

CISN Earthquake Early Warning: Testing of Seismological Algorithms

Final Report: August 2006 – July 2009

Prepared by the
CISN Earthquake Early Warning Development Group

October 30, 2009

Cooperative agreements:

UC Berkeley: 06HQAG0147

Caltech: 06HQAG0149

USC/SCEC: 06HQAG0148

Content	page
Executive Summary	2
The T_c -Pd (On-site) Warning Algorithm	10
Development, testing and performance of ElarmS in California and Japan	24
Summary of Virtual Seismologist status and performance as of October 2009	47
Research on Finite Faults: Early Warning for Large Magnitude Earthquakes	60
CISN Testing Center Development and operations	69
Project Publications and Conference Abstracts	81

For questions contact Richard Allen, UC Berkeley (rallen@berkeley.edu)

CISN Earthquake Early Warning Testing – Final Report – November 2009

Executive summary

1. Introduction

In July 2009, the California Integrated Seismic Network (CISN) completed a three-year research and development effort to evaluate the performance and feasibility of earthquake early warning (EEW) in California. EEW systems are algorithms that detect the initial P-waves from an earthquake, rapidly estimate the hazard potential, and send out alert messages to the surrounding population centers. EEW systems offer the potential for a few seconds to a few tens of seconds warning before hazardous ground shaking occurs: enough time for automated safety measures such as shutdown of utilities to minimize fire hazard from broken fuel lines, or slowing of trains to reduce the risk of derailment. In some cases warnings can be issued to the general public, allowing people to get to a safe location, such as under a sturdy table, before ground shaking occurs at their location. Active EEW systems have already been deployed in Mexico, Japan, Taiwan, Turkey, and Romania; the systems in Japan and Mexico provide publically available warnings. Realtime testing of EEW systems is underway in Italy, Switzerland, China and Korea, in addition to California.

During the CISN study three independent EEW algorithms were tested: OnSite, Virtual Seismologist, and ElarmS. OnSite (Kanamori, 2005; Wu *et al.*, 2007; Böse *et al.*, 2009a; Böse *et al.*, 2009b) produces earthquake information from P-wave parameters observed at a single station, providing a very fast warning. Virtual Seismologist (VS)(Cua and Heaton, 2007; Cua *et al.*, 2009) and ElarmS (Allen and Kanamori, 2003; Wurman *et al.*, 2007; Allen *et al.*, 2009) combine information from several stations to produce warnings which are potentially slower than those from OnSite, but are potentially more accurate. Results from each algorithm were reported to a central database run by the Southern California Earthquake Center (SCEC) for analysis and comparison. Detailed description of the research efforts and results are contained in the following sections of this Final Project Report. Here we summarize the results of the testing.

2. Results

During the three-year study all three algorithms successfully detected many earthquakes and in some cases predicted ground shaking before it occurred. Each algorithm, originally designed for a specific region of the state, was expanded to process data from throughout California, to run continuously in realtime, and to provide realtime alert messages. The statewide testing made use of 382 stations with a total of 585 broadband and strong motion instruments. The realtime data was provided by the Anza Network (AZ), the Berkeley Digital Seismic Network (BK), the Southern

California Seismic Network (CI), the Northern California Seismic Network (NC), and National Strong Motion Network (NP) all of which are cooperating partners in the California Integrated Seismic Network (CISN). The data streamed to and was processed at three network facilities at Caltech/Pasadena, UC Berkeley, and USGS Menlo Park (Table 1 and Figure 1).

Table 1. CISN channels used for EEW testing in California.

	Caltech, Pasadena		UC Berkeley	USGS Menlo Park	
	CI	AZ	BK	NC	NP
Stations	186	14	33	87	62
HH channels	158	14	29	20	0
HN/HL channels	182	2	31	87	62

HH: broadband channels; HN/HL: strong motion channels. AZ: Anza Network, BK: Berkeley Digital Seismic Network, CI: Southern California Seismic Network, NC: Northern California Seismic Network, NP: National Strong Motion Network

Of particular note is the performance of the algorithms during the two largest earthquakes during the testing period. The October 30, 2007 M_w 5.4 Alum Rock earthquake in the San Francisco Bay Area was initially detected simultaneously by two stations close to the epicenter and the initial location and magnitude estimate of M5.2 was available 5 seconds before the peak shaking in San Francisco. This made it possible to generate an accurate map of predicted ground shaking across the region before shaking was felt in San Francisco, Oakland and other cities in the north bay. The July 28, 2008 M_w 5.4 Chino Hills earthquake in the Los Angeles region was detected by a total of 60 CISN stations and magnitude estimates generated by individual station detections ranged from M4.4 to M6.5 with a median of M5.6. The first estimate was available 10 sec after the origin time corresponding to a “no-warning” blind zone of ~35km. Recent improvements to the processing system mean that notifications would now be available within 5-6 sec of the origin time.

2.1 Speed

As shown in table 2 the delay of the Onsite method for $M > 3.5$ earthquakes ranges from 9-16 sec from the P-wave trigger until a magnitude and PGV estimate at the site are available. This has recently been reduced to 4-11 sec by software improvements. The median time of the first VS estimate for $M > 3.0$ is 21 sec after the origin time in southern California and 31 sec in northern California. ElarmS has a variable alert criteria requiring different numbers of station to trigger as a function of location

before an alert is issued. The median delay statewide is currently 32 sec but is reduced in regions with dense instrumentation such as the Bay Area where the median delay is 27 sec.

Table 2. Summary of algorithm performance. Error distributions are represented as mean ± standard deviation.

	Onsite	Virtual Seismologist	ElarmS
Speed	Median time of the magnitude estimate in southern California is 11 sec, and in northern California is 13 sec.	Median time of 1 st estimate is 21 sec after the origin time in southern CA and 31 sec in northern CA.	Median time of 1 st alert is 32 sec after the origin time for all of CA, 27 sec for the San Francisco Bay Area.
Magnitude accuracy	For ~140 CA earthquakes $3.5 \leq M \leq 5.4$ the error is 0.0 ± 0.5 magnitude units. The error in the Intensity estimates derived from PGV estimates at the triggered stations is 0.0 ± 0.7	For $M \geq 3.0$ events the magnitude error of the first VS estimate is 0.0 ± 0.23 in southern CA and 0.12 ± 0.33 in northern CA.	For $M \geq 3.0$ earthquakes reaching the alert criteria, the magnitude error across CA is -0.2 ± 0.7 . A study of 84 earthquakes M4.0 to M8.0 in Japan (to look at the accuracy for large earthquakes) showed a magnitude error of 0.0 ± 0.4
Detections, false and missed earthquakes	~140 $M > 3.5$ earthquakes detected across CA over 3 years.	Looking at $M \geq 3.0$ in southern California during the last 6 months of the project 95 earthquakes were correctly detected and 15 false events.	Looking at $M \geq 3.0$ earthquakes in northern CA after the alert criteria was introduced (Aug-Oct 2009) there were 19 accurate detections, 3 false and 4 missed. In the San Francisco Bay Area where station spacing is 20km there were 8 detections with 0 false and 0 missed.

These delays are due mostly to data packetization at the station, telemetry to the network center, and processing. The current median latency to receive waveform data at the network processing centers across the state is 5.23 sec but ranges from less than 1 to ~15 sec. The wide range is due to the different dataloggers currently in use. The K2 and Q330 dataloggers are the fastest and the early warning group anticipates a significant reduction in the latencies with the planned upgrade of older dataloggers to Q330s.

Both ElarmS and VS currently require 4 stations to be associated with an earthquake before issuing an alert. The median delays for these alerts range from 21 to 32 sec. Some of this additional delay is waiting for additional stations to trigger, but most is due to processing time. For example it currently takes ElarmS 5 sec to reduce waveforms to parameters, which is just the first step in the processing sequence and there is no reason why it should not take a fraction of a second. Many lessons have been learned about the best way to implement the EEW algorithms during this testing and the implementation of improved code frameworks to minimize these processing delays is already underway.

2.2 Accuracy

The uncertainty in the magnitude estimates was assessed by comparing the algorithm magnitude estimates to the CISN catalog magnitude. For the Onsite method the mean error is 0.0 and the standard deviation of the error is 0.5 magnitude units, i.e. an error distribution of 0.0 ± 0.5 . Onsite also provided at-station PGV predicted values which when converted to intensity have an error distribution of 0.0 ± 0.7 . The first VS estimate has a magnitude error of 0.0 ± 0.2 in southern California and 0.1 ± 0.3 in northern California. Finally the ElarmS magnitude estimates once the alert criteria is reached is -0.2 ± 0.7 for all events in California. The ElarmS methodology was also applied offline to a set of 84 earthquakes in Japan with magnitudes ranging from 4.0 to 8.0, 43 of the events had $M \geq 6.0$. The uncertainty in the magnitude estimates was 0.0 ± 0.4 for all events, 0.0 ± 0.5 for $M \geq 6.0$ events and -0.2 ± 0.5 for $M \geq 7.0$ events indicating a slight saturation effect in the magnitude estimates for the largest earthquakes.

2.3 False alarms and missed events

Rates of false alarms and missed events were directly dependent on station density. In regions where the stations were approximately 20 km apart, the EEW algorithms successfully detected 95% to 100% of events. In regions where the stations were approximately 20-100 km apart, performance was reduced and the EEW algorithms detected 85% of the events, missing 15% and producing false alarms at a rate of about 15% of the total number of real events. In regions where station spacing is greater than 100km, the EEW algorithms performed sporadically. For each earthquake in these

regions, there may or may not be a seismic station near enough to the epicenter to detect the P-wave at all. If there are nearby stations, there may not be enough of them to satisfy the station criteria for the network-based algorithms (Virtual Seismologist and ElarmS). In other cases the station criteria were eventually met, but the alerts were too delayed to be considered practical for a production EEW system. These results show that it is highly advisable to have a station density of 20 km (or higher) to provide useful, reliable, realtime warnings for earthquakes.

To improve EEW performance in the highest-hazard regions of California, the existing seismic networks would have to be augmented by approximately 100 new seismic stations. Figure 1 shows a map of the current station coverage in California plus the proposed 100 new stations. The background colors represent the ground shaking hazard, at 10% exceedence in the next 50 years. The red, blue and green squares are current EEW-capable stations in California. The only regions which meet the 20 km station spacing criteria are the San Francisco Bay Area, Los Angeles, and a portion of the southern section of the San Andreas Fault. Upgrading the seismic networks with 100 new stations (white squares) can increase the density along the high hazard zones to a 20 km station spacing.

The EEW algorithms use both strong motion acceleration and weak motion velocity instruments. Most of the UC Berkeley stations (blue on Figure 1) and Caltech/USGS Pasadena stations (red) have both instrumentation types. Most of the USGS Menlo Park stations (green) are acceleration only, with the exception of the northern-most stations. The best system performance would be obtained by putting both velocity and acceleration instruments at all 100 new sites. While strong motion acceleration instrumentation is a requirement at all sites, reducing the number of stations that also have velocity instruments would not provide adequate performance. Standard National Strong Motion Program or California Geological Survey strong motion stations tend to be located at noisy sites, and thus provide data that are hard to process, and would sometimes lead to false alerts for EEW algorithms. We need the additional 100 stations to be of reasonable seismological quality to provide warnings both for moderate-sized and large quakes. This means that the installation needs to meet ANSS specifications for broadband sites. The added benefit would be that these data could be used for continued algorithm development and provide EEW for moderate-sized quakes, which has no practical value, but has a significant public relations value.

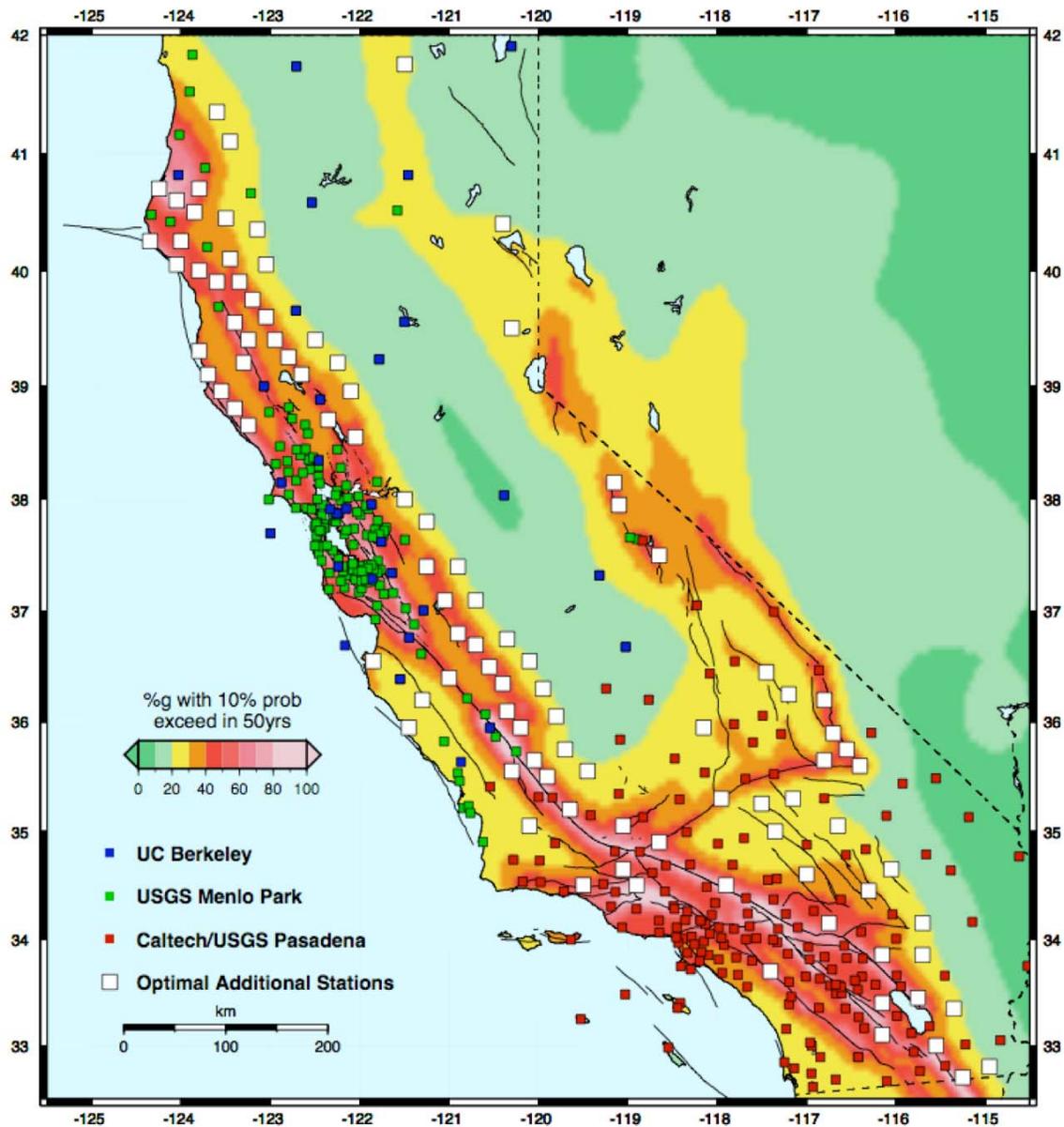


Figure 1. Map of current realtime seismic stations used for the EEW testing (green, blue and red squares) overlaid on a seismic hazard map. The white squares are the ~100 additional seismic stations needed to have a 20km station spacing in the high hazard zones (red and orange) and 40km station spacing in the adjacent zones (yellow) regions along the San Andreas Fault System.

2.4 Big Earthquakes

Of primary concern is EEW performance for large earthquakes. As large earthquakes did not occur in California during the three-year testing period, one of the algorithms, ElarmS, was tested offline with a dataset of large magnitude events from Japan, recorded by Japan's Kyoshin-NET (K-NET) strong-

motion seismic network. K-NET has a dense station coverage of 20-25km spacing throughout the country, and the ElarmS algorithm successfully detected all 83 events in the dataset. For magnitudes 4 up to 6.9, the mean magnitude error was 0.0, with a standard deviation of 0.5 magnitude units. For the seven events with magnitudes greater than 7, the mean magnitude error was -0.2, with a standard deviation of 0.5. Therefore we must ask whether this magnitude saturation is a serious impediment to the success of an EEW system. The saturation effect is not seen until M7 events. One approach is therefore to issue an alert to an entire metropolitan region when a $M \geq 6.5$ earthquake is detected. As everyone in the region would feel the shaking this would not be perceived as a false alarm. At the same time more research and development into techniques to detect and map large magnitude earthquakes would be advisable.

Currently all three EEW algorithms tested in California, and all EEW algorithms implemented worldwide, treat all earthquakes as a point source. $M \leq 7.0$ earthquakes are probably adequately described by a point source, but the extended ruptures of larger earthquakes mean that detection of the fault sections rupturing would enhance EEW systems. The CISN EEW testing team are in the process of developing a finite fault capability, to recognize ongoing fault rupture and adjust hazard predictions accordingly. Fault rupture dimensions are estimated by realtime classification of near-source and far-source records, identification of envelopes of acceleration, usage of slip amplitudes and prior knowledge of smoothness factors for regional faults. While these new methodologies seem quite promising for EEW systems, they only work well when there is an adequate distribution of near-source stations. As with the false alarm rates, 20km spacing is expected to significantly improve performance.

3. Next Steps

In August 2009 a second three-year study was initiated, to integrate the three test algorithms into a single prototype EEW system and provide realtime warning to a small group of test users by the end of the study in summer 2012. The CISN is currently identifying a small group of about 10 test users who will soon start to receive alerts from the new prototype system, called the CISN ShakeAlert System.

In the next two years ARRA stimulus funding will be used to upgrade many of the older, slower data loggers throughout the CISN. These older data loggers currently have latencies (time till waveform data is available at the processing centers) of 4 to 12 seconds. They will be upgraded to Q330 data loggers with a latency of 1-2 sec once both the planned hardware and software upgrades are implemented. This will result in a reduction in the latency at these sites of between 2 and 11

seconds. The median latency of all data streams into the EEW test systems is currently 5.2 sec and we anticipate a reduction of this statewide median to 2-3 sec.

The testing clearly illustrated the benefits of dense station spacing to reduce the potential of false and missed alerts and also to improve the speed of alert delivery. The station spacing currently used for EEW in Japan is ~20 km. Performance of EEW in California is shown to be best in the San Francisco Bay Area and the great Los Angeles area where stations spacing is ~20 km. Enhancing the networks to provide a similar 20 km station spacing throughout the earthquake source region of the San Andreas Fault system would require an additional ~100 seismic stations (Figure 1).

References

- Allen, R. M., and H. Kanamori (2003), The potential for earthquake early warning in southern California, *Science*, **300**, 786-789.
- Allen, R. M., H. Brown, M. Hellweg, O. Khainovski, P. Lombard, and D. Neuhauser (2009), Real-time earthquake detection and hazard assessment by ElarmS across California, *Geophys. Res. Lett.*, **36**, L00B08 doi:10.1029/2008GL036766.
- Böse, M., E. Hauksson, K. Solanki, H. Kanamori, and T. H. Heaton (2009a), Real-time testing of the on-site warning algorithm in southern California and its performance during the July 29 2008 Mw5.4 Chino Hills earthquake, *Geophys. Res. Lett.*, L00B03, doi:10.1029/2008GL036366.
- Böse, M., E. Hauksson, K. Solanki, H. Kanamori, Y.-M. Wu, and T. Heaton (2009b), A new trigger criterion for improved real-time performance of on-site early warning in southern California, *Bull. Seismol. Soc. Am.*, **99**, 897-905.
- Cua, G., and T. Heaton (2007), The Virtual Seismologist (VS) method: A Bayesian approach to earthquake early warning, in *Earthquake early Warning Systems*, edited by P. Gasparini, G. Manfredi and J. Zschau, pp. 97-132, Springer.
- Cua, G., M. Fischer, T. Heaton, S. Wiemer, and D. Giardini (2009), Real-time performance of the Virtual Seismologist method in southern California, *Seismo. Res. Lett.*, **80**, 740-748 doi: 710.1785/gssrl.1780.1785.1740.
- Kanamori, H. (2005), Real-time seismology and earthquake damage mitigation, *Annu. Rev. Earth Planet. Sci.*, **33**, 195-214.
- Wu, Y.-M., H. Kanamori, R. M. Allen, and E. Hauksson (2007), Experiment using the tau-c and Pd method for earthquake early warning in southern California, *Geophys. J. Int.*, **170**, 711-717.
- Wurman, G., R. M. Allen, and P. Lombard (2007), Toward earthquake early warning in northern California, *J. geophys. Res.*, **112**, B08311, doi:08310.01029/02006JB004830.

The τ_c - P_d (On-site) Warning Algorithm

Maren Böse, Tom Heaton, Egill Hauksson, Kalpesh Solanki, California Institute of Technology

Summary

Over the past three years, the τ_c - P_d algorithm has detected ~140 local earthquakes in California and Mexico Baja with moment magnitudes of $3.5 \leq M_w \leq 5.4$. Combined with newly derived station corrections the algorithm rapidly determined moment magnitudes M_w and Modified Mercalli Intensities (derived from peak ground velocity, PGV) with uncertainties of ± 0.5 and ± 0.7 units, respectively. Mostly reporting delays range from 9 sec to 16 sec (including 3 sec waveform data required by the current algorithm; Böse *et al.*, 2009a), but were recently reduced to 4 sec to 11 sec. The largest event, the July 29 2008 M_w 5.4 Chino Hills earthquake, triggered a total of 60 CISN stations at epicentral distances up to 250 km. Magnitude predictions at these stations ranged from M_w 4.4 to M_w 6.5 with a median of M_w 5.6. The first estimate (M_w 6.1), available 10 sec after the origin time, would have provided 6 sec notification time at Los Angeles City Hall, had an alerting system been operational. Recent improvements to the algorithm have significantly increased the processing speed. Were another earthquake to occur near the Chino Hills location, notifications would presumably be available within 5-6 sec after the origin time.

1. Brief Description of the τ_c - P_d Algorithm

One of the major elements of EEW is the rapid and reliable determination of earthquake magnitudes. To determine the size of an earthquake, it is important to find out whether the earthquake rupture has stopped or keeps growing. This is generally reflected in the period of the initial ground motion. Kanamori (2005) extended the method of Nakamura (1988) and Allen and Kanamori (2003) to determine a period parameter τ_c from the initial few seconds of P waves. τ_c is defined as $\tau_c = 2\pi/\sqrt{r}$ where $r = \left[\int_0^{\tau_0} \dot{u}^2(t) dt \right] / \left[\int_0^{\tau_0} u^2(t) dt \right]$, $u(t)$ is the ground-motion displacement, and τ_0 is the duration of the record used. In a series of studies (Wu and Kanamori, 2005a, 2005b, 2008a, 2008b; Wu *et al.*, 2006; 2007) τ_0 is set at 3 seconds. Wu *et al.* (2007) systematically studied the off-line records from earthquakes in southern California to determine a $\log_{10}(\tau_c)$ - M_w scaling relation for EEW.

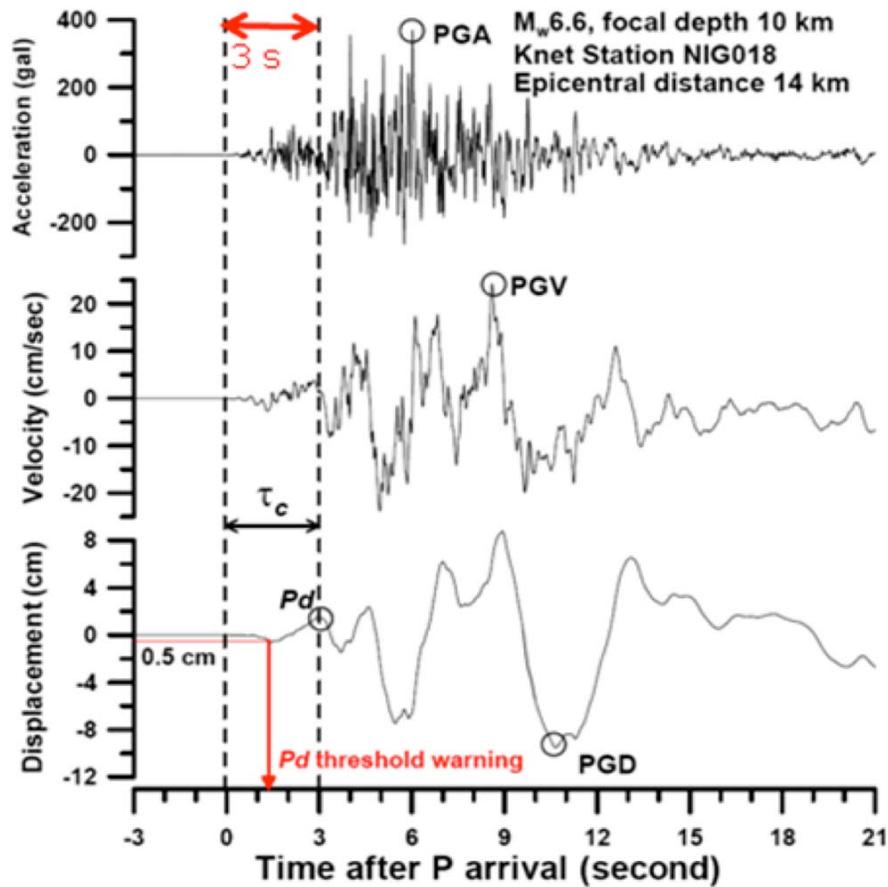


Figure 1 Parameters used by the τ_c - P_d algorithm (Wu and Kanamori, 2008).

Another important element in EEW is the estimation of the strength of S wave shaking. Wu and Kanamori (2005b) showed that the maximum amplitude of the high-pass filtered vertical displacement during the initial 3 seconds of the P wave, P_d , can be used to estimate the peak ground velocity (PGV) at the same site (Figure 1).

2. Developments and Real-Time Performance of the τ_c - P_d Algorithm

Implementation and initial testing of the τ_c - P_d algorithm in southern California started in 2007. A state-wide implementation was achieved in early 2009 including broadband (HH) and strong motion (HN) instrumentation. An overview about the most important developments of the τ_c - P_d algorithm during the past three years is given in Table 1.

Table 1. Major Developments of the τ_c - P_d Algorithm at Caltech.

Time	Step/Event	CISN Channels	Observation	Solution/Action
2007-2008	Implementation and initial testing of the τ_c - P_d on-site algorithm in southern CA	CI, AZ southern CA 172 HH channels	<ul style="list-style-type: none"> - Many false triggers - Scattering in M_w estimates for small earthquakes 	<ul style="list-style-type: none"> - τ_c-P_d trigger criterion - Testing server
July 29, 2008	Chino Hills M_w5.4		<ul style="list-style-type: none"> - Data latencies ... depend on datalogger ... are variable with time - τ_c / P_d site-dependent 	<ul style="list-style-type: none"> - Native Q330 wave packets processing - Enhanced log-files - Site corrections
2008-2009	State-wide implementation, including strong motion sensors	CI, AZ BK, NC, NP state-wide 221 HH & 364 HN at 382 locations	<ul style="list-style-type: none"> - Reduced proc. speed - Similar results HH/HN - More false triggers at HN 	<ul style="list-style-type: none"> - Optimized source code (compact instead of modular design) - Higher thresholds for HN/coupling with HH (in progress)
Since Nov. 2008	Deployment of SLATE field processors and on-site processing software (in progress)			
April 2009	Capability to “replay” off-line data as real-time data stream (e.g., Northridge, Hector Mine, Landers, simulated waveform data /scenarios, noise records) for improved error assessment			
June/July 2009	Localization capability Two-station approach Grey/black lists, automatic notification in the case of mass-recentering			

2.1 Initial Implementation and Testing Phase

For the real-time testing of the τ_c - P_d method within the CISN, we have implemented the algorithm in an UNIX environment. The processing steps are as follows: (1) retrieve velocity/acceleration data from the CISN; (2) remove baseline and apply gain correction; (3) convert to displacement by recursive integration; apply high-pass Butterworth filter (> 0.075 Hz); and (4) calculate τ_c and P_d from the initial 3 seconds of waveform data. For the picking of the seismic P phase we make use of a modification of an algorithm proposed by Allen (1978) in a combination with a simple P/S wave discriminator which is based on the ratios of horizontal to vertical ground motions.

