UNIVERSITY OF CALIFORNIA SAN DIEGO

Seismogeodetic Methods for Earthquake and Tsunami Warning and Response

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Earth Sciences

by

Dara Elyse Goldberg

Committee in charge:

Yehuda Bock, Chair
Gabi Laske
Xanthippi Markenscoff
Diego Melgar
David Sandwell
Michael Todd

2018

The dissertation of Dara Elyse Goldberg is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____

_____

Chair

University of California San Diego

2018

DEDICATION

For Pa

EPIGRAPH

So many things are possible just as long as you don't know they're impossible

—Norton Juster, The Phantom Tollbooth

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I come from a family of educators, and I would not be where I am today if it were not for their commitment to questioning, learning, and teaching. I have been extraordinarily lucky to have their support and encouragement in all of my endeavors, academic and otherwise. Mom, Dad, Ally, Nannie, Pa, Grandma and Grandpa: I am so grateful for your guidance and unwavering support.

This dissertation would not be possible without the help of my dissertation committee. My advisor, Yehuda Bock, has mentored me throughout my PhD and helped me to become a better science communicator by providing opportunities to present my research at conferences, workshops, and universities around the world. Diego Melgar has been a constant source of enthusiasm, guidance, and encouragement. Gabi Laske, Xanthippi Markenscoff, David Sandwell, and Michael Todd, thank you for the many conversations and nudges to keep me moving toward this goal.

I am very thankful for the other members (past and present) of the Scripps Orbital and Permanent Array Center, each of whom has had a profound impact on the direction and success of my studies. Thank you Dana Caccamise, Peng Fang, Jianghui Geng, Jennifer Haase, Songnian Jiang, Emilie Klein, Allen Nance, Matt Norenberg, Glen Offield, Anne Sullivan, Mindy Squibb, and Maria Turingan.

My San Diego family has been my support system for many years now. My geophysics cohort: Matthew Cook, John Desanto, Adrian Doran, and Jessie Saunders, I am grateful for the years of friendship and look forward to our work intertwining in the future. Sarah, Maddie, Matt, and Dox: thank you for making a home with me on the other side of the country.

Portions of this dissertation have been published or are expected to be published in peer reviewed journals.

Chapter 2 in full, is a reprint of the material as it appears in Journal of Geophysical

Research: Goldberg, D.E. and Bock, Y. (2017). Self-contained local broadband seismogeodetic early warning system: Detection and location. *Journal of Geophysical Research: Solid Earth*, 122(4):3197–3220. The dissertation author was the primary investigator and author of this material.

Portions of Chapter 3 are in preparation for publication as: Goldberg, D.E., Melgar, D. and Bock, Y. Evaluation of earthquake determinism with seismogeodetic displacements, *in prep*. The dissertation author is the primary investigator and author of this material.

Chapter 4 is currently in revision for publication in Journal of Geophysical Research as: Goldberg, D.E., Melgar, D., Bock, Y., and Allen, R.M. (2018). Geodetic observations of rupture evolution. *Journal of Geophysical Research: Solid Earth*, *in revision*. The dissertation author is the primary investigator and author of this material.

Additional work from Chapter 5 is currently being prepared for publication. This work is coauthored with Yehuda Bock, Falko Kuester, and Eric Lo. The dissertation author is the primary investigator and author of this material.

VITA

2011	B. S. with Honors in Geology-Physics/Math, Brown University
2011	B. A. in Engineering, Brown University
2014	M. S. in Earth Scinces, University of California San Diego
2018	Ph. D. in Earth Sciences, University of California San Diego

PUBLICATIONS

Goldberg, D.E., Melgar, D. and Bock, Y. Evaluation of earthquake determinism with seismogeodetic displacements, *in prep*.

Goldberg, D.E., Melgar, D., Bock, Y., and Allen, R.M. (2018). Geodetic observations of rupture evolution. *Journal of Geophysical Research: Solid Earth*, *in revision.*

Ruhl, C.J., Melgar, D., Geng, J., **Goldberg, D.E.**, Crowell, B.W., Allen, R.M., Bock, Y., Barrientos, S., Riquelme, S., Baez, J.C., Cabral-Cano, E., Pérez-Campos, X., Hill, E.M., Protti, M., Ganas, A., Ruiz, M., Mothes, P., Jarrín, P., Nocquet, J.-M., Avouac, J.-P., and D'Anastassio, E. (2018). A global database of strong motion displacement GNSS recordings and an example application to PGD scaling. *Seismological Research Letters*, *accepted.*

Goldberg, D.E. and Bock, Y. (2017). Self-contained local broadband seismogeodetic early warning system: Detection and location. *Journal of Geophysical Research: Solid Earth*, 122(4):3197–3220.

Saunders, J.K., **Goldberg, D.E.**, Haase, J.S., Bock, Y., Offield, D.G., Melgar, D., Restrepo, J., Fleischman, R.B., Nema, A., Geng, J., Walls, C., Mann, D., and Mattioli, G.S. (2016) Seismogeodesy using GNSS and low-cost MEMS accelerometers: Perspectives for earthquake early warning and rapid response. *Bulletin of the Seismological Society of America*, 106(6):2469–2489.

ABSTRACT OF THE DISSERTATION

Seismogeodetic Methods for Earthquake and Tsunami Warning and Response

by

Dara Elyse Goldberg

Doctor of Philosophy in Earth Sciences

University of California San Diego, 2018

Yehuda Bock, Chair

This dissertation presents a series of algorithms and plans for implementation of a prototype earthquake early warning system by merging two seismic observation tools: strong-motion accelerometers and Global Navigation Satellite Systems (GNSS). The *seismogeodetic* approach, that which optimally merges these disparate data types, allows a reliable estimation of the earthquake source for early warning procedures from event detection to magnitude estimation and higher order products. We address implementation of this system in a real-time, automated environment, designed to operate without manual oversight. We assess the capabilities of low-cost micro-electro-mechanical systems (MEMS) accelerometers merged with observatory-grade GNSS for seismic P-wave detection, and suggest statistical methods to determine whether P-wave

detections are consistent between stations, which allows for the automated removal of poor quality detections to avoid propagating these errors to higher order warning products.

In line with these goals, we demonstrate how seismogeodetic observations contribute to our view of initial rupture dynamics of large earthquakes. We address an open question in seismology regarding the deterministic nature of earthquakes, assessing how early it is possible to fully characterize the source parameters of large, damaging earthquakes. Our seismogeodetic observations suggest that final earthquake magnitude is not discernable from the first few seconds of observation and therefore the earthquake rupture process is not strongly deterministic. We further investigate the complete temporal evolution of seismic moment release to identify the earliest magnitude-dependent features in our observational dataset. We create synthetic rupture models to identify the physical basis responsible for the timing of observed magnitude-dependent qualities. Our findings suggest that earthquake magnitude can be estimated prior to rupture completion, consistent with a weakly deterministic rupture process. These results provide new insights into the best practices for early warning and rapid response, and suggest limitations on the timeliness of earthquake source characterization.

Finally, we turn to the built environment to demonstrate the applicability of these multi-instrument sensors to long-term structural health and seismic monitoring, to bolster earthquake and other natural hazard response practices. We describe efforts to characterize local buildings by creating a baseline model for the healthy structure and propose methods by which structural health can be evaluated over time or following a major seismic event to assess building safety without the need for manual inspection. We again assess the capabilities of low-cost MEMS accelerometers and describe the limitations of this lower-quality instrumentation for structural monitoring applications.

Chapter 1

Introduction

1.1 Background

Earthquake hazard mitigation is important for protecting infrastructure, economic interests, and human life in seismically active areas. The last couple of decades have seen an enormous effort to advance hazard mitigation strategies through the development of seismic building codes, technologies to allow infrastructure to cross active fault lines with reduced risk, and earthquake and tsunami early warning systems. Earthquake early warning is a misnomer; in fact, the goal is not to warn in advance of an earthquake, but rather to rapidly identify the earthquake source and warn users at some distance prior to the arrival of heavy shaking.

Presently, there are a number of countries around the world with an operational earthquake early warning system, including Japan, Mexico, Romania, and Taiwan (e.g. Allen et al., 2009). The United States is testing the ShakeAlert earthquake early warning system, aimed at notifying the West Coast of seismic activity (Kohler et al., 2017). Traditionally, earthquake early warning systems are designed using broadband seismometers and strong-motion accelerometers. These instruments are well suited to pick up the small, early seismic wave arrivals, contributing the the timeliness of a warning, and allow the identification of even small or distant earthquakes. However,

both instruments experience challenges within close distance of large earthquakes. Broadband seismometers have a limited dynamic range, rendering them unable to fully record large amplitude ground motions (e.g. Bock and Melgar, 2016). Strong motion seismometers are designed with lower gains, which allow these instruments to stay on-scale during heavy, large amplitude shaking. However, strong-motion accelerometers are unable to differentiate between translational and rotational motions (Trifunac and Todorovska, 2001), thus using the recorded accelerations to estimate velocity (single integration) or displacement (double integration) amplifies errors that grow unbounded with time (Boore et al., 2002). Removal of this unphysical drift in real-time is accomplished through high-pass filtering, which results in the loss of the record of any permanent displacements.

Earthquake magnitude is arguably the single most important metric for anticipating the severity of an earthquake's impact on public safety. For earthquakes offshore, magnitude is the main predictor variable for estimating tsunamigenic potential of an earthquake (e.g. Tatehata, 1997). Large magnitude megathrust subduction zone earthquakes can lead to significant economic and human losses through tsunamigenesis (Figure 1.1). Onshore, the magnitude, coupled with a detailed understanding of the regional geology, will provide a first order estimate of the intensity of expected ground motion (e.g. Boore et al., 2014). As a result of the seismic instrumentation limitations, the magnitude of large earthquakes is typically underestimated in real-time (e.g. Hoshiba and Ozaki, 2014). Geodetic estimates of slip, for example from Global Navigation Satellite System (GNSS) observations (Global Positioning System, GPS, and other navigation satellite constellations) provide a more accurate estimation of earthquake magnitude (Melgar et al., 2015a). To this day, GNSS datasets remain underutilized for early warning and rapid response.

Each earthquake observation instrument performs best for a particular set of conditions. For example, broadband seismometers have low noise floors and are therefore well designed for small or distant earthquakes, whereas GNSS observations are not precise enough to observe

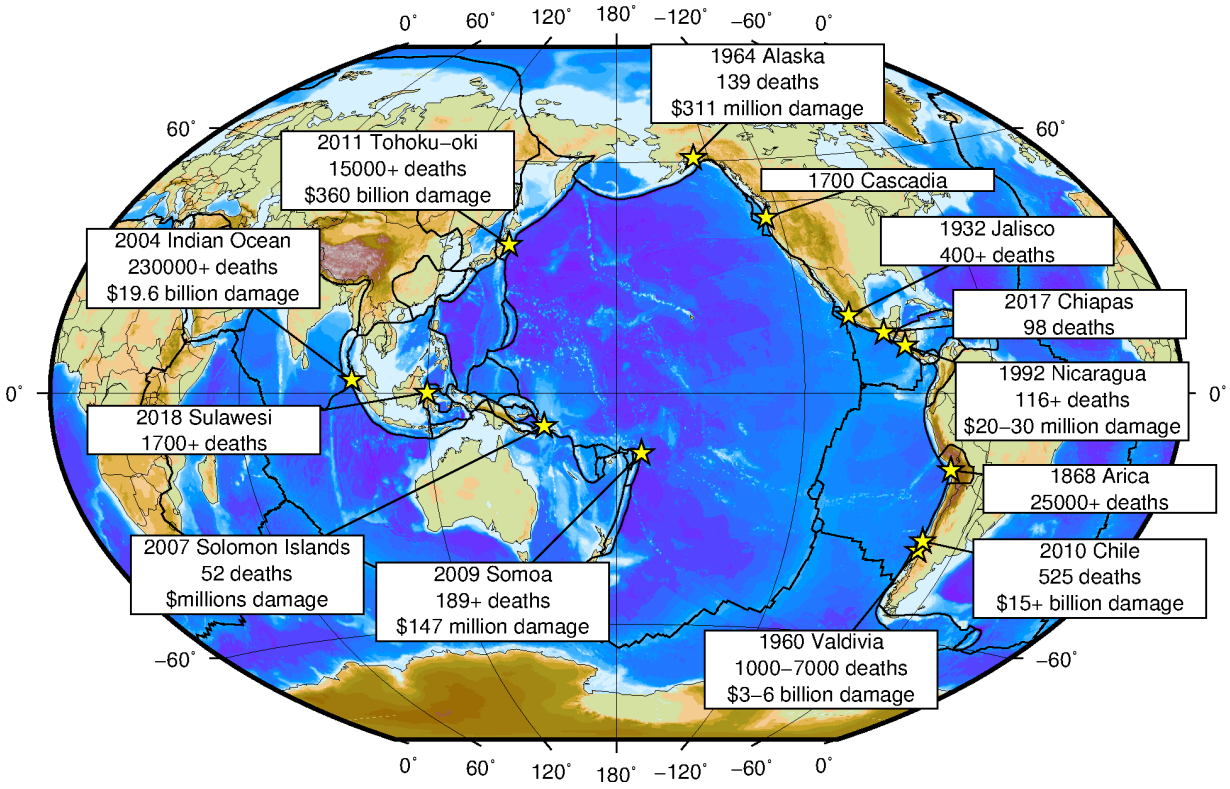


Figure 1.1: Historical tsunamigenic earthquakes around the Pacific Rim. The Cascadia Subduction Zone has not experienced a damaging tsunami since 1700, and is a critical region for implementation of an effective local tsunami warning system in the United States.

such events. Conversely, GNSS provides the most accurate estimation of slip from very large earthquakes, whereas broadband seismometers will exceed their dynamic range when subjected to heavy shaking (making observation of the full earthquake motions impossible) and strong-motion accelerometers will confuse rotations and translations, resulting in an inaccurate estimate of ground motions. **Seismogeodesy** is the optimal combination of collocated seismic and geodetic datasets such that the merged solution provides a truly broadband time series that captures the small initial seismic arrivals as well the largest amplitude signals, while retaining a record of any permanent displacement resulting from the earthquake (Bock et al., 2011). This method of course requires that seismic and geodetic (GNSS) instruments be collocated.

Historically, seismic and geodetic networks have evolved independently. Seismic systems, originally developed for earth science applications, were prioritized at coastlines and near active faults for earthquake monitoring. There are few limitations on the installation locations of seismic instruments, though a site is preferred to be away from any infrastructure that could introduce noise. GNSS networks were originally designed for the military and navigation. Once the application of GNSS for geodesy became clear, network design was then motivated by geophysical research priorities to observe tectonic plate motions. GNSS stations require a clear view of the satellite orbits, and therefore must be installed where there is no overhead obstruction. The utility of GNSS for seismology applications was later made possible with the development of high-rate and real-time continuous GNSS stations, illuminating the complementary nature of seismic and GNSS networks to observe geophysical processes. However, the disconnect between the historical development and location requirements unfortunately has resulted in a relatively sparse network of collocated seismic and GNSS instruments. The expansion of the collocated network is limited by the high cost of observatory-grade seismic instruments and GNSS stations. Recently, upgrades of existing GNSS stations with new, low-cost accelerometers has increased the seismogeodetic station density in the western United States (Saunders et al., 2016; Goldberg and Bock, 2017).

Throughout this dissertation, we describe the utility of the seismogeodetic approach for geophysical and engineering applications. As a direct outcome of this work, we are providing real-time seismogeodetic data streams from a number of collocated sites in the western United States to the National Oceanic and Atmospheric Administration (NOAA) Tsunami Warning Centers (TWCs) to improve current tsunami early warning capabilities. Testing and implementation of the algorithms and methods described in the following chapters within the NOAA tsunami early warning operational framework is ongoing. NOAA currently provides reliable basin-wide tsunami warnings, designed to alert far-field coastal regions of the timing and expected amplitude of tsunami inundation. However, providing accurate and timely warnings for the coastal region directly adjacent to a subduction zone remains challenging. Through this collaboration, the incorporation of seismogeodetic methods is expected to advance local tsunami forecasting capabilities.

1.2 Scientific Questions

This dissertation describes a multi-sensor approach to earthquake and tsunami early warning and structural health monitoring. The analyses herein expand the existing seismogeodetic dataset and knowledge base with new seismic and geodetic records and analysis methods. The objectives of the dissertation are:

1. Devise the basis of a fully seismogeodetic approach to earthquake and tsunami early warning systems,

2. Investigate the rupture nucleation process to address the open question of earthquake determinism,

3. Consider best practices for rapid estimation of earthquake magnitude and source parameters, and

4. Assess the capability of low-cost seismic instrumentation for earthquake early warning and rapid response.

1.3 Outline of the Dissertation

The dissertation is comprised of six chapters. Chapters two through six were written as stand-alone documents, and therefore will have some overlapping material. **Chapter 2** describes the algorithms which comprise the basis of a self-contained seismogeodetic earthquake early warning system. This work was published in the *Journal for Geophysical Research: Solid Earth* in March, 2017. The prototype system described therein is undergoing implementation and testing at the Scripps Orbit and Permanent Array Center (SOPAC) at Scripps Institution of Oceanography (SIO). **Chapter 3** begins an investigation into the open question of earthquake determinism, evaluating the use of seismogeodetically-derived displacement amplitude for rapid magnitude estimation. This chapter is in preparation for journal submission. **Chapter 4** delves deeper into deterministic rupture observations, investigating the rupture physics of large earthquakes to explain geodetic observations at the Earth's surface. This chapter has been submitted for publication to the *Journal for Geophysical Research: Solid Earth*, where it is currently in revision. **Chapter 5** explores the utility of the seismogeodetic system for structural health monitoring and earthquake rapid response. Concluding remarks and perspective for future work are provided in **Chapter 6**.

Chapter 2

Seismogeodetic Earthquake Early Warning

2.1 Background

Earthquake early warning (EEW) and rapid response systems, such as for local tsunami warnings (LTW), require near-field data to minimize the time between earthquake origin and issuance of a warning. Initial warnings are often comprised of the earthquake location, and an estimate of moment magnitude. The most critical parameters for LTW systems are earthquake magnitude and rupture mechanism. Current EEW and LTW systems rely on seismic instrumentation (broadband seismometers and strong-motion accelerometers) to make these estimates. For large magnitude earthquakes, ground motions can exceed the dynamic range of broadband seismometers (clipping) in the epicentral region where permanent deformation is present. Therefore, strong motion sensors are deployed at seismic stations, with lower gains, allowing the on-scale recording of ground accelerations, but these instruments cannot distinguish between rotational and translational motions (Trifunac and Todorovska, 2001), obscuring the true record (Boore and Bommer, 2005). Rapid estimation of earthquake magnitude from seismic instruments requires the double integration of acceleration waveforms to displacement waveforms, which amplifies small errors, known as baseline offsets, introducing unrealistic drift. The cause of these baseline

offsets remains largely unsolved, but many methods have been suggested to remove their effects (Iwan et al., 1985; Boore, 1999; Wu and Wu, 2007; Chao et al., 2010). The simplest method involves high-pass filtering the integrated waveforms (Boore and Bommer, 2005), but filtering results in the loss of the static (i.e. 0 Hz) offset, which is an impediment to rapid magnitude estimation. Moreover, the relationship between magnitude and ground displacement derived from seismic instrument recordings in the epicentral region saturates (e.g. Crowell et al., 2013), thus, discerning earthquake magnitude above a certain threshold is problematic (Brown et al., 2011; Hoshiba and Ozaki, 2014). More reliable estimates are available with teleseismic data, but an early warning cannot be effective with such methods because of the inherent latency. During the 2011 M_w9.0 Tohoku-oki earthquake, the Japan Meteorological Agency's EEW system, which relies on seismic instrumentation, estimated the event to be a M_w7.2 8.6 s after first detection, amending to a M_w8.1 after 116.8 s (Hoshiba and Ozaki, 2014). This was the upper limit of magnitude scaling in the near-source region, and thus an accurate magnitude was unavailable until the seismic waves were recorded at teleseismic distances. After the fact, Colombelli et al. (2012) used seismic instrumentation to estimate the magnitude, achieving M_w=8.4 after 35 s. However, the true M_w9.0 is still eight times larger than M_w8.4 in terms of energy release, so while this result represents an improvement, it is still a significant underestimation.

GNSS observations are typically used to measure static displacements, but, high-rate GNSS observations have also been used to estimate dynamic displacements (Nikolaidis et al., 2001; Larson et al., 2003; Bock et al., 2004; Langbein and Bock, 2004; Genrich and Bock, 2006; Larson, 2009), and have been integrated into earthquake early warning methods (Crowell et al., 2009; Grapenthin et al., 2014; Crowell et al., 2016). GNSS is especially well-suited in the near-field of large earthquakes because arbitrarily large displacements can be recorded without clipping (Bock et al., 2004). Furthermore, magnitude saturation for large earthquakes is avoided when using displacement directly, without filtering (Crowell et al., 2013; Melgar et al., 2015a). Instantaneous GNSS-derived displacements have a precision of 5-10 mm in the horizontal and

10-20 mm in the vertical component (Genrich and Bock, 2006). The vertical is the direction most applicable to P-wave detection since compressive motion is parallel to the propagation direction. As a result, GNSS estimates are not precise enough to detect P-wave arrivals, even at near-fault distances for large earthquakes. Larger displacements from subsequent arrivals (e.g. S waves) which exceed the typical instrument noise could be used for event detection, but this reduces the effectiveness of earthquake early warning because of the added latency. Furthermore, the precision of GNSS positions can be affected by a number of factors including the available satellite coverage, atmospheric delays, and multipath signals.

It is possible to remove much of the GNSS error contributed by multipath signals with sidereal filtering, taking advantage of the repeating satellite orbits to remove signals that are consistent over repeating sidereal days (Bock, 1991; Genrich and Bock, 1992; Choi et al., 2004; Agnew and Larson, 2007). Such filtering may help to reduce unreliable signals that could cause false detections using only GNSS data during quiescent periods. However, reducing multipath effects, which is challenging in a real-time environment, is still not sufficient to allow P-wave detection, compared to accelerometer or velocity waveforms. The noise level of instantaneous GNSS positioning is reduced during seismic shaking relative to the static periods (Genrich and Bock, 2006; Saunders et al., 2016). Bock et al. (2004) postulated that accelerations experienced by the geodetic monument temporarily randomize the multipath effects, which explains the improvement in precision. This is supported by the study of Pesyna Jr et al. (2014) of GPS positioning with smartphones. The antenna's poor multipath suppression and irregular gain pattern result in large time-correlated phase errors. However, when the smart phone experiences gentle wavelength-scale random motion, the time to integer-cycle phase ambiguity resolution is significantly reduced. Once phase ambiguities are resolved then centimeter-level precision can be achieved. In any case, seismic observations, whose resolution is far better than GNSS, are better suited to earthquake detection and early warning.

It is clear, then, why current earthquake early warning systems have developed to rely

strictly on seismic data. However, because of the aforementioned problems, seismic stations require both broadband instruments and accelerometers to record the full velocity or acceleration waveforms. These waveforms are then either singly- or doubly-integrated to dynamic displacements; as indicated, recovering the static displacement is problematic. In using broadband instruments and strong-motion accelerometers, there is a reliance on two non-overlapping types of observations to construct the full spectrum of dynamic displacement; that is, one source of information with no degrees of freedom. The optimal combination of the GNSS and strong-motion accelerometer data with a Kalman filter (e.g. Smyth and Wu, 2007) produces broadband displacement and velocity waveforms that span the full spectrum of displacement, both dynamic and static. In essence, the seismogeodetic combination provides a broadband seismometer that is immune to clipping, has the temporal resolution of a strong-motion accelerometer, and retains the permanent (static) displacement (Bock et al., 2011; Geng et al., 2013; Saunders et al., 2016). The knowledge that there is no static offset, for example, could be used to minimize false alarms. Furthermore, throughout observed motion, two independent sources of information (one degree of freedom) are being used: accelerations and GNSS observations. Since the seismogeodetic displacement solution matches the GNSS-only solution at long periods, it provides a smooth and self-contained transition to downstream products (Figure 2.1). Thus, seismogeodesy provides an effective, practical, and straightforward implementation of earthquake and local tsunami early warning. Before seismic shaking dissipates, an estimate of magnitude can be obtained based on peak ground displacement scaling relationships (Melgar et al., 2015a). The seismogeodetic solution has shown promise when using P-wave amplitude to scale for magnitude, without saturation experienced with seismic-only recordings (Crowell et al., 2013). The seismogeodetic approach helps to rapidly identify the coseismic offset when the seismogeodetic velocity falls below a given threshold value. Once the permanent displacements are estimated, they can be input to centroid moment tensor solutions (Melgar et al., 2012), finite fault slip models (Crowell et al., 2012; Melgar et al., 2013) and predictions of tsunami run-up and inundation (Melgar and

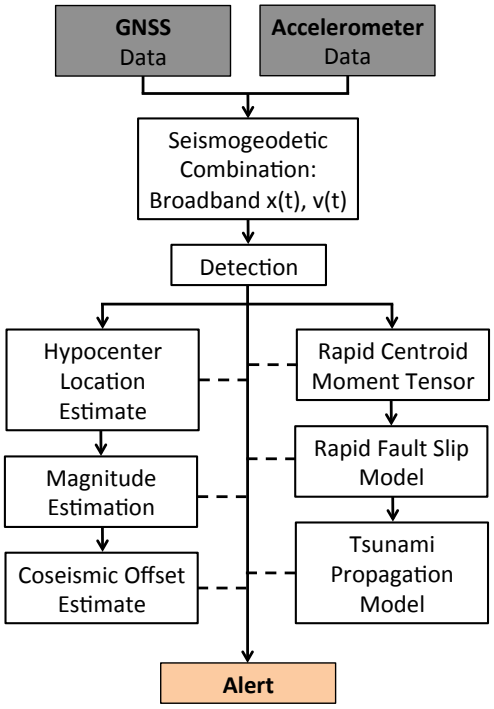


Figure 2.1: Early warning products undergoing implementation in SOPAC's early warning system. Data from collocated GNSS and accelerometer stations are combined in real time, resulting in broadband displacement, $x(t)$, and velocity, $v(t)$. Detection and hypocenter location are discussed here. Magnitude scaling (Crowell et al., 2013; Melgar et al., 2015a), centroid moment tensor inversion (Melgar et al., 2012, 2013), fault slip inversion (Crowell et al., 2012; Melgar et al., 2013) and tsunami propagation modeling (Melgar and Bock, 2013, 2015) as they apply to our system are described in detail in the cited works. The system as designed can provide a hierarchy of alerts from S-wave arrivals to tsunami run-up and inundation.

Bock, 2013, 2015).

The implementation of real-time seismogeodetic analysis requires collocated seismic and geodetic instrumentation, ideally within meters, but at least within 4 km (Emore et al., 2007; Melgar et al., 2015b). Unfortunately, the independent development of seismic and geodetic networks has resulted in a relatively sparse set of collocated stations. In this chapter, we discuss the capabilities of a self-contained seismogeodetic approach to earthquake early warning including event detection and location. We describe the current implementation of our system (Figure 2.1) operated by the Scripps Orbit and Permanent Array Center (SOPAC) at the Scripps Institution of Oceanography at University of California, San Diego. Currently, several real-time seismogeodetic

networks span major faults in the western United States (Figure 2.2), including 41 stations in Southern California (San Andreas, San Jacinto, and Elsinore faults), 19 stations in the San Francisco Bay Area (Hayward and Rodger's Creek faults), and 40 stations in Oregon and Washington (Cascadia Subduction Zone). Collocated stations also exist in western Chile and Japan. We focus on three medium to large earthquakes that occurred in regions of collocated station coverage: two primarily strike-slip events, the 2014 M_w6.0 Napa, California earthquake and the 2010 M_w7.2 El Mayor-Cucapah, Mexico earthquake; and the 2014 M_w8.2 Iquique, Chile thrust earthquake. We perform retrospective analysis of these three events using the algorithms currently implemented in real time at SOPAC. The collocated instruments in each of these three examples include observatory-grade accelerometers.

We use these examples as a proof of concept to test our seismogeodetic EEW approach. Our objective is to be able to apply these methods using inexpensive MEMS accelerometer upgrades to existing GNSS stations. To that end, we demonstrate the seismogeodetic EEW approach using two recent, smaller magnitude earthquakes in California observed by GPS and MEMS accelerometers: the 2014 M_w4.0 Piedmont and 2016 M_w5.2 Borrego Springs earthquakes. Successful proof of concept of the seismogeodetic system supports the upgrade of existing single-instrument stations to seismogeodetic capability in order to improve current early warning methods.

2.2 Event Detection

The detection requirements for EEW depend on the tectonic and geographic settings. The Mexican EEW system, for example, is designed to detect seismic arrivals near the fault and transmit a warning ahead of the propagating S- and surface waves to highly populated Mexico City (e.g. PérezCampos et al., 2013). In the Western U.S., fault systems capable of sustaining large magnitude events often coincide with large cities. Thus, rapid detection using the first seismic

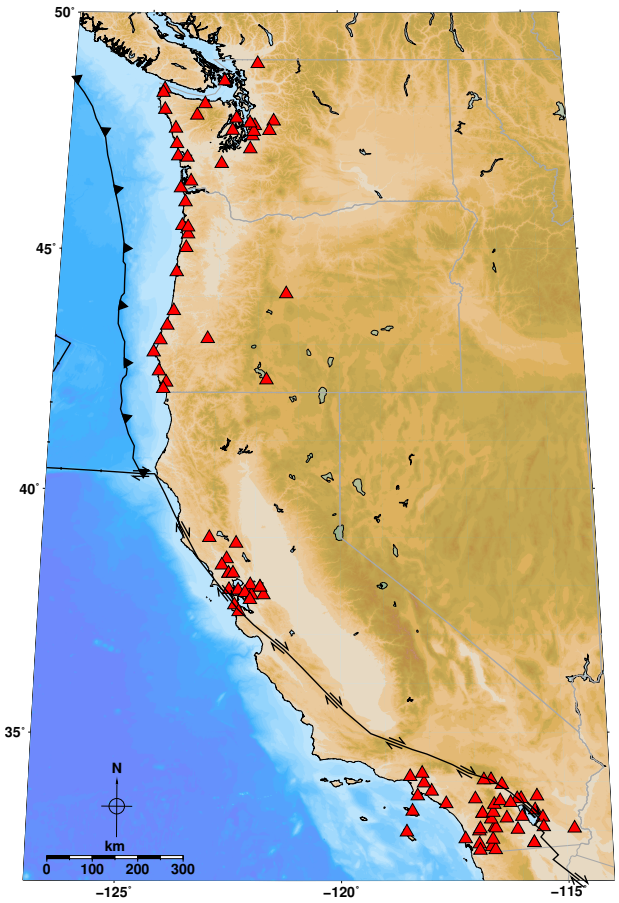


Figure 2.2: Collocated GPS and accelerometer station coverage (red triangles) in regions of high seismic hazard in the western United States. Major plate boundaries are shown in black. The southern California network includes stations from the Southern California Integrated GPS Network (SCIGN), the Plate Boundary Observatory (PBO) and the Southern California Seismic Network (SCSN). The San Francisco Bay Area stations are part of PBO and the Northern California Seismic Network (NCSN). The Cascadia stations include PBO, the Pacific Northwest Geodetic Array (PANGA), and the Pacific Northwest Seismic Network (PNSN).

phase arrivals (P-waves) is imperative for the timeliness of EEW. Many methods for automation of P-wave detection are used. The Quake-Catcher Network (QCN) compares current acceleration to the average acceleration from the previous minute of data, and defines a trigger when the current acceleration deviates from the average by more than three standard deviations (Cochran et al., 2009). An event is defined if both the number of triggers in a geographic region exceeds six standard deviations from the average rate of triggers and those triggers can be explained by circular spreading of seismic waves. The community seismic network of southern California employs the more common short time average- long time average (STA/LTA) method (Allen, 1982; Earle and Shearer, 1994; Clayton et al., 2012), in which a detection is defined when the ratio of average acceleration in a short window to average acceleration in a longer window exceeds a previously determined threshold. An event is declared when the number of detections exceeds a certain number (Clayton et al., 2012). The STA/LTA method is also employed by ElarmS (Kuyuk et al., 2013), one of the main algorithms of ShakeAlert, an EEW system in development in the Western U.S. The capabilities of smartphone-based EEW networks are rapidly improving. Kong et al. (2016) employ the STA/LTA method for the smartphone application MyShake, however the built-in smartphone accelerometers are expected to trigger on larger amplitude shaking behind the initial P-wave arrival. Although the instrument quality required for each of these networks varies (QCN and MyShake use low-cost MEMS accelerometers while ElarmS relies on observatory-grade instruments), each is based on seismic-only instrumentation.

