discrepancies acceptably small. Guided by this view, *Caprio et al.* [2011] developed and validated a real-time evolutionary magnitude determination scheme that utilizes the actual spectrum of the early part of the seismogram. In their scheme, the displacement spectra are updated repeatedly every second for increasing time windows, and each time are inverted for the low-frequency plateau, the corner frequency and the (frequency-independent) attenuation factor. Then, a simple expression is employed [*Aki and Richards*, 2002] that relates the seismic moment to the low-frequency plateau, the hypocentral distance, the rock densities and the shear wave speeds. The main advantage of this approach with respect to τ_{log} is that it does not rely on empirically derived relationships and is therefore readily implementable without having to go through a calibration phase. The obvious disadvantage of this scheme with respect to the τ_{log} proxy is that it requires knowledge of the hypocentral location. Consequently, the accuracy of this magnitude estimate is limited by the accuracy of the real-time hypocenter. That the vast majority of the τ_{log} reported above are smaller than a second and only a tiny fraction of them is greater than 3 s, implies that the magnitude of most earthquakes may be estimated after just 1 or 2 s, and that the incorporation of τ_{log} into available evolutionary schemes may enhance their performances.

Acknowledgments
The facilities of the Southern California Earthquake Data Center (SCEDC), and the Southern California Seismic Network (SCSN), as well as KiK-net and K-NET online databases were used for access to waveforms and parametric data used in this study. Comments from H. Wust-Bloch helped to improve the clarity of this article. I thank the Editor, Andrew Newman, Aldo Zollo, and an anonymous reviewer for their very constructive remarks.

Andrew Newman thanks Anthony Lomax and Aldo Zollo for their assistance in evaluating this paper.

References

Allen, R. M., and H. Kanamori (2003), The potential for earthquake early warning in Southern California, *Science*, *300*, 786–789.

Aki, K., and P. Richards (2002), *Quantitative Seismology,* 2nd ed., Univ. Sci., Sausalito, Calif.

Brown, H. M., R. M. Allen, M. Hellweg, O. Khainovski, D. Neuhauser, and A. Souf (2011), Development of the ElarmS methodology for earthquake early warning: Realtime application in California and offline testing in Japan, *Soil Dyn. Earthquake Eng.*, *31*(2), 188–200.

Caprio, M., M. Lancieri, G. B. Cua, A. Zollo, and S. Wiemer (2011), An evolutionary approach to real-time moment magnitude estimation via inversion of displacement spectra, *Geophys. Res. Lett.*, *38*, L02301, doi:10.1029/2010GL045403.

Colombeli, S., O. Amoroso, A. Zollo, and H. Kanamori (2012), Test of a threshold-based earthquake early-warning method using Japanese data, *Bull. Seismol. Soc. Am.*, *102*(3), 1266–1275.

Lancieri, M., A. Fuenzalida, S. Ruiz, and R. Madariaga (2011), Magnitude scaling of early-warning parameters for the Mw 7.8 Tocopilla, Chile, earthquake and its aftershocks, *Bull. Seismol. Soc. Am.*, *101*(2), 447–463.

Nakamura, Y. (1988), On the urgent earthquake detection and alarm system (UrEDAS), paper presented at 9th World Conference on Earthquake Engineering VII, Tokyo, 673–678.

Olson, E. L., and R. M. Allen (2005), The deterministic nature of earthquake rupture, *Nature*, *438*, 212–215, doi:10.1038/nature04214.

Olson, E. L., and R. M. Allen (2006), Is earthquake rupture deterministic? (Reply), *Nature*, *442*, E6, doi:10.1038/nature04964.

Sadeh, M., A. Ziv, and H. Wust-Bloch (2013), Real-time magnitude proxies for earthquake early warning in Israel, *Geophys. J. Int.*, *196*(2), 939–950.

Shieh, J. T., Y. M. Wu, and R. M. Allen (2008), A comparison of τ_c and τ_p^{max} for magnitude estimation in earthquake early warning, *J. Geophys. Res.*, *35*, L20301, doi:10.1029/2008GL035611.

Wolfe, C. J. (2006), On the properties of predominant-period estimators for earthquake early warning, *Bull. Seismol. Soc. Am.*, *96*(5), 1961–1965.

Wu, Y. M., and H. Kanamori (2005a), Experiment on an onsite early warning method for the Taiwan early warning system, *Bull. Seismol. Soc. Am.*, *95*(1), 347–353.

Wu, Y. M., and H. Kanamori (2005b), Rapid assessment of damage potential of earthquakes in Taiwan from the beginning of the P wave, *Bull. Seismol. Soc. Am.*, *95*(3), 1181–1185.

Zollo, A., O. Amoroso, M. Lancieri, Y. M. Wu, and H. Kanamori (2010), A threshold-based earthquake early warning using dense accelerometer networks, *Geophys. J. Int.*, *183*, 354–365.