# Quantifying the Value of Real-Time Geodetic Constraints for Earthquake Early Warning Using a Global Seismic and Geodetic Data Set

C. J. Ruhl[1,2] ID, D. Melgar[3] ID, A. I. Chung[1] ID, R. Grapenthin[4] ID, and R. M. Allen[1] ID

[1]Berkeley Seismological Laboratory, University of California, Berkeley, CA, USA, [2]Department of Geosciences, The University of Tulsa, Tulsa, OK, USA, [3]Department of Earth Sciences, University of Oregon, Eugene, OR, USA, [4]Department of Earth and Environmental Science, New Mexico Institute of Mining and Technology, Socorro, NM, USA

**Abstract** Geodetic earthquake early warning (EEW) algorithms complement point-source seismic systems by estimating fault-finiteness and unsaturated moment magnitude for the largest, most damaging earthquakes. Because such earthquakes are rare, it has been difficult to demonstrate that geodetic warnings improve ground motion estimation significantly. Here, we quantify and compare timeliness and accuracy of magnitude and ground motion estimates in simulated real time from seismic and geodetic observations for a suite of globally distributed, large earthquakes. Magnitude solutions saturate for the seismic EEW algorithm (we use ElarmS) while the ElarmS-triggered Geodetic Alarm System (G-larmS) reduces the error even for its first solutions. Shaking intensity (Modified Mercalli Intensity, MMI) time series calculated for each station and each event are assessed based on MMI threshold crossings, allowing us to accurately characterize warning times per station. We classify alerts and find that MMI 4 thresholds result in true positive alerts for only 13.7% of sites exceeding MMI 4 with a median warning time of 18.9 s for ElarmS, while G-larmS issues true positive alerts for 52.3% of all sites exceeding MMI 4 with a significantly longer median warning time of 55.8 s. The geodetic EEW system reduces the number of missed alerts for a threshold of MMI 4 from 48.7% to 19.2% for all sites, but also increases the number of false positive alerts from 1.2% to 13.4% of all sites. By quantifying increased accuracy in magnitude, ground motion estimation, and alert timeliness, we demonstrate that finite-fault geodetic algorithms add significant value, including better cost savings performance, to point-source seismic EEW systems for large earthquakes.

## 1. Introduction

The concept underpinning earthquake early warning (EEW) is to detect and characterize earthquakes as soon as possible after they initiate in order to warn ahead of the arrival of strong ground shaking (Allen, Gasparini, et al., 2009). Ideally, EEW should be a ground motion (GM) warning system as it is knowledge of the expected intensity of shaking that is most important to a user. In fact, that user's actions often depend on the level of GM expected at their site. The United States' ShakeAlert EEW system currently uses earthquake source parameters in combination with a GM model to provide warnings. Therefore, the success of the EEW system depends on accurate earthquake source characteristics (origin time, location, magnitude, and fault-finiteness) in order for the GM estimates themselves to be accurate. Here, we use data from 32 large (M > 6) globally distributed earthquakes to evaluate the accuracy and timeliness of earthquake magnitude and GM estimates when comparing seismic and geodetic EEW systems. We find that for large earthquakes the performance is much improved when including information from geodetic algorithms.

Traditionally, EEW systems use features of elastic waves recorded on inertial seismometers to estimate the magnitude and epicenter of an earthquake (Allen, Gasparini, et al., 2009). The ShakeAlert system's Earthquake Point-source Integrated Code (EPIC), for example, uses the amplitude of the first few seconds of the *P* wave arrival on several seismic stations to estimate source parameters (Chung et al., 2017). It has been noted that this approach leads to saturation, an underestimation of the magnitudes of large events (e.g., Hoshiba & Ozaki, 2014). This is due to saturation of accelerations at higher frequencies in the epicentral region of an earthquake, that is, very strong shaking. When integrating acceleration records to obtain displacement, errors are introduced because of this instrumental saturation. Inertial sensors, which are used by EEW algorithms, therefore provide unreliable measurement of very-low-frequency displacements (Boore & Bommer, 2005; Melgar et al., 2013). Meier et al. (2016) found that the first few

seconds of the *P* wave, as recorded by inertial sensors, do not contain enough information to forecast growth of the earthquake into a very large M8+ event. For example, during the 2011 $M_w$9.0 Tohoku-Oki, Japan, earthquake, first and final alerts issued 8.6 and 116.8 s after origin time underestimated the final magnitude by 1.8 and 0.9 magnitude units, respectively (Hoshiba et al., 2011). This saturation resulted in underestimated GMs in the greater Tokyo area and timely, but severely underestimated, tsunami warnings—including amplitudes and geographic extent (Hoshiba & Ozaki, 2014).

Blewitt et al. (2006) first proposed using geodetic measurements to overcome magnitude saturation after severe underestimation of the 2004 $M_w$9.3 Sumatra earthquake by long-period seismic observations within the first hour after the event. Based on this principle, instead of traditional seismic data (velocity and acceleration), geodetic EEW algorithms use observations collected by Global Navigation Satellite Systems (GNSS). GNSS can be conceptualized as strong-motion displacement sensors capable of measurement at the longest periods down to the static or permanent offset at 0 Hz (Melgar et al., 2013). Since the 2004 Sumatra earthquake and tsunami, many GNSS-based techniques have been developed and improved to estimate source properties for earthquake and tsunami early warning in real time (e.g., Allen & Ziv, 2011; Colombelli et al., 2013; Crowell et al., 2016; Grapenthin et al., 2014a, 2014b; Grapenthin & Freymueller, 2011; Kawamoto et al., 2017; Minson et al., 2014); a history of geodetic early warning methods and their development can be found in Bock and Melgar (2016).

Seismic systems, as compared to geodetic, are also more limited by network configuration as they require good azimuthal coverage and dense station spacing. Out-of-network and edge-of-network events, that is, those with poor azimuthal coverage, are often severely mischaracterized both in location and magnitude. This limitation is demonstrated by the performance of seismic point-source (e.g., ElarmS, Allen, Brown, et al., 2009) and seismic finite-fault algorithms (e.g., FinDer, Böse et al., 2012) during replays of the out-of-network $M_w$7.2 El Mayor-Cucapah event occurring south of the United States-Mexico border. When replaying the earthquake using only the stations operating in real time within the United States in 2010, both ElarmS and FinDer resulted in severe event mislocation and magnitude underestimation (Ruhl et al., 2017). Geodetic systems, however, successfully characterized the El Mayor-Cucapah event using a similar (United States-only) network geometry (Allen & Ziv, 2011; Grapenthin et al., 2014a; Ruhl et al., 2017) and have been shown to contribute to alarms in sparse seismic networks through S wave detections (Grapenthin et al., 2017). With the addition of near-source stations, seismic-finite-source algorithms like FinDer are better able to model the 2010 $M_w$7.2 El Mayor earthquake (Böse et al., 2015). Using asymmetrical GM templates, FinDer is also now able to model offshore subduction zone events. For the magnitude of the 2011 $M_w$9.0 Tohoku-oki earthquake FinDer predicted $M_w$8.5 240 s after origin time (Böse et al., 2015). Recent FinDer results on moderate magnitude events (e.g., for the 2016 $M_w$7.0 Kumamoto earthquake) are encouraging as well (Böse et al., 2018), and future comparisons between seismic and geodetic finite-fault algorithms are important.

Chung et al. (2017) presented EEW results for recent earthquakes using two versions of ElarmS (E2 and E3). While the number of missed and false events is substantially reduced in E3, both versions demonstrate that missed and false events are primarily those which originate outside of network boundaries (e.g., offshore) or in sparse network areas (e.g., eastern Oregon and Washington). Similarly, during replays of a simulated $M_w$8.7 megathrust earthquake on the offshore Cascadia subduction zone, ElarmS first locates the event offshore with an initial magnitude of ~8 before relocating it to within the network and lowering the magnitude to ~7. Geodetic finite-fault results for this simulation as well as the El Mayor-Cucapah event are more robust in this regard and demonstrate its ability to accurately estimate magnitudes of offshore, out-of-network events based on the first alerts produced from the seismic algorithm (Ruhl et al., 2017). It is worth noting that in the western United States several M > 7 earthquake hazards exist near the edges of and beyond the footprint of the real-time EEW seismic network; for example, in the Cascadia subduction zone to the west, on crustal faults in the Basin and Range and Walker Lane to the east, and to the south and north of the contiguous United States. To provide a true west-coast-wide system, we must extend the seismic network beyond the footprint of interest or include geodetic data to ameliorate some of the issues with limited network configurations.

To provide coverage for the full range of damaging earthquakes (M6+) in the United States, several groups have developed EEW algorithms that make use of GNSS data. Three of these are currently being tested for implementation into ShakeAlert: G-larmS, BEFORES, and G-FAST (Crowell et al., 2016; Grapenthin et al.,

2014a; Minson et al., 2014). In this work, we will focus on performance of the G-larmS algorithm (see Murray et al., 2018, for a comparison) and make the data freely available so that other algorithm developers can conduct similar evaluations. The Geodetic Alarm System (G-larmS) was the first operational real-time geodetic system in the United States (Grapenthin et al., 2014a), and has been running in real time since the beginning of May 2014. G-larmS analyzes GNSS position time series in real time, determines static offsets, and performs a least-squares inversion for slip using a priori fault geometries determined by the initial event location and magnitude provided by the ShakeAlert seismic algorithms (Colombelli et al., 2013). Fault configuration files are specific to each region, see section 3.2 for details.

G-larmS operates as a triggered system and is coupled to the seismic point-source algorithms of ShakeAlert. Thus, our goal is to conduct end-to-end tests of the seismic only and the coupled seismic and geodetic systems with real data. We use a suite of large (M > 6) earthquakes worldwide for which we have waveforms from both seismic and geodetic sites to test both the seismic (ElarmS) and geodetic systems (G-larmS). We quantify the timeliness and accuracy of seismic and geodetic magnitude and GM EEW alerts. We then show that the additional information and accuracy achieved by using available real-time GNSS data has substantial added value and that geodesy has an important role to play in providing warnings for the largest, most damaging earthquakes and their associated hazards.

## 2. Data

We test the EEW algorithms using 32 earthquakes from around the world ranging in magnitude from $M_w6.0$ (2004 Parkfield) to $M_w9.0$ (2011 Tohoku-oki) and with variable quantity and quality of seismic and geodetic data (Table 1). The database is dominated by subduction zone megathrust events but includes continental strike-slip (e.g., 2016 $M_w7.0$ Kumamoto), intraplate normal (e.g., 2017 $M_w8.2$ Tehuantepec), and other non-subduction zone events (e.g., 2015 $M_w7.8$ Nepal). The number of stations of each data type vary from a few to hundreds; both seismic and geodetic records (either real or synthetic) exist for 29 of the 32 earthquakes. We do not have seismic data for the 2010 $M_w7.7$ Mentawai earthquake, the 2014 $M_w7.7$ Iquique, Chile, aftershock, or the 2015 $M_w7.3$ Nepal aftershock, for completeness we include these events as part of the geodetic analysis. The $M_w8.7$ "Cascadia001300" and $M_w7.0$ "Hayward4Hz" earthquakes are simulations of scenario events for which we have synthetic seismic and geodetic data (Melgar et al., 2016; Rodgers et al., 2018). For the $M_w6.9$ Nisqually 2001 event we have actual recorded seismic data as well as synthetic GNSS data from a slip inversion (Crowell et al., 2016). In the following two sections, we discuss the details of seismic and geodetic data used in this study.

### 2.1. Seismic Data

Seismic data were collected from various sources for a total of 29 out of 32 earthquakes. The 2014 $M_w7.7$ Iquique, Chile, aftershock and the 2015 $M_w7.3$ Nepal aftershock do not have seismic data and we were unable to include seismic data from the 2010 $M_w7.7$ Mentawai, Indonesia, earthquake. We format the data into miniseed format in SI units (cm/s or cm/s/s) and create channel files specifying, among other things, the units, sample rates, and gains of each channel. Using the accompanying channel files, all waveforms for each event are combined and rewritten into one or more Earthworm tank-player files to be used for real-time replays.

For the 2001 $M_w6.9$ Nisqually, 2004 $M_w6.0$ Parkfield, 2010 $M_w7.2$ El Mayor-Cucapah, and 2014 $M_w6.1$ Napa earthquakes, we downloaded acceleration and velocity waveforms in miniseed format along with channel files directly from the official ShakeAlert test suite (Cochran et al., 2017). For each of these, waveforms begin 2 min prior to origin time and are a total of 7 min long. Sampling rates range from 40 to 200 samples per second (sps), depending on the instrument type.

