Real-time testing of the on-site warning algorithm in southern California and its performance during the July 29 2008 Mw 5.4 Chino Hills earthquake

M. Böse,1 E. Hauksson,1 K. Solanki,1 H. Kanamori,1 and T. H. Heaton1

Received 17 October 2008; revised 26 November 2008; accepted 10 December 2008; published 5 February 2009.

[1] The real-time performance of the τc-Pd on-site early warning algorithm currently is being tested within the California Integrated Seismic Network (CISN). Since January 2007, the algorithm has detected 58 local earthquakes in southern California and Baja with moment magnitudes of 3.0 ≤ Mw ≤ 5.4. Combined with newly derived station corrections the algorithm allowed for rapid determination of moment magnitudes and Modified Mercalli Intensity (MMI) with uncertainties of ±0.5 and ±0.7 units, respectively. The majority of reporting delays ranged from 9 to 16 s. The largest event, the July 29 2008 Mw 5.4 Chino Hills earthquake, triggered a total of 60 CISN stations in epicentral distances of up to 250 km. Magnitude predictions at these stations ranged from Mw 4.4 to Mw 6.5 with a median of Mw 5.6. The closest station would have provided up to 6 s warning at Los Angeles City Hall, located 50 km to the west-northwest of Chino Hills.


1. Introduction

[2] The purpose of earthquake early warning (EEW) is to provide real-time information about earthquakes to distant sites before the seismic S or surface waves arrive. Because warning times usually are extremely short, EEW systems must recognize the severity of expected ground motions within seconds after the P-wave arrival at the EEW sensors. If warnings can be issued in a timely manner, suitable actions for damage mitigation can be initiated and executed [e.g., Goltz, 2002].

[3] Currently, the performance of the τc-Pd on-site warning algorithm [Kanamori, 2005; Wu and Kanamori, 2005a, 2005b; Wu et al., 2007; Wu and Kanamori, 2008a, 2008b] is being real-time tested within the California Integrated Seismic Network (CISN) [Hauksson et al., 2006]. The algorithm is based on single sensor observations using two parameters: period parameter τc and high-pass filtered displacement amplitude Pd. Both parameters are determined from the vertical components of velocity and/or displacement data, \( \dot{u} \) and \( u \), using the first \( t_0 = 3 \) s of P-waveforms. The period parameter \( \tau_c \), computed by

\[
\tau_c = 2\pi \sqrt{\frac{\int_0^t \dot{u}^2(t)dt}{\int_0^t u^2(t)dt}},
\]

approximately represents the P-wave pulse width [Wu et al., 2008b]. Previous studies have determined empirical relationships between \( \tau_c \) and the moment magnitudes Mw, and between Pd and the peak ground velocities (PGV) at the sites of observation [Kanamori, 2005; Wu and Kanamori, 2005a, 2005b]. Wu et al. [2007] established corresponding relationships for earthquakes in southern California using seismic off-line data. Observed and estimated values of PGV can be transformed into Modified Mercalli Intensity (MMI) scale using empirical relationships developed by Wald et al. [1999].

[4] For the real-time testing, we have implemented the \( \tau_c \cdot P_d \) algorithm in an UNIX environment, using existing software components developed by the California Institute of Technology (Caltech), the U.S. Geological Survey (USGS), and UC Berkeley, that are built on software systems developed for the CISN and the Advanced National Seismic System (ANSS). The processing steps are as follows [Solanki et al., 2007]: (1) retrieve velocity data from the CISN; (2) set the baseline to 0 by using average values continuously determined from the real-time data streams in intervals of 60 s, and apply gain correction; (3) convert velocity to displacement data by recursive integration; apply high-pass Butterworth filter (>0.075 Hz); (4) calculate \( \tau_c \) and Pd from the initial 3 s of waveform data; (5) keep only triggers with \( \tau_c \cdot P_d \) combinations that are characteristic of a local earthquake [Böse et al., 2009]; for a local earthquake with period \( \tau_c \) in a rupture-to-site distance \( r \), \( r_{min} \leq r \leq r_{max} \), we expect \( P_{d,min} \leq P_d \leq P_{d,max} \). Böse et al. [2009] determined displacement amplitudes \( P_{d,min} \) and \( P_{d,max} \) from empirical attenuation relations for earthquakes in southern California with \( r_{min} = 1 \) km and \( r_{max} = 100 \) km. To avoid false alerts, we currently require the triggering of at least 3 stations before an earthquake is processed.