GNSS alone is too noisy to accurately detect the first seismic wave arrivals, hampering its use as stand-alone instrumentation for earthquake early warning. We demonstrate that when GNSS is used in combination with accelerometers, the resulting seismogeodetic velocity solution is capable of detection of P-wave arrivals for earthquakes with magnitudes of interest to early warning, comparable to that of seismic-only instrumentation. By using the combination dataset for event detection, we can simplify our EEW system by using a single source of data for all early warning products including detection, magnitude estimation and rapid source modeling, as well

as LTW modeling and prediction. We identify collocated seismic and GNSS stations (within 2.5 km, based on earlier studies (e.g. Emore et al., 2007; Melgar et al., 2015b), in the vicinity of historical earthquakes with magnitudes of interest to early warning with which to test early warning algorithms. Distances between the instruments used in a single collocation can be found in the supporting information (Section 2.6). From these collocations, we produce seismogeodetic velocity and displacement waveforms retrospectively, using the Kalman filter described in Bock et al. (2011).

2.2.1 Method

We apply the STA/LTA method (Allen, 1982) to automate identification of P-waves at seismogeodetic stations using the velocity waveforms from our earthquake dataset. We calculate the short and long averages using a recursive mean to eliminate the need to retain many seconds worth of data, easing the computational requirements on our real-time system. The relationships are defined as

$$\begin{aligned} \mathrm{STA}_x &= \left(\frac{n_{\mathrm{STA}}-1}{n_{\mathrm{STA}}}\right)\mathrm{STA}_{x-1} + \left(\frac{1}{n_{\mathrm{STA}}}\right)v_x \\ \mathrm{LTA}_x &= \left(\frac{n_{\mathrm{LTA}}-1}{n_{\mathrm{LTA}}}\right)\mathrm{LTA}_{x-1} + \left(\frac{1}{n_{\mathrm{LTA}}}\right)v_x \end{aligned} \tag{2.1}$$

where v_x is the seismogeodetic velocity of the current epoch, n_{STA} is the length of the short time window and n_{LTA} is the length of the long time window.

Velocities are bandpass filtered between 1 and 3 Hz, which helps expose the P-wave arrival. Data transmission processes impose additional constraints on real-time implementation. Data arrives in packets of 1 s regardless of sampling rate (1-100 Hz), thus our window lengths must be simple integer seconds. Once the algorithm determines that the trigger threshold has been reached, it reinvestigates the most recent second-long packet to determine more precisely

the P-wave arrival time (to the precision of the accelerometer sampling rate).

Typically, regional systems set a short window between 1 and 2 seconds. The long window can be more than ten times that length (Trnkoczy, 2012). Our choices of window length and threshold were determined by testing reasonable combinations (e.g. STA = 1s, 2s, 3s; LTA = STA+1s, STA+2s...; and threshold = 1.5, 1.6... 5.0) and finding the combination that minimized the difference between manual and automated picks for the available data. For our purposes, we impose the same parameters for the STA/LTA algorithm at every station. In practice, the algorithm could be implemented independently at individual stations, thus window lengths and the detection threshold could be adjusted to reflect an individual station's noise characteristics, although we have not done so here.

Detections must be corroborated by nearby stations in order to trigger the succession of additional early warning products (Figure 2.1). We therefore compare arrival times at nearby stations and calculate whether each detection may be due to the same event based on inter-station distances and a conservative estimate of P-wave velocity. In the current implementation, the apparent P-wave velocity is assumed to be 5.5 km/s, such that a detection at station 2, X km from the first triggered station, must occur within X/5.5 s to signify a corroborated detection. The subsequent algorithms (e.g., event location, centroid moment tensor solution) depend on the availability of arrival times at at least four stations, and we assume that an event large enough to be of interest to early warning will be observed by at least four stations in the network, given the typical inter-station spacing of $\sim$20 km. Thus, if fewer than four station detections occur within the limits prescribed by their distance, the detections are discarded and the triggering algorithm is reset.

2.2.2 Results and Accuracy

Using observatory-grade accelerometers, it is easy to manually select the P-wave arrival from near-source monitoring stations. We analyze the reliability of our real-time seismogeodetic

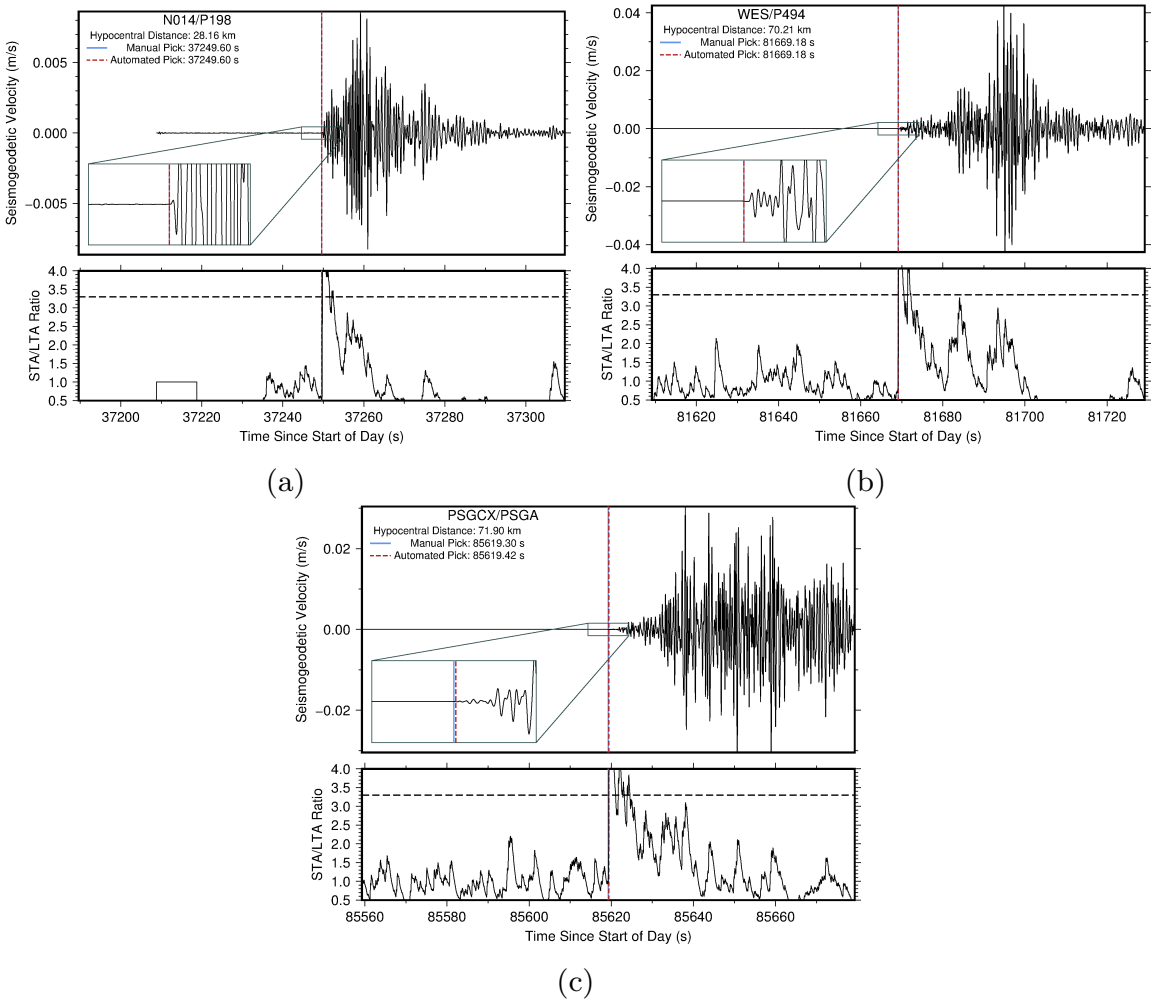


Figure 2.3: Automated detection algorithm example at nearest seismogeodetic stations for medium to large magnitude events (a) 2014 M_w6.0 South Napa, California: N014/P198, (b) 2010 M_w7.2 El Mayor-Cucapah, Mexico: WES/P494, and (c) 2014 M_w8.2 Iquique, Chile: PSGCX/PSGA. The name of the stations are according to accelerometer station code/GPS four-character station code. Top panel shows bandpassed seismogeodetic velocity with manual pick (blue) and automated pick (dashed red), with an inset highlighting the P-wave arrival. Bottom panel shows the STA/LTA ratio, with horizontal dashed line indicating the detection threshold. In (a), the top panel shows the start of available data, at which point the algorithm is initiated. The bottom panel shows a boxcar shape during algorithm initialization. Eight of the stations used in analysis of the Napa event were operated in triggered mode, which accounts for the timing of the earliest data availability (see supplementary material). In all cases, algorithm initialization is complete prior to the P-wave arrival.

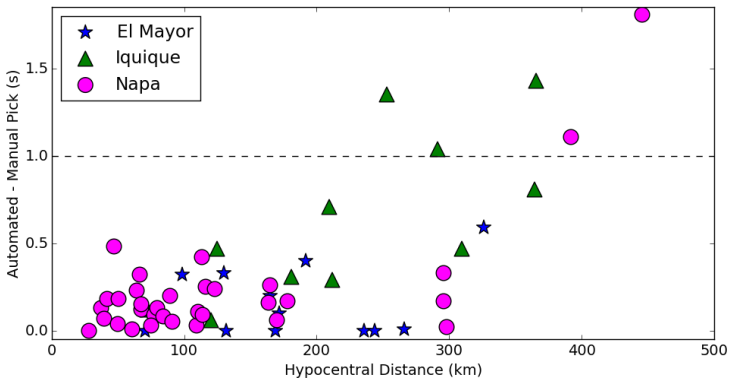


Figure 2.4: Automated P-wave pick error compared to hypocentral distance. Above the dotted line, the difference between manual and automated pick exceeds 1 second.

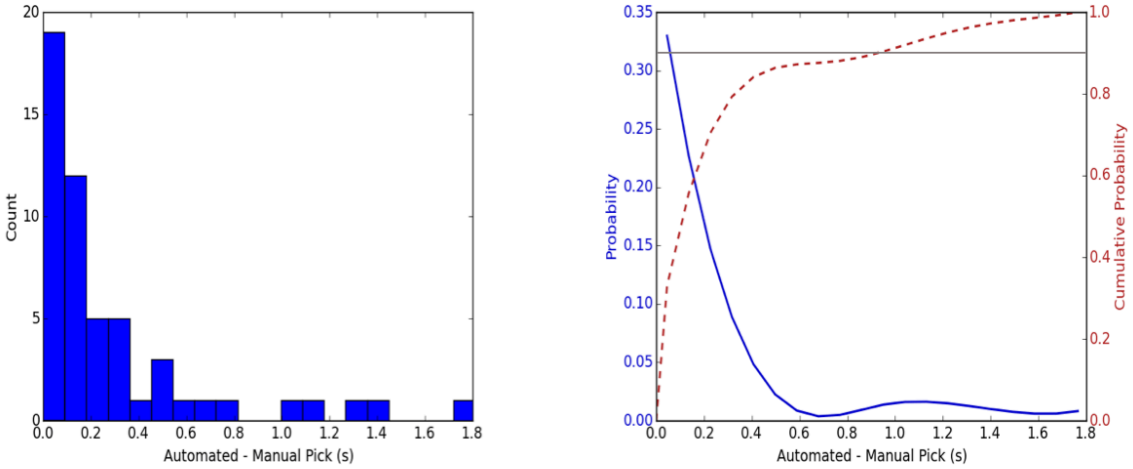


Figure 2.5: Automated P-wave detection error statistics: (a) histogram of difference between manual and automated picks including 56 stations from the 3 earthquakes (Napa: 33 stations, El Mayor-Cucapah: 12 stations, Iquique: 11 stations), and (b) probability density function created by spline interpolation and normalization of histogram (blue) and cumulative density function (dashed red). 90% probability is denoted by solid grey line.

velocity P-wave picks with respect to manual P-wave picks from observatory-grade acceleration waveforms. This comparison is illustrated for the 2014 M_w6.0 South Napa, 2010 M_w7.2 El Mayor-Cucapah, and 2014 M_w8.2 Iquique earthquakes in Figure 2.3. The complete set of picked waveforms are provided in the supporting information (Section 2.6). From these 3 events and 56 total stations, we have determined that a short time window of 1 s and a long time window of 5 s is ideal for P-wave detection on seismogeodetic velocity time series. We find that a detection threshold of 3.3 for the ratio STA_x/LTA_x is effective for the window lengths selected. Ideally, the ratio should be close to 1.0 during periods of quiescence, and because we have chosen windows with a small difference in length compared to the literature (Trnkoczy, 2012), a comparatively low detection threshold is appropriate. Station MNMCX, $\sim$136 km from the M_w8.2 Iquique hypocenter, recorded errant signals in the accelerometer, causing it to trigger often (see Figure S2.9a). This station was therefore removed from subsequent analyses. The algorithm, with a single set of STA/LTA parameters, is effective up to at least $\sim$ 300 km away from the source of a relatively shallow M_w6.0 earthquake (as is typical in California), using the 2014 South Napa earthquake as a test case (seismogeodetic station HATC/HCRO had a P-wave pick error of 0.02 s at a hypocentral distance of 298 km, see Figure 2.4). The algorithm's performance with smaller events and/or more distant stations will become clearer as new events occur and their data become available. The limited historical earthquake dataset recorded by seismogeodetic stations impedes our ability to characterize the extent more fully. The performance of the algorithm with GPS and inexpensive MEMS accelerometers will be addressed in Section 2.3.4 and Section 2.4.

We use the differences between the manual and automated picks for the three earthquakes to construct probability and cumulative density functions of the errors (Figure 2.5), which will be used later for evaluating the precision of our hypocenter estimates (Section 2.3). The mean difference between manual and automated picks for the three earthquakes is 0.30 s with a median of 0.16 s and a variance of 0.15 s^2. Ninety percent of the picks are within 1.0 s (Figure 2.5b). The skewness of the error distribution is 2.13, indicating that the left tail of the distribution is more

heavily populated. The excess kurtosis is 1.17, implying that much of the variance comes from outliers (i.e. values in the right tail). From Figure 2.4, we see that the furthest stations also tend to have some of the largest detection errors. This is likely due to seismic wave attenuation that leaves the P-wave arrival obscured by station noise. Picks at farther stations may also trigger on later, larger seismic wave arrivals, e.g. S-waves. The nearest stations are generally most reliable and most important, since those will be used for the earliest dissemination of warning. The discussion of our hypocenter location algorithm (Section 2.3) will further assess the errors in our P-wave picks.

2.3 Event Location

Hypocenter location is a complex problem in real time, as location has a nonlinear dependence on seismic travel times and can be obscured by a poorly constrained velocity structure. Using only stations in the epicentral region, it is difficult to constrain the depth of the earthquake, especially if the event is outside the network (Lee and Stewart, 2981; Lomax et al., 2009), because the station arrival times for a shallow earthquake will look virtually identical to a deeper earthquake with earlier origin time. Later reflected phases are useful for constraining the earthquake depth (Billings et al., 1994), but reliable automated selection of the later-phase onset from within the signal of previously occurring phases is difficult. Thus, our algorithm relies on P-wave arrivals, a side effect of which is poor sensitivity to changes in hypocenter depth, since a change in depth has little effect on the ray path length of direct P-waves. The assumed velocity structure introduces more error to the depth estimate than can be resolved by the P-phase (Billings et al., 1994). For this reason, many earthquake catalogs assign a default ("fixed") depth (e.g. USGS-FAQs, 2015). Similarly, we include a variety of fixed-depth options based upon the tectonic region to more accurately determine each station's hypocentral distance, a required parameter for rapid magnitude scaling (Crowell et al., 2013; Melgar et al., 2015a).

Many of the most destructive earthquakes are associated with subduction zone events. These deeper events (e.g. the 2014 M_w8.2 Iquique earthquake at 25 km depth) require consideration of the variability of seismic wave speed through the layered Earth. Thus, in all cases, we account for a 1D crustal velocity structure (Laske et al., 2013) that is automatically selected at the location of the first station to detect an event, since that station is the best estimate of the hypocenter without any additional information. We choose to implement a 1D velocity structure for our real-time algorithm, rather than a 3D velocity structure due to uncertainty in 3D velocity models and extra computing power necessary for a rapid 3D implementation (Lienert et al., 1986; Wald et al., 1995). A hypocenter can be computed with arrival times from a minimum of 4 stations, and is recalculated and updated with each subsequent station arrival. Accurate and rapid hypocenter determination is critical to forecasting areas at greatest risk (i.e. in the epicentral region). We discuss later in this section how the accuracy affects downstream warning products.

2.3.1 Method

The location estimation algorithm implemented at SOPAC is based on Geiger's method of earthquake location (Geiger, 1912) and the automatic P-wave arrival times (Section 3). Briefly, this method involves making an initial guess of the earthquake location parameters $\boldsymbol{x} = [x_0, y_0, z_0, \tau_0]^T$, where $[x_0, y_0]$ is the epicenter, z_0 is the depth and τ_0 is the origin time, then calculating the arrival times expected at each station from an event at that hypocenter and comparing to observed arrivals. For the small area of near-field stations, we use a local Cartesian reference frame, in which the current hypocenter estimate represents the origin of the coordinate system, x is positive East, y is positive North, and z is positive up. Cartesian coordinates are converted back into geographic coordinates (x and y to longitude and latitude, respectively) at the end of the calculation. In our operational system, the initial guess has an epicenter at the location of the first triggered station, with fixed depth chosen based on the tectonic environment, i.e. average depth of the seismogenic zone. We define the residuals, $\boldsymbol{r}$ between the calculated and observed arrival times (calculated

arrival time is the sum of the calculated travel time, $\boldsymbol{t}^{calc}$, and origin time, τ):

$$\boldsymbol{r} = \boldsymbol{t}^{\text{obs}} - (\boldsymbol{t}^{\text{calc}} + \boldsymbol{\tau}) \quad . \tag{2.2}$$

The arrival time function, $\boldsymbol{t}$, at the i^{th} station due to an earthquake at hypocenter location $\boldsymbol{x}$ is equal to the arrival time due to an earthquake at the location of the initial guess, $\boldsymbol{x}_0$, plus a location improvement, $\boldsymbol{\delta x}$ where $\boldsymbol{x} = \boldsymbol{x}_0 + \delta\boldsymbol{x}$ and

$$\boldsymbol{x}_0 = [x_0, y_0, z_0, \tau_0]^T$$
$$\boldsymbol{\delta x} = [\delta x, \delta y, \delta z, \delta\tau]^T \quad .$$

By the first degree Taylor expansion,

$$t_i(\boldsymbol{x}) = t_i(x_0 + \delta x) = t_i(x_0) + \frac{\partial t_i}{\partial x}\delta x + \frac{\partial t_i}{\partial y}\delta y + \frac{\partial t_i}{\partial z}\delta z + \frac{\partial t_i}{\partial \tau}\delta\tau \quad . \tag{2.3}$$

The last four terms express the location improvement, a function of the partial derivatives of the hypocenter parameters. We assume that this improvement is the requirement to correct for the difference between the calculated and observed arrival times (Geiger, 1912; Havskov and Ottemöller, 2010). From Equation 2.2:

$$r_i = \frac{\partial t_i}{\partial x}\delta x + \frac{\partial t_i}{\partial y}\delta y + \frac{\partial t_i}{\partial z}\delta z + \frac{\partial t_i}{\partial \tau}\delta\tau \quad . \tag{2.4}$$

In Matrix Notation,

$$\boldsymbol{r} = \boldsymbol{A}\boldsymbol{\delta x} \tag{2.5}$$

where:

$$\boldsymbol{r} = \begin{bmatrix} r_1 \\ \vdots \\ r_n \end{bmatrix} \quad \boldsymbol{A} = \begin{bmatrix} \frac{\partial t_1}{\partial x} & \frac{\partial t_1}{\partial y} & \frac{\partial t_1}{\partial z} & \frac{\partial t_1}{\partial \tau} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial t_n}{\partial x} & \frac{\partial t_n}{\partial y} & \frac{\partial t_n}{\partial z} & \frac{\partial t_n}{\partial \tau} \end{bmatrix} \quad \boldsymbol{\delta x} = \begin{bmatrix} \delta x \\ \delta y \\ \delta z \\ \delta \tau \end{bmatrix} \quad .$$

The improvement factor, $\boldsymbol{\delta x}$ is added to the initial hypocenter guess, $\boldsymbol{x}_0$, to get an updated hypocenter estimate, $\boldsymbol{x}$, and the process is repeated until the residual between calculated and observed arrival times converges to a predefined tolerance. Because we impose a fixed depth based upon the tectonic setting, the estimation is simplified by constraining $\delta z = 0$. In our operational system, iteration is complete when the L2-norm of the residuals changes by less than a predetermined (user-defined) threshold, meaning that iteration will conclude when the update to the hypocenter is small.

To indicate that the observations are subject to errors that are fully described by the first-order (mean) and second-order (variance) moments, we rewrite (2.5) as

$$\boldsymbol{r} = \boldsymbol{A}\boldsymbol{\delta x} + \boldsymbol{\varepsilon}; \qquad E\{\boldsymbol{\varepsilon}\} = 0; \qquad D\{\boldsymbol{\varepsilon}\} = \sigma_0^2 C_\varepsilon \tag{2.6}$$

where E denotes statistical expectation, D denotes statistical dispersion, C_ε is the covariance matrix of observation errors, and σ_0^2 is an *a priori* variance factor. If we assume that the observations are uncorrelated in space and time, the covariance matrix C_ε is diagonal. If we assign a weight matrix, $\boldsymbol{W} = C_\varepsilon^{-1}$, the weighted least squares solution for the inversion of (2.6) is derived by minimizing the Euclidian (L2) norm of the residual vector such that

$$\min||\boldsymbol{\varepsilon}^T \boldsymbol{W} \boldsymbol{\varepsilon}|| = \min||(\boldsymbol{r} - \boldsymbol{A}\boldsymbol{\delta x})^T \boldsymbol{W} (\boldsymbol{r} - \boldsymbol{A}\boldsymbol{\delta x})|| \tag{2.7}$$

with the weighted least squares solution $\widehat{\boldsymbol{x}}$ and the estimated covariance matrix $\widehat{\Sigma}_{\widehat{x}}$ given, respec-

tively, by:

$$\widehat{\boldsymbol{x}} = \left(\boldsymbol{A}^T \boldsymbol{W} \boldsymbol{A}\right)^{-1} \boldsymbol{A}^T \boldsymbol{W} \boldsymbol{r} \tag{2.8}$$

$$\widehat{\Sigma}_{\widehat{x}} = \widehat{\sigma}_0^2 (\boldsymbol{A}^T \boldsymbol{W} \boldsymbol{A})^{-1}; \qquad \widehat{\sigma}_0^2 = \frac{\widehat{\boldsymbol{\varepsilon}}^T \boldsymbol{W} \widehat{\boldsymbol{\varepsilon}}}{n-u} \tag{2.9}$$

The degrees of freedom, $n-u$, is given by the number of observations, n, minus the number of estimated parameters, u. The hat denotes an estimated quantity. In practice, we implement a simple weighting scheme to favor nearer stations, since nearer stations tend to have the largest, most impulsive, first arrivals. This is especially important for smaller events, where first arrivals at further stations may not be noticeable above the typical noise level, causing stations to trigger on larger, later arrivals, e.g. S-waves. Stations with later arrival times (i.e. greater hypocentral distance) receive less weight in the estimation:

$$\boldsymbol{W} = \text{diag}\left(\frac{1}{(t_i - (t_0 - 1))^2}\right) \tag{2.10}$$

where "diag" represents a square matrix with the input arguments on the diagonal. The denominator in the weighting scheme is $(t_i - (t_0 - 1))^2$ rather than $(t_i - t_0)^2$ to avoid possible division by zero. Of course, by setting $\boldsymbol{W} = \boldsymbol{I}$, this reduces to the ordinary least squares solution. We will examine both the ordinary and weighted estimates and denote them by L2 and WL2, respectively. It can be shown that

$$E(\widehat{\boldsymbol{\varepsilon}}^T \boldsymbol{W} \widehat{\boldsymbol{\varepsilon}}) = (n-u)\widehat{\sigma}_0^2 \tag{2.11}$$

(Hamilton, 1964) so that an unbiased estimate of σ_0^2, the *a posteriori* variance factor, often called the "chi-squared per degrees of freedom" or simply "chi-squared", is given above as Equation 2.9. This metric provides an estimate of the goodness of fit of the model to the data. We will use the goodness of fit estimate for assessing the quality of the input observations $\boldsymbol{r}$ for each of the earthquakes studied.

The formal error ellipse for the epicenter estimate is obtained from the upper 2x2 portion of the covariance matrix $\widehat{\mathbf{\Sigma}}_{\hat{x}}$ representing the longitude (x) and latitude (y) (Havskov and Ottemöller, 2010). The orientation and dimensions of the error ellipse is determined by singular value decomposition.

The L1 norm minimization may be useful because it effectively dampens the influence of outliers on the resulting solution (Shearer, 1997). However, there is no simple analytical solution. Rather, L1 optimization is obtained by sweeping through possible solutions. We use software package CVXOPT (cvxopt.org) for convex optimization to find the solution with minimum error. We use the following for the L1 norm metric to assess goodness of fit:

$$\gamma = \frac{\Sigma_1^n |\widehat{r}|}{(n-u)\sigma} \tag{2.12}$$

assuming that the observations are equally weighted with a standard deviation of σ.

The formal error ellipse for the L2-norm (Equation 2.9) is not easily applied to the L1-norm. To measure and compare the accuracy and precision of different solvers, we use a bootstrapping method following Shearer (1997). The automated arrival times are randomly adjusted using the error distribution described in Figure 2.5b, and the event is relocated 500 times to estimate the probable scatter of the resulting hypocenter locations due to P-wave pick uncertainties. We take the mean of the bootstrapped results as the hypocenter estimate, with error ellipse defined by the spread of the solutions. The 95% confidence ellipse that describes the bootstrapping results are summarized under "Error Ellipse" in tables 2.1, 2.2, and 2.3.

We explore all three solution types (L1, L2, and WL2 norm minimization) in Section 2.3.2. For L2 and WL2 we can use the chi-squared metric to characterize the quality of observations as newly triggered stations are added in the epicenter estimation. That is, we may be able to discern when a new observation contributes to a poor epicenter fit, and develop criteria to justify elimination of such stations from the inversion. In Section 2.3.2, we simply state the value of

goodness of fit associated with the final hypocenter results from each solution type using all available stations. In Section 2.4, we describe in greater depth the potential utility of these criteria using the example of the M_w5.2 Borrego Springs earthquake, for which the quality of P-wave picks at some stations is limited due to the comparatively small event magnitude and the use of MEMS accelerometers.

Finally, to further convey the accuracy of the location estimate, we calculate the azimuthal gap, a common metric for expressing confidence in an epicenter estimate. The azimuthal gap is defined as the largest distance (in degrees) between azimuthally adjacent stations, i.e. the largest angle formed between a station, the epicenter estimate (vertex), and a second azimuthally adjacent station. Generally, the smaller the azimuthal gap, the more reliable the estimate is assumed to be, as a small gap indicates high azimuthal station coverage. An azimuthal gap larger than 180° denotes one-sided station coverage.

2.3.2 Results and Accuracy

Table 2.1: Hypocenter location: 2014 M_w6.0 South Napa, California, USA. Results of location algorithm for different solver types: L1, L2, and WL2. Depth is assumed to be 10 km for crustal California. The USGS published hypocenter is also listed for comparison. The last row shows the results from simple PGD magnitude scaling relations using the seismogeodetic displacement records from the 33 available stations and the listed hypocenter as input (Melgar et al., 2015a).

Solver		**USGS**	**L1**	**L2**	**WL2**
Longitude (°E)		-122.31	-122.29	-122.28	-122.32
Latitude (°N)		38.22	38.23	38.22	38.21
Origin Time (UTC)		10:20:44	10:20:43.45	10:20:43.64	10:20:43.75
Depth (km)		10	10	10	10
Azimuthal Gap (°)		—	72.10	72.36	70.83
Goodness of Fit		—	0.03	0.07	0.03
Error Ellipse (95%)	Azimuth (°)	—	232.9	260.1	253.4
	Major (km)	—	4.02	2.66	3.75
	Minor (km)	—	2.22	1.10	2.26
	± Origin Time	—	0.32	0.15	0.31
Distance from USGS (km)		—	1.84	2.63	1.55
PGD Magnitude		5.82	5.82	5.82	5.82

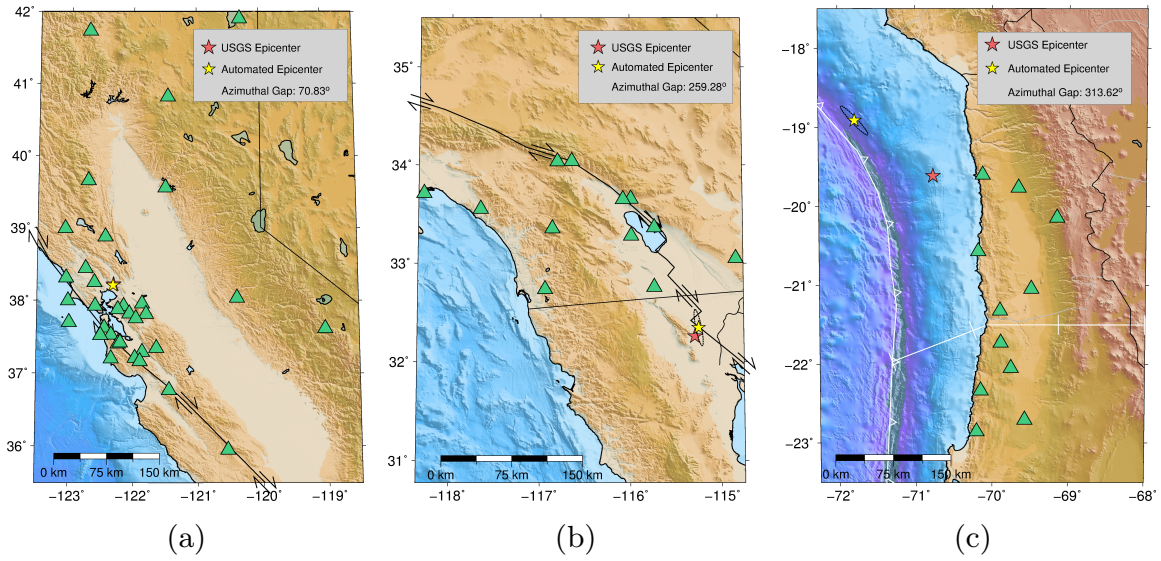


Figure 2.6: Epicenter estimate and 95% error ellipse using automated P-wave arrivals and WL2 norm minimization. Examples with good azimuthal station coverage, (a) 2014 M_w6.1 South Napa, California, and with one-sided station coverage, (b) 2010 M_w7.2 El Mayor-Cucapah, Mexico and (c) 2010 M_w8.2 Iquique, Chile. The azimuthal gap associated with the estimate in (a) is 70.8°, and the distance between our estimate and the USGS published epicenter is 1.55 km. The small error ellipse is obscured by the epicenter symbols (see Table 2.1 for details). The distance from the published hypocenter to the nearest station is about 28 km. The azimuthal gap associated with the estimate in (b) is 259.3°, and the distance between our estimate and the USGS published epicenter is about 10.49 km. The distance from the published hypocenter to the nearest station is about 72 km. The azimuthal gap associated with the estimate in (c) is 313.4°, and the distance between our estimate and the USGS published epicenter is about 130.50 km. The distance from the published hypocenter to the nearest station is about 70 km.

The locations estimated using the WL2 solution are shown in Figure 2.6 for the three investigated earthquakes. A full description of solver types and results are listed in tables 2.1, 2.2, and 2.3. For the 2014 M_w6.0 South Napa, California earthquake, good azimuthal coverage (azimuthal gap of 70.8°) results in a hypocenter estimate only 1.55 km from the USGS published hypocenter (Figure 2.6a). For this earthquake, all three estimators provide nearly identical results. This is encouraging, but not surprising, given that the difference between the manual and automatic picks and the azimuthal gap are small. The nearly identical goodness of fit metrics indicate that all picks are of the same quality, there is no need for distance weighting, and there

Table 2.2: Hypocenter location: 2010 M_w7.2 El Mayor-Cucapah, Mexico. Results of location algorithm for different solver types: L1, L2, and WL2. Depth is assumed to be 10 km for crustal California. The USGS published hypocenter is also listed for comparison. The last row shows the results from simple PGD magnitude scaling relations using the seismogeodetic displacement records from the 12 available stations and the listed hypocenter as input (Melgar et al., 2015a).

Solver		**USGS**	**L1**	**L2**	**WL2**
Longitude (°E)		-115.29	-115.34	-115.30	-115.24
Latitude (°N)		32.26	32.42	32.33	32.34
Origin Time (UTC)		22:40:57	22:40:57.46	22:40:56.10	22:40:56.74
Depth (km)		10	10	10	10
Azimuthal Gap (°)		—	249.45	257.68	259.28
Goodness of Fit		—	0.37	1.96	1.63
Error Ellipse (95%)	Azimuth (°)	—	-29.97	-24.81	-10.04
	Major (km)	—	12.94	11.76	41.53
	Minor (km)	—	6.07	5.32	5.88
	± Origin Time	—	1.56	1.45	7.18
Distance from USGS (km)		—	18.36	8.24	10.49
PGD Magnitude		7.26	7.21	7.24	7.24

are no apparent outliers in the data. However, the goodness of fit metric is on the low end ($\sim$0.05) compared to the expected value of unity indicating that we may have underestimated the precision of our observations.