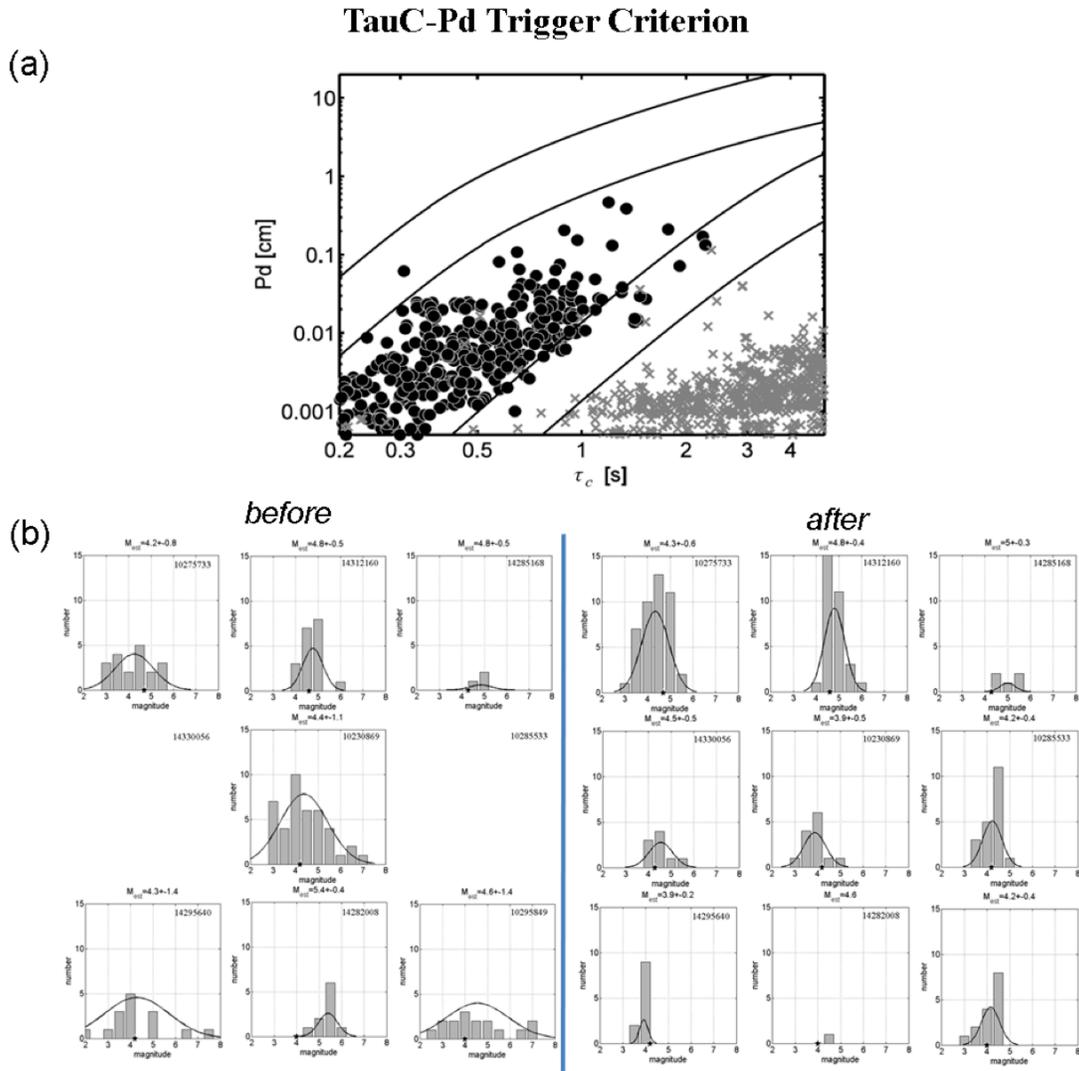


Figure 2. The $\tau_c - P_d$ trigger criterion (Böse et al., 2009a).

During the initial testing period, we observed a high number of false triggers, i.e. triggers that cannot be associated with any local earthquakes, and a significant scattering in estimated magnitudes for small- to moderate-sized earthquakes due to poor S/N ratios. To improve the real-time performance of the $\tau_c - P_d$ algorithm, in particular for small and moderate-sized events, we developed the so-called $\tau_c - P_d$ trigger criterion. This criterion is based on τ_c -dependent and therewith magnitude-dependent P_d thresholds: for a local earthquake with period τ_c in a rupture-to-site distance r , $r_{min} \leq r \leq r_{max}$, we expect $P_{d,min} \leq P_d \leq P_{d,max}$ (Figure 2a). Böse *et al.* (2009a) 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. The $\tau_c - P_d$ trigger criterion removed 97% of previous false triggers at the broadband stations and led to a significant reduction of the scatter in magnitude estimates for small earthquakes (Figure 2b).

2.2 Observations During the 2008 Mw5.4 Chino Hills Earthquake

The July 29th, 2008 M_w 5.4 Chino Hills mainshock in the eastern 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 M_w 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 M_w and PGV values by the τ_c - P_d algorithm: 60 during the M_w 5.4 mainshock and between 8 and 15 during the three largest aftershocks (Böse *et al.*, 2009b).

The first magnitude prediction of the M_w 5.4 Chino Hills mainshock was available 10 s after O.T. (M_w 5.6 at station CI.PSR). Subsequent (independent) estimates based on data from other stations ranged from M_w 4.4 to M_w 6.5 with a median value of M_w 5.6. 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 3).

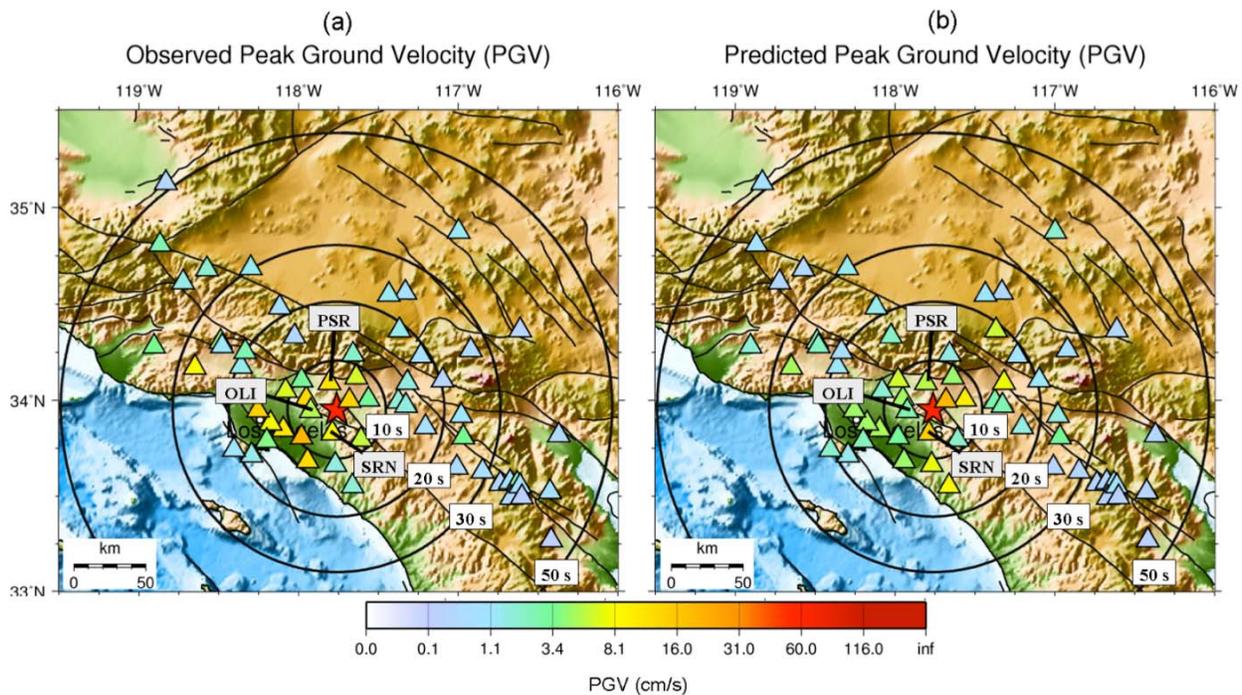


Figure 3. Distribution of (a) observed and (b) predicted values of peak ground velocity (PGV) at 60 CISN stations triggered by the EEW software during the July 29 2008 M_w 5.4 Chino Hills earthquake. Neglecting telemetry and processing delays, each single estimate in (b) could have been made available within 3 seconds after P-wave arrival. Circles show S-wave arrivals at 10 sec, 20 sec, 30 sec and 50 sec after origin time (Böse *et al.*, 2009b).

Neglecting the telemetry and processing delays of the system then, each single estimate in Figure 3b could have been made available within 3 seconds after the P-wave arrival (This is the time required by the τ_c - P_d algorithm.). The entire map in Figure 3b 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).

The 2008 Chino Hills earthquake (and other previous events) demonstrated that τ_c and P_d are site-dependent parameters (see e.g. Figure 3b). At CISN stations with a reasonable number of observations during past earthquakes (at least during 3 events with $M_w \geq 3.0$ at close epicentral distances) we determined station corrections from the median value of observed residuals. The station corrections reduced the uncertainties in predicted moment magnitudes M_w and Modified Mercalli Intensities (derived from peak ground velocity, PGV) to ± 0.5 and ± 0.7 units, respectively (Figure 4). Note that the scattering still is significant and further improvements and error assessment needs to be done (see Conclusions and Outlook).

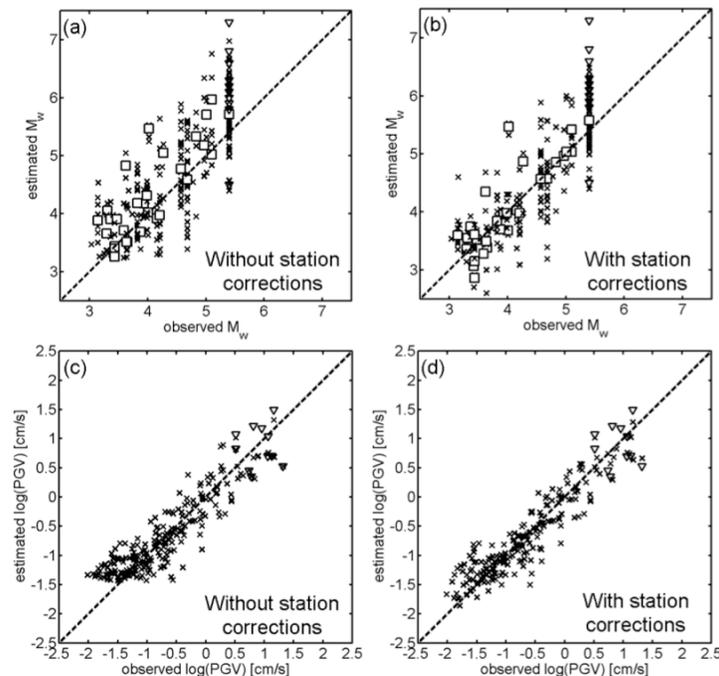


Figure 4. Estimated vs. observed source and ground motion parameters of 58 local earthquakes ($3.0 \leq M_w \leq 5.4$) in southern California and Baja between 2007 and 2008: (a) and (b) moment magnitudes M_w ; (c) and (d) peak ground velocities (PGV) (Böse *et al.*, 2009b).

2.3 State-wide Implementation of the τ_c - P_d Algorithm

In early 2009 we deployed our on-site algorithm testing system at two additional processing sites in northern California (UC Berkeley and USGS Menlo Park). This enables us to process earthquake data state-wide using multiple computer systems, databases and notification systems.

The design of the EEW test system is distributed (Figure 5). We are using UDP (User Datagram Protocol) based communication between software agents that detect the P-wave triggers and software agents that forward the message to different warning and analysis applications. The division of the tasks allows a simple deployment into the network.

The software agents that do real-time waveform processing are closer to the real-time streams. The EEW Message Server receives UDP messages (triggers) from the real-time processing agents deployed in different networks and forward the messages to the CISN Messaging System (Message Oriented Middleware) for the distribution to other applications. It also stores the triggers into the database for later analysis.

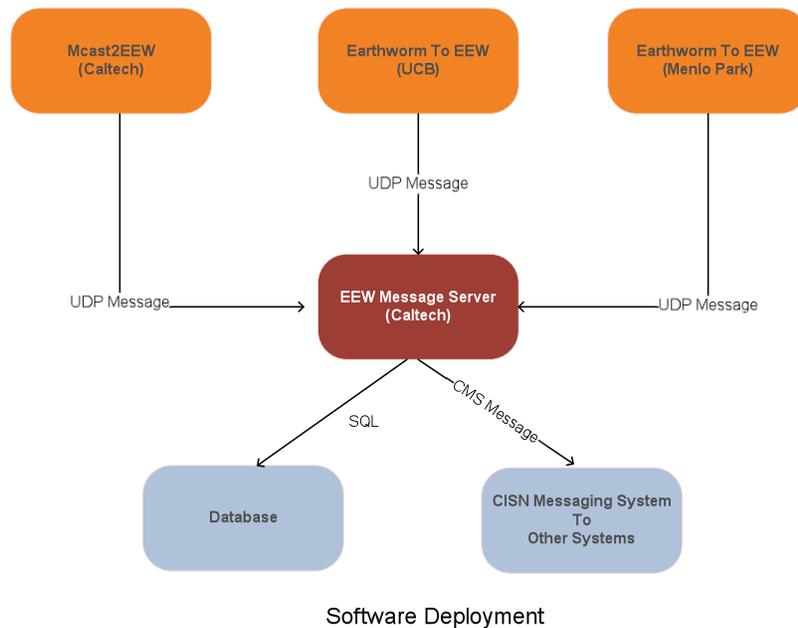


Figure 5. State-wide implementation of the τ_c - P_d algorithm (Solanki *et al.*, 2009).

Since January 2009 we process both broadband (HH) and strong motion (HN/HL) channels of the CISN network. Usually there is a good agreement between the τ_c and P_d observations at both types of channels (Figure 6).

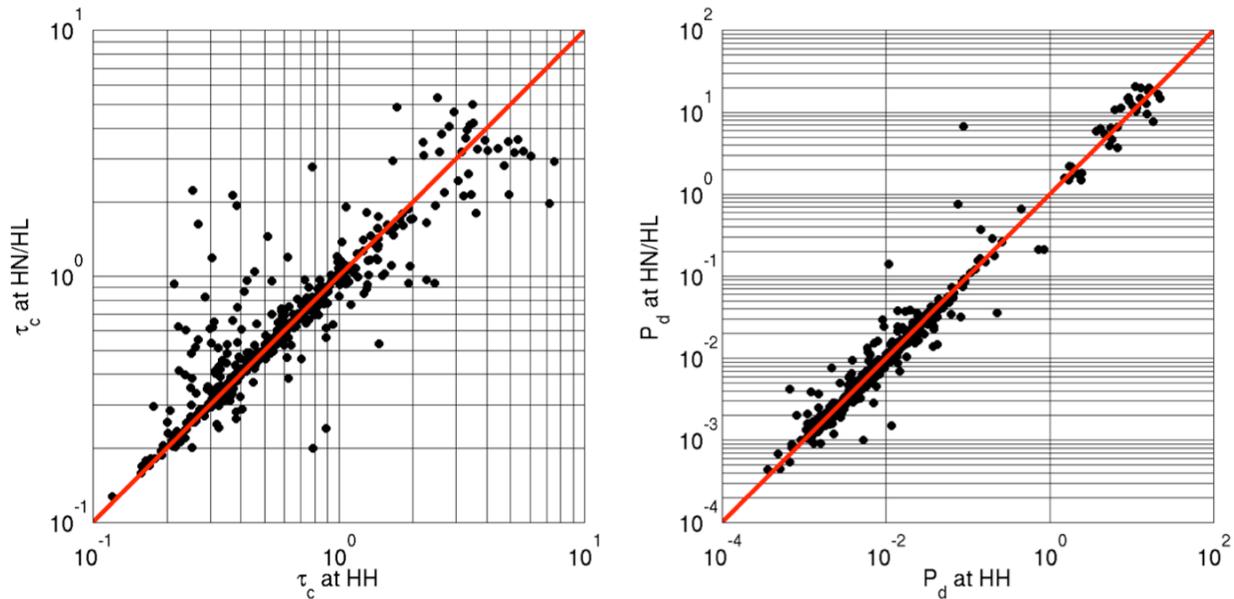


Figure 6. Comparison of τ_c and P_d observations at HH and HN/HL channels.

2.4 Further Developments

Furthermore, we have added the following improvements/capabilities to our code. Note, that many of these capabilities have been developed only recently and require further testing and enhancements in the future.

- **Clipping criterion:** For very close and strong earthquakes, amplitudes at HH channels might be clipped within the first 3 seconds after P-wave detection. We have implemented a routine to recognize, if the amplitudes at a HH channel exceed a sensor-specific clipping level. In this case, the system will add a specific flag to the estimated M_w and PGV values. Reports including flags are not sent to the SCEC webpage, but are stored in our database for later analyses.
- **Two-station approach:** In order to reduce the number of false triggers (that mostly are caused by low trigger thresholds for algorithm testing during small to moderate earthquakes), we have developed a two-station approach: a trigger must be confirmed by at least one further station

within a certain time window before a report is sent. As the τ_c - P_d algorithm requires 3 seconds of waveform data, there is a high probability that during these 3 seconds at least one further station will also detect the P wave. We expect no reductions of warning time compared to the single-station approach, but a significant stabilization of the system.

- Grey List: The “Grey List” is a list to which we can add noisy HN/HL channels. Triggers at these channels are processed only if the P wave is detected also at a co-located HH sensor.
- Black List: The majority of false triggers that we observed in the past were caused either (a) by temporarily noisy sites (e.g., caused by constructions close to a sensor or malfunctioning station equipment), or (b) by the mass-recentering by the network maintenance crew. A malfunctioning sensor/datalogger can produce up to hundreds or thousands of false triggers during one single day leading to bad performance statistics. Also mass-recentering at broadband sensors by the maintenance crew is a problem for a single-sensor based algorithm: this process causes signals with very long periods and large amplitudes (typical for $M > 7.0$ events) that the system will recognize as “false alerts”. Both (a) and (b) are not expected to cause major problems in an operational EEW system, because we usually know when these problems occur. In order to improve the performance of the τ_c - P_d algorithm, we recently have developed a “Black List”, to which we can manually add noisy channels that will temporarily not being processed. Once the malfunctioning sensor/datalogger is replaced, we can remove this station from the Black List. The Black List can be up-dated without restarting the entire system. In addition, we have developed a script that can be executed by the maintenance crew before starting the mass-recentering process. All triggers at the corresponding sensor will be ignored for the next 5 minutes.
- P/S-wave discrimination: For stations at close epicentral distances the first 3 seconds of waveform data as currently is used by the τ_c - P_d algorithm, may be contaminated by the seismic S wave. In this case, the P_d amplitudes are too large and the prediction of PGV incorrect. We therefore have developed an algorithm based on the ratios of horizontal and vertical ground motions to recognize the arrival of the seismic S wave so that the time window for the determination of P_d can be automatically decreased. The P/S wave discrimination is fairly sensitive and we will put more efforts into its optimization (In the past we observed sometimes that P waves were incorrectly classified as S waves and thus triggers ignored).

- **Localization capability:** We have implemented two methods to localize earthquakes based on (a) the particle motions at a single station; (b) mapping between two stations corresponding to time delays in P-wave detection. Usually, the localization results are fairly crude, but still give a reasonable approximation of the earthquake source region. The localization procedure will need further calibration in the next months.
- **On-site processing:** Recently, we have started to implement the τ_c - P_d algorithm on Kinometrics Slate Field Processors at 20 CISN stations along the southern San Andreas Fault (Figure 7). This way we want to increase the robustness of the system and check for the possible reduction of data latencies by transmitting the τ_c and P_d parameters only rather than waveform data.



Figure 7. Installation of Slate Field Processors along the southern San Andreas Fault.

- **Replay capability:** We are developing a tool based on the Earthworm Tankplayer to replay the records of past earthquakes from SCEDC/NCEDC and synthetic waveforms (e.g., of the Great Southern California ShakeOut) as simulated real-time data streams. The streams are processed by the same code that also is applied to the real-time CISN waveform data. This way we can study the possible performance of the system and assess uncertainties in the future.
- During very large earthquakes, e.g. during the 2008 M_w 7.9 Wenchuan mainshock or the recent 2009 M_w 6.9 earthquake in the Gulf of California, surface waves triggered the τ_c - P_d algorithm at some tens of CISN stations. At some stations, the high-pass filtered amplitudes exceeded 1 cm at periods of 1.5 seconds and longer. The majority of these teleseismic events can be recognized and

removed by the τ_c-P_d trigger criterion (Böse et al., 2009a), by we need to refine the algorithm in the case of very large magnitudes in the future.

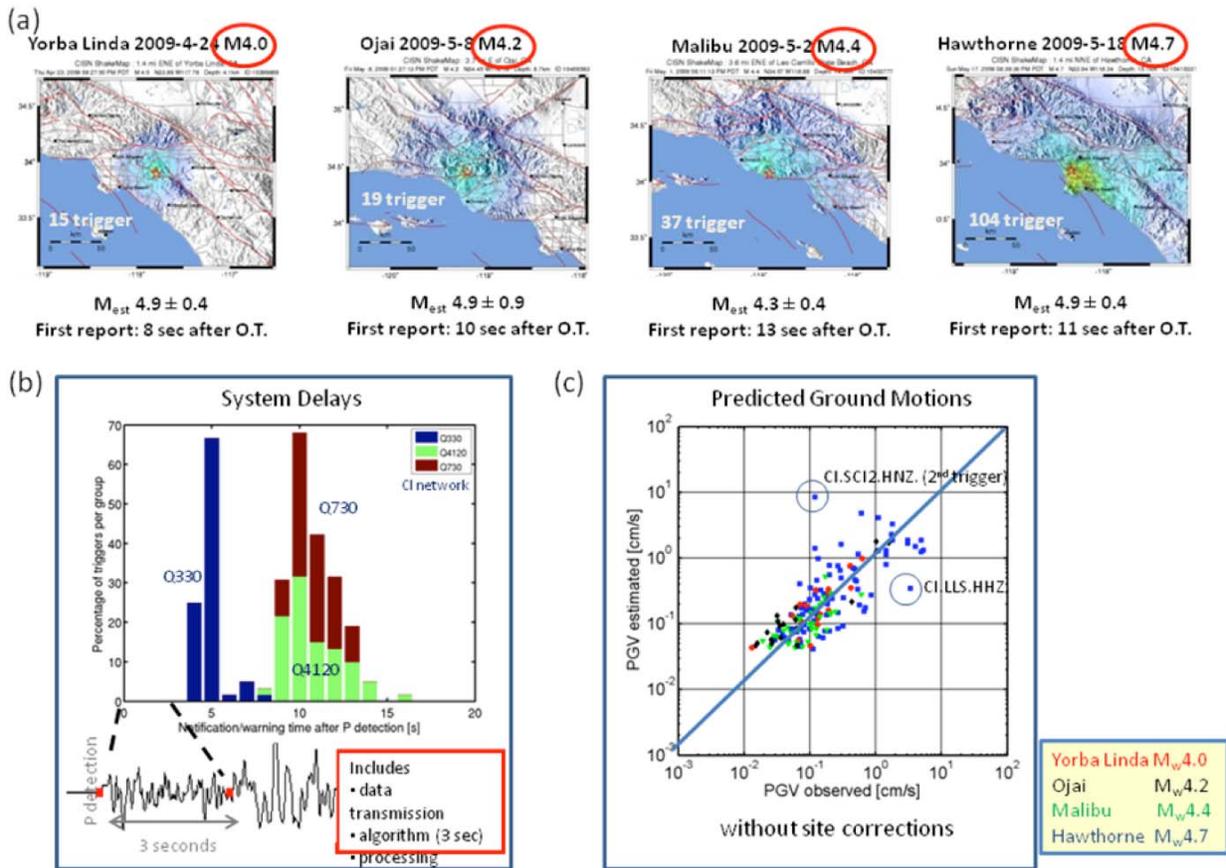


Figure 8. Performance of the τ_c-P_d algorithm during recent earthquakes.

3. Conclusions and Outlook

Over the past three years, the τ_c-P_d algorithm has detected ~140 local earthquakes in California and Mexico Baja with moment magnitudes of $3.5 \leq M_w \leq 5.4$. Combined with newly derived station corrections the algorithm rapidly determined moment magnitudes M_w and Modified Mercalli Intensities (derived from peak ground velocity, PGV) with uncertainties of ± 0.5 and ± 0.7 units, respectively. Mostly reporting delays range from 9 sec to 16 sec (including 3 sec waveform data required by the current algorithm; Böse *et al.*, 2009a), but were recently reduced to 4 sec to 11 sec.

Since based on single-sensor observations, the τ_c-P_d on-site warning algorithm usually provides faster estimates of M_w , PGV and epicenter locations than the regional warning approaches (ElarmS and

VS), in particular in regions with sparse instrumentation. However, the algorithm has (1) a higher probability of false triggers, and (2) there is a significant amount of scattering in the $\log_{10}(\tau_c)$ - M_w relation (see e.g. Figure 4). As discussed previously, the main causes for false triggers are temporarily increased levels of noise, e.g. caused by construction, malfunctioning stations, or mass-recentering. However, we do not think that this will be a severe problem in an operational EEW system.

By the introduction of the “Black List” (see atop) we expect a significant reduction of false triggers in the future. A further stabilization of the system will be achieved by the usage of the newly developed two-station approach for the confirmation of triggers. We do not expect a reduction of warning times compared to the single-station approach.

One of the open questions that we will address soon by the usage of the newly developed replay capability of past records (see atop), is the assessment of uncertainties in the prediction of source and ground motion parameters. Recently, we discovered systematic trends in the spatial distribution of the τ_c residuals in California: e.g. the magnitudes of earthquakes with epicenters in the Imperial Valley or LA Basin usually are overestimated by up to 1.5 magnitude units, whereas earthquakes in the Big Bend of the San Andreas Fault show the opposite picture. A likely explanation is the existence of spatial differences in the stress drops in California, as reported, e.g. by Shearer *et al.* (2006). We are currently working on a more detailed analysis of the uncertainties and try to develop tools for a possible reduction. The results will be included in our system.

Both Japanese EEW systems, JMA and UrEDAS, have a mechanism to issue warning at about 1 sec after a large acceleration is detected. Although no systematic study has been made, it would be probably better to use the displacement amplitude rather than the acceleration amplitude. Another method is to issue a warning as soon as the displacement exceeds a prescribed threshold (e.g., 0.5 cm) without waiting for the end of the 3 s window. From our experience with the Japan, Taiwan and southern California data, if P_d exceeds 0.5 cm, the PGV at the site most likely exceeds the damaging level, i.e., 20 cm/s. One possible approach for faster warning is to monitor P_d , and issue a warning as soon as it has exceeded 0.5 cm. As shown in Figure 9 for the 2007 Niigata Chuetsu-Oki earthquake, at the nearest stations, NIG018 ($\Delta=14$ km), the threshold value of $P_d=0.5$ cm was reached at 1.36 sec from the arrival of P wave. If we issue a warning at a threshold of $P_d \geq 0.5$ cm, a warning will be issued at 1.36 sec after the P arrival and several seconds before the occurrences of PGA and PGV. This type of early warning approach will become effective especially for close-in sites where warnings are most needed (Wu and Kanamori, 2008, Sensors).

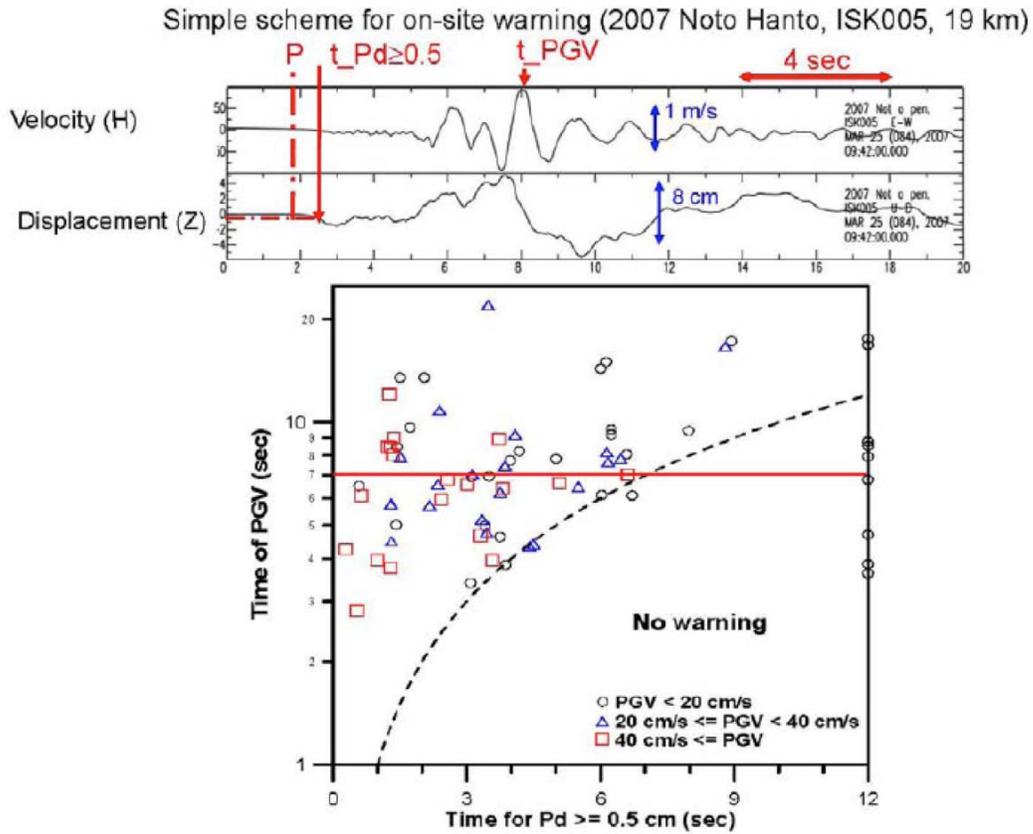


Figure 9. Warning is issued when the displacement exceeds 0.5 cm after P arrival. The peak ground motion velocity occurred at about 3 s after the warning (top). Time of PGV vs. time of $P_d \geq 0.5$ cm when a warning is issued (bottom).