[5] To improve the accuracy of Mw and PGV estimates, we refined in this study the \( \tau_c \cdot M_w \) and \( \tau_c \cdot P_d \cdot PGV \) relations by Wu et al. [2007] with new station corrections. We determined these factors from the median of residuals from (1) 431 off-line estimates of Mw during 27 earthquakes (4.0 ≤ Mw ≤ 7.3) [Wu et al., 2007], and from (2) 257 real-time estimates of Mw during 58 earthquakes (3.0 ≤ Mw ≤ 5.4) as analyzed in this paper. Correction factors were determined and applied in this study only for stations for which at least 2 records were available.

1 Seismological Laboratory, California Institute of Technology, Pasadena, California, USA.
the largest earthquake to occur in the greater Los Angeles Basin (33.95°N, −117.76°W, 14.7 km depth) was the largest earthquake to occur in the greater Los Angeles metropolitan area since the Mw 6.7 Northridge earthquake in 1994. The event was widely felt across southern California, but caused only minor damage [Hauksson et al., 2008]. The Chino Hills earthquake sequence produced 97 estimates of Mw and PGV values by the τ-P algorithm: 60 during the Mw 5.4 mainshock (Table 1) and between 8 and 15 during the three largest aftershocks (Mw 2.8, Mw 3.8, and Mw 3.6).

Table 1. Performance of the On-Site Warning Algorithm at Caltech During 9 Local Earthquakes (Mw > 4.5) in Southern California and Baja

<table>
<thead>
<tr>
<th>Origin Time (PST)</th>
<th>Latitude and Longitude (deg)</th>
<th>Mw</th>
<th>Number of Reports</th>
<th>Time of First Report</th>
<th>First Reporting Station</th>
<th>Estimated Mw at First Reporting Station</th>
<th>Median and Scattering in Mw Over All Reports</th>
</tr>
</thead>
<tbody>
<tr>
<td>2008-7-29 11:42:15</td>
<td>33.95, −117.76</td>
<td>5.4</td>
<td>60</td>
<td>11:42:25</td>
<td>CLPSRb</td>
<td>5.6</td>
<td>5.6</td>
</tr>
<tr>
<td>2008-2-10 10:29:30</td>
<td>32.33, 5.1</td>
<td>2</td>
<td>2</td>
<td>10:29:57</td>
<td>CLDRE</td>
<td>5.3</td>
<td>5.0</td>
</tr>
<tr>
<td>2008-2-1 23:12:30</td>
<td>32.43, 5.0</td>
<td>5</td>
<td>5</td>
<td>23:12:30</td>
<td>CLDRE</td>
<td>5.4</td>
<td>5.4</td>
</tr>
<tr>
<td>2008-2-19 14:41:52</td>
<td>32.45, 5.0</td>
<td>6</td>
<td>6</td>
<td>14:41:52</td>
<td>CLDRE</td>
<td>4.8</td>
<td>4.8</td>
</tr>
<tr>
<td>2008-2-23 23:33:03</td>
<td>32.55, 4.8</td>
<td>4</td>
<td>4</td>
<td>23:33:03</td>
<td>CLDRE</td>
<td>4.7</td>
<td>4.7</td>
</tr>
<tr>
<td>2008-2-29 11:31:29</td>
<td>32.43, 4.8</td>
<td>27</td>
<td>27</td>
<td>11:31:29</td>
<td>CLDRE</td>
<td>4.0</td>
<td>4.0</td>
</tr>
<tr>
<td>2008-3-7 14:41:52</td>
<td>32.55, 4.8</td>
<td>1</td>
<td>1</td>
<td>14:41:52</td>
<td>CLDRE</td>
<td>4.6</td>
<td>4.6</td>
</tr>
<tr>
<td>2008-7-29 23:12:05</td>
<td>32.55, 4.8</td>
<td>1</td>
<td>1</td>
<td>23:12:05</td>
<td>CLDRE</td>
<td>3.9</td>
<td>3.9</td>
</tr>
<tr>
<td>2008-7-29 23:12:05</td>
<td>32.55, 4.8</td>
<td>27</td>
<td>27</td>
<td>23:12:05</td>
<td>CLDRE</td>
<td>3.2</td>
<td>3.2</td>
</tr>
<tr>
<td>2008-7-29 23:12:05</td>
<td>32.55, 4.8</td>
<td>26</td>
<td>26</td>
<td>23:12:05</td>
<td>CLDRE</td>
<td>4.6</td>
<td>4.6</td>
</tr>
</tbody>
</table>