Two events have only synthetic seismic data. The $M_w8.7$ Cascadia001300 event is a simulated megathrust earthquake on the Cascadia Subduction zone offshore of Oregon, Washington, and California (Ruhl et al., 2017). Acceleration waveforms (50 sps) begin 1 min prior to origin time and are a total of 8.66 min long. The $M_w7.0$ Hayward4Hz earthquake is a simulated strike-slip rupture initiating on the downdip extent of the Hayward fault in Northern California (Rodgers et al., 2018). Velocity waveforms sampled at 40 sps and with frequencies up to 4 Hz were obtained from Rodgers et al. (2018); each begins 2.0 s before origin time and has a total duration of approximately 1.5 min.

**Table 1**
*List of Earthquakes and Their Source Parameters Used in This Study*

| | Event name, country | Origin time (UTC)[a] | Longitude | Latitude | $z$ (km) | Mw | Number of geodetic sites | Number of seismic sites | Mechanism |
|---|---|---|---|---|---|---|---|---|---|
| 1 | Tohoku2011, Japan | 2011-03-11T05:46:24 | 142.3720 | 38.2970 | 30.0 | 9.0 | 815 (288) | 211 | Reverse |
| 2 | Maule2010, Chile | 2010-02-27T06:34:14 | −72.7330 | −35.9090 | 35.0 | 8.8 | 27 | 7 | Reverse |
| 3 | Cascadia001300, United States (Synthetic) | 2016-09-07T07:00:00 | −124.6160 | 45.8638 | 19.8 | 8.7 | 62 | 40 | Reverse |
| 4 | Illapel2015, Chile | 2015-09-16T22:54:33 | −71.6540 | −31.5700 | 29.0 | 8.3 | 58 | 40 | Reverse |
| 5 | Tokachi2003, Japan | 2003-09-25T19:50:06 | 143.9040 | 41.7750 | 27.0 | 8.3 | 368 (189) | 313 | Reverse |
| 6 | Tehuantepec2017, Mexico | 2017-09-08T04:49:19 | −93.8990 | 15.0220 | 47.4 | 8.2 | 7 | 88 | Normal |
| 7 | Iquique2014, Chile | 2014-04-01T23:46:47 | −70.7690 | −19.6100 | 25.0 | 8.1 | 40 | 55 | Reverse |
| 8 | Ecuador2016, Ecuador | 2016-04-16T23:58:36 | −79.9220 | 0.3820 | 20.6 | 7.8 | 21 | 21 | Reverse |
| 9 | Kaikoura2016, New Zealand | 2016-11-13T11:02:56 | 173.0540 | −42.7370 | 15.0 | 7.8 | 39 | 34 | Strike-slip |
| 10 | Nepal2015, Nepal | 2015-04-25T06:11:25 | 84.7310 | 28.2310 | 8.2 | 7.8 | 7 | 4 | Reverse |
| 11 | Ibaraki2011, Japan | 2011-03-11T06:15:34 | 141.2653 | 36.1083 | 43.2 | 7.7 | 1,149 (432) | 278 | Reverse |
| 12 | Iquique_aftershock2014, Chile | 2014-04-03T02:43:13 | −70.4930 | −20.5710 | 22.4 | 7.7 | 17 | 0 | Reverse |
| 13 | Mentawai2010, Indonesia | 2010-10-25T14:42:22 | 100.1140 | −3.4840 | 20.0 | 7.7 | 13 | 0 | Reverse |
| 14 | N.Honshu2011, Japan | 2011-03-11T06:25:44 | 144.8940 | 37.8367 | 34.0 | 7.7 | 1,148 (230) | 387 | Normal |
| 15 | Melinka2016, Chile | 2016-12-25T14:22:26 | −74.3910 | −43.5170 | 30.0 | 7.6 | 58 | 12 | Reverse |
| 16 | Nicoya2012, Costa Rica | 2012-09-05T14:42:08 | −85.3050 | 10.0860 | 40.0 | 7.6 | 9 | 14 | Reverse |
| 17 | Iwate2011, Japan | 2011-03-11T06:08:53 | 142.7815 | 39.8390 | 31.7 | 7.4 | 1,149 (338) | 216 | Reverse |
| 18 | Miyagi2011A, Japan | 2011-03-09T02:45:12 | 143.2798 | 38.3285 | 8.3 | 7.3 | 892 (263) | 294 | Reverse |
| 19 | N.Honshu2012, Japan | 2012-12-07T08:18:20 | 144.3153 | 37.8158 | 46.0 | 7.3 | 978 (196) | 430 | Reverse |
| 20 | Nepal_aftershock2015, Nepal | 2015-05-12T07:05:19 | 86.0660 | 27.8090 | 15.0 | 7.3 | 5 | 0 | Reverse |
| 21 | ElMayor2010, Mexico | 2010-04-04T22:40:42 | −115.2800 | 32.2590 | 10.0 | 7.2 | 137 | 465 | Strike-slip |
| 22 | Miyagi2011B, Japan | 2011-04-07T14:32:43 | 141.9237 | 38.2028 | 60.7 | 7.1 | 1,137 (381) | 386 | Reverse |
| 23 | N.Honshu2013, Japan | 2013-10-25T17:10:18 | 144.5687 | 37.1963 | 56.0 | 7.1 | 59 (59) | 349 | Reverse |
| 24 | Puebla2017, Mexico | 2017-09-19T18:14:38 | −98.4890 | 18.5500 | 48.0 | 7.1 | 18 | 79 | Normal |
| 25 | Hayward4Hz, United States (Synthetic) | 2017-01-01T00:00:02 | −122.2850 | 37.9638 | 17.1 | 7.0 | 2,301 (231) | 2301 | Strike-slip |
| 26 | Kumamoto2016, Japan | 2016-04-15T16:25:05 | 130.7630 | 32.7545 | 12.5 | 7.0 | 277 (245) | 230 | Strike-slip |
| 27 | Aegean2014, Greece | 2014-05-24T09:25:02 | 25.3890 | 40.2890 | 12.0 | 6.9 | 6 | 139 | Strike-slip |
| 28 | Nisqually2001, United States (Synthetic Disp.) | 2001-02-28T18:54:32 | −122.7270 | 47.1490 | 51.8 | 6.9 | 26 | 63 | Normal |
| 29 | E.Fukushima2011, Japan | 2011-04-11T08:16:12 | 140.6727 | 36.9457 | 6.4 | 6.6 | 1,146 (476) | 260 | Normal |
| 30 | Lefkada2015, Greece | 2015-11-17T07:10:07 | 20.6002 | 38.6650 | 10.7 | 6.5 | 23 | 4 | Strike-slip |
| 31 | Napa2014, United States | 2014-08-24T10:20:44 | −122.3100 | 38.2150 | 11.0 | 6.1 | 224 (222) | 560 | Strike-slip |
| 32 | Parkfield2004, United States | 2004-09-28T17:15:24 | −120.3700 | 35.8150 | 7.9 | 6.0 | 13 | 309 | Strike-slip |

[a]Dates are formatted as year-month-day.

The remaining events are downloaded or obtained from local earthquake authorities in each country of origin (see Acknowledgments). Waveform lengths and sampling rates vary on a station-by-station and network-by-network basis; some are triggered stations that begin after the *P* wave. We did not apply a strict distance or GM amplitude cutoff beyond which we no longer consider records. In many cases, we collected all available seismic data associated with an event. For Japanese events, with significantly more stations recording each event, we use up to ~400 stations per event starting with the highest GMs based on the stations associated with each event. In total, we use 7,589 three-component seismic records for the ElarmS analysis.

### 2.2. Geodetic Data

The geodetic data set consists of high-rate GNSS observations for 29 real earthquakes worldwide from the open data set of Melgar and Ruhl (2018). The displacement waveforms were calculated in a uniform fashion using the precise point positioning approach of Geng et al. (2013). The overwhelming majority of the recordings are collected at 1 sps but a few (2010 $M_w$7.2 El Mayor-Cucapah, 2012 $M_w$7.6 Nicoya, 2014 $M_w$6.1 Napa, and 2015 $M_w$7.8 Nepal) have some 5 sps recordings. These data were resampled to 1 sps for use with G-larmS which currently processes data by the integer-epoch. The data were processed into 6-hr-per-channel text files to mimic the real-time trackRT format previously used at the Berkeley Seismological Laboratory.

We also use synthetic displacement data for three additional earthquakes. The $M_w$7.0 Hayward4Hz displacement data were created by integrating the seismic data simulated in Rodgers et al. (2018) and described in section 2.1. The Cascadia001300 synthetic data were developed using a hybrid semistochastic approach

developed by Melgar et al. (2016) and described in Ruhl et al. (2017). The Nisqually 2001 earthquake is a real event in Washington state that was recorded seismically, but displacements were simulated by Crowell et al. (2016). For these events, data from multiple stations are rewritten into one time-ordered horizontal and one vertical component text file per event. In total, we use 4,545 three-component geodetic records for the G-larmS analyses.

## 3. Methods

First, we replay seismic data from each earthquake through the ElarmS EEW algorithm in simulated real time to estimate event magnitudes, epicentral locations, and origin times. We then use the seismic first-alerts, as well as "perfect" alerts (i.e., true origin time and location), to trigger the Geodetic Alarm System (G-larmS) and generate distributed slip and magnitude evolution time series. Using those results, we predict shaking intensity (Modified Mercalli Intensity, MMI) time series for each seismic station for each event to compare to the observations. Finally, we employ an MMI threshold approach (Meier, 2017) to accurately characterize warning times (WTs) on a per station basis, thus enabling classification of true positive (TP), true negative (TN), false positive (FP), and false negative (FN) alerts for each event. Below we discuss the details of ElarmS (section 3.1), G-larmS (section 3.2), and the MMI threshold method used for classifying our simulated real-time alerts (section 3.3).

### 3.1. Seismic Alerts: ElarmS

ShakeAlert′s seismic point-source algorithm (EPIC) is a derivative of the Earthquake Alarm System (ElarmS), a network-based EEW algorithm developed at the Berkeley Seismological Laboratory over the last 10 years (Allen, Gasparini, et al., 2009; Kuyuk et al., 2013). Because there are only minor functional differences between EPIC and ElarmS, we test our data set using the latest version of ElarmS currently operating at the Berkeley Seismological Laboratory (Chung et al., 2017). ElarmS identifies and associates triggers and locates events epicentrally assuming a fixed depth or set of depths (8 and 20 km used in this study). Next, the algorithm estimates event magnitudes based on $P$ wave amplitudes and distances to its estimated epicenter. ElarmS then generates earthquake alerts when a minimum of three stations with at least 0.2 s of data meet region-specific spatial constraints (e.g., station density is taken into consideration).

We create earthquake tank-player files and channel files containing all data for each earthquake and run them through ElarmS in simulated real time. When ElarmS identifies an event, it outputs estimates of origin time, magnitude, and location, as well as solution information such as number of stations. As additional stations trigger and seismic data develop, ElarmS refines and adjusts its source parameters and issues updated alerts. We retain a list of alert parameters in a separate log file for each event.

### 3.2. Geodetic Alerts: G-larmS

The Geodetic Alarm System (G-larmS) incorporates real-time GNSS data into EEW systems (Grapenthin et al., 2014a). In real time operation, G-larmS continuously analyzes positioning time series and is capable of ingesting both relative displacements (baselines) and absolute positions from precise-point-positioning solutions. During an earthquake, the ShakeAlert seismic system issues event messages containing hypocenter and origin time that trigger G-larmS to estimate static offsets epoch-by-epoch at each site. Simultaneously, it inverts these static offsets for distributed slip on a finite-fault. The latest version of G-larmS builds a linear fault using region-specific a priori geometries and, in addition, attempts to fit the event by imposing slip onto nearby known faults, allowing for complex geometries. In the first case, G-larmS centers the model fault plane on the earthquake hypocenter provided by ShakeAlert and allows the fault to grow symmetrically based on scaling relationships (Wells & Coppersmith, 1994). Model fault plane orientations are predefined for expected tectonic regimes based on location and Green′s functions are calculated in real time. This means that, for instance, for an event in the San Francisco Bay Area, it will be modeled using linear San Andreas fault (SAF) parallel, SAF conjugate, and SAF splay ($\pm 5°$ from SAF) geometries. For the latter case, so-called "catalog faults" are built into the system by simplifying models of large faults (e.g., UCERF3, Field et al., 2014; Slab1.0, Hayes et al., 2012). Therefore, for the San Francisco Bay Area example, an event in Oakland, CA, is modeled with slip imposed onto the San Andreas and Hayward faults (separately) as well as on the growing, linear regional geometries. One benefit of using catalog faults is that Green′s functions can be precomputed for fixed station sets saving computation time during inversion.