4. References

- Allen, R. M., and H. Kanamori (2003). The potential for earthquake early warning in Southern California, *Science* 300 786-789.
- Allen, R. V. (1978). Automatic earthquake recognition and timing from single traces, *Bull. Seism. Soc. Am.* 68 1521-1532.
- Böse, M., E. Hauksson, K. Solanki, H. Kanamori, Y.-M. Wu, and T.H. Heaton (2009a), A New Trigger Criterion for Improved Real-Time Performance of On-Site Early Warning in Southern California, *Bull. Seism. Soc. Am.*, 99-2A.
- Böse, M., E. Hauksson, K. Solanki, H. Kanamori, and T. H. Heaton (2009a). Real-time testing of the on-site warning algorithm in southern California and its performance during the July 29 2008 M_w5.4 Chino Hills earthquake, L00B03, doi:10.1029/2008GL036366.
- Hauksson, E., K. Felzer, D. Given, M. Giveon, S. Hough, K. Hutton, H. Kanamori, V. Sevilgen, S. Wei, and A. Yong, and (2008), Preliminary Report on the 29 July 2008 Mw5.4 Chino Hills, Eastern Los Angeles Basin, California, Earthquake Sequence, *Seism. Res. Lett.*, 79, 855- 868.
- Kanamori, H. (2005), Real-time seismology and earthquake damage mitigation, *Annual Review of Earth and Planetary Sciences*, 33, 195-214, doi: 10.1146/annurev.earth.33.092203.122626.
- Nakamura, Y. (1988). On the urgent earthquake detection and alarm system (UrEDAS), *Proceeding of the 9th world conference on earthquake engineering* 1988, Tokyo-Kyoto, Japan.
- Shearer, P.M., G.A. Prieto, and E. Hauksson (2006), Comprehensive analysis of earthquake source spectra in southern California, *J. Geophys. Res.* 111, B06303, doi:10.1029/2005JB003979.
- Solanki, K., M. Böse, E. Hauksson, H. Kanamori, and T. Heaton (2009), CISN Earthquake Early Warning: Implementation and Testing of the Tau_c-Pd Onsite Warning Algorithm, 2nd International Workshop on Earthquake Early Warning, April 21-22, 2009 in Kyoto, Japan.
- Wald, D., V. Quitoriano, T. Heaton, H. Kanamori, C. Scrivner, and C. Worden (1999), TriNet ShakeMaps: Rapid generation of instrumental ground motion and intensity maps for earthquakes in southern California, *Earthquake Spectra*, 15, 537-556.
- Wu, Y. M., and H. Kanamori (2005a), Experiment on an onsite early warning method for the Taiwan early warning system, *Bull. Seism. Soc. Am.*, 95, 347-353.
- Wu, Y. M., and H. Kanamori (2005b), Rapid assessment of damaging potential of earthquakes in Taiwan from the beginning of P Waves, *Bull. Seism. Soc. Am.*, 95, 1181-1185.
- Wu, Y.-M., H. Kanamori, R.M. Allen, and E. Hauksson (2007), Determination of earthquake early warning parameters, τ_c and P_d , for southern California, *Geophys. J. Int.*, 170, 711-717, doi: 10.1111/j.1365-246X.2007.03430.x.
- Wu, Y. M., and H. Kanamori (2008a), Development of an Earthquake Early Warning System Using Real-Time Strong Motion Signals, *Sensors*, 8, 1-9.
- Wu, Y. M., and H. Kanamori (2008b), Exploring the feasibility of on-site earthquake early warning using close-in records of the 2007 Noto Hanto earthquake, *Earth, Planets and Space*, 60, 155-160.

Development, testing and performance of ElarmS in California and Japan

Holly M. Brown, Richard M. Allen, Margaret Hellweg, Oleg Khainovski,

Douglas Neuhauser, and Adeline Souf

Seismological Laboratory, University of California, Berkeley

Summary

In July 2009 the California Integrated Seismic Network concluded a three-year study of earthquake early warning systems in California. Three algorithms were expanded and examined during the study. Here we discuss the history, methodology and performance of one of the algorithms, ElarmS. Earthquake Alarm Systems, or ElarmS, uses peak displacement and maximum predominant frequency of the P-wave to detect earthquakes and quantify their hazard in the seconds after rupture begins. ElarmS was developed for Northern and Southern California, and now processes waveforms in realtime from 603 seismic sensors across the state. We outline the methodology as currently implemented, present several example events from different regions of California, and summarize the performance in terms of false and missed alarms. ElarmS was also tested offline with a dataset of 84 large magnitude earthquakes from Japan. The results from the Japan dataset were used to create a statistical error model for the algorithm. The model can be used to provide realtime uncertainty estimates at any stage in processing. In August 2009 the CISN embarked on a second three-year study of earthquake early warning. As part of this ongoing research, we identify the technological and methodological challenges facing ElarmS. Telemetry latencies and false alarm rates are two key opportunities for improvement.

1. Introduction

Earthquake early warning (EEW) systems are algorithms that detect the initial P-waves from an earthquake, rapidly estimate the location and magnitude of the event, and then predict subsequent ground shaking in the surrounding region. EEW systems offer the potential for a few seconds to a few tens of seconds warning prior to hazardous ground shaking: enough time for individuals to get to a safe location, perhaps under a sturdy table, for shutdown of utilities, slowing of trains, and other automated steps to reduce hazards from ground shaking.

In July 2009, the California Integrated Seismic Network (CISN) completed a three-year investigation into the viability of an EEW system in California. Three algorithms were expanded, tested, and compared during the study: Onsite, a single-station method that uses τ_c and P_d [Böse, et al., 2009], Virtual Seismologist, a network-based method that uses peak amplitudes and Bayesian statistics

[Cua, et al., 2009], and ElarmS, a network-based method that uses τ_p^{\max} and $P_{d/v}$ [Allen, et al., 2009].

The goal of the three-year project was to determine whether EEW is feasible in California. Results from each algorithm were continuously reported to a central database run by the Southern California Earthquake Center (SCEC) for analysis. By the end of the three years, all three algorithms had successfully predicted ground shaking before it was felt for many earthquakes in the state. At the end of the study the CISN determined that EEW is feasible, potentially desirable, and within reach for California. In August 2009 a second three-year study was initiated, to integrate the three test algorithms into a single prototype EEW system and provide realtime warning to a small group of test users by the end of the study in summer 2012.

Here we delineate the methodology, progress, and results of the ElarmS algorithm, which is now an integral part of the forthcoming prototype CISN EEW system. The ElarmS algorithms for magnitude and location estimation were developed offline with two datasets of events from Northern and Southern California. Those algorithms are now used in realtime, continuously processing waveforms from throughout the state of California and producing predictions of ground shaking within seconds of event detection. A separate dataset of events from Japan was processed offline to test ElarmS' performance for large events. From the Japan results we developed an error model which can be used in realtime to estimate the uncertainty in any ElarmS prediction.

2. Development and Methodology

2.1 Overview

Earthquake Alarm Systems, or ElarmS, is a network-based EEW system. The algorithm detects P-wave arrivals at several stations around an event epicenter and uses the amplitude and frequency content of the P-wave to rapidly estimate the magnitude and hypocenter of the event. Estimates from several stations are combined to improve accuracy and minimize the chance of a false alarm. ElarmS then applies the estimated magnitude and location to NEIC ShakeMap regional attenuation relations to produce a realtime prediction of impending ground shaking. Predictions above a certain threshold prompt an automatic alert message that can be sent to users.

The ElarmS algorithm is divided into a waveform processing module and an event monitoring module. The waveform processing module analyzes raw waveforms from all contributing stations, detects P-wave arrivals, and calculates the necessary ElarmS parameters: τ_p^{\max} , peak amplitudes, signal-to-noise ratio (SNR), peak ground acceleration and velocity (PGA and PGV), and trigger times. These parameters are then passed to the event monitor, which associates the triggers into an event, estimates the event location, estimates the magnitude, and predicts ground shaking. As additional

stations record P-wave arrivals, the waveform processing module passes their parameters to the event monitor, which includes them into the event analysis [Allen, 2007; Allen, et al., 2009].

2.2 Location

Event location is estimated by a four-stage algorithm, defined by the number of station triggers. When a single station triggers, the event is located directly beneath the station, at a depth of 8km. When two stations have triggered, the event is located between them based on arrival times, again at a depth of 8km. When three stations have triggered, ElarmS uses a two-dimensional grid search at a depth of 8km to determine the hypocenter and origin time that minimizes arrival time residuals. Finally, once four or more stations have triggered, ElarmS performs a three-dimensional grid search, with depth intervals every 10km, to estimate the hypocenter and origin time that minimizes arrival time residuals. In California, most events occur at depths of 5-15 km and the average depth is 8km [Hill, et al., 1990]. Rather than determining depth, ElarmS sets the depth of all California earthquakes to 8km. When processing events in Japan, all four stages are used including the depth determination.

2.3 Magnitude

ElarmS was originally developed from an empirically observed relationship between maximum predominant period, τ_p^{\max} , and final event magnitude [Allen and Kanamori, 2003]. For any vertical channel (HHZ, HLZ, HNZ), the predominant period time series is defined recursively by:

$$\tau_{p,i} = 2\pi (X_i/D_i)^{1/2}$$

where $X = \alpha X_{i-1} + x_i^2$ and $D_i = \alpha D_{i-1} + (dx/dt)_i^2$. The constant α is a smoothing factor equal to $1 - dt$, where dt is the sample interval, and x_i is the ground velocity of the last sample. Acceleration waveforms are integrated to velocity first, and all waveforms are filtered with a causal 3 Hz low-pass Butterworth filter. τ_p^{\max} is then the maximum observed τ_p value during the first four seconds of P-wave arrival.

To determine the empirical scaling relations, all τ_p^{\max} values for a given region are plotted against the final magnitude of each event. A least squares fit to the data produces the scaling relation, which is then used in realtime to estimate magnitude (see section 3.1).

In 2007 ElarmS was updated to utilize a second P-wave parameter, the peak amplitude [Wurman, et al., 2007]. Similar to the process for τ_p^{\max} , peak amplitudes observed during the first four seconds of P-wave arrival are compared to the final catalog magnitude for the event, and a least squares fit to the data provides a scaling relation. In California, peak displacement is used for broadband (HH) instruments and peak velocity is used for strong motion (HL and HN) instruments. While peak displacement has a theoretically longer period signal and thus less high frequency noise, integrating acceleration instruments twice to displacement can introduce sufficient errors to discount the

advantage of using displacement. Thus for acceleration instruments, peak velocity provides a more robust scaling relation than does peak displacement. In Japan, peak displacement produced the strongest scaling relation for all instruments, despite the double integration from acceleration. In general, we refer to the peak amplitude scaling relations as $P_{d/v}$ with the understanding that we may use P_d or P_v for any given site.

Although the scaling relations for τ_p^{\max} and $P_{d/v}$ are determined using four seconds of P-wave arrival, waiting for a full four seconds of P-wave to be available during realtime processing wastes valuable seconds of potential warning time. Instead ElarmS begins to apply the scaling relations and estimate magnitude as soon as a single station has observed a single full second of P-wave arrival (the first half-second is discarded). As additional seconds of P-wave become available, ElarmS recalculates τ_p^{\max} and $P_{d/v}$ accordingly. Since both τ_p^{\max} and $P_{d/v}$ are the maximum or peak values, they can only increase with additional seconds of data. The initial one-second magnitude estimate is therefore always a minimum estimate.

For each triggering station, τ_p^{\max} and $P_{d/v}$ are scaled separately to create two independent estimates of magnitude. The estimates are then averaged to form a single event magnitude for that station. As additional stations report P-wave triggers, their magnitude estimates are averaged into the event magnitude, to provide an increasingly accurate description of the event as time passes.

2.4 Ground Motions

Once location and magnitude have been estimated for an event, ground motion is predicted at each triggered station by applying the location and magnitude to NEIC-defined ShakeMap attenuation relations for the region [Wald, et al., 1999]. The resulting “AlertMap” displays predicted ground shaking in the familiar ShakeMap format, i.e. a map of predicted shaking intensity. As peak ground shaking is observed at individual stations, the observations are integrated into the shaking intensity map. Eventually, when all stations have reported peak ground shaking, the AlertMap looks much the same as the post-event ShakeMap.

The ElarmS algorithm has been tested with datasets from Northern California, Southern California, and Japan [Allen and Kanamori, 2003; Allen, 2007; Wurman et al., 2007; Tsang et al., 2007; Allen, et al., 2009; Brown, et al., 2009]. Each test dataset provided regional scaling relations for τ_p^{\max} and $P_{d/v}$, and utilized attenuation relations specific to that location. Most recently ElarmS has been adapted to run in realtime throughout the state of California.

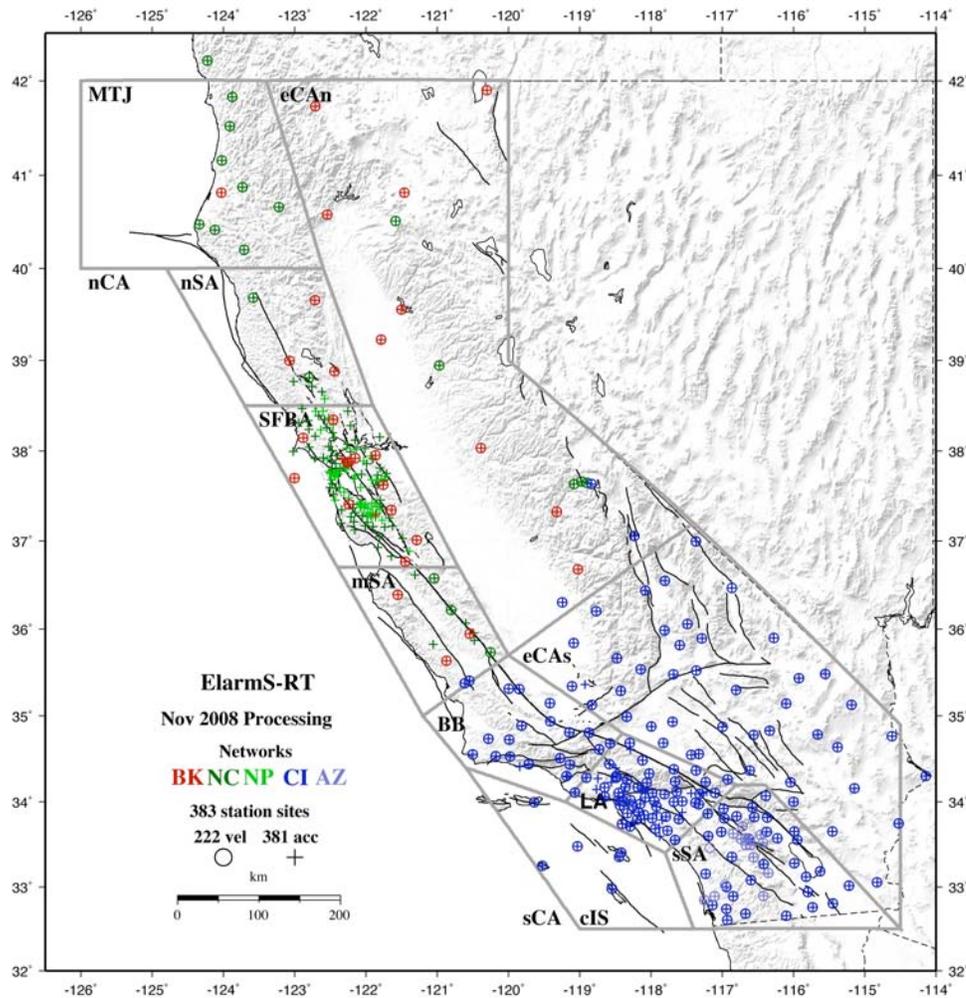


Figure 1: Realtime seismic stations used by ElarmS in California. Circles are velocity instruments, and crosses are accelerometers. Many stations have co-located velocity and acceleration sensors. The grey boxes indicate regions used for alert requirements: Mendocino Triple Junction (MTJ), north San Andreas (nSA), San Francisco Bay Area (SFBA), middle San Andreas (mSA), Big Bend (BB), Los Angeles (LA), south San Andreas (sSA), Channel Islands (cIS), east California south (eCAs), and east California north (eCAn). The straight line between regions mSA/eCAn and BB/eCAs is the Gutenberg-Byerly line dividing northern and southern California.

3. Application of ElarmS to California

3.1 Scaling and attenuation relations

Offline tests of California earthquake datasets have produced separate scaling relations for Northern and Southern California events [Wurman, et al., 2007; Tsang, et al., 2007]. The magnitude scaling relations are determined empirically by comparing observed τ_p^{\max} and $P_{d/v}$ values to final catalog magnitude for a dataset of test events, with as wide a range of magnitudes as possible. Once

determined, the scaling relations are used in realtime to estimate event magnitude, based on realtime observations of P-wave frequency and amplitude.

For northern California, Wurman, et al., [2007] analyzed a dataset of 43 events recorded by Berkeley Digital Seismic Network (BK) and Northern California Seismic Network (NC) seismometers (Figure 1) between 2001 and 2007, with magnitudes ranging from 3.0 to 7.1. The analysis resulted in the following scaling relations:

$$M_w = 5.22 + 6.66 * \log_{10}(\tau_p^{\max}) \quad \text{for } \tau_p^{\max} \text{ on HHZ, HLZ, HNZ channels}$$

$$M_w = 1.04 * \log_{10}(P_d) + 1.27 * \log_{10}(R) + 5.16 \quad \text{for } P_d \text{ on HH channels}$$

$$M_w = 1.37 * \log_{10}(P_v) + 1.57 * \log_{10}(R) + 4.25 \quad \text{for } P_v \text{ on HL channels}$$

$$M_w = 1.63 * \log_{10}(P_v) + 1.65 * \log_{10}(R) + 4.40 \quad \text{for } P_v \text{ on HN channels}$$

where R is the hypocentral distance to the station. The τ_p^{\max} and P_d relations are shown in Figure 2ab. These scaling relations are now used by ElarmS for all events north of the Gutenberg-Byerly line (shown on Figure 1 as the line between regions mSA/eCAn and BB/eCAs).

For southern California, Tsang, et al., [2007] analyzed a dataset of 59 earthquakes recorded by the Southern California Seismic Network (CI) between 1992 and 2003, with magnitudes ranging from 3.0 to 7.3. The analysis resulted in the following scaling relations (Figure 2cd):

$$M_w = 6.36 + 6.83 * \log_{10}(\tau_p^{\max}) \quad \text{for } \tau_p^{\max} \text{ on HHZ, HLZ, HNZ channels}$$

$$M_w = 1.24 * \log_{10}(P_d) + 1.65 * \log_{10}(R) + 5.07 \quad \text{for } P_d \text{ on HH, HL, HN channels}$$

These scaling relations are used by ElarmS for all events south of the Gutenberg-Byerly line.

Ground motions in California are predicted using the Boore, et. al., [1997] attenuation relations, as preferred by NEIC ShakeMap version 3.1 for events in California:

$$\log_{10}(PGA) = -0.313 + 0.527 * (M-6) + 0.778 * \ln(R) - B_v * \ln(V_s, V_a)$$

where $R = (R_{jb}^2 + h^2)^{1/2}$. R_{jb} is the "Joyner-Boore" distance to the surface projection of the fault, in km, and h is the hypocentral depth. The numerical coefficients are chosen for strike-slip faults [Wald, et al., 2005], and we currently substitute the epicentral distance for R_{jb} .

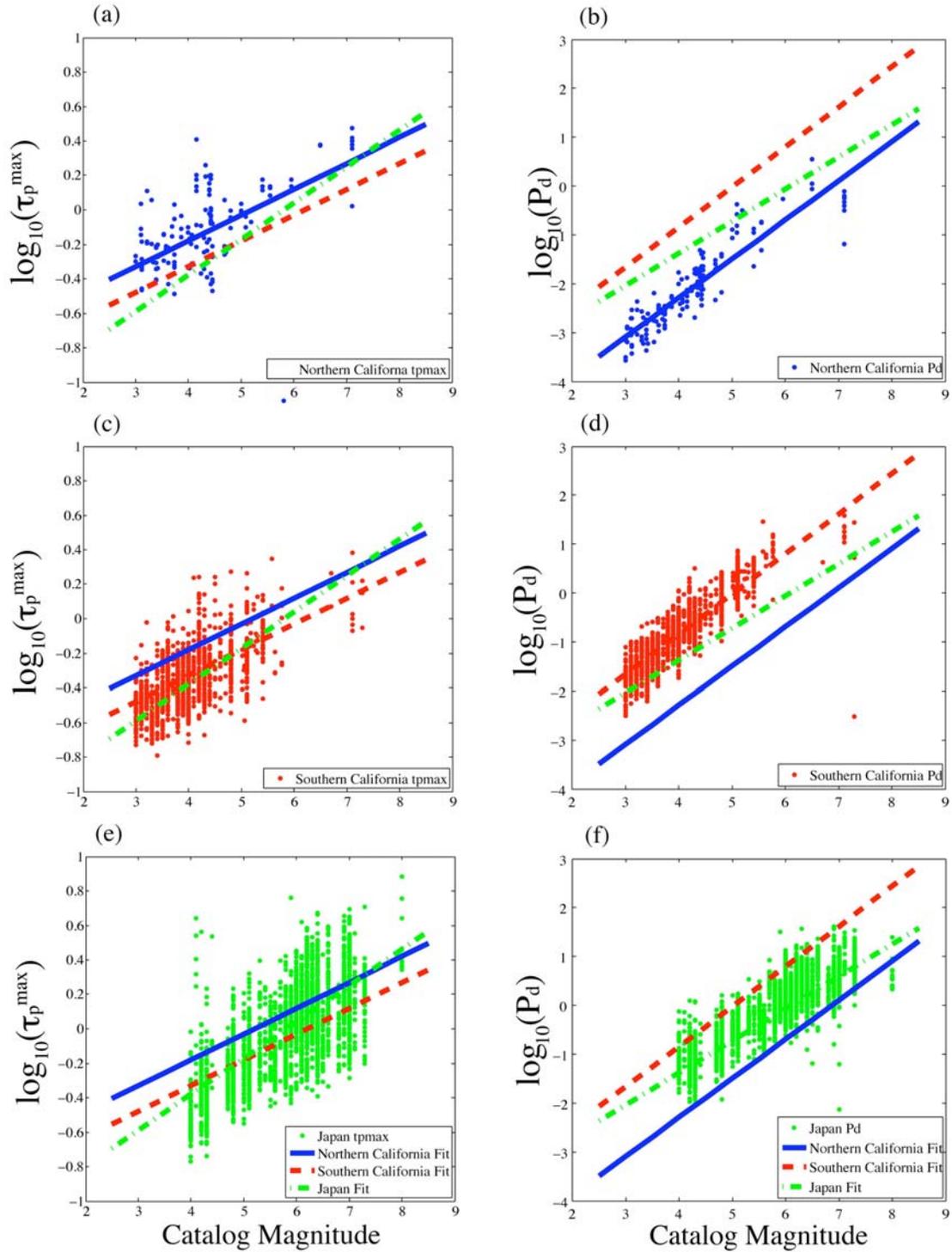


Figure 2: Scaling relations. (a) τ_p^{\max} , northern California; (b) P_d , northern California; (c) τ_p^{\max} , southern California; (d) P_d , southern California; (e) τ_p^{\max} , Japan; (f) P_d , Japan. Circles are individual station observations of τ_p^{\max} or P_d . Lines are regional scaling relations defined by the linear best fit to the data. The best-fit linear relations for all three regions are shown on all plots.

3.2 Realtime processing

ElarmS was adapted to run in realtime in Northern California in October 2007, and expanded statewide in November 2008 [Allen, et al., 2009]. The system now processes waveforms from all realtime-capable stations in the state: a total of 603 velocity and accelerations sensors at 383 sites (Figure 1). The ElarmS waveform processing module is distributed among three regional processing centers, which receive the continuously streamed waveforms. Data from the Berkeley Digital Seismic Network (BK) is streamed to UC Berkeley, data from the Northern California Seismic Network (NC) and from some stations in the USGS Strong Motion Network (NP) is streamed to USGS Menlo Park, and data from the Southern California Seismic Network (CI), the Anza Network (AZ), and the remaining NP stations is streamed to Caltech/USGS Pasadena. At these regional processing centers, the waveform processing module distills the waveforms to their essential parameters: trigger times, peak predominant period, peak amplitudes (acceleration, velocity, and displacement), peak ground shaking observations, and signal-to-noise ratio. These parameters are then forwarded to UC Berkeley, where a single event monitor integrates data from all of California to identify and analyze earthquakes in realtime. When an event is determined to be above a certain magnitude threshold, an alert message can be sent to users notifying them of the event location, origin time, estimated magnitude, and number of triggers. Currently alerts are sent to the authors and the SCEC database for CISN EEW analysis.

3.3 System latency

The total ElarmS processing time, from when a P-wave arrives at a station until ElarmS outputs event information, can be divided into two types: telemetry of data and computer analysis time. Data telemetry includes the time while a station collects data into a packet for transmission, transit from individual stations to the regional processing centers, where the waveforms are processed, and transit time from the processing centers to UC Berkeley where the single event monitor is located. The primary source of telemetry latencies is the packetization of data by station data loggers. A data logger will not send its data to the waveform processing module until the data packet is full. Packet sizes are usually of a configurable byte size, but many station data loggers are currently set for packet sizes equivalent to 4-6 seconds of data. Manually reconfiguring these data loggers to require packets equivalent to 1-2 seconds of data would greatly decrease the delays. In addition, all BK data loggers and most CI data loggers will be upgraded to data loggers with short 1 second packets in the next two years with recently provided US Federal ARRA stimulus funding.

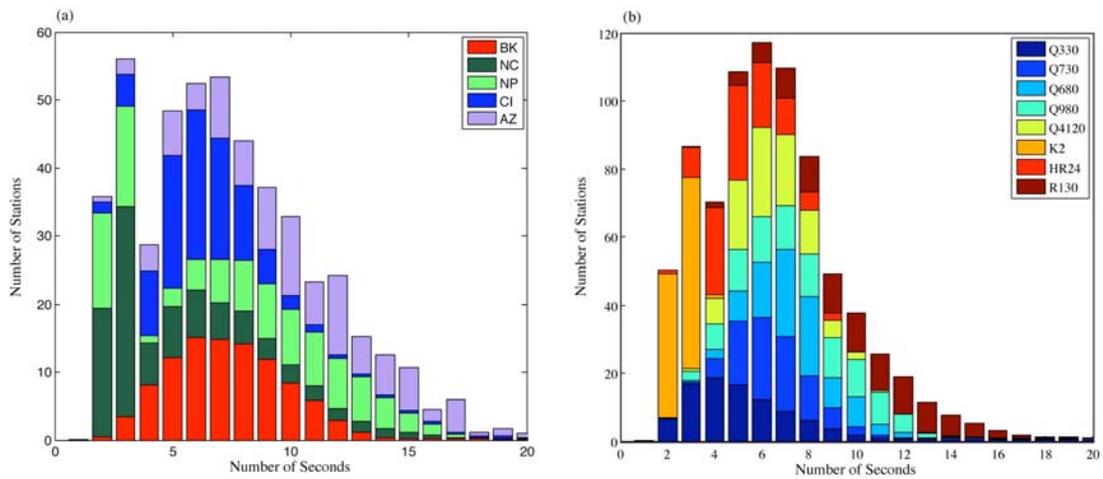


Figure 3: A stacked histogram of latencies by (a) network, and (b) data logger type. Both histograms are truncated at 20 seconds for clarity, but the long tail to the histogram continues, with columns of 0-2 data points, up to as much as 200 seconds.

Table 1: Median, mean and standard deviation values for the telemetry latencies shown in Figure 3.

		Median	Mean	Std. Dev.
Network	BK	6.23	6.48	2.76
	NC	2.54	10.62	29.78
	NP	7.38	7.09	4.54
	CI	5.17	9.49	28.62
	AZ	9.29	9.50	4.21
Data Logger	Q330	3.96	6.51	10.86
	Q730	5.54	6.94	7.14
	Q680	6.29	6.45	1.73
	Q980	6.60	6.73	2.49
	Q4120	5.30	9.93	42.37
	K2	1.57	1.58	0.37
	HR24	4.04	4.21	1.45
	R130	9.10	9.23	3.18

Figure 3a shows the data latencies for transmission to the waveform processing site by each seismic network. These delays are the difference in seconds between when a P-wave arrives at a station and when the waveform packet it is received by the regional processing center. They are thus composed of the time for a packet to fill and the time in transit to the regional processing center. The median latencies for each network are shown in Table 1. The median latency for all stations is 5.23 seconds. Each histogram is characterized by an extended tail at the high latencies (the figure is truncated at 20 seconds for clarity, but the distributions continue to higher latencies, up to several hundred seconds,

for a small number of stations). The tail indicates stations that are drastically delayed, due to poor telemetry availability, temporary telemetry failure or station disruption.

NC has the fastest median of 2.5 sec due to a large number of NC station data loggers configured for a packet size equivalent to 1-2 seconds of data. The large mean and standard deviation for NC show that the remaining stations are significantly slower, and there is an extensive tail to the distribution. The median and mean for BK are nearly identical, 6.2 and 6.5 sec respectively, with a low standard deviation. This indicates a nearly uniform hardware, software and telemetry configuration for all stations in the network, with few excessively delayed stations. CI uses much of the same equipment as BK but also has more variability in the equipment used resulting in a slightly faster median of 5.1 sec but slower mean of 9.5 sec. NP is uniform and a little slower with mean and median of 7.1 and 7.4 sec. AZ has the highest median latency which is due to an extra telemetry step as the data is forwarded through the Scripps Oceanographic Institute before arriving at the Caltech regional processing center.