a Magnitude estimates are station corrected. O.T. is origin time, Δ epicentral distance.
b Direct ethernet radio link to Caltech.
c Virtual Private Network (VPN) over microwave.

d Based on the performance between January 2007 and September 2008, we want to analyze in this paper (1) the real-time applicability of the τ-P algorithm for earthquakes in southern California, and (2) the suitability of the current CISN instrumentation and telemetry for issuing EEW. Of course, small and moderate-sized earthquakes (Mw < 6.0) as analyzed in this paper usually will not cause damage and therefore do not require EEW. To gain experience more quickly and to have working algorithms when large earthquakes occur, we use the more frequent small events for testing and calibrating of EEW algorithms and systems [Böse et al., 2009].

3. Results

Our Mw and PGV values of the 58 local earthquakes were estimated automatically by our software from the observed τ and P values using the relations proposed by Wu et al. [2007]. The real-time estimated parameters correlate well with the corresponding values reported by CISN (Figures 1a and 1c). However, the scatter is often quite large, sometimes with outliers of as much as two magnitude units. The median values, taken for each earthquake over all available Mw estimates, usually show a slight overestimation of Mw by 0.3 units. The uncertainties in the predictions of magnitude (Figure 1a) and the logarithmic values of PGV (Figure 1c) are ±0.6 and ±0.3 units, respectively. The latter is equivalent to an uncertainty of ±0.75 MMI intensity units [Wald et al., 1999].

[9] Both Mw and PGV values of the 58 local earthquakes were estimated automatically by our software from the observed τ and P values using the relations proposed by Wu et al. [2007]. The real-time estimated parameters correlate well with the corresponding values reported by CISN (Figures 1a and 1c). However, the scatter is often quite large, sometimes with outliers of as much as two magnitude units. The median values, taken for each earthquake over all available Mw estimates, usually show a slight overestimation of Mw by 0.3 units. The uncertainties in the predictions of magnitude (Figure 1a) and the logarithmic values of PGV (Figure 1c) are ±0.6 and ±0.3 units, respectively. The latter is equivalent to an uncertainty of ±0.75 MMI intensity units [Wald et al., 1999].

[10] The newly derived station corrections lead to significant improvement of the predictions, but some outliers still remain (Figures 1b and 1d); the majority of them are associated with events with small magnitudes (Mw < 4.5) and large epicentral distances (Δ > 150 km), i.e. are caused by poor signal-to-noise ratios. The station corrections reduce the errors in magnitude and intensity estimates to ±0.5 and ±0.7 units, respectively (Figures 1b and 1d). In general, the Mw and PGV values estimated from the off-line processed strong motion data agree well with the real-time processed broadband observations of the Chino Hills earth-
quake (triangles in Figures 1a–1d), but show more scatter, possibly due to poor signal-to-noise ratios. Note that the strong motion data in Figures 1a–1d are not station corrected.