Another benefit is that curving faults such as megathrusts or complex strike-slip faults (e.g., big bend of SAF) can be modeled more accurately than with the linear tectonic regime faults. At each epoch, the geometry that minimizes the model misfit to the data is selected as the preferred solution. A detailed description of the original algorithm can be found in Grapenthin et al. (2014a, 2014b) and previous performance with synthetics offshore Cascadia can be found in Ruhl et al. (2017).

In both replay (real events) and simulation mode (synthetic events), G-larmS is run in two separate steps, rather than simultaneously estimating offsets and finite-fault parameters as in the real-time system. The first module is the Offset Estimator (OE), which calculates and stores the coseismic (static) offsets, and the second is the Parameter Estimator (PE) that actually inverts the offsets for slip on a finite fault. This separation is more efficient in an offline, personal computer-based implementation. We run both the OE and PE twice: once using ShakeAlert XML messages created from the seismic first alert to trigger G-larmS (described in section 3.2.1), and then again using messages containing the exact hypocenter and origin time as a perfect alert to trigger G-larmS (section 3.2.2).

The G-larmS OE uses the ShakeAlert style XML event message to determine a start time for offset estimation at each station within a specific radius based on the event location, magnitude, origin time, and a configurable wave speed. Because static offsets typically arrive with the $S$ wave, choosing a shear-wave velocity (~3 km/s) is often the preferred or recommended approach. However, a comparison of finite-fault solutions based on offset estimations started at estimated $P$ wave and $S$ wave arrival, respectively, showed that starting the offset estimation earlier resulted in damping of early offsets and, therefore, damping of the finite-fault solutions (Ruhl et al., 2017). This is acceptable since near- and intermediate-field oscillating dynamic displacements can sometimes inflate initial static displacement (i.e., offset) estimates. Also, using a faster velocity can account for error in origin times and locations (i.e., prevent missing initial offsets) and may be more representative of average crustal velocities for deeper events. In this paper, we use a velocity of 5.2 km/s for all events, regardless of location or tectonic setting. G-larmS then calculates and stores the mean displacement amplitude before the calculated start time. These are subtracted from average displacements following the start time to estimate static offsets.

For the real earthquakes, we reformat the data into 6-hr GPS time series for each station-component and store them in text files in GPS time units (i.e., without leap second adjustments). These are ingested in batches by the OE in faster-than-real-time replays and offsets are written to additional log files to be ingested later by the PE. For the simulated displacement data sets, we store horizontal displacement data for all stations in one time-ordered text file and vertical displacements in another. Random noise ($\pm$2.5 and $\pm$4.0 cm for horizontal and vertical components, respectively) is added to the displacements as they are read and offsets are estimated and written into individual log files. Once all offsets are calculated, the PE reads the offset logs and begins the slip inversion in real time based on the ShakeAlert XML message and the region-specific fault configurations.

G-larmS calculates earthquake magnitudes at each epoch based on the overall fault geometry and amount of slip imposed on it. Outputs include subfault geometries and the magnitudes of strike-slip and dip-slip components of slip per subfault. We simplify this information by calculating a surface-projected perimeter around the subfault patches that have slip greater than 10% of the maximum slip; we ensure this perimeter includes the hypocenter, even if it is located on a subfault that has less than 10% of the maximum subfault slip amount. This value (10% of maximum slip) was determined empirically by Ruhl et al. (2017) using 1,300 synthetic earthquakes in the Cascadia subduction zone. The perimeter is used to calculate the closest epicentral distance, $R_{jb}$, to the fault necessary for GM prediction. The actual subfault patches within this perimeter are then used to calculate $R_{Rup}$, the closest hypocentral distance to the fault rupture itself. We then calculate the dominant rake of the overall fault based on the amount of strike-slip and dip-slip and characterize the solution as a rectangular fault with pure and uniform reverse (90), normal (−90), dextral (−180), or sinistral (180) slip. This information is used for GM prediction (see section 3.3).

### 3.2.1. ElarmS-Triggered G-larmS

For a true real-time comparison, we first trigger G-larmS using the ElarmS first alerts. These are referred throughout the text and in figures as "ElarmS-Triggered G-larmS" solutions. We use the magnitude, epicentral location, and origin time of the first ElarmS solution to build the XML event messages for each event; depths are fixed to either 8 or 20 km, depending on the first alert. G-larmS replays always begin at the

origin time, therefore solutions are calculated for each epoch as soon as the estimated *P* wave reaches the closest GNSS station. Because geodetic sites may be closer than seismic sites, this may result in unrealistically timed solutions (i.e., before the ElarmS solution exists). Therefore, we remove all solutions before the ElarmS first alert plus one epoch.

### 3.2.2. Perfectly Triggered G-larmS

In addition, we calculate perfect G-larmS solutions assuming that we know exactly where each earthquake occurs at exactly the origin time. These are referred throughout the text and figures as "Perfectly Triggered G-larmS" solutions. We use M6.0 for the initial magnitude of each event as well as the exact depth as reported in the published catalog locations (Table 1). We employ this approach to assess how much the simulated real-time environment degrades a perfect solution. For discussion of these results, please see the supporting information.

### 3.3. Real-Time Alert Classification: MMI Threshold Approach

EEW is inherently a GM warning system because users' actions depend on the level of shaking intensity expected at their individual sites, rather than on the magnitude or location of the earthquake. Because most algorithms provide the latter information without GM estimates, it is difficult to assess an EEW system based on source parameters alone. Instead, Meier (2017) suggested to develop quantitative metrics, such as WT, and to classify alerts using shaking intensity thresholds on a per station basis. We follow that approach and assess our results with respect to GM in addition to comparing magnitude estimates.

First, we process the observed data by converting each seismic waveform to envelopes of peak ground acceleration (PGA) and peak ground velocity (PGV). We combine the individual components into one PGA and one PGV envelope by taking the maximum of the three components. We then use both PGA and PGV to convert the data to observed maximum instrumental MMI time series using the method of Worden et al. (2012). A total of 5,151 sites associated with ElarmS-alerted events are used.

Next, we compute predicted PGA and PGV for each station from the ElarmS, ElarmS-Triggered G-larmS, and Perfectly Triggered G-larmS solution time series at each epoch. For all three solution types and for all global earthquakes in our study, we use the GM prediction equations of Abrahamson et al. (ASK14, 2014). For both G-larmS and ElarmS we use the same site-specific $V_S30$ value extracted from a slope-based global database from the USGS (Wald & Allen, 2007). We use dip and rake angles simplified from the G-larmS finite-fault solutions based on the average rake of the subfaults. The dips are 15°, 90°, 90°, and 60° for any fault with primarily reverse dip-slip, normal dip-slip, sinistral strike-slip, and dextral strike-slip, respectively. The simplified rakes are 90°, 180°, −180°, and −90° for each fault type, respectively. The GM estimates are controlled by varying only three input parameters: each algorithm uses its own $M_w$ estimate per-epoch, site-specific distance metric $R_{JB}$, and fault width W. $R_{JB}$ is the closest horizontal distance to the surface projection of the rupture (G-larmS) or point-source location (ElarmS). For the G-larmS solutions, we calculate $R_{JB}$ and $R_{Rup}$ using the finite-fault described in section 3.2. We then calculate width W for all solutions based on $M_w$ estimates and the tectonic regime inferred from the simplified rakes using empirical relationships from Wells and Coppersmith (1994). All other parameters are held constant for GM predictions calculated for all three algorithms.

We calculate and store predicted PGA and PGV values as time series based on the solutions that update each epoch. Predicted PGA and PGV are then converted and combined into a maximum-MMI envelope time series in the same manner as the observations. We compare the observed and predicted MMI time series with respect to a specific MMI alerting threshold on a per station basis at all the available seismic sites for any given event. If both the observations and predictions exceed the specified MMI threshold, the WT is defined as the time difference between the observed threshold crossing and the time at which the threshold crossing was first predicted. For instance, long WTs indicate that the prediction came well before the site experienced shaking equivalent to the MMI threshold. A short WT, on the other hand, means that MMI threshold-exceeding shaking follows quickly after its prediction. If the WT is negative, the threshold-crossing GM alert was issued after actual GM already exceeded the threshold. Only alerts with positive WTs are considered TP alerts. Late warnings (negative WTs) are classified as FNs or missed alerts. We calculate median WTs for all TP sites, as well as for all sites alerted, even if late (Table S3). If neither the observations nor predictions cross the threshold, no WT is calculated, no alert is issued, and it is classified as a TN end-user experience. If the observation crosses the threshold, but the prediction does not, no WT is calculated and it is classified as a FN
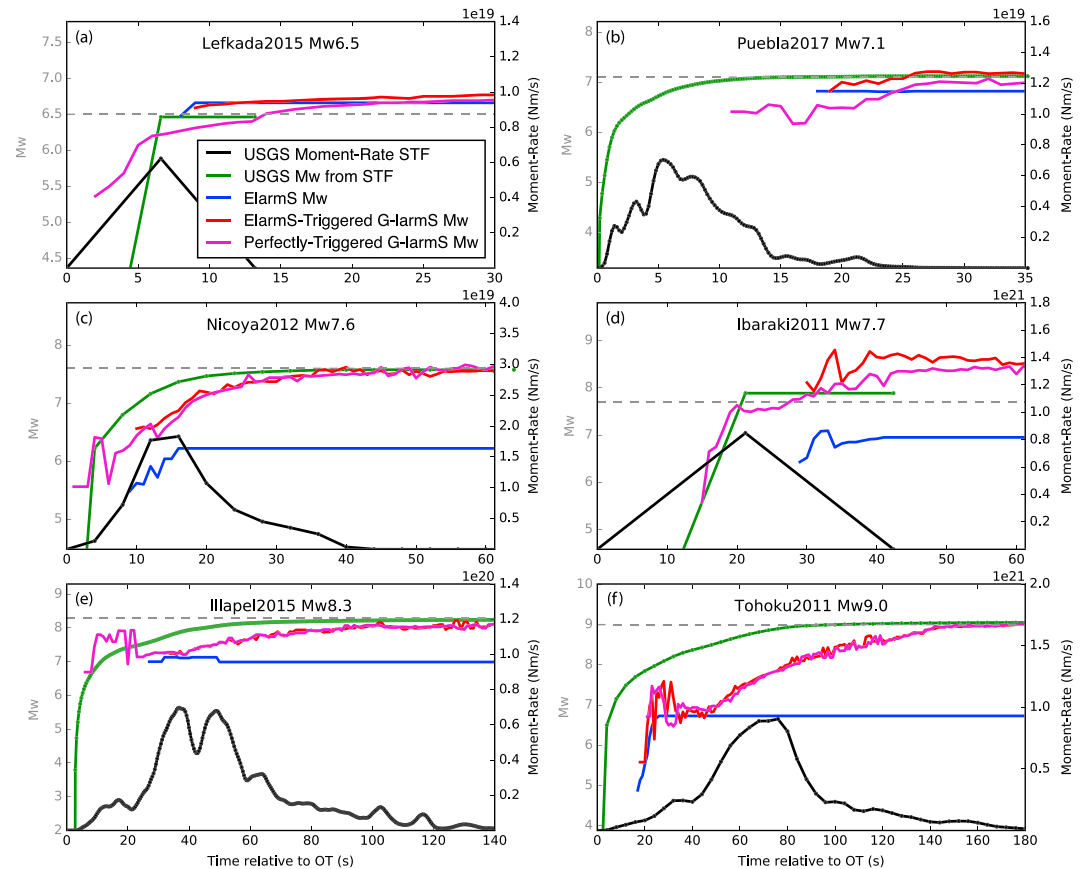
**Figure 1.** Magnitude estimate evolutions for six earthquakes used as examples throughout this paper. The magnitude of the examples increases from the top left to the bottom right and includes the (a) 2015 $M_w$6.5 Lefkada, (b) 2017 $M_w$7.1 Puebla, (c) 2012 $M_w$7.6 Nicoya, (d) 2011 $M_w$7.7 Ibaraki, (e) 2015 $M_w$8.3 Illapel, and (f) 2011 $M_w$9.0 Tohoku earthquakes. In each panel, the black curve is the USGS Moment-rate function derived from finite-fault source inversion or a triangle set to the width of twice the half-duration from moment tensors solutions published by the USGS. Green curves are the corresponding moment magnitude evolution for the black STF curves. Blue curves are from ElarmS, red curves are from ElarmS-triggered G-larmS solutions, and magenta curves are for Perfectly Triggered G-larmS solutions. The dashed gray line shows the final Mw.

along with the late alerts. And finally, if the prediction crosses the threshold, but the observations do not, no WT is calculated and it is classified as a FP. Classifying alerts into these four categories allows quantification of the performance of an EEW system in GM space (Meier, 2017; Minson et al., 2019). We repeat the MMI threshold calculations for all seismic station sites for all events using the ElarmS, ElarmS-Triggered G-larmS, and Perfectly Triggered G-larmS solutions for thresholds from MMI 3 to 7.