Figure 3b shows the delays by data logger type, independent of network. The distribution statistics are shown in Table 1. The fastest data logger is the K2 used at many of the USGS sites and sends 1 sec data packets. The Quanterra Q330 comes second again due to the fact that it sends out 1 sec data packets, although there is a wider range of the total telemetry latencies which is likely due to software discrepancies between the different networks. The Berkeley processing software was designed for the older model data loggers and has not yet been updated to accommodate the Q330. This software will be upgraded by Spring 2010. The older Quanterra data loggers (the Q730, Q680, Q980 and Q4120) are slower. In the network upgrade that is now underway the majority of these older and slower data loggers are being upgraded to Q330s. The combined effect of new dataloggers and revised software will reduce the latencies at these stations by 3 to 5 sec.

3.4 Alert criteria

The station distribution in California is not uniform (Figure 1). Not surprisingly, the performance of a network-based system is directly related to the density of the network. Accuracy improves when more stations contribute to an event estimate, but potential warning time is lost while waiting for those stations to trigger, especially when the stations are far apart. ElarmS performs best in the heavily instrumented regions around Los Angeles, San Diego, and San Francisco (LA, sSA, SFBA in Figure 1). In these regions the mean station separation is only 20km, and the system often receives two or three triggers in the first second after an earthquake begins. In regions with lower station density the system must wait, as valuable seconds pass, until enough stations have reported P-wave arrivals. Regions with minimal instrumentation also suffer from higher false alarm rates, as there are fewer stations to contradict a false trigger. We therefore tailor the alert requirements to each region.

In regions SFBA, LA and sSA where station density is approximately 20km, the system requires at least 4 triggers within 30 km of the epicenter before an alert can be sent for an event. In regions eCAs, BB, mSA, and nSA, where stations are separated by 20-100km, we require 5 or more stations within 100km to trigger before an alert is generated. And in MTJ, eCAn, and cIS, where stations are more than 100km apart, we require 10 or more stations (at any epicentral distance) to trigger. These regional boundaries and requirements continue to be honed as we monitor the realtime system.

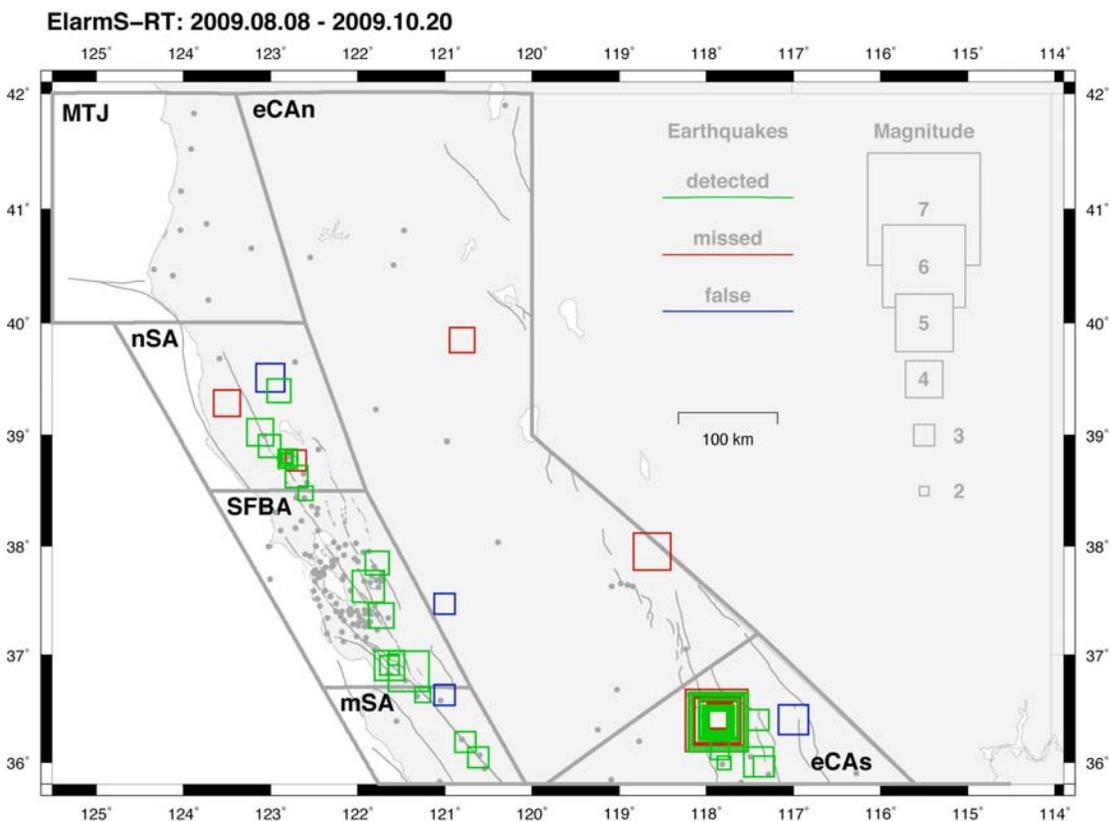


Figure 4: Map showing all ElarmsS-detected earthquakes with $M > 3$, and all false and missed alerts in Northern California, during a ten-week test period from 8 August 2009 until 20 October 2009. Green, red, and blue boxes are detected, missed, and false alerts, respectively. Grey dots are seismic stations.

3.5 False and Missed Alerts

Figure 4 shows all detected, false, and missed alerts with magnitude 3 or greater which occurred in Northern California during a ten-week test period from 8 August 2009 and 20 October 2009. A false alert is defined as an ElarmsS event that meets the alert criteria for its region but does not correspond

to a USGS-recognized event. A missed alert is a USGS $M > 3$ event for which no ElarmS alert message was issued; ElarmS may have not detected the event, or it may have detected the event but not satisfied the criteria required to issue an alert. For this ten-week test period there were 63 real events $M > 3$. ElarmS detected 45 of them and missed 18. Eleven of the missed events were part of an aftershock sequence described below. ElarmS also sent four false alert messages for nonexistent events.

The false and missed alarm rate is related to two factors: the station density, and whether an earthquake is occurring during a swarm such as during an aftershock sequence. In the SFBA region, where station density is approximately 20km, there were 8 detected events and no false or missed alerts for this time period (Figure 4). In mSA there were 3 detected events and 1 false alert. In nSA there were 8 detected events, 1 false alert and 2 missed alerts. Performance is moderate in the mSA and nSA regions as the station spacing is 20-100km.

In the eCAn and eCAs regions performance is much poorer due to the much lower station density. In eCAn there were two missed alerts and one false alert. In the eCAs region in the lower right of the map there is a cluster of green (detected) and red (missed) squares. These represent two $M5$ events on October 1st and 3rd, and their aftershock sequences. ElarmS successfully detected the $M5.1$ event on October 1st, but missed the $M5.2$ event two days later. It caught 20 out of 31 total aftershocks of magnitude 3 or greater. ElarmS missed the second large event due to concurrent aftershock activity from the first event.

This illustrates the challenge of defining optimal alert criteria for each region. Criteria which are too strict (requiring too many stations to trigger) may fail to be met by a moderate size event, resulting in no alert message even though the event is real, or will slow down the time until an alert is issued. Criteria which are too loose (requiring too few stations) may be met by unrelated, erroneous triggers, resulting in an alert message when there is no real event. As with all associators the performance is also reduced during swarms of seismicity or aftershock sequences. Improvements to the associator scheme specifically for early warning applications would be beneficial.

4. Sample Events

We illustrate ElarmS performance in California with three sample events from different regions of the state, all processed by the realtime system.

4.1 $M_w 5.4$ Alum Rock, SFBA region

Figure 5 shows the $M_w 5.4$ Alum Rock event, which occurred on 30 October, 2007. This was the largest event in the San Francisco Bay Area since the 1989 Loma Prieta $M_w 6.9$ event. At the time of

the Alum Rock earthquake, ElarmS had been running in realtime for less than a month and used only stations from the BK network. The event begins in Figure 5a when two stations trigger simultaneously. The location is estimated between the stations, at a depth of 8km. One second later (Figure 5b), the magnitude is estimated at 5.2, using the observed τ_p^{\max} and $P_{d/v}$ values from the two triggered stations. A third station triggers and the location is triangulated based on the arrival times at the three stations. The estimated location and magnitude are applied to local attenuation relations to produce a prediction of ground shaking around the epicenter. The errors in the PGA and PGV predictions are -0.2 and -0.3, respectively, at this time. PGA and PGV errors are the difference of the logarithm of the observed minus the predicted ground motions; a factor of two difference between the predicted and observed PGA corresponds to an error of 0.7, and a factor of 10 to an error of 2.3. One second later (Figure 5c), the τ_p^{\max} and $P_{d/v}$ values from the third station are incorporated, and the

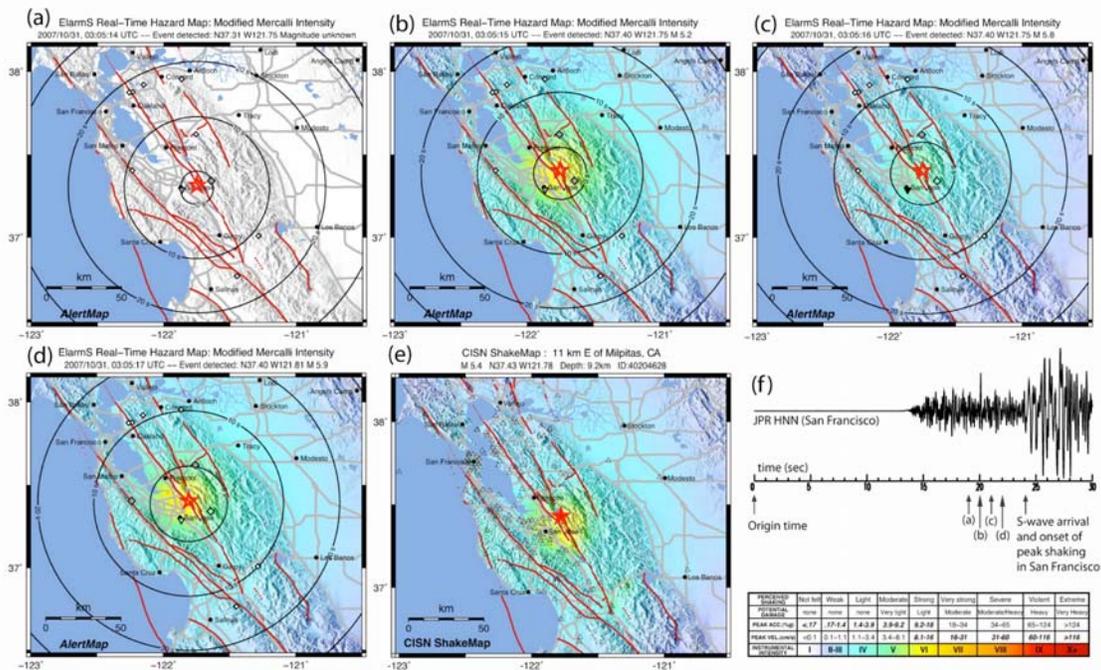


Figure 5: Example of ElarmS event processing for the 30 October 2007 Alum Rock M_w 5.4 earthquake. (a-d) Progressive AlertMaps as stations trigger and the event is analyzed in realtime. The AlertMaps themselves were produced after the event, but the data used to create them was available at the time indicated on the map. (e) NEIC ShakeMap published after the event. (f) Timeline comparing when the data used to create the AlertMaps was available with respect to the arrival of peak ground shaking in San Francisco.

magnitude estimate rises to M5.8. The errors in PGA and PGV change to 0.0 and -0.4. One second later (Figure 5d), a fourth station triggers, the location is adjusted, and the magnitude estimate rises to M5.9. The predictions of peak ground shaking are adjusted to account for the new location and magnitude estimates. The PGA and PGV errors change to 0.1 and -0.2. As additional seconds pass,

more stations trigger and their P-wave parameters are incorporated into the evolving estimates of location and magnitude, and the predictions of ground shaking.

Figure 5e shows the CISN ShakeMap published after the Alum Rock event. From the time of the first magnitude estimate, one second after the first P-wave detection, the predictive AlertMap (Figure 5b) is a close match to the ShakeMap. Figure 5f shows a seismogram recorded in San Francisco during the Alum Rock earthquake. The timeline denotes the times at which the data used in (a),(b),(c), and (d) was available. At the time ElarmS applied a 15 second buffer to the incoming waveforms, to reduce latency differences between stations. Despite the 15 second buffer, the data used to create (b-d) was available four to two seconds before the S-waves reached San Francisco and peak ground shaking began. This event represented the first “proof-of-concept” event for the realtime ElarmS system as it illustrates that hazard information is available before shaking is felt.

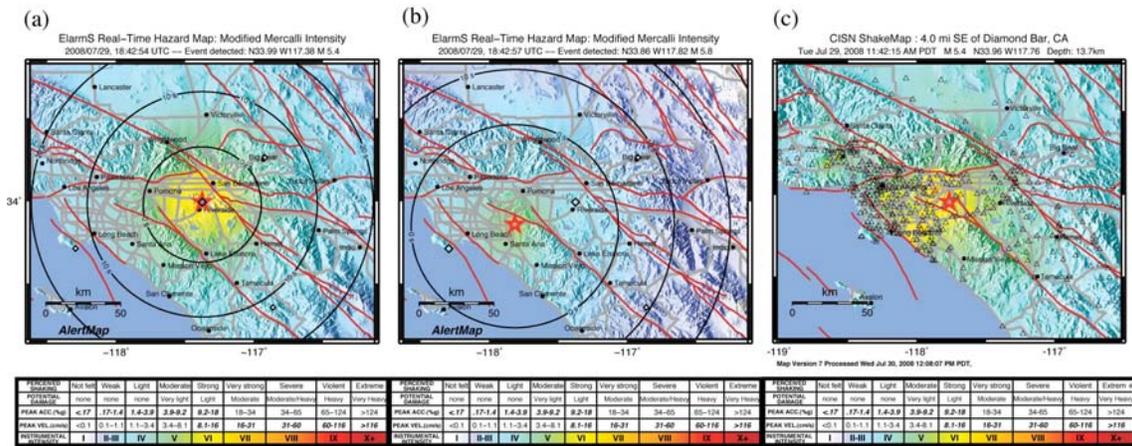


Figure 6: Example of ElarmS processing for the 29 July 2008 Chino Hills Mw5.4 earthquake. (a) AlertMap showing predictions of ground shaking after one station had triggered. (b) AlertMap showing adjusted predictions after two stations had triggered. (c) NEIC ShakeMap published after the event.

4.2 Mw5.4 Chino Hills, LA region

Figure 6 shows the Mw5.4 Chino Hills event, which occurred on 29 July, 2008. At the time ElarmS was midway through the conversion to statewide coverage, and was receiving data from only 15 southern California stations. ElarmS was still able to estimate magnitude, location and ground shaking using only the three stations within 100km of the epicenter. When the first station triggered, the event was located directly beneath the station at a depth of 8km. The observed τ_p^{max} and $P_{d/v}$ values were used to estimate a magnitude of 5.4. From that location and magnitude, local attenuation relations were used to predict peak ground shaking in the region (Figure 6a). After a second station triggered, the location was adjusted between the stations based on arrival times, at a depth of 8km. The τ_p^{max} and $P_{d/v}$ magnitudes for the second station were averaged together with those from the first station, producing a new event magnitude of M5.8. The new location and magnitude were used to update the

predictions of ground shaking (Figure 6b). Figure 6c shows the NEIC ShakeMap for comparison. The ShakeMap is published after the event, using observations from all available stations. The ElarmsS predictive AlertMap is reasonably similar to the ShakeMap, considering ElarmsS used data from only two stations.

4.3 M_w4.4 Lone Pine, eCAs region

The Lone Pine M_w4.4 occurred on October 3, 2009, in the eCAs region. In this region the stations are separated by 20-100km, so ElarmsS requires at least 5 station to trigger before issuing an alert. In Figure 7a the event is detected when two stations trigger simultaneously, four seconds after the event origin time. One second later (7b) the event magnitude is estimated at 4.0. Four more seconds pass before a third station triggers, at which point the location is adjusted and the magnitude estimate is raised to 4.1 (7c). The thin station coverage necessitates waiting longer in this region than

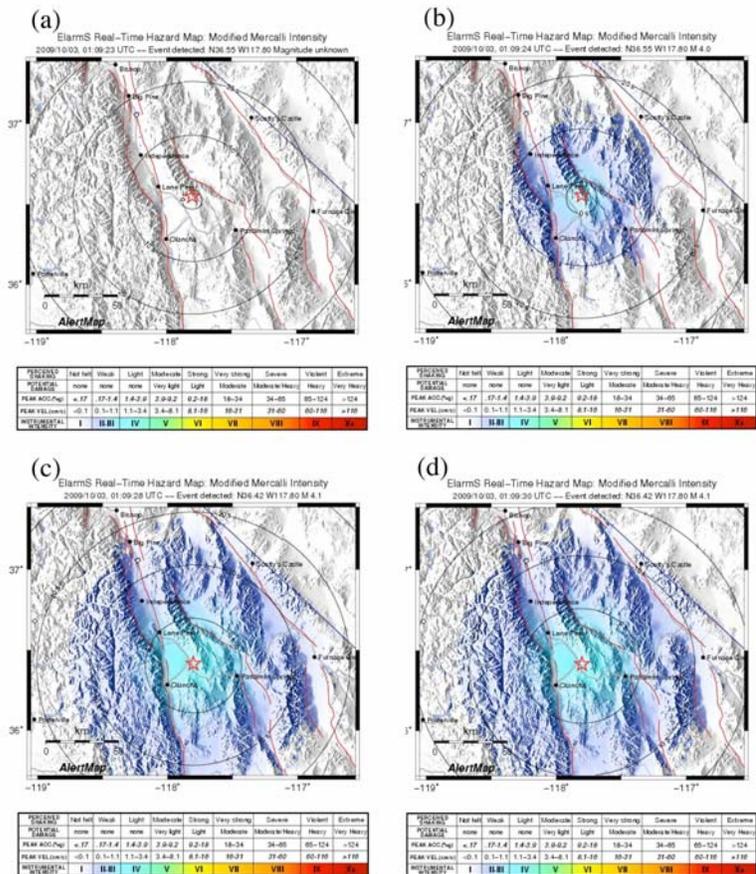


Figure 7: Example of ElarmsS processing for the 3 October 2009 Lone Pine M_w4.4 earthquake. (a) Hypocenter was estimated when two stations triggered, 4 seconds after the event began. (b) One second later (OT + 5 seconds) magnitude was estimated at 4.0, using P-wave parameters from the two triggering stations. (c) Four seconds later (OT + 9) a third station triggered. Location, magnitude, and ground shaking predictions were adjusted. (d) One second later (OT + 11), the five station requirement was met and an alert was issued (to the authors) for this event.

in the previous examples. The five station requirement for alert issuance is not met until two seconds later (7d), eleven seconds after the event begins. The fourth and fifth stations did not appreciably change the magnitude, location, or ground motion predictions in this case, but they ensured that the event was real.

5. Application of ElarmS to Japan

5.1 Scaling and attenuation relations

While ElarmS has been tested with many datasets in California, there are few recent, well-recorded, large earthquakes in California. Since an early warning system is designed specifically to warn people of large events, we are especially interested in its performance for these events. Thus we tested the system with a dataset of large events from Japan [Brown, et al., 2009]. The Japanese events also provided insight into ElarmS' performance in a subduction zone environment.

The dataset included 84 Japanese events that occurred between September 1996 and June 2008 (Figure 8). The magnitudes ranged from 4.0 to 8.0, with 43 events of magnitude 6.0 or greater. The largest event was the M8.0 Tokachi-Oki earthquake of 26 September, 2003. The events were recorded by Japan's Kyoshin Net (K-NET) strong-motion seismic network. K-NET consists of 1,000 digital strong motion seismometers, distributed across Japan with approximately 25km spacing. Each station is capable of recording accelerations up to 2,000 cm/s^2 , with a sampling frequency of 100 samples per second and a dynamic range of 100dB.

The events were processed offline, using all available data, using the same methodology as described above. The first step is to determine scaling relations between the predominant period and peak amplitudes of the P-waves and the magnitude for the event dataset. The observed scaling relations for Japan are shown in Figure 2ef and are:

$$M_{\text{JMA}} = 4.76 * \log_{10}(\tau_p^{\text{max}}) + 5.81$$

$$M_{\text{JMA}} = 5.82 + 1.52 * \log_{10}(P_d) + 1.39 * \log_{10}(R)$$

where M_{JMA} is the JMA catalog magnitude and R is the hypocentral distance. The predominant periods observed in Japan are of similar values to those of Northern and Southern California, but the best-fit slope is steeper in Japan. The peak amplitude values are higher than those in Northern California and lower than those in Southern California, with a slightly shallower slope in Japan.

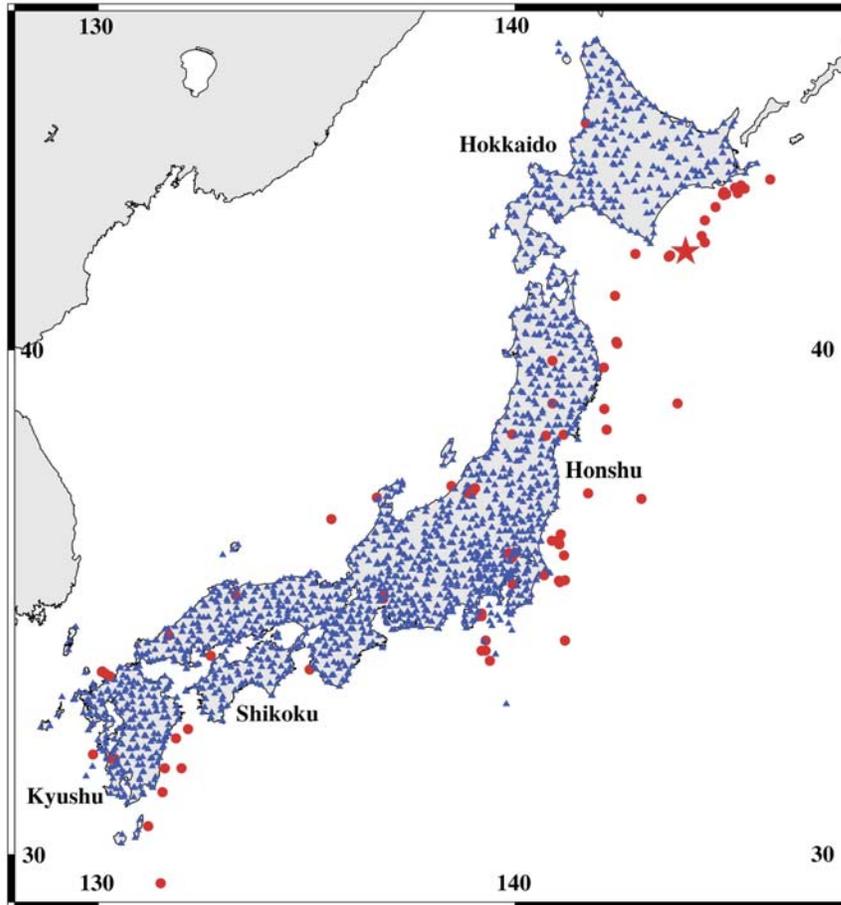


Figure 8: Events and stations used in the Japan test dataset. Red circles are events, blue triangles are K-NET stations. The red star is the largest event in the study, the M8.0 Tokachi-Oki earthquake of Sept 26, 2003.

For the prediction of peak ground shaking, we used the attenuation relations that NEIC ShakeMap uses for Japanese events. The NEIC global attenuation relations use either the Boore, et al., [1997] or the Youngs, et al., [1997] model, depending on depth and magnitude of the event. For events shallower than 20km or smaller than magnitude 7.7, the relations are defined by Boore, et al., [1997], with numerical coefficients specified for reverse faulting:

$$\ln(PGA) = -0.117 + 0.527 * (M-6) + 0.778 * \ln(R)$$

where R is defined by

$$R = (R_{jb}^2 + h^2)^{1/2}$$

R_{jb} is the distance in km to the surface projection of the fault and h is the hypocentral depth. We substitute the epicentral distance for R_{jb} . The coefficients used in Japan are different from those used in California, where strike-slip faulting is predominant.

For events deeper than 20km or greater than magnitude 7.7, the NEIC and ElarmS use the attenuation relations defined by Youngs, et al., [1997]:

$$\ln(PGA) = 0.2418 + 1.414 * M - 2.552 * \log_{10}(R_{rup} + 1.7818 \exp(0.554 * M)) + 0.00607 * H$$

where R_{rup} is the distance to the surface projection of the rupture and H is the hypocentral depth. We substitute the epicentral distance for R_{rup} .

5.2 Performance for large magnitudes

Once the necessary scaling relations had been developed all 84 events were processed in a simulated realtime environment to provide ElarmS predictions of ground shaking. We assumed zero data latency and processed data sequentially according to the time-stamp on the waveform data. After the events were processed we analyzed ElarmS performance for different magnitude ranges. Figure 9 shows the resulting ElarmS magnitude error histograms. The blue histogram is the magnitude error for all events in the Japanese dataset, with magnitudes from 4.0 to 8.0. The mean error for all events was 0.04 magnitude units, with a standard deviation of 0.4. The green histogram is the magnitude error for all events magnitude 6.0 or greater (of which there are 43). The mean error for this distribution is again 0.04, with a standard deviation of 0.5. This is a similar distribution statistically to that for all events. The red histogram is the magnitude error for events magnitude 7.0 or greater (of which there are seven in this dataset). Of the seven events $M > 7$, four of the magnitudes are underestimated, two are overestimated, and one is accurately estimated. The mean error for this distribution is -0.2 magnitude units, with a standard deviation of 0.5. This lower mean error means that ElarmS underestimates the magnitude of the largest events by 0.2 magnitude units on average. An underestimation of 0.2 magnitude units is within our tolerance for ElarmS magnitude estimates, but we recognize that the magnitude algorithm may need to be adjusted to prevent underestimation in the future. A first step may be to weight the average of τ_p^{\max} and $P_{d/v}$ in favor of τ_p^{\max} for high magnitude events, since τ_p^{\max} is less prone to saturation effects at the highest magnitudes [Brown, et al., 2009].

5.3 Methodological improvements

The Japanese dataset provided some methodological challenges. The majority of the events were offshore. The resulting limited azimuthal coverage (all stations are onshore) slowed down our location algorithm, requiring more station trigger times and therefore more seconds to produce a reasonable epicentral estimate. Many of the events were also deep. The original California location algorithm assumed a depth of 8km for all events, and found the hypocenter on a 2D grid at that depth. For the subduction zone events we expanded the algorithm into a 3D grid search, finding hypocenters at depths down to 80km, in 10km increments.

5.4 Error model

As part of the Japan dataset testing, we developed an error model to analyze the errors in Elarms' output [Brown, et al., 2009]. We separated the algorithm into its location, magnitude, and ground motion steps, and isolated the errors produced during each step. Errors were calculated by comparing the estimated location or magnitude to the catalog location or magnitude, and the predicted ground shaking at all stations and times prior to recording ground shaking to the eventual observation of peak ground shaking at that station. Predictions of peak ground shaking at stations

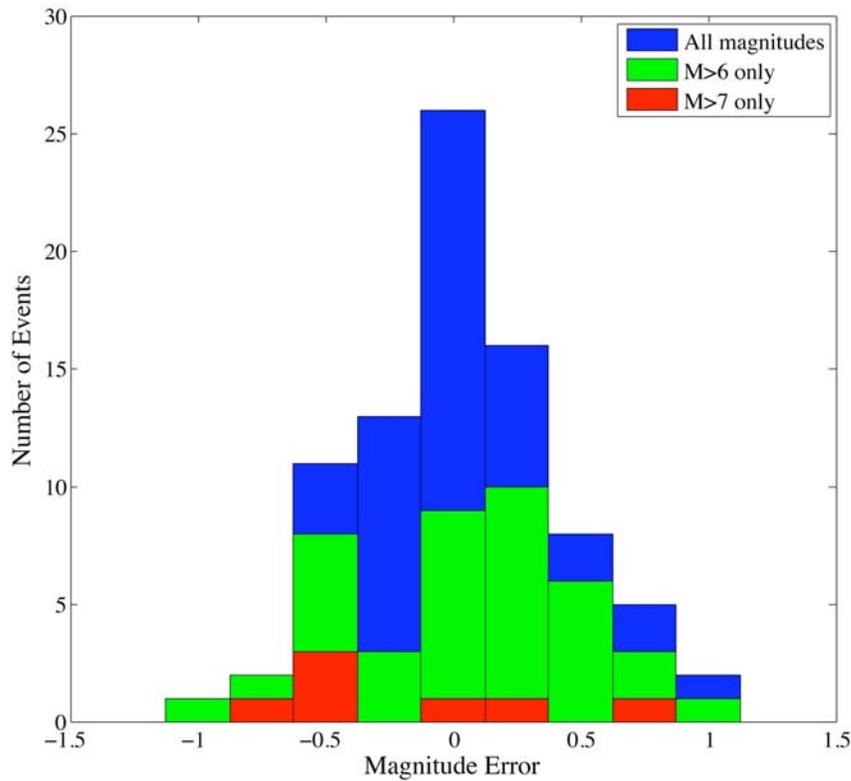


Figure 9: Histogram of magnitude errors for Japan dataset. The blue histogram is the distribution of magnitude error for all 84 events in the Japan dataset, M4.0 to M8.0. The green histogram is the distribution for the subset of 43 events with magnitude 6.0 or greater (up to and including magnitude 8.0). The red histogram is for the subset of 7 events with magnitude 7.0 or greater (up to and including magnitude 8.0).

