AGU PUBLICATIONS

JGR

Journal of Geophysical Research: Solid Earth

RESEARCH ARTICLE

10.1002/2014JB011400

**Key Points:**
- End-to-end system for real-time GPS EEW provides finite fault slip models
- Sensitivity tests scrutinize a priori assumptions that linearize the inversion
- Exhaustive catalog of test events necessary to validate operational RT-GPS EEW

**Supporting Information:**
- Readme
- Movie S1
- Movie S2
- Movie S3
- Movie S4

**Correspondence to:**
R. Grapenthin,
rg@nmt.edu

**Citation:**
Grapenthin, R., I. A. Johanson, and R. M. Allen (2014), Operational real-time GPS-enhanced earthquake early warning, *J. Geophys. Res. Solid Earth*, *119*, doi:10.1002/2014JB011400.

Received 20 JUN 2014
Accepted 19 SEP 2014
Accepted article online 7 OCT 2014

# Operational real-time GPS-enhanced earthquake early warning

**R. Grapenthin[1,2], I. A. Johanson[1], and R. M. Allen[1]**

[1]Berkeley Seismological Laboratory, University of California, Berkeley, California, USA, [2]Department of Earth and Environmental Science, New Mexico Institute of Mining and Technology, Socorro, New Mexico, USA

**Abstract** Moment magnitudes for large earthquakes ($M_w \geq 7.0$) derived in real time from near-field seismic data can be underestimated due to instrument limitations, ground tilting, and saturation of frequency/amplitude-magnitude relationships. Real-time high-rate GPS resolves the buildup of static surface displacements with the *S* wave arrival (assuming nonsupershear rupture), thus enabling the estimation of slip on a finite fault and the event's geodetic moment. Recently, a range of high-rate GPS strategies have been demonstrated on off-line data. Here we present the first operational system for real-time GPS-enhanced earthquake early warning as implemented at the Berkeley Seismological Laboratory (BSL) and currently analyzing real-time data for Northern California. The BSL generates real-time position estimates operationally using data from 62 GPS stations in Northern California. A fully triangulated network defines 170+ station pairs processed with the software trackRT. The BSL uses G-larmS, the Geodetic Alarm System, to analyze these positioning time series and determine static offsets and preevent quality parameters. G-larmS derives and broadcasts finite fault and magnitude information through least-squares inversion of the static offsets for slip based on a priori fault orientation and location information. This system tightly integrates seismic alarm systems (CISN-ShakeAlert, ElarmS-2) as it uses their *P* wave detections to trigger its processing; quality control runs continuously. We use a synthetic Hayward Fault earthquake scenario on real-time streams to demonstrate recovery of slip and magnitude. Reanalysis of the $M_w$7.2 El Mayor-Cucapah earthquake tests the impact of dynamic motions on offset estimation. Using these test cases, we explore sensitivities to disturbances of a priori constraints (origin time, location, and fault strike/dip).

## 1. Introduction

Earthquake early warning (EEW) exploits the differential propagation velocities of seismic body waves. By detecting the less destructive *P* wave (pressure wave) rapidly in real time, a warning is issued before the damaging *S* Wave (shear wave) arrives [*Allen and Kanamori*, 2003]. Location and moment magnitude estimates derived from *P* wave records of a few seismometers are key inputs into models that, for example, predict ground shaking at specific locations (e.g., USGS ShakeMap [*Wald et al.*, 2006]) or forecast tsunami threats [e.g., *Kerr*, 2005; *Blewitt et al.*, 2006]. However, moment magnitudes for large earthquakes ($M_w \geq 7.0$), inferred in real time from near-field seismic data, are often significantly underestimated [e.g., *Brown et al.*, 2009; *Hoshiba et al.*, 2011] due to instrument limitations, ground tilting [e.g., *Bock et al.*, 2011], and saturation of frequency/amplitude-magnitude relationships [e.g., *Brown et al.*, 2009]. Real-time high-rate GPS, on the other hand, resolves static surface displacements, which begin to build up with the *S* wave arrival (assuming nonsupershear rupture), thus enabling the estimation of slip on a finite fault and the event's geodetic moment.

Within the last decade the benefits of high-rate GPS data (sample rate ≥1 Hz) in earthquake research have been widely recognized. *Larson et al.* [2003] first demonstrated the ability of GPS to resolve surface waves. *Grapenthin and Freymueller* [2011] show *S* wave and surface wave radiation in 2-D during the 2011 Tohoku-oki earthquake. *Yue and Lay* [2011] derive the slip history of the Tohoku-oki event from time-dependent inversions of 1 Hz GPS data and *Avallone et al.* [2011] use static and dynamic motions recorded by 10 Hz GPS to reconstruct the kinematic rupture history of the 2009 L'Aquila earthquake.

Recently, a range of high-rate GPS strategies for EEW and rapid response have been demonstrated on off-line data. Using high-rate GPS waveforms, *Crowell et al.* [2009] demonstrate a purely GPS-based EEW system. Based on analysis of the 2010 El Mayor-Cucapah event, *Allen and Ziv* [2011] conclude that the most

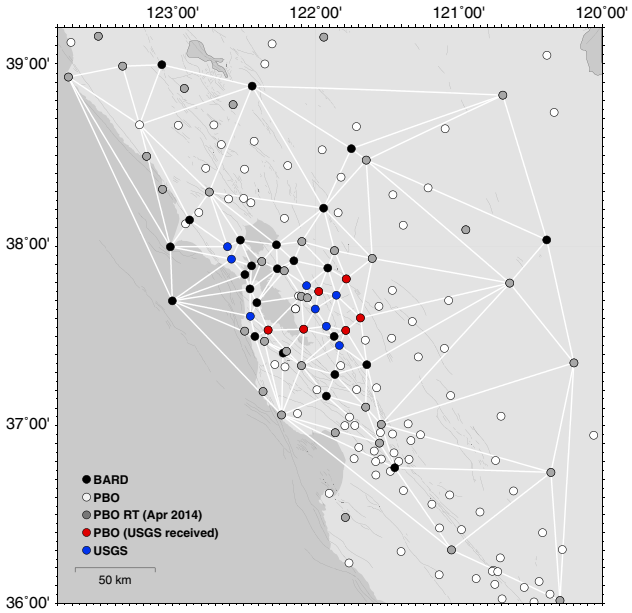


**Figure 1.** Map of the greater San Francisco Bay Area with high-rate GPS stations (dot color represents operator). Lines connecting the stations indicate real-time station pairs (172) we use in the processing. White dots are non real-time PBO stations (April 2014). PBO RT site list has been obtained through their NTRIP caster table most recently in April 2014. Some recent RT PBO sites between Point Reyes and Cape Mendocino still have to be incorporated into the processing (P182, P183, P203, and P206). P188 (white dot connected to baselines in NW) is still indicated to be streaming real-time data.

effective strategy would be the integration of GPS-derived information (there magnitude) and seismically derived information (detection and location of epicenter). *Melgar et al.* [2012] demonstrate rapid centroid moment tensor determination from high-rate GPS data at time scales similar to moment tensor derivation from long-period (> 10 s) seismic waves [*Guilhem and Dreger*, 2011]. *Wright et al.* [2012] reanalyze the 2011 Tohoku-oki earthquake data in real-time mode and find robust magnitude estimates derived from static slip inversions that stabilize before the rupture completes. Monitoring the difference in ratio of short-term average (STA) to long-term average (LTA) displacements, *Ohta et al.* [2012] devise an algorithm that automatically identifies permanent displacements and triggers fault model and magnitude determination for tsunami hazard forecasts. Using off-line data for a few recent large earthquakes ($M_w$9.0 Tohoku-oki, 2011; $M_w$8.3 Tokachi-oki, 2003; and $M_w$7.2 El Mayor-Cucapah, 2010), *Colombelli et al.* [2013] show that magnitude and first-order slip distribution on a predefined fault can be derived reliably using rapid coseismic static offset estimates in a least-squares finite fault inversion. Using similar off-line data, *Minson et al.* [2014] apply an analytical Bayesian approach in the inversion step.

While GPS basically has unlimited dynamic range to record coseismic or static offsets induced by earthquakes, it lacks sensitivity to small movements that seismometers were developed for. Hence, using current technologies, GPS alone is unable to resolve *P* wave displacements in real time (which induce mainly vertical motions). *Bock et al.* [2011] solve this through a combination of data from GPS and colocated broadband seismometers or simple accelerometers attached to GPS monuments. However, the main benefit gained from integrating GPS into earthquake early warning remains the precise estimation of coseismic/static offsets.

Real-time coseismic offsets are critical for the real-time magnitude estimation of large earthquakes ($M_w > 7.0$) as the current empirical *P* wave frequency- (or amplitude-) magnitude relationships for seismic records tend to saturate above this magnitude. This may be due to an undersampled population of large earthquakes or a fundamental difference in the rupture mechanism (e.g., "double slip" during 2011 Tohoku-oki [*Bilek et al.*, 2011; *Meng et al.*, 2011; *Grapenthin*, 2012]) occurring for large events, or is lost in the records of small events. Static offsets can be readily inverted for slip along a finite fault model which, in turn, gives a geodetic magnitude. A range of studies demonstrate the possibility to use high-rate GPS this way [e.g., *Wright et al.*, 2012; *Colombelli et al.*, 2013]; however, all of them operate in off-line mode on replayed data.

Here we describe an operational system for real-time GPS-enhanced EEW as implemented at the Berkeley Seismological Laboratory (BSL) and currently running on data for Northern California. We roughly implement the approach of *Colombelli et al.* [2013] who describe a distributed slip inversion on a finite fault (configured from fault catalog information) based on static displacements extracted from high-rate GPS data. Similarly to *Colombelli et al.* [2013], we do not implement a full nonlinear inversion, which is computationally expensive. Instead of relying on fault catalogs to linearize the inversion step, we derive fault

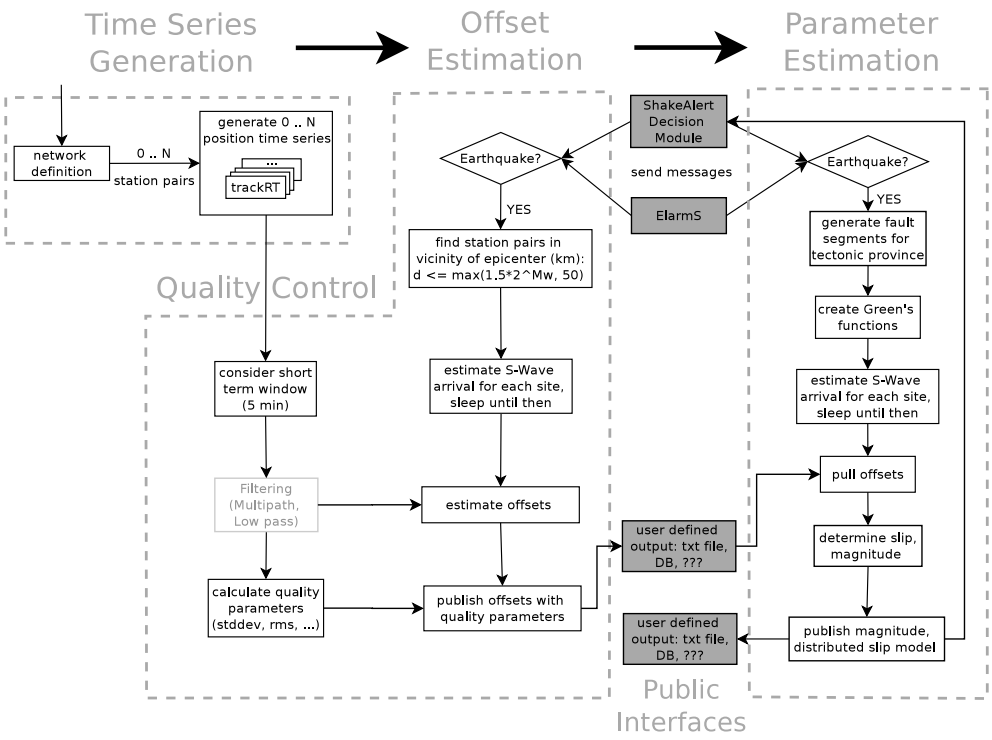


**Figure 2.** Flowchart for real-time GPS processing at the Berkeley Seismological Laboratory. Dashed lines enclose physically separate components. Grey boxes mark the interface level. Light grey filtering box indicates planned but not yet implemented feature.