In the case of the 2010 M_w7.2 El Mayor-Cucapah, Mexico earthquake, station coverage is one-sided (all stations northward with respect to the hypocenter), with the best coverage within the northwest quadrant of the strike slip fault. However, since the event is mainly strike-slip, even one-sided coverage covers two quadrants of motion. The azimuthal gap is wide, about 260 degrees, and the distance from the hypocenter location to the nearest recording stations is over 70 km. This poor azimuthal coverage is likely to negatively impact the accuracy of the epicenter estimate, and any estimate accompanied by an azimuthal gap greater than 180° should be considered with appropriate skepticism. Nevertheless, we are able to resolve the hypocenter location within about 10 km compared to the published USGS location with minor variations among the three estimators (Figure 2.6b, Table 2.2). The goodness of fit parameter is 0.4 for the L1 norm, and about 1.8 to 2.0 for the L2 norm and WL2 norm. This appears to indicate that the

Table 2.3: Hypocenter location: 2014 M_w8.2 Iquique, Chile. Results of location algorithm for different solver types: L1, L2, and WL2. Depth is assumed to be 25 km for offshore western Chile. The USGS published hypocenter is also listed for comparison. The last row shows the results from simple PGD magnitude scaling relations using the seismogeodetic displacement records from the 11 available stations and the listed hypocenter as input (Melgar et al., 2015a).

Solver		**USGS**	**L1**	**L2**	**WL2**
Longitude (°E)		-70.77	-71.81	-71.82	-71.77
Latitude (°N)		-19.61	-18.92	-18.90	-18.92
Origin Time (UTC)		23:46:47	23:46:29.50	23:46:29.38	23:46:29.97
Depth (km)		25	25	25	25
Azimuthal Gap (°)		—	313.42	313.82	313.62
Goodness of Fit		—	0.08	0.08	0.06
Error Ellipse (95%)	Azimuth (°)	—	-43.78	-41.35	-44.46
	Major (km)	—	56.77	52.21	60.09
	Minor (km)	—	8.80	7.58	9.34
	± Origin Time	—	5.74	6.02	6.38
Distance from USGS (km)		—	133.41	135.53	130.50
PGD Magnitude		7.40	7.65	7.65	7.65

arrival times are not of the same quality as for the South Napa earthquake, and that there may be some picks for El Mayor-Cucapah that have greater error. The 1D velocity structure used may also be insufficient for the El Mayor-Cucapah event, given its location in a complicated rift zone environment (Hauksson et al., 2011). Considering that the South Napa network geometry is optimal, and after taking into account the low goodness of fit, it is apparent that the El Mayor-Cucapah prior variances have been underestimated. In this case, the L2 solution is closer than the WL2 solution with respect to the USGS location, but not significantly different (Table 2.2).

For the M_w8.2 Iquique earthquake, all stations are within the same quadrant of motion for the subduction zone off western Chile, with a distance of over 70 km from the epicenter to the closest station and a very wide azimuthal gap of $\sim$313 degrees. Furthermore, the estimated hypocenter location is $\sim$130 km from the USGS published solution. Here the goodness of fit is low, $\sim$0.1. It appears then that the precision of the observations is high but that there is a large bias in the estimate so that the accuracy is low. This is most likely due to the poor network coverage for this earthquake (Figure 2.6c).

While the error ellipse denotes precision of our estimate, the azimuthal gap is a more robust metric for signifying confidence in our solution's accuracy (see Figure 2.6c). The accuracy of the epicenter estimate is affected by several factors including P-wave detection accuracy (see Section 2.2.2), assumed velocity structure, and station configuration. Station configuration appears to be a significant source of error for the situation in which regional station coverage is one-sided. For example, the P-wave arrival times from a nearby event, at distance, d, with some origin time, τ, could be very similar to those from an event at a greater distance, $d + \delta d$, with earlier origin time, $\tau - \delta\tau$. Thus, in the case of an earthquake with one-sided station coverage, simply fixing the depth (as described in Section 2.2.1), may not be sufficient for accurately calculating the hypocenter location, and formal and/or bootstrapped error calculations may not be sufficient for demonstrating the solution accuracy.

The accuracy of the hypocenter estimate is important in downstream products, especially rapid magnitude scaling. Our system currently employs the simple scaling relation described in Melgar et al. (2015a), which relates the peak ground displacement (PGD) and hypocentral distance of each recording station to quickly estimate the moment magnitude. Any error in the hypocenter location will affect the calculated hypocentral distance and therefore propagate to the magnitude estimate. Since the relation is logarithmic, small errors do not have a large effect on the resulting moment magnitude. For example, for the same PGD measurements, a magnitude 5 earthquake could be mistaken for a magnitude 6 only if mislocation exceeds 100 km. For larger magnitudes, still a larger mislocation required for a an error that exceeds ± 1 magnitude unit. Using the Iquique event as an example, we estimated the moment magnitude using PGD measured at our 11 stations and calculating hypocentral distance both from the USGS published hypocenter and from the L1, L2, and WL2 norm minimizations ($>$100 km error compared to the USGS published hypocenter). This discrepancy results in a difference of only $<$0.3 magnitude units using the PGD scaling relation (Table 2.3).

2.3.3 Resolving the Earthquake Location

Station distribution has a significant impact on hypocenter estimation accuracy. Lomax et al. (2009) includes a detailed discussion of hypocenter uncertainty from few stations, or poor station coverage. For two of our case studies, the 2010 M_w7.2 El Mayor-Cucapah and 2014 M_w8.2 Iquique earthquakes, the available stations are located only to one side of the epicenter. In California, as station coverage improves, we are less likely to encounter this issue, but may continue to be affected by inability to receive real-time data streams from northern Mexico (see Section 2.3.4). Furthermore, for subduction zone events, the seismogeodetic network will typically only be landward of the hypocenter (the Sumatra megathrust is one exception, where observations are available on the Mentawai Islands located on the forearc). Thus, we must assume that station coverage could impact our ability to rapidly resolve the hypocenter location, and characterize the capabilities of our method based on the station configuration.

We performed synthetic testing to quantify the expected error in a hypocenter estimate given the available station coverage. For a grid of potential hypocenter locations in the region of the 2011 M_w8.2 Iquique, Chile earthquake, at 25 km depth, we calculate the expected arrival times at each available seismogeodetic station using 1D ray tracing, and add error to the arrival time pick based on the typical error distribution of the automatic P-wave detection algorithm (see Section 2.2.2, and Figure 2.5b). Then, we use these synthetic arrival times to invert for the epicenter location, to determine how reliably the algorithm returns the initial input epicenter. The size of our dataset limits our ability to characterize the dependence of source-station distance on P-wave pick error, so although there is likely some correlation (Figure 2.4), such dependence is not included in this synthetic test. Therefore, we view the L2 to be more appropriate for such testing than the WL2 or L1 norm minimization, which are only applicable if it is assumed that arrival times are more accurate at nearby stations, where the signal is larger.

An example of this process is given in Figure 2.7a. Arrows connect the input epicenter (tail) to the epicenter resulting from the location algorithm (head). Short arrows indicate high

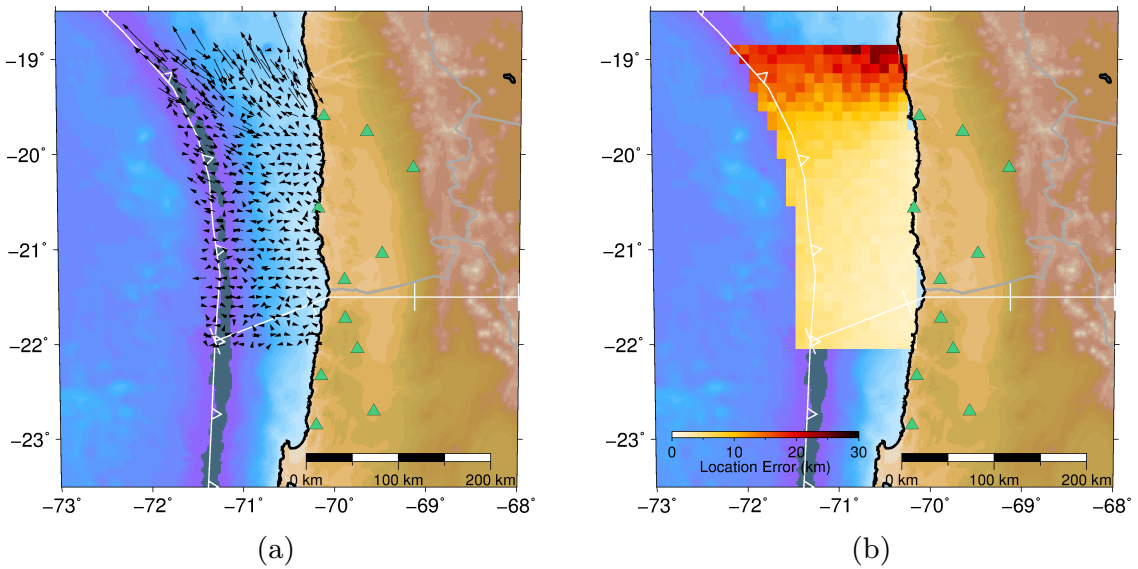
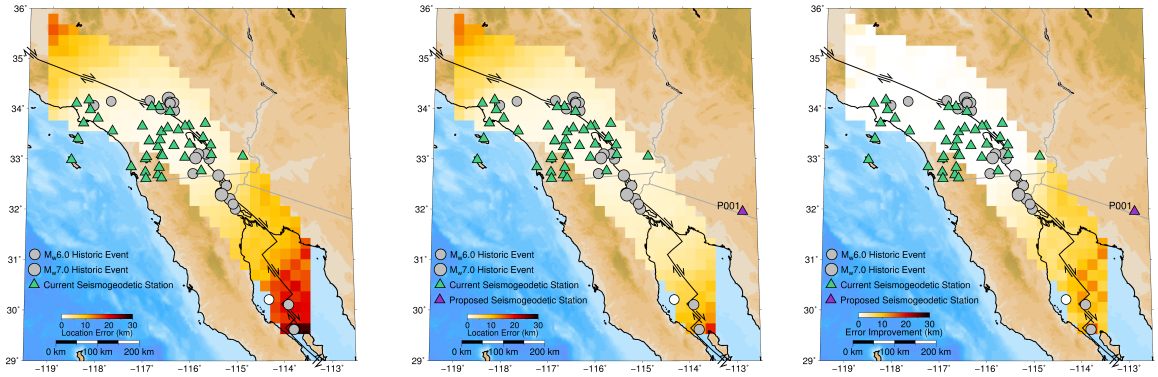


Figure 2.7: Monte Carlo test of epicenter location. (a) Single run of Monte Carlo test to determine reliability of epicenter location in the subduction zone region west of the Iquique, Chile earthquake given geometry of seismogeodetic stations and (b) average epicenter location error for 100 iterations.

accuracy, long arrows indicate low accuracy, and the arrow length signifies the magnitude of inaccuracy in the solution. The process is repeated 100 times and error at each input location is averaged over the many iterations to obtain the typical inaccuracy for an event at that location (Figure 2.7b). Note that this method shows only the robustness of our estimate with respect to P-wave pick errors and station geometry, but does not account for additional error associated with the selection of the velocity model (the velocity model may be a significant source of error, but a 3D velocity model can be difficult to constrain and is beyond the scope of this study).

Many iterations are required to identify persistent patterns in location solution accuracy. However, individual iterations of the sensitivity test highlight some of the recurring features. For example, the single iteration for the region near the Iquique, Chile earthquake shown in Figure 2.7a indicates that an epicenter location will likely be well-resolved within the latitudes covered by the seismogeodetic network, but less well-resolved immediately north of the station coverage. Particularly on the northern and western sides of the grid, there is a geographical bias pushing



(a) (b) (c)

Figure 2.8: Automated epicenter location reliability in Southern California (a) using the current seismogeodetic network, (b) with the addition of proposed seismogeodetic upgrade to Arizona GPS station P001, and (c) the improvement from (a) to (b). Current seismogeodetic stations are shown as green triangles, with proposed seismogeodetic upgrade at station P001 in purple (b and c). Locations of $>M_w5.5$ historical earthquakes along the major fault are shown as grey circles (from USGS). The recent March 2016 $M_w5.5$ event is in white.

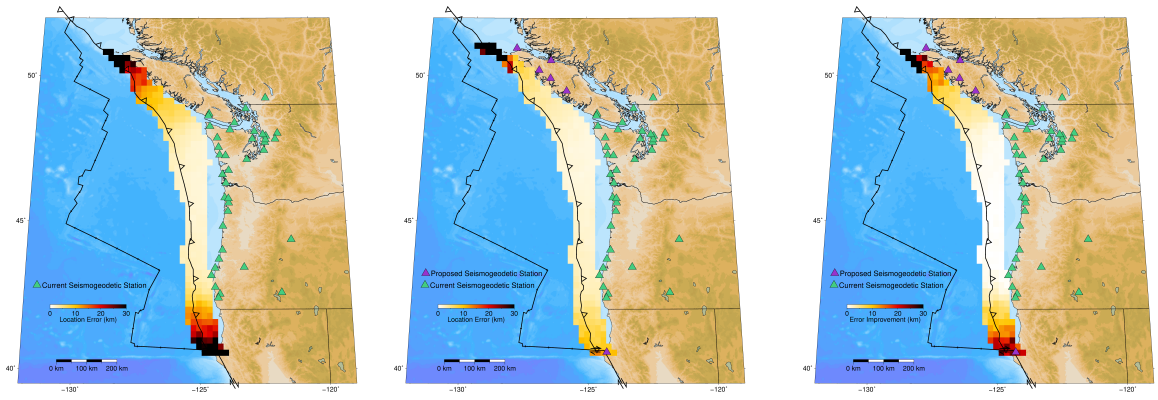
the epicenter further northwest. This geographic pattern is characteristic of each iteration of the sensitivity study and represents the introduced bias in the estimated epicenter due to geographic distribution of the available near-source seismogeodetic stations. It also may help to explain the overestimate of the epicenter toward the northwest (trading off with an earlier origin time). Furthermore, arrow orientation suggests that the epicenter locations are least resolved in the array-normal direction, and better constrained in the direction parallel to the network. In Figure 2.7b, we have averaged the the results from 100 iterations at each input location. Again, the earthquake location is shown to be best resolved within the latitudes covered by the network, and less resolved to the north and west of the tested locations.

2.3.4 Seismogeodetic Upgrades

We can use this method of synthetic testing to determine the locations of possible additional stations that would have the greatest improvement on epicenter resolution. We include an

example for the southern California seismogeodetic network. Using the 2010 El Mayor-Cucapah event as a test, we demonstrated the ability to recover the hypocenter within 8 km using the L2 norm minimization (Table 2.2). Since then, the deployment of 16 SIO GAPs (SIO Geodetic Module and low-cost MEMS Accelerometer Package, designed as a simple plug-in for existing GPS stations (Saunders et al., 2016)) have improved seismogeodetic station coverage. It would be prudent to select the locations of further upgrades based on their ability to improve station configuration with respect to epicenter estimation. Therefore, we complete Monte Carlo testing along the southern San Andreas Fault using the methodology described in Section 2.3.3 for 100 iterations, averaging the location error at each input location over the many iterations. It is clear that given the network geometry, we can easily resolve the epicenter location within the network (Figure 2.8a). However, our ability to resolve the epicenter location is especially poor in the Gulf of California, where reasonably sized earthquakes have occurred historically, most recently including a M_w5.5 in March 2016 about 100 km SSE of San Felipe, Mexico. A repeat of this synthetic test supposing the upgrade of GPS station P001, in Arizona, to a real-time seismogeodetic station demonstrates a marked improvement in the Gulf of California, including the region of the recent M_w5.5 earthquake, suggesting that upgrade of this station is a strategically important improvement to the current array (Figures 2.8b, 2.8c). Station P001 was upgraded with real-time radio telemetry and an SIO GAP by UNAVCO in November 2016 based on this analysis. For robustness, additional stations should be chosen to improve the network design.

The same method can be used to prioritize locations for upgrade in the Cascadia region of western North America. We consider the location resolution along the locked portion of the subduction zone, where we might expect the largest, most destructive, and possibly tsunamigenic events to occur. Figure 2.9a demonstrates that our station coverage should allow location accuracy within 10 km along the locked portion of the Cascadia subduction zone within the area directly west of our station coverage (remember, this estimate does not account for additional error due to velocity structure). However, the subduction area extends southward to northern California



(a) (b) (c)

Figure 2.9: Automated epicenter location reliability in Cascadia using (a) the current seismogeodetic network, (b) with the addition of proposed seismogeodetic upgrade to stations in the Vancouver area (HOLB, ELIZ, NTKA, TGNO, and WOST) and in northern California (P157) and (c) the improvement from a to b. Current seismogeodetic stations are shown as green triangles, with proposed seismogeodetic upgrades in purple (b and c).

and northward into Canada, and these areas remain poorly resolved given current coverage. In fact, at the northernmost part of the grid, there are fewer than four seismogeodetic stations within 400 km of these potential hypocenters; therefore these locations are severely limited by a direct P-wave arrival inversion method. In Figure 2.9b, we demonstrate that by upgrading five existing GNSS stations in Vancouver to include strong motion sensors, we begin to improve hypocenter location resolution in the northern part of the grid.

The northernmost row of the grid remains poorly constrained, with an average location error between 20 and 25 km, and represents an area where additional seismogeodetic upgrades in British Columbia would be appropriate to better resolve the hypocenter location. Furthermore, if only one seismogeodetic station upgrade were available to help constrain epicenter location along the southernmost part of the subduction zone, we demonstrate the improvement possible by adding an accelerometer to GPS station P157 in northern California. We have demonstrated the improvement attained with minimal upgrades, however further additions to the ones suggested here would allow an even greater improvement.

Figure 2.10: Seismogeodetic station SIO5 in La Jolla, California. SIO MEMS accelerometer is visible on the vertical brace of the continuous GPS station.

2.4 Seismogeodetic EEW with MEMS Accelerometers

We suggest that the upgrade of continuous GPS stations with MEMS accelerometers such as those deployed with the SIO GAP (Saunders et al., 2016) provides an inexpensive alternative that does not impede early warning applications. The SIO GAP is a low-cost solution that affixes to the vertical brace of the continuous GPS monument (Figure 2.10). The geodetic module time tags the acceleration data using the GPS clock and can also be used to combine the separate seismic and geodetic data streams into the merged seismogeodetic product on site. GPS networks alone are not sufficient for EEW, requiring a seismic trigger to be effective. The SIO GAP upgrade transforms the network into a standalone seismogeodetic EEW system.

Figure 2.11 shows data from seismogeodetic station P224, equipped with observatory-grade GPS and SIO GAP, used for detection of a M4 earthquake near Piedmont, California on August 17, 2015. This station is only 2.85 km from the epicenter, and the event is detected at this station 0.48s after the manual pick. Adjusting the detection parameters to a threshold of 2.9 brings the automated pick to within 0.3 s of the manual P-wave pick, which is equal to the

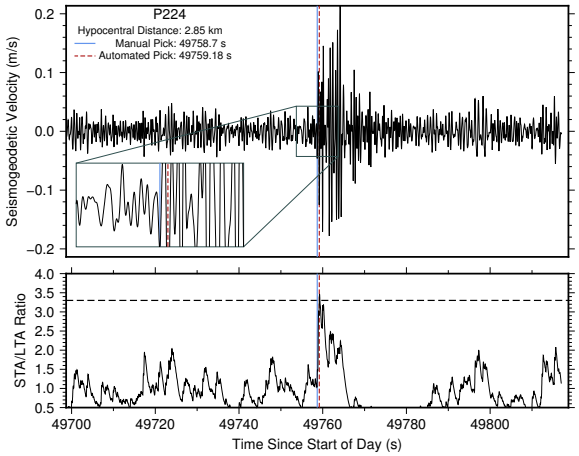


Figure 2.11: M4 earthquake detection of first seismic wave arrival at a station recently upgraded with an inexpensive SIO GAP. The station is located at a hypocentral distance of 2.85 km from the 2014 M4 Piedmont earthquake. Top panel shows bandpassed seismogeodetic velocity with manual pick (blue) and automated pick (dashed red), with an inset highlighting the P-wave arrival. Bottom panel shows the STA/LTA ratio, with horizontal dashed line indicating the detection threshold.

average pick error (Figure 2.5). An M4 earthquake is below the size of interest to early warning, so this example is the extreme case. Nonetheless, it provides confidence that the inexpensive upgrades will be sufficient for these applications.

On June 10, 2016, a M_w5.2 earthquake struck Borrego Springs, California within our network of collocated GPS with SIO GAPs. This provided an additional dataset for testing our current capabilities using seismogeodetic early warning algorithms, fed by GPS and SIO MEMS accelerometer data. The nearest SIO GAP-equipped station was 20.36 km from the hypocenter and detected the P-wave arrival to within 0.14 s of the manual pick (Figure 2.12).

Station SIO5, at a hypocentral distance of 100 km had been experiencing data collection issues with its MEMS accelerometer, logging 0 m/s^2 for significant segments of the time series leading up the earthquake onset. The seismogeodetic combination was therefore also contaminated by this instrument issue (see Figure S2.9b). As such, we decided to eliminate it from further analysis and plan to implement quality control measures that will remove malfunctioning instruments from the EEW products. Figure 2.13 shows the final early warning products for the

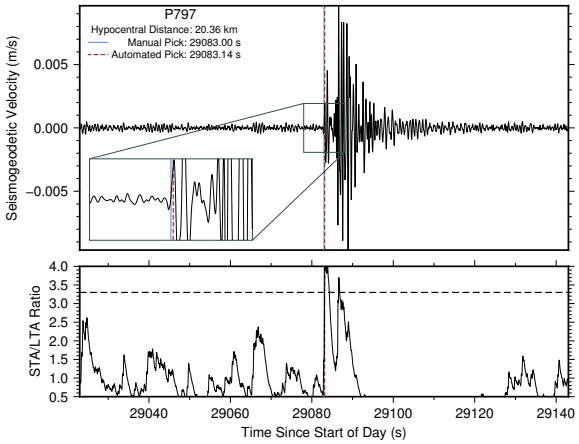


Figure 2.12: Automated P-wave detection at GPS station P797, recently equipped with an SIO GAP, demonstrates that detection of first seismic arrivals using seismogeodetic velocity time series is possible with GNSS and inexpensive SIO MEMS accelerometer. The station is located at an hypocentral distance of ∼20 km from the 2016 M_w5.2 Borrego Springs earthquake. Top panel shows bandpassed seismogeodetic velocity with manual pick (blue) and automated pick (dashed red), with an inset highlighting the P-wave arrival. Bottom panel shows the STA/LTA ratio, with horizontal dashed line indicating the detection threshold.

Borrego Springs event including detection, hypocenter estimation via WL2 norm minimization, and PGD magnitude scaling. It should be noted that displacements from this earthquake were not large enough to be discernible above the typical noise at most stations. Typically, a magnitude should not be calculated unless the PGD exceeds the noise at at least four stations. If this criterion is not met, the maximum possible magnitude can be estimated by assuming the PGD is equivalent to the amplitude of the noise at the nearest four stations. From this logic, the PGD magnitude estimate calculated here should be considered the upper limit estimate. Furthermore, the PGD scaling relation of Melgar et al. (2015a) was derived using events $>M_w$5.9, so this event is outside the limits for which the relation was designed.

Since the Borrego Springs event was relatively small— the very low end of earthquake size of interest to early warning—some of the P-wave arrivals at regional stations were not detectable above the station noise; some stations triggered on the larger, later seismic arrivals (Stations P483, RAAP, and DESC) and some stations did not detect the event at all (GLRS, P494, and POTR). Figures detailing detection at each of these stations are available in the supporting

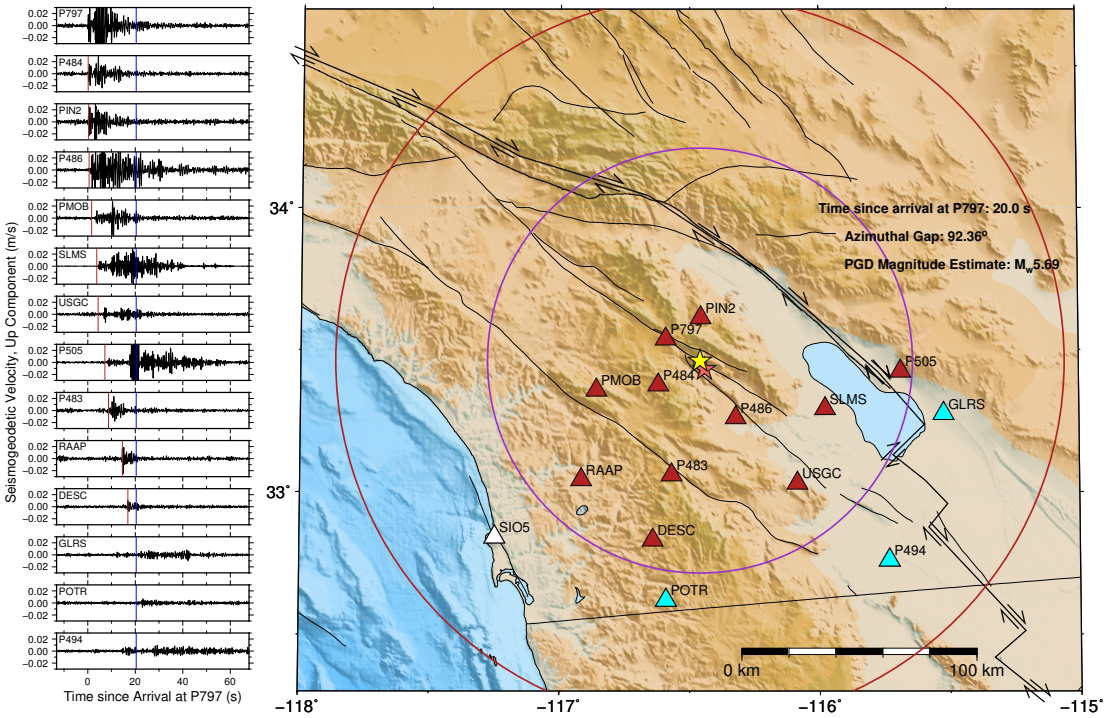


Figure 2.13: Seismogeodetic earthquake early warning products for 2016 M_w5.2 Borrego Springs event. Left: Seismogeodetic velocity time series for local seismogeodetic stations (GPS with SIO GAPs). Blue line denotes current epoch, red line denotes automated detection time at each station. Right: USGS published epicenter (pink star), automated epicenter estimate (yellow star) from WL2 solution and seismogeodetic station coverage (triangles) for this event. Red triangles denote stations that recorded a trigger. Light blue stations did not trigger given their signal-to-noise ratio. The white triangle shows the location of malfunctioning station SIO5, which was removed from the analysis. This frame shows the warning information available 20.0 s after the first detection at station P797. Concentric circles indicate theoretical P-wave (red) and trailing S-wave (magenta) fronts based on hypocenter location and calculated earthquake origin time. The apparent P- and S-wave velocities are assumed to be 6.0 and 3.5 km/s, respectively. S-wave front is used as a proxy for shaking onset time.

Table 2.4: **Hypocenter location: 2016 M_w5.2 Borrego Springs, California, USA.** Results of hypocenter location algorithm for 2016 M_w5.2 Borrego Springs earthquake using different solver types: L1, L2, and WL2. Depth is assumed to be 10 km for crustal California. The USGS published hypocenter is also listed for comparison. The last row shows the results from simple PGD magnitude scaling relations using the seismogeodetic displacement records from the 11 available stations and the listed hypocenter as input (Melgar et al., 2015a).

Solver		**USGS**	**L1**	**L2**	**WL2**
Longitude (°E)		-116.44	-116.44	-116.34	-116.46
Latitude (°N)		33.43	33.50	33.72	33.49
Origin Time (UTC)		08:04:38	08:04:37.99	08:04:36.61	08:04:38.84
Depth (km)		10	10	10	10
Azimuthal Gap (°)		—	106.86	248.91	92.43
Goodness of Fit		—	0.63	5.20	1.97
Error Ellipse (95%)	Azimuth (°)	—	18.39	-1.49	244.89
	Major (km)	—	7.60	8.32	4.98
	Minor (km)	—	6.32	3.84	2.71
	± Origin Time	—	0.65	1.02	0.25
Distance from USGS (km)		—	8.16	33.89	3.41
PGD Magnitude		5.54	5.59	5.87	5.56

information). Regardless, using a WL2 norm minimization, we resolved the epicenter location in this region to within 3.41 km. Our implementation using the unweighted L2 norm minimization resulted in an inaccurate location more than 30 km from the USGS published epicenter. The L1 and WL2 approaches are more reasonable for small events, where inaccurate picks on later, larger arrivals are more likely, because the former is better equipped to handle outliers, and the latter relies more on closer stations, which are more likely to have accurate arrival time picks. From this and the previous results, we suggest that the WL2 is the most consistently reliable method for the epicenter inversion.

We propose a method to improve all our estimates, but specifically the L2 norm minimization solution, in real-time through calculation of the goodness of fit with each additional observation. In doing so, we observe that there are notable increases to the value of chi-squared with the influence of stations with late detections, specifically with the arrival at station P483 where the chi-squared value increases more than ten-fold (Table 2.5). From Figure S2.7i, it is clear

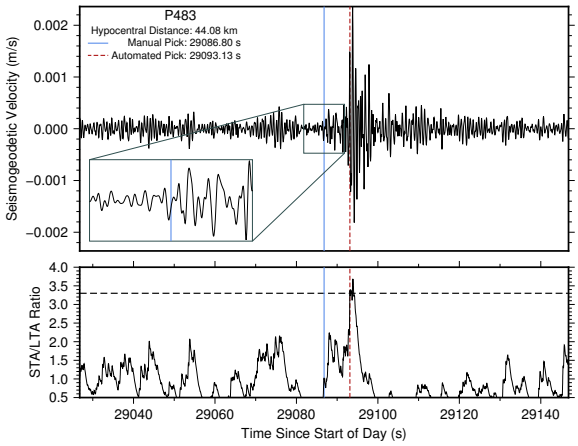


Figure 2.14: Automated detection performance at station P483, where a low signal-to-noise ratio prevents the accurate detection of the first seismic arrival. Instead, the detection occurs on a larger, later arrival. Top panel shows bandpassed seismogeodetic velocity with manual pick (blue) and automated pick (dashed red), with an inset highlighting the period near the P-wave arrival. Bottom panel shows the STA/LTA ratio, with horizontal dashed line indicating the detection threshold.

Table 2.5: Epicenter estimate evolution 2016 M_w5.2 Borrego Springs, CA, USA, L2 norm minimization. Effect of addition of stations on the epicenter estimate and goodness of fit using the unweighted L2 norm minimization. Once 4 stations register a detection, we invert for the epicentral location using the method of Section 2.3.1. Depth is assumed to be fixed at 10 km for crustal California events. The fit of the model to the data worsens with the addition of station P483, a station which has clearly triggered on the larger S-wave arrival (see Figure S2.7i).

New Station	**Longitude (°E)**	**Latitude (°N)**	**Origin Time (UTC)**	**Azimuthal Gap (°)**	**Chi-Squared ($\widehat{\sigma}_0^2$)**	**Distance from USGS (km)**
P797	—	—	—	—	—	—
P484	—	—	—	—	—	—
PIN2	—	—	—	—	—	—
P486	–116.4726	33.4532	08:04:37.94	142.42	0.24	3.69
PMOB	-116.4642	33.4549	08:04:37.96	146.61	0.23	3.30
SLMS	-116.4442	33.4539	08:04:37.77	116.47	0.28	2.54
USGC	-116.4509	33.4611	08:04:37.88	115.28	0.31	3.54
P505	-116.4533	33.4622	08:04:37.89	94.61	0.24	3.59
P483	-116.4153	33.5370	08:04:38.05	125.46	2.82	12.04
RAAP	-116.3593	33.5962	08:04:37.91	185.83	4.37	19.90
DESC	-116.3424	33.7242	08:04:36.61	248.91	5.20	33.89

Table 2.6: Epicenter estimate evolution 2016 M_w5.2 Borrego Springs, CA, USA, WL2 norm minimization. Effect of addition of stations on the epicenter estimate and goodness of fit using the WL2 norm minimization. Once 4 stations register a detection, we invert for the epicentral location using the method of Section 2.3.1. Depth is assumed to be fixed at 10 km for crustal California events. The fit of the model to the data worsens with the addition of station P483, however the value of χ^2 associated with the solution confirms that the WL2 norm minimization fares better with the addition of poor automated picks compared to the L2 solution (Table 2.5).