[11] The largest outlier in Figures 1a and 1b is observed for a Mw4.0 earthquake which occurred on March 30, 2007 at six kilometers east of Coso Junction, an area known for the frequent occurrence of earthquake swarms: the magnitude of this event was overestimated by 1.5 units. Böse et al. [2009] suspected that this overestimation might have been caused by a small foreshock which occurred 48 s before the mainshock resulting in a high background noise level at the majority of close EEW stations before and during the arrival of the seismic P phase from the mainshock. Although such foreshocks are relatively rare, the real-time identification of such events will pose a major challenge in the future developments of EEW systems [Böse et al., 2009].

[12] The real-time performance of the \( \tau_r-P_d \) on-site warning algorithm for all analyzed earthquakes with \( M_w > 4.5 \) is summarized in Table 1. The errors in the station corrected estimates of \( M_w \) at the first reporting station reach from 0.0 to 0.8 units. Six of the events occurred within an earthquake swarm near the Cerro Prieto Geothermal field at the U.S./Mexican border, which started \( \approx \)20 miles southeast of Calexico with a Mw5.1 earthquake on February 8, 2008 (http://www.scsn.org/2008bajaaddendum.html). Note that the first reporting station CI.DRE is located at approximately 40 to 55 km north, i.e. relatively far away from the epicenters of the large events in the swarm. The first magnitude and PGV estimates thus were not available until 23 s after origin time (O.T.).

[13] The first magnitude prediction of the Mw5.4 Chino Hills mainshock was available 10 s after O.T. (Mw5.6 at station CI.PSR). Subsequent (independent) estimates based on data from other stations ranged from Mw4.4 to Mw6.5 with a median value of Mw5.6 (Table 1). The MMI intensities determined from PGV [Wald et al., 1999] were slightly underestimated by 0.2 ± 0.8 units. The largest prediction errors occurred in the western part of the Los Angeles basin where seismic wave amplitudes were strongly amplified due to basin effects (Figures 2a and 2b). Because the current database is insufficient for the determination of correction terms for many of the stations shown in Figure 2b, the PGV estimates in the map are not station corrected.

[14] Neglecting the telemetry and processing delays of the current system, each single estimate in Figure 2b could have been made available within 3 s after the P-wave arrival (This is the time required by the \( \tau_r-P_d \) algorithm.). The entire map in Figure 2b was available in less than 1 minute after O.T. For comparison the first automatically generated CISN ShakeMap [Wald et al., 1999] of the Chino Hills earthquake was released about 12 minutes after O.T. [Hauksson et al., 2008].

[15] To illustrate the effect of EEW delays in more detail, we have drawn in Figure 2 circles centered at the epicenter of the Mw5.4 Chino Hills earthquake. These circles show

---

Figure 1. Estimated vs. observed source and ground motion parameters of 58 local earthquakes (3.0 \( \leq M_w \) \( \leq 5.4 \)) in southern California and Baja: (a and b) moment magnitudes \( M_w \) and (c and d) peak ground velocities (PGV). Estimates in Figures 1b and 1d include correction factors, which are applied to stations with observations during at least two earthquakes. Crosses show real-time processed data, triangles off-line processed strong motion data of the July 29 2008 Mw5.4 Chino Hills earthquake. The median \( M_w \) values in Figures 1a and 1b, taken over the magnitude estimates of each earthquake, are shown by squares. Values along the dashed lines show a perfect correlation.
the predicted locations of the S-wave arrivals at 10 s, 20 s, 30 s and 50 s after O.T., assuming a constant shear wave velocity of 3.2 km/s. Depending on the time, at which a warning or estimate is delivered, only sites outside of the corresponding circle can obtain a warning before the S wave arrives. This delay depends on (a) the P-wave arrival time at the reporting EEW sensor, which is controlled by the station density, and (b) the so-called warning delay between P-wave arrival and the reporting of parameters and/or warnings. The warning delays include delays caused by station equipment (in particular by the type of datalogger), by telemetry of waveform data to the central processing facility at Caltech, 3 s waveform data required by the \( t_{c-P_d} \) algorithm, and processing delays. In an operational system, additional delays will be caused by the transmission of warnings to users.