Table 2: Statistical parameters (mean ± standard deviation) of error distributions for magnitude, location, and ground motion.

	0 stations	1 station	2 stations	3 stations	4 stations	5 stations
Mag, 1 sec	-	-0.38 ± 0.63	-0.33 ± 0.56	-0.37 ± 0.57	-0.39 ± 0.56	-0.41 ± 0.56
Mag, 2 sec	-	-0.2 ± 0.57	-0.15 ± 0.5	-0.18 ± 0.54	-0.21 ± 0.52	-0.22 ± 0.50
Mag, 3 sec	-	-0.09 ± 0.53	-0.05 ± 0.48	-0.08 ± 0.52	-0.10 ± 0.49	-0.10 ± 0.47
Mag, 4 sec	-	0.01 ± 0.52	0.04 ± 0.46	0.03 ± 0.48	0.03 ± 0.44	0.02 ± 0.43
Mag, 5 sec	-	0.04 ± 0.50	0.07 ± 0.45	0.07 ± 0.48	0.07 ± 0.43	0.06 ± 0.42
Location	-	33.6 ± 17.9	32.1 ± 21.4	32.5 ± 18.7	18.8 ± 13.6	21.1 ± 16.8
PGA	0.11 ± 0.30	0.09 ± 0.35	0.08 ± 0.37	0.06 ± 0.29	0.10 ± 0.28	0.03 ± 0.30

after the peak shaking had occurred were not included in the error analysis. The errors of each component of the system are shown in Table 2.

The accuracy of any given step is dependent on the amount of data available. The error in the location estimate, for example, is dependent on the number of stations reporting P-wave arrivals. The error in the magnitude estimate is dependent on both the number of stations providing information and the number of seconds of P-wave that have arrived at each station. The error in the prediction of peak ground shaking is dependent on the number of stations whose observations of peak ground shaking have been used to adjust the prediction.

The errors calculated (Table 2) were then used to produce an error model for ElarmS' final prediction of ground shaking, given any combination of inputs. If there were no errors at all in the system, then the ElarmS prediction of ground shaking would be based on the same magnitude and location that the catalog uses. Since ElarmS uses the same attenuation relations as the NEIC, an error-free ElarmS AlertMap should look exactly like the NEIC ShakeMap. Therefore, the error contributed by ElarmS is the difference between the ShakeMap calculation of ground shaking and the AlertMap prediction of ground shaking. The ideal, error-free output is defined by the attenuation relations for an event. For example, for an event shallower than 20km depth with a magnitude less than 7.7, the error-free output would simply be the Boore, et al., [1997] attenuation relation. For peak ground acceleration (PGA):

$$\ln(\text{PGA})_{\text{ideal}} = -0.117 + 0.527 * (M-6) + 0.778 * \ln(R) \quad \text{Ideal, error-free output}$$

where M is magnitude and R is the distance from the event epicenter to the location where PGA is being predicted.

We then introduce errors into the calculation, using the error distributions we observed for our Japan dataset.

$$\ln(\hat{\text{PGA}}) = -0.117 + 0.527 * (M + \epsilon_M - 6) + 0.778 * \ln(R \pm \epsilon_R) + \epsilon_{\text{Att}} \quad \text{ElarmS output}$$

where M is the catalog magnitude, R is the hypocentral distance, and ϵ_M , ϵ_R , and ϵ_{Att} are the errors in magnitude, location, and attenuation relations, respectively.

The difference between $\text{PGA}_{\text{ideal}}$ and $\hat{\text{PGA}}$ is the error in our final prediction of ground shaking.

$$\epsilon_{\text{PGA}} = \ln(\text{PGA})_{\text{ideal}} - \ln(\hat{\text{PGA}}) \quad \text{Error}$$

This represents the total error in the entire algorithm. ϵ_{PGA} is a unitless value; a factor of two difference between the ideal and estimated PGA corresponds to an error of 0.7, and a factor of 10 to an error of 2.3.

The errors for each step (ϵ_M , ϵ_R , ϵ_{Att}) are dependent on the quantity of data included (the number of trigger times, the number of τ_p^{\max} and $P_{d/v}$ values, etc.) and vary within the probability distributions defined in Table 2. Thus the error model is similarly dependent. We calculated ϵ_{PGA} 1000 times for every combination of data inputs, 1086 combinations, each time choosing the error values by a Monte Carlo simulation based on the mean and standard deviation of the error distributions (Table 2). The resulting 1000 values for ϵ_{PGA} are used to create a probability distribution for ϵ_{PGA} given that specific combination of data inputs. Figure 10a shows three sample ϵ_{PGA} distributions, and Figure 10b shows all 1086 ϵ_{PGA} distributions, corresponding to 1086 unique combinations of data inputs

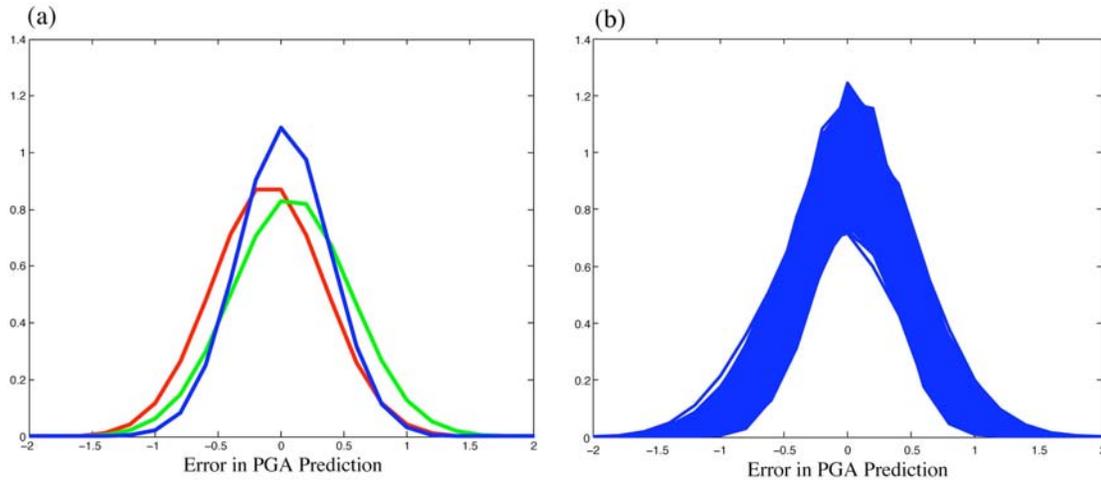


Figure 10: Error model distributions. (a) Three examples, showing best-fit Gaussian distributions for errors in ground motion estimation, given various quantities of data input. The red line is the error if two stations contribute to a location estimate, two stations contribute to the magnitude estimate (one using 1 second of P-wave data, one using 2 seconds), and zero stations report PGA observations. The green line is error if three stations contribute to the location estimate, two stations contribute to the magnitude estimate (one with 2 seconds of P-wave data, one with 3 seconds), and one station reports a PGA observation. The blue line is error if five stations contribute to the location estimate, five stations contribute to the magnitude estimate (4 with four seconds of P-wave data, one with 3 seconds), and three stations report PGA observations. (b) All 1086 error distributions resulting from the error model. Each line represents a unique combination of data inputs.

(number of stations contributing to location estimate, number stations contributing to magnitude estimate, number of seconds of P-wave for each station, and number of observations of peak ground shaking). The mean errors for all ϵ_{PGA} distributions range from -0.17 to 0.20, with a median error of 0.04. Standard deviations range from 0.32 to 0.56, with a median of 0.39. The standard deviations of all error distributions are less than 0.7 meaning less than a factor of 2 error in the PGA prediction. These calculated error distributions are stored in an internal library, accessible during processing for realtime estimates of uncertainty in PGA predictions.

6. Conclusion

The three-year CISN project gave us the opportunity to combine the offline development of ElarmS in California with the error analysis performed in Japan and produce a statewide realtime system. Already we have integrated data from five disparate networks, adapted our algorithms to run in realtime using data that is unevenly delayed by telemetry, and added the ability to send alert messages within seconds of event detection. While improvements to the seismic networks in California would improve ElarmS performance, ElarmS has successfully predicted ground shaking for many events even with the current network of stations.

There are opportunities for improvement in the next three-year phase of the project. Our algorithm continues to struggle with false alarms, especially in the regions with low station density. Honing our regional trigger requirements may be the primary step needed to reduce the false alarm rate. In addition, the event associator needs to be improved to better tolerate aftershock sequences, so that we don't risk missing a large mainshock due to its foreshock or other nearby events.

Data latencies are also a significant problem, claiming much of the potential warning time. Some latencies may be reduced by more efficient code design, such as updating the BK network software to accommodate the faster Q330 data loggers. Others require reformatting individual station data loggers, or upgrading data logger hardware. In the next two years ARRA stimulus funding will be used to upgrade many data loggers throughout the CISN, reducing latency by 3-5 seconds at these sites. The current statistical median latency is 5.2 seconds. With these upgrades we anticipate this will be reduced to 2-3 seconds.

We expect many improvements to the ElarmS code and the CISN networks during the coming three years. Learning from the realtime experience of the last three years, the ElarmS, VS, and Onsite methodologies will be integrated into a single prototype system. New code is being written to reduce processing delays and hardware upgrades will reduce data transmission latencies. The CISN is currently identifying a small group of about 10 test users who will soon start to receive alerts from the new prototype system, called the CISN ShakeAlert System.

References

- [1] R. M. Allen (2004), Rapid magnitude determination for earthquake early warning, in *The Many Facets of Seismic Risk*, G. Manfredi et al. (Editors), 15– 24, Univ. degli Studi di Napoli "Federico II", Naples, Italy.
- [2] R. M. Allen (2006), Probabilistic warning times for earthquake ground shaking in the San Francisco Bay Area, *Seismol. Res. Lett.* 77(3), 371–376.
- [3] R. M. Allen (2007), The ElarmS earthquake early warning methodology and application across

- California, in *Earthquake Early Warning*, P. Gasparini, et al. (Editors), 21– 44, Springer Ital., Milan, Italy.
- [4] R. M. Allen and H. Kanamori (2003), The potential for earthquake early warning in southern California, *Science* 300, 786– 789.
- [5] R. M. Allen, H. Brown, M. Hellweg, O. Khainovski, P. Lombard, and D. Neuhauser (2009), Real-time earthquake detection and hazard assessment by ElarmS across California, *Geophys. Res. Lett.*, 36, L00B08.
- [6] D. M. Boore, W. B. Joyner, and T. E. Fumal (1997), Equations for estimating horizontal response spectra and peak accelerations from western North American earthquakes: A summary of recent work, *Seismol. Res. Lett.* 68(1), 128– 153.
- [7] M. Böse, E. Hauksson, K. Solanki, H. Kanamori, and T. H. Heaton (2009), Real-time testing of the on-site warning algorithm in southern California and its performance during the July 29 2008 M_w 5.4 Chino Hills earthquake, L00B03.
- [8] H. M. Brown, R. M. Allen, and V. F. Grasso (2009), Testing ElarmS in Japan, *Seismol. Res. Lett.*
- [9] G. Cua, M. Fischer, T. Heaton, and S. Wiemer (2009), Real-time performance of the Virtual Seismologist earthquake early warning algorithm in southern California, *Seismol. Res. Lett.*
- [10] D. P. Hill, J. P. Eaton, and L. M. Jones (1990), Seismicity, 1980– 86, in *The San Andreas Fault System, California*, R. E. Wallace (Editor), U.S. Geol. Surv. Prof. Pap., 1515, 115–152.
- [11] A. B. Lockman and R. M. Allen (2005), Single-station earthquake characterization for early warning, *Bull. Seismol. Soc. Am.* 95(6), 2029 – 2039.
- [12] E. L. Olson and R. M. Allen (2005), The deterministic nature of earthquake rupture, *Nature* 438(7065), 212– 215.
- [13] L. Tsang, R.M. Allen and G. Wurman (2007) Magnitude scaling relations from P-waves in southern California, *Geophys. Res. Lett.* 34 L19304.
- [14] D. J. Wald, V. Quitoriano, T. H. Heaton, H. Kanamori, C. W. Scrivner, and B. C. Worden (1999a), TriNet ShakeMaps: Rapid generation of instrumental ground motion and intensity maps for earthquakes in southern California, *Earthquake Spectra* 15(3), 537–556.
- [15] D. J. Wald, B. C. Worden, V. Quitoriano, and K. L. Pankow (2005), ShakeMap1 Manual: Technical manual, users guide, and software guide, Tech. Methods 12–A1, U.S. Geol. Surv., Reston, Va.
- [16] G. Wurman, R. M. Allen, and P. Lombard (2007), Toward earthquake early warning in northern California, *J. Geophys. Res.*, 112, B08311.
- [17] R. R. Youngs, S.-J. Chiou, W.J. Silva, and J.R. Humphrey (1997), Strong ground-motion relationships for subduction zones, *Seism. Res. Lett.* 68, No.1, 58-73

Summary of Virtual Seismologist status and performance as of October 2009**Georgia Cua, Michael Fischer, Thomas Heaton**

The Virtual Seismologist (VS) algorithm began real-time data processing at the Southern California Seismic Network (SCSN) in July 2008. Since then, the VS algorithm has continued to run in real-time at SCSN, and has been installed in at the Berkeley Digital Seismic Network (BDSN) and the Menlo Park strong motion network.

Details on the methodology and real-time implementation of the VS algorithm are documented in the publication “Real-time performance of the Virtual Seismologist earthquake early warning algorithm in Southern California” (Cua, Fischer, Heaton and Wiemer, 2009, *Seismol. Res. Lett.*, 30, 5). While we are continually adding enhancements and new features to the VS algorithm, we choose to be somewhat conservative in the code versions that we allow to run in real-time. The SRL paper is an accurate description of the current version of the VS codes installed at SCSN, BDSN, and Menlo Park. However, only the SCSN installation was configured to send reports to the SCEC EEW Testing Center, due to a lack of funding for software developer time at ETH. We maintained our own logs of real-time reports for the BDSN and Menlo Park installations.

Southern California installation

In this document, we provide an update of the real-time performance analysis of the VS algorithm at the SCSN, spanning the period July 2008 through October 2009 (6 months longer than the period described in the SRL paper).

Figure 1 shows the real-time VS detections and missed events (relative to the SCSN catalogue) during the 15 month study period. Open blue circles are real-time VS detections with $M < 3$ (1970 events), filled green squares are real-time VS detections with $M \geq 3$ (182 events), and open red squares are missed $M \geq 3.0$ events. There

are concentrations of missed events south of the California-Mexico border, and northwest of the northern edge of the SCSN AOR. It should be noted that, in the SCSN catalogue, many of these missed (by the VS algorithm) events are lower quality solutions. Within the interior of the network, there are very few missed $M > 3$ events. Generally, we can expect the VS algorithm to perform well in regions where the network also performs well. That is, in a given region, there is a strong correlation between the quality of the EEW and the network solutions.

Based on real-time performance data collected over the study period, the initial VS estimate time is normally distributed about a mean value of 20.1 seconds, with a standard deviation of 5.9 seconds (Figure 2). This initial estimate time is the time relative to the earthquake origin time at which a VS estimate could have been sent out, were this an operational system. It is an alert time (definition in March 2008 EEW Testing Document), and thus includes the effects of telemetry delay, processing time, and algorithm time. Note, VS algorithm time is the time required for P-waves to be detected at 4 stations.

Figure 3 shows a map of all real-time VS detections for which reports were sent to the SCEC EEW Testing Center. Our criteria for sending a report to SCEC was that the initial VS magnitude estimate was $M = 2.7$ or greater (since we expected that the initial VS magnitude estimate would be a lower bound for the actual magnitude). A total of 421 reports were sent to the SCEC EEW Testing Center over the 15 month study period. 351 out of 421 of these reports matched an entry in the SCSN catalogue. 70 out of 421 of these reports were false alerts that did not match an entry in the catalogue. 24 of these false alerts were withdrawn by the VS algorithm (identified as “phantom events”) once additional data was available.

While these numbers can be used to derive false and missed alert rates, there is expected to be a strong time-dependence of these false and missed alert rates. It is likely that many of the false and missed alerts are found towards the beginning of our real-time operations (ca. July 2008).

The performance statistics on the various algorithms collected by the SCEC EEW Testing Center is an alternative data source for deriving VS false alert rates. Based on data collected by the SCEC EEW Testing website, the false alert rate of VS at the $M>3.0$ level over the last 6 months is 43 false triggers / (43 false triggers + 534 correct triggers) = .074 (or 7.4% probability of a false alert given a real-time VS report with $M>3.0$) and at the $M>4.0$ level is 5 false triggers / (5 false triggers + 132 correct triggers) = .036 (or 3.6% probability of a false alert given a real-time VS report with $M>4.0$). Missed alert rates cannot be derived from the SCEC data since the SCEC “Missed Triggers” take into account both northern and southern California events, while the VS algorithm only submitted real-time reports for southern California.

Characterizing VS estimate uncertainty in Southern California

We examined the VS magnitude and epicentral location errors as a function of time for $M\geq 3.0$ events in regions where we expect reasonably good VS performance (ie. excluding regions on the outskirts or outside of the network) (Figure 4). The mean magnitude error tends to increase with time, since we currently do not take into account site amplification effects. Any site amplification effects are currently mapped into larger magnitude estimates. However, the mean magnitude error for the 25th VS estimate is quite reasonable, at approximately 0.1 magnitude unit. The standard deviation σ of the VS magnitude estimate also has some time dependence, with $\sigma=0.23$ for the 1st VS estimate, and $\sigma=0.15$ for the 25th VS estimate. (VS reports typically have 20-30 updates of magnitude, location, and peak ground motion for a given event.). In general, we found that the VS magnitude errors are well-described by a Gaussian distribution.

Epicentral location errors are log-normally distributed. Most VS location estimates are within 10 km of the network epicenter (Figure 6). However, VS location

estimates can be more than 15 or 30 km away from the network epicenter at the outskirts and outside of the network. That VS location estimates are not as good in these regions compared to performance in the interior of the network is expected, since station coverage is lower.

Statewide implementation

The VS codes are currently running statewide via 3 separate installations at SCSN, BDSN, and Menlo Park.

Northern California installation

Figure 7 summarizes VS real-time performance in Northern California between March and October 2009. The VS installations at BDSN and Menlo Park generated 72 real-time reports during this period (not sent to the SCEC EEW Testing Center). Similar to the SCSN installation, the generation criteria is that the initial VS magnitude estimate is $M \geq 2.7$. 51 of these real-time reports matched entries in the ANSS catalogue (within ± 1 magnitude unit). At the $M \geq 3.0$ level, 26 VS reports matched an ANSS catalogue entry, and 14 VS reports had no corresponding match. Based on these values, VS has a $14 / (14 + 26) = .35$ false alert rate at the $M \geq 3.0$ level (or 35% probability of false alert given a VS $M \geq 3.0$ report) in Northern California. At the $M \geq 4.0$ level, 2 VS reports matched ANSS catalogue entries, and 4 did not. Thus, VS has a $4 / (2 + 4) = .66$ false alert rate at the $M \geq 4.0$ level (or 66% probability of false alert given a VS $M \geq 4.0$ report) in Northern California. In contrast, the false alert probabilities are 7.4% and 3.6% at the $M \geq 3.0$ and $M \geq 4.0$ levels, respectively, in Southern California.

The mean time to the initial VS estimate in Northern California is 31 seconds, with a standard deviation of 12 seconds (Figure 8). This is in contrast to the mean initial estimate time of 21 seconds, with 6 second standard deviation, in Southern

California. This difference is due primarily to differences in station density and network topology.

Figure 9 shows histograms of the initial and final VS epicentral location errors. Most initial VS location estimates are within 15 km of the ANSS epicenter.

Figure 10 shows the error of the initial and final VS magnitude estimates as a function of the ANSS magnitude. The initial VS magnitude estimates have a mean error of 0.12 and standard deviation of 0.33 magnitude units. The final VS magnitude estimates have a mean error of 0.17 and standard deviation of 0.37 magnitude units. As with the SCSN installation, the Northern California installations do not correct for site amplification effects. Any site amplification effects are mapped to larger magnitude estimates. The standard deviations of the VS magnitude errors are larger in Northern California than in Southern California.

The Northern California installations are relatively new, and performance levels are currently not comparably to the more mature SCSN installation. Potential reasons for the discrepancies in the performance are: 1) the BDSN and Menlo Park installations are separate installations, and currently do not exchange or share data, 2) the Menlo Park network is composed of only strong motion stations. The noise level at these stations is relatively high, 3) network geometry, with a large cluster of stations near San Francisco and the Bay Area, and large regions with sparse station coverage. It is also possible that sampling issues come into play, since the SCSN installation has been running for twice as long as the Northern California installations.

Combining the data from the BDSN and Menlo Park networks should improve VS performance in northern California and will be addressed in the next stage of the project.

Future enhancements

Among the future enhancements related to the VS methodology that will be undertaken as part of the CISN ShakeAlert project are:

- Multiple-threshold implementation allowing for generation of alerts with fewer stations (current VS minimum number of stations is 4) if signals are large enough
- Including prior information (previously observed seismicity and Gutenberg-Richter relationship, to start with)
- Correcting for site amplification effects
- Using network Ml relationship once peak S-wave amplitudes are available at a station, with the goal of improving the convergence of EEW information to the network solutions

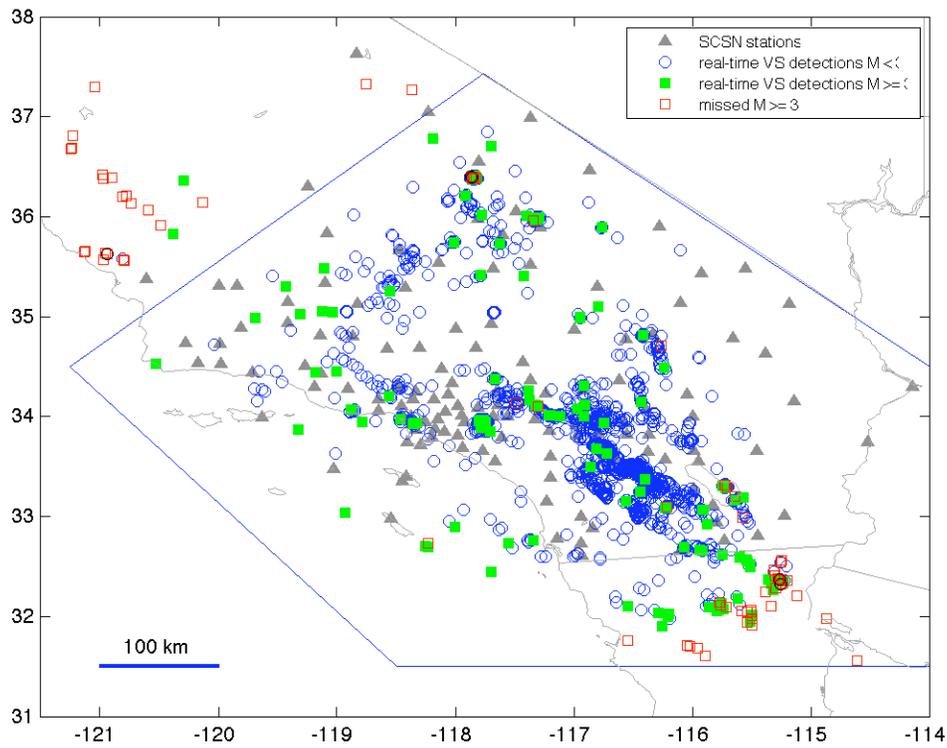


Figure 1: VS real-time performance in Southern California from July 2008 through October 2009. The blue polygon encloses the SCSN area of responsibility (AOR).

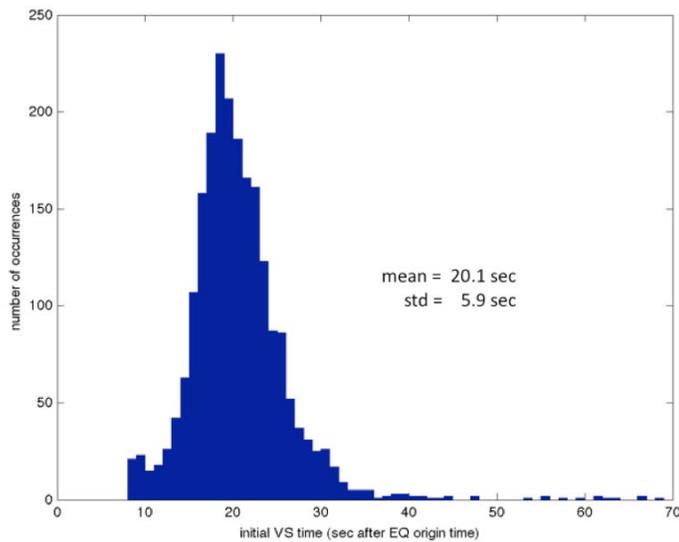


Figure 2: Histogram of the initial VS estimate time. The initial VS estimate time includes the effects of telemetry delay, processing time, and algorithm time.

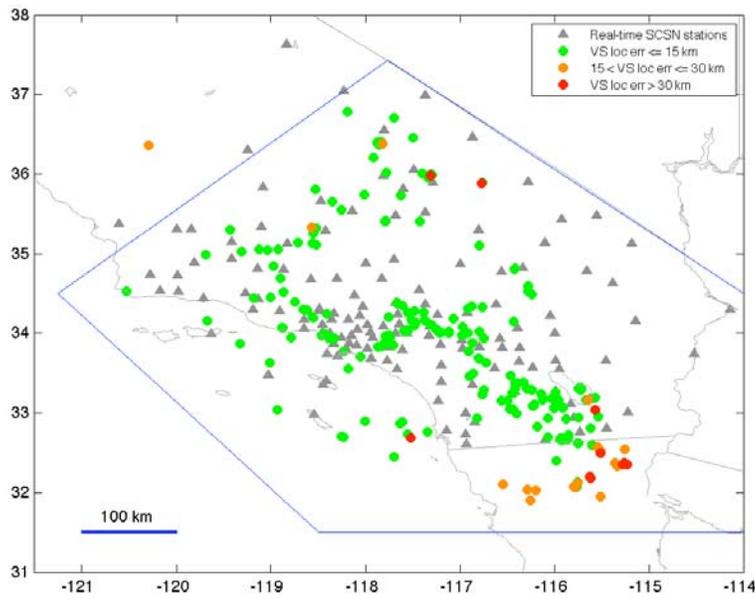


Figure 3: Map of real-time VS detections with $M > 2.7$, color-coded according to epicentral location error of the initial VS location estimate. VS location estimates in the interior are relatively good, and decrease in quality at the edges of the network.

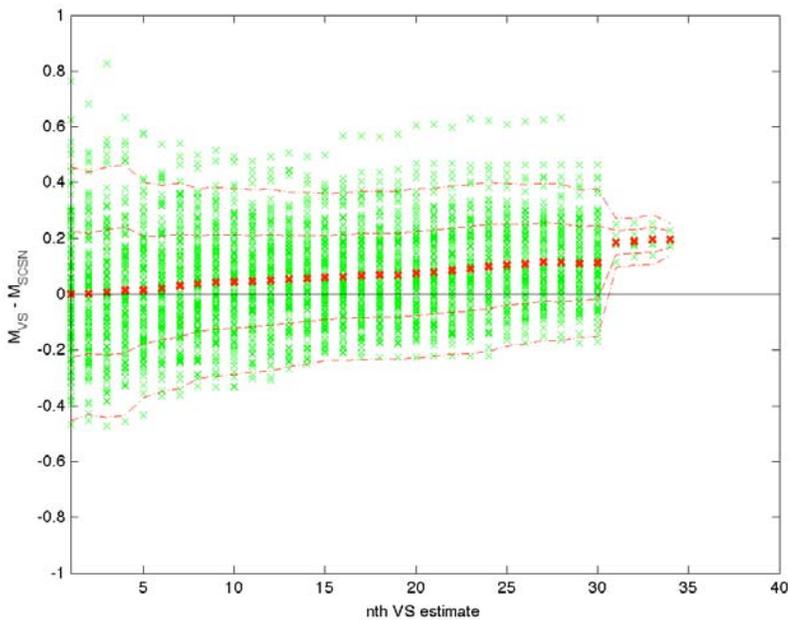


Figure 4: The VS magnitude error as a function of rank of VS estimate (ie, 1st estimate, 2nd estimate, etc) for 137 $M \geq 3.0$ events in regions where we expect reasonable VS performance (green circles in Figure 3). The red crosses denote the mean magnitude error for the nth VS estimate. The red dashed lines denote the $\pm\sigma$ and $\pm 2\sigma$ levels.