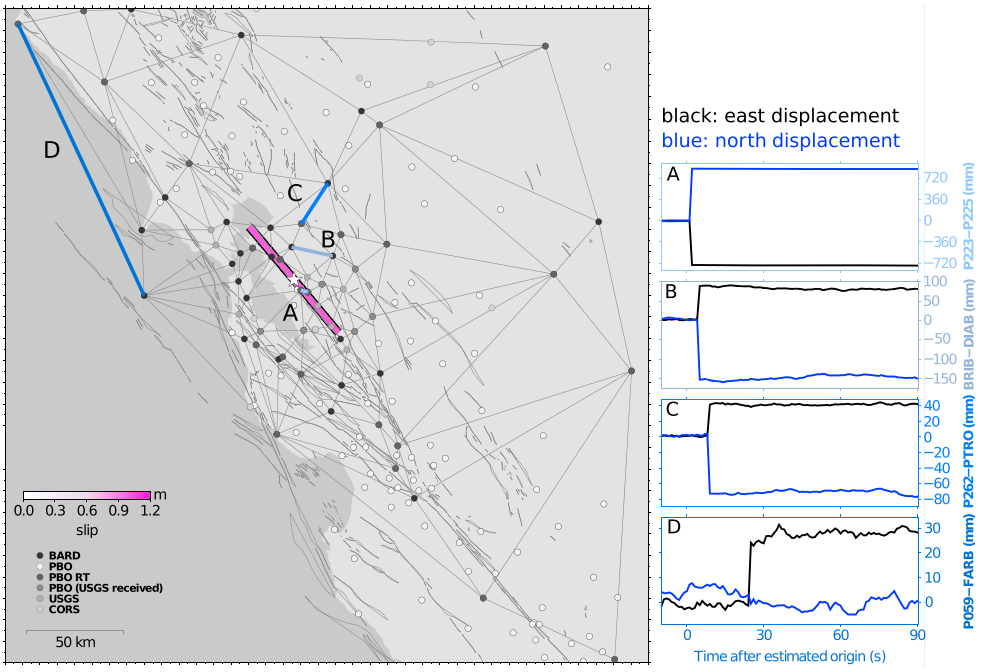


**Figure 3.** Network used for the synthetic Hayward test. A surface projection of the 90° dipping forward model fault is shown as pink patch (1.25 m of slip). The white star indicates the synthetic epicenter. (a–d) Baselines shown in shades of blue and labeled A-D mark locations of the time series in the right panels. The time series (blue: north displacement, black: east displacement) show position change along the baselines with respect to the origin time of the synthetic event. Signal-to-noise ratio in Figure 3d is low enough to get a good intuition about the noise in the real-time GPS data. The step in all time series is the simulated offset; the move out from near to far field is simulated based on the expected effective *S* wave travel time. Upon expected *S* wave arrival time, the synthetic offsets are added to the respective base lines.

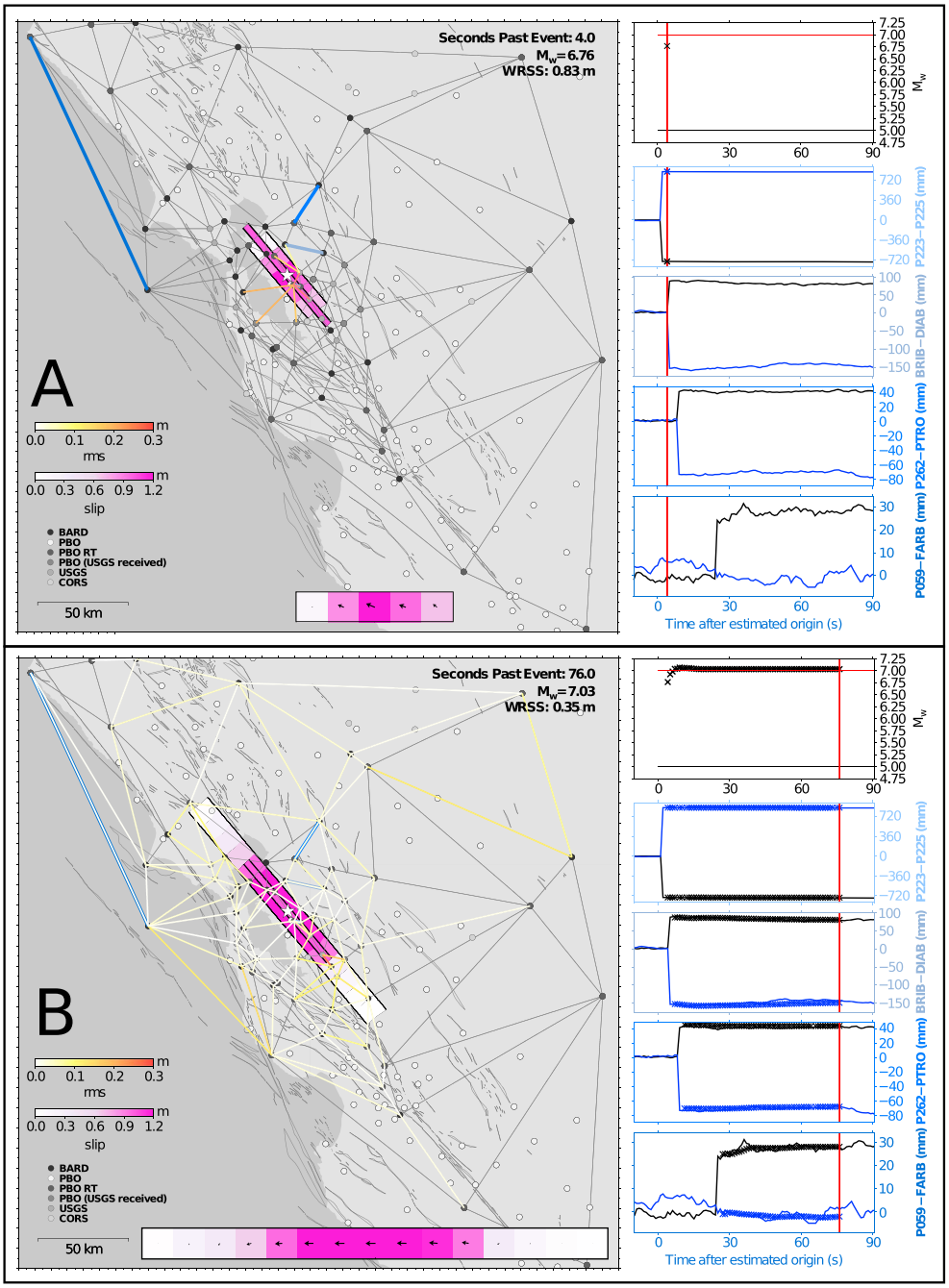


**Figure 4.** (a) Initial recovered slip model. Similar to Figure 3, but color scale of baselines on the map indicates RMS misfit of individual baselines, here at 4 s after event onset (time shown by red vertical line in time series panels). Baselines that show no offset yet or provide no data remain gray. Total WRSS misfit is given in the upper right corner of the map. Wide fault patches inside the network (projection of vertical fault) show distributed slip model inferred from the available offsets at 4 s (original forward model plotted on top for comparison). The slip model in the inset at the bottom of the map shows a north-south (left to right) cross section through the fault. One vector for each fault patch in the cross section shows direction and relative magnitude of slip (normalized to maximum slip value 90 s postevent onset); patch colors show slip magnitude. $M_w$ plot shows time series of inferred magnitude (black crosses) versus forward model magnitude (red line); the black line represents the input magnitude from the alarm message. Baseline time series show recovered offsets in east and north direction as black and blue crosses, respectively. These offset time series are used by the Parameter Estimator to infer slip model. (b) Final recovered slip model; shows recovered offsets and slip model at 76 s after event onset.

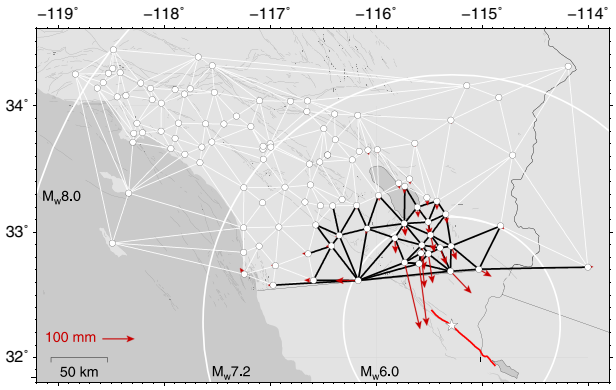


**Figure 5.** Set of baselines (lines connecting white dots) used for the reanalysis of the 2010 $M_w$7.2 El Mayor-Cucapah earthquake (white star is USGS epicenter, red line is surface rupture trace). Red arrows show absolute horizontal displacements (published by SOPAC). For the test of impact of dynamic shaking on G-larmS' performance, we estimated offsets and inverted these for slip for all baselines (see Figure 7). Black baselines form a subset used for the time-based reanalysis (see Figure 9). White circles mark G-larmS site selection threshold radii for events with a seismic alert magnitude of $M_w$ 6.0, 7.2, and 8.0, respectively.

geometry information from tectonic provinces (e.g., regions of "San Andreas Fault parallel" faults, etc.) and take location information from the early warning information provided through the CISN ShakeAlert [*Böse et al.*, 2014]. We test this bare-bones approach, optimized for speed, in an operational real-time environment and investigate its sensitivities to our assumptions.

Simulated permanent offsets along the Hayward fault help us validate the implemented algorithms and data flow. We demonstrate how a large earthquake in Northern California that includes dynamic motion due to shaking would be processed by replaying and analyzing data from the 2010 $M_w$7.2 El Mayor-Cucapah event in simulated real time. As our methodology is based on triggers and information from the seismic system and a priori assumptions about fault geometry, we evaluate our dependence on uncertainties in origin time, event location, and fault strike and dip for both the synthetic Hayward fault case and the replayed El-Mayor event. This explores how much the (parallelized) real-time parameter space exploration can be limited to fit available computer resources while producing physically sensible first-order rupture approximations. Furthermore, most currently published feasibility studies of EEW with real-time GPS use examples in which the rupture usually occurs at significant distance from the network (e.g., Tohoku-oki 2011) as opposed to under the network, which is the case in California. Comparing the sensitivities to location and fault geometry between synthetic ruptures on the Hayward fault (under the network) and El Mayor (outside the Southern California network) sheds light on the impact of observation distance on resolution of subtleties in slip distribution and fault geometry.

## 2. Operational Real-Time Monitoring Using G-larmS

We use the G-larmS infrastructure for real-time analysis of displacement time series. Details on functionality and implementation of G-larmS will be given in a companion paper that is currently being prepared by the authors. Here we provide a brief overview of some of its functionality and focus on specific algorithms we developed or adapted for use in the BSL production system.

The main tasks to include GPS into EEW discussed below are (1) generation of displacement time series, (2) quality control that flags bad data (excessive data gaps, noisy, outliers), (3) extraction of static offsets and/or waveform parameters from the displacement time series in case of an earthquake, and (4) inversion of these offsets for earthquake parameters such as moment-magnitude, fault geometry, slip distribution, etc. Figure 2 shows how these tasks translate into processing and data flow at the BSL.

### 2.1. Network Design and Time Series Generation

The BSL currently generates real-time position estimates for about 62 1 Hz GPS stations in the greater San Francisco Bay Area (Figure 1). The BSL operates 26 of these stations as the BARD network, seven are operated by the USGS, and the remaining 29 stations are part of the Plate Boundary Observatory (PBO) operated through UNAVCO, who currently make real-time data streams available through the public Internet. While more high-rate (1 Hz) stations exist in the area, especially in the critical Northern San Andreas Fault region just south of Cape Mendocino, some have not yet been upgraded to real-time streaming or included into our processing. Hence, Figure 1 should be considered a snap shot.

To create positioning time series from real-time data streams, we employ a fully triangulated network scheme (Figure 1). Neighboring station pairs (identified using GMT's Delauney Triangulation [*Wessel and*

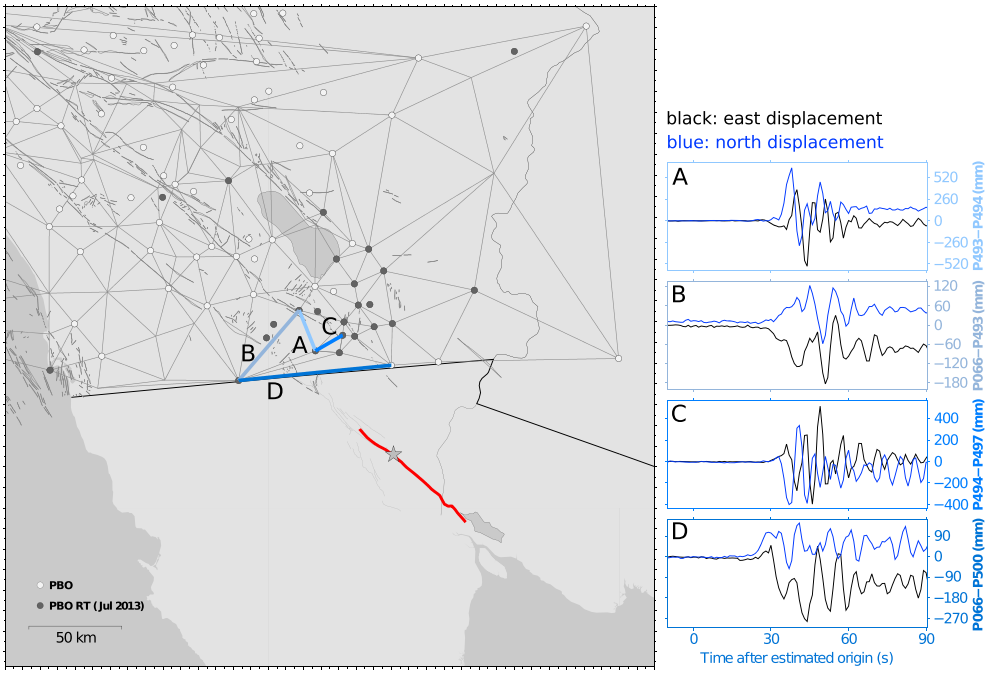


**Figure 6.** Similar to Figure 3, but for the El Mayor-Cucapah event showing baselines for Southern California. The time series show dynamic motion (shaking) due to the earthquake as resolved for the highlighted baselines at the sites.