New Station	Longitude (°E)	Latitude (°N)	Origin Time (UTC)	Azimuthal Gap (°)	Chi-Squared ($\widehat{\sigma}_0^2$)	Distance from USGS (km)
P797	—	—	—	—	—	—
P484	—	—	—	—	—	—
PIN2	—	—	—	—	—	—
P486	–116.4720	33.4526	08:04:38.87	88.07	0.27	3.68
PMOB	-116.4713	33.4525	08:04:38.87	88.18	0.20	3.55
SLMS	-116.4603	33.4503	08:04:38.78	91.34	0.25	2.67
USGC	-116.4620	33.4511	08:04:38.80	90.89	0.21	2.85
P505	-116.4613	33.4517	08:04:38.78	91.17	0.19	2.86
P483	-116.4604	33.4552	08:04:38.82	91.71	1.18	3.13
RAAP	-116.4605	33.4574	08:04:38.83	91.89	1.85	3.35
DESC	-116.4591	33.487	08:04:38.84	92.43	2.35	3.41

that the P-wave at station P483 was not distinguishable above the noise, and detection occurred on the S-wave arrival. Similar associations can be drawn from the arrivals at RAAP and DESC and increase to the chi-squared value. Details for detection at these two stations are available in the supplementary material. By rejecting the new station(s) associated with a large increase to chi-squared (below double line in Table 2.5), we settle on an epicenter that better aligns with the known location. Compared to the same analysis using the WL2 norm minimization, we again see an increase to the value of chi-squared (by 5x) with the influence of the stations with poor automated detections. However, the WL2 solution is associated with a more favorable goodness of fit compared to the unweighted L2 solution, because the weighting scheme is such that further stations have less influence.

We can perform a simple F-test to justify removal of individual stations. If all station observations are of the same quality, then the variance of the fits between the observations and corresponding model should be statistically equivalent regardless of the number of stations used

to obtain the model. In other words, by comparing the chi squared goodness of fits between subsequent solutions as newly triggered stations are added to the hypocenter inversion, we can state whether the data fits come from statistically identical distributions.

$$F = \frac{s_9^2}{s_8^2} = \frac{\chi_9^2/\nu_9}{\chi_8^2/\nu_8} \tag{2.13}$$

where s^2 is the sample variance, ν is the degrees of freedom, and the subscripts denote the number of stations used in the inversion. The degrees of freedom is again $\nu = n - u$. Since we solve for three parameters (x_0, y_0, and τ_0), $u = 3$, and n is of course the number of stations. In this example, we consider the case with 8 stations (before the inclusion of station P483) compared to that with 9 stations (the first inversion to include station P483). The hypothesis we test is that $s_9^2 = s_8^2$ or $\frac{s_9^2}{s_8^2} = 1$. The value $F_{\nu_9,\nu_8,\alpha}$ is the F value such that the probability that F exceeds $F_{\nu_9,\nu_8,\alpha}$ is α. We use $\alpha = 0.05$ to test 95% confidence.

$$F = \frac{\chi_9^2/\nu_9}{\chi_8^2/\nu_8} = \frac{2.82}{0.24} = 11.75$$

$$F_{6,5,0.05} = 4.95$$

$$F > F_{6,5,0.05} \longleftarrow \text{Hypothesis is Rejected}$$

From this analysis, we see that the distributions are statistically inequivalent, which may justify elimination of station P483. Note that by this analysis, station P483 would be rejected using the WL2 solution as well, $F = \frac{1.18}{0.19} = 6.21$, although the value of F exceeds $F_{6,5,0.05}$ by far less in the WL2 case. If station P483 is eliminated, the F value associated with including stations RAAP and DESC results in grounds for their elimination by this method as well. A full description of this evolution of chi squared and F-test rejection criteria is available in tables 2.7 and 2.8. Of course, it is not certain that the first stations are most accurate and the error necessarily lies in a more distant station. In this case, analysis of all permutations of stations could help to isolate the offending observation, in particular for the stations closest to the earthquake epicenter.

Table 2.7: Evolution of chi-squared parameter for 2016 M_w5.2 Borrego Springs, California, L2 epicenter solver. Note that because station P483 is rejected, the number of stations when RAAP is added remains at 9. The same is true for station DESC.

2016 M_w5.2 Borrego Springs, California, USA, L2 norm minimization						
Newest Station	**# Stations (n)**	**Degrees of Freedom**	**Chi-Squared ($\widehat{\sigma}_0^2$)**	$F = \frac{s_n^2}{s_{n-1}^2}$	$F_{n,n-1,0.05}$	**Reject?**
P486	4	1	0.24	–	–	–
PMOB	5	2	0.23	0.958	199.500	Fail to Reject
SLMS	6	3	0.28	1.212	19.164	Fail to Reject
USGC	7	4	0.31	1.101	9.117	Fail to Reject
P505	8	5	0.24	0.774	6.256	Fail to Reject
P483	9	6	2.82	11.750	4.950	Reject
RAAP	9	6	4.43	18.458	4.950	Reject
DESC	9	6	5.61	22.958	4.950	Reject

Table 2.8: Evolution of chi-squared parameter for 2016 M_w5.2 Borrego Springs, California, WL2 epicenter solver. Note that because station P483 is rejected, the number of stations when RAAP is added remains at 9. The same is true for station DESC.

2016 M_w5.2 Borrego Springs, California, USA, WL2 norm minimization						
Newest Station	**# Stations (n)**	**Degrees of Freedom**	**Chi-Squared ($\widehat{\sigma}_0^2$)**	$F = \frac{s_n^2}{s_{n-1}^2}$	$F_{n,n-1,0.05}$	**Reject?**
P486	4	1	0.27	–	–	–
PMOB	5	2	0.20	0.747	199.500	Fail to Reject
SLMS	6	3	0.25	1.250	19.164	Fail to Reject
USGC	7	4	0.21	0.840	9.117	Fail to Reject
P505	8	5	0.19	0.905	6.256	Fail to Reject
P483	9	6	1.18	6.210	4.950	Reject
RAAP	9	6	1.24	6.526	4.950	Reject
DESC	9	6	1.27	6.684	4.950	Reject

Additional work will be required to ensure that this pattern is consistent when applied to future medium magnitude events, where late detections are more likely. Given that this example describes the low end of earthquake size of importance to EEW systems, our analysis provides confidence in the performance of seismogeodesy using the SIO GAPs, and reinforces the conclusions of Saunders et al. (2016), that this instrumentation is sufficient for earthquakes of significance to EEW.

2.5 Conclusions

We have described the elements of our seismogeodetic earthquake and tsunami early warning system, where the detection and location algorithms initiate higher-order products such as earthquake magnitude scaling, and centroid moment tensor and finite-fault slip solutions. These higher order products can then be used as input to local tsunami prediction and modeling algorithms. The seismogeodetic method allows for a self-contained earthquake early warning system that utilizes the best qualities of seismic and geodetic approaches to earthquake and local tsunami early warning. Specifically, determining with confidence that no permanent displacement has occurred is a valuable piece of information. A GNSS-only system is not sufficiently precise for early warning based on P-wave metrics. The seismogeodetic approach is capable of detection of first seismic arrivals, comparable to manual P-wave picks from observatory-grade accelerometers and broadband seismometers. We have not discussed the issue of false detections and do not have sufficient statistics to be able to say that seismogeodesy will reduce them. However, seismogeodesy may be more robust than either seismic or GNSS because it has an additional degree of freedom.

The seismogeodetic combination allows rapid magnitude scaling without saturation, a critical metric for EEW and LTW, and one of the major limitations of seismic-only systems. Rapid estimation of the coseismic offset, required for timely centroid moment tensor solutions, is made

simpler with the seismogeodetic combination; the velocity time series can immediately detect the timing of cessation of shaking, and the corresponding broadband displacement waveform contains the static (0 Hz) displacement. The seismogeodetic approach initiates robust EEW and LTW products from displacement and velocity waveforms without the requirement for external seismic location and magnitude input to start the process. The performance of a self-contained seismogeodetic approach is still limited by sparse collocated instrument coverage in some areas, due to independent development of seismic and GNSS networks. Purposefully designed collocations have recently begun being prioritized for installation. Due to the advantages of the seismogeodetic approach to early warning, we suggest the upgrade of currently-available single-instrument stations to seismogeodetic capability. From the 2015 M_w5.2 Borrego Springs earthquake, we have demonstrated that the MEMS accelerometer is sufficient for local early warning systems, and therefore provides a cost effective option for expanding seismogeodetic station coverage. We suggest prioritization of seismogeodetic upgrades based upon the effect of station distribution on location resolution to most efficiently fill gaps in the coverage of high-risk areas. The results of our automated P-wave detection and location algorithms provide confidence in the performance of the self-contained system, especially in California, where station coverage around high-risk faults is improving. With increased attention to the creation of collocated strong motion and geodetic stations, the seismogeodetic approach will be an important addition to early warning efforts.

2.6 Supporting Information

2.6.1 2014 M_w6.0 Napa Supporting Information

Table S2.1: Collocated stations used to analyze the 2014 M_w6.0 Napa, California earthquake.

Accelerometer	Longitude	Latitude	GPS	Longitude	Latitude	Separation (km)
N014	-122.6099	38.2473	P198	-122.6075	38.2599	1.41
BRIB	-122.1518	37.9189	BRIB	-122.1526	37.9194	0.08
BL67	-122.2432	37.8749	LRA3	-122.2437	37.8754	0.08
NTAC	-122.5965	37.9237	T3RP	-122.5880	37.9280	0.85
68329	-122.7476	38.4413	P197	-122.7674	38.4286	2.21
BDM	-121.8655	37.9540	P248	-121.8687	37.9756	2.41
CBR	-122.0630	37.8165	ROCP	-122.0630	37.8160	0.05
CMOB	-121.9620	37.7496	P229	-121.9780	37.7494	1.41
C045	-121.8027	37.8104	P230	-121.7864	37.8190	1.74
NPRB	-123.0195	37.9964	PTRB	-123.0187	37.9962	0.04
69039	-123.0527	38.3110	P183	-123.0689	38.3137	1.42
J039	-122.4528	37.6212	SWEP	-122.4580	37.6120	1.11
J023	-122.4428	38.8787	P178	-122.4428	38.8787	0.02
MNRC	-122.3469	37.5255	MNRC	-122.3324	37.5345	1.63
J009	-122.5117	37.5197	P177	-122.4951	37.5282	1.77
FARB	-123.0011	37.6978	FARB	-123.0008	37.6972	0.07
1787	-122.2061	37.4179	SLAC	-122.2043	37.4165	0.21
JRSC	-122.2387	37.4037	JRSC	-122.2274	37.4062	1.07
HOPS	-123.0723	38.9935	HOPB	-123.0747	38.9952	0.30
SCCB	-121.8642	37.2874	LUTZ	-121.8652	37.2869	0.13
MHC	-121.6426	37.3416	MHCB	-121.6426	37.3415	0.04
JPSB	-122.3467	37.1986	P227	-121.7896	37.5330	1.89
JJO	-121.9773	37.2063	P213	-121.9908	37.2017	1.33
JUM	-121.8990	37.1607	SODB	-121.9255	37.1664	2.44
ORV	-121.5004	39.5545	ORVB	-121.5003	39.5546	0.03
CMB	-120.3865	38.0346	CMBB	-120.3860	38.0342	0.09
GASB	-122.7160	39.6547	GASB	-122.7159	39.6547	0.01
SAO	-121.4472	36.7640	SAOB	-121.4472	36.7653	0.14
PKD	-120.5416	35.9452	PKDB	-120.5416	35.9452	0.04
HATC	-121.4612	40.8161	HCRO	-121.4699	40.8159	0.75
MQ1P	-119.0213	37.6128	P630	-119.0004	37.6130	1.82
YBH	-122.7104	41.7320	YBHB	-122.7107	41.7317	0.08
MOD	-120.3030	41.9025	MODB	-120.3028	41.9023	0.02

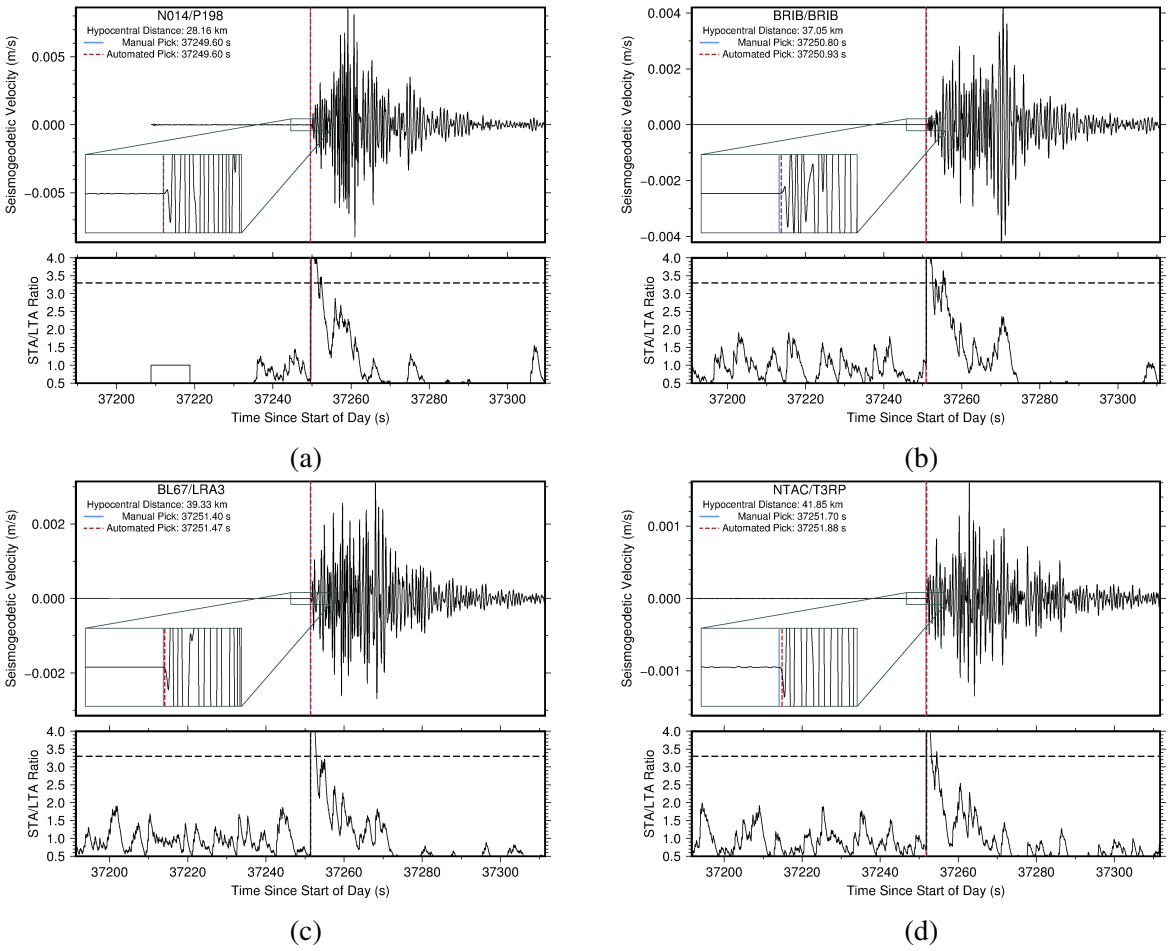


Figure S2.1: 2014 M_w6.0 Napa, California seismogeodetic P-wave detections. Short time average, long time average, and detection threshold are 1 second, 5 seconds, and 3.3, respectively. Top panel shows bandpassed seismogeodetic velocity with manual pick (blue) and automated pick (dashed red), with an inset highlighting the period near the P-wave arrival. Bottom panel shows the STA/LTA ratio, with horizontal dashed line indicating the detection threshold. For some of these stations, acceleration data was unavailable until soon before the event (subfigures S2.1a, S2.1e, S2.1i, S2.1k, S2.1l, S2.1m, S2.1o, and S2.1r). In these cases, the bottom panel shows a boxcar shape during algorithm initialization. Initialization is complete prior to the P-wave arrival in each case.

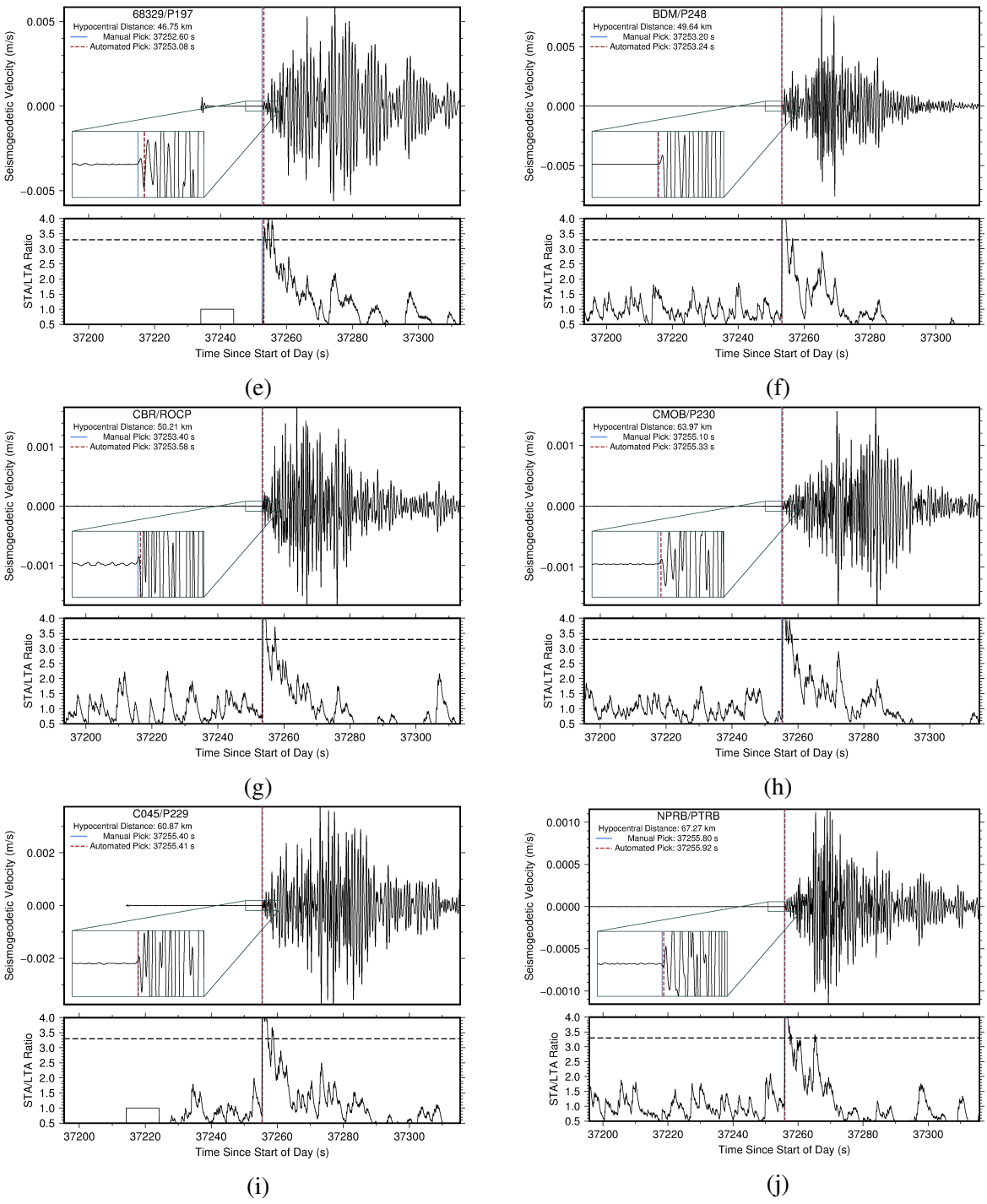


Figure S2.1: 2014 M_w6.0 Napa, California seismogeodetic P-wave detections, continued.

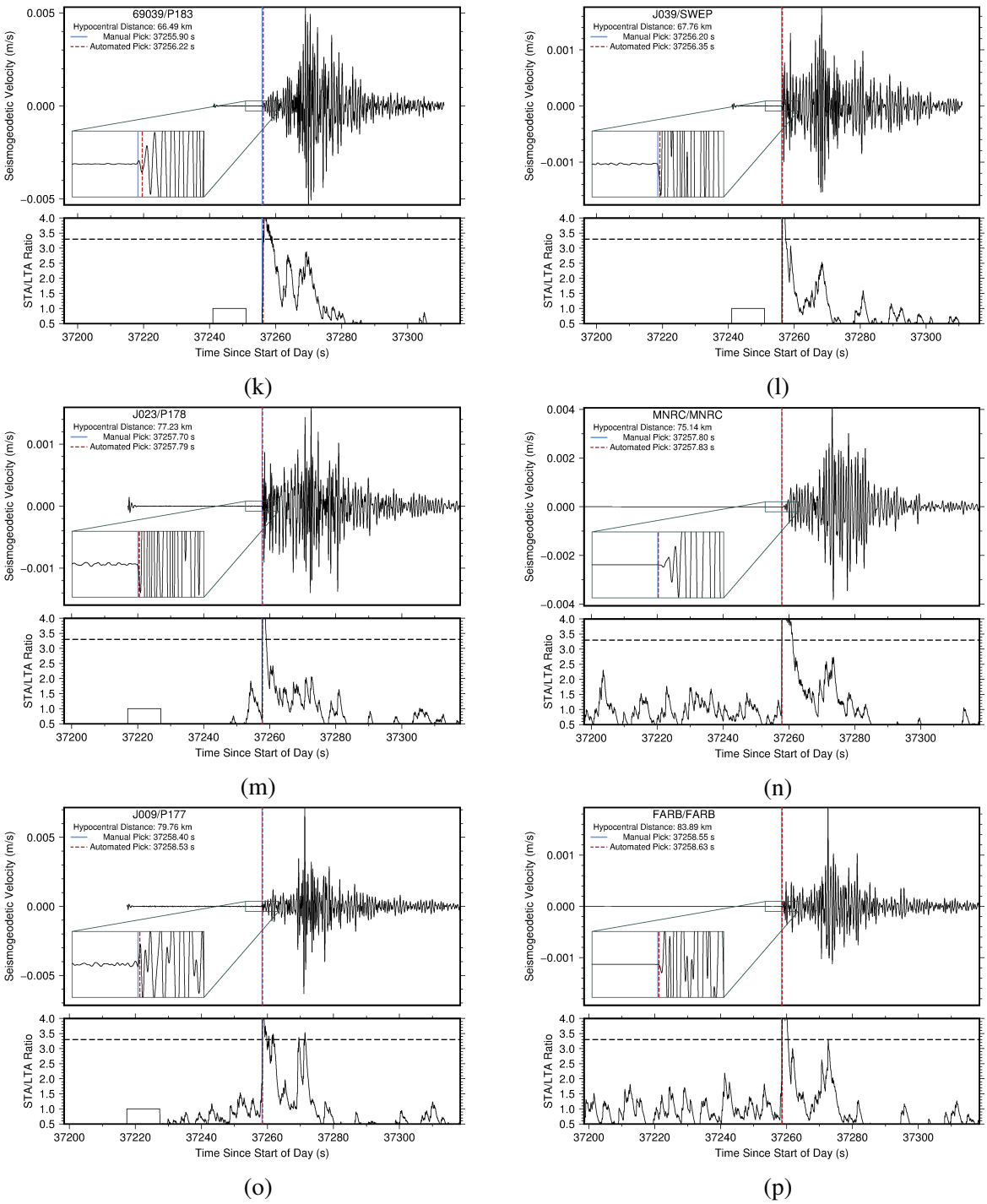


Figure S2.1: 2014 M_w6.0 Napa, California seismogeodetic P-wave detections, continued.

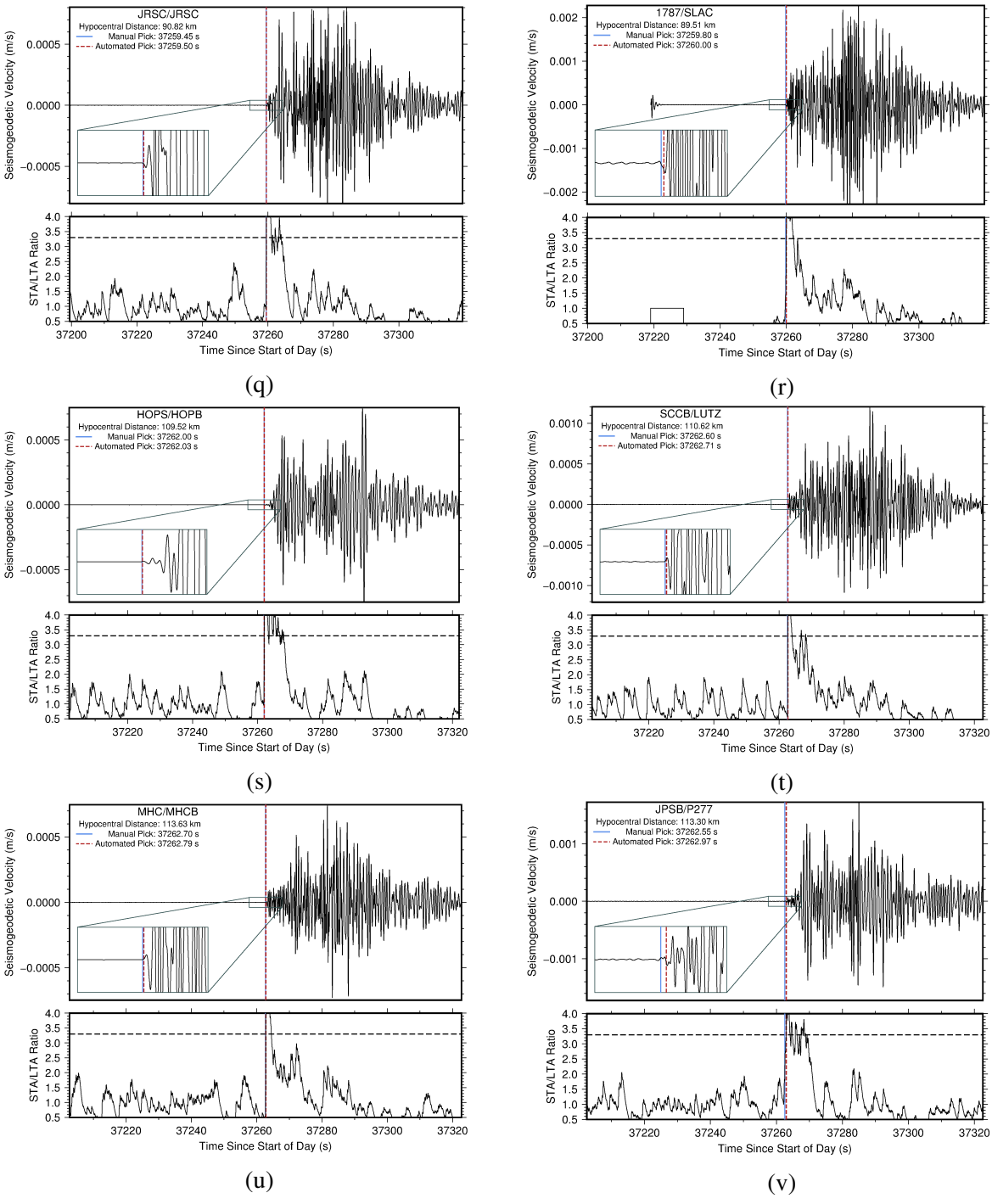
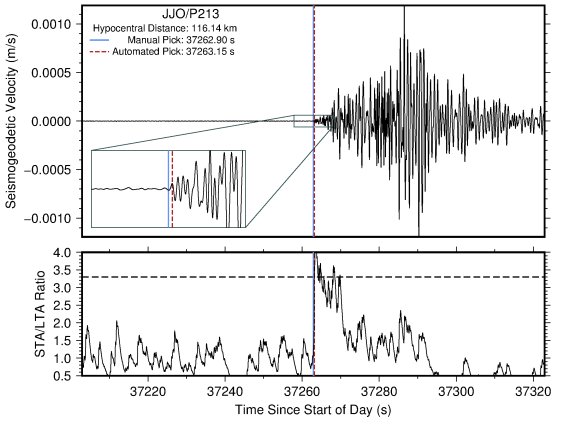
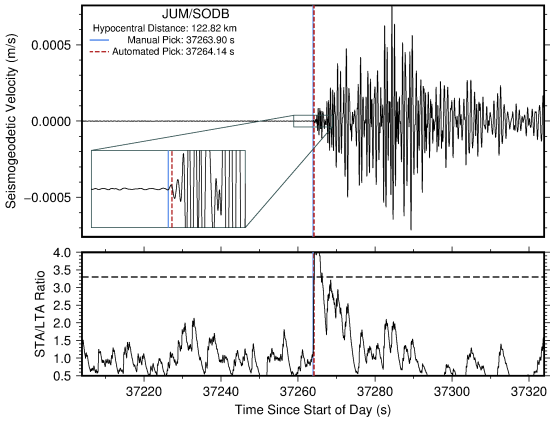


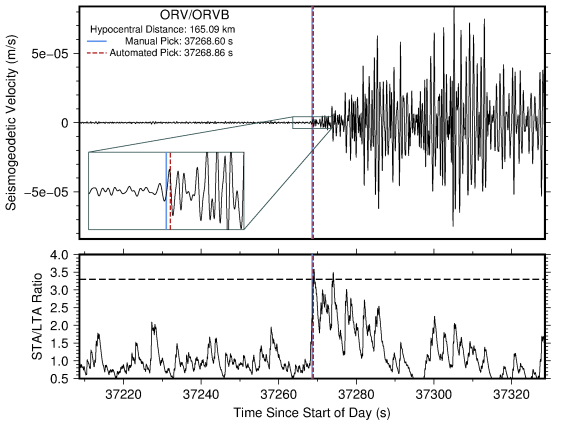
Figure S2.1: 2014 M_w6.0 Napa, California seismogeodetic P-wave detections, continued.



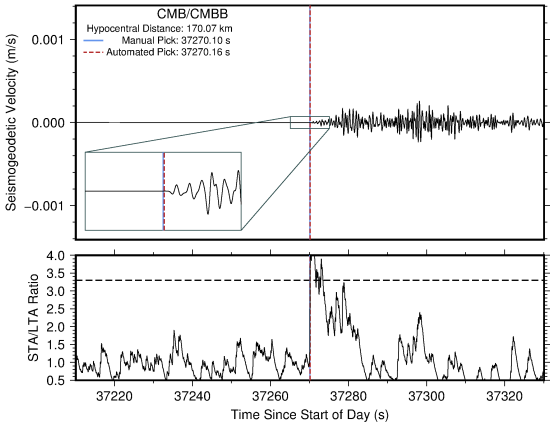
(w)



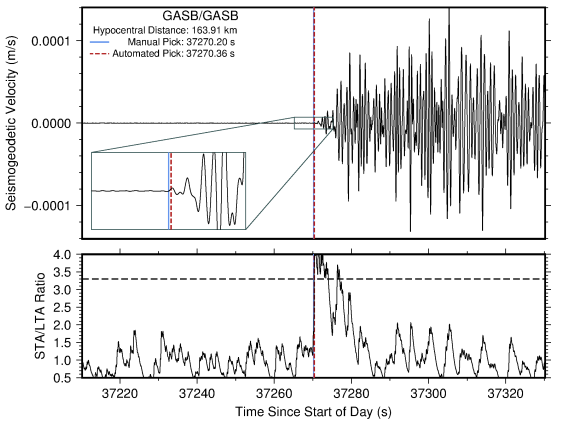
(x)



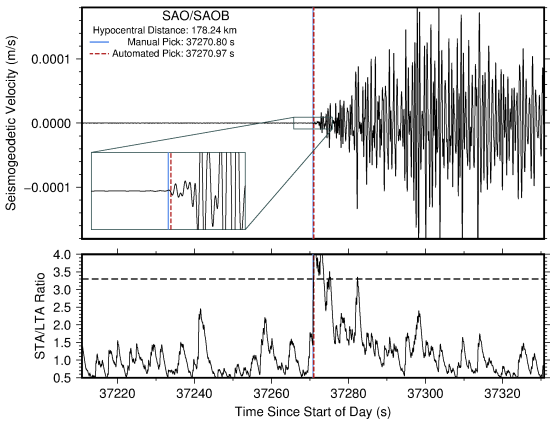
(y)



(z)

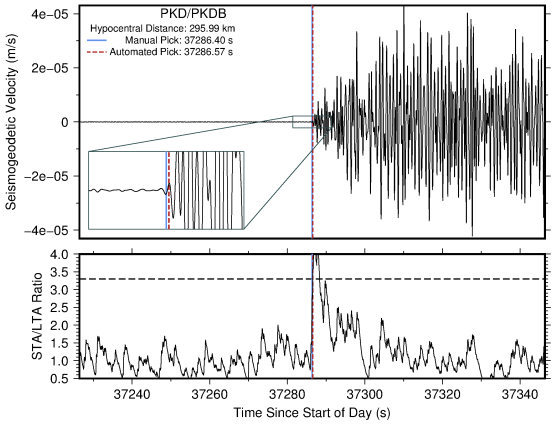


(aa)

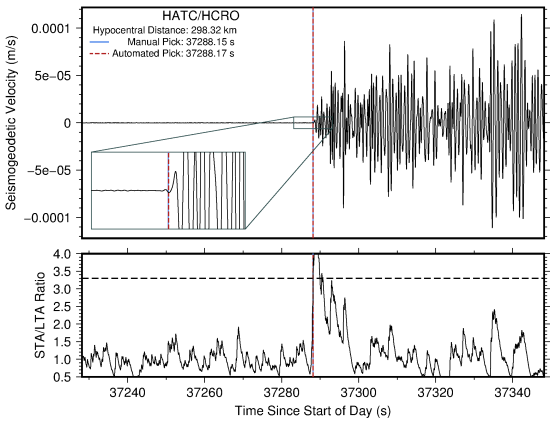


(ab)

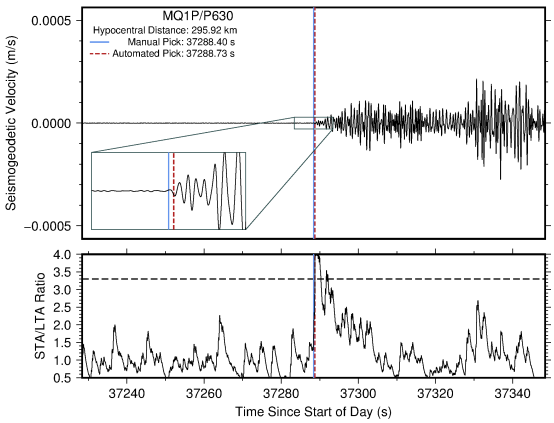
Figure S2.1: 2014 M_w6.0 Napa, California seismogeodetic P-wave detections, continued.