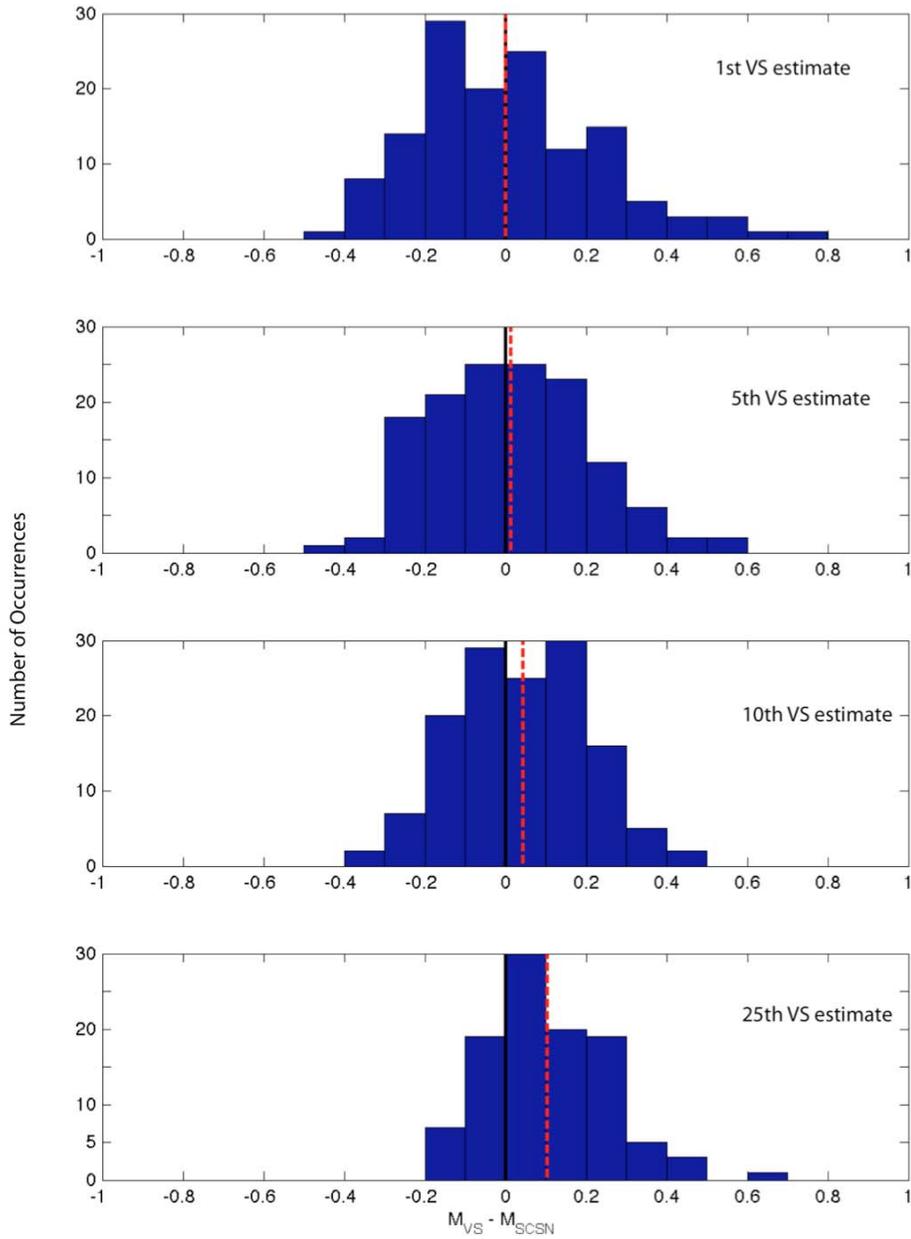


Figure 5: Histograms of VS magnitude error for the 1st, 5th, 10th, and 25th VS estimate. The thick black line denotes 0 error, while the red dashed line is the mean VS magnitude error at a given time.

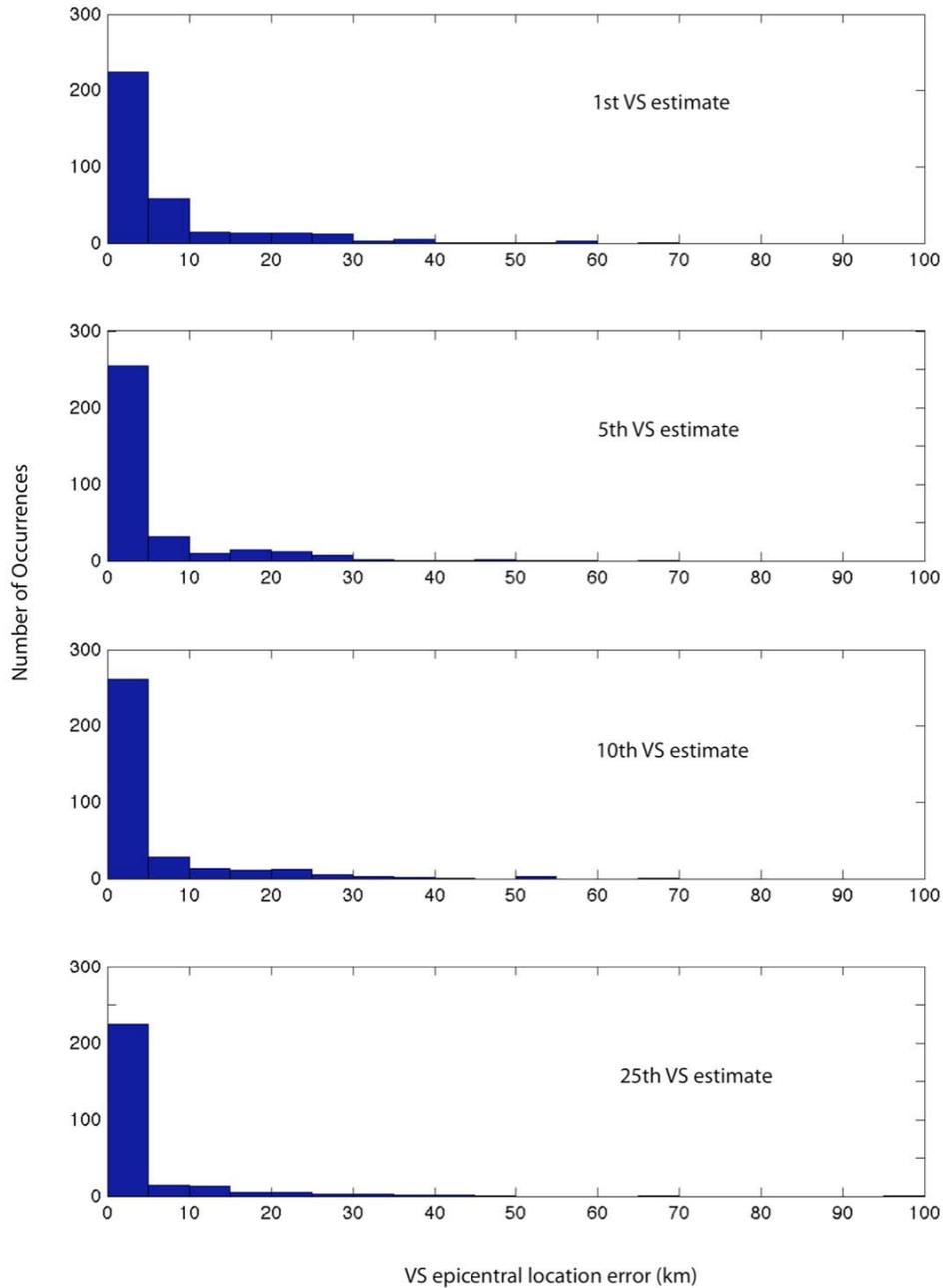


Figure 6: Histogram for VS epicentral location error (for all $M \geq 3.0$ events) for the 1st, 5th, 10th, and 25th VS estimates. VS epicentral location errors are log-normally distributed. Most location estimates are typically within 10 km of the network epicenter. The larger epicentral location errors are from events on the outskirts or outside of the network (red circles in Figure 3), such as northwest of the northern edge of the SCSN AOR, and south of the California-Mexico border.

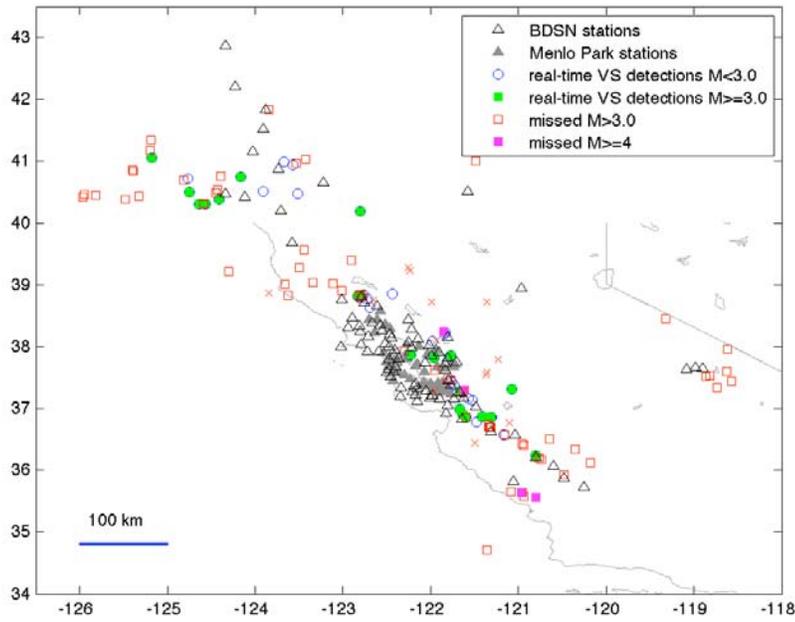


Figure 7: VS real-time performance in Northern California from February through October 2009.

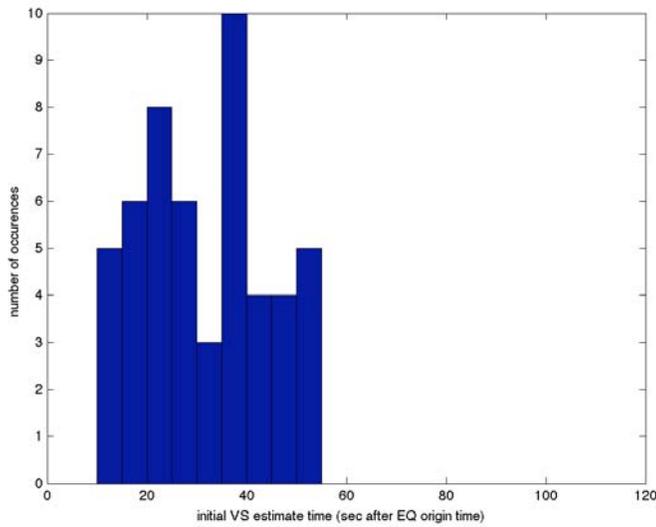


Figure 8: Histogram of initial VS estimate time in Northern California. The mean initial VS estimate time is 31 seconds, with a standard deviation of 12 seconds.

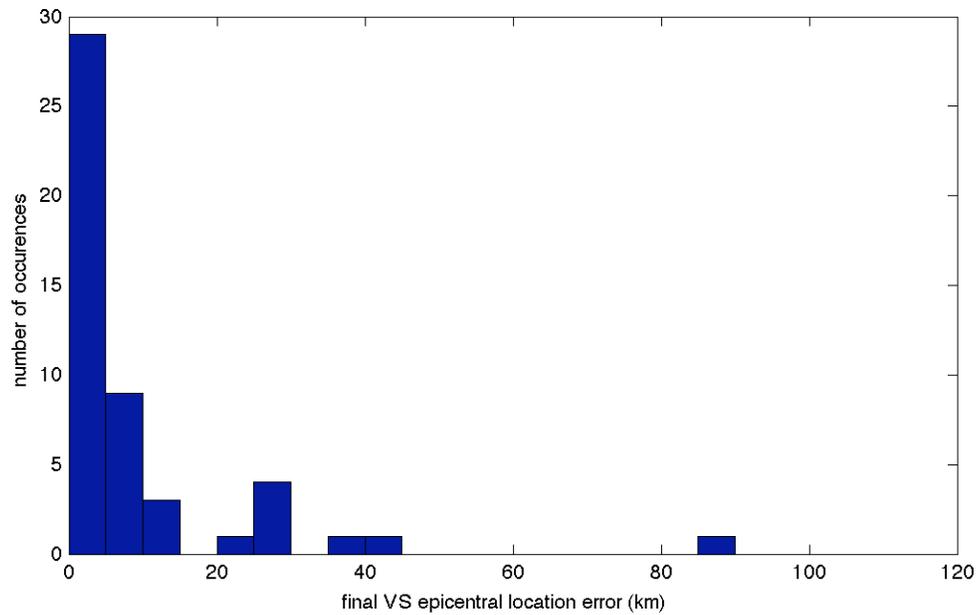
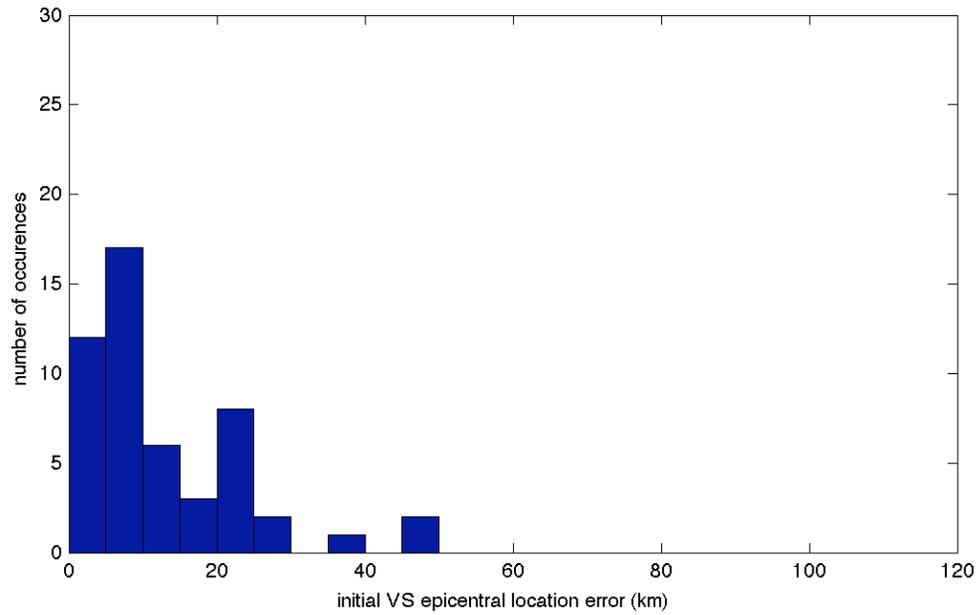


Figure 9: Histograms of initial and final VS epicentral location errors. Most VS location estimates are within 15 km of the ANSS epicenter.

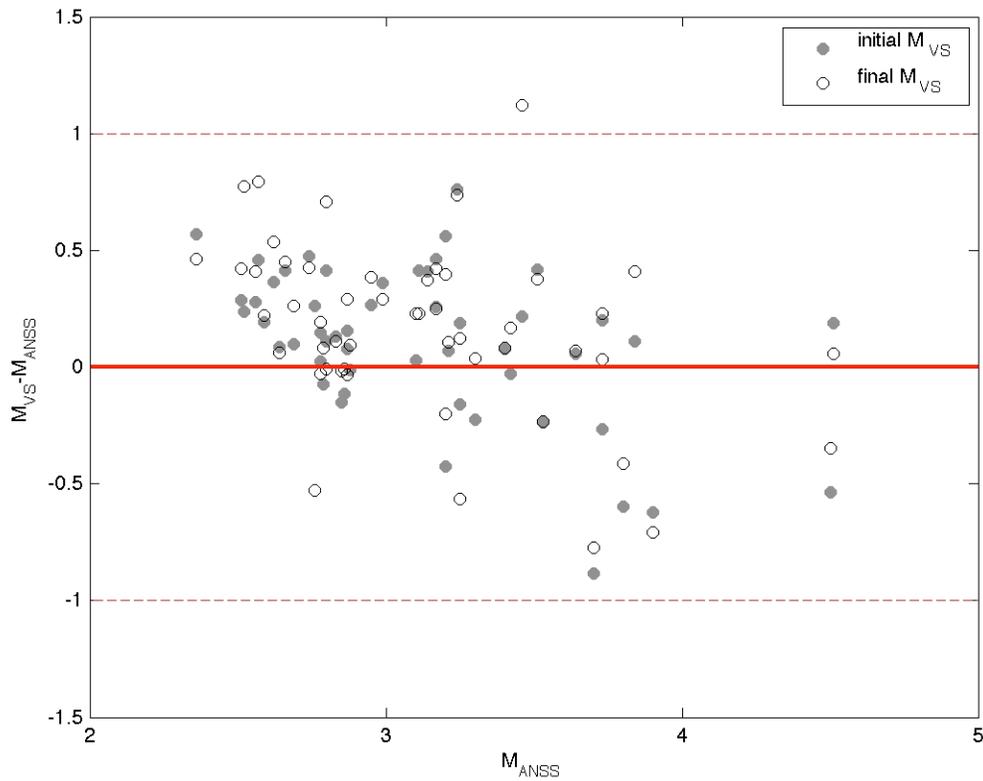


Figure 10: Initial and final VS magnitude error in Northern California as a function of the catalogue magnitude. The red dashed lines indicate the ± 1 magnitude unit acceptable error criteria of the SCEC EEW Testing Center. There are systematic offsets in the mean error levels, due to (thus far) uncorrected site amplification effects. Standard deviations of the VS magnitude error are slightly higher in Northern California than in Southern California.

Research on Finite Faults: Early Warning for Large Magnitude Earthquakes

**Maren Böse, Tom Heaton, Masumi Yamada,
California Institute of Technology**

Introduction

The majority of current approaches developed for Earthquake Early Warning (EEW) around the world considers the earthquake as a point source, i.e. neglects rupture finiteness. Large magnitude earthquakes ($M >> 7.0$), with rupture lengths of up to hundreds of kilometers, cause damaging ground shaking in much larger areas than moderate to strong events. Despite their rare occurrence, many more people and users could benefit from an EEW system during large earthquakes (Heaton, 1985; Allen, 2006). Due to the low propagation speed of seismic ruptures (~80% of the seismic S-wave speed), warning times to heavily shaken areas may exceed more than one minute, while during moderate earthquakes users will usually receive warnings of a few seconds only.

The suitability of an EEW system for large earthquakes in California was recently demonstrated by the USGS and partners for the Great California ShakeOut M7.8 scenario earthquake with a seismic rupture starting at the Southern portion of the San Andreas Fault (SAF) and then propagating over a distance of ~300 km towards North-Western direction (Jones *et al.*, 2008). Warning times for large cities and communities along the SAF, such as Palm Springs, San Bernardino or Los Angeles, would be in the order of some tens of seconds to more than one minute (Figure 1).

A key to EEW for large magnitude earthquakes is the rapid prediction of the final dimensions of the ongoing rupture. This is because the level and distribution of seismic ground motions, such as of peak ground velocity (PGV) or seismic intensity, are largely controlled by the earthquake magnitude and the rupture-to-site distance, which both can be estimated from the predicted rupture dimensions. The rupture of a large earthquake propagates mainly in the horizontal direction along a fault, i.e. rapid prediction of the rupture length is especially important for EEW. Ideally, these predictions are given in a probabilistic manner to allow for making appropriate decisions when available data is sparse and the prediction of source and ground motion parameters uncertain.

Location	Seconds after start of earthquake that strong shaking begins at this location
Palm Springs	25
San Bernardino	45
Los Angeles (downtown)	70
Orange County	70
Santa Monica	85
Palmdale	75
Ventura	105

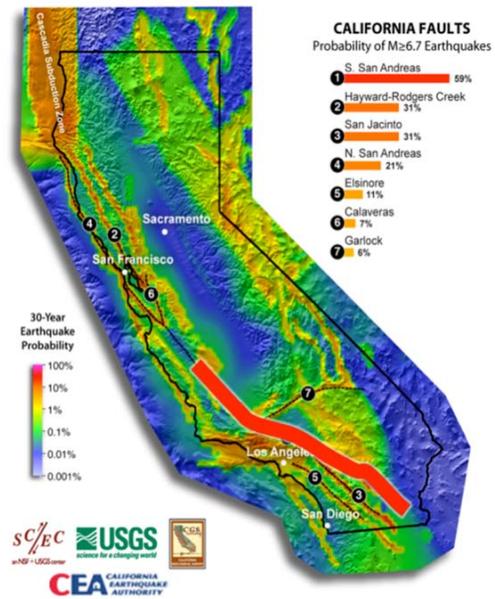


Figure 1. Expected warning times for the Great Southern California ShakeOut M7.8 scenario earthquake (courtesy of L. Jones, 2008).

During the past three years we have carried out several studies to develop methods allowing for the real-time estimation of fault rupture extend

1. using near-source vs. far-source classification (Yamada *et al.*, 2007)
2. using envelopes of acceleration (extension of the Virtual Seismologist for finite faults by Yamada and Heaton, 2008)
3. using slip amplitudes and the “smoothness” of the causative fault (Böse and Heaton, in prep.)

In the following we will give a brief summary of the main results of these studies. References are given for further details.

1. Real-Time Estimation of Fault Rupture Extent Using Near-Source versus Far-Source Classification

To estimate the fault dimension of an earthquake in real-time, Yamada *et al.* (2007) developed a methodology to classify seismic records into near-source or far-source records. Characteristics of ground motion, such as peak ground acceleration, have a strong correlation with the distance from a fault rupture for large earthquakes. This study analyzes peak ground motions and finds the function that best classifies near-source and far-source records based on these parameters. Yamada *et al.* (2007) performed (1) Fisher's linear discriminant analysis and two different Bayesian methods to find the coefficients of the linear discriminant function and (2) Bayesian model class selection to find the best combination of the peak ground-motion parameters. Bayesian model class selection shows that the combination of vertical acceleration and horizontal velocity produces the best performance for the classification (Figure 2). The linear discriminant function produced by the three methods classifies near-source and far-source data, and in addition, the Bayesian methods give the probability for a station to be near-source, based on the ground-motion measurements. This discriminant function is useful to estimate the fault rupture dimension in real-time, especially for large earthquakes.

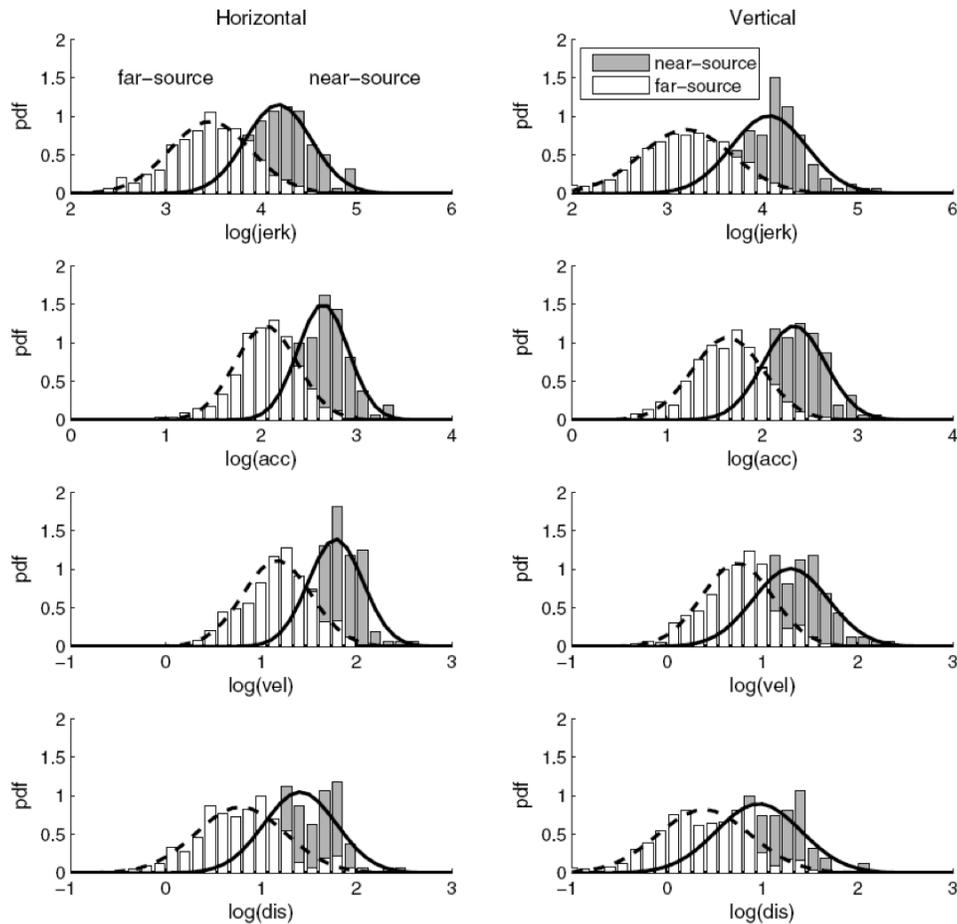


Figure 2. Histograms and Gaussian densities based on the sample means and standard deviations of the log of ground motions for the near-source and far-source records. These are distributions for jerk, acceleration, velocity, and displacement from the top (Yamada *et al.*, 2007).

2. Real-Time Estimation of Fault Rupture Extent Using Envelopes of Acceleration

Yamada and Heaton (2008) proposed a new strategy to estimate the geometry of a rupture on a finite fault in real-time for EEW as an extension of the Virtual Seismologist (VS) method. They developed a new model to simulate high-frequency motions from earthquakes with large rupture dimension: the envelope of high-frequency ground motion from a large earthquake can be expressed as a root-mean-squared combination of envelope functions from smaller earthquakes (Figure 3). They used simulated envelopes of ground acceleration to estimate the direction and length of a rupture in real time. Using the 1999 Chi-Chi earthquake dataset, they ran simulations with different parameters to discover which parameters best describe the rupture geometry as a function of time. They parameterized the fault geometry with an epicenter, a fault strike, and two

along-strike rupture lengths. The simulation results show that the azimuthal angle of the fault line converges to the minimum uniquely, and the estimation agrees with the actual Chi-Chi earthquake fault geometry quite well. The rupture direction can be estimated at 10 s after the event onset, and the final solution is achieved after 20 s. While this methodology seems quite promising for warning systems, it only works well when there is an adequate distribution of near-source stations.

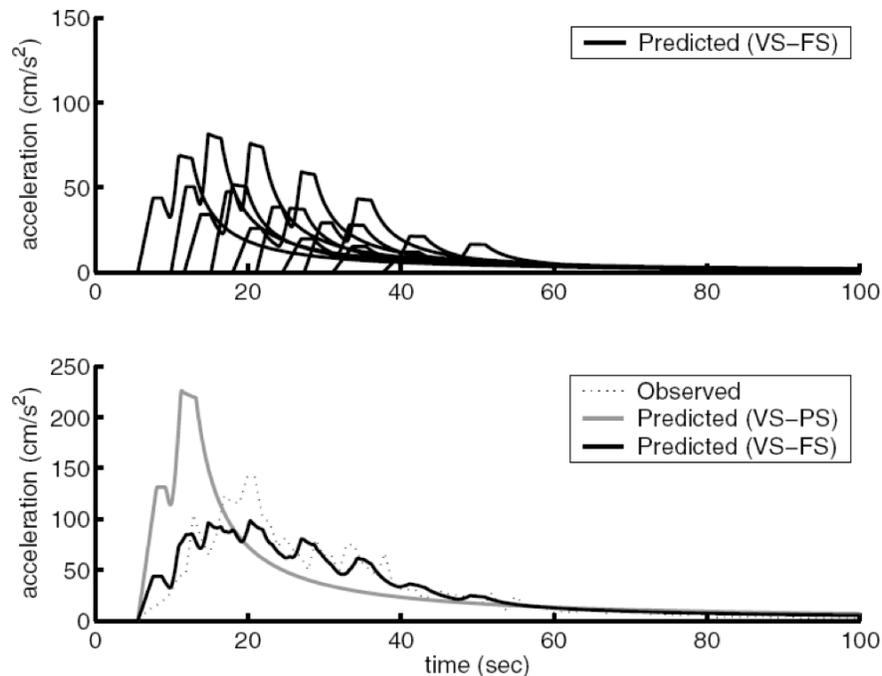


Figure 3. Envelopes of vertical acceleration recorded at the station C024 for the Chi-Chi earthquake. Top: Predicted envelopes of the vertical acceleration record for each subsurface with M 6.0. Bottom: Observed envelope (in dotted black line) and predicted envelopes of the point source model in VS-PS method (in solid gray line) and of the multiple source model in VS-FS method (in solid black line) (Yamada and Heaton, 2008).

3. Real-Time Estimation of Fault Rupture Extent Using Slip Amplitudes and “Smoothness” of the Causative Fault

Predicting the shaking from large earthquakes requires some estimate of the likelihood of the future evolution of an ongoing rupture. An EEW system that anticipates future rupture using the current magnitude together with the Gutenberg-Richter frequency-size statistics, e.g., will likely never predict a large earthquake, because of their rare occurrence. However, it seems to be a reasonable assumption that (1) an ongoing rupture with a large present slip amplitude D_p is more likely to rupture a large distance than an event with a small D_p , and (2) an earthquake that

is rupturing along a mature through-going fault (e.g., the San Andreas Fault, SAF) is more likely to continue rupturing than an event on a short fault with little total geologic offset.

Böse and Heaton (in prep.) investigated how this information may be used to estimate the eventual size of an ongoing rupture using estimates of D_p and the “smoothness” of the causative fault. They simulated suites of evolving ruptures using a 1-D stochastic model of spatially heterogeneous slip. They found that while large slip amplitudes D_p increase the probability for the continuation of a rupture and the possible evolution into a large earthquake, the recognition that rupture is occurring on a spatially smooth fault could have an even stronger effect (Figure 4). They concluded that an EEW system for large magnitude earthquakes should have some mechanism for recognizing a rupturing fault (e.g., is the earthquake on the SAF?). This information can be obtained, e.g., from real-time GPS measurements.