*Smith*, 1995]) are processed independently by individual `trackRT` (part of GAMIT/GLOBK [*Herring et al.*, 2010]) instances (about 170 processes). In this double-differencing approach, one station is assumed stationary and the other is treated as moving with respect to this site. Hence, motion at the base station is mapped into the time series of the respective rover station. We use ultrarapid orbits (predicted, sp3 format) distributed by the International GNSS Service (IGS) [*Dow et al.*, 2009]. The Global Mapping Function (GMF) [*Boehm et al.*, 2006] is used to map hydrostatic and wet delays to satellite elevation angles. The Global Pressure and Temperature model (GPT) [*Boehm et al.*, 2007] provides a priori dry tropospheric delay estimates. It is important to mention that we do not include any kind of network adjustment to relate motion within the network to one or a few individual stations, which are assumed stable.

### 2.2. Quality Control

While we have implemented quality control, our current inversion strategy does not discard or downweight any data based on this information yet. However, the next generation of our processing will perform quality-based data selection. Therefore, we include a brief summary of this part of the processing (Figure 2, Quality Control column). Monitored quality measures include the standard deviation of each displacement time series component (east, north, and up), and the mean position solution misfit (RMS) and its standard deviation reported by the GPS processor over this time period. The standard deviation is a measure of stability, i.e., whether the position time series is mostly flat lined or experiences large swings, for example, due to multipath, atmospheric disturbances, incorrectly resolved ambiguities, etc. The average RMS is a similar, more crude, and harder to evaluate measure that indicates how well the modeled phases of the GPS processor match recorded phases coming from the GPS satellites, which is a measure of position solution quality. The formal uncertainties reported by GPS processors also indicate solution quality, but we propagate these with individual offset estimates (see following section). We track the data gaps of the position time series in percent of buffer length and plan to pass on the average connectivity of the gaps in terms of solution epochs.

### 2.3. Offset Estimation

While displacement time series quality is evaluated continuously by G-larmS, estimation of static offsets is triggered through the seismic system, CISN ShakeAlert [*Böse et al.*, 2014], which sends out alerts based on *P* wave detection and therefore is faster than GPS-based transient detection would be (Figure 2, Offset

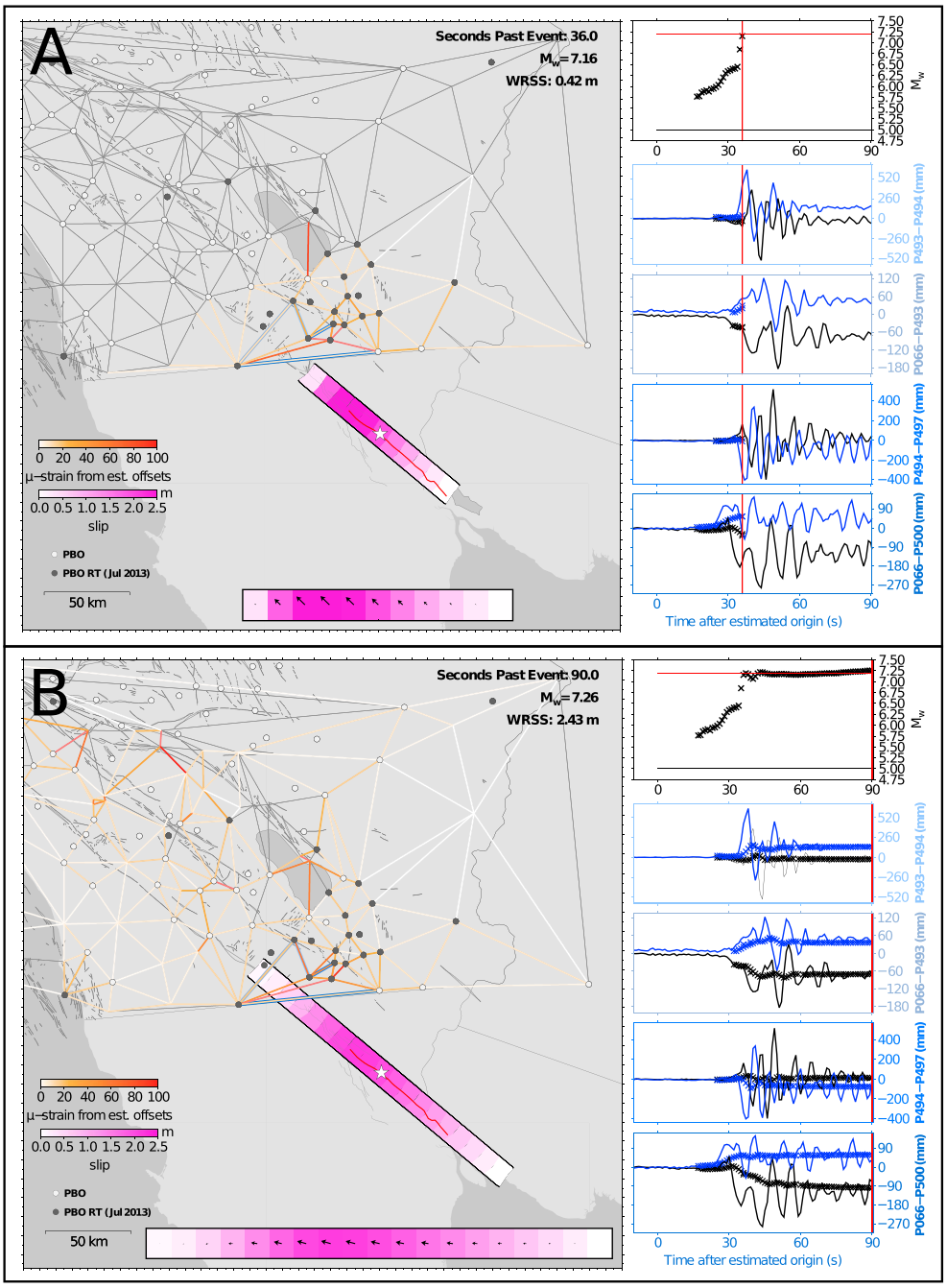


**Figure 7.** (a) Initial recovered slip model for El Mayor-Cucapah, 36 s after official event onset time. Similar to Figure 4, but color of baselines indicates microstrain along the respective baselines. Offset estimation (blue and black crosses in time series panels) begins too early due to strike slip rupture not starting until 15 s after event origin [*Wei et al.*, 2011]; hence the early and low initial magnitude estimates. This snap shot shows a $M_w$ = 7.16 solution 2 s after *Colombelli et al.* [2013] indicate static offset arrival and give their first solution at $M_w$ = 7.08. (b) Final recovered slip model and estimated offsets 90 s after event onset (see supporting information for full animation). Offset time series now clearly show the small impact of dynamic motion on offset estimation and hence magnitude estimation.

Estimation column). Once an alarm is received, G-larmS selects baselines within a generous radius, $r \leq \max(1.5 \times 2^{M_w}, 50)$, for which we expect static offsets (e.g., r($M_w$6.0) = 96 km and r($M_w$7.0) = 192 km). This depends on the magnitude broadcast by CISN ShakeAlert and is at least 50 km.

To estimate coseismic offsets for the east, north, and vertical components of each displacement time series, we implement the following algorithm:

1. Predict *S* wave arrival time using CISN ShakeAlert epicenter and apparent *S* wave velocity (v = 3.0 km/s) [e.g., *Grapenthin and Freymueller*, 2011];
2. Determine preevent position and uncertainties by averaging time series and formal uncertainties in the buffer up to the seismically derived event origin time (max. 5 min);
3. Wait until *S* wave is expected to arrive at closest (or farthest) site from epicenter (configurable, this paper uses arrival at closest site to minimize wait time, which can result in additional dynamic motion in the time series as *S* wave and surface waves pass the more distant site);
4. Average each new datum and its formal uncertainty after predicted *S* wave arrival to determine postevent position
5. Coseismic offset is the average postevent position minus average preevent position at each epoch (should approach stable value, which will be used as a termination criterion in a future implementation);
6. Offset uncertainty is the sum of preevent and postevent uncertainty.

The offset estimates and uncertainties are currently streamed to the file system, where processes working on magnitude and finite fault dimensions consume them.

The key element is the wait for the *S* wave arrival, which, certainly in the near field (e.g., $< 2°$ epicentral distance), can be associated with the arrival of coseismic displacements [e.g., *Grapenthin and Freymueller*, 2011; *Grapenthin*, 2012]. We do this to avoid underestimating coseismic offset estimates by including preevent data into a moving average estimate of the postevent position. The trade-off is that, if the *S* wave arrival is misestimated, we might still significantly overestimate or underestimate the coseismic offset for an individual baseline. However, we find that (a) the offset estimate stabilizes after about 1.5 wavelengths of dynamic motion [see also *Allen and Ziv*, 2011] and (b) that the magnitude estimation is robust and not much affected by overestimates of individual baselines (see sensitivity test in section 5.1).

### 2.4. Offset Inversion/Earthquake Parameter Estimation

The offsets estimated for individual events and the preevent time series quality measures are consumed by G-larmS' Parameter Estimator module, which implements least-squares offset inversion for distributed strike slip and dip slip along multiple fault segments.

Similarly to the Offset Estimator, CISN ShakeAlert events trigger the Parameter Estimator (Figure 2, Parameter Estimation column). Upon the receipt of an event message, G-larmS generates fault segments centered on the event epicenter. We follow *Colombelli et al.* [2013] in that the initial fault length is at least 3 times the length suggested by empirical relations [*Wells and Coppersmith*, 1994] based on the initial seismic magnitude (minimum number of fault segments is 5). This fault length estimation repeats at each epoch but uses the geodetic magnitude after this initialization. Deviating from *Colombelli et al.* [2013], rather than dividing the fault into seven segments, we fix the fault segment length to be 10 km long. The width of each segment reaches from the surface to 12 km depth (and hence depends on dip); the predominant seismogenic depth in California [*Sibson*, 1982; *Bürgmann et al.*, 2000; *Hauksson*, 2011]. Fixing these values allows us to grow the fault throughout the analysis and calculate Green's functions only for new segments. Strike and dip are predefined for fault regimes, i.e., regions of predominant strike and dip (such as San Andreas fault parallel: strike = 320°N, dip = 90°).

To generate the Green's functions for each fault segment, we use the analytical expressions for strike slip and dip slip on a fault embedded in an elastic homogeneous half space given by *Okada* [1985]. As we do not have a universal reference, we propagate our network of rover-base station pairs into the inversion scheme. The base stations' Green's functions are subtracted from those of the respective rovers to set-up an inversion according to baseline convention. Once the Green's functions are set up, the Parameter Estimator begins consuming offset estimates as they become available.

We use data for all baselines that provide offsets to invert for strike slip and dip slip along the fault segments. No weighting of the contributions based on quality has been implemented yet. The least-squares inversion applies Laplacian smoothing with a constant smoothing factor to regularize the solution and smoothes the fault ends to zero. The individual slip estimates combine to a total geodetic moment and hence a geodetic magnitude. The results (segments with slip, $M_w$, uncertainties, fault portion that accommodates 90% of total slip), including misfit measures (WRSS) are published in different formats (GMT plot files, ASCII tables, and ActiveMQ messages to ShakeAlert's decision module).