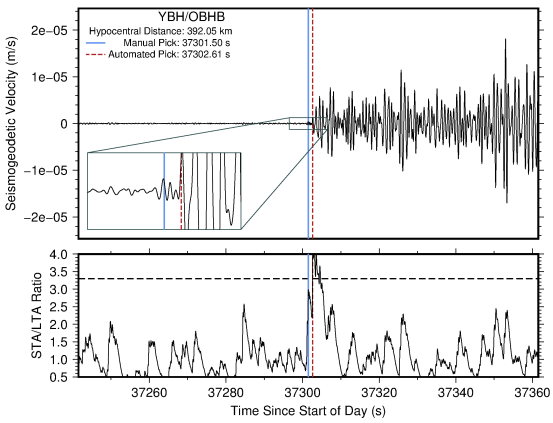
(ac)



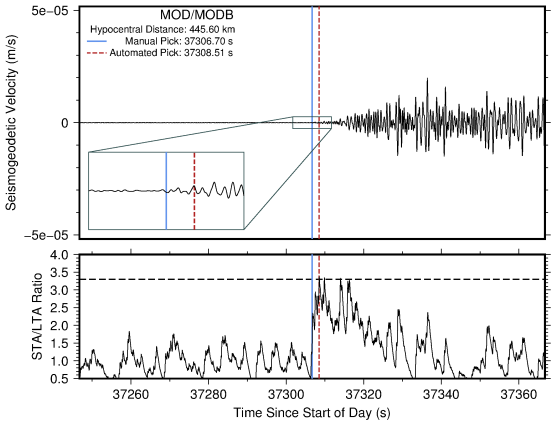
(ad)



(ae)

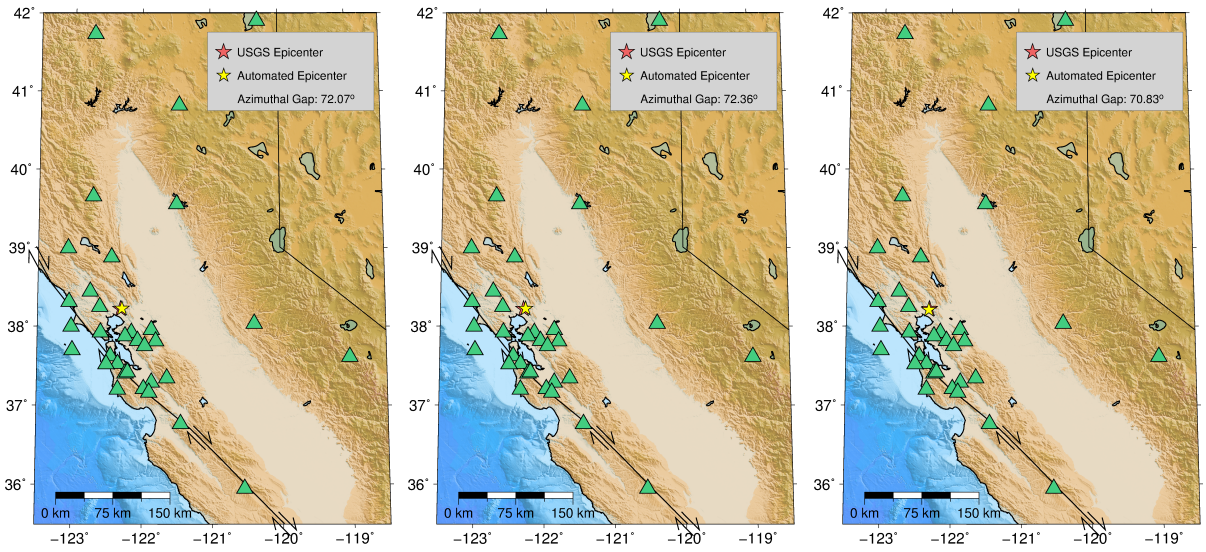


(af)



(ag)

Figure S2.1: 2014 M_w6.0 Napa, California seismogeodetic P-wave detections, continued.



(a) (b) (c)

Figure S2.2: 2014 M_w6.0 Napa, California epicenter location estimates and error ellipses (from bootstrapping procedure) for different solver types: (a) L1 minimization, (b) L2 (least squares) minimization, and (c) WL2 (weighted least squares) minimization. The small error ellipses are obscured by the epicenter symbols. See Table 2.1 for details.

2.6.2 2010 M_w7.2 El Mayor-Cucapah Supporting Information

Table S2.2: Collocated stations used to analyze the 2010 M_w7.2 El Mayor- Cucapah, Mexico earthquake.

Accelerometer	Longitude	Latitude	GPS	Longitude	Latitude	Separation (km)
WES	-115.7316	32.7590	P494	-115.7321	32.7697	0.08
GLA	-114.8271	33.0512	GMPK	-114.8273	33.0511	0.05
SAL	-115.9859	33.2801	SLMS	-115.9778	33.2922	1.50
BOM	-115.7296	33.3647	BOMG	-115.7297	33.3646	0.01
CTC	-115.9901	33.6551	CACT	-115.9900	33.6551	0.01
THM	-116.0773	33.6507	THMG	-116.0773	33.6506	0.01
SDR	-116.9424	32.7356	P473	-116.9495	32.7338	0.70
PLM	-116.8626	33.3536	PMOB	-116.8595	33.3572	0.50
MSC	-116.6480	34.0385	MSCG	-116.6480	34.0385	0.00
SNO	-116.8078	34.0352	SNOG	-116.8078	34.0352	0.01
SDD	-117.6617	33.5526	SBCC	-117.6615	33.5530	0.05
FMP	-118.2938	33.7126	VTIS	-118.2938	33.7126	0.00

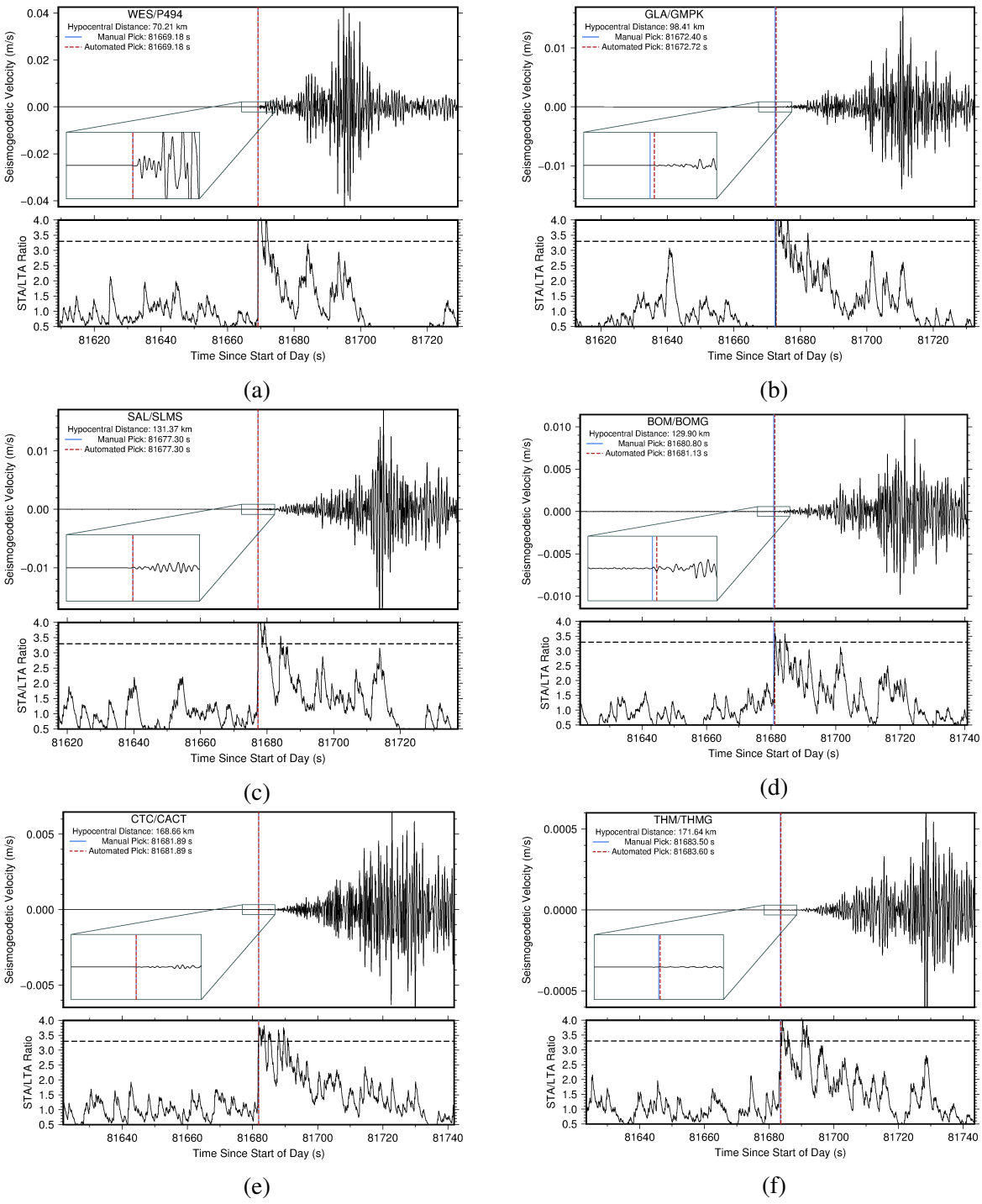


Figure S2.3: 2010 M_w7.2 El Mayor-Cucapah, Mexico seismogeodetic P-wave detections. Short time average, long time average, and detection threshold are 1 second, 5 seconds, and 3.3, respectively. Top panel shows bandpassed seismogeodetic velocity with manual pick (blue) and automated pick (dashed red), with an inset highlighting the period near the P-wave arrival. Bottom panel shows the STA/LTA ratio, with horizontal dashed line indicating the detection threshold.

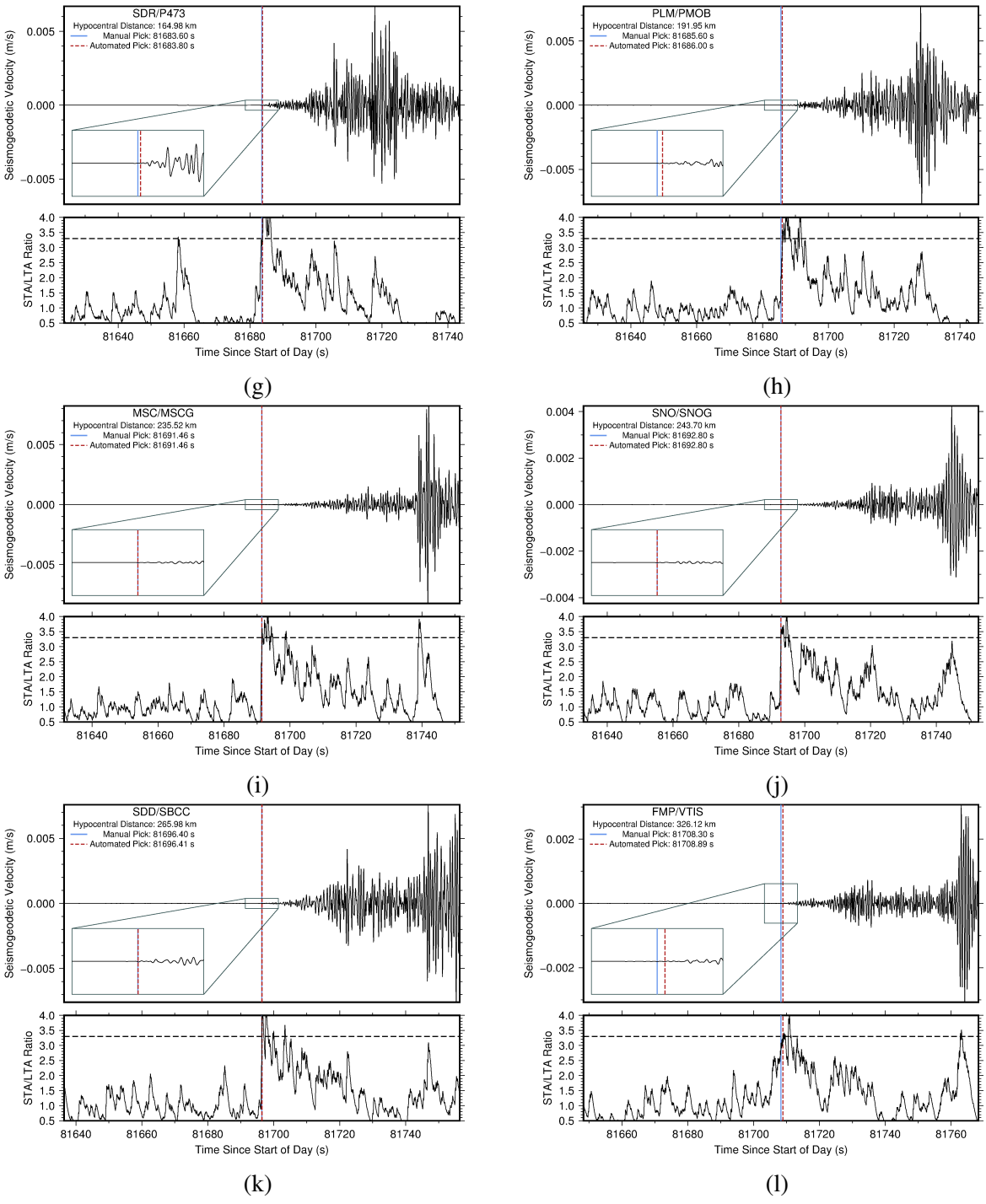


Figure S2.3: 2010 M_w7.2 El Mayor-Cucapah, Mexico seismogeodetic P-wave detections, continued.

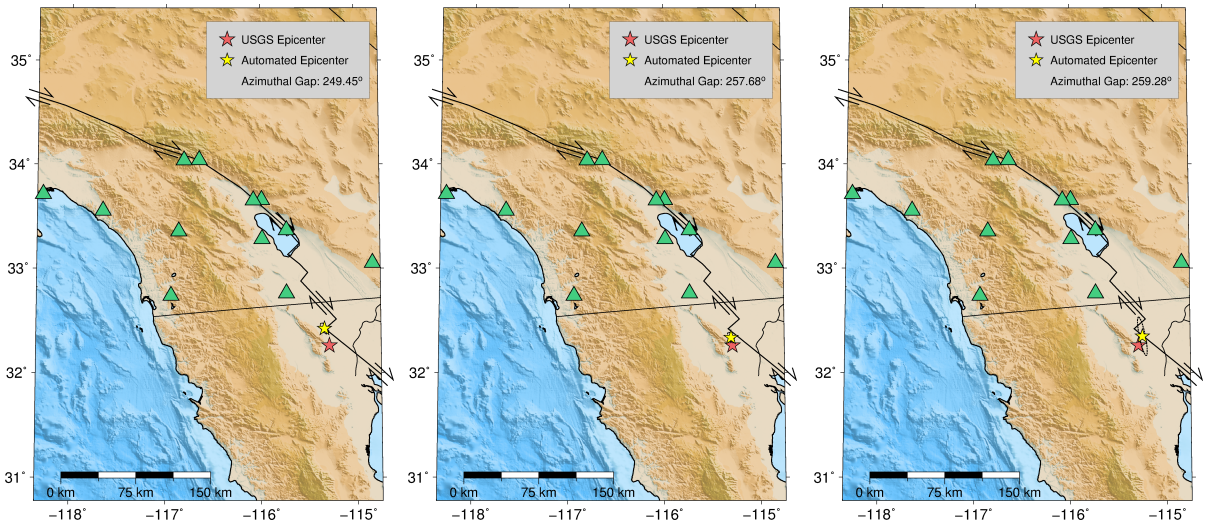


Figure S2.4: 2010 M_w7.2 El Mayor-Cucapah, Mexico epicenter location estimates and error ellipses (from bootstrapping procedure) for different solver types: (a) L1 minimization, (b) L2 (least squares) minimization, and (c) WL2 (weighted least squares) minimization. The small error ellipses (especially in (b)) are obscured by the epicenter symbols. See Table 2.2 for details.

2.6.3 2014 M_w8.2 Iquique Supporting Information

Table S2.3: Collocated stations used to analyze the 2014 M_w8.2 Iquique, Chile earthquake.

Accelerometer	Longitude	Latitude	GPS	Longitude	Latitude	Separation (km)
PSGCX	-70.1231	-19.5972	PSGA	-70.1230	-19.5974	0.02
PB11	-69.6558	-19.7610	PB11	-69.6557	-19.7613	0.04
PB08	-69.1534	-20.1411	PBO8	-69.1609	-20.1430	0.82
TA01	-70.1807	-20.5656	AEDA	-70.1777	-20.5462	2.17
PB01	-69.4874	-21.0432	PB01	-69.4879	-21.0435	0.06
PB02	-69.8960	-21.3197	PB02	-69.8928	-21.3153	0.60
PB07	-69.8862	-21.7267	PB07	-69.8860	-21.7269	0.03
PB03	-69.7531	-22.0485	PB03	-69.7520	-22.0489	0.13
PB04	-70.1492	-22.3337	PB04	-70.1497	-22.3350	0.16
PB06	-69.5719	-22.7058	PB06	-69.5721	-22.7060	0.03
PB05	-70.2024	-22.8528	PB05	-70.2030	-22.8532	0.08

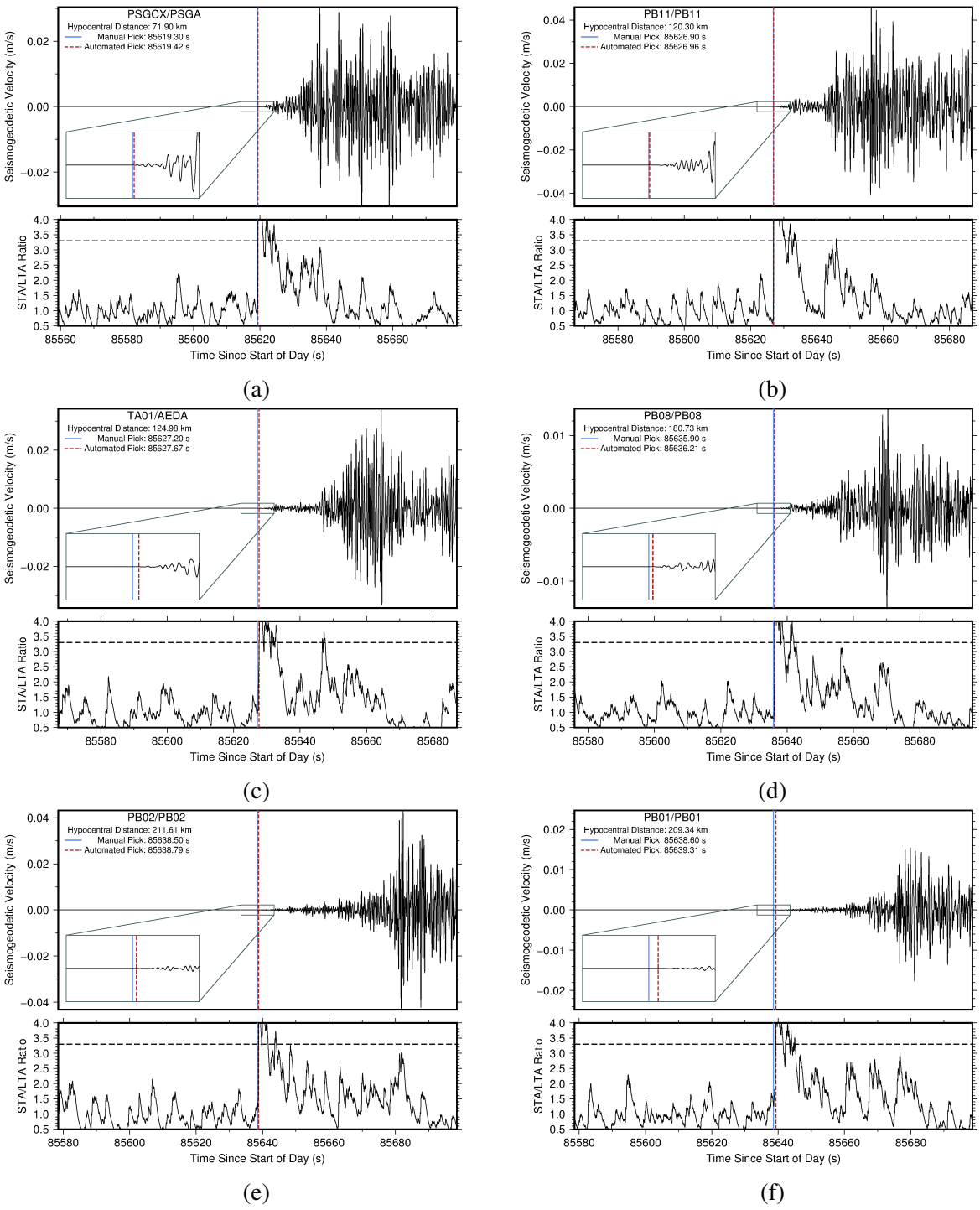
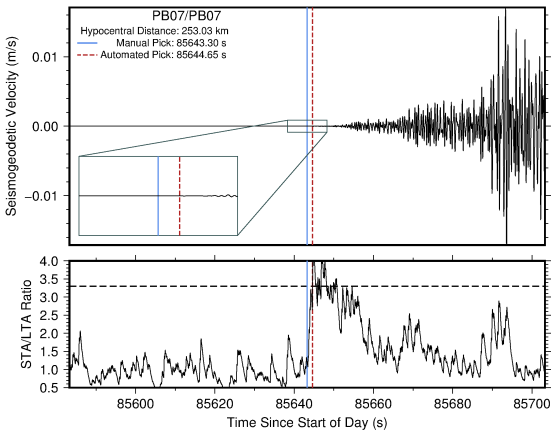
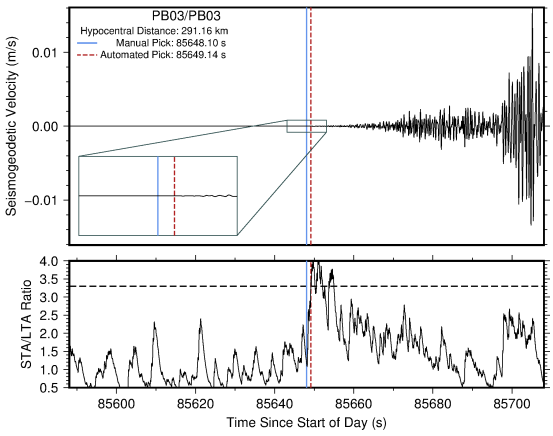


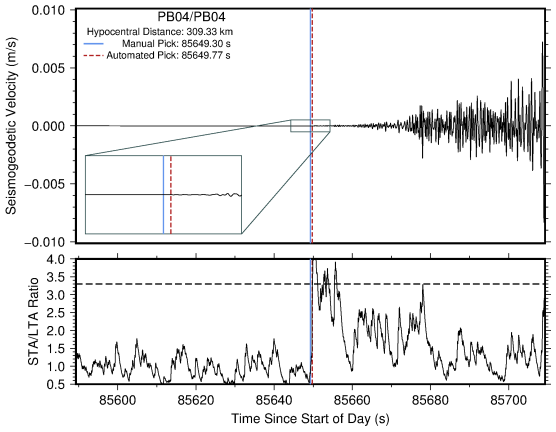
Figure S2.5: 2014 M_w8.2 Iquique, Chile seismogeodetic P-wave detections. Short time average, long time average, and detection threshold are 1 second, 5 seconds, and 3.3, respectively. Top panel shows bandpassed seismogeodetic velocity with manual pick (blue) and automated pick (dashed red), with an inset highlighting the period near the P-wave arrival. Bottom panel shows the STA/LTA ratio, with horizontal dashed line indicating the detection threshold. Station MNMCX, shown in S2.9a was rejected due to continuous detections from errant signals in the accelerometer.



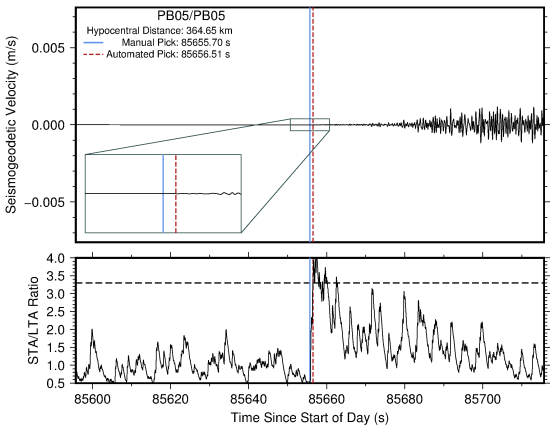
(g)



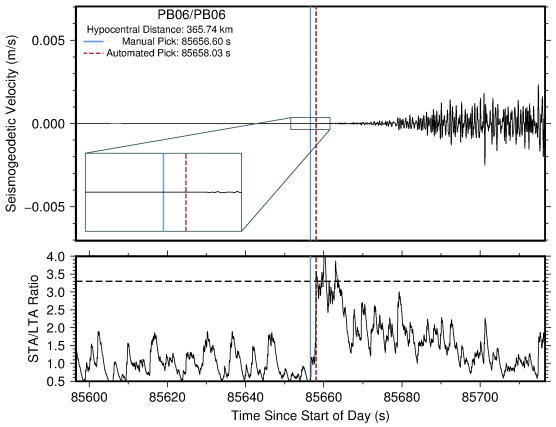
(h)



(i)

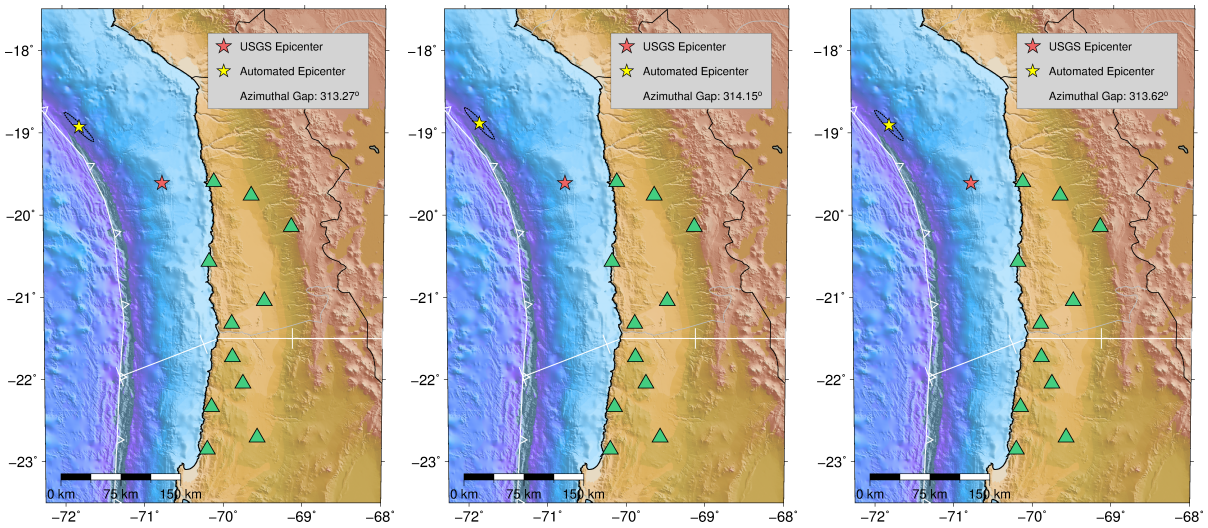


(j)



(k)

Figure S2.5: 2014 M_w8.2 Iquique, Chile seismogeodetic P-wave detections, continued.



(a) (b) (c)

Figure S2.6: 2014 M_w8.2 Iquique, Chile epicenter location estimates and error ellipses (from bootstrapping procedure) for different solver types: (a) L1 minimization, (b) L2 (least squares) minimization, and (c) WL2 (weighted least squares) minimization.

2.6.4 June 10, 2016 M_w5.2 Borrego Springs Supporting Information

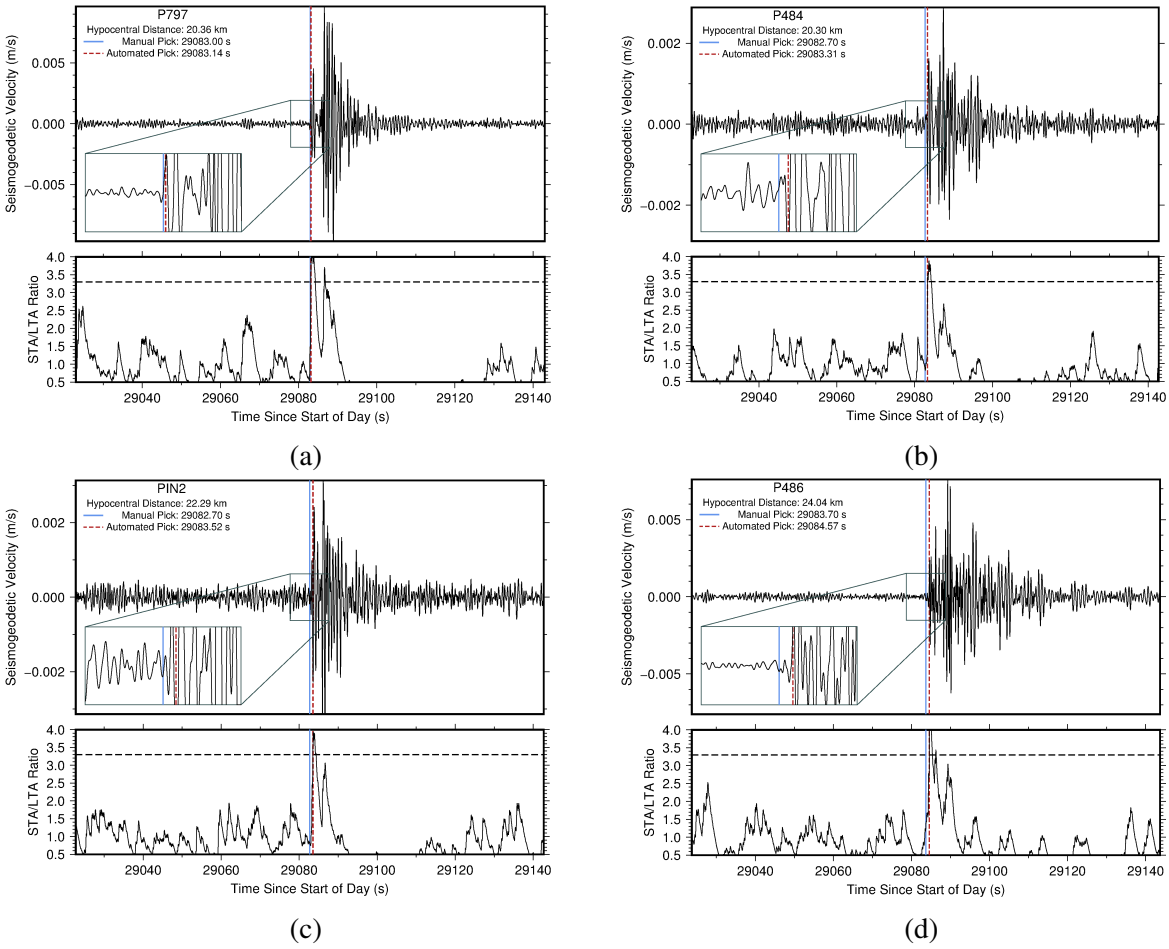


Figure S2.7: 2016 M_w5.2 Borrego Springs, California seismogeodetic P-wave detections. Short time average, long time average, and detection threshold are 1 second, 5 seconds, and 3.3, respectively. Top panel shows bandpassed seismogeodetic velocity with manual pick (blue) and automated pick (dashed red), with an inset highlighting the period near the P-wave arrival. Bottom panel shows the STA/LTA ratio, with horizontal dashed line indicating the detection threshold.