Böse and Heaton (in prep.) also demonstrated, how probabilistic estimates of rupture length and future slip can be used for a probabilistic prediction of seismic ground motions along the evolving rupture. An example is shown in Figure 5.

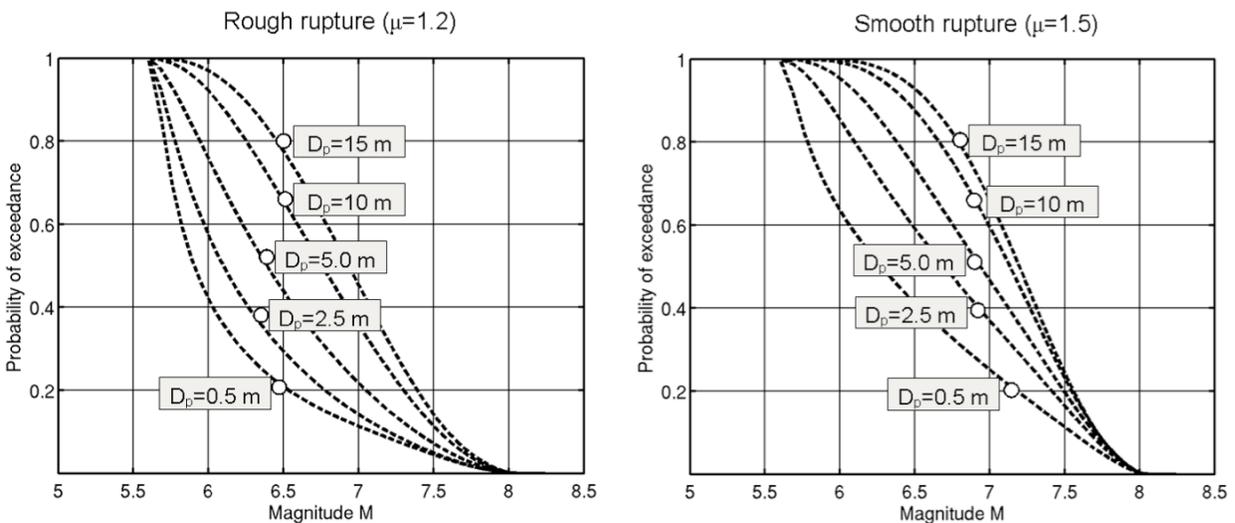


Figure 4. Probability of exceedance of different magnitude levels M for an ongoing rupture with a present magnitude of $M_p=5.5$ and present slip amplitude D_p . The probabilities were derived from a stochastic slip model after Liu-Zeng *et al.* (2006) for rough (left) and smooth ruptures (right) on generic and mature faults, respectively (Böse and Heaton, in prep.).

The majority of the algorithms, which are currently applied to EEW, use a fixed time window of a couple of seconds of the seismic P wave for a rapid estimation of the earthquake magnitude, using, e.g., the average or the predominant period of shaking (e.g. Nakamura, 1988; Allen and Kanamori, 2003; Kanamori, 2005; Wu *et al.*, 2007). This approach is heavily disputed, because the rupture process of earthquakes with $M > 6.0$ take much longer than 3 seconds and it appears questionable, why the final rupture length and therewith the magnitude should be predetermined at this early stage of the rupture process (e.g., Rydelek and Horiuchi, 2006; Rydelek *et al.*, 2007). Only if a large amount of the seismic energy of the earthquake is radiated at the initial stage, i.e. only if seismic ruptures start in patches of high seismic slip (as proposed, e.g., by Mai *et al.*, 2005), it might be feasible to give a good magnitude prediction within a couple of seconds after the rupture initiation. If this is not the case, the magnitude might be underestimated, if we use a limited time window of a few seconds.

The statistical analyses carried out by Böse and Heaton (in prep.) demonstrate that it seems appropriate to assign higher probabilities that the rupture on a mature fault (such as the SAF) will continue and finally evolve into a “Big One”, than on a generic fault, for which the probability that the rupture will continue is quite low. This implies that the observation of an (apparent) moderate earthquake on a mature fault (judging from the first few seconds of waveform data) should result in a stronger warning than for a similar observation made for a generic fault.

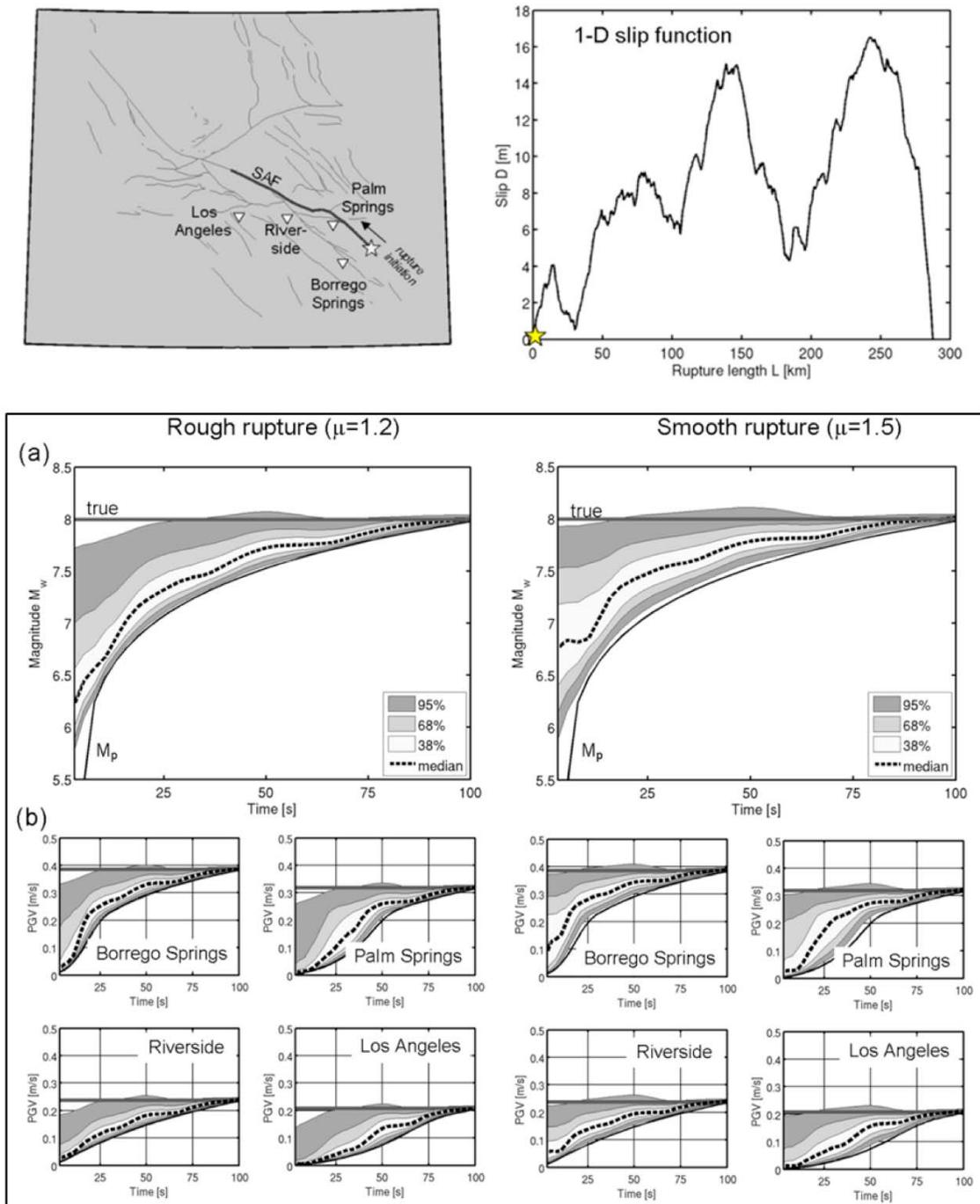


Figure 5. Example of a M8.0 scenario earthquake along the (mature) San Andreas Fault. Slip amplitudes are used for the real-time probabilistic prediction of (a) magnitude and (b) peak ground velocity along the evolving rupture. Shown are the 38%, 68% and 95% confidence intervals. There are clear differences in the prediction depending on the assumed slip heterogeneity of the causative fault: predictions for a generic fault are always more conservative than for a mature fault (Böse and Heaton, in prep.).

References

- Allen, R.M. (2006). Probabilistic warning times for earthquake ground shaking in the San Francisco Bay Area, *Seismo. Res. Lett.* 77 (3), 371-376.
- Böse, M. and T. H. Heaton (in prep). Probabilistic Prediction of Rupture Length, Slip and Seismic Ground Motions for an Ongoing Rupture – Implications for Early Warning for Large Earthquakes.
- Heaton, T. (1985). A model for a seismic computerized alert network, *Science* 228 (4702), 987-990.
- Jones, L.M., R. Bernknopf, D. Cox, J. Goltz, K. Hudnut, D. Mileti, S. Perry, D. Ponti, K. Porter, M. Reichle, H. Seligson, K. Shoaf, J. Treiman, and A. Wein (2008). The ShakeOut Scenario, U.S. Geological Survey Open File Report 2008-1150, California Geological Survey.
- Yamada, M., T. Heaton, and J. Beck (2007). Real-Time Estimation of Fault Rupture Extent Using Near-Source versus Far-Source Classification, *Bull. Seismol. Soc. Am.*, Vol. 97, No. 6, pp. 1890–1910, doi: 10.1785/0120060243.
- Yamada, M. and T. Heaton (2008). Real-Time Estimation of Fault Rupture Extent Using Envelopes of Acceleration, *Bull. Seismol. Soc. Am.*, Vol. 98, No. 2, pp. 607–619, doi: 10.1785/0120060218.

1. CISN EEW TESTING CENTER DEVELOPMENT

As part of the CISN EEW algorithm development, the CISN and the Southern California Earthquake Center (SCEC) have established an independent an EEW algorithm performance testing system. The CISN EEW testing center addresses several important aspects involved in evaluating the seismological and speed of operation performance for the CISN EEW algorithms. Development of a common testing environment has helped the CISN EEW algorithm developers standardize on specific testing issue such as the region for which CISN EEW should be tested, and the types of performance summaries that should be produced on a nightly basis. The CISN EEW testing activity has helped to define common algorithm performance information such as standardized timestamps, and standardize algorithm parameters (e.g. origin time, estimated magnitude). The CISN EEW testing center activity has also developed standardized data reporting formats.

The CISN EEW testing center is designed to perform prospective testing in order to eliminate the possibility of bias towards a known correct answer which is possible in retrospective testing. In the case of EEW, prospective testing requires that the CISN EEW algorithms report (or log to a file) predicted parameters such as final magnitude while the event is in progress. Currently both the CIT TauC-PD and the Virtual Seismologist algorithms report prospective predicted parameters to the testing center. The UCB ElarmS currently reports retrospective information to the testing center, although prospective information may be available to the UCB development groups.

2. CISN EEW TESTING CENTER OPERATIONS:

The CISN EEW testing center has now been in operation for over one year. On a nightly basis, the CISN EEW Testing Center software retrieves current ANSS catalog information for California, and compares CISN EEW triggers against events in the ANSS catalog. These comparisons are used to produce CISN EEW performance summaries which are then automatically posted on the CISN EEW Testing Center web site. (<http://www.scec.org/eeew>) (Username: guest) (pwd: cisnew). The performance summaries on the CISN EEW web site are updated on a daily basis at about midnight PST.

In the CISN EEW performance summaries, the ANSS catalog is used as the reference data set. CISN EEW trigger information, including origin time, location, and magnitude, are compared against the ANSS catalog information. The CISN EEW testing center software framework that runs on a nightly basis is derived from an open-source software framework created at SCEC for the Collaboratory for the Study of Earthquake Predictability (CSEP) (<http://www.scec.org/csep>). Use of the CSEP software has helped reduce the required software development needed to support the CISN EEW testing center.

3. CISN EEW PERFORMANCE SUMMARY RESULTS:

The CISN EEW testing center provides a variety of performance summaries as specified in a CISN EEW Testing specification document [Bose et al 2008]. These performance summaries are based on two input data sets. One, performance reports are created by the real-time or near real-time, algorithms and transferred to the CISN EEW testing center. Second, the CISN EEW testing center retrieves earthquake catalog data from the ANSS catalog. This ANSS data is consider the reference data. EEW forecasts are compared against the ANSS to determine their accuracy.

These CISN EEW performance summaries are produced automatically on a daily basis. An example performance summary that is available on the SCEC web site is a trigger summary. An example trigger performance summary for events > M14.0 is shown below. Information that can be gleaned from this summary includes the following.

Yellow rows indicate ANSS events. Green rows following an ANSS event indicate EEW triggers for that event that contained accurate information (such as event origin time and magnitude). This example table shows several ANSS events with EEW triggers. It also shows that for several ANSS event, in both NC and SC, there were ANSS M13.0+ events with no EEW triggers.

For the first ANSS event (M13.9), this table shows that ElarmS triggered on the event. ElarmS produces a timeseries of information during the event. The multiple green rows show that the ElarmS information about the event evolved with time. ElarmS reports only Algorithm time to the testing center. With this time, ElarmS reported good (within 1 Magnitude unit and -10 to +30 seconds) origin time and magnitude information with 11 seconds of network data.

EEW algorithm changes were not rigorously tracked during this project. This should be changed in future testing efforts. During this project, several significant EEW algorithm changes were implemented but the performance measures do not indicate when those changes were made.

In our current testing results, we address this by limiting the length of time for which we examine performance. While not completely rigorous, we suggest that the algorithms under test were reasonably stable during the last year of the project (10/1/2009 – 9/30/2009). The system contains some performance measures prior to this time; however, the results from earlier time periods will probably not be representative of the algorithms performance near the end of the project.

Need to Track Active Channels:

As we evaluated the EEW performance metrics, we recognized that performance of the algorithms depends on which stations are being used by the EEW algorithms. For some metrics, such as “missed events”, the metrics isn’t meaningful unless the “active channels” are known. Ideally, to compare the “missed event” performance for two algorithms, both algorithms would use the same active channels. If they do not use the same list of channels, the performance may not be comparable.

In the current implementation of the testing center, we used a manual procedure in which EEW algorithms sent a list of active channels to the testing center. Our experience is that this is not a reliable and efficient approach. The active channel lists were often out of date. A new, improved, automatic reporting of active channels in use by each EEW algorithm would improve the testing center.

A related issue to this we encountered is the reporting of site specific information. ElarmS forecast site specific intensity at sites where CISN operated seismic stations. The list of sites for which information was reported was manually coordinated between UCB and SCEC. Again, automating this type of configuration information is important in the future. We found that the lists of reports were out of sync with actual CISN network so that we often received site-specific information for stations that were no longer operational, or that we did not receive site-specific reports for new stations.

Universal Metrics versus Catalog-specific Metrics:

During this development, we recognized that some EEW metrics are “event specific” and some are not. For example, a metrics of clear interest is “how long the system takes to produce a warning” (i.e. we call this the Warning Delay). It turns out that we cannot provide a universal answer to this. We can measure how an EEW system performs for a series of earthquakes and we can estimate how it would perform for scenario earthquake. But this type of metric must always be related to specific earthquakes.

We have learned that as we define metrics, we should consider whether these metrics are “Universal” or whether they must be measured only for a specific set of earthquakes.

Defining a Testing Region and a Service Region:

We determined that some EEW metrics of interest require agreement on a testing region. In order to define, for example, the rate of missed events, the scientific group must agree on a testing region. For the CISN, the state of California might serve. However, it is possible that a subset of this region would be selected. For the performance reports given in this summary, we used a Testing Region definition that included all of California and a small additional buffer region around the state which matches the California Testing region used in RELM and CSEP testing.

In a similar way, once an EEW system is distributing warnings, the region for which warnings are to be distributed must be defined. The CISN groups should consider the goals of the EEW development and then define both a CISN EEW Testing Region and a CISN EEW Service Region for future testing.

Event Metrics versus Ground Motion Metrics:

During our testing center developments, we recognized the importance of distinguishing between Event Metrics and Ground Motion Metrics. For example, for a trigger, and EEW algorithm may forecast the final magnitude (an event metric). An algorithm may also forecast site-specific information such as peak intensity.

Event metrics are easier to track. Site-specific metrics are problematic particularly due to the infinite number of sites for which information might be reported. If site-specific metrics are measured in the future, a list of sites of interest needs to be coordinated. Reporting of site-specific information at seismic station locations is probably a useful approach. In this case, we could use the seismic recordings to confirm the site-specific forecasts.

Reporting of EEW Forecast Parameters as Time Series:

The two network oriented algorithms (ElarmS and Virtual Seismologist) report their forecast parameters (e.g. final magnitude) as time series. Initial reports are calculated, logged, then updated on a 1 second, or 5 second basis. This presents challenges to the testing center as we try to evaluate the accuracy of the predicted parameters. The key question is “which reported value do we use in evaluation?” After the event, we can evaluate the time series, and determine at what time it begins to report accurate information.

The testing center (and users) does not have any way to evaluate whether a report is “good enough” during the event. We propose to address this by placing the burden of filtering time series reports onto the EEW systems. To address this, we believe the testing center will need to evaluate each sample in the time series as its own report. If the reported data rate is too high, we can evaluate reports on a decimated time series. Then, assume each sample in the time series is considered a separate report, metrics such as false alarms and magnitude or ground motion accuracy will include measurements from the first samples received. This places the burden of delaying the time series reports until they meet quality expectations onto the EEW systems, and this removes any responsibility on the part of the testing center to determine which sample in the time series are “good enough” to use in metrics.

Defining False Alarms and Missed Events:

We anticipate that EEW system users will want to know how reliable the CISN EEW system is using metrics such as false alarms and missed events. We have learned that these metrics require significant scientific qualification in their definitions.

To begin, we need to decide whether we are going to use event-specific information, or site-specific information. An “event-specific” definition for a false alarm might state that the EEW system will not declare any triggers unless there is an event in the testing region. A “site-specific” definition for false alarms might state that the EEW system will not declare a ground motion alert for a specific site unless strong ground motion occurs at the site. Event specific definitions are simpler; however, site-specific information may be more valuable to system users. Similarly, missed events can be defined in terms of earthquakes, or site-specific ground motions.

There are further complications in these definitions. For example, if a false alarm metric is used, an algorithm must be used to match ANSS events with CISN EEW triggers. A trigger without an event is a false alarm. However, this matching process is non-trivial. For example, we may get a trigger in San Diego at the same time there is an earthquake in San Francisco. Even though the time is right, the San Diego Trigger may be a false alarm. What tolerance in time and location should we give as we try to classify a trigger as “real” or a “false alarm.” As we continue EEW testing, we should clarify our definition of these measures.

Further CISN scientific agreement on appropriate definitions for these measures should be developed. We report on both false alarms and missed events in the current testing center. However, the definitions for these metrics are not clearly and widely known.

False alarm definitions seem like they will require time, location, magnitude tolerances. For false alarm rates to be comparable, these tolerances must be known. Also, false alarm metrics need to include an active channel list. False alarm rates will not be comparable unless the active channels in use are also known.

We also note that false alarm metrics trade-off with metrics about the accuracy of seismological parameters. If an algorithm triggers on something seismic, over-estimates the magnitude, and produces an alert, we must decide how to classify this. Is this a false alarm? Or is this a valid trigger, and an inaccurate magnitude estimate? In the future, we recommend that false alarm definition moves towards a more seismological meaningful definition. We suggest that if seismic ground motions cause a trigger, it is a correct trigger, and is not considered a false alarm. Implementing this definition may be challenges as we'd need to review waveforms, but such a metric would probably be more meaningful than current approach.

Need to Access Waveforms:

The current testing center uses ANSS event catalog as “reference” data and compares CISN EEW triggers against this data. In future testing, the CISN EEW algorithms will forecast site-specific information such as peak intensity. To evaluate this information, the testing center will need access to site-specific parametric data. This information is currently not available through the ANSS data center. The Testing Center will need to interface to either SCEC Data Center or Northern California Data Center. The complexity of the testing center processing increases significantly if retrieval of waveforms and processing of waveforms into parameters is required.

5. PRESERVATION OF COLLECTED DATA:

The automated performance summaries provide some limited analysis of the CISN EEW testing results. Further analysis of the CISN EEW testing data collected during the project may provide some additional value. To support analysis efforts, we have produced data archives at SCEC containing both the original EEW algorithm performance reports and the contents of the testing center database for the final year of the project (10/1/2008 through 9/30/2009). This full data set can be provided to any group interested in further analysis of the system.

6. PRELIMINARY OBSERVATIONS ON CISN EEW PERFORMANCE:

The data in the CISN Testing Center can be analyzed to evaluate the performance of the CISN EEW algorithms. We find it is often helpful to distinguish between seismological performance measures and speed of operation measures. We summarize seismological and speed of operation results from the CISN EEW testing center in the next few tables. The information in these tables is from the CISN EEW testing center and is available through the CISN EEW testing center although it may not be in the form we present.

A metric we propose for future testing is EEW system “uptime.” Basically, we’d like to know whether the EEW system is operational at all times. The following reports try to get at this by comparing triggers to ANSS events on a daily basis. In these plots, we show days on which ANSS triggers occurred during the final testing year (10/1/2008 – 9/30/2009) and we show which days each algorithm triggered. Missing triggers may be caused by a number of reasons including network coverage issues, EEW algorithm down, the testing center down, or other possible causes. However, these plots may give some idea of consistency of reporting by the various algorithms.

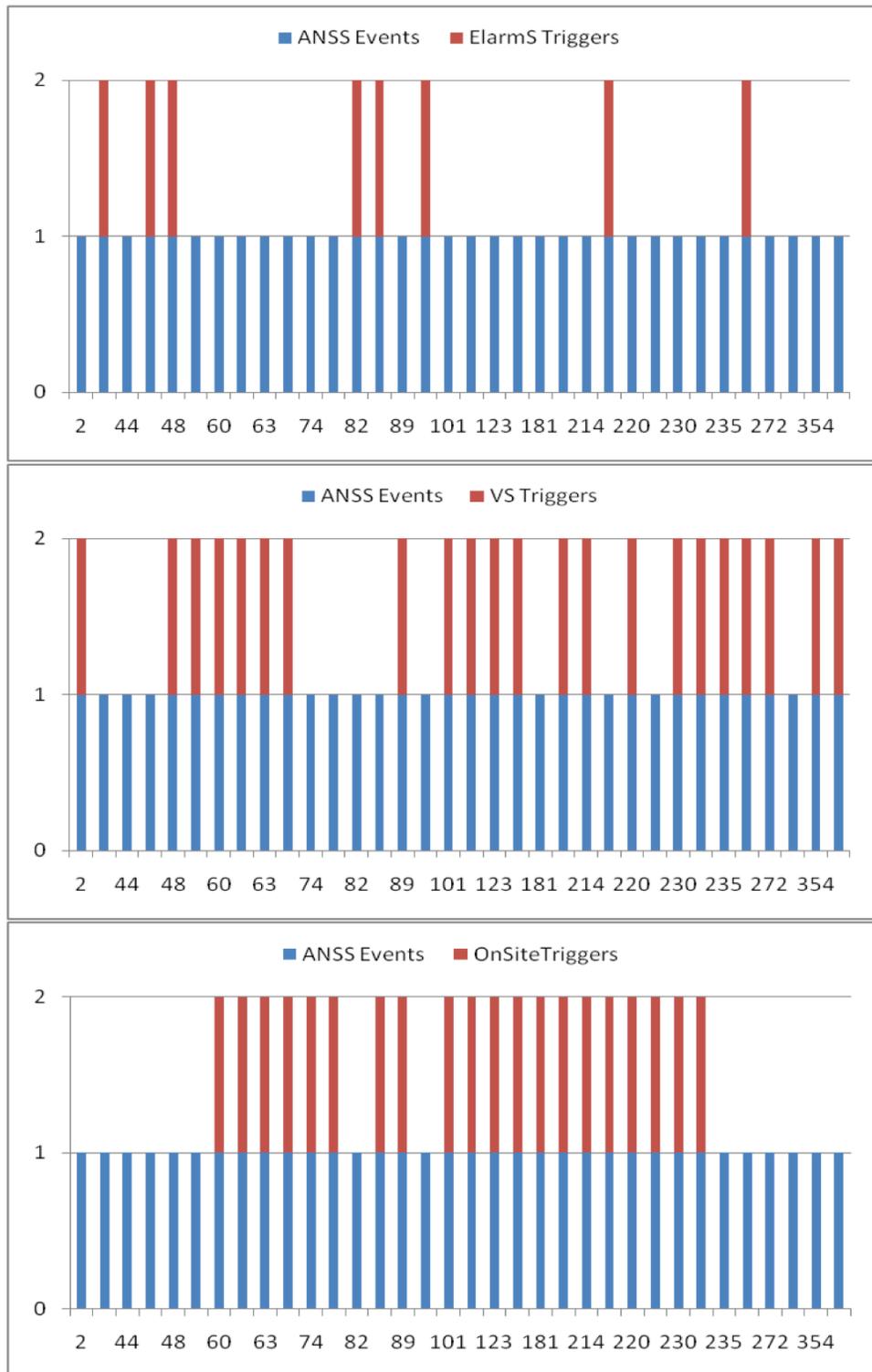


Figure 2: Horizontal axis indicates day of Project Year 3 in which there is at least one M4.0 event in the ANSS catalog for the California Testing Region. The red marks indicate the days on which EEW Algorithms produced triggers forecasting M4.0+ events. Days with events but no triggers may be due to a number of reasons including testing center problems.

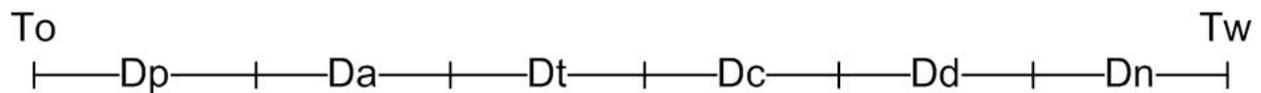
7. CISN EEW TIME MEASUREMENT MODEL

During the CISN EEW testing center development, we recognized that CISN EEW performance metrics need to include both speed of operation as well as accuracy of seismological parameters. Both categories of performance metrics are of interest to both scientists and potential users.

With regards to speed of operation, the group spent significant effort defining specific time measurements. The testing document discussed two main types of timestamps which we called algorithm time, and alert time and these are defined in the testing document. While these measures represented reasonable starting place, further work has led to an improved definition of time measurements. Our improved time measurement model replaces these other terms. These other terms can now be identified as variations in our current time measurement model.

The current time measurement model is based on tracking where delays occur during EEW processing as the system works to produce an alert. Sources of delay may include wave propagation time, transmission delay, and others.

Fast operation helps reduce the blind zone for an EEW system. To help us understand the speed of operation of the current CISN EEW algorithm, we first define the goal of an EEW as delivering ground motion warning to users at a specific site. Then we note that speed of operation changes event-by-event based on location of earthquake, station density, telemetry and processing delays. We identify where delays may occur delivering a warning to a user in the following timeline and define Speed of Operation in the following terms:



The terms in this formula are (expressed in its single channel form):

- To – Origin time of earthquake
- Dp – Duration of P-wave propagation time to sensors
- Da – Duration of data after P-wave arrival needed by algorithm
- Dt – Duration of transmission (packaging and telemetry) delay
- Dc – Duration of computation time on EEW processing system
- Dd – Duration to deliver data to users warning device
- Dn – Duration of time to produce user notification
- Tw – Time point warning notification is available to user

Then the warning time (a point in time after the event origin) is:

$$Tw = To + Dp + Da + Dt + Dc + Dd + Dn \quad (\text{Equation 1})$$

The duration to produce a warning (which we want to minimize) can be expressed as:

$$Dw \text{ (i.e. Duration to Produce Warning)} = Tw - To \quad (\text{Equation 2})$$

We want to use this Dw value to evaluate the speed of operation of the current CISN EEW algorithms. We must make a few assumptions. First, CISN EEW system does not currently deliver the warnings to users so the terms Dd and Dn are not known. We will set these to zero for now.

We can now identify the algorithm and alert times using this definition as shown below as EEW delay measurements without some of the terms in the full equation.

Origin Time	Ptt to sensor	Algorithm	Transmission	Computation	Delivery	Notification
To	Dp	Da	Dt	Dc	Dd	Dn
To	Alert Duration				0	0
To	Algorithm Duration		0	0	0	0

Figure 3: This shows how the time measures described in the CISN EEW testing document (alert time and algorithm time) can be defined using a more complete definition of EEW Warning Delays.

8. SUMMARY OF CISN EEW SEISMOLOGICAL PERFORMANCE:

In order to evaluate the seismological performance of the CISN EEW algorithms, we try to address three basic questions: (1) Are the algorithms triggering for all M4.0 and larger events in California? (2) Are the predicted magnitudes within 1.0 unit of the final ANSS magnitude? (3) Are the algorithms triggering on non-events (false triggers)? The following tables try to address these basic questions.

The following tables contain a few assumptions and qualification criteria. We only examine results from a 6 month period of time for a number of reasons (11/01/08 to 05/20/09). All three algorithms were operating and reporting to the testing center during that time. Prior to that time, the algorithms were changing rapidly and the performance was changing significantly. We only consider magnitude M4.0 and larger events based on the belief that system performance changes for larger magnitude events. Also, only events within the California testing region are considered. These performance summaries could be repeated with a more complete data set if needed. We present the following as a representative sample of results.