We do not include Hayward4Hz in the MMI analysis because it has >2,000 stations only at very close distances and would dominate the results. The Cascadia001300 data, on the other hand, includes only 40 seismic sites over a range of distances and their influence on statistics is therefore more representative of real observations.

It is worth noting that aside from transmission delays (both in receiving data and issuing alerts), the timing of our results is entirely realistic, and, indeed, these added latencies are typically quite short (~1s). Our GM estimates were calculated offline, but eqInfo2GM, a GM module currently under development by ShakeAlert, reports negligible calculation times less than 1.5 s, even for finite-source models.

## 4. Results

ElarmS results were obtained for 26 out of the 29 earthquakes with seismic data. Thus, we have a total of 26 first-alert triggers from which we recovered ElarmS-Triggered G-larmS solutions. The 2016 $M_w$7.8
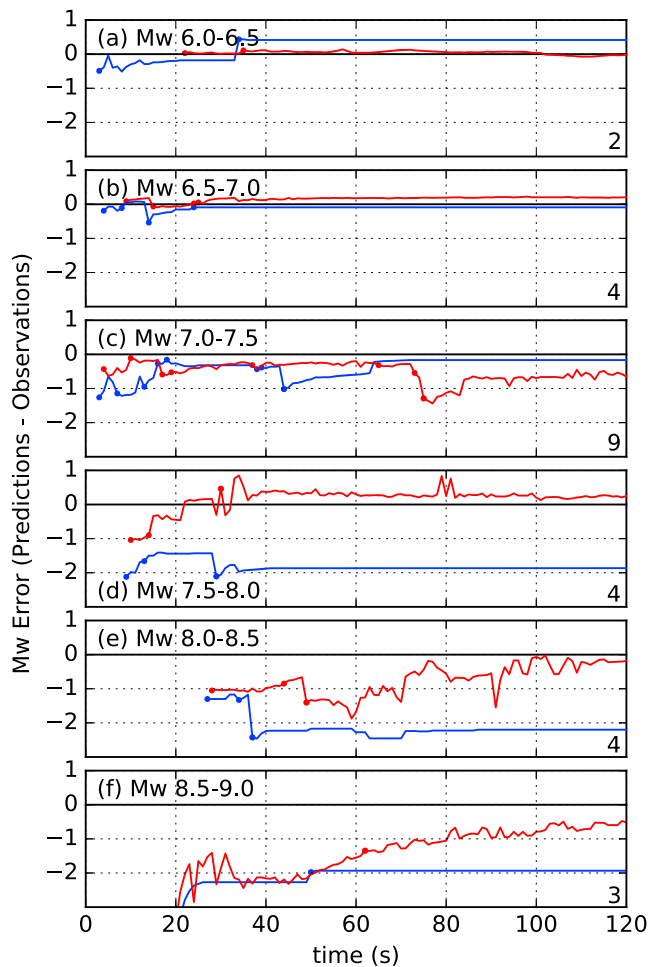
**Figure 2.** Mean magnitude error evolutions for six magnitude bins (a) $6.0 \leq M < 6.5$, (b) $6.5 \leq M < 7.0$, (c) $7.0 \leq M < 7.5$, (d) $7.5 \leq M < 8.0$, (e) $8.0 \leq M < 8.5$, (f) $8.5 \leq M \leq 9.0$. Errors are calculated as the predictions minus the observations such that a negative number shows magnitude saturation. Solutions were averaged for all events within the magnitude bin (total labeled in bottom right of each panel) for ElarmS (blue) and ElarmS-Triggered G-larmS (red) solutions as the alerts came in. Points on each curve show the first alert times of additional events included in each mean.

Ecuador, 2015 $M_w7.8$ Nepal, and 2011 $M_w7.7$ N. Honshu, Japan, earthquakes did not produce seismic alerts due to poor station coverage or very far distances to the few closest stations. In the following sections, we first discuss the timeliness and accuracy of magnitude estimates from the algorithms, then present the MMI threshold results and discuss the timeliness and accuracy of the simulated real-time alerts. Throughout this section, we use the 2015 $M_w6.5$ Lefkada, 2017 Mw7.1 Puebla, 2012 Mw7.6 Nicoya, 2011 Mw7.7 Ibaraki, 2015 M8.3 Illapel, and 2011 $M_w9.0$ Tohoku-oki earthquakes as examples to illustrate the performance of the various algorithms over a range of magnitudes. In addition to spanning the magnitude range that we tested, these events are recorded on as few as four and up to hundreds of stations and also exhibit different focal mechanisms: Lefkada is a continental strike-slip earthquake, Puebla is a relatively deep intraslab normal earthquake, and the remaining four are reverse events. Individual results for these six events are shown in Figures 1, 5, and 6; results for all other events are available in the supporting information (Table S1 and Figures S1–S32). Figures 2–6 and 8–10 compile results for all events from which statistics are computed.

### 4.1. Accuracy and Timeliness of Magnitude Estimates

Magnitude time series plotted for ElarmS (blue), ElarmS-Triggered G-larmS (red), and Perfectly Triggered G-larmS (magenta) solutions in Figure 1 demonstrate the algorithm performance for the six example earthquakes. Geodetic magnitude estimates tend to approach the final magnitude (green), tracking modeled magnitude evolutions derived from published moment rate functions (black, Figure 1). The difference between the blue ElarmS magnitude estimates and the red and magenta G-larmS solutions tends to increase with increasing magnitude (top-left to bottom-right in Figure 1), revealing saturation in the seismic-only point-source solutions. Magnitude-binned and averaged errors also show a significant increase in ElarmS magnitude saturation above M7.5, while magnitude errors for G-larmS are more stable with respect to increasing magnitude (Figure 2).

Figure 3 demonstrates the significant magnitude accuracy improvement by comparing estimates at three stages: the first time an alert is available (i.e., the first alert), the alert at 30 s after origin time (if available), and the final alert in the rep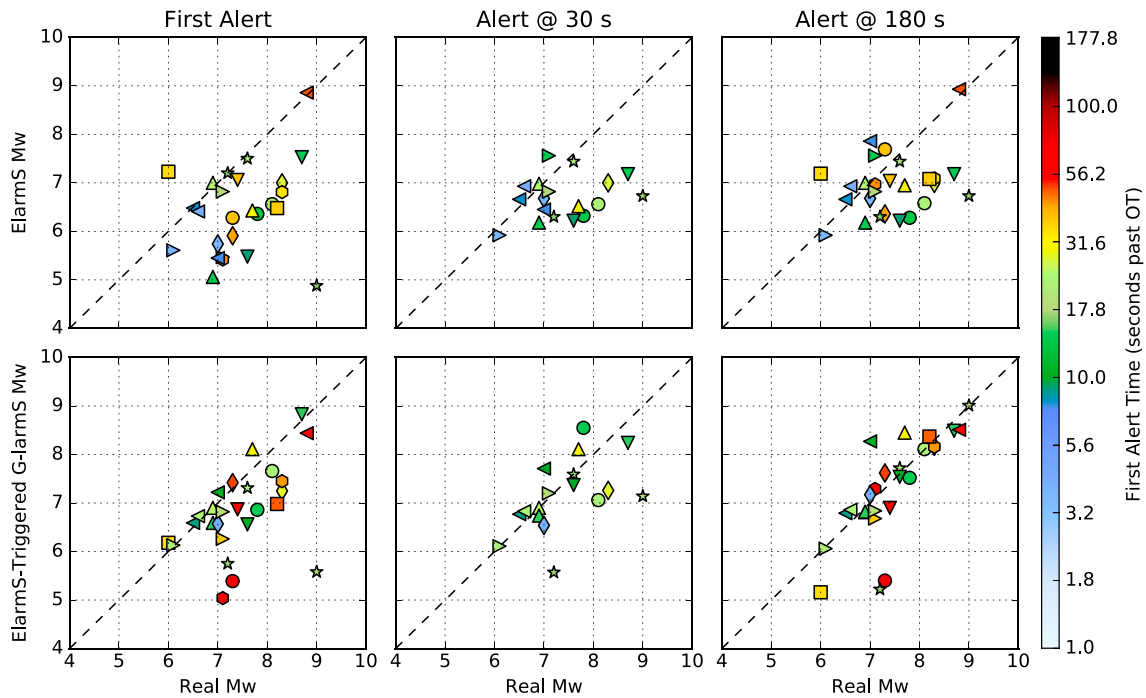lay (Figure 1 and Table S1). Alert times are calculated relative to origin time. For ElarmS, the final alert time is the last update produced shortly after the last triggered station arrives. For G-larmS, all solutions were estimated until 180 s after origin time, which is the final solution time. ElarmS magnitude errors are $-1.0 \pm 1.0$ (mean ± sample standard deviation) for the first alert, $-0.71 \pm 0.75$ at 30 s, and $-0.50 \pm 0.83$ at the final update. ElarmS-Triggered G-larmS magnitude errors are $-0.62 \pm 0.86$ for the first alert, $-0.26 \pm 0.73$ at 30 s, and $-0.14 \pm 0.65$ at the final update around 180 s. G-larmS provides an improvement of ~0.5 magnitude units, on average, by 30 s after origin time for all events (Figure 3). This improves with increasing magnitude as shown in Figure 2 and is often present in first alerts far earlier than 30 s. G-larmS first- and final-alert magnitude estimates are both statistically more accurate than ElarmS magnitude estimates, but as expected, the triggered geodetic algorithm takes longer, on average, to issue its first alerts than the seismic system (Figure 4). Mean first alert times for ElarmS and ElarmS-Triggered G-larmS are $22 \pm 13.7$ s (mean ± sample standard deviation) and $31 \pm 20.5$ s, respectively (Table S1). Figure 4 does show that short first-alert times are achievable for the triggered geodetic system and that the bulk of the distribution indicates comparable first alert times for the respective algorithms. The better final accuracy of the G-larmS results is reflected by the nearly 1:1 ratio in the bottom-right panel of Figure 3 (see also Figure S36).