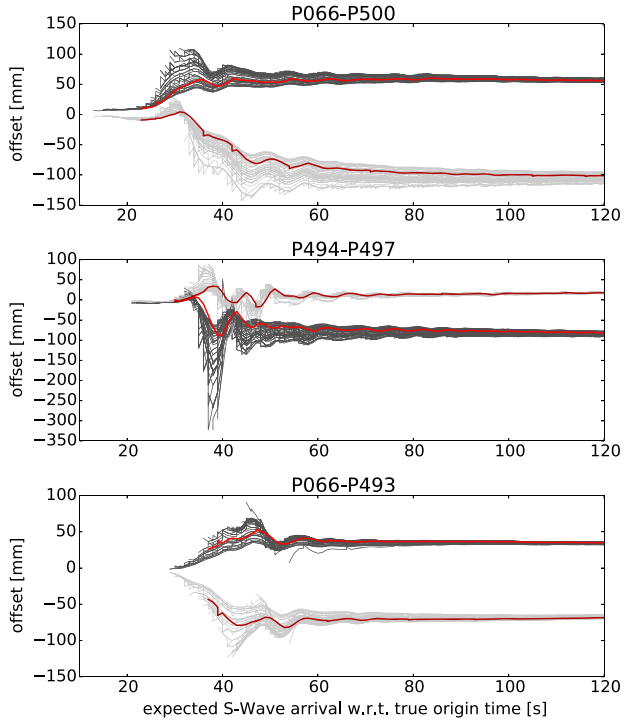


**Figure 8.** Stack of offset time series for three baselines; light grey is east offset and dark grey is north offset with respect to base station. Figure shows the variation and respective time evolution of offsets depending on event origin time uncertainties. Offset evolution per baseline is given relative to USGS event origin time (22:40:47 UTC, red lines are offset evolution triggered given this origin time). Panels are sorted by distance from the epicenter, hence the shift in onset of offset estimation, which depends on the estimated *S* wave arrival time at the base line (see text for details).

The inversion repeats every epoch, or as often as processing delays allow. The implemented approach has no memory, i.e., previous solutions are not taken into account to perform temporal smoothing commonly applied to postprocessed, high spatial resolution inversions [e.g., *Yue and Lay*, 2011]. While G-larmS provides options to select only the most recent offset data, or offset data that are not older than a certain time interval, we take in all available data and plan to downweight data based on their age: 1/(current_time - last_offset_time) in the inversion.

## 3. Synthetic Test: Simulated Offsets on the Hayward Fault

To test our implementation, we simulate a $M_w$=7.0 earthquake along the Hayward fault through 1.25 m of static, right lateral offset (purple patch in Figure 3 is surface projection of 70 km long, 320° strike, 90° dip fault). In this initial test we rupture the fault from 12 km depth to the surface, using the analytical expressions from *Okada* [1985] in the forward model. We want to stress that this model is unrealistic, as the Hayward Fault is not locked in the upper 3–5 km [e.g., *Bürgmann et al.*, 2000; *Aagaard et al.*, 2010]. However, it provides a useful test of the implementation. In section 5.2, we compare this surface rupture scenario to (still simplified) buried fault rupture scenarios.

We model absolute station displacements induced by the model and convert these into baseline displacements by subtracting base station offsets from the respective rover station offsets. A synthetic ShakeAlert message (event magnitude is "underestimated" as $M_w$ = 5.0, issued at 04:17:00 UTC) triggers G-larmS' Offset Estimator, Parameter Estimator, and the simulator. After the hypothetical *S* wave arrival at the sites closest to the epicenter (star in Figure 3, 37.77°N and 122.139°W) the simulator adds the respective baseline offset to real-time positioning solutions (5 September 2013, 04:12:00–04:24:00 UTC), which capture real-time noise. G-larmS' performance can be evaluated by how well the Offset Estimator recovers the synthetic static offset and how well the slip suggested by the Parameter Estimator matches the actual slip. This test employs the entire workflow shown in Figure 2. Although working on archived real-time data from the file system, G-larmS operates very similar to real-time mode.

Figures 3a–3d show the time series, including the simulated slip (step function), for four baselines (subplot labels of time series correspond to baseline labels on the map). The simulated move out of the static offset arrival is very obvious; so is the decrease in displacement amplitude with distance from the epicenter even when integrated over long baselines. Figure 3d gives an impression of typical real-time noise in the horizontal time series. This noise is similar across the network but does not show up as clearly in Figures 3a–3c as it is clearly dominated by the static offsets.

Running the simulation through G-larmS gives the solutions shown in Figure 4. The map view shows the forward model (skinny purple patch) and the respective first and final solutions for recovered slip along the fault patches 4 and 76 s after the event origin time. The baselines are colored by RMS misfit comparing

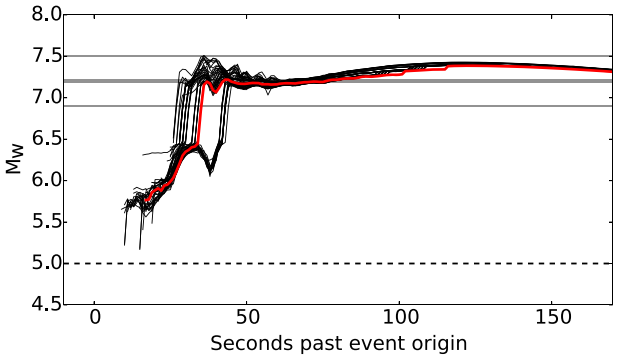


**Figure 9.** Stack of magnitude time series and their dependence on event origin time uncertainties. Each magnitude time series is derived from offsets estimated between −7 and +13 s of USGS event origin time (22:40:47 UTC). Premature onset of offset estimation can lead to delays in reaching the final magnitude. Late onset of offset estimation can result in slight oscillation of magnitude estimates due to unsmoothed dynamic motion at the GPS sites due to shaking. Target magnitude of $M_w$ = 7.2 is reached between about 28 and 46 s after the event origin time.

baseline displacements going into G-larmS (synthetic $M_w$7.0 test case) to baseline displacements induced by the recovered slip model (note that baselines in the original gray indicate missing data for these baselines). Even with few baselines after 4 s (Figure 4), the magnitude is well recovered at $M_w$ = 6.76. Slip distribution and rake indicate the characteristics of the forward model, but the recovered fault (still only the initial five patches) is too short which explains the moderate fit to the data. Within 2 more seconds the input model and magnitude are stably recovered at good fit to the data (see Figure 4b, and supporting information). The minor differences are due to smoothing and real-time noise.

Movie S1 steps through the time series of slip estimation and shows that it is stable even without enforced temporal smoothing. As the fault grows initially, overestimated slip gets spread out over more fault panels, but we do not observe erratic jumping of slip between different fault segments. The colors of the baselines in Movie S2 indicate microstrain derived from the estimated offsets instead of RMS. This gives a sense for the simulated spreading of static offsets through the network.

This test shows that the offset estimation is insensitive to the average background noise in real-time data if the signal dominates the noise. The input offsets are well recovered (Figure 4b), which results in a quick derivation of a satisfying distributed slip model.

## 4. Data Test: The 2010 El Mayor-Cucapah Earthquake

The synthetic test of the previous section instills confidence in the implementation of our system. However, to understand how dynamic motions from *S* waves and surface waves during an actual earthquake affect the results, we replay data from the $M_w$7.2 El Mayor-Cucapah earthquake [e.g., *Hauksson et al.*, 2010] through G-larmS. Though sampled only in the far field, the event is the closest analogy to a Northern California earthquake captured by high-rate GPS. The fault geometry and slip history for this event, however, have been shown to be more complex than simple right-lateral strike slip on a single fault plane [see *Wei et al.*, 2011], which provides an implicit test of our simplifying assumptions.

We obtained high-rate rinex data for the El Mayor-Cucapah earthquake from the SOPAC archive. Triangulation of the station network that recorded the event to define the baselines is all that is necessary to translate our strategy to Southern California. Figure 5 shows the resulting network of baselines including vectors representing absolute horizontal offsets due to the event at some GPS stations. We use `trackRT` rewind (`trackRTr`) to simulate real-time displacement time series; `trackRTr` mimics the behavior of `trackRT` but runs off-line on rinex data rather than real-time RTCM streams. We use the same settings and models as in real-time processing; including the lower quality, predicted IGS orbits.

Similarly to the synthetic test above, we replay the displacement time series (e.g., Figures 6a–6d) through G-larmS but without adding any synthetic offsets. Instead, the dynamic motions due to *S* wave and surface waves (e.g., Figures 6a–6d) will challenge the offset estimation algorithm. This test is triggered by a synthetic CISN ShakeAlert message with the final USGS parameters for origin time (4 April 2010 at 22:40:42 UTC), location (32.259°N, 115.287°E), and an underestimated magnitude estimate of $M_w$ = 5.0. The use of an underestimated magnitude is to show that the system is insensitive to this value, and faults will grow if they have to accommodate larger slip. Use of the final USGS location and origin time reflects the expected performance of a seismic EEW system [e.g., *Kuyuk et al.*, 2014] given the required station density [*Kuyuk and Allen*, 2013].

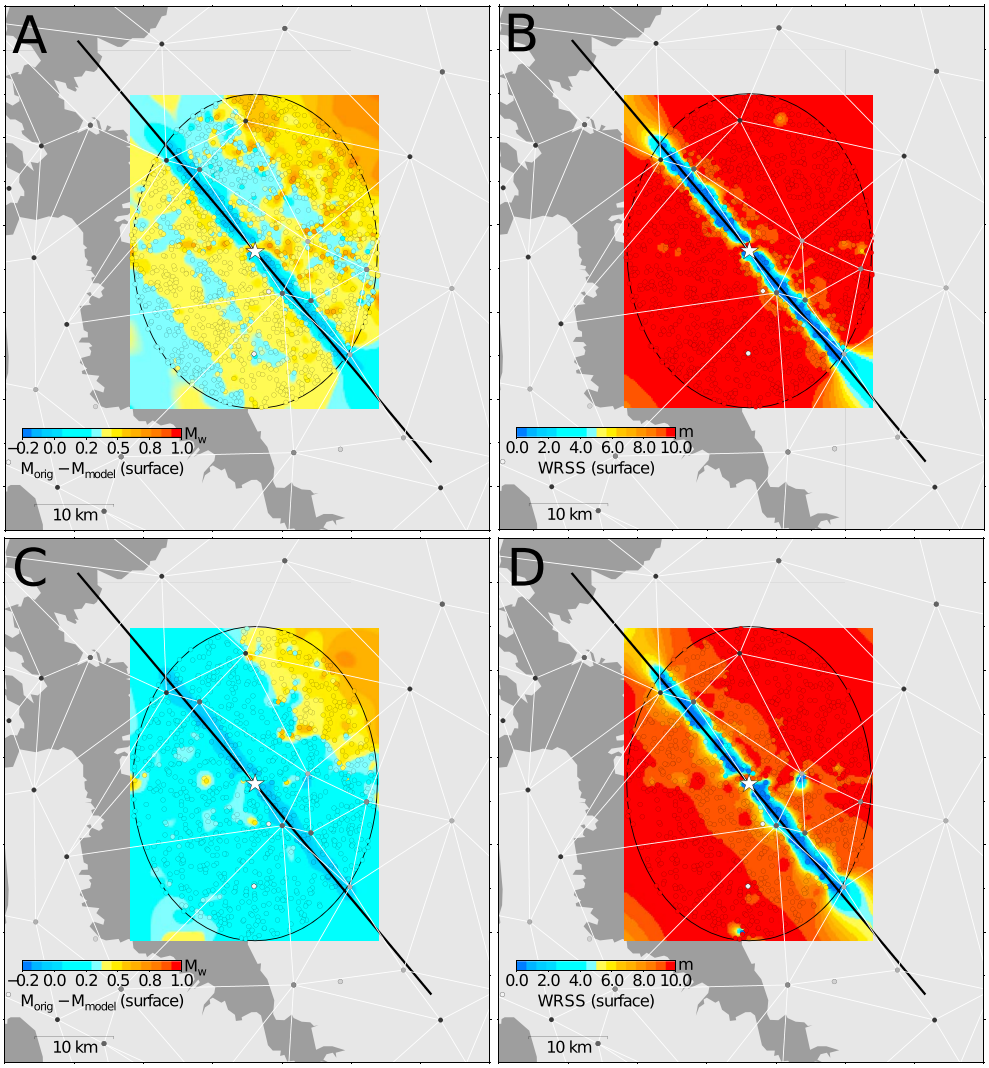


**Figure 10.** (a and c) Magnitude error and (b and d) model misfit to data (WRSS) for synthetic Hayward fault event that ruptures from the surface to 12 km depth at about 75 s after event. The event origin for the slip inversion is randomly changed to locations indicated by circles within the ellipse. Laplacian smoothing was higher in the upper row than in the lower row. Ellipse is about 37 km around event. Background colors are interpolated values derived at the circles. Growing fault can accommodate along-strike mislocation. Sensitive to off-strike mislocations with magnitude errors up to 1.0 and significantly increased misfit to data.

One of the first solutions, showing the slip distribution 36 s after the event origin time, is given in Figure 7a. The estimated magnitude at $M_w$ = 7.16 is similar to the final magnitude ($M_w$ = 7.2) derived after the event. The slip, however, is concentrated somewhat at the northern end of the rupture plane. Given the one-sided, far-field GPS observations, this is no surprise. The offset estimation (Figure 7a) begins too early due to strike slip rupture not starting until 15 s after event origin [*Wei et al.*, 2011]. This explains the early, low initial magnitude estimates. This snap shot shows a $M_w$ = 7.16 solution 2 s after *Colombelli et al.* [2013] indicate static offset arrival (shown by the step in the magnitude solution) and give their first solution at $M_w$ = 7.08.