The following performance reports have several qualifications. First, the CISN EEW algorithms run on only part of the CISN stations. Typically UCB ElarmS-NI runs on Northern California stations, and TauC-PD and VS-RT run on Southern California stations. Efforts are underway to run the algorithms state-wide. Also, the CISN EEW algorithm developer may have versions of their codes running which do not report to the CISN EEW testing center. If the CISN EEW algorithm performance reports differ from performance reports from the Algorithm groups, it may be due to the fact the algorithm groups are reporting on performance of preliminary code implementations that are not yet reporting to the CISN EEW testing center.

Table 1: Summary of CISN EEW Performance on M4.0 Events (11/01/08 to 05/20/09)

Data Sources	ANSS Catalog Event	CISN Totals	ElarmS-NI	TauC-PD	VS-RT
Correct Event Info	26	23	4	14	16
Errors/Misses					
No Reports			22	11	10
Mags Too High (> 1.0+)			2	1	2
False Triggers			0	1000+	1

Table 1 show that the ANSS catalog contains 26 M4.0+ in California over the last 6 months. Of those, only three events did not produce valid triggers with accurate predicted magnitude information from the CISN EEW algorithms. Table 2 below shows the list of 26 ANSS events during this time periods and shows which algorithms triggered correctly for each event. In Table 2 below nr = no report, and mth= magnitude too high. The three events which did not produce a trigger are shown in bold italics.

Table 2: CISN EEW Seismological Performance Summary for M4.0+ Events (11/01/08 to 05/20/09)

Event IDs	Event Locations	Mag	ElarmS-NI	TauC-P	VS
10411545	33.93 -118.34	4.04	nr	1	1
10410337	33.94 -118.34	4.70	nr	1	1
10406593	34.44 -119.18	4.17	nr	mth	1
51220943	40.75 -124.16	4.17	1	nr	nr
10403777	34.07 -118.88	4.42	nr	1	1
14443704	32.07 -115.76	4.30	nr	nr	1
14443616	32.33 -115.25	4.22	nr	1	nr
40234037	37.28 -121.61	4.30	nr	nr	nr
14433456	33.32 -115.73	4.77	nr	1	1
14418600	35.41 -117.79	4.39	nr	1	1
10374021	32.69 -118.23	4.19	nr	1	1
10370141	34.11 -117.30	4.45	nr	nr	1
51214595	38.78 -122.77	4.30	nr	nr	nr
10368325	32.57 -115.54	4.52	nr	1	1
200812262043	39.96 -120.87	4.50	1	nr	nr
51213534	36.67 -121.30	4.00	1	nr	nr
10366249	32.55 -115.54	4.03	nr	1	nr
10366101	35.97 -117.32	4.01	nr	1	nr
14408052	34.81 -116.42	5.06	nr	1	1
14407020	35.97 -117.32	4.03	nr	1	1
14406304	35.97 -117.32	4.04	nr	1	1
14406196	35.97 -117.33	4.03	nr	1	1
14404512	32.33 -115.33	4.98	nr	nr	1
14403732	33.50 -116.86	4.11	nr	nr	1
51211307	40.31 -124.60	4.60	1	nr	nr
51211113	40.44 -125.27	4.00	nr	nr	nr
Totals Correct Triggers			4	14	16

The seismological reports in Table 1 above also identify which CISN EEW algorithms are producing false triggers. The UCB ElarmS, because it is reporting retrospective data, should never produce a false trigger. This does not represent its real-time performance. The VS only rarely reported false triggers. The TauC-PD, as a single station algorithm, produces many false triggers. This might be partially addressed by adjusting the algorithm to only trigger on larger events (e.g. M5.0+)

A wide variety of other seismological performance measures are reported by individual research groups. In the future, we propose to select useful seismological performance measures used by individual research groups and implement those within the CISN testing center.

9. SUMMARY OF CISN EEW SPEED OF OPERATION:

Along with seismological performance, the CISN Testing Center also reports metrics on system speed of operation. As noted above, EEW speed of operation measures are typically event specific. The system speed of operation varies on an event-by-event basis.

We focus initially on measurements of the parameter we call Warning Delay (Dw) as defined above. This is intended to be a real-time measurement of how long after origin time that a user receives notification that an event is in progress. Shorter values are better.

In the following table, we list the Dw for the first correct trigger from the algorithm. A correct report contains a predicted magnitude within 1.0 unit of the final ANSS magnitude. In a real-time system, there is no way of knowing which of the incoming reports contains a correct magnitude.

The TauC-PD reports single station triggers. In this table, we report Dw for the first correct TauC-PD trigger. Single station triggers may not be practical due to increased possibility of false triggers.

ElarmS-NI reports only Dp + Da. In this table, Dw values for ElarmS-NI assume that Dt and Dc are zero. These terms vary significantly for the other algorithms, so Dw for a real-time ElarmS implementation is likely to add 5-10 seconds to the reported values.

Given these considerations, the table shows that the CISN EEW algorithms, running on the existing network, with existing telemetry and central processing systems, are capable of predicting accurate magnitudes in less than 20 seconds in many cases. A 20 second Dw corresponds to a blind zone (in California with Vs of 3.6 km/s) of about 72 km.

Table 3: CISN EEW Speed of Operation Summary for M4.0+ Events (11/01/08 to 05/20/09)

Event IDs	Event Locations	Mag	Dw (ElarmS-NI) Seconds	Dw (TauC-P) Seconds	Dw (VS) Seconds
10411545	33.93 -118.34	4.04		14	15
10410337	33.94 -118.34	4.70		15	15
10406593	34.44 -119.18	4.17			19
51220943	40.75 -124.16	4.17	10		
10403777	34.07 -118.88	4.42		16	19
14443704	32.07 -115.76	4.30			32
14443616	32.33 -115.25	4.22		20	
40234037	37.28 -121.61	4.30			
14433456	33.32 -115.73	4.77		15	18
14418600	35.41 -117.79	4.39		19	20
10374021	32.69 -118.23	4.19		27	28
10370141	34.11 -117.30	4.45			18
51214595	38.78 -122.77	4.30			
10368325	32.57 -115.54	4.52		19	29
200812262043	39.96 -120.87	4.50	15		
51213534	36.67 -121.30	4.00	5		
10366249	32.55 -115.54	4.03		19	
10366101	35.97 -117.32	4.01		16	
14408052	34.81 -116.42	5.06		15	24
14407020	35.97 -117.32	4.03		16	25
14406304	35.97 -117.32	4.04		28	38
14406196	35.97 -117.33	4.03		29	26
14404512	32.33 -115.33	4.98			29
14403732	33.50 -116.86	4.11			20
51211307	40.31 -124.60	4.60	15		
51211113	40.44 -125.27	4.00			

We have plotted these results graphically in the figures below. These figures seem to indicate that for the current CISN EEW system, longer Dw values occur for events near the edges of the network as might be expected. We suggest that it would be instructive to map the events, the delays, and the location of the seismic stations used by the EEW system.

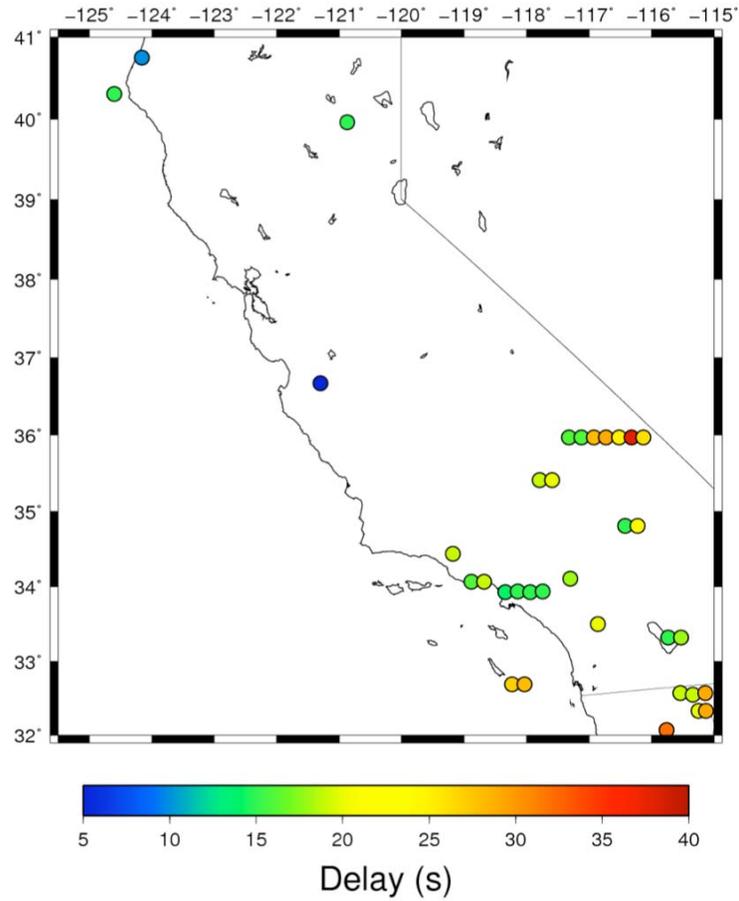


Figure 4: This shows the location of M4.0+ ANSS events for which the CISN EEW algorithms triggered. This shows the performance of all three CISN EEW algorithms. Colors indicate the Warning Delay (in seconds) for each event. The ElarmsS algorithm in Northern California reported Warning delays to the testing center without including P-wave travel time and telemetry delays so the delays for this algorithm will be greater when these delays are included.

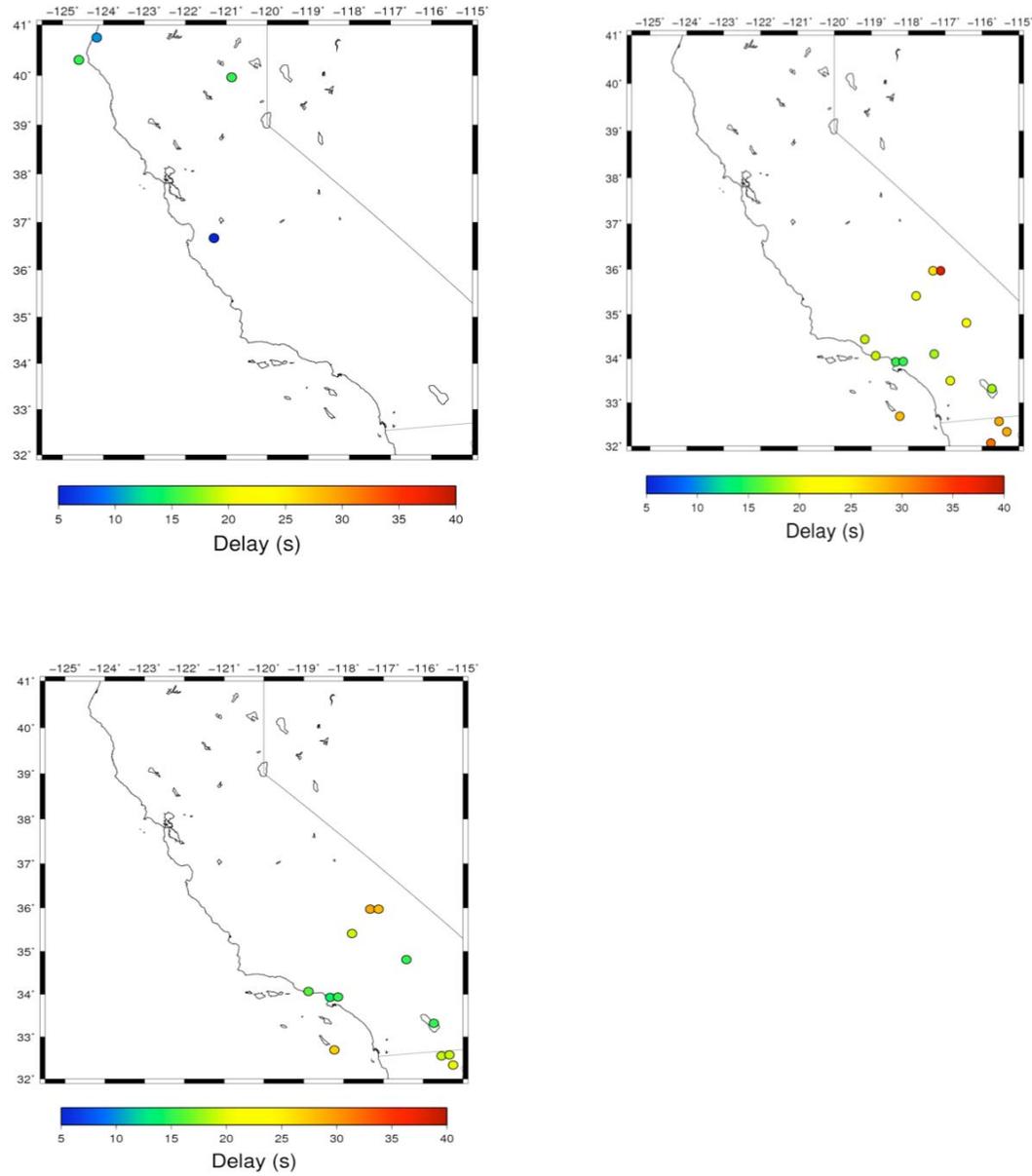


Figure 5: Warning Delay (in seconds) by CISN EEW Algorithm for 6 months of ANSS M4.0+ events in California. UCB- ElarmS (top left) ETH-Virtual Seismologist (top right) CIT-On-Site (bottom left)

CISN Earthquake Early Warning: Testing of Seismological Algorithms - Publications

Articles about the project in broad-interest publications

- Allen, R.M., 2008. At First Jolt: Will we have warnings for the next big earthquake? *Geotimes/EARTH*, 53 (10), p52-59.
- Cua, G., 2008. Earthquake early warning systems: an investment that pays off in seconds. *Natural Hazards Observer*, volume 32, number 5, pp1-3. Natural Hazards Research Center. Univ. of Colorado, Boulder. Invited comment.
- Vance, E. , 2008. Do you want a 20-second warning? *San Francisco Magazine*, p 60, November 2008.

Journal Publications

- Allen, R.M., 2007. The ElarmS earthquake early warning methodology and its application across California. In "Earthquake Early Warning Systems", P. Gasparini, G. Manfredi, J. Zschau (Eds), p. 21-44, Springer, ISBN-13 978-3-540-72240-3.
- Allen, R.M., 2007. Earthquake hazard mitigation: New directions and opportunities. In "Treatise on Geophysics", G. Schubert (ed.), Vol 4 (H. Kanamori ed.), p607-648, Elsevier.
- Allen, R.M., P. Gasparini and O. Kamigaichi (eds), New Methods and Applications of Earthquake Early Warning. Special Section, *Geophys. Res. Lett.*, **36**, L00B01, L00B02, L00B03, L00B04, L00B05, L00B06, L00B07, L00B08, 2009.
- Allen, R.M., H. Brown, M. Hellweg, O. Khainovski, P. Lombard, D. Neuhauser, Real-time earthquake detection and hazard assessment by ElarmS across California *Geophys. Res. Lett.*, **36** L00B08, doi:10.1029/2008GL036766, 2009.
- Allen, R.M., P. Gasparini and O. Kamigaichi (eds), Earthquake Early Warning, Special Issue, *Seismo. Res. Lett.*, **80**, (5) p682-782, 2009.
- Allen, R.M., P. Gasparini, O. Kamigaichi, M. Böse, The Status of Earthquake Early Warning around the World: An Introductory Overview, *Seismo. Res. Lett.*, **80**, (5) p682-693, doi: 10.1785/gssrl.80.5.682, 2009.
- Böse, M., E. Hauksson, K. Solanki, H. Kanamori, Y.-M. Wu, and T.H. Heaton, 2009. A New Trigger Criterion for Improved Real-time Performance of On-site Earthquake Early Warning in Southern California, *Bull. Seism. Soc. Am.* 99-2A.
- Böse, M., E. Hauksson, K. Solanki, H. Kanamori, and T.H. Heaton, 2009. Real-Time Testing of the On-site Warning Algorithm in Southern California and Its Performance During the July 29 2008 Mw5.4 Chino Hills Earthquakes, *Geophys. Res. Lett.* 36, L00B03, doi:10.1029/2008GL036366.
- Böse, M. and T.H. Heaton: Probabilistic Prediction of Rupture Length, Slip and Seismic Ground Motions for an Ongoing Rupture - Implications for Early Warning for Large Earthquake, in prep.
- Brown, H., R.M. Allen, V. Grasso, Testing ElarmS in Japan, *Seismo. Res. Lett.*, **80**, (5) p727-739, doi: 10.1785/gssrl.80.5.727, 2009.
- Brown, H.M., R.M. Allen, M. Hellweg, O. Khainovski, D. Neuhauser, A. Souf, in review. Development of the ElarmS methodology for earthquake early warning: Realtime application in California and offline testing in Japan, *Soil Dynamics and Earthquake Engineering*.

- Cua, G., M. Fischer, T. Heaton, S. Wiemer, D. Giardini, 2009. Real-time performance of the Virtual Seismologist method in southern California. *Seismol. Res. Lett.* 80 (5) p740-747.
- Cua, G. and T. H. Heaton, 2007. The Virtual Seismologist (VS) method: a Bayesian approach to earthquake early warning. In "Earthquake Early Warning Systems", P. Gasparini, G. Manfredi, J. Zschau (Eds), Springer, ISBN-13 978-3-540-72240-3.
- Cua, G., and T. H. Heaton, *in review* Characterizing average properties of southern California ground motion amplitudes and envelopes. *Bull. Seism. Soc. Am.*
- Köhler, N., Cua, G., Wenzel, F., and M. Böse, 2009: Rapid Source Parameter Estimations of Southern California Earthquakes Using PreSEIS, *Seism. Res. Lett.* 80 (5), pp. 748-754, doi: 10.1785/gssrl.80.5.748.
- Lockman, A.B. and R.M. Allen, 2007. Magnitude-period scaling relations for Japan and the Pacific Northwest: Implications for earthquake early warning *Bull. seism. Soc. Am.* **97** (1), 140-150, doi: 10.1785/0120040091
- Olivieri, M., R.M. Allen and G. Wurman, 2008. The potential for earthquake early warning in Italy using ElarmS. *Bull. seism. Soc. Am.* **98**, p495-503, doi: 10.1785/0120070054.
- Oth, A., M. Böse, F. Wenzel, N. Köhler and M. Erdik: Optimizing Seismic Networks for Earthquake Early Warning: the Istanbul case, in prep.
- Shieh, J.T., Y.M. Wu, R.M. Allen, 2008. A comparison of tau-c and tau-p-max for magnitude estimation in earthquake early warning *Geophys. Res. Lett.*, **35**, L20301, doi:10.1029/2008GL035611.
- Tsang, L., R.M. Allen and G. Wurman, 2007. Magnitude scaling relations from P-waves in southern California. *Geophys. Res. Lett.* **34** L19304, doi:10.1029/2007GL031077.
- Wurman, G., R.M. Allen and P. Lombard, 2007. Toward earthquake early warning in northern California. *J. Geophys. Res.* **112**, B08311, doi:10.1029/2006JB004830.
- Wu, Y.-M., H. Kanamori, R.M. Allen and E. Hauksson, 2007. Determination of earthquake early warning parameters, tau-c and Pd, for southern California *Geophys. J. Int.* **170**, 711-717, doi: 10.1111/j.1365-246X.2007.03430.x
- Yamada, M. 2007. Early warning for earthquakes with large rupture dimension. Ph.D. thesis, California Institute of Technology.
- Yamada, M. and T. Heaton, 2008. Real-time Estimation of Fault Rupture Extent using Envelopes of Acceleration. *Bull. seism. Soc. Am.* **98** (2), 607-619.
- Yamada, M., T. Heaton and J. Beck, 2007. Real-time Estimation of Fault Rupture Extent Using Near-source versus Far-source Classification. *Bull. seism. Soc. Am.* **97** (6), 1890-1910.
- Yamada, M., T. Heaton and J. Beck, *in press*. Early Warning Systems for Large Earthquakes: Classification of Near-source and Far-source Stations by using the Bayesian Model Class Selection. *ICASP10*.

Project Conference Abstracts (Chronological)

- Böse, M., E. Hauksson, K. Solanki, H. Kanamori, T. Heaton and the CISN EEW Group. Recent Developments in CISN Earthquake Early Warning Testing for California. Annual SCEC Meeting, Palm Springs, CA. September 2009.

- Allen, R.M. Early warning testing and development in California with a focus on ElarmS. 2nd International Workshop on Earthquake Early Warning, Kyoto, Japan. April 2009.
- Böse, M., E. Hauksson, K. Solanki, H. Kanamori, T. Heaton. Updates on Early Warning Testing and Finite Fault Research at Caltech. 2nd International Workshop on Earthquake Early Warning, Kyoto, Japan. April 2009.
- Cua, G., M. Fischer, T. Heaton. Real-time performance of the Virtual Seismologist earthquake early warning algorithm. 2nd International Workshop on Earthquake Early Warning, Kyoto, Japan. April 2009.
- Heaton, T. and M. Böse. Probabilistic Prediction of Rupture Length, Slip and Seismic Ground Motion of an ongoing Rupture. 2nd International Workshop on Earthquake Early Warning, Kyoto, Japan. April 2009.
- Cua, G., M. Fischer, T. Heaton, S. Wiemer. Evaluating the real-time and offline performance of the Virtual Seismologist earthquake early warning algorithm. European Geophysical Union Annual Meeting, Vienna, Austria, April 2009.
- Cua, G., M. Fischer, T. Heaton, S. Wiemer, D. Giardini. Real-time and off-line performance of the Virtual Seismologist earthquake early warning method in southern California and Switzerland. American Geophysical Union Fall Meeting, San Francisco, California, December 2008.
- Böse, M., E. Hauksson, K. Solanki, H. Kanamori, T. Heaton, and Y.-M. Wu. Real-time Testing of On-site Earthquake Early Warning within the California Integrated Seismic Network (CISN) Using Statewide Distributed and On-site Processing, AGU Fall Meeting, San Francisco, December 2008.
- Brown, H. and R.M. Allen, Analyzing the Capabilities of the ElarmS Methodology Using a Japanese Dataset, San Francisco, December 2008.
- Heaton, T.H. and M. Böse. Earthquake Early Warning for Large Events using Probabilistic Models for Seismic Rupture Prediction. AGU Fall Meeting, San Francisco, December 2008.
- Hellweg, M., R.M. Allen, M. Bose, H. Brown, G. Cua, D. Given, E. Hauksson, T. Heaton, T. Jordan, O. Khainovski, P. Maechling, D. Neuhauser, D. Oppenheimer, K. Solanki, M. Zeleznik, Earthquake early warning in California: Evaluating Hardware and Software, Eos Trans. AGU, 89(53), Fall Meet. Suppl., Abstract, S14B-05, 2008.
- Neuhauser, D.S., R.M. Allen, H. Brown, M. Hellweg, O. Khainovski and P. Lombard, Realtime Earthquake Detection and Hazard Assessment by ElarmS Across California, Eos Trans. AGU, 89(53), Fall Meet. Suppl., Abstract S11A-1725, 2008
- Hellweg, M., Allen, R.M., Brown, H. and Neuhauser, D., Earthquake Early Warning for California, Third Conference on Earthquake Hazards in the Eastern San Francisco Bay Area: Science, Hazard, Engineering and Risk, October 22-24, 2008, Hayward, California, 2008.
- Böse, M.. Seconds Can Make a Difference - Recent Progress in Earthquake Early Warning, Annual SCEC Meeting, Palm Springs, September 2008.
- Allen, R.M. ElarmS AlertMaps: An earthquake in California and another near Istanbul. SAFER (Seismic Early Warning for Europe) Annual Meeting, Istanbul, Turkey, June 2008.
- Allen, R.M., M. Böse, H. Brown, G. Cua, D. Given, E. Hauksson, T. Heaton, M. Hellweg, T. Jordan, A. Kireev, P. Maechling, D. Neuhauser, D. Oppenheimer, K. Solanki,

- M. Zeleznik. Rapid telemetry and earthquake early warning. American Geophysical Union Joint Meeting, Fort Lauderdale, Florida, May 2008.
- Houlié, N., R.M. Allen. The Instantaneous Displacement (ID) Method: Application of rapid displacement estimates to earthquake early warning alerts. American Geophysical Union Joint Meeting, Fort Lauderdale, Florida, May 2008.
- Cua, G., P. Maechling, R.M. Allen, M. Böes, E. Hauksson, T. Heaton, P. Hellweg, A. Kireev, D. Neuhauser, K. Solanki, S. Wiemer, J. Woessner, M. Zeleznik. Comparison and testing of earthquake early warning algorithm performance. European Geophysical Union meeting, April 2008.
- Hellweg, M, A. Chung, D. Dreger, A. Kim and J. Boatwright, Mapping the rupture of the MW 5.4 Alum Rock earthquake. *Seismol. Res. Lett.* 79, 353, 2008.
- Allen, R.M., G. Wurman, P. Hellweg, A. Kireev and D. Neuhauser. Earthquake early warning across California: Performance of ElarmS on the existing seismic networks. AGU Fall Meeting, San Francisco, December 2007.
- Brown, H. and R.M. Allen. Application of ElarmS to earthquakes in Japan. AGU Fall Meeting, San Francisco, December 2007.
- Cua, G., M. Fischer, J. Clinton, T. Heaton, S. Wiemer, and D. Giardini. Calibrating and implementing the Virtual Seismologist method in Switzerland. AGU Fall Meeting, San Francisco, December 2007.
- Cua, G. and T. H. Heaton. New ground motion prediction equations spanning weak and strong ground motion levels. AGU Fall Meeting, San Francisco, December 2007.
- Hellweg, M., R. Uhrhammer, K. Hutton, A. Walter, P. Lombard, E. Hauksson, Recalibrating ML for the California Integrated Seismic Network, *Eos Trans. AGU*, 88(52), Fall Meet. Suppl., Abstract S43A-1057, 2007.
- Neuhauser, D., A. Kireev, G. Wurman, P Hellweg, R.M. Allen. A Real-Time CISN Test Bed for the ElarmS Early Warning Algorithm. AGU Fall Meeting, San Francisco, December 2007.
- Olivieri, M., R. Basili and R.M. Allen. ElarmS and the next large earthquake in Italy. AGU Fall Meeting, San Francisco, December 2007.
- Solanki, K., E. Hauksson, H. Kanamori, Y.-M. Wu, T. Heaton, and M. Böse. Earthquake Early Warning. Real-time Testing of an On-site Method Using Waveform Data from the Southern California Seismic Network. AGU Fall Meeting, San Francisco, December 2007.
- Tsang, L.L.H., R.M. Allen and G. Wurman. Calibration of ElarmS using earthquakes in southern California. AGU Fall Meeting, San Francisco, December 2007.
- Wenzel, F., N. Köhler, G. Cua, and M. Böse. Neural network methodology for earthquake early warning applications. American Geophysical Union Fall Meeting, San Francisco, December 2007.
- Zeleznik, M., P. Maechling, G. Wurman, A. Kireev, K. Solanki, R. Allen, D. Neuhauser, E. Hauksson, P. Hellweg, G. Cua, T. Heaton. Development of the CISN Earthquake Early Warning Web Site: Establishing a Basis for Comparison of Algorithms. AGU Fall Meeting, San Francisco, December 2007.
- Cua, G., T. Heaton, and S. Wiemer. Bayesian earthquake early warning: a dispatch from one of many frontiers of real-time seismology. 5th International Workshop on Statistical Seismology, Erice, July 2007.
- Allen, R.M. Toward earthquake early warning for California. SAFER (Seismic Early Warning for Europe) Meeting, Athens, Greece, June 2007.

- Cua, G., M. Fischer, D. Eberhard, K. Solanki. Implementing the Virtual Seismologist expert system. SAFER (Seismic Early Warning for Europe) Meeting, Athens, Greece, June 2007.
- Wurman, G., R. M. Allen, P. Lombard, D. Neuhauser, A. Kireev, and M. Hellweg. Real-Time Implementation of ElarmS for Earthquake Early Warning in Northern California. Seismological Society of America, Hawaii, April 2007.
- Wurman, G. and R.M. Allen. Toward Earthquake Early Warning in Northern California. AGU Fall Meeting, San Francisco, December 2006.
- Allen, R.M. and G. Wurman. Earthquake early warning in northern California. SCEC Annual Meeting, September 2006.
- Allen, R.M. ElarmS: An earthquake early warning methodology for California and elsewhere. SAFER Kickoff meeting, Capri, Italy, July 2006.
- Allen, R.M. and A. Limaye. Applications and Benefits of Earthquake Early Warning: Implementation and Alert Times Across California. 1906 Earthquake Conference/SSA, April 2006.
- Grasso, V. and R.M. Allen. Earthquake Warning Systems from the User Perspective. 1906 Earthquake Conference/SSA, April 2006.
- Hauksson, E., K. Solanki, D. Given, P. Maechling, D. Oppenheimer, D. Neuhauser and M. Hellweg, Implementation of Real-Time Testing of Earthquake Early Warning Algorithms: Using the California Integrated Seismic Network (CISN) Infrastructure as a Test Bed. *Seismol. Res. Lett.* 77(2), 307, 2006.
- Wurman, G. and R.M. Allen. ElarmS Earthquake Alarm Systems: Early Results in Northern California. 1906 Earthquake Conference/SSA, April 2006.