**Figure 3.** Magnitude estimates compared to real magnitudes for each algorithm for the first alerts (left column), alert at 30 s (middle column, where available), and at 180 s (or the final alert, right column). Top row subplots show ElarmS magnitude estimates and bottom panels show ElarmS-triggered G-larmS magnitude estimates. Dashed lines show 1:1 ratio in all subplots. Events are colored by the First Alert time in all three columns and maintain the same shape. Note that if the first alert was after 30 s, the event will not be plotted in the second column.
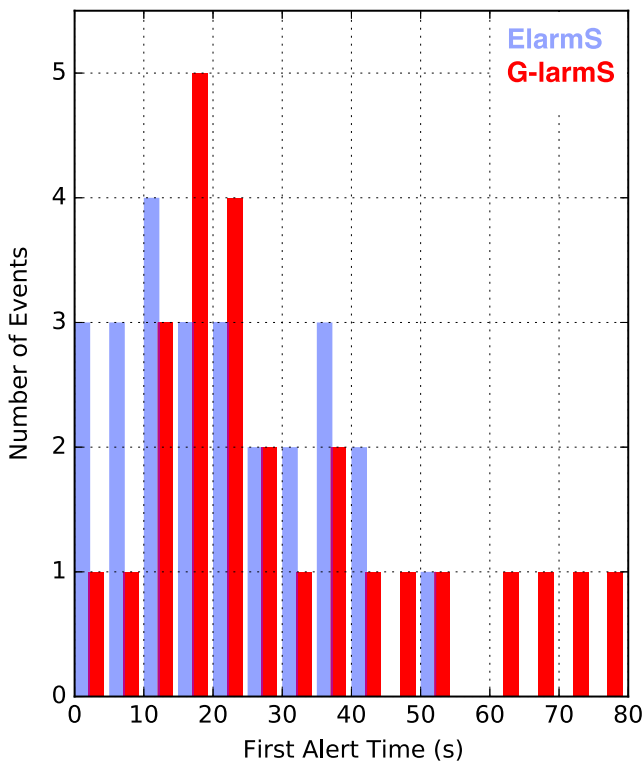


**Figure 4.** Histogram of first alert times (per event) for ElarmS (blue) and ElarmS-triggered G-larmS (red) results.

### 4.2. Accuracy and Timeliness of GM Estimates

In order to understand, from the end-user perspective, whether the improved, but delayed, geodetic earthquake characterizations are useful, we study the MMI performance of each solution. To demonstrate the technique, we show the MMI envelopes for two close-in stations per example earthquake with distances noted in Figure 5. The multicolored curve with a black center line is the MMI envelope of each station and the gray shaded area shows the WT for the end-to-end test using a threshold of MMI 4. The shaded area spans from the time that either ElarmS or ElarmS-Triggered G-larmS MMI predictions exceed the MMI threshold to the time when the data exceeds the same threshold. Sometimes the final shaking is over-estimated (Figure 5a) and other times it is underestimated (Figure 5b), but the threshold approach enables us to look at accuracy in terms of binary alert classification. Even though GMs may be overestimated or underestimated, as long as they are above or below the user's threshold of interest, the alert is useful and considered a success or true alert (made up of TPs and TNs). Geodetically inferred TP alerts are correctly issued for all 12 site examples in Figure 5; ElarmS, however, did not predict ground shaking stronger than the threshold at the bottom-right five stations. This means that, at least at these locations, it never issues a warning for users who will experience shaking greater than MMI 4.

Alert classifications for all stations recording the six example events are shown in Figure 6 with an MMI 4 threshold. For the smallest event, the 2015 $M_w$6.5 Lefkada earthquake, there is no significant improvement from the coupled seismic-geodetic solution (Figure 6a). This is not surprising since seismic saturation is not an issue at this magnitude; however, for
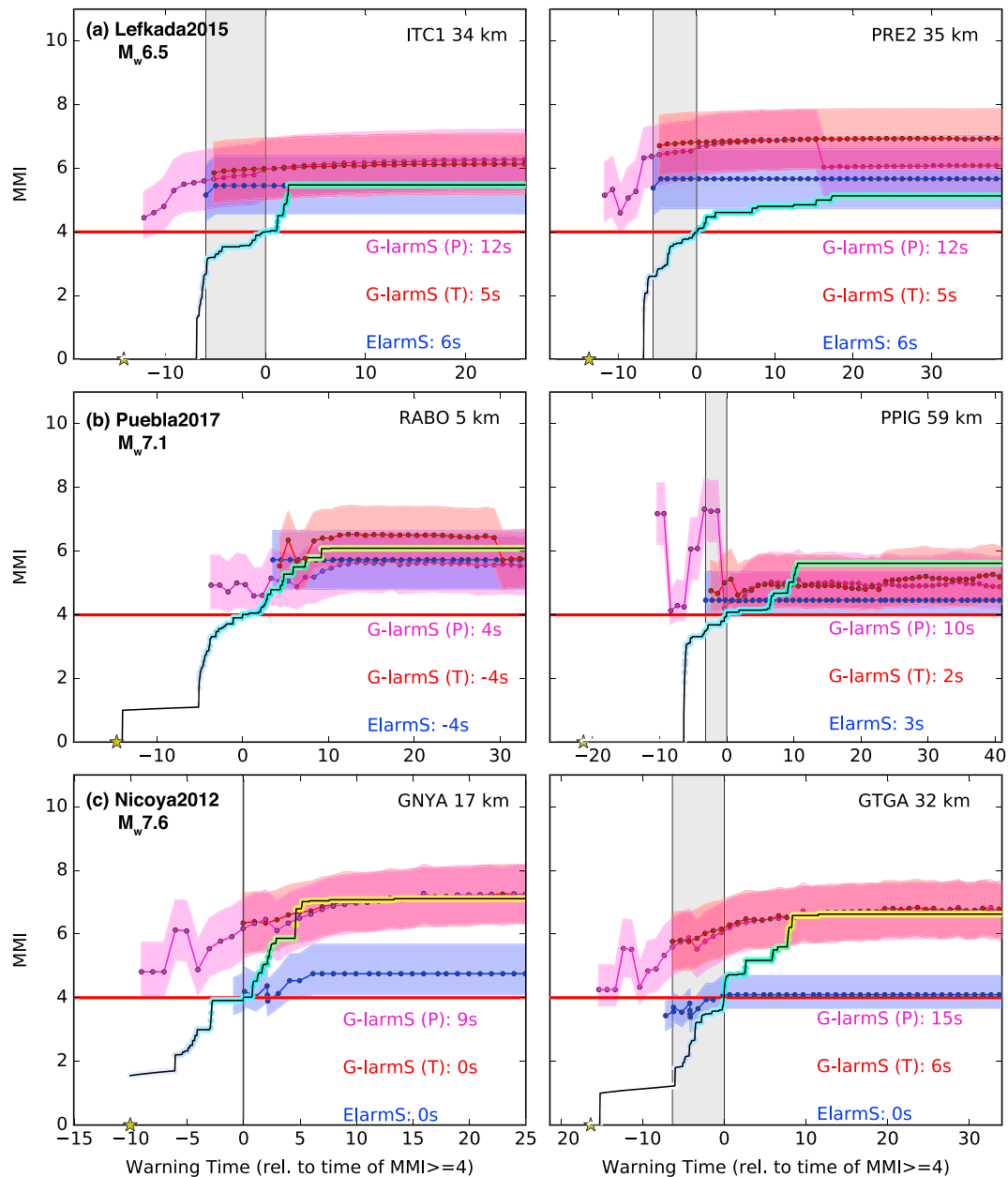
**Figure 5.** Examples of the MMI threshold method of alert classification for individual stations. We show two example stations for each of the six example earthquakes in Figure 1 including the (a) 2015 $M_w6.5$ Lefkada, (b) 2017 $M_w7.1$ Puebla, (c) 2012 $M_w7.6$ Nicoya, (d) 2011 $M_w7.7$ Ibaraki, (e) 2015 $M_w8.3$ Illapel, and (f) 2011 Mw9.0 Tohoku earthquakes. The station name and hypocentral distance are labeled in the top right corner of each figure. Red horizontal line shows the threshold of MMI 4. Colored curve shows the maximum MMI envelope of three component seismic data for each station. Blue curves with shaded regions show MMI estimates from ElarmS $\pm1\sigma$. Red curves with shaded region show MMI estimates from ElarmS-triggered G-larmS solutions $\pm1\sigma$. Magenta curves show MMI estimates from Perfectly Triggered G-larmS solutions $\pm1\sigma$. Warning times calculated based on the crossing of the MMI threshold are shown in bottom-right-hand corner of each subplot for the three algorithms. The gray shaded regions show the maximum warning times achieved from either ElarmS or the ElarmS-triggered G-larmS solutions. MMI = Modified Mercalli Intensity.

the larger events, there is remarkable improvement in the GM estimates. To further understand the performance of the algorithms, we synthesize the results from all 26 events recorded on 5,151 seismic stations. As described in section 3.3, we classify TP, TN, FP, and FN user-alerts (Figure 8 and Table S2) and compute WTs for all TP sites for all events (Figures 7, 8, and 9 and Table S3). Here, we focus on ElarmS-Triggered G-larmS results and use "G-larmS" to mean this version of the algorithm. For discussion of Perfectly Triggered G-larmS results, please see supporting information text and Figures S33–36.
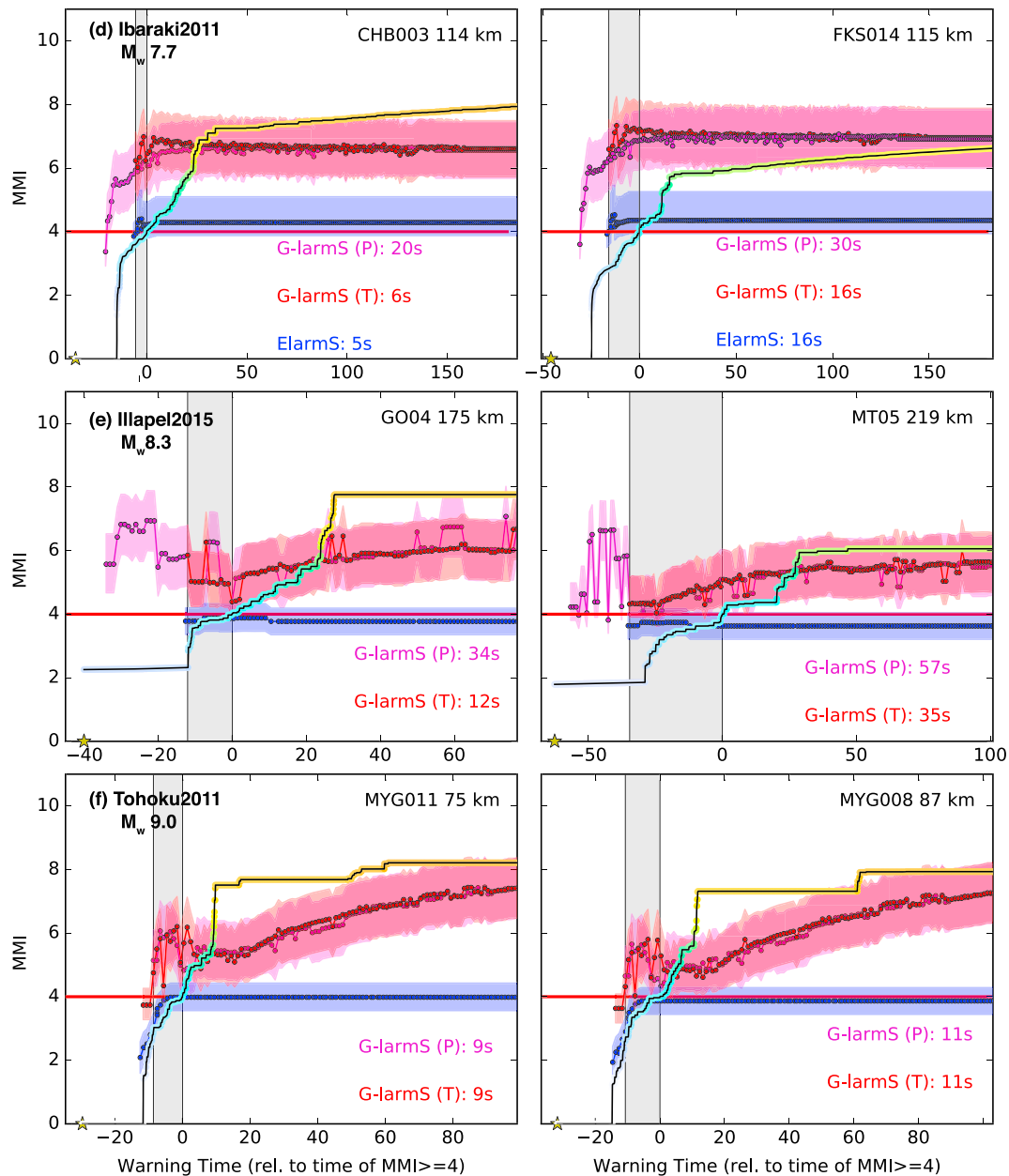
**Figure 5.** (continued)