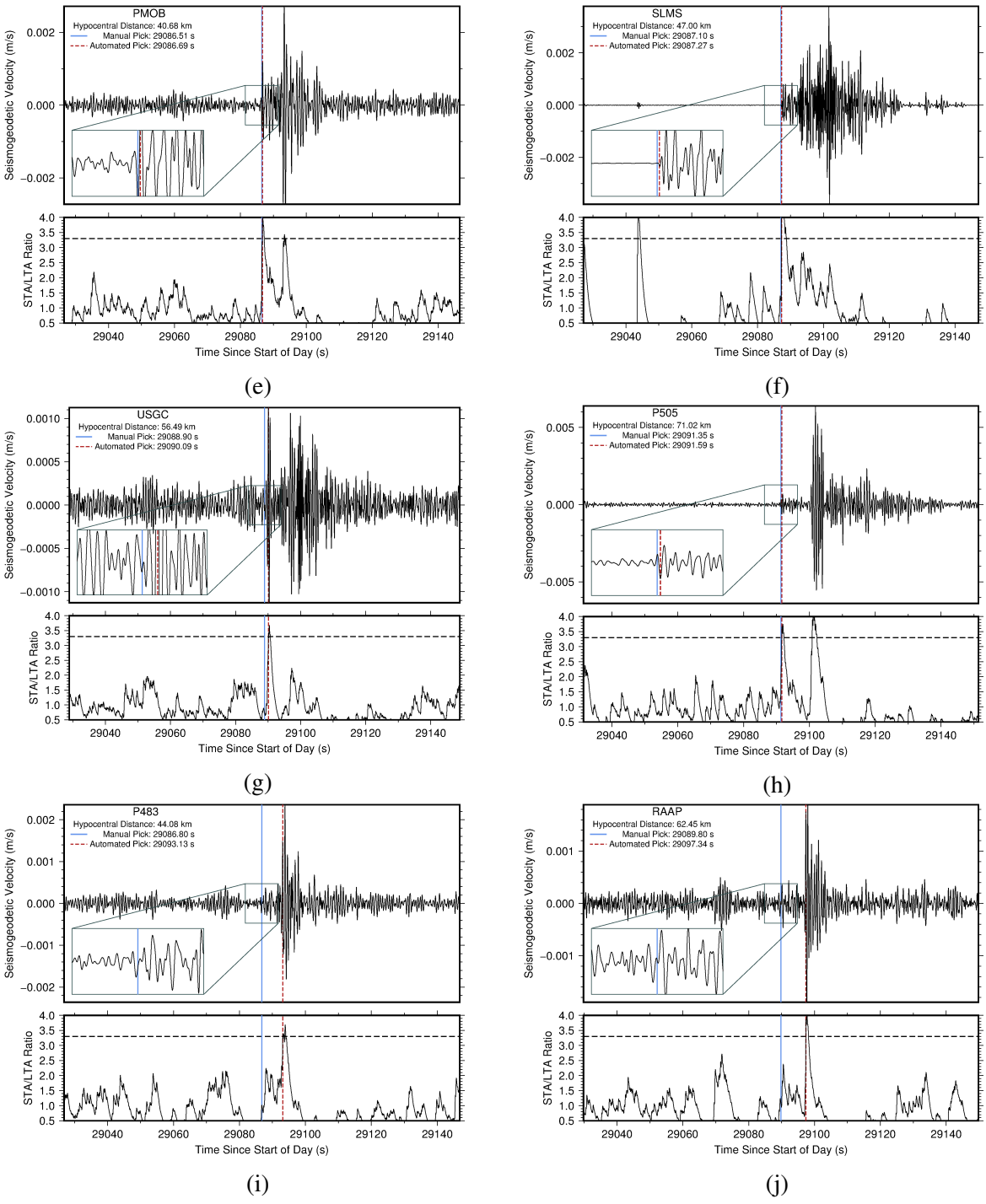


Figure S2.7: 2016 M_w5.2 Borrego Springs, California seismogeodetic P-wave detections, continued.

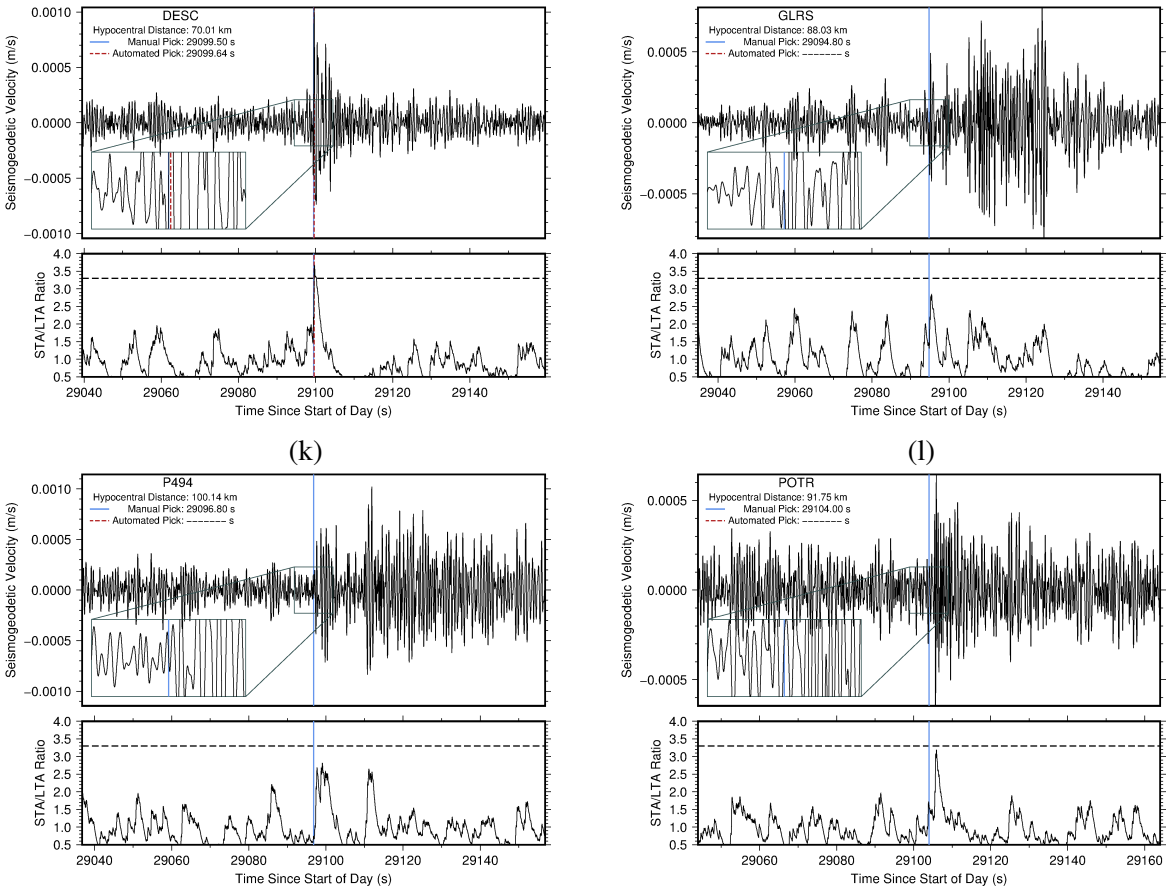
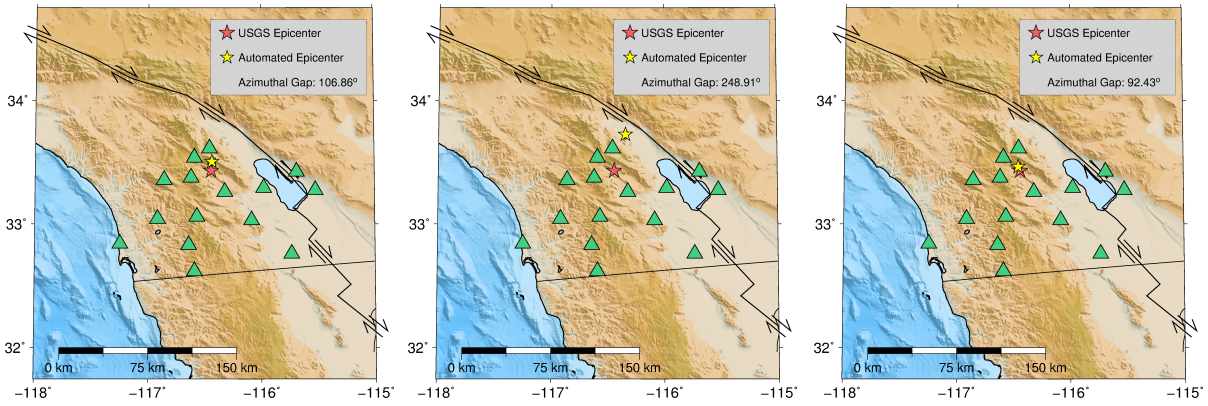


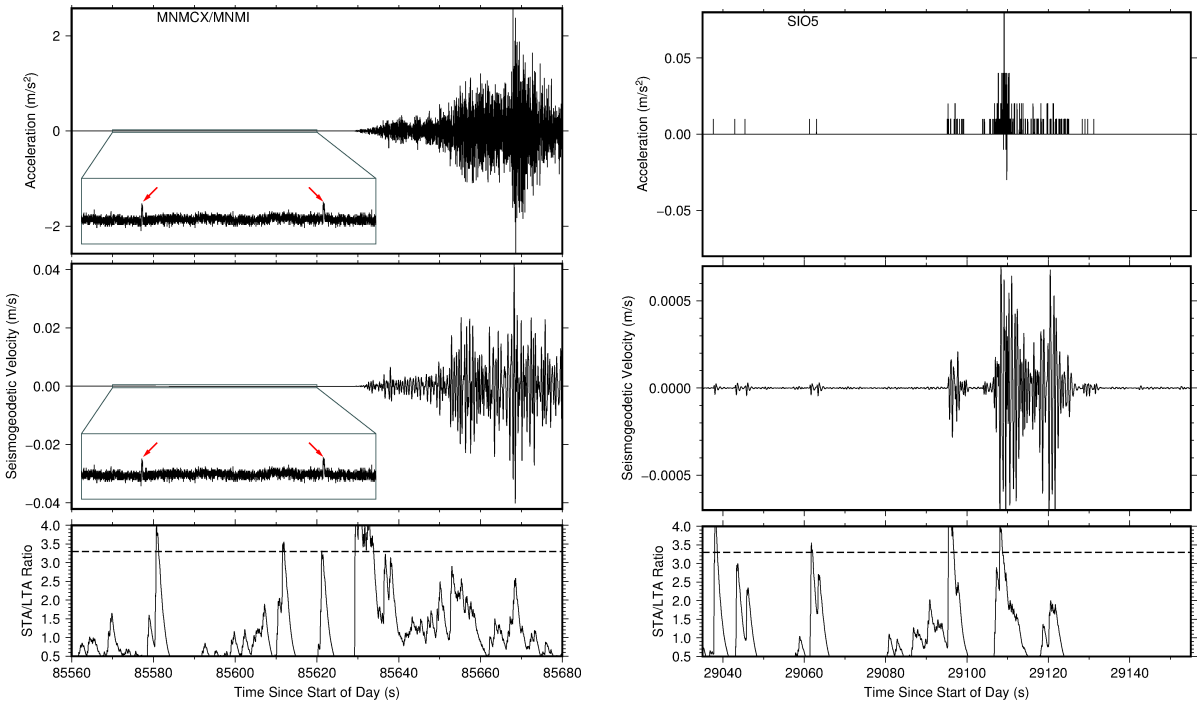
Figure S2.7: 2016 M_w5.2 Borrego Springs, California seismogeodetic P-wave detections, continued.



(a) (b) (c)

Figure S2.8: 2016 M_w5.2 Borrego Springs, California epicenter location estimates and error ellipses (from bootstrapping procedure) for different solver types: (a) L1 minimization, (b) L2 (least squares) minimization, and (c) WL2 (weighted least squares) minimization. The small error ellipses are obscured by the epicenter symbols. See Table 2.4 for details.

2.6.5 Problematic Stations



(a) (b)

Figure S2.9: Problematic stations, eliminated from analysis due to spurious recordings. (a) MNMCX, 136km from the Iquique, Chile hypocenter had small, impulsive signals throughout the time series which repeatedly crossed the threshold for detection. (b) Station SIO5, 100 km from the Borrego Springs event experienced a malfunction in the MEMS accelerometer which recorded 0 m/s^2 (no noise) spuriously for portions of the time series, and appears one-sided (only positive accelerations recorded) which resulted in a wildly fluctuating STA/LTA ratio. Every non-zero recording resulted in a sharp increase to the STA/LTA ratio which would have caused false detections had it been included in the analysis.

2.7 Acknowledgements

This work is published in **Goldberg, D.E., and Bock, Y. (2017). Self-contained local broadband seismogeodetic early warning system: Detection and location.** ***Journal of Geophysical Research: Solid Earth*****, 122(4):3197-3220. doi:10.1002/2016JB013766**

The dissertation author was the primary investigator and author of this material. We thank two anonymous reviewers for their thoughtful comments. Thanks to Peng Fang, Jennifer Haase, Allen Nance, Glen Offield, Jessie Saunders, Anne Sullivan, and Melinda Squibb at SOPAC, and Chris Walls, Doerte Mann, and Glen Mattioli at UNAVCO for their contributions to this research. Engineering services were provided by the EarthScope Plate Boundary Observatory operated by UNAVCO, and supported by the National Science Foundation (EAR-0732947 and EAR-1261833 for the GAGE Facility). We gratefully acknowledge funding support from the National Aeronautics and Space Administration grants ROSES Earth Science Applications: Disasters NNX14AT33G, ROSES Earth Surface and Interior NNX14AQ53G and Earth Science Research Program, ACCESS NNX16AM04A; the Southern California Earthquake Center, University of Southern California NSF and USGS subaward Y865564. Figures were produced using Generic Mapping Tools (Wessel et al., 2013).

Chapter 3

Evaluation of Earthquake Determinism with Seismogeodetic Displacements

3.1 Background

Chapter 2 addressed the importance of robust earthquake detection from a seismogeodetic network and proper hypocenter location. Hypocentral distance (source-to-station distance) is a required variable for estimating magnitude. Tsunami evacuation plans are based on three earthquake source properties: 1. hypocenter location (see Chapter 2), 2. faulting mechanism, and 3. magnitude. Very generally, a tsunami would be expected if the earthquake is located offshore, the faulting mechanism is primarily thrust, and the magnitude is greater than $M_w 7$. The Japanese Meteorological Agency (JMA) uses a database of precomputed scenarios of different earthquake locations and magnitudes, selecting the most similar scenario from which to create a local tsunami warning (Tatehata, 1997; Ozaki, 2011). From prior knowledge about the tectonic setting, the faulting mechanism is commonly assumed. Rapid estimation of earthquake magnitude is therefore the major outstanding metric required for local tsunami early warning.

In regions with risk of large earthquakes, seismologists rely on strong-motion accelerom-

eters to record the full dynamic range of the ground motion (Trifunac and Todorovska, 2001). However, strong-motion accelerometers are limited in their ability to accurately describe ground displacements. First, they are unable to distinguish between rotational and translational motions, leading to spurious recordings of translations that are in fact due to torsional motion. Second, observations include small errors known as baseline offsets. For rapid magnitude estimation, acceleration waveforms are doubly integrated to displacements, which amplifies baseline offsets, introducing a quadratic drift. The simplest way to remove the errors in real time involves a high-pass filter (e.g Boore et al., 2002), which results in a physically realistically looking displacement waveform devoid of drift, but is inaccurate due to removal of the long-period information, which reduces the observed amplitude of displacement, and eliminates the record of permanent displacement (e.g. Melgar et al., 2015a). The magnitude of large, destructive earthquakes is therefore often underestimated– even after rupture is complete– because seismic instrumentation in the near-field is unable to properly capture the full broadband spectrum of ground motion.

Magnitude underestimation, referred to as saturation, inhibits the effectiveness of an early warning, particularly an evacuation order, because the damage and affected areas will be underestimated in turn. During the 2011 M_w9.1 Tohoku-oki earthquake offshore Japan, the early warning system operated by the Japan Meteorological Agency (JMA) worked as designed, disseminating information to the public about the rapidly modeled earthquake and tsunami hazard. Due to its reliance on typical seismic instrumentation, the system significantly underestimated the earthquake size. The event was rapidly estimated as a M_w8.1, 122 seconds after origin, upgrading to a M_w8.4 after 74 minutes, at which point the tsunami had already inundated parts of the coast. It wasn't until the earthquake was observed at teleseismic distances that it was ultimately upgraded to a magnitude 9 (Hoshiba and Ozaki, 2014), resulting in insufficient rapid delineation of the evacuation zone (Yun and Hamada, 2014).

In fact, the current limitation in the National Oceanic and Atmospheric Administration's (NOAA) Tsunami Warning System for local tsunamis is a rapid magnitude estimation (Paul

Whitmore, personal correspondence, NOAA National Tsunami Warning Center). Accurate and rapid magnitude estimation is critical for populations and infrastructure in the near-source region of large subduction zone earthquakes with risk of tsunamigenesis. In the Northwestern U.S., the Cascadia subduction zone is of particular concern. The most recent large earthquake in the area occurred in 1700, and the resulting tsunami was observed across the Pacific Basin in Japan, reaching 2-3m high in some locations (Satake et al., 1996). Geologic evidence suggests the region is due for another large magnitude event, but seismically, this zone is generally quiet. The location is of particular societal concern because it includes the highly populated coastal city of Seattle. Furthermore, the subduction zone is quite close to the shoreline, allowing only $\sim$15 minutes between earthquake origin time and initially coastal inundation, as compared to 30-60 minutes for coastal Japan after the 2011 Tohoku-oki event. To facilitate a fast and accurate warning, it is imperative to evaluate the true magnitude quickly.

Presently, the most accurate magnitude estimates are available only after rupture has completed. The earthquake rupture process is often described by the cascade model, in which slip on one patch will propagate to adjacent patches until there is insufficient energy to continue. The cascade model is consistent with the relationship of seismic moment, M_0, to shear modulus, μ, rupture area, A, and average slip, $\bar{D}$. It implies that rupture must complete before before $\bar{D}$ can be measured, such that earthquake magnitude, M_{w}, which scales with $\log_{10}M_0$, can be resolved.

$$M_0 = \mu A \bar{D} \tag{3.1}$$

$$M_{\mathrm{w}} = \frac{2}{3}\log_{10}M_0 + 9.1 \tag{3.2}$$

However, theories have been proposed to suggest that initial rupture observations can be used to predict the final rupture extent. For example, Zollo et al. (2006) suggests the possibility that the dynamic stress drop at rupture initiation (or initial rupture surface area) is proportional to the energy flow that advances rupture, and therefore initial rupture observations predict the

probability of rupture extent and eventual magnitude. Different qualities of early observations have been exploited to relate first observations to final magnitude: predominant period (Wu and Kanamori, 2005), maximum predominant period (Nakamura, 1988; Olson and Allen, 2005), displacement amplitude (Wu and Zhao, 2006), and other nucleation phase properties (Ellsworth and Beroza, 1995; Colombelli et al., 2012). However, observational support for this mechanism has not been entirely convincing, either because of limited datasets or because of evidence of magnitude saturation above ∼M7. Furthermore, other studies document evidence to the contrary; that early rupture properties are statistically indistinguishable, regardless of eventual magnitude (Rydelek and Horiuchi, 2006; Meier et al., 2016, 2017).

These studies are limited in that they have been conducted with traditional seismic instrumentation. Magnitude is a measure of slip, $\bar{D}$, which requires low-frequency observations of the earthquake, i.e. the 0 Hz static offset. However, seismic instrumentation is not well-designed to measure low frequencies. This problem is only exacerbated for large ground motions typical of great earthquakes, where early warning is arguably most important. Seismic observations are routinely high-pass filtered to eliminate drift, which effectively reduces the observed displacement amplitudes and limits the measurement of predominant period. The result is that the drift-corrected waveforms of large magnitude earthquakes cannot be distinguished from those of earthquakes of lesser magnitude.

Recently, Global Navigation Satellite System (GNSS) measurements have gained traction for measuring the dynamic and permanent displacements resulting from medium to large earthquakes. GNSS instrumentation measures displacement directly, which avoids the integration issues that plague traditional seismic measurements and allows accurate representation of ground motions regardless of size. However, GNSS has lower sensitivity than seismic instrumentation, and typically much lower sampling rates. Therefore, it is not sensitive to very small ground motions at the start of the seismic record. Seismogeodesy, the optimal combination of collocated seismic and geodetic data, provides the most favorable dataset for exploring the subtleties of early

observations (Bock et al., 2011; Crowell et al., 2013). The combination dataset has the temporal resolution of the seismic instrumentation, and results in a broadband displacement time series that is more accurate than that using integrated broadband velocities or strong-motion accelerations, but has reduced noise compared to GNSS-only, improving the sensitivity. Importantly, the seismogeodetic combination is also sensitive enough to detect seismic P-wave arrivals, improving the timeliness of a warning (Goldberg and Bock, 2017).

Crowell et al. (2013) first noted that while strong-motion derived P-wave amplitude measurements are insufficient to distinguish the magnitude of large earthquakes, the seismogeodetic solution results in a more accurate measurement of P-wave displacement amplitude and thus suggested this metric may in fact be indicative of final magnitude. That study demonstrated that the displacements derived from doubly-integrated strong-motion records can result in either over- or under-estimations of displacement amplitude due to baseline drift or removal of the long-period signal through filtering. This principle study was limited to a dataset of five globally distributed events, including two events with only three observing stations, as few observations of medium to large earthquakes at collocated strong-motion and GNSS stations were available.

Availability of new seismogeodetic datasets now allows us to investigate this P-wave amplitude scaling metric more fully. Here, we present the results from 14 medium to large magnitude events (M_w5.7-9.1) in Japan, each observed by at least 24 collocated GNSS and strong-motion seismic stations. We explore the feasibility of earthquake magnitude estimation from early onset observations, and discuss the implications of our results on the open question of earthquake determinism. We later investigate the temporal evolution of maximum displacement amplitude, test the the ability to continuously estimate magnitude throughout the observation of rupture, and discuss best practices for early warning.

3.2 Dataset

We address the possibility of using P-wave amplitude as a magnitude-predictive metric using the dense strong-motion and GNSS networks available in Japan. The strong-motion acceleration data are available from National Research Institute for Earth Science and Disaster Resilience (NIED, www.kyoshin.bosai.go.jp) KiK-net and K-NET datasets. GNSS data was retrieved from the Geospatial Information Authority (GIS, www.gsi.go.jp), which operates the GNSS Earth Observation Network System (GEONET). We consider strong-motion and GNSS station pairs to be collocated if they are located within 2 km of one another. In Japan, more than 300 such station pairs exist. This inter-instrument distance is a conservative criterion, based upon previous work by Emore et al. (2007) demonstrating agreement between instrument pairs up to 4 km separation.

We processed the high-rate (1 Hz) GEONET observations using precise point positioning (PPP) (Geng et al., 2013; Zumberge et al., 1997). Although we will use the term "GNSS" to describe the GEONET observations, note that our processing ingests only the Global Positioning System (GPS) observations to derive the displacement time series. The processed GNSS data were then merged with the strong-motion accelerations from collocated KiK-net/K-NET sites using the Kalman-filter method of Bock et al. (2011) to create the seismogeodetic solution. The hypocentral distance, R, of each source-station pair was computed from the hypocenter location determined by the NIED (http://www.fnet.bosai.go.jp). The P-wave arrival times were selected manually from the vertical strong-motion waveforms. The Japanese national strong-motion networks of $\sim$1700 stations are operated in triggered-mode. Therefore, for the earthquakes we studied, some stations begin recording after the first seismic wave arrivals. If a record appears to begin after the first seismic wave arrivals, the station pair is eliminated from the analysis. As a result, the number of collocated stations used to study each earthquake is reduced to between 24 and 84 sites (Table 3.1).

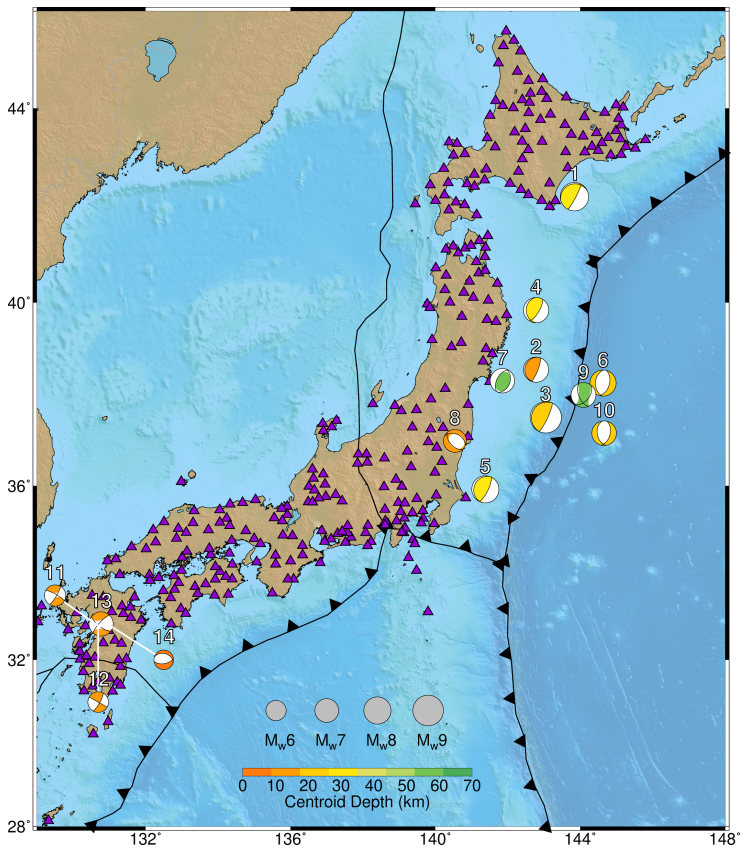


Figure 3.1: Seismogeodetic station locations in Japan. Location of collocated stations pairs (purple triangles) and earthquakes (colored circles) in the analyzed dataset. Circle size corresponds to magnitude and color corresponds to hypocenter depth.

Figure 3.1 shows the locations of the seismogeodetic stations (purple triangles) and the 14 earthquakes included in the study, with details in Table 3.1. The earthquakes occurred between 2003 and 2016, with magnitudes of M_w5.7 to M_w9.1. The seismogeodetic combination has a reduced noise level in comparison to GNSS alone, and therefore is suitable for observation of the earliest seismic signals.

3.3 Method

We concentrate on the maximum P-wave amplitude metric for magnitude scaling. We examine observation periods of 3 and 5 seconds after P-wave arrival time, and consider peak

Table 3.1: Earthquake event information for the 14 events included in the seismogeodetic dataset analysis. Earthquake origin time and hypocenter location are from NIED (http://www.fnet.bosai.go.jp). Origin time is given as year-month-day hour:minute:second. Magnitude is from Global CMT (globalCMT.org).

Location	M_w	Origin Time (UTC)	Hypocenter			No.
			Lon (°E)	Lat (°N)	Depth (km)	Stations
Tokachi-oki[1]	8.3	2003-09-25 19:50:07.64	144.0785	41.7797	42.0	65
Miyagi[2]	7.3	2011-03-09 02:45:12.97	143.2798	38.3285	8.3	43
Tohoku-oki[3]	9.1	2011-03-11 05:46:18.12	142.3730	38.2970	29.0	54
Iwate[4]	7.4	2011-03-11 06:08:53.05	142.7815	39.8390	31.7	24
Ibaraki[5]	7.9	2011-03-11 06:15:34.46	141.2653	36.1083	43.2	25
N. Honshu[6]	7.6	2011-03-11 06:25:44.04	144.8940	37.8367	34.0	64
Miyagi[7]	7.1	2011-04-07 14:32:43.43	141.9237	38.2028	60.7	84
E. Fukushima[8]	6.7	2011-04-11 08:16:12.02	140.6727	36.9457	6.4	27
N. Honshu[9]	7.2	2012-12-07 08:18:20.28	144.3153	37.8158	46.0	60
N. Honshu[10]	7.1	2013-10-25 17:10:18.39	144.5687	37.1963	56.0	25
Kumamoto[11]	6.2	2016-04-14 12:26:34.43	130.8087	32.7414	11.4	43
Kumamoto[12]	6.0	2016-04-14 15:03:46.45	130.7777	32/7007	6.7	34
Kumamoto[13]	7.0	2016-04-15 16:25:05.47	130.7630	32.7545	12.5	59
Kumamoto[14]	5.7	2016-04-15 16:45:55.45	130.8990	32.8632	10.6	28

horizontal amplitude, P_h, and peak vertical amplitude, P_v, as well as the peak three-component sum of squares, P_3, where

$$\begin{aligned} P_{\mathrm{h}} &= \max\left(\sqrt{\mathrm{N}(t)^2 + \mathrm{E}(t)^2}\right) \\ P_{\mathrm{v}} &= \max\left(\sqrt{\mathrm{U}(t)^2}\right) \\ P_3 &= \max\left(\sqrt{\mathrm{N}(t)^2 + \mathrm{E}(t)^2 + \mathrm{U}(t)^2}\right) \end{aligned} \tag{3.3}$$

where N(t), E(t), and U(t) are the north-south, east-west, and up-down component displacement waveforms, respectively. We consider search windows of $0 < t \leq 3$ and $0 < t \leq 5$ where 0 is the P-wave arrival time at each individual station.

We use the scaling laws proposed by Wu and Zhao (2006) and Crowell et al. (2013), Equation (3.4), which relates peak amplitude, P, to magnitude, M_w, and hypocentral distance, R, assuming that P attenuates linearly on a log-log scale. We apply a weighting matrix wherein

earthquakes are divided into magnitude bins ($M_w \leq 7$, $7 < M_w \leq 7.5$, $7.5 < M_w \leq 8$, $8 < M_w \leq 8.5$, $8.5 < M_w$), and weighted by the sum of the amplitudes in that bin. In doing so, each magnitude bin is equally weighted, regardless of the amount of data in the bin. The seismogeodetic combination using lower-quality MEMS accelerometers is sensitive down to 8 mm in displacement (Saunders et al., 2016). Our dataset is made up of observatory-grade geodetic and strong-motion instrumentation, and we consider any seismogeodetic displacement amplitude that exceeds 1 mm.

$$\log_{10}(P) = A + B \cdot M_w + C \cdot \log_{10}(R) \tag{3.4}$$

We solve for the regression coefficients, A, B, and C with an L1 norm, to reduce sensitivity to outliers. In order to quantify the uncertainties of our regression coefficients, we take a bootstrapping approach, since formal errors for L1 norm minimizing inversions cannot be calculated algebraically. We remove 10% of the stations at random and calculate the regression coefficients, repeating this process 100 times. From the 100 iterations, we define the best fitting coefficients A, B, and C as the mean $\pm$ one standard deviation. In each iteration, the unused 10% of the data are used to estimate the magnitude with the best fitting relation derived from the other 90%. The average difference between the known and estimated magnitudes of each event are recorded. The error in the magnitude estimate is described as one standard deviation of this difference between the known and estimated magnitudes from the unused 10% of data. The process is completed for two search windows, 3 and 5 seconds, following the P-wave arrival at an individual station.