The final solution (90 s after the origin time) for the replay of the El Mayor-Cucapah data is shown in Figure 7b. Note that the high strain in the north-west corner of Figure 7b is due to noise at some sites and not related to the earthquake. The final slip distribution spreads over a longer fault than observed through the surface rupture trace (red line in map view of Figures 6 and 7b). Hence, peak slip of ≈2 m is lower than the 6 m derived by *Wei et al.* [2011], but our fault discretization is coarser and, as opposed to *Wei et al.* [2011], we are limited to only far-field GPS data. The one-sidedness of the observations also explains why the slip is closer to the network than the actual surface rupture. The offset estimation (black and blue crosses in

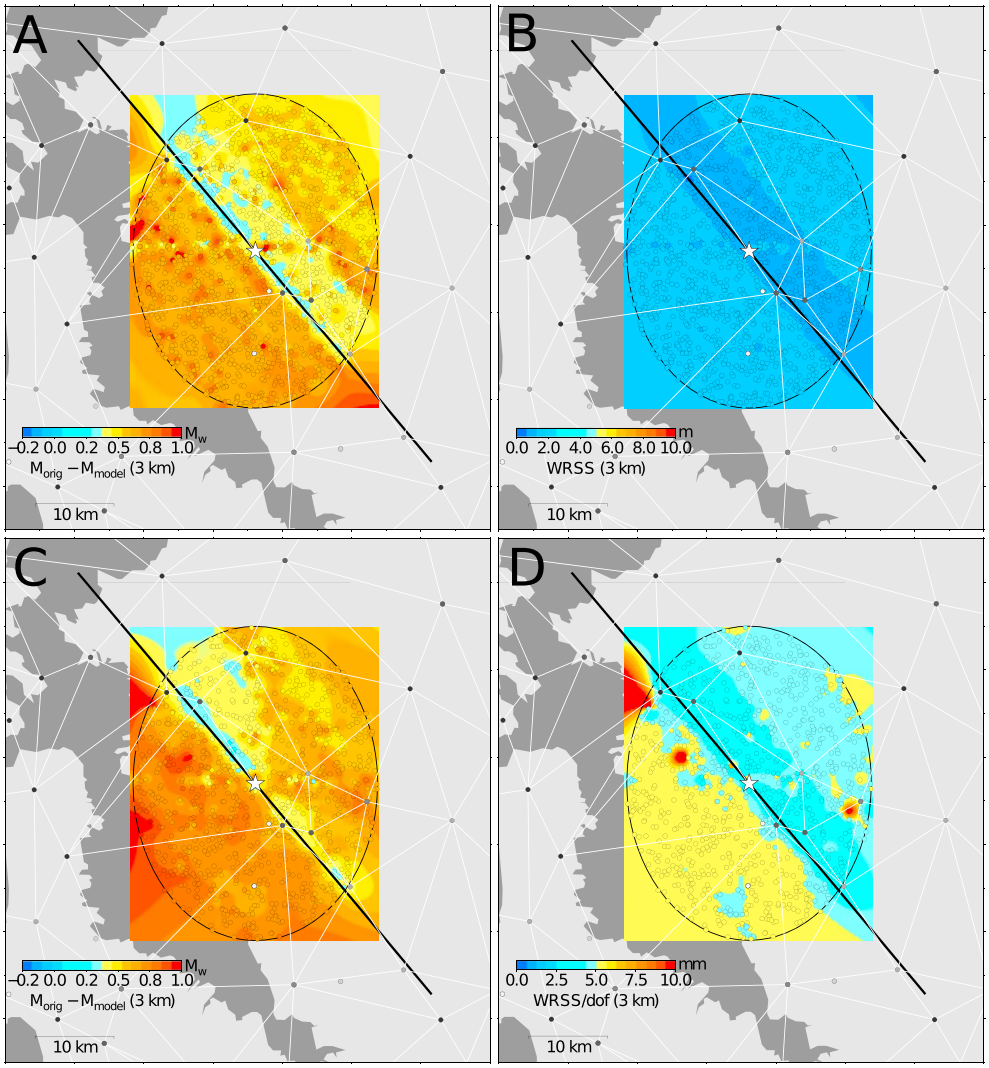


**Figure 11.** Same as Figure 10 except that the fault in the forward model ruptures from 3 to 12 km depth (inversion assumes surface rupture). Magnitude recovery is significantly decreased at generally good model fit. This is due to minimal or unphysical slip (with less Laplacian smoothing) being put in the model fault in the inversion.

baseline time series in Figure 7b) is rather insensitive to the dynamic motion, and the static offsets are quickly recovered as previously suggested by *Allen and Ziv* [2011].

Movies S3 and S4 step through the time series of slip estimation and show a coherent picture of stable magnitude and therefore stable total slip estimates. The slip distribution, however, depends on the location of available static offsets. At first, with measured offsets only in the near field, but many baselines included in the inversion, the slip gets pulled close to stations that indicate static offset. The slip then gets pushed farther out until it stabilizes to the distribution shown in Figure 7.

## 5. Sensitivity Analyses

We simplified the slip estimation to a bare bones inversion to limit complexities and currently do not estimate event location, fault strike, and fault dip. This results in a purely linear inverse problem which can be easily solved at each epoch. However, it is important to get an intuition about the sensitivity of this method to uncertainties in our assumptions about strike and dip, and to uncertainties in input values such as origin time and event location. *Kuyuk et al.* [2014] give an overview of empirical location and timing uncertainties for ElarmS-2, one of the algorithms contributing to CISN ShakeAlert. In the following section, we test the

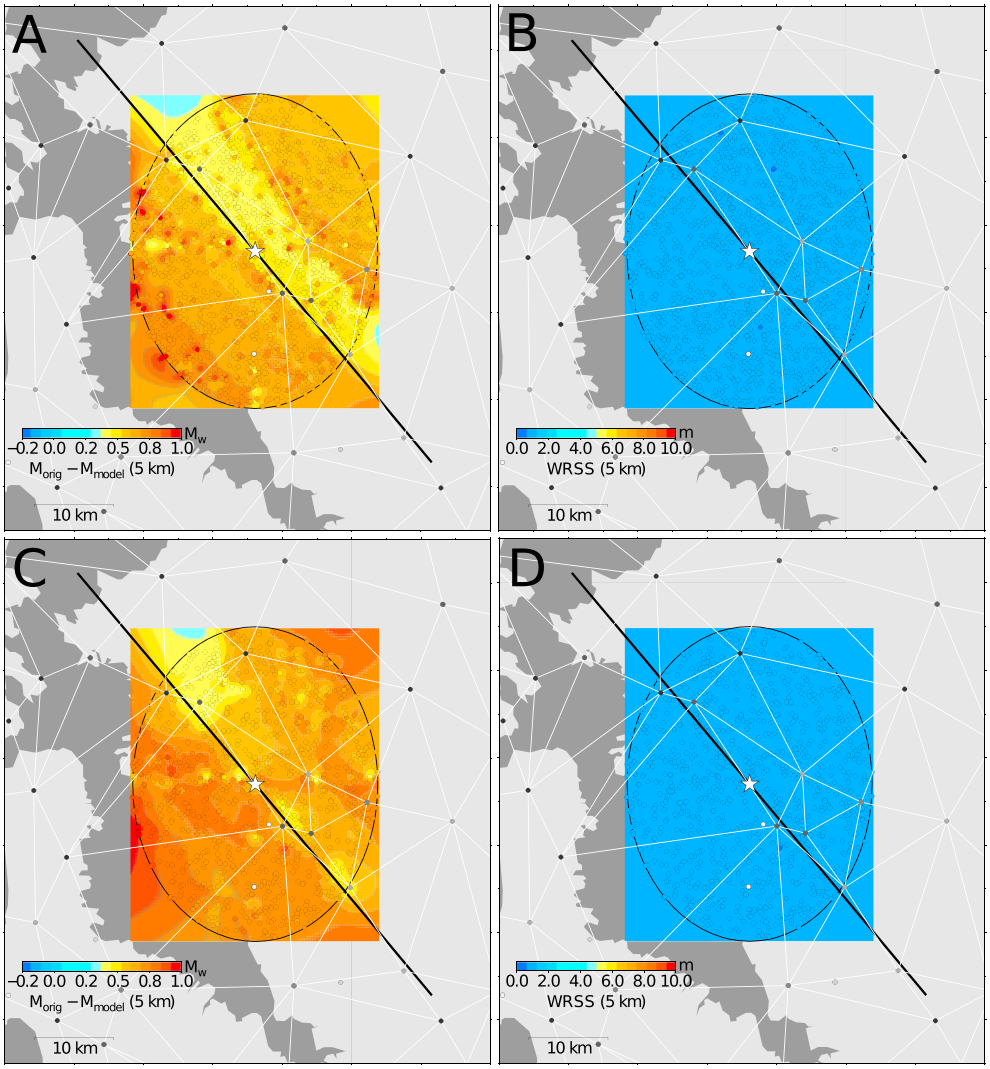


**Figure 12.** Same as Figure 11 except that the fault in the forward model ruptures from 5 to 12 km depth (inversion assumes surface rupture). Magnitude error increases and misfit field homogenizes even more compared to Figure 11.

sensitivity of four scenarios to variation of location and origin time within these empirical uncertainties and we perturb strike and dip of the fault in the inversion.

### 5.1. Offset Estimation Sensitivity to Origin Time

G-larmS' Offset Estimator depends on an assumption of effective *S* wave travel time from the epicenter to the baseline station closest to the event. Especially in the near field, the assumed velocity depends highly on the depth of the event which we currently do not take into account (ElarmS-2 fixes depth at 8 km). This introduces uncertainty in the expected onset of transient and dynamic motion, which may result in postevent position estimation to begin too early or too late. While the former contaminates initial postevent positions with preevent data, and hence delays the time at which a final estimate of the coseismic offset is reached, the latter may result in significant biases depending on the phase of a dynamic feature (*S* waves, surface waves) at which postevent position averaging begins.

To get an intuition of how origin time uncertainties and hence onset of postevent position estimation affect the slip model and magnitude estimates, we repeatedly replay the El-Mayor Cucapah earthquake as described in section 4 but on a smaller network of 48 baselines (see black baselines in Figure 5). This exercise also approximates the impact incorrect effective *S* wave travel times have on the offset estimations. In each of 50 runs, we perturb the onset of postevent position estimation randomly within an asymmetric 20 s interval around the USGS origin time at 22:40:42 UTC. We assume that the timing error is at

most −7 and +13 s, which more than doubles observed uncertainties in origin time determination for *Kuyuk et al.* [2014] for ElarmS-2. Our approach mimics the effect of changes in event origin time while keeping the preevent baseline positions constant within data noise (both, preevent and postevent positions depend on event origin time, we want to alter only one). All other parameters are held constant.

Figure 8 shows static offset estimate ensembles for three baselines (compare to Figure 5 for location). The red lines represent the respective static offset time series estimated in section 4 using the USGS event origin time. We observe spread of offsets depending on onset time. Higher initial offsets are estimated at later offset determination onset. This is no surprise as later offset determination can catch the displacements at maxima of dynamic motion. The individual offsets converge to a common value representing the static displacements over about 1 wavelength of the displacement time series (≈ 10 s).

The resulting magnitude time series with respect to true origin time are given in Figure 9 (red line uses USGS origin time). The target range of $M_w$7.2 ± 0.3 is indicated by the horizontal gray lines; the hatched line shows the input magnitude $M_w$ = 5.0. We see a distinct inverse correlation that gives the target magnitude quicker for later origin times. Not surprisingly, if too many preevent data are averaged into the postevent position, it takes longer for the magnitude to reach the target value. This is obvious looking at the red line (based on USGS origin time, but does not represent the onset of slip on the strike-slip segments, which was 15 s later [*Wei et al.*, 2011]), which takes significantly longer to jump to $M_w$ = 7.2 due to preevent position contamination of the coseismic offset estimates. More importantly, the impact of the dynamic motion does not overwhelm the solution. We observe some overshooting, up to $M_w$7.5 due to undamped dynamic offsets, but generally we remain within target range and converge after some minor oscillations. The long wavelength bump between 75 and 150 s is due to noise in the NW part of the network (see Figure 7). The operational version does not use data from stations that far away.

### 5.2. Parameter Estimation Sensitivity to Event Location and Patch Width

The event location is another important parameter that we currently do not solve for but instead use the solution from the seismic system. To test the sensitivity to errors in event location and a priori defined fault geometry, we consider the estimated offsets fixed. While the estimated event location affects the estimated *S* wave arrival time and hence the onset of offset estimation in our algorithm (see section 5.1), we investigate only late magnitude estimates to isolate the impact event mislocation will have on our slip model.