For this large magnitude data set, G-larmS produces more positive alerts overall, both FP and TP, and has longer WTs than ElarmS. Using an alerting threshold of MMI 4, out of a total of 5,151 stations, ElarmS produced 10.4% TPs while ElarmS-Triggered G-larmS resulted in 39.9% TPs (Table S2 and Figure 6). The total amount of true alert classifications (TP + TN) increases by ~17 percentage points when using ElarmS-Triggered G-larmS over ElarmS alone. This primarily represents a large decrease in missed alerts (FN), from 48.7% with ElarmS to 19.2% with G-larmS, but it is important to note that FP alerts increase from 1.2% of all sites with ElarmS to 13.4% of all sites with G-larmS. Of the 3,917 sites that actually experienced shaking exceeding MMI 4, ElarmS alerted only 536 (13.7%), while G-larmS successfully alerted 2,055 (52.3%). The number of sites experiencing each MMI level is compared to the number of sites accurately warned by each algorithm using alerting thresholds of MMI 3 and MMI 4 in Figure S37.
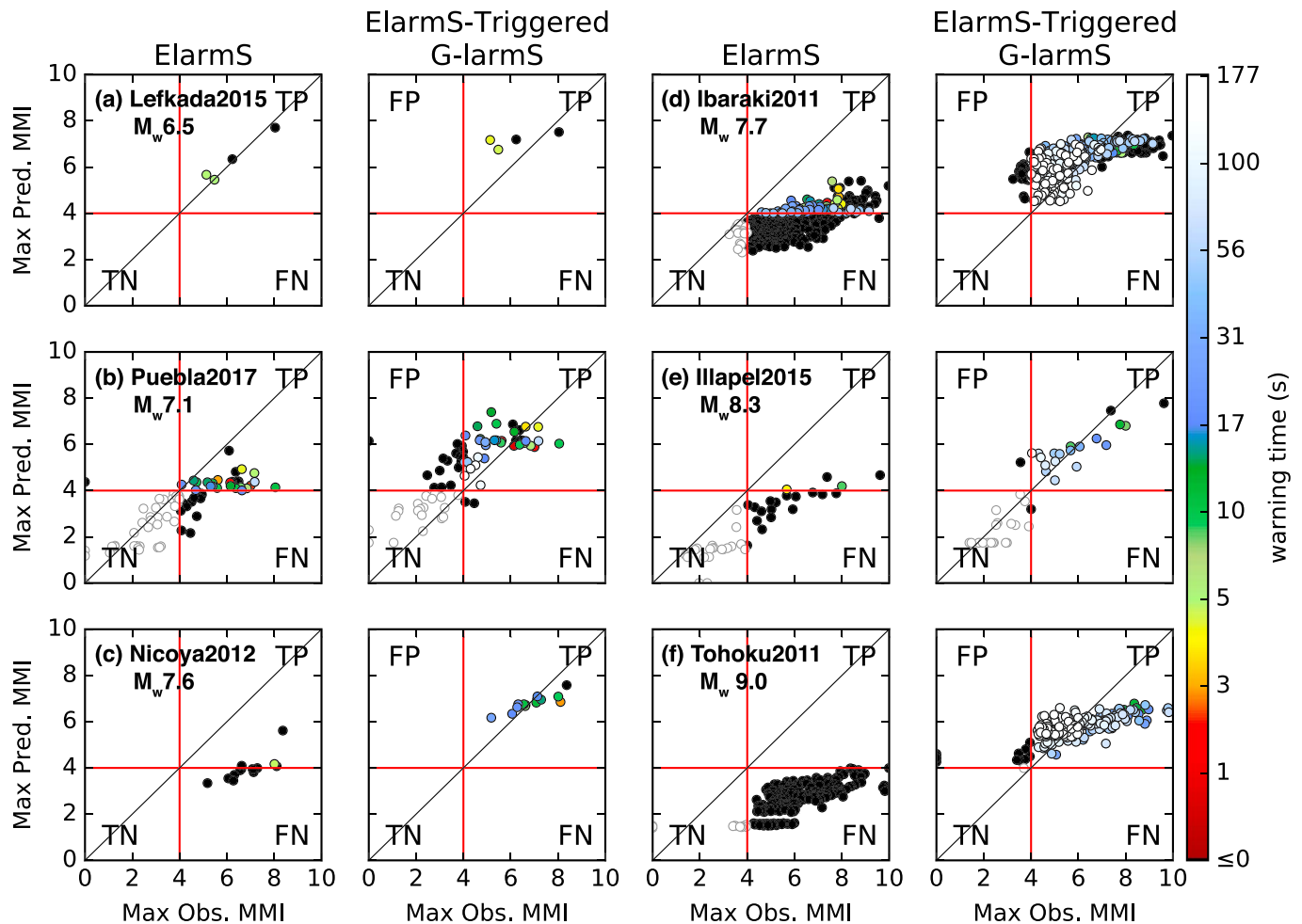
**Figure 6.** Real-time alert classification plots for the six example earthquakes shown in Figures 1 and 5 including the (a) 2015 $M_w6.5$ Lefkada, (b) 2017 $M_w7.1$ Puebla, (c) 2012 $M_w7.6$ Nicoya, (d) 2011 $M_w7.7$ Ibaraki, (e) 2015 $M_w8.3$ Illapel, and (f) 2011 $M_w9.0$ Tohoku earthquakes. Each event has two subplots showing alert classifications for ElarmS and G-larmS solutions. Each data point is colored by the warning time at that station based on a threshold of MMI 4 (red lines). Quadrants are labeled as TP, TN, FP, or FN. Note that the Max. predicted MMI is the maximum predicted after the final solution, not the maximum predicted MMI before the alerting threshold is exceeded by the station. TP = true positive; TN = true negative; FP = false positive; FN = false negative; MMI = Modified Mercalli Intensity.

Choosing alerting thresholds of MMI 3 or 5, instead of MMI 4, also leads to more TPs when using G-larmS compared to ElarmS alone. The numbers of TPs are higher for both algorithms when using an MMI 3 threshold (43.5% for ElarmS compared to 65.2% for G-larmS) and lower when using an MMI 5 threshold (1.8% for ElarmS compared to 19.9% for G-larmS). Using an alerting threshold of MMI 3, the number of sites experiencing shaking greater than MMI 5 for ElarmS and G-larmS is actually very similar and represents a high percentage of the observations (Figure S37). This supports using a lower alerting threshold to improve alert accuracy for larger MMIs for the seismic algorithm.

The huge increase in the number of TP alerts when comparing ElarmS and G-larmS is exemplified by the WT distributions shown in Figure 7. Using a threshold of MMI 4, ElarmS has a median WT of $18.9 \pm 19.7$ s. With G-larmS alerts, median WT increased to $55.8 \pm 46.3$ s—more than twice the WT of ElarmS alone. For ElarmS, median WTs increase to 35.0 s using MMI 3 and decrease to 7.6 s using MMI 5. This shows that for a seismic system detecting large earthquakes, a lower shaking intensity threshold will result in more TPs with longer WTs than a higher threshold with fewer TPs and shorter WTs. Notably, G-larmS results do not show the same relationship to MMI threshold choice: median WTs decrease from 55.8 s with a threshold of MMI 4 to 52.4 and 54.4 s when using thresholds of MMI
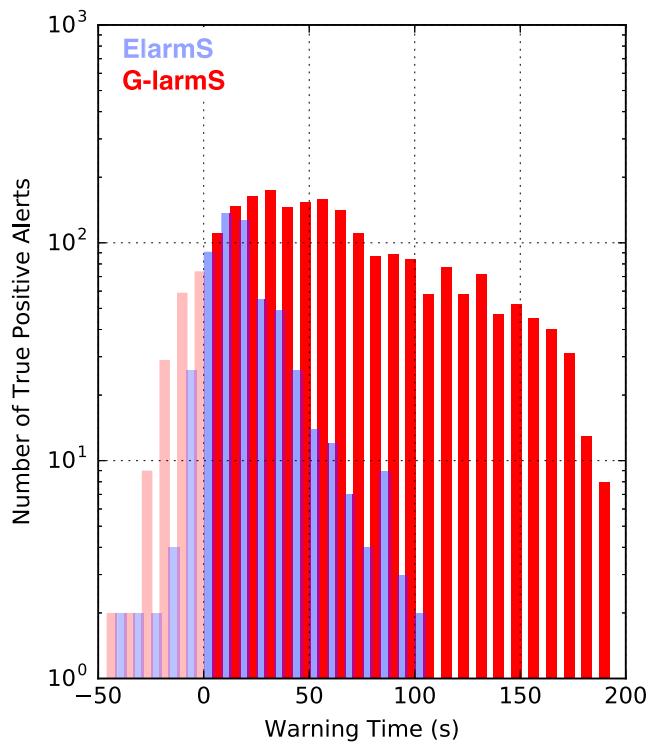
**Figure 7.** True positive warning times for ElarmS (blue) and ElarmS-triggered G-larmS (red) results with a Modified Mercalli Intensity 4 threshold.

3 and 5, respectively. This suggests that G-larmS' WTs are considerably less sensitive to MMI threshold than ElarmS.

To further explore timeliness, we plot the distributions of WTs at stations exceeding various shaking intensity levels for each algorithm (Figure 9). This approach exhibits the maximum possible WTs for each shaking intensity level for our results. In general, maximum possible WTs decrease with increasing observed MMI. These results are consistent with those of Meier (2017) and Minson et al. (2019). Notably, the long WTs for large MMI stations shown in Figure 9 are quite similar to probabilistic WTs calculated for the San Francisco Bay Area (Allen, 2006).

## 5. Discussion

As shown in Figure 9, when weighed against real data, the performance of a coupled seismic-geodetic system is better than a seismic-only point-source system; useful WTs are routinely achievable for large events with high GMs. Earthquake source processes can be complex and GM prediction equations have substantial uncertainties attached to them, thus some proportion of false alerts remain (Table S2); EEW systems and users that consider and effectively deal with these uncertainties based on some previously defined tolerance level will be important going forward. For example, a user's tolerance can be quantified by a cost ratio $r$ defined as the ratio between the cost of damage to the cost of taking action to avoid said damage (Aagaard et al., 2018; Minson et al., 2017; Minson et al., 2019). A value of $r \gg 1$ represents a user that is false alert tolerant, while $r < 1$ is representative of a user that is intolerant to false alerts (it costs more to take action than to sustain damage). Considering a nuclear power plant, it may be very costly to shut down, therefore, the cost of action is very high in comparison to the cost of potential damage from uncertain shaking estimates; this would result in a very small $r$ value for that particular user. For students in a classroom, however, the cost of getting under a desk for a few seconds to minutes has a very small cost compared to the cost of injury or death that may result during very strong shaking; this scenario results in a very high cost ratio and a higher tolerance for false alerts. Likewise, slowing or stopping a train is a relatively low cost of action compared to the train potentially derailing as a result of strong ground shaking; there are many such scenarios where the cost ratio is much greater than one.

One way to quantify the success or usefulness of an EEW system is to calculate the Cost Savings Performance Metric Q for theoretical users with different cost ratios (Aagaard et al., 2018; Minson et al., 2017; Minson et al., 2019). Q is defined as a function of the cost ratio and real-time alert classifications:

$$Q = \frac{TP - \frac{FP + F_{ALL}}{r-1}}{TP + FN}. \tag{1}$$

Note that this metric is user-dependent, as its value will change based on each user's cost ratio, $r$, and thus, when the respective Q is greater than zero, the EEW system is useful for that user and provides a cost savings. As an absolute performance metric, Q also depends on the data set used. Because each event is not uniformly sampled (i.e., station spacing, distances, and completeness is varied), the result is affected by the "missing" data. Nonetheless, comparisons made herein are valid as both algorithms are using the same flawed set of stations.