3.4 Results

3.4.1 Seismogeodetic Solution

Figure 3.2 shows an example of the seismogeodetic solution for collocated station pair MYG011/0050, the closest to the hypocenter of the 2011 M_w9.1 Tohoku-oki event, at a distance

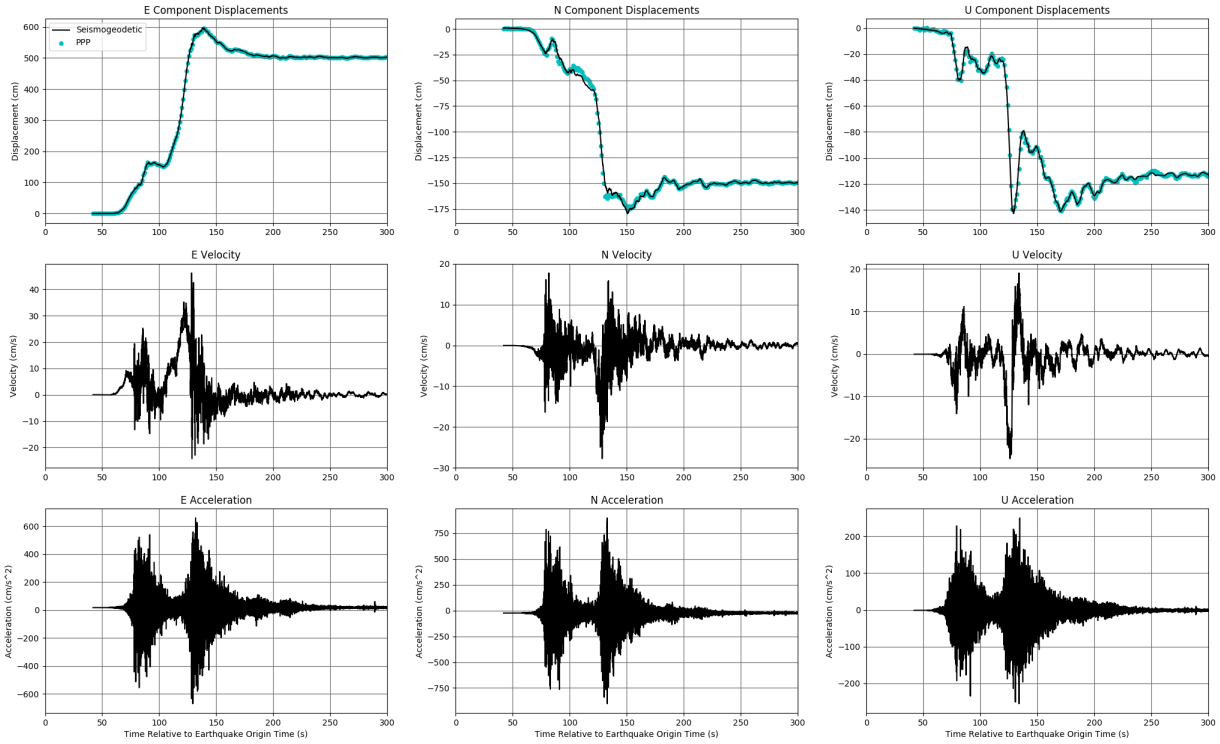


Figure 3.2: Example Kalman filter seismogeodetic combination of collocated GNSS instrument 0550 and strong-motion accelerometer MYG011 to achieve 3-component broadband velocity and displacement data at the temporal resolution of the accelerometer (100 Hz), informed by the long period information of the GNSS instrument. The result is immune to clipping, and retains the permanent (0 Hz) deformation in the displacement time series. Top: PPP solution (blue dots) and seismogeodetic displacement (black line), middle: seismogeodetic velocity, bottom: accelerations from strong motion instrument for East (left), North (center), and Up (right) components.

of 123.6 km. Note that station pairs are identified by the accelerometer station code/GNSS four-character station code. The seismogeodetic solution produces broadband velocity and displacement time series that are informed by the high-frequency content of the accelerometer and the long-period content of the GNSS instrument. The solution is immune to clipping, has the temporal resolution of the accelerometer (100 Hz), and follows the GNSS solution at long periods.

3.4.2 P-wave Amplitude

For an earthquake with a large rupture duration, like the M_w9.1 Tohoku-oki mainshock, the peak ground displacement is achieved almost 90 seconds after the P-wave arrival even at the closest available stations. This is the limiting factor for solution speed, and the ability to distinguish between magnitudes prior to achievement of peak displacement would be advantageous. Figure 3.3 shows an example of maximum P-wave amplitude selection for station pair MYG011/0550 during the Tohoku-oki event (see Figure 3.2). The top panels show the horizontal displacement amplitude (extended time series on the left, detail of first 10 seconds after P-wave arrival on the right), while the middle panel shows the same for the vertical amplitude and the bottom panel shows the same for the three-component sum of squares. Red dots on the right-hand panels show the selected P-wave amplitude for an example search window of 5 seconds after P-wave arrival. Since the vertical amplitudes are much smaller than horizontal in this case, the three-component sum of squares time series looks quite similar to the horizontal panel at this scale. Figure 3.4 shows the same information for station pair AOM021/0153. In this example, the influence of differing definitions of P-wave amplitude (P_h, P_v, or P_3 with 3 or 5 seconds of observation after the P-wave arrival) become clear. Unlike the previous example, the vertical component of station AOM021/0153 dominates the P-wave amplitude measurement. Furthermore, the timing of peak displacement within this 5 second window is notably different depending on the component of motion considered.

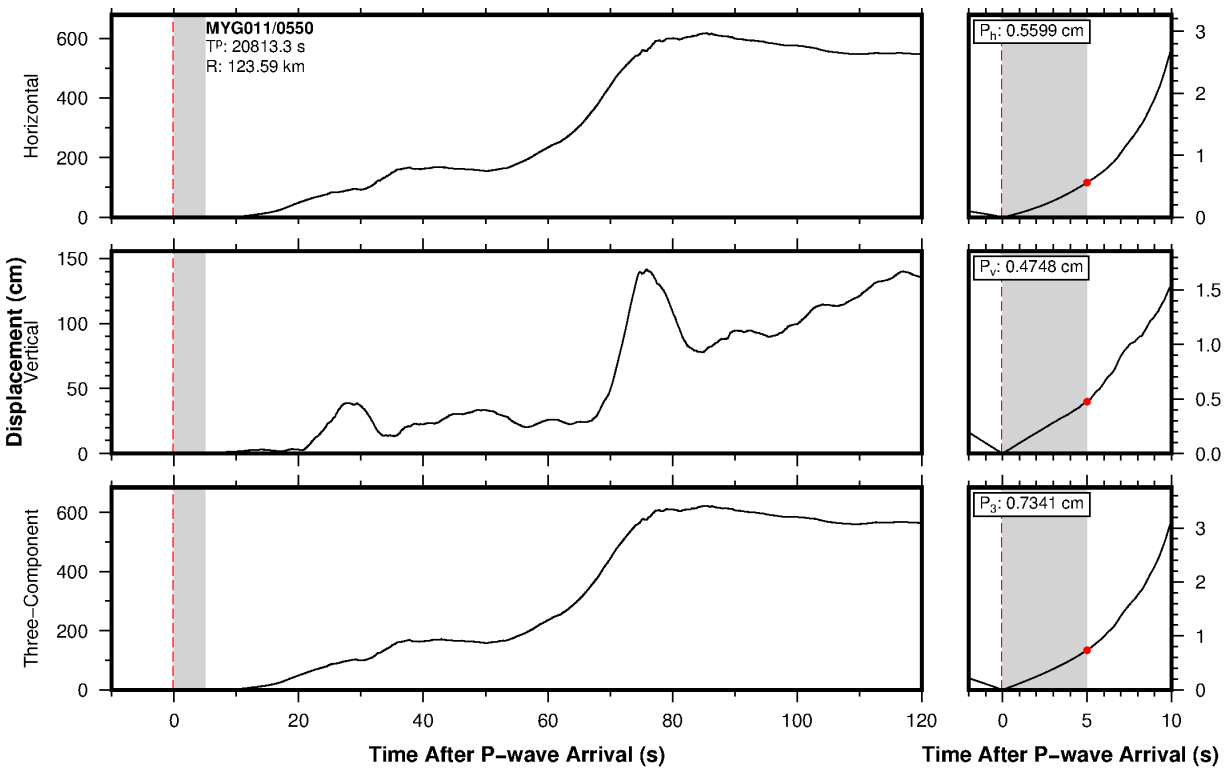


Figure 3.3: Example of maximum P-wave amplitude selection for a 5 second search window for station pair MYG011/0550 during the 2011 M_w9.1 Tohoku-oki earthquake. Left: Time series of horizontal (top), vertical (middle), and three-component (bottom) seismogeodetic displacement. Red dashed line denotes P-wave arrival time at the station. The grey shaded region is the allowed search window, shown here for 5 seconds after P-wave arrival. T^P is the P-wave arrival time in seconds after start of day and R is the hypocentral distance. Right: Detailed view of first 10 seconds of time series (same component as left), showing the P-wave amplitude selection. P_h, P_v, and P_3 are the horizontal, vertical, and three-component sum of squares amplitudes, respectively. Red dot denotes the maximum displacement value within the shaded region.

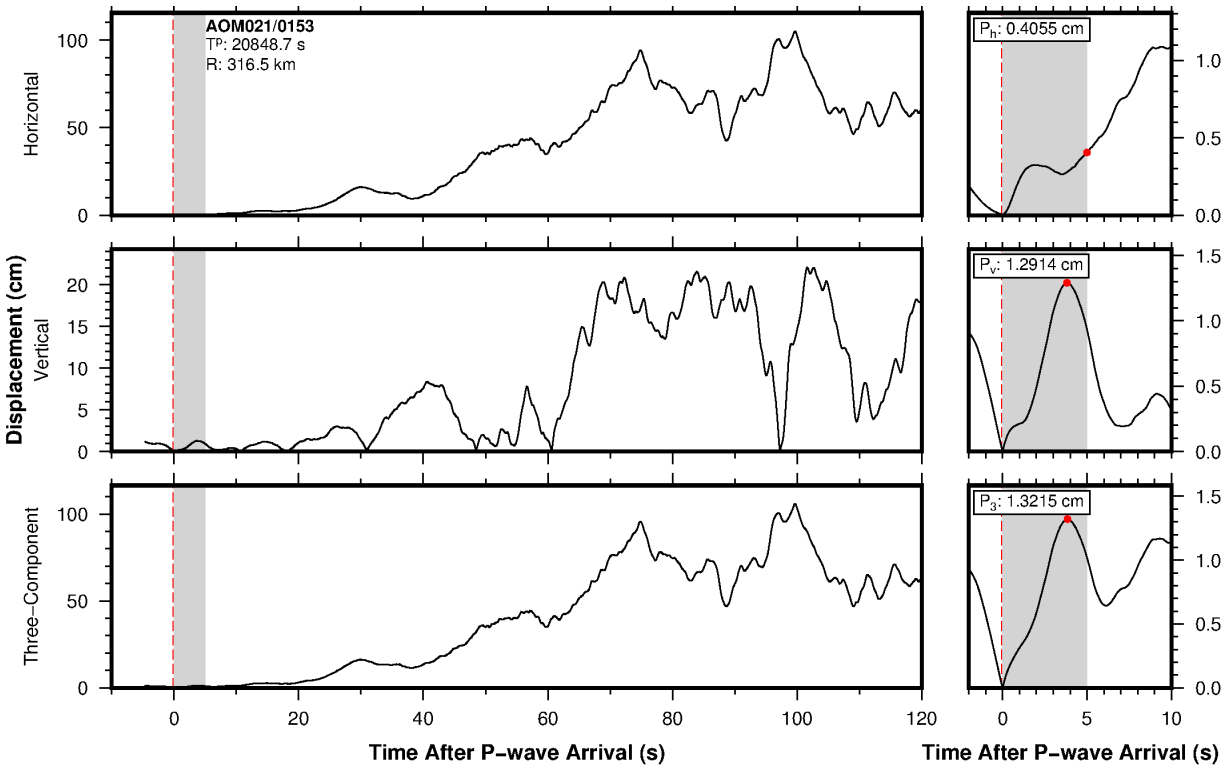


Figure 3.4: Example of maximum P-wave amplitude selection for a 5 second search window for station pair AOM021/0153 during the 2011 M_w9.1 Tohoku-oki earthquake. Left: Time series of horizontal (top), vertical (middle), and three-component (bottom) seismogeodetic displacement. Red dashed line denotes P-wave arrival time at the station. The grey shaded region is the allowed search window, shown here for 5 seconds after P-wave arrival. T^P is the P-wave arrival time in seconds after start of day and R is the hypocentral distance. Right: Detailed view of first 10 seconds of time series (same component as left), showing the P-wave amplitude selection. P_h, P_v, and P_3 are the horizontal, vertical, and three-component sum of squares amplitudes, respectively. Red dot denotes the maximum displacement value within the shaded region.

3.4.3 Magnitude Scaling Relation

Table 3.2: Summary of P-wave amplitude magnitude scaling best fitting relation using Equation 3.4 for seismogeodetic P_h, P_v, and P_3 in search windows of 3 and 5 seconds after P-wave arrival.

	Window	A$\pm\sigma_A$	B$\pm\sigma_B$	C$\pm\sigma_C$	M_w Error
P_h	3 s	-2.925 $\pm$ 0.048	0.180 $\pm$ 0.006	-0.401 $\pm$ 0.016	1.343
	5 s	-2.577 $\pm$ 0.029	0.168 $\pm$ 0.006	-0.458 $\pm$ 0.015	1.447
P_v	3 s	-3.048 $\pm$ 0.055	0.144 $\pm$ 0.008	-0.205 $\pm$ 0.019	1.446
	5 s	-2.774 $\pm$ 0.048	0.157 $\pm$ 0.011	-0.282 $\pm$ 0.028	1.454
P_3	3 s	-2.803 $\pm$ 0.044	0.138 $\pm$ 0.006	-0.259 $\pm$ 0.014	1.647
	5 s	-2.530 $\pm$ 0.037	0.173 $\pm$ 0.005	-0.399 $\pm$ 0.023	1.290

Figure 3.5 shows the P-wave amplitude vs. hypocentral distance, colored by earthquake magnitude. Visually, there is no discernible difference between maximum amplitude of events of different magnitude measured in either the first 3 or 5 seconds after P-wave arrival for any component (horizontal, vertical, or three-component). The observations of earthquakes of different magnitudes overlap, and no clear pattern is visible. The vertical-only displacement amplitude does not appear to be useful for the purpose of magnitude scaling. In fact, the P_v scaling relation results in a higher residual in comparison to the P_h at both 3 and 5 second search windows. This is surprising, because the P-wave arrival is expected to be most prominent in the vertical component (see Chapter 2). Figure 3.5 makes clear that the difference between the horizontal and three-component displacement amplitude is small when considered on a log scale (Figure 3.5a vs. 3.5e and 3.5b vs. 3.5f).

Over the 100 solution iterations, the standard deviation of the error between known and estimated magnitude is $\geq$1.29 magnitude units for each definition of the P-wave amplitude tested. Our dataset spans a range of 3.4 magnitude units, which corresponds to less than 3 standard deviations, meaning that we cannot reliably use P-wave amplitude to estimate magnitude. This result is in contrast with the previous study by Crowell et al. (2013), which observed separation between a much smaller seismogeodetic dataset of globally distributed earthquakes. In that study, the two largest earthquakes were subduction zone events to the east of Japan, while the three smaller events were strike-slip events in southern California and northern Mexico. It may therefore be

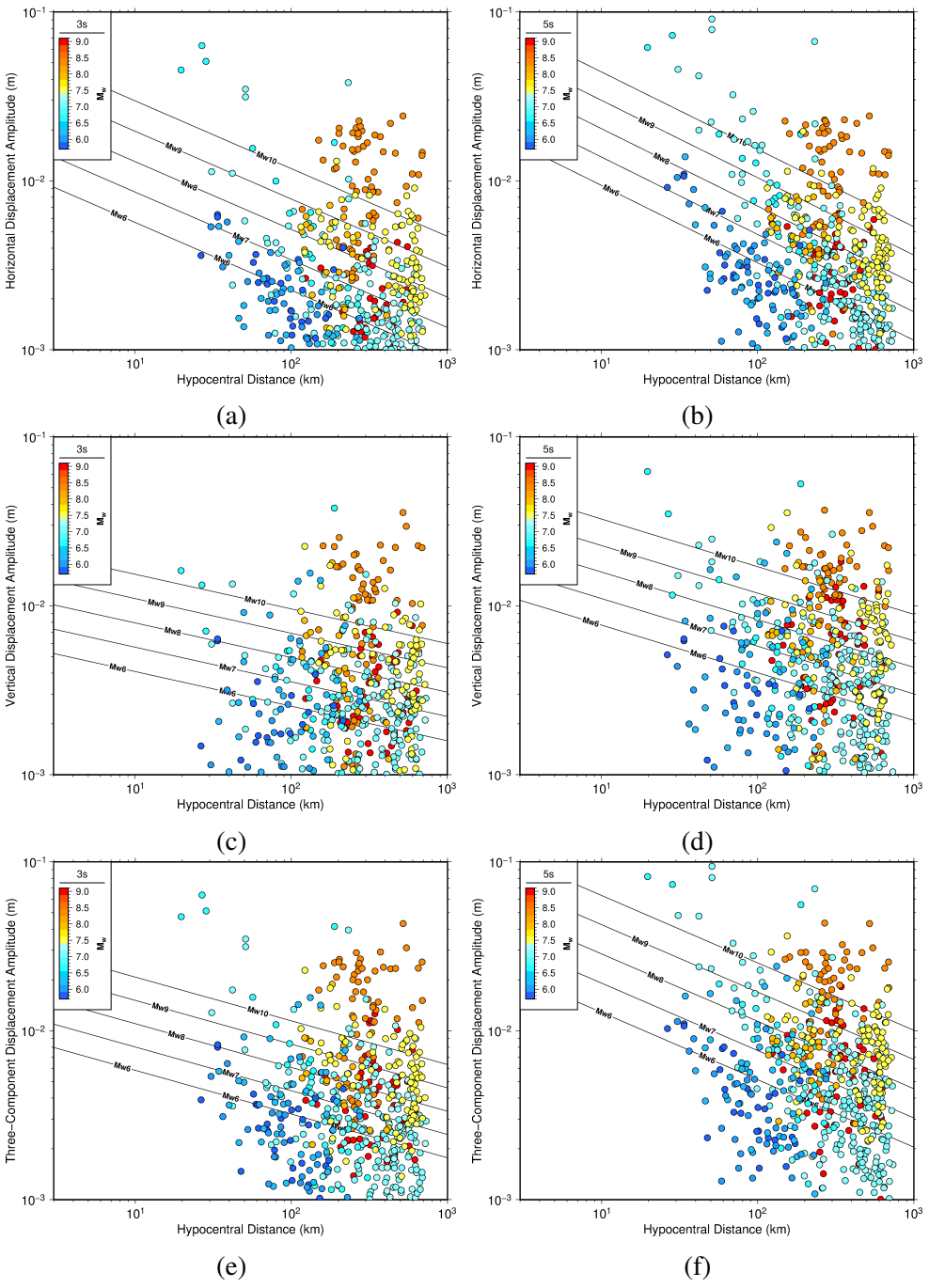


Figure 3.5: Magnitude scaling relation of seismogeodetic P-wave amplitude. Best fitting relation between maximum displacement amplitude and hypocentral distance with Equation 3.4 (black lines). (a) maximum horizontal-only displacement within 3 seconds after P-wave arrival, (b) maximum horizontal-only displacement within 5 seconds after P-wave arrival, (c) maximum vertical-only displacement within 3 seconds after P-wave arrival, (d) maximum vertical-only displacement within 5 seconds after P-wave arrival, (e) maximum three-component sum of squares displacement within 3 seconds after P-wave arrival, and (f) maximum three-component sum of squares displacement within 5 seconds after P-wave arrival.

possible that the results of the Crowell et al. (2013) study were influenced by regional effects or faulting mechanism. The limited geographic scope of the current study reduces any regional effect.

3.5 Peak Amplitude Throughout Time Series

The seismogeodetic P-wave amplitude is shown to be insufficient for identifying the eventual earthquake magnitude. Thus, we explore how the maximum displacement amplitude grows throughout earthquake observation for events of different magnitude, with the goal of identifying magnitude-dependent patterns of maximum displacement with time. To do so, we first modify the magnitude scaling relation (Equation 3.4) to better fit the expected behavior of rupture throughout the full earthquake observation. As rupture expands spatially over the fault, it becomes invalid to approximate the earthquake source as a point source as we had in Equation 3.4, and requires that we consider the magnitude-dependent dimensions of the finite fault. Following Crowell et al. (2013) and Melgar et al. (2015a), we introduce a magnitude dependence in the distance-attenuation term to account for this change in source size:

$$\log_{10}(P) = D + E \cdot M_{\mathrm{w}} + F \cdot M_{\mathrm{w}} \cdot \log_{10}(R) \tag{3.5}$$

We rename the coefficients D, E, and F, to avoid confusion about which scaling relation is being employed (Equation 3.4 or 3.5). We expect that Equation 3.4 will be a better fit to the early observations, when the point source assumption holds, and that at later observations there will be a transition point where Equation 3.5 becomes the better fit to the data. Using these empirically derived relations, and solving for the best fitting parameters A, B, C and D, E, F at 5-second intervals, t, we evaluate the potential of using a series of time-evolving parameters $A(t)$, $B(t)$, $C(t)$ and $D(t)$, $E(t)$, $F(t)$ that describe magnitude-dependent growth patterns, to determine whether the observations at a particular station over time resemble the signal expected from a

M_w6, M_w7, M_w8, etc.

We complete this analysis in 5-second intervals beginning at P-wave arrival $< t \leq 5$ s and ending at P-wave arrival $< t \leq 170$ s, at which point the maximum displacement has been reached for every station considered. For brevity, we focus on the three-component sum of squares definition of displacement amplitude, P_3, in the analysis and discussion and show fits to both the point source and finite fault scaling relations (Equations 3.4 and 3.5 respectively). Comprehensive results which consider the horizontal and vertical components of displacement separately can be found in the Supporting Information (see Section 3.7, Tables S3.3-3.4), however the discussion regarding the three-component sum of squares is broadly applicable to these metrics as well.

Equation 3.4 appears to fit the data better only with $\leq$15 s of data. After 15 s, the errors in the righthand column of Table 3.3 exceed those in Table 3.4. This is in agreement with the previous discussion which suggested that Equation 3.4 is better suited for early observations when the point-source approximation holds, whereas Equation 3.5 is better suited later in rupture. This analysis shows that magnitude estimates come within one magnitude unit only after 15 s of data is available following the P-wave arrival. To come within $\pm$0.3 magnitude units requires at least 55 s of data following the P-wave arrival.

The best fitting scaling relations at are presented at 15-second intervals in Figure 3.6. While initially the events are largely indistinguishable (Figure 3.6a), the smaller magnitude events can be visually distinguished by Figure 3.6b or 3.6c, and the largest events are separated by Figure 3.6e. By the time the final peak ground displacement has been observed, there is clear separation, consistent with previous studies (Crowell et al., 2013; Melgar et al., 2015a).

3.5.1 Kolmogorov-Smirnov Test

To clarify the time at which separation between events of different magnitude occurs, we employ a two-dimensional paired Kolmogorov-Smirnov (K-S) test (Peacock 1983). The null

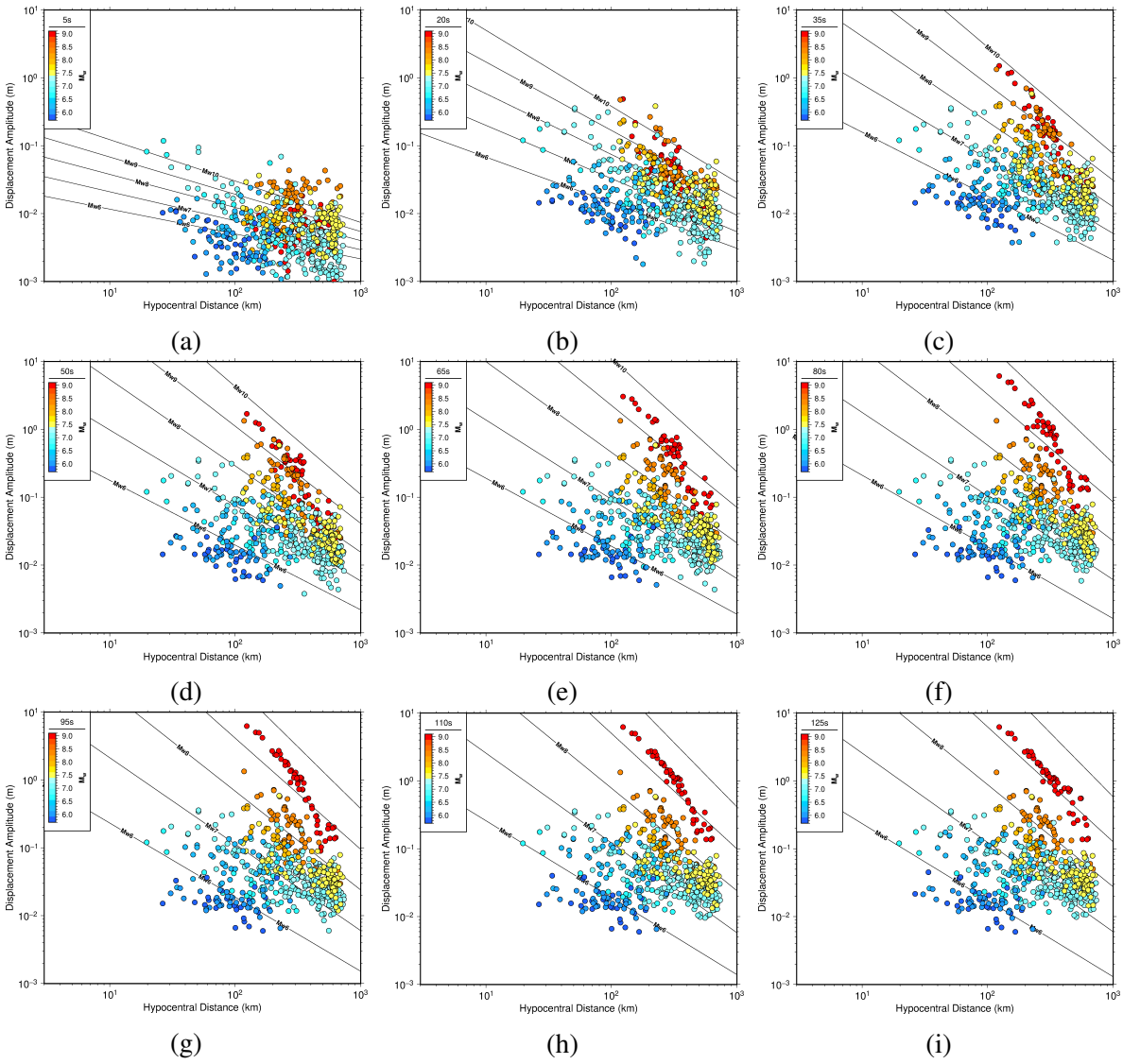


Figure 3.6: Seismogeodetic displacement evolution, using the three-component displacement amplitude. Best fitting scaling relations for the 14 considered events using Equation 3.5 (finite fault source assumption), presented at 15 s intervals. Amplitudes are derived from three-component sum of squares of motion. Each data point represents the observation at a single seismogeodetic station pair, where the color of the data point denotes the magnitude of the observed event. Black lines denote the best fitting magnitude scaling relation, from Equation 3.5.

Table 3.3: Magnitude scaling coefficients, three-component sum-of-squares displacement amplitude and point source assumption (Equation 3.4).

Window (s)	A$\pm\sigma_A$	B$\pm\sigma_B$	C$\pm\sigma_C$	M_w Error
5	-2.530 $\pm$ 0.037	0.173 $\pm$ 0.005	-0.399 $\pm$ 0.023	1.290
10	-2.444 $\pm$ 0.072	0.199 $\pm$ 0.007	-0.414 $\pm$ 0.028	1.087
15	-2.305 $\pm$ 0.058	0.268 $\pm$ 0.007	-0.624 $\pm$ 0.019	0.704
20	-2.234 $\pm$ 0.026	0.312 $\pm$ 0.006	-0.735 $\pm$ 0.018	0.618
25	-2.241 $\pm$ 0.064	0.359 $\pm$ 0.010	-0.844 $\pm$ 0.020	0.551
30	-2.387 $\pm$ 0.029	0.413 $\pm$ 0.009	-0.916 $\pm$ 0.022	0.472
35	-2.459 $\pm$ 0.046	0.477 $\pm$ 0.006	-1.060 $\pm$ 0.019	0.405
40	-2.539 $\pm$ 0.059	0.486 $\pm$ 0.009	-1.037 $\pm$ 0.019	0.390
45	-2.608 $\pm$ 0.056	0.488 $\pm$ 0.007	-1.003 $\pm$ 0.021	0.386
50	-2.758 $\pm$ 0.044	0.522 $\pm$ 0.007	-1.033 $\pm$ 0.014	0.358
55	-3.027 $\pm$ 0.061	0.581 $\pm$ 0.006	-1.087 $\pm$ 0.022	0.311
60	-3.168 $\pm$ 0.062	0.599 $\pm$ 0.010	-1.071 $\pm$ 0.026	0.285
65	-3.243 $\pm$ 0.068	0.619 $\pm$ 0.012	-1.092 $\pm$ 0.026	0.286
70	-3.357 $\pm$ 0.059	0.658 $\pm$ 0.007	-1.155 $\pm$ 0.025	0.272
75	-3.522 $\pm$ 0.060	0.690 $\pm$ 0.010	-1.176 $\pm$ 0.022	0.265
80	-3.495 $\pm$ 0.056	0.684 $\pm$ 0.011	-1.164 $\pm$ 0.016	0.270
85	-3.504 $\pm$ 0.067	0.692 $\pm$ 0.013	-1.182 $\pm$ 0.029	0.266
90	-3.556 $\pm$ 0.078	0.703 $\pm$ 0.013	-1.192 $\pm$ 0.030	0.267
95	-3.580 $\pm$ 0.079	0.708 $\pm$ 0.010	-1.193 $\pm$ 0.029	0.272
100	-3.555 $\pm$ 0.083	0.713 $\pm$ 0.011	-1.215 $\pm$ 0.026	0.275
105	-3.540 $\pm$ 0.078	0.718 $\pm$ 0.009	-1.235 $\pm$ 0.028	0.270
110	-3.511 $\pm$ 0.091	0.723 $\pm$ 0.009	-1.258 $\pm$ 0.025	0.271
115	-3.684 $\pm$ 0.063	0.744 $\pm$ 0.007	-1.249 $\pm$ 0.034	0.271
120	-3.809 $\pm$ 0.042	0.760 $\pm$ 0.011	-1.240 $\pm$ 0.029	0.272
125	-3.825 $\pm$ 0.025	0.771 $\pm$ 0.004	-1.262 $\pm$ 0.017	0.272
130	-3.805 $\pm$ 0.042	0.774 $\pm$ 0.005	-1.282 $\pm$ 0.024	0.268
135	-3.793 $\pm$ 0.049	0.774 $\pm$ 0.004	-1.287 $\pm$ 0.026	0.262
140	-3.826 $\pm$ 0.047	0.775 $\pm$ 0.005	-1.275 $\pm$ 0.025	0.280
145	-3.854 $\pm$ 0.034	0.774 $\pm$ 0.005	-1.261 $\pm$ 0.024	0.273
150	-3.859 $\pm$ 0.036	0.774 $\pm$ 0.005	-1.257 $\pm$ 0.023	0.272
155	-3.862 $\pm$ 0.034	0.773 $\pm$ 0.005	-1.253 $\pm$ 0.021	0.263
160	-3.866 $\pm$ 0.034	0.773 $\pm$ 0.005	-1.250 $\pm$ 0.023	0.270
165	-3.867 $\pm$ 0.037	0.772 $\pm$ 0.005	-1.248 $\pm$ 0.022	0.269
170	-3.870 $\pm$ 0.032	0.770 $\pm$ 0.005	-1.241 $\pm$ 0.022	0.275

Table 3.4: Magnitude scaling coefficients, three-component sum-of-squares displacement amplitude and finite fault source assumption (Equation 3.5).