We investigate four cases, for each we vary the amount of Laplacian smoothing of slip in the inversion: the already introduced synthetic Hayward model and the El Mayor-Cucapah event, as well as two variations of the Hayward fault model, which assume scenarios of interseismic fault creep for the upper 3 and 5 km, respectively. This means that the uppermost sections of the fault do not slip during the event. For these buried faults we adjust the amount of slip in the lower section of the fault such that the magnitude still amounts to $M_w$ = 7.0 (1.70 m and 2.20 m, respectively). The inversion of slip on the buried faults, however, still assumes patches that rupture from the surface to a depth of 12 km (current default for California). This is a demonstration of the impact of simplified patch width assumptions. We also investigate effects of location uncertainty on the El Mayor-Cucapah event, which is a good proxy for far field effects, though the rupture complexities of this event might bias the results.

We replay the offsets estimated for these cases through the Parameter Estimator as described in sections 3 and 4, respectively. For each run in our test series, we randomly perturb the a priori location in our alert message within an approximately 20 km radius of the true event location. Hence, the model faults in each test run are centered on the perturbed location and parallel to the true fault. Considering, for example, ElarmS-2' empirical location uncertainties [*Kuyuk et al.*, 2014], this very generously doubles the mislocation region for $M_w > 3.5$ events. We repeat this experiment for each of the four input data sets 1000 times to thoroughly sample the area.

Figure 10 zooms in on the Hayward fault scenario that ruptures the surface (black line is surface trace of forward model) and shows the error of the final recovered magnitude (Figures 10a and 10c) and misfit of modeled GPS displacement to recorded displacements (Figure 10b and 10d) (40 s after event onset). Each circle indicates one of the 1000 perturbed locations on which the inversion fault is centered (striking parallel to forward model fault). The colored background is an interpolation of the values at these circles. The inversions in the upper row were subject to higher Laplacian smoothing than those in the lower row.

The recovered magnitude in Figure 10a fits the modeled magnitude well at perturbed locations along strike of the original model fault. The total misfit of the respective inversion result compared to the input data is also insignificant at these locations. This is due to the growth of the fault in the inversion which results in faults that cover the area of slip from the synthetic input, except not centered on the perturbed event location.

As soon as we move a critical distance perpendicular to the forward model fault, we increase the error in recovered magnitude to larger than 0.5 magnitude units (underestimate) accompanied by a large increase in model misfit. This critical distance (blue band in Figure 10a) coincides with the location of GPS sites that are close to the fault. The resulting misfit can be explained by significant offsets along baselines associated with these near-field sites due to the surface rupture in the forward model. Sites close to the original fault show the wrong sense of motion when on the opposing side of a strike slip fault. Hence, the inversion has little chance to find slip values that fit the observations.

The case of low smoothing in Figure 10c appears to recover the magnitude much better over a larger region than the high-smoothing case in Figure 10a. The misfit to the data shows the same pattern as Figure 10b. Using Laplacian smoothing to regularize the slip model, i.e., get all fault patches to slip in roughly the same direction, slip has to be minimal to minimize the misfit. Minimal slip inhibits growth of the fault (as growth is controlled by magnitude.) and hence limits the maximum magnitude that can be recovered. When we lower the smoothing in Figure 10c, we get unphysical slip distributions at distances outside the narrow band of good fit in Figure 10a, i.e., right-lateral and left-lateral slip for this very simplistic synthetic event. While this results in higher total slip and therefore a higher magnitude, which results in a longer fault, it does not provide a good fit to the data (Figure 10d).

Figures 11 and 12 are very similar to Figure 10 but instead of assuming surface rupture, the forward models use a buried fault reaching from 3-12 km and 5-12 km depth, respectively. The color scales are the same in Figures 10-12. The difference between Figures 11 and 12 is marginal. Neither recovers the magnitude well at any point in the region of interest (0.5–0.8 magnitude units underestimate, with a small band of about $M_w$0.2 error on top of the synthetic fault buried at 3 km). The amounts of smoothing do not make as big a difference as in Figure 10. However, both suggest a significantly improved, almost uniform misfit field (slightly better for the more deeply buried fault) over the surface rupture model. This is explained by the fact that a rupture on a buried fault induces a more distributed displacement field with overall smaller magnitudes. Assuming the wrong fault patch width (i.e., surface rupture) with sites close to the ruptured fault will constrain slip to be small to match the observations and overall fit the observations better. (Keep in mind that small offsets in the far field may be due to real-time noise and hence contribute positively to the inversion results; additionally this depends on magnitude)

The sensitivity to location misestimates is much decreased in the El Mayor-Cucapah case (Figure 13). When much smoothing is applied (Figures 13a and 13b), good magnitude recovery and fit to data are obtained within the north-western quadrant of the mislocation ellipse. This is to be expected considering that the GPS network observes this event one-sidedly from this direction. The fit to the data for the rest of the ellipse is good except for a few outliers, but generally at poor magnitude recovery. Once the smoothing is reduced (Figure 13c and 13d), the region of good magnitude recovery increases as increased smoothing tends to inhibit fault growth and therefore the ability to accumulate slip. Regions of gross misfit in Figure 13d correlate with points of magnitude overestimation combined with unphysical slip on the fault interface in Figure 13c.

### 5.3. Parameter Estimation Sensitivity to Fault Geometry

To evaluate the sensitivity of our inversion strategy to fault geometry we investigate changes in strike and dip (effects of patch width misestimates are obvious from the previous section). The results in Figure 14 address the four less smoothed test cases and mostly reiterate observations for location sensitivities.

Figure 14, first row shows results for the synthetic Hayward scenario with surface rupture (0 km burial depth). We find very good magnitude recovery and fit to the data from 87 to 93° and a somewhat poorer magnitude recovery with significantly decreased fit to the data for 85–97° and 93–95°. Beyond that the magnitude recovery drops to an underestimate of 0.2–0.4 magnitude units with off-the-chart misfit to the data. Between 40–60° and 120–140° the magnitude recovery seems to improve. However, recall that the figure shows final magnitude recovery. In some of these apparently good cases, the initial slip models give

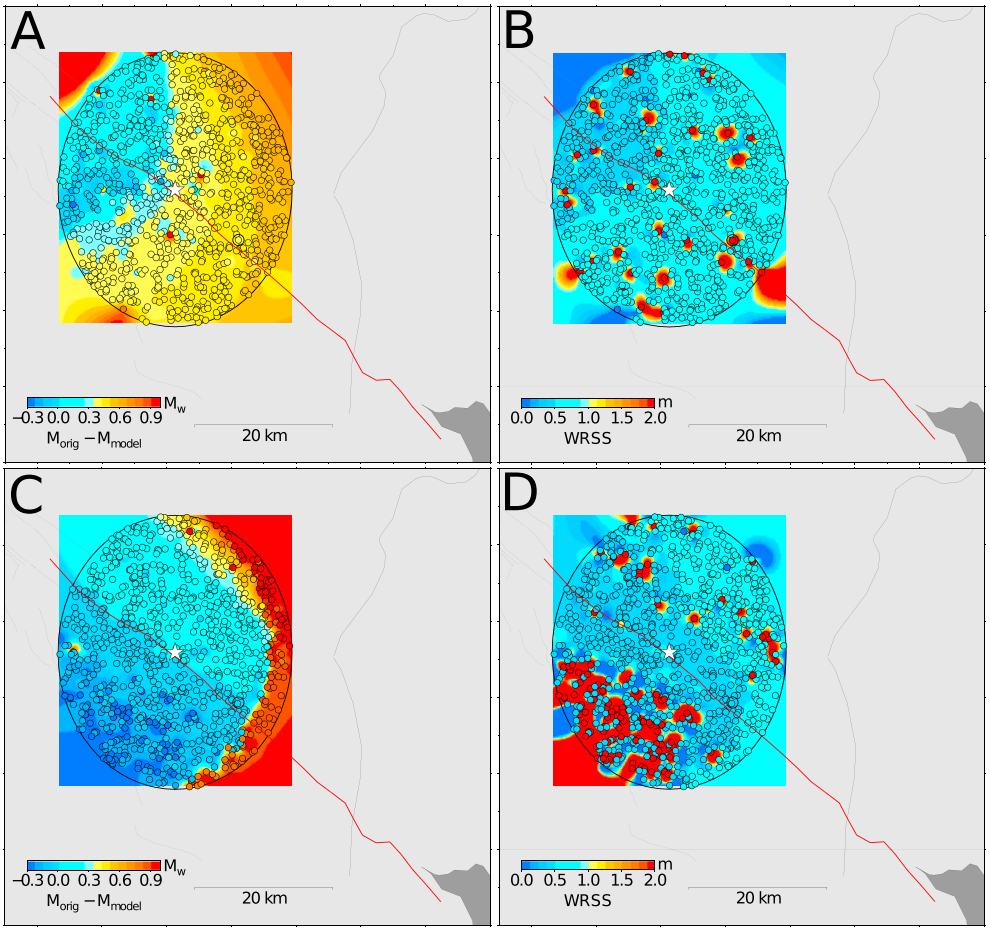


**Figure 13.** Setup is similar to Figures 10–12, but for the El Mayor-Cucapah event (scales are changed) at about 120 s after the event onset when alerted event origin is randomly changed to locations indicated by circles within the ellipse. Ellipse is about 40 km around the actual event epicenter. Less Laplacian smoothing results in broader area of good magnitude recovery with good fit to data. This is due to the distance of the event from the network and hence decreased sensitivity to location uncertainties. More smoothing tends to inhibit fault growth and hence the ability to accumulate slip. This explains the triangle pattern of better fit of erroneous locations toward the network in the upper row. Regions of gross misfit in the lower row correlate with points of magnitude overestimation combined with unphysical slip on the fault interface.

left lateral slip with magnitudes up to $M_w$ = 8.0, which results in very long faults. The slip distribution is at times split into two patches with opposing sense of slip. At later points, with more data available in the network, the slip on these patches decreases but is still large enough to explain an event of a magnitude close to $M_w$ = 7.0. However, the fit to the data is not good at any point in time. The models with largest misfits tend to produce slip distributions with both right-lateral and left-lateral slip (maximum slip about 30–60% of forward model, with back slip) even with higher smoothing. Some of these effects are due to the fixed fault depth from 0 km to 12 km. An increase in dip results in an increase in fault width, which means less slip does not necessarily decrease the moment.

The strike for this first Hayward scenario (Figure 14, left column) shows a similar pattern of good fit and magnitude recovery for a strike of 320 ± 5°. Outside of this range, the fit deteriorates significantly at about similar magnitude recovery. Just like with the dip, we find that back slip is introduced into slip models the more the strike deviates from the nominal value. While 0.3–0.4 magnitude units may seem a decent recovery, considering the size and number of fault patches, we have to keep in mind that $M_w$6.5 values can be produced by fitting slip on sufficiently large patches to real-time noise.

The second and third rows in Figure 14 show results for the Hayward fault scenarios that are buried at 3 and 5 km, respectively. The plots reiterate observations from the previous section investigating mislocations: the

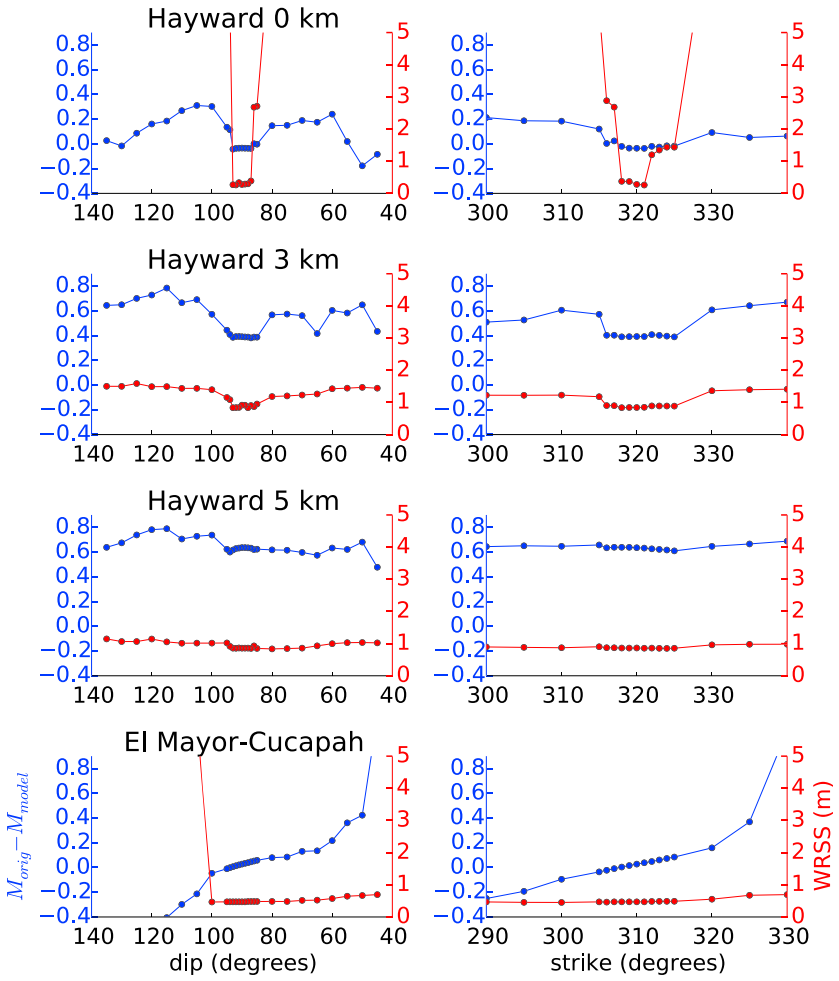


**Figure 14.** Sensitivities to a priori assumptions about strike and dip in the inversion. Three forward models (fault from 0, 3, and 5 km to 12 km depth, respectively) with 90 dip creating $M_w$7.0 equivalent slip and the data from the El Mayor-Cucapah event are inverted multiple times. Each time we vary dip (left column) and strike (right column) of fault plane for the inversion. All panels show in blue the final error in magnitude recovery after the inversion (≈4 min after event origin), while the red graphs show the respective WRSS misfit between input data and recovered model. The left column displays changes in dip which we varied from 45 to 135° for all cases (note that the plots order degrees of dip in decreasing order as this translates more intuitively to map view: 135–91° dip to the SW and 89–45° dip to NE). The right column shows changes in strike, which we varied from 290 to 330° for El Mayor-Cucapah and 300 to 340° for the synthetic Hayward cases. For all tests we increased the angles in 5° steps but decreased this interval to 1° steps around the nominally correct values. The dots in the panels mark the discrete test values. We connect the dots to guide the reader's eye.

data can be fit well (not as well as near the nominal values in the surface rupture case) with any model while the initial magnitude is not being recovered (0.4–0.7 magnitude units, maximum slip for most dip/strike ranges from 10 to 40% of forward model).