[16] The current CISN configuration provided a first estimate 10 s after O.T. of the Chino Hills mainshock (Table 1). In an operational EEW system, only sites outside of the smallest circle in Figure 2 with a radius of 30 km could have obtained a warning. As an example, at Los Angeles City Hall, located 50 km to the west-northwest of Chino Hills, the current system would have provided up to 6 s warning before S-wave arrival. For comparison the first M_\text{L} estimate (M_\text{L}5.6) by CISN was automatically released ~80 s after O.T., an up-dated estimate (M_\text{L}5.8) around 60 s later. The automatic moment tensor and M_\text{w} estimate (M_\text{w}5.4) were available ~10 min after O.T. [Hauksson et al., 2008].

[17] From January to September 2008, the testing system at Caltech recorded around 50,000 triggers, which were mostly not caused by local earthquakes and not processed by the EEW algorithm. Fortuitously, these triggers provide data for comprehensive statistics of warning delays. The majority of warning delays range from 9 to 16 s (Figure 3). The older generation Refteks in the Anza network use both microwave and internet based telemetry. While older generations of CISN dataloggers (Q4120 and Q730) transmit 3 s of demultiplexed miniSEED waveform packets, the new Q330 dataloggers transmit data packets of 1 s multiplexed waveform data. We have developed the capability of capturing and analyzing these 1 s packets and therewith were able to reduce the warning delays by several seconds.

4. Discussion and Conclusions

[18] Between January 2007 and September 2008, 58 local earthquakes with 3.0 \( \leq M_\text{w} \leq 5.4 \) were real-time processed by the EEW software at Caltech, including the July 29 2008 M_\text{w}5.4 Chino Hills earthquake as the largest event. The performance during these events demonstrates the real-time applicability of the \( t_{c-P_d} \) algorithm.

[19] A scatter in the real-time predicted values of M_\text{w} for both on- and off-line processed earthquakes (Figure 1a) is unavoidable, because \( t_{c} \) is affected by many factors, including source (radiation patterns, directivity, and stress drop), propagation and site effects. Poor signal-to-noise ratios pose an additional problem for small magnitude events (M_\text{w} < 4.5) and events at large epicentral distances (\( \Delta > 150 \) km). The scatter can be reduced by the application of station correc-
The delays include times required for the delivery mainly depends on the type of datalogger and station equipment (in particular types of dataloggers, see legend), the telemetry, 3 s required by the τc-Pd algorithm, and the centralized waveform processing at Caltech. The histogram shows triggers obtained between Jan. and Sept. 2008.

Figure 3. Warning delays between P-wave arrivals at the CISN stations and the reporting of Mw and PGV values at Caltech. The delays include times required for the processing and communication of the waveform data, i.e. delays caused by station equipment (in particular types of dataloggers, see legend), the telemetry, 3 s required by the τc-Pd algorithm, and the centralized waveform processing at Caltech. The histogram shows triggers obtained between Jan. and Sept. 2008.

<table>
<thead>
<tr>
<th>τc</th>
<th>Pd</th>
<th>Mw</th>
<th>PGV</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>τc</th>
<th>Pd</th>
<th>Mw</th>
<th>PGV</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
</tr>
</tbody>
</table>

Acknowledgments. We thank two anonymous reviewers for their positive comments on an earlier version of our manuscript. This work is funded through contract 06HQAG0149 from USGS/ANSS to the California Institute of Technology (Caltech). The Southern California Seismic Network (SCSN) and the Southern California Earthquake Data Center (SCEDC) are funded through contracts with USGS/ANSS, the California Office of Emergency Services (OES), and the Southern California Earthquake Center (SCEC). This contribution 10010 of the Seismological Laboratory, Geophysical and Planetary Sciences at Caltech.

References


M. Böse, E. Hauksson, T. H. Heaton, H. Kanamori, and K. Solanki, Seismological Laboratory, California Institute of Technology, 1200 East California Boulevard, Mail Code 252-21, Pasadena, CA 91125, USA. (mbose@caltech.edu)