Window (s)	D$\pm\sigma_D$	E$\pm\sigma_E$	F$\pm\sigma_F$	M_w Error
5	-3.470 $\pm$ 0.059	0.317 $\pm$ 0.017	-0.061 $\pm$ 0.004	1.329
10	-3.397 $\pm$ 0.073	0.348 $\pm$ 0.022	-0.064 $\pm$ 0.006	1.116
15	-3.688 $\pm$ 0.056	0.476 $\pm$ 0.013	-0.092 $\pm$ 0.003	0.712
20	-3.969 $\pm$ 0.044	0.579 $\pm$ 0.010	-0.112 $\pm$ 0.003	0.584
25	-4.290 $\pm$ 0.065	0.675 $\pm$ 0.018	-0.129 $\pm$ 0.004	0.514
30	-4.642 $\pm$ 0.047	0.765 $\pm$ 0.018	-0.142 $\pm$ 0.005	0.422
35	-5.037 $\pm$ 0.026	0.863 $\pm$ 0.007	-0.157 $\pm$ 0.002	0.351
40	-5.053 $\pm$ 0.064	0.867 $\pm$ 0.013	-0.157 $\pm$ 0.003	0.352
45	-5.002 $\pm$ 0.061	0.856 $\pm$ 0.014	-0.153 $\pm$ 0.004	0.364
50	-5.209 $\pm$ 0.054	0.902 $\pm$ 0.013	-0.159 $\pm$ 0.003	0.317
55	-5.563 $\pm$ 0.057	0.973 $\pm$ 0.014	-0.166 $\pm$ 0.004	0.292
60	-5.692 $\pm$ 0.051	0.993 $\pm$ 0.012	-0.166 $\pm$ 0.003	0.276
65	-5.874 $\pm$ 0.054	1.026 $\pm$ 0.014	-0.167 $\pm$ 0.003	0.274
70	-6.060 $\pm$ 0.035	1.066 $\pm$ 0.008	-0.172 $\pm$ 0.002	0.260
75	-6.200 $\pm$ 0.094	1.093 $\pm$ 0.020	-0.175 $\pm$ 0.003	0.265
80	-6.226 $\pm$ 0.086	1.101 $\pm$ 0.020	-0.176 $\pm$ 0.004	0.265
85	-6.308 $\pm$ 0.069	1.119 $\pm$ 0.014	-0.178 $\pm$ 0.003	0.262
90	-6.387 $\pm$ 0.071	1.142 $\pm$ 0.018	-0.183 $\pm$ 0.004	0.274
95	-6.413 $\pm$ 0.055	1.145 $\pm$ 0.014	-0.182 $\pm$ 0.004	0.273
100	-6.449 $\pm$ 0.051	1.152 $\pm$ 0.013	-0.183 $\pm$ 0.003	0.272
105	-6.471 $\pm$ 0.071	1.161 $\pm$ 0.015	-0.185 $\pm$ 0.003	0.277
110	-6.563 $\pm$ 0.082	1.185 $\pm$ 0.024	-0.189 $\pm$ 0.006	0.276
115	-6.774 $\pm$ 0.050	1.221 $\pm$ 0.016	-0.191 $\pm$ 0.004	0.271
120	-6.835 $\pm$ 0.048	1.222 $\pm$ 0.016	-0.188 $\pm$ 0.004	0.276
125	-6.864 $\pm$ 0.041	1.227 $\pm$ 0.014	-0.188 $\pm$ 0.004	0.273
130	-6.885 $\pm$ 0.033	1.224 $\pm$ 0.011	-0.186 $\pm$ 0.003	0.271
135	-6.907 $\pm$ 0.031	1.220 $\pm$ 0.009	-0.183 $\pm$ 0.003	0.284
140	-6.912 $\pm$ 0.032	1.217 $\pm$ 0.011	-0.181 $\pm$ 0.004	0.276
145	-6.906 $\pm$ 0.032	1.217 $\pm$ 0.011	-0.181 $\pm$ 0.004	0.270
150	-6.901 $\pm$ 0.028	1.213 $\pm$ 0.011	-0.180 $\pm$ 0.004	0.274
155	-6.902 $\pm$ 0.033	1.216 $\pm$ 0.010	-0.181 $\pm$ 0.004	0.275
160	-6.884 $\pm$ 0.032	1.205 $\pm$ 0.014	-0.178 $\pm$ 0.005	0.278
165	-6.891 $\pm$ 0.034	1.207 $\pm$ 0.014	-0.178 $\pm$ 0.005	0.271
170	-6.870 $\pm$ 0.031	1.200 $\pm$ 0.016	-0.176 $\pm$ 0.006	0.267

Table 3.5: Two dimensional paired Kolmogorov-Smirnov test.

Earthquake 1		Earthquake 2		Time to Separation (s)
2011 M_w6.7	East Fukushima	2011 M_w7.4	Iwate	20
2011 M_w7.4	Iwate	2011 M_w7.9	Ibaraki	45
2011 M_w7.9	Ibaraki	2003 M_w8.3	Tokachi-oki	55
2003 M_w8.3	Tokachi-oki	2011 M_w9.1	Tohoku-oki	60

hypothesis of the K-S test is that the two datasets are part of the same continuous distribution. We apply the test to a subset of the earthquakes in our study to determine how many seconds of recorded data are necessary to distinguish between earthquakes of large magnitude. For brevity, we show only the three-component sum of squares displacement, P_3. We find that no statistical separation exists between the different earthquakes with only 5 s of data, which is consistent with our previous analysis suggesting that magnitude estimates for large earthquakes is unreliable using first seismic observations. Findings are summarized in Table 3.5. We determine that while the April 2011 M_w6.7 becomes statistically different from the March 2011 M_w7.4 with only 20 s of observations, 45 s of observations is required before separation is established between the March 2011 M_w7.4 and March 2011 M_w7.9. With one minute of data, it is first possible to statistically separate the 2003 M_w8.3 from the 2011 M_w9.1.

3.5.2 Simulated Real-Time Magnitude Estimates

Using the best fitting relations for each search window (5 s to 170 s), we present a simulated real-time magnitude estimation for each of the earthquakes considered with respect to earthquake origin time (OT). We use the P-wave arrival time selected from the seismogeodetic velocity waveforms to determine how much data has been recorded at any time, t_i, after earthquake origin. For each individual station, we measure the maximum amplitude observed between P-wave arrival and OT+t_i, and use the corresponding scaling relation to derive a magnitude estimate for that station. For example, a station with 5-10 s of available data following the P-wave arrival estimates magnitude using the coefficients found in the first line of Table 3.4. Once it has recorded

10 s of data, the selected peak displacement becomes subject to the coefficients in line 2 of Table 3.4. Once 15 s of data has accrued, the coefficients in line 3 are invoked, etc. A weighted average of the individual station estimates is calculated such that the station's weight is defined by the amount of available data (i.e. stations with more recorded data have more weight). In this way, each station's individual magnitude estimate will be allowed to update until it reaches 170 s of available data following the P-wave arrival.

Figure 3.7 shows the estimated magnitude of each event in the simulated real-time environment described above. The grey points show the magnitude estimate updates using only the final coefficients, $D(170\text{ s})$, $E(170\text{ s})$, and $F(170\text{ s})$, which fits the peak ground displacement observed over the entire waveform (similar to the process of Melgar et al. (2015a)). The blue points show the estimated magnitude using the time-evolving coefficients as described in Table 3.4. The pink shapes show the source time function for each event, allowing a clear understanding of when the magnitude estimate becomes reliable relative to rupture duration. For events either farther off-shore or with short rupture durations, it is common that no data are available prior to rupture completion (see Figures 3.7c, e, g, h, j, k, l, n).

The estimates using only final coefficients (grey points) will always initially underestimate the magnitude, and grow as maximum observed ground displacement increases, until reaching final peak ground displacement. Fluctuations in the time-varying estimates (blue points) are expected in the early estimates. Eventually, as enough data are observed subsequent to the P-wave arrival, the two estimates will meet and become identical. In some cases, notably the 2011 M_w9.1 Tohoku-oki event, using the time-evolving parameters appears to speed the process of obtaining an accurate magnitude. However, this may be misleading. We know from Section 3.5.1 that the M_w9.1 Tohoku-oki event does not appear statistically different from the M_w8.3 Tokachi-oki event until 60 seconds of data has been observed. Looking closely at Figures 3.7a and 3.6b, this is confirmed. Although the proper magnitude for the Tohoku-oki event is reached with time-evolving coefficients $\sim$30 s after first P-wave arrival, with the same length of observation

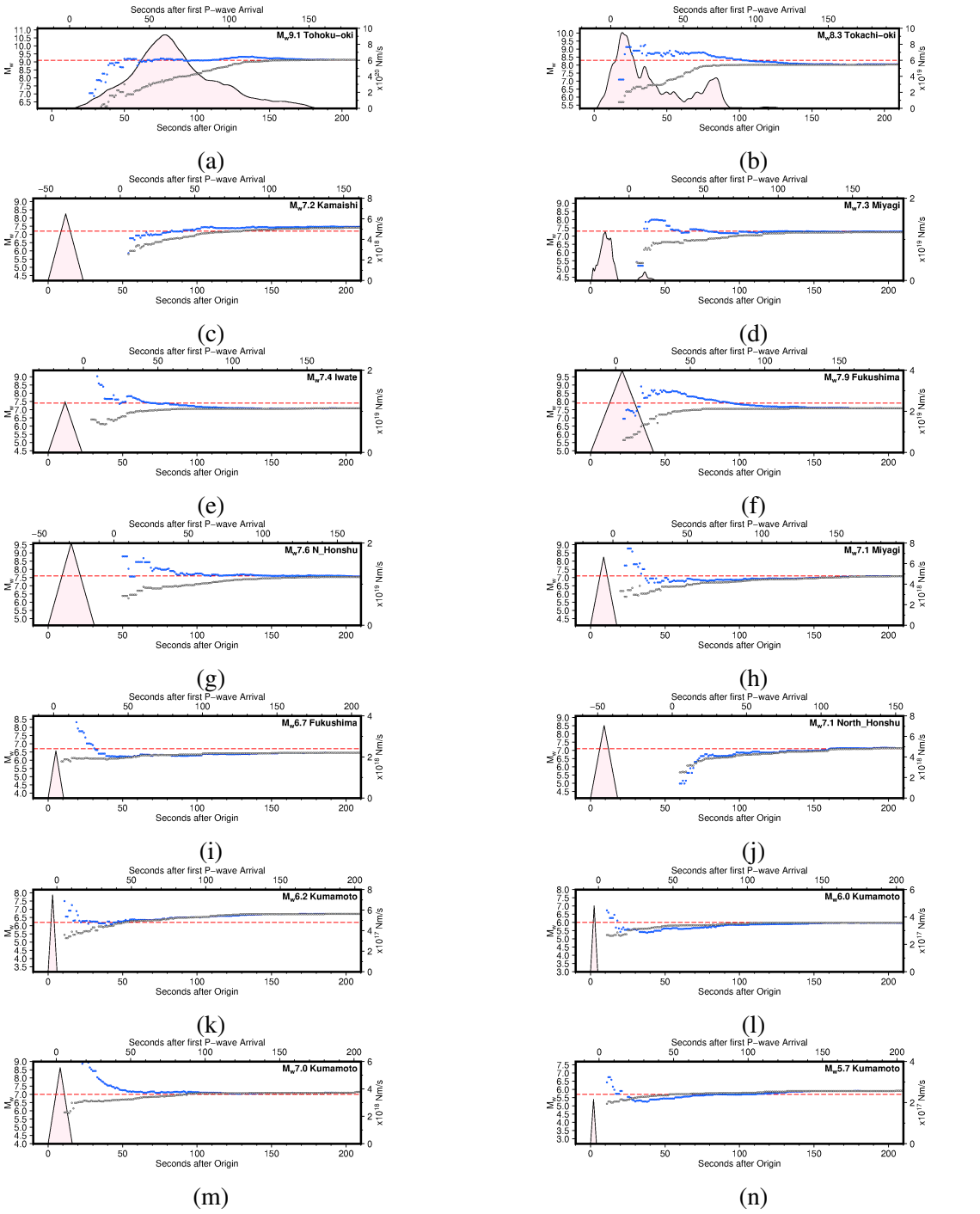


Figure 3.7: Simulated real-time magnitude estimation using final peak ground displacement coefficients, $D(170)$, $E(170)$, and $F(170)$, (grey dots) and time-evolving coefficients $D(t)$, $E(t)$, and $F(t)$, (blue dots). The red dashed line denotes the target, or known magnitude. The light pink shape describes the source time function of the earthquake, as given by the kinematic slip model (Melgar and Bock, 2015 or National Earthquake Information Center, NEIC). If unknown, the source time function is denoted by a triangle function based on the half-duration from United States Geological Survey (USGS) NEIC.

the Tokachi-oki event is also estimated to be a $\sim M_w 9$ event, dropping to its proper magnitude 60+ seconds after the first P-wave arrival. This result confirms that ability to properly discriminate between the M8 and M9 event requires $\sim$60 s of observation.

3.6 Discussion and Conclusions

The results of this study (see Figure 3.5) are in contrast with those of Crowell et al. (2013), which suggested that maximum P-wave amplitude, observed in the first 3-5 seconds after P-wave arrival time, appears to scale without saturation when measured from the seismogeodetic dataset. In this study, the larger events (>M7) appear indistinguishable this early in the rupture process, which is consistent with the magnitude saturation typical of seismic-only instrumentation. The Crowell et al. (2013) study was completed with a significantly smaller, yet globally distributed, dataset due to the availability of historical seismogeodetic datasets at the time, and therefore we consider the results of this larger study to be more robust.

We further investigated the evolution of maximum displacement amplitude across magnitudes to determine if the growth rate of different magnitude events diverges, and whether an evolving set of parameters over the course of observation allows more timely magnitude estimation. The smaller magnitude events that have very rapid convergence to the correct magnitude (Figure 3.7) also have very short rupture durations such that rupture is complete prior to any observation, making it difficult to investigate any magnitude-dependent signature of rupture nucleation. This approach does not appear to hasten the proper evaluation of magnitude for the largest events, requiring about a minute of data observation before both the $M_w 8.3$ and $M_w 9.1$ can be properly distinguished from one another.

We acknowledge that availability of observational data is a major limitation for this kind of study. Despite the large and overlapping seismic and GNSS datasets in Japan, the number of collocated stations that recorded the full earthquake from P-wave arrival through cessation of

shaking was limited. Furthermore, there are very few large (>M8) earthquakes which have been recorded with collocated seismic and GNSS instrumentation. Presently, the M_w9.1 Tohoku-oki earthquake is the only well-instrumented M9 event, and therefore studies such as this one are forced to assume that these large megathrust events share similar properties. However, the 2004 M_w9.1 Sumatra earthquake (observed with mainly teleseismic data) famously began with slow rupture and little slip for nearly a minute prior to ramping up to become a great, long duration earthquake (Ammon et al., 2005). It is unlikely that this topic will be fully resolved without additional observational data from future large earthquakes.

Still, there are lessons learned toward improving our capability of delivering a rapid and accurate warning. In an ideal scenario, the two horizontal axes on Figure 3.7 would be nearly aligned, indicating that observation at seismogeodetic instrumentation occurs almost immediately after earthquake origin time. Instead, we have three events in which first observation occurred $\sim$50 s behind the earthquake origin time, limiting the timeliness of solution. This observation implies that it may become crucial to incorporate seafloor geodetic techniques into the real-time system in order to accommodate rapid observation of large earthquakes offshore to facilitate early warning and rapid response. However, the notion that closer instrumentation will lead to faster earthquake magnitude calculation is not unequivocal. Our seismogeodetic observations are not indicative of strong earthquake determinism. If earthquakes are not deterministic, we must wait until rupture is complete prior to observing useful quantities for identifying magnitude. In the case of large, long duration earthquakes, rapid identification of earthquake signature would improve timeliness of earthquake detection, but not necessarily estimation of earthquake size. We explore this topic of earthquake determinism and timeliness of earthquake magnitude estimation in greater detail in the following chapter.

3.7 Supporting Information

Table S3.1: Magnitude scaling coefficients, horizontal displacement amplitude and point source assumption (Equation 3.4).

Window (s)	A$\pm\sigma_A$	B$\pm\sigma_B$	C$\pm\sigma_C$	M_w Error
5	-2.577 $\pm$ 0.029	0.168 $\pm$ 0.006	-0.458 $\pm$ 0.015	1.447
10	-2.397 $\pm$ 0.083	0.228 $\pm$ 0.006	-0.627 $\pm$ 0.024	0.976
15	-2.429 $\pm$ 0.042	0.338 $\pm$ 0.008	-0.870 $\pm$ 0.018	0.575
20	-2.310 $\pm$ 0.030	0.379 $\pm$ 0.007	-0.998 $\pm$ 0.017	0.562
25	-2.457 $\pm$ 0.047	0.444 $\pm$ 0.007	-1.094 $\pm$ 0.021	0.512
30	-2.610 $\pm$ 0.061	0.522 $\pm$ 0.008	-1.234 $\pm$ 0.028	0.402
35	-2.734 $\pm$ 0.059	0.591 $\pm$ 0.005	-1.368 $\pm$ 0.027	0.317
40	-2.810 $\pm$ 0.050	0.603 $\pm$ 0.009	-1.358 $\pm$ 0.019	0.312
45	-2.788 $\pm$ 0.036	0.597 $\pm$ 0.004	-1.338 $\pm$ 0.013	0.319
50	-2.922 $\pm$ 0.063	0.624 $\pm$ 0.012	-1.357 $\pm$ 0.024	0.290
55	-3.168 $\pm$ 0.047	0.692 $\pm$ 0.007	-1.444 $\pm$ 0.016	0.258
60	-3.317 $\pm$ 0.044	0.713 $\pm$ 0.010	-1.432 $\pm$ 0.021	0.239
65	-3.459 $\pm$ 0.037	0.743 $\pm$ 0.005	-1.453 $\pm$ 0.011	0.220
70	-3.639 $\pm$ 0.062	0.761 $\pm$ 0.008	-1.425 $\pm$ 0.020	0.211
75	-3.911 $\pm$ 0.057	0.796 $\pm$ 0.006	-1.407 $\pm$ 0.021	0.201
80	-3.925 $\pm$ 0.090	0.798 $\pm$ 0.014	-1.400 $\pm$ 0.018	0.201
85	-4.014 $\pm$ 0.077	0.803 $\pm$ 0.010	-1.374 $\pm$ 0.016	0.202
90	-4.110 $\pm$ 0.057	0.811 $\pm$ 0.007	-1.353 $\pm$ 0.023	0.196
95	-4.147 $\pm$ 0.069	0.811 $\pm$ 0.009	-1.336 $\pm$ 0.022	0.200
100	-4.232 $\pm$ 0.073	0.822 $\pm$ 0.010	-1.333 $\pm$ 0.019	0.197
105	-4.297 $\pm$ 0.094	0.833 $\pm$ 0.016	-1.336 $\pm$ 0.021	0.199
110	-4.319 $\pm$ 0.079	0.845 $\pm$ 0.011	-1.363 $\pm$ 0.025	0.195
115	-4.404 $\pm$ 0.041	0.858 $\pm$ 0.008	-1.368 $\pm$ 0.024	0.193
120	-4.530 $\pm$ 0.023	0.880 $\pm$ 0.005	-1.381 $\pm$ 0.017	0.193
125	-4.564 $\pm$ 0.027	0.883 $\pm$ 0.003	-1.376 $\pm$ 0.016	0.195
130	-4.557 $\pm$ 0.023	0.885 $\pm$ 0.003	-1.384 $\pm$ 0.013	0.190
135	-4.562 $\pm$ 0.018	0.885 $\pm$ 0.004	-1.382 $\pm$ 0.014	0.196
140	-4.568 $\pm$ 0.026	0.885 $\pm$ 0.003	-1.380 $\pm$ 0.014	0.193
145	-4.611 $\pm$ 0.046	0.885 $\pm$ 0.003	-1.359 $\pm$ 0.020	0.189
150	-4.618 $\pm$ 0.049	0.884 $\pm$ 0.003	-1.355 $\pm$ 0.020	0.192
155	-4.604 $\pm$ 0.045	0.885 $\pm$ 0.004	-1.362 $\pm$ 0.021	0.188
160	-4.633 $\pm$ 0.050	0.883 $\pm$ 0.004	-1.346 $\pm$ 0.023	0.188
165	-4.626 $\pm$ 0.045	0.884 $\pm$ 0.003	-1.350 $\pm$ 0.021	0.185
170	-4.624 $\pm$ 0.045	0.883 $\pm$ 0.003	-1.349 $\pm$ 0.020	0.186

Table S3.2: Magnitude scaling coefficients, horizontal displacement amplitude and finite fault source assumption (Equation 3.5).

Window (s)	A$\pm\sigma_A$	B$\pm\sigma_B$	C$\pm\sigma_C$	M_w Error
5	-3.594 $\pm$ 0.073	0.317 $\pm$ 0.018	-0.066 $\pm$ 0.004	1.599
10	-3.788 $\pm$ 0.048	0.433 $\pm$ 0.009	-0.091 $\pm$ 0.004	1.020
15	-4.452 $\pm$ 0.074	0.646 $\pm$ 0.019	-0.131 $\pm$ 0.004	0.585
20	-4.712 $\pm$ 0.054	0.745 $\pm$ 0.017	-0.150 $\pm$ 0.004	0.498
25	-5.107 $\pm$ 0.050	0.857 $\pm$ 0.013	-0.169 $\pm$ 0.003	0.440
30	-5.529 $\pm$ 0.066	0.969 $\pm$ 0.019	-0.187 $\pm$ 0.004	0.358
35	-5.985 $\pm$ 0.050	1.085 $\pm$ 0.013	-0.206 $\pm$ 0.003	0.282
40	-5.964 $\pm$ 0.039	1.076 $\pm$ 0.009	-0.201 $\pm$ 0.003	0.276
45	-5.985 $\pm$ 0.037	1.081 $\pm$ 0.010	-0.201 $\pm$ 0.002	0.268
50	-6.130 $\pm$ 0.060	1.105 $\pm$ 0.014	-0.201 $\pm$ 0.003	0.251
55	-6.526 $\pm$ 0.023	1.189 $\pm$ 0.005	-0.211 $\pm$ 0.001	0.229
60	-6.729 $\pm$ 0.030	1.224 $\pm$ 0.006	-0.213 $\pm$ 0.002	0.220
65	-6.902 $\pm$ 0.051	1.252 $\pm$ 0.010	-0.213 $\pm$ 0.002	0.209
70	-7.061 $\pm$ 0.064	1.278 $\pm$ 0.014	-0.214 $\pm$ 0.004	0.209
75	-7.218 $\pm$ 0.062	1.291 $\pm$ 0.015	-0.209 $\pm$ 0.003	0.194
80	-7.241 $\pm$ 0.053	1.288 $\pm$ 0.013	-0.205 $\pm$ 0.003	0.200
85	-7.244 $\pm$ 0.059	1.281 $\pm$ 0.016	-0.201 $\pm$ 0.004	0.202
90	-7.288 $\pm$ 0.094	1.286 $\pm$ 0.021	-0.200 $\pm$ 0.004	0.194
95	-7.317 $\pm$ 0.083	1.287 $\pm$ 0.019	-0.199 $\pm$ 0.003	0.195
100	-7.411 $\pm$ 0.066	1.303 $\pm$ 0.014	-0.200 $\pm$ 0.003	0.205
105	-7.448 $\pm$ 0.075	1.310 $\pm$ 0.016	-0.200 $\pm$ 0.003	0.200
110	-7.587 $\pm$ 0.078	1.338 $\pm$ 0.018	-0.203 $\pm$ 0.004	0.203
115	-7.772 $\pm$ 0.045	1.361 $\pm$ 0.011	-0.203 $\pm$ 0.004	0.197
120	-7.847 $\pm$ 0.039	1.367 $\pm$ 0.010	-0.201 $\pm$ 0.003	0.195
125	-7.864 $\pm$ 0.034	1.369 $\pm$ 0.008	-0.200 $\pm$ 0.002	0.199
130	-7.883 $\pm$ 0.030	1.367 $\pm$ 0.009	-0.198 $\pm$ 0.003	0.192
135	-7.893 $\pm$ 0.027	1.368 $\pm$ 0.007	-0.198 $\pm$ 0.002	0.200
140	-7.892 $\pm$ 0.027	1.363 $\pm$ 0.008	-0.196 $\pm$ 0.003	0.195
145	-7.879 $\pm$ 0.024	1.361 $\pm$ 0.009	-0.196 $\pm$ 0.003	0.197
150	-7.887 $\pm$ 0.027	1.362 $\pm$ 0.008	-0.196 $\pm$ 0.003	0.195
155	-7.884 $\pm$ 0.026	1.361 $\pm$ 0.009	-0.196 $\pm$ 0.003	0.197
160	-7.875 $\pm$ 0.030	1.357 $\pm$ 0.010	-0.194 $\pm$ 0.003	0.190
165	-7.875 $\pm$ 0.031	1.357 $\pm$ 0.009	-0.194 $\pm$ 0.003	0.194
170	-7.872 $\pm$ 0.029	1.357 $\pm$ 0.009	-0.195 $\pm$ 0.003	0.195

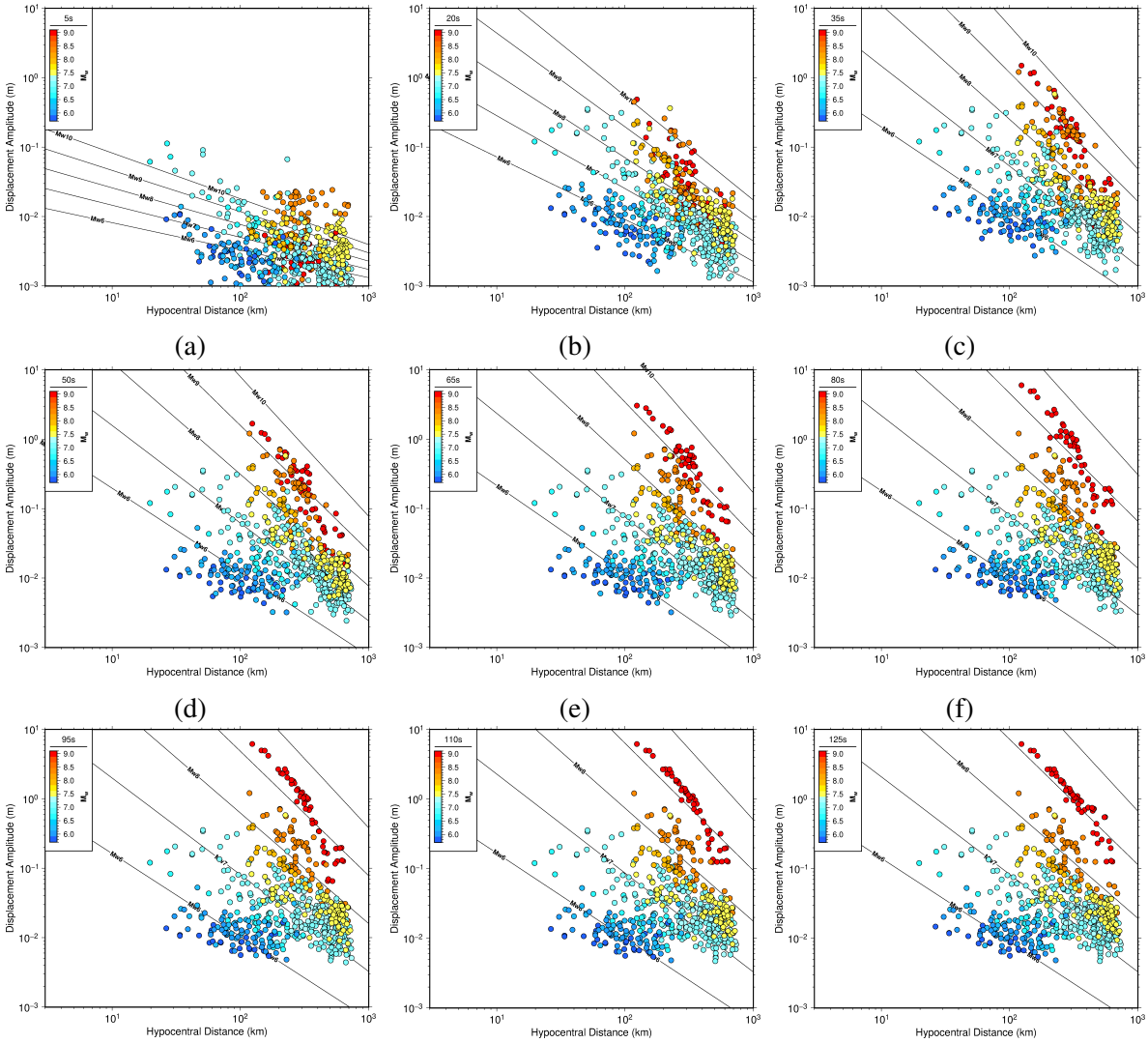


Figure S3.1: Seismogeodetic displacement evolution from horizontal displacement amplitude. Best fitting scaling relations for the 14 considered events using Equation 3.5 (finite fault source assumption), presented at 15 s intervals. Amplitudes are derived from horizontal-only component of motion. Each data point represents the observation at a single seismogeodetic station pair, where the color of the data point denotes the magnitude of the observed event. Black lines denote the best fitting magnitude scaling relation, from Equation 3.5.

Table S3.3: Magnitude scaling coefficients, vertical displacement amplitude and point source assumption (Equation 3.4).

Window (s)	A$\pm\sigma_A$	B$\pm\sigma_B$	C$\pm\sigma_C$	M_w Error
5	-2.774 $\pm$ 0.048	0.157 $\pm$ 0.011	-0.282 $\pm$ 0.028	1.454
10	-2.919 $\pm$ 0.069	0.207 $\pm$ 0.014	-0.291 $\pm$ 0.026	1.069
15	-2.552 $\pm$ 0.049	0.190 $\pm$ 0.006	-0.337 $\pm$ 0.018	1.098
20	-2.356 $\pm$ 0.043	0.178 $\pm$ 0.006	-0.344 $\pm$ 0.014	1.206
25	-2.188 $\pm$ 0.040	0.200 $\pm$ 0.007	-0.450 $\pm$ 0.015	1.073
30	-2.401 $\pm$ 0.036	0.266 $\pm$ 0.008	-0.521 $\pm$ 0.021	0.804
35	-2.336 $\pm$ 0.033	0.265 $\pm$ 0.006	-0.535 $\pm$ 0.014	0.781
40	-2.438 $\pm$ 0.048	0.317 $\pm$ 0.006	-0.628 $\pm$ 0.020	0.658
45	-2.459 $\pm$ 0.063	0.332 $\pm$ 0.013	-0.650 $\pm$ 0.027	0.589
50	-2.466 $\pm$ 0.059	0.349 $\pm$ 0.006	-0.692 $\pm$ 0.022	0.582
55	-2.556 $\pm$ 0.082	0.356 $\pm$ 0.007	-0.668 $\pm$ 0.020	0.563
60	-2.593 $\pm$ 0.041	0.359 $\pm$ 0.005	-0.650 $\pm$ 0.014	0.542
65	-2.659 $\pm$ 0.059	0.375 $\pm$ 0.007	-0.662 $\pm$ 0.025	0.546
70	-2.713 $\pm$ 0.076	0.385 $\pm$ 0.003	-0.661 $\pm$ 0.025	0.504
75	-2.924 $\pm$ 0.064	0.425 $\pm$ 0.011	-0.683 $\pm$ 0.026	0.439
80	-2.998 $\pm$ 0.074	0.455 $\pm$ 0.015	-0.736 $\pm$ 0.030	0.425
85	-3.074 $\pm$ 0.075	0.462 $\pm$ 0.011	-0.720 $\pm$ 0.027	0.407
90	-3.057 $\pm$ 0.058	0.473 $\pm$ 0.004	-0.755 $\pm$ 0.021	0.406
95	-3.163 $\pm$ 0.054	0.493 $\pm$ 0.008	-0.770 $\pm$ 0.018	0.395
100	-3.194 $\pm$ 0.044	0.499 $\pm$ 0.005	-0.771 $\pm$ 0.013	0.385
105	-3.201 $\pm$ 0.056	0.503 $\pm$ 0.007	-0.778 $\pm$ 0.016	0.383
110	-3.168 $\pm$ 0.064	0.505 $\pm$ 0.006	-0.794 $\pm$ 0.021	0.388
115	-3.174 $\pm$ 0.058	0.508 $\pm$ 0.004	-0.802 $\pm$ 0.020	0.392
120	-3.248 $\pm$ 0.047	0.511 $\pm$ 0.003	-0.778 $\pm$ 0.018	0.390
125	-3.346 $\pm$ 0.048	0.516 $\pm$ 0.005	-0.752 $\pm$ 0.018	0.390
130	-3.383 $\pm$ 0.046	0.522 $\pm$ 0.006	-0.753 $\pm$ 0.015	0.392
135	-3.377 $\pm$ 0.048	0.522 $\pm$ 0.006	-0.755 $\pm$ 0.017	0.384
140	-3.368 $\pm$ 0.059	0.520 $\pm$ 0.006	-0.751 $\pm$ 0.017	0.400
145	-3.393 $\pm$ 0.040	0.520 $\pm$ 0.006	-0.740 $\pm$ 0.018	0.393
150	-3.400 $\pm$ 0.036	0.519 $\pm$ 0.005	-0.735 $\pm$ 0.017	0.391
155	-3.398 $\pm$ 0.034	0.520 $\pm$ 0.004	-0.736 $\pm$ 0.015	0.391
160	-3.392 $\pm$ 0.045	0.517 $\pm$ 0.005	-0.731 $\pm$ 0.019	0.392
165	-3.397 $\pm$ 0.039	0.515 $\pm$ 0.007	-0.722 $\pm$ 0.025	0.386
170	-3.401 $\pm$ 0.041	0.513 $\pm$ 0.007	-0.712 $\pm$ 0.027	0.381

Table S3.4: Magnitude scaling coefficients, vertical displacement amplitude and finite fault source assumption (Equation 3.5).

Window (s)	A$\pm\sigma_A$	B$\pm\sigma_B$	C$\pm\sigma_C$	M_w Error
5	-3.450 $\pm$ 0.063	0.260 $\pm$ 0.020	-0.042 $\pm$ 0.005	1.521
10	-3.592 $\pm$ 