The last row treats the El Mayor-Cucapah event. We want to restate the actual complexities of the rupture as presented by *Wei et al.* [2011]: four rupture planes, one of which is an extensional feature on which the event nucleates and then spreads after 15 s onto the other planes which dip to opposing directions (northern segments dip to east, southern segment to west) at 310° strike. While our "best guess" single dip and strike model does not compare to results from a high-resolution study, this case provides a test of whether such simplifications can be adequate for real-time analysis.

The bottom row in Figure 14 shows that the far field GPS data are more sensitive to the northern part of the rupture, which dips to the east. For dips from 60 to 100°, we get reasonable magnitude recovery at low misfits. However, dips larger than 100° (i.e., dipping to west by more than 10° from vertical) result in large

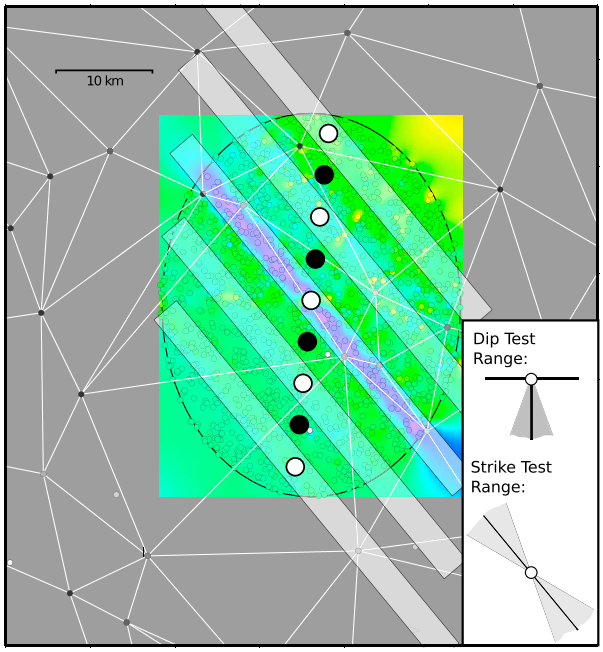


**Figure 15.** Proposed solution scheme that reduces the impact of uncertainties in a priori assumptions at reasonable computational cost. Main fault plane orientations as defined for respective fault province (here 320°) will be tested at nine locations (black and white circles), which cover the location uncertainties in the seismic system very well. For each of these, strike and dip variations will be evaluated as shown in the inset.

misfits and magnitude overestimates (paired with unphysical slip models). The strike misfit is generally good and hence does not indicate much about the quality of the derived magnitude, which is good to within 10° of the nominal 310°. The El Mayor-Cucapah results show that keeping to the tectonic regime for far-field observations is a prudent assumption and can result in reasonable magnitude and slip distribution recovery.

## 6. Discussion

The results for the synthetic Hayward Fault scenarios as well as for the El Mayor-Cucapah event show that G-larmS' algorithms for offset estimation and slip inversion give reasonable results. We want to emphasize that all tests were initialized with underestimated magnitudes, which results in initially short faults. The algorithms presented here currently analyze data for Northern California operationally in real time at the BSL. R. Grapenthin et al. (The 2014 $M_w$ 6.0 Napa Earthquake, California: Observations from real-time GPS-enhanced earthquake early warning, *Geophysical Research Letters*, in review, 2014) summarize their performance during the $M_w$ 6.0 South Napa earthquake and provide a preliminary latency analysis (≈ 6 s data latency, virtually no processing latency). The companion manuscript, that details G-larmS and is currently being prepared by the authors, will provide a more in-depth performance analysis and performance projection under network expansion.

Laplacian smoothing, as a way to regularize the solution, should be kept to a minimum, as it can inhibit fault growth. Oversmoothing of slip on short faults that are smoothed to zero at the ends will delay or completely inhibit the overcoming of thresholds [*Wells and Coppersmith*, 1994; *Colombelli et al.*, 2013] to lengthen the fault based on derived magnitude. Too little smoothing, in turn, will allow back slip. As the physical reasonableness of a slip model can be tested after its production, we reduced the smoothing in G-larmS to an acceptable minimum. However, systematic investigation of this important parameter is necessary. Ideally, smoothing will be based on network coverage and implemented in a manner that adapts depending on the slip models the inversion produces. Instead of applying Laplacian smoothing, *Duputel et al.* [2014] present a promising approach that includes model uncertainty into fault slip inversions. However, this comes at large computational cost, which makes it currently unsuitable for real-time applications. Experimentation with the performance of the genetic algorithm used by *Colombelli et al.* [2013] in a real-time environment is another option for future exploration.

Our investigation of the El Mayor-Cucapah event is complicated by the discrepancy between origin time and the actual onset of large strike slip 15 s later [*Wei et al.*, 2011]. The static displacements arrive later than G-larmS assumes. Early offset estimations average many samples of the preevent position into the postevent position. This smears out the buildup of static offset, which is an effect we intended to minimize. An improvement over the current approach would be for G-larmS to begin offset estimation based on its own detections of displacements emerging out of the noise (e.g., ratio of short-term average displacement to long-term average displacement (STA/LTA), *Allen and Ziv* [2011] and *Ohta et al.* [2012]) within a radius around the location from ShakeAlert.

The sensitivity analyses with respect to location uncertainty (Figures 10–12) show that a wrong initial location estimate impacts the results in the form of magnitude underestimates. This is mediated as the distance

between network and event increases (burying a fault is one way to increase the distance to the network). However, this must not be misunderstood as a recommendation for sparser networks! Instead, our results show that, in case of near-field ruptures, we have to be very careful with assumptions that simplify the inverse problem. We will have to invest in appropriate computing methods and technology to address the high resolution of the rupture process captured in the data. The parameters that are nonlinear in the slip inversion (location, strike, dip, and patch width) pose a parallel problem that can easily be addressed through, for example, a MapReduce [*Dean and Ghemawat*, 2008] approach.

The near-field cases (Hayward scenarios) show that if the earthquake matches our assumptions within small bounds (strike = 320 ± 5°, dip = 90 ± 3°, mislocation within 3 km orthogonal to surface trace in a dense network, surface rupture; for ElarmS-2 *Kuyuk et al.* [2014] infer an empirical, median location error of 2 km for events $M_w \geq 3.5$), we recover both slip distribution and magnitude well. Our sensitivity tests, however, show that further improvements of the inversion strategy are necessary to limit our exposure in unforeseen cases and poor fault discretization.

The sensitivities to geometry (Figures 11, 12, and 14) indicate that we do not need to know dip angles precisely, but in a case of a surface rupture under the network, we still need to know strike and dip to within 5°. While the use of only one patch stretching from the surface along dip works well for far-field events [*Colombelli et al.*, 2013], it is a poor discretization of faults under the network that have slip distributed along dip. We will replace this with a discretization of the fault surface that is based on data in a future version.

The tests with wrong assumptions about the slip distribution along dip present important implications for real-time applications. The examples shown in Figures 11, 12, and 14 (Hayward Fault buried at 3 km and 5 km) are an analog to the assumption that a fault creeps in certain regions and would not rupture during an event, yet it may. We get an overall reasonable fit to the data (through low slip, unphysical models including back slip, etc.) at a very poor magnitude recovery. Constraining a model too tightly in either way can result in misestimates. This is well known and usually easily corrected for in a research environment, but may be left unnoticed in a real-time implementation until a corner case event occurs.

## 7. Conclusion and Outlook

This paper follows two objectives: (1) to introduce processing details of our implementation of G-larmS in Northern California and validate it with synthetic and real data tests and (2) to scrutinize our current simplifying assumptions in terms of fault geometry and sensitivities to uncertainties in event location and event onset provided by the seismic system.

To test our approach, we add offsets due to synthetic slip on the Hayward fault to real-time data. We are able to recover input slip and magnitude. Additionally, we replay high-rate GPS data from Southern California that recorded the 2010 $M_w$ = 7.2 El Mayor-Cucapah earthquake. These data include dynamic motion due to *S* waves and surface waves. Again, magnitude and slip are well recovered.

To scrutinize our assumptions about strike, dip, timing, and quality of location estimates (from the seismic system), we perturb these parameters at the respective stages in the processing. The findings suggest that slip for an earthquake occurring within small bounds of our assumptions, i.e., strike = 320 ± 5°, dip = 90 ± 3°, mislocation within 3 km orthogonal to surface trace, and surface rupture, will be well recovered for ruptures within the network (depending on station distribution). Outside of these bounds, however, we may face unphysical slip models and poor magnitude recovery at relatively good fit of the model to the data. The latter can be particularly impactful when the along-dip partitioning of the fault surface is flawed. This is particularly important to bear in mind when multiple finite fault models are to be combined based on misfit.

In the case of California and other places with onshore crustal faults, the worst case scenario is undoubtedly the surface rupture of a strike-slip fault under a dense network. While dense station coverage provides high resolution and, most importantly fast observations, it requires us to sample the parameter space for the nonlinear finite slip inversion problem more thoroughly (compare results for surface rupture Hayward test case to El Mayor-Cucapah results). The most problematic issue is that of choosing any parameter such that individual (near field) stations end up on the wrong side of the model fault. While this highlights a problem with the fault province approach G-larmS currently implements, it also shows that relying solely on a

fault catalog (e.g., UCERF3 fault catalog [*Field et al.*, 2013]) and selecting the nearest fault to the ShakeAlert location can cause similar problems. As the parameter space exploration in real time cannot be exhaustive unless heavily parallelized (each inversion must complete within one data epoch), we have to exploit this property in the future. As real-time GPS is sought to contribute improved magnitude estimates within ±0.3 magnitude units and an approximate fault length for ground motion prediction to earthquake early warning, two simplifications are possible: (1) ignoring/downweighting near-field data at single outlier stations in a dense network that produce significant misfits (see Figure 10) and (2) placing model faults between station pairs with highest strain. These can be combined, and both reduce the complexity of the solution while providing approximate solutions. The tradeoff between speed and accuracy of the solutions requires thorough investigation.

Figure 15 outlines a scheme we will implement in the near future to reduce dependencies on a priori assumptions and external uncertainties. Within the radius of maximum location uncertainty, we place nine model faults parallel to the predominant strike and solve for slip on these (white rectangles with white dots and gaps between them indicated by black dots). Each of the model faults gives a decent solution for a swath width that depends to some degree on the fault orthogonal distance to the closest GPS stations (see Figure 10), but we fix this to 3 km based on the Hayward fault (which is the most densely instrumented). As the model faults grow based on slip that needs to be accommodated [*Colombelli et al.*, 2013], we will find a good slip distribution somewhere along the fault. The centroid of the slip maximum could be returned as geodetic location estimate. In each location we will vary dip and strike within the ranges indicated in the inset of Figure 15 (±20° to be tested in 5° steps).