Using the simulated real-time alert classifications for each algorithm shown in Table S2, we calculate Q using two cost ratios ($r = 2$ and $r = 10$) at the three MMI thresholds for each algorithm. For a cost ratio of two, ElarmS alerts not useful at the MMI thresholds tested (negative Q values). For a cost ratio of 10, seismic alerts result in positive Q values for an MMI 3 threshold, but decrease with increasing MMI to the point of no longer being useful at MMI 5 (Table S2 and Figure 10). Geodetic solution metrics, on the other hand, are positive for all MMIs between 3 and 5 using a cost ratio of 10 and are positive for a ratio of two only at an MMI threshold of 3. The highest Q values are found for the geodetic algorithms at MMI 3 (~0.7),
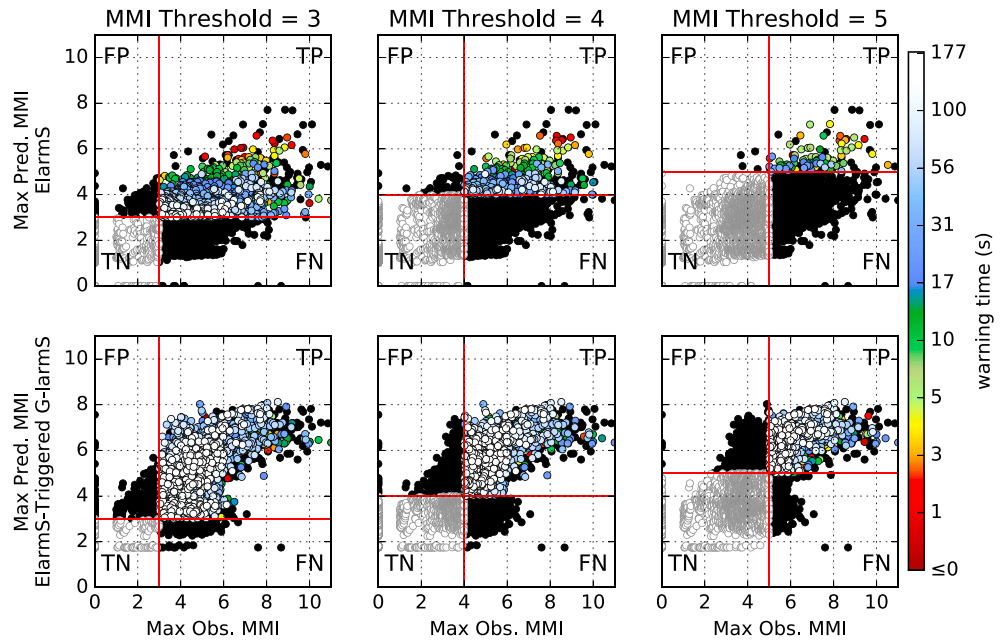
**Figure 8.** Real-time classification plots for thresholds of MMI 3 (left column), MMI 4 (middle column), and MMI 5 (right column) for all 5,151 individual station records for 31 earthquakes total ("Hayward4Hz" synthetic excluded, see text for discussion). Top row shows ElarmS results and bottom row shows ElarnS-Triggered G-larmS results. Data are colored by warning time based on the specific thresholds used (red lines). Quadrants are labeled as TP, TN, FP, or FN. Note that the Max. predicted MMI is the maximum predicted after the final solution, not the maximum predicted MMI before the alerting threshold is exceeded by the station. TP = true positive; TN = true negative; FP = false positive; FN = false negative; MMI = Modified Mercalli Intensity.
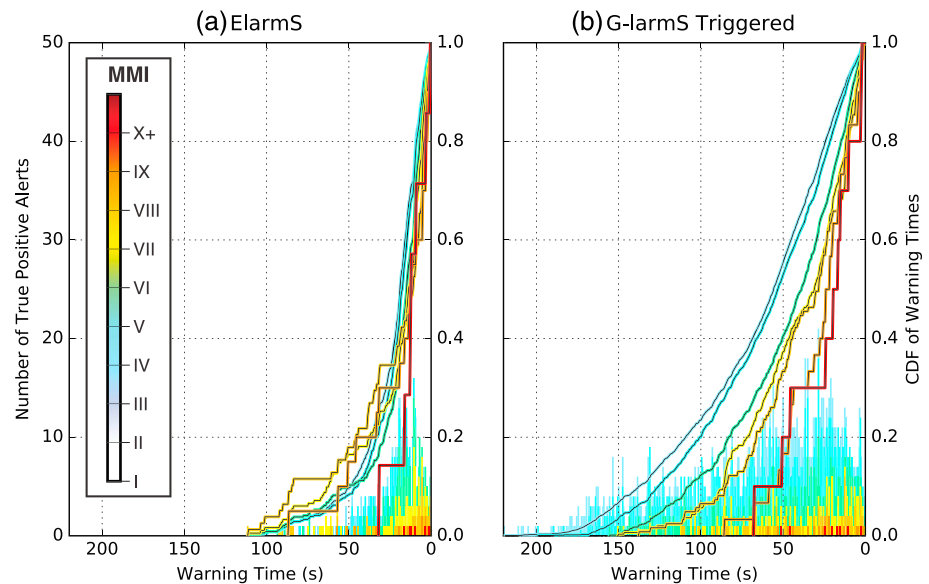


**Figure 9.** Empirical CDFs for all warning times at stations with MMI greater than or equal to each threshold (line colors, right *y* axes) and histograms of warning times binned and colored by observed maximum MMI (left *y* axes) for (a) ElarmS and (b) G-larmS results using a warning threshold of MMI 4. CDF = cumulative distribution function; MMI = Modified Mercalli Intensity.
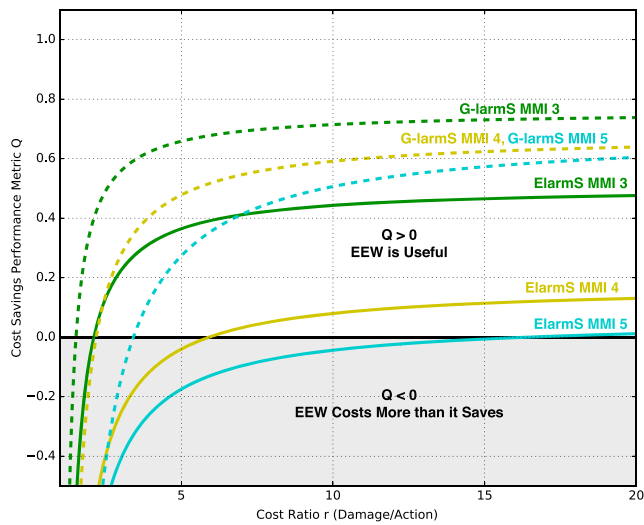
**Figure 10.** Cost Savings Performance Metric Q as a function of cost ratio *r* for ElarmS (solid) and ElarmS-Triggered G-larmS (dashed) real-time alert classification results using thresholds of MMI 3 (green), MMI 4 (yellow), and MMI 5 (cyan). MMI = Modified Mercalli Intensity; EEW = earthquake early warning.

suggesting that G-larmS will be more successful at alerting for larger shaking intensities accompanying large events when using a lower intensity alert threshold.

Additionally, using a large cost ratio (e.g., 10 or greater) representative of low-cost actions such as personal response (getting under a desk) or mass industrial response (stopping trains and elevators) shows that geodetic algorithm solutions are always useful, with a minimum Q value of 0.51 using a threshold of MMI 5. These numbers also show that geodetic solutions are much less affected by increasing the MMI threshold than ElarmS is for large events. We did not include smaller magnitude earthquakes (M < 6) in our analysis, therefore these statistics may only apply to higher magnitude events with large GMs.

Figure 10 shows the performance metric results for a large range of cost ratios for the same MMI thresholds. Most striking is that the coupled seismic-geodetic approach generally results in higher cost savings at smaller cost ratios for similar MMI thresholds. This is particularly important as *r* approaches 1, emphasizing that geodetic analyses provide significant value to users with small false alert tolerances.

Future studies should include a Q analysis of the full range of magnitudes to make proper recommendations about the use of the ShakeAlert system. Of course, much work is also needed in order to quantify the actual values of the cost ratio, *r*, for a variety of users; for the present, we can only speculate as to what these might be. However, using the framework of equation (1) it is clear that geodetic algorithms provide substantial added value to EEW for M > 7 earthquakes.

Geodetic finite-fault algorithms provide more accurate GM estimates (Figure 9) and they also provide better user-alert timeliness (Figure 7). If we consider the true alerts, the geodetic system provides improved WTs and accuracy under all thresholds tested against for a larger portion of sites than ElarmS does (Table S2 and Figure 7). Geodetic WTs can also be very long (>30 s), even for high MMI thresholds (Figures 5, 7, and 9).

Not only are the median WTs longer when incorporating geodetic finite-fault solutions, but, perhaps more importantly, for some of the stations with the largest GM, ElarmS never reaches the MMI threshold (e.g., Figure 5f). Thus, ElarmS never issues an alert for some of the strongest ground shaking. This behavior results in more missed alerts or FNs than the geodetic algorithm and a worse overall performance for the seismic system, especially for users sensitive to FNs.

Because they rely on the first few seconds of the *P* wave, initial alerts from seismic EEW algorithms are faster than those from geodetic algorithms (Figure 4). The higher noise levels in real-time geodetic time series (1–3 cm) preclude detection of the early onset signals and increase the delay until measured GM exceeds the noise. However, seismic *P* wave methods have important limitations. As shown in the previous section, they underestimate the magnitudes of very large earthquakes, and as a result of this and the point source assumption, are not as useful for providing accurate GM alerts to users for these events. As illustrated during the Mw9.0 Tohoku-oki earthquake, using seismic algorithms alone would mean that the region of predicted strong shaking is significantly smaller than the actual area of strong shaking and many users would not receive an alert. In an effort to ameliorate this, Hoshiba and Aoki (2015) proposed the PLUM method which uses a very dense network of strong motion sensors and uses present observations of the seismic wavefield to forecast its likely intensity some time into the future. This algorithm, now operational in Japan, improves EEW performance during large events but produces only short WTs and is limited to very dense networks. Similarly, FinDer (Böse et al., 2018) attempts to use the spatial patterns of ground acceleration to produce finite-fault source estimates, but performs best within dense networks.

Thus, there is a need for faster unsaturated magnitude calculations that can be translated to GM estimates. GNSS fills this niche: though slower to create an initial alert, the geodetic EEW algorithms provide significantly more accurate magnitude estimation, slip distribution estimation, and, in turn, more accurate GM prediction for large earthquakes. A common concern is that GNSS source estimates are too slow to be

useful in EEW, our results show that while this is true for moderate magnitude events around ~M6.5, for larger earthquakes, GNSS provides a substantial and very valuable improvement in the timeliness and reliability of the alerts. Geodetic algorithms can be used to correctly warn a significantly larger portion of the population than seismic algorithms alone, and to provide substantial WTs for users in areas that will experience strong shaking. Our results show that the current ShakeAlert point-source system alone will not be sufficient to forecast the strongest shaking due to the largest events. Instead, combining seismic and geodetic approaches can capture the earliest shaking with shorter WTs using the seismic data and predict more distant shaking with longer WTs from the geodetic data. We recommend a combined system as well as one that continually updates to account for the growth of large earthquakes. Large earthquakes are complex geophysical phenomena, and the best outcome for EEW is obtained when they are measured by geophysical instrumentation with complementary strengths. Our results conclusively show that GNSS data add substantial value to seismic point-source EEW systems.

Finally, we note that we have only produced testing results for one candidate geodetic finite-fault algorithm. As discussed in section 2, there are other candidate algorithms and we hope that they can be tested in a similar fashion with the data we make available so that we can make objective comparisons between the proposed algorithms and determine which features of the respective methods produce the most reliable GM estimates with the longest WTs. Future work should also include a comparison of geodetic algorithms to more advanced seismic EEW methods such as PLUM (Hoshiba & Aoki, 2015) and FinDer (Böse et al., 2012). Much of the improvement shown here is likely more a result of the fault-finiteness than of the more accurate moment magnitudes, and more testing is required to verify this.

## 6. Conclusions

Here, we quantified the timeliness and accuracy of seismic and geodetic magnitude and GM EEW alerts by testing a suite of large (M > 6) earthquakes worldwide. ElarmS magnitude errors indicate magnitude saturation for large events and are $-1.0 \pm 1.0$ and $-0.50 \pm 0.83$ units for the first and final alerts, respectively. ElarmS-triggered G-larmS magnitude errors are $-0.62 \pm 0.86$ and $-0.14 \pm 0.65$ units at the first and final update, respectively. We calculated shaking intensity time series for each station for each event using the simulated real-time solutions. Applying an MMI threshold approach to accurately characterize WTs on a per station basis, we classified TP, TN, FP, and FN alerts for each event. Using a threshold of MMI 4, ElarmS produced only 10.4% TP alerts with a median WT of $18.9 \pm 19.7$ s, while ElarmS-triggered G-larmS solutions result in 39.9% TP alerts with a longer median WT of $55.8 \pm 46.3$ s. The number of missed alerts (FN) using thresholds of MMI 3 and 4 is reduced significantly with the seismically triggered geodetic EEW system. Perfectly triggered G-larmS results provided similar statistics as the triggered system, with slightly longer WTs and more accurate final magnitudes (see supporting information for detailed comparison).