The observation that low misfit does not guarantee good magnitude recovery but may imply wrong model assumptions (Figures 11 and 12) is well known and can be treated on a case-by-case basis in research environments. However, the underlying issue is that relying purely on misfit to data when exploring the parameter space can force us into a local minimum (e.g., one single dip angle for the entire fault plane). This is likely to happen with a rupture similar to El Mayor-Cucapah directly under the network in the Bay Area. We have to devise methodologies that simplify the problem of complex near-field ruptures such that it can be solved in real time. It is important to bear in mind that we are trying to solve a problem in real time that in an ideal post-processed world with many data still produces a variety of differing slip models and magnitude estimates. Data and model uncertainty are just now being separated [*Duputel et al.*, 2014], which will improve the scientific results, but will remain too expensive for a while to run in real time.

Through the test of sensitivity to onset time we have found that the current way of estimating postevent positions based on expected *S* wave arrival can add unnecessary delays and biases. Future implementations of G-larmS will include autonomous event detection through, e.g., STA/LTA based approaches [e.g., *Allen and Ziv*, 2011; *Ohta et al.*, 2012]. The triggering will likely remain to determine a preevent position for the baselines that are expected to experience coseismic offsets. After that, they will report offsets based on displacements emerging from the real-time noise above a statistically significant threshold.

Our results show that it is important for the community to work toward an extensive catalog of test events for real-time GPS-enhanced earthquake early warning. These catalogs must contain both real events and simulations that include shaking and static offsets. Unfortunately, few large events have been recorded by high-rate GPS, which makes realistic simulations our only viable option. The few examples and combinations of parameter perturbations, impact of smoothing, and distance between event and observations shown here stress the need for a huge array of such simulations. Of course, this raises the question of how constraints imposed on such models affect our results. This is likely negligible compared to the alternative of relying on the few events we have been using (El Mayor, Tohoku, and Tokachi-oki), which can result in tailoring our systems to characteristics of these few events (e.g., one-sided observations, far from network). Operational real-time GPS EEW systems will have to be benchmarked against such massive catalogs. The catalogs should include all imaginable edge cases that may even be more exotic than El Mayor-Cucapah or the 2012 Off-Sumatra event [*Meng et al.*, 2012]. We should anticipate complex ruptures in the near field and demonstrate how our finite fault systems will fare considering that most recent big earthquakes surprised the community in one way or another.

**Acknowledgments**
We thank Tom Herring's group at MIT for quick responses and accommodation of our requests regarding track/trackRT(r). Generous discussions with Sarah Minson, Jessica Murray, John Langbein, and others involved in the West Coast EEW effort improved this work. Roland Bürgmann provided input through discussions and suggestions with regard to test cases. Simona Colombelli shared code and insights from her [*Colombelli et al.*, 2013] work. Mario Arahna provides programming support at the BSL. Many figures were made with GMT [*Wessel and Smith*, 1995]. Real-time positioning solutions used for the Hayward Fault test cases and G-larmS are available from the authors. The El Mayor-Cucapah rinex data and offset solutions are available through the SOPAC archive at http://garner.ucsd.edu/pub/highrate/historical/rinex/2010/094/. Some data for this study come from the Bay Area Regional Deformation Network (BARD), doi:10.7932/BARD, operated by the UC Berkeley Seismological Laboratory, which is archived at the Northern California Earthquake Data Center (NCEDC), doi: 10.7932/NCEDC. Some real-time GPS data are provided by the Plate Boundary Observatory operated by UNAVCO for EarthScope and supported by NSF (EAR-0350028 and EAR-0732947). This work was funded by the Gordon and Betty Moore Foundation through grant GBMF3024 to UC Berkeley, and the USGS/National Earthquake Hazards Reduction Program Award G12AC20348.

## References

Aagaard, B. T., R. W. Graves, D. P. Schwartz, D. a. Ponce, and R. W. Graymer (2010), Ground-motion modeling of Hayward fault scenario earthquakes, Part I: Construction of the suite of scenarios, *Bull. Seismol. Soc. Am.*, *100*(6), 2927–2944, doi:10.1785/0120090324.

Allen, R. M., and H. Kanamori (2003), The potential for earthquake early warning in southern California, *Science*, *300*(5620), 786–789, doi:10.1126/science.1080912.

Allen, R. M., and A. Ziv (2011), Application of real-time GPS to earthquake early warning, *Geophys. Res. Lett.*, *38*, L16310, doi:10.1029/2011GL047947.

Avallone, A., et al. (2011), Very high rate (10 Hz) GPS seismology for moderate-magnitude earthquakes: The case of the $M_w$ 6.3 L'Aquila (central Italy) event, *J. Geophys. Res.*, *116*, B02305, doi:10.1029/2010JB007834.

Bilek, S. L., H. R. DeShon, and E. R. Engdahl (2011), Spatial variations in earthquake source characteristics within the 2011 $M_w = 9.0$ Tohoku, Japan rupture zone, *Geophys. Res. Lett.*, *39*, L09304, doi:10.1029/2012GL051399.

Blewitt, G., C. Kreemer, W. C. Hammond, H. P. Plag, S. Stein, and E. Okal (2006), Rapid determination of earthquake magnitude using GPS for tsunami warning systems, *Geophys. Res. Lett.*, *33*, L11309, doi:10.1029/2006GL026145.

Bock, Y., D. Melgar, and B. W. Crowell (2011), Real-time strong-motion broadband displacements from collocated GPS and accelerometers, *Bull. Seismol. Soc. Am.*, *101*(6), 2904–2925, doi:10.1785/0120110007.

Boehm, J., A. Niell, P. Tregoning, and H. Schuh (2006), Global mapping function (GMF): A new empirical mapping function based on numerical weather model data, *Geophys. Res. Lett.*, *33*, L07304, doi:10.1029/2005GL025546.

Boehm, J., R. Heinkelmann, and H. Schuh (2007), Short note: A global model of pressure and temperature for geodetic applications, *J. Geod.*, *81*, 679–683, doi:10.1007/s00190-007-0135-3.

Böse, M., et al. (2014), Early warning for geological disasters, in *Early Warning for Geological Disasters, Advanced Technologies in Earth Sciences*, edited by F. Wenzel and J. Zschau, chap. 3, pp. 49–69, Springer, Berlin, Heidelberg, doi:10.1007/978-3-642-12233-0.

Brown, H. M., R. M. Allen, and V. F. Grasso (2009), Testing ElarmS in Japan, *Seismol. Res. Lett.*, *80*(5), 727–739, doi:10.1785/gssrl.80.5.727.

Bürgmann, R., D. Schmidt, R. M. Nadeau, M. D'Alessio, E. Fielding, D. Manaker, T. V. McEvilly, and M. H. Murray (2000), Earthquake potential along the Northern Hayward Fault, California, *Science*, *289*(5482), 1178–1182, doi:10.1126/science.289.5482.1178.

Colombelli, S., R. M. Allen, and A. Zollo (2013), Application of real-time GPS to earthquake early warning in subduction and strike-slip environments, *J. Geophys. Res. Solid Earth*, *118*, 3448–3461, doi:10.1002/jgrb.50242.

Crowell, B. W., Y. Bock, and M. B. Squibb (2009), Demonstration of earthquake early warning using total displacement waveforms from real-time GPS networks, *Seismol. Res. Lett.*, *80*(5), 772–782, doi:10.1785/gssrl.80.5.772.

Dean, J., and S. Ghemawat (2008), MapReduce: Simplified data processing on large clusters, *Commun. ACM*, *51*(1), 107–113.

Dow, J. M., R. E. Neilan, and C. Rizos (2009), The international GNSS service in a changing landscape of global navigation satellite systems, *J. Geod.*, *83*, 191–198, doi:10.1007/s00190-008-0300-3.

Duputel, Z., P. S. Agram, M. Simons, S. E. Minson, and J. L. Beck (2014), Accounting for prediction uncertainty when inferring subsurface fault slip, *Geophys. J. Int.*, *197*(1), 464–482, doi:10.1093/gji/ggt517.

Field, E. H., et al. (2013), The uniform California earthquake rupture forecast, version 3 (UCERF3): The time-independent model, *U.S. Open File Rep., 2013-1165*, 97, U.S. Geol. Surv., Reston, Va.

Grapenthin, R. (2012), Volcano deformation and subdaily GPS products, PhD thesis, Univ. of Alaska Fairbanks, Fairbanks.

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

Guilhem, A., and D. S. Dreger (2011), Rapid detection and characterization of large earthquakes using quasi-finite-source Green's functions in continuous moment tensor inversion, *Geophys. Res. Lett.*, *38*, L13318, doi:10.1029/2011GL047550.

Hauksson, E. (2011), Crustal geophysics and seismicity in southern California, *Geophys. J. Int.*, *186*(1), 82–98, doi:10.1111/j.1365-246X.2011.05042.x.

Hauksson, E., J. Stock, K. Hutton, W. Yang, J. A. Vidal-Villegas, and H. Kanamori (2010), The 2010 $M_w$ 7.2 El Mayor-Cucapah earthquake sequence, Baja California, Mexico and Southernmost California, USA: Active seismotectonics along the Mexican Pacific Margin, *Pure Appl. Geophys.*, *168*(8–9), 1255–1277, doi:10.1007/s00024-010-0209-7.

Herring, T. A., R. W. King, and S. C. McClusky (2010), *GAMIT/GLOBK Reference Manuals, Release 10.4*. Massachusetts Institute of Technology, Cambridge, U. K.

Hoshiba, M., K. Iwakiri, N. Hayashimoto, and T. Shimoyama (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 Space*, *63*(7), 547–551, doi:10.5047/eps.2011.05.031.

Kerr, R. A. (2005), Failure to gauge the quake crippled the warning effort, *Science*, *307*, 201.

Kuyuk, H. S., and R. M. Allen (2013), Optimal seismic network density for earthquake early warning: A case study from California, *Seismol. Res. Lett.*, *84*(6), 946–954, doi:10.1785/0220130043.

Kuyuk, H. S., R. M. Allen, H. Brown, M. Hellweg, I. Henson, and D. Neuhauser (2014), Designing a network-based earthquake early warning algorithm for California: ElarmS-2, *Bull. Seismol. Soc. Am.*, *104*(1), 162–173, doi:10.1785/0120130146.

Larson, K. M., P. Bodin, and J. Gomberg (2003), Using 1-Hz GPS data to measure deformations caused by the Denali fault earthquake, *Science*, *300*(5624), 1421–1424, doi:10.1126/science.1084531.

Melgar, D., Y. Bock, and B. W. Crowell (2012), Real-time centroid moment tensor determination for large earthquakes from local and regional displacement records, *Geophys. J. Int.*, *188*(2), 703–718, doi:10.1111/j.1365-246X.2011.05297.x.

Meng, L., A. Inbal, and J.-P. Ampuero (2011), A window into the complexity of the dynamic rupture of the 2011 $M_w$ 9 Tohoku-Oki earthquake, *Geophys. Res. Lett.*, *38*, L00G07, doi:10.1029/2011GL048118.

Meng, L., J.-P. Ampuero, J. Stock, Z. Duputel, Y. Luo, and V. C. Tsai (2012), Earthquake in a maze: Compressional rupture branching during the 2012 M(w) 8.6 Sumatra earthquake, *Science*, *337*(6095), 724–726, doi:10.1126/science.1224030.

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

Ohta, Y., et al. (2012), Quasi real-time fault model estimation for near-field tsunami forecasting based on RTK-GPS analysis: Application to the 2011 Tohoku-Oki earthquake ($M_w$ 9.0), *J. Geophys. Res.*, *117*, B02311, doi:10.1029/2011JB008750.

Okada, Y. (1985), Surface deformation due to shear and tensile faults in a half-space, *Bull. Seismol. Soc. Am.*, *75*, 1135–1154.

Sibson, R. H. (1982), Fault zone models, heat flow, and the depth distribution of earthquakes in the continental crust of the United States, *Bull. Seismol. Soc. Am.*, *72*(1), 151–163.

Wald, D. J., B. C. Worden, V. Quitoriano, and K. L. Pankow (2006), ShakeMap manual, *Tech. Rep.*, U.S. Geological Survey Reston, Va.

Wei, S., et al. (2011), Superficial simplicity of the 2010 El Mayor-Cucapah earthquake of Baja California in Mexico, *Nat. Geosci.*, *4*(9), 615–618, doi:10.1038/ngeo1213.

Wells, D. L., and K. J. Coppersmith (1994), New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement, *Bull. Seismol. Soc. Am.*, *84*(4), 974–1002.

Wessel, P., and W. H. F. Smith (1995), New version of the generic mapping tools released, *Eos Trans. AGU*, *76*(33), 329, doi:10.1029/95EO00198.

Wright, T. J., N. Houlié, M. Hildyard, and T. Iwabuchi (2012), Real-time, reliable magnitudes for large earthquakes from 1 Hz GPS precise point positioning: The 2011 Tohoku-Oki (Japan) earthquake, *Geophys. Res. Lett.*, *39*, L12302, doi:10.1029/2012GL051894.

Yue, H., and T. Lay (2011), Inversion of high-rate (1 sps) GPS data for rupture process of the 11 March 2011 Tohoku earthquake ($M_w$ 9.1), *Geophys. Res. Lett.*, *38*, L00G09, doi:10.1029/2011GL048700.