Analysis of the cost savings performance metric Q showed that the geodetic solutions provide a higher cost savings value to users with a variety of cost ratios $r$ when compared to the seismic-only point-source system, particularly for less false alert tolerant users. It also suggests that both systems will be more successful at alerting for larger shaking intensities accompanying large events when using a lower intensity alert threshold. These results demonstrate the added value of a geodetic EEW system, quantifying improvements in magnitude accuracy, GM accuracy, and alert timeliness in GM space. Permanent, static surface displacements are an essential part of earthquake observation and must be incorporated, having been measured with fidelity via GNSS, into EEW systems to ensure their success for the largest, most damaging earthquakes.

## References

Aagaard, B. T., Minson, S. E., Baltay, A. S., Cochran, E. S., & Hanks, T. C. (2018). Seismology of the Americas Meeting. *Seismological Research Letters*, *89*(2B), 717–966. https://doi.org/10.1785/0220180082

Allen, R. M. (2006). Probabilistic warning times for earthquake ground shaking in the San Francisco Bay Area. *Seismological Research Letters*, *77*(3), 374–379. https://doi.org/10.1785/gssrl.77.3.374

Allen, R. M., Brown, H., Hellweg, M., Khainovski, O., Lombard, P., & Neuhauser, D. (2009). Real-time earthquake detection and hazard assessment by ElarmS across California. *Geophysical Research Letters*, *36*, L00B08. https://doi.org/10.1029/2008GL036766

Allen, R. M., Gasparini, P., Kamigaichi, O., & Bose, M. (2009). The status of earthquake early warning around the world: An introductory overview. *Seismological Research Letters*, *80*(5), 682–693. https://doi.org/10.1785/gssrl.80.5.682

Allen, R. M., & Ziv, A. (2011). Application of real-time GPS to earthquake early warning. *Geophysical Research Letters*, *38*, L16310. https://doi.org/10.1029/2011GL047947

Blewitt, G., Kreemer, C., Hammond, W. C., Plag, H., Stein, S., & Okal, E. (2006). Rapid determination of earthquake magnitude using GNSS for tsunami warning systems. *Geophysical Research Letters*, *33*, L11309. https://doi.org/10.1029/2006GL026145

Bock, Y., & Melgar, D. (2016). Physical applications of GPS geodesy: A review. *Reports on Progress in Physics*, *79*(10), 106801. https://doi.org/10.1088/0034-4885/79/10/106801

Boore, D. M., & Bommer, J. J. (2005). Processing of strong-motion accelerograms: Needs, options and consequences. *Soil Dynamics and Earthquake Engineering*, *25*(2), 93–115. https://doi.org/10.1016/j.soildyn.2004.10.007

Böse, M., Felizardo, C., & Heaton, T. H. (2015). Finite-fault rupture detector (FinDer): Going real-time in Californian ShakeAlert warning system. *Seismological Research Letters*, *86*(6), 1692–1704. https://doi.org/10.1785/0220150154

Böse, M., Heaton, T., & Hauksson, E. (2012). Rapid estimation of earthquake source and ground-motion parameters for earthquake early warning using data from a single three-component broadband or strong-motion sensor. *Bulletin of the Seismological Society of America*, *102*(2), 738–750. https://doi.org/10.1785/0120110152

Böse, M., Smith, D. E., Felizardo, C., Meier, M.-A., Heaton, T. H., & Clinton, J. F. (2018). FinDer v.2: Improved real-time ground-motion predictions for M2-M9 with seismic finite-source characterization. *Geophysical Journal International*, *212*(1), 725–742. https://doi.org/10.1093/gji/ggx430

Chung, A. I., Henson, I. H., Allen, R. M., Hellweg, M., & Neuhauser, D. S. (2017), ElarmS earthquake early warning system: 2017 Performance and New ElarmS Version 3.0 (E3). In AGU Fall Meeting Abstracts.

Cochran, E. S., Kohler, M. D., Given, D. D., Guiwits, S., Andrews, J., Meier, M. A., et al. (2017). Earthquake early warning ShakeAlert system: Testing and certification platform. *Seismological Research Letters*, *89*(1), 108–117.

Colombelli, S., Allen, R. M., & Zollo, A. (2013). Application of real-time GNSS to earthquake early warning in subduction and strike-slip environments. *Journal of Geophysical Research: Solid Earth*, *118*, 3448–3461. https://doi.org/10.1002/jgrb.50242

Crowell, B. W., Schmidt, D. A., Bodin, P., Vidale, J. E., Gomberg, J., Hartog, J. R., et al. (2016). Demonstration of the Cascadia G-FAST geodetic earthquake early warning system for the Nisqually, Washington, earthquake. *Seismological Research Letters*, *87*(4), 930–943. https://doi.org/10.1785/0220150255

Field, E. H., Biasi, G. P., Bird, P., Dawson, T. E., Felzer, K. R., Jackson, D. D., et al. (2014). Uniform California earthquake rupture forecast, version 3 (UCERF3)—The time-independent model: U.S. Geological Survey Open-File Report 2013–1165, 97 p., California Geological Survey Special Report 228, and Southern California Earthquake Center Publication 1792. Retrieved from http://pubs.usgs.gov/of/2013/1165/

Galetzka, J., Melgar, D., Genrich, J. F., Geng, J., Owen, S., Lindsey, E. O., et al. (2015). Slip pulse and resonance of the Kathmandu basin during the 2015 Gorkha earthquake, Nepal. *Science*, *349*(6252), 1091–1095.

Geng, J., Bock, Y., Melgar, D., Crowell, B. W., & Haase, J. S. (2013). A new seismogeodetic approach applied to GPS and accelerometer observations of the 2012 Brawley seismic swarm: Implications for earthquake early warning. *Geochemistry, Geophysics, Geosystems*, *14*, 2124–2142. https://doi.org/10.1002/ggge.20144

Grapenthin, R., & Freymueller, J. T. (2011). The dynamics of a seismic wave field: Animation and analysis of kinematic GPS data recorded during the 2011 Tohoku-oki earthquake, Japan. *Geophysical Research Letters*, *38*, L18308. https://doi.org/10.1029/2011GL048405

Grapenthin, R., Johanson, I., & Allen, R. M. (2014a). Operational real-time GNSS-enhanced earthquake early warning. *Journal of Geophysical Research: Solid Earth*, *119*, 7944–7965. https://doi.org/10.1002/2014JB011400

Grapenthin, R., Johanson, I., & Allen, R. M. (2014b). The 2014 Mw6.0 Napa earthquake, California: Observations from real-time GPS-enhanced earthquake early warning. *Geophysical Research Letters*, *41*, 8269–8276. https://doi.org/10.1002/2014GL061923

Grapenthin, R., West, M., & Freymueller, J. (2017). The utility of GNSS for earthquake early warning in regions with sparse seismic networks. *Bulletin of the Seismological Society of America*, *107*(4), 1883–1890. https://doi.org/10.1785/0120160317

Hayes, G. P., Wald, D. J., & Johnson, R. L. (2012). Slab1.0: A three-dimensional model of global subduction zone geometries. *Journal of Geophysical Research*, *117*, B01302. https://doi.org/10.1029/2011JB008524

Hoshiba, M., & Aoki, S. (2015). Numerical shake prediction for earthquake early warning: Data assimilation, real-time shake mapping, and simulation of wave propagation. *Bulletin of the Seismological Society of America*, *105*(3), 1324–1338. https://doi.org/10.1785/0120140280

Hoshiba, M., Iwakiri, K., Hayashimoto, N., Shimoyama, T., Hirano, K., Yamada, Y., et al. (2011). Outline of the 2011 off the Pacific coast of Tohoku earthquake ($M_w$9.0)—Earthquake early warning and observed seismic intensity. *Earth, Planets and Space*, *63*(7), 547–551. https://doi.org/10.5047/eps.2011.05.031

Hoshiba, M., & Ozaki, T. (2014). Earthquake early warning and tsunami warning of the Japan meteorological agency, and their performance in the 2011 off the Pacific coast of Tohoku earthquake (Mw 9.0). In *Early Warning for Geological Disasters* (pp. 1–28). Berlin: Springer.

Kawamoto, S., Ohta, Y., Hiyama, Y., Todoriki, M., Nishimura, T., Furuya, T., et al. (2017). REGARD: A new GNSS-based real-time finite fault modeling system for GEONET. *Journal of Geophysical Research: Solid Earth*, *122*, 1324–1349. https://doi.org/10.1002/2016JB013485

Kuyuk, H. S., Allen, R. M., Brown, H., Hellweg, M., Henson, I., & Neuhauser, D. (2013). Designing a network-based earthquake early warning algorithm for California: ElarmS-2. *Bulletin of the Seismological Society of America*, *104*(1), 162–173.

Meier, M. A., Heaton, T., & Clinton, J. (2016). Evidence for universal earthquake rupture initiation behavior. *Geophysical Research Letters*, *43*, 7991–7996. https://doi.org/10.1002/2016GL070081

Meier, M.-A. (2017). How "good" are real-time ground motion predictions from earthquake early warning systems? *Journal of Geophysical Research: Solid Earth*, *122*, 5561–5577. https://doi.org/10.1002/2017JB014025

Melgar, D., Bock, Y., Sanchez, D., & Crowell, B. W. (2013). On robust and reliable automated baseline corrections for strong motion seismology. *Journal of Geophysical Research: Solid Earth*, *118*, 1177–1187. https://doi.org/10.1002/jgrb.50135

Melgar, D., LeVeque, R. J., Dreger, D. S., & Allen, R. M. (2016). Kinematic rupture scenarios and synthetic displacement data: An example application to the Cascadia subduction zone. *Journal of Geophysical Research: Solid Earth*, *121*, 6658–6674. https://doi.org/10.1002/2016JB013314

Melgar, D. & Ruhl, C. J. (2018). High-rate GNSS displacement waveforms for large earthquakes version 2.0 (Version 2.0) [Data set]. Zenodo. http://doi.org/10.5281/zenodo.1434374

Minson, S. E., Baltay, A. S., Cochran, E. S., Hanks, T. C., Page, M. T., McBride, S. K., et al. (2019). The limits of earthquake early warning accuracy and best alerting strategy. *Scientific Reports*, *9*(1), 2478. https://doi.org/10.1038/s41598-019-39384-y

Minson, S. E., Baltay, A. S., Meier, M.-A., Hanks, T. C., & Cochran, E. S. (2017). The effect of ground-motion variability on the accuracy of earthquake early warning. *Seismological Research Letters*, *88*.

Minson, S. E., Murray, J. R., Langbein, J. O., & Gomberg, J. S. (2014). Real-time inversions for finite fault slip models and rupture geometry based on high-rate GPS data. *Journal of Geophysical Research: Solid Earth*, *119*, 3201–3231. https://doi.org/10.1002/2013JB010622

Murray, J. R., Crowell, B. W., Grapenthin, R., Hodgekinson, K., Langbein, J. O., Melbourne, T., et al. (2018). Development of a geodetic component for the U. S. west coast earthquake early warning system. *Seismological Research Letters*, *89*(6), 2322–2336. https://doi.org/10.1785/0220180162

Rodgers, A. J., Pitarka, A., Petersson, N. A., Sjögreen, B., & McCallen, D. B. (2018). Broadband (0–4 Hz) ground motions for a magnitude 7.0 Hayward fault earthquake with three-dimensional structure and topography. *Geophysical Research Letters*, *45*, 739–747. https://doi.org/10.1002/2017GL076505

Ruhl, C. J., Melgar, D., Grapenthin, R., & Allen, R. M. (2017). The value of real-time GNSS to earthquake early warning. *Geophysical Research Letters*, *44*, 8311–8319. https://doi.org/10.1002/2017GL074502

Wald, D. J., & Allen, T. I. (2007). Topographic slope as a proxy for seismic site conditions and amplification. *Bulletin of the Seismological Society of America*, *97*(5), 1379–1395. https://doi.org/10.1785/0120060267

Wells, D. L., & Coppersmith, K. J. (1994). New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement. *Bulletin of the seismological Society of America*, *84*, 974–1002.

Worden, C. B., Gerstenberger, M. C., Rhoades, D. A., & Wald, D. J. (2012). Probabilistic relationships between ground-motion parameters and modified Mercalli intensity in California. *Bulletin of the Seismological Society of America*, *102*(1), 204–221. https://doi.org/10.1785/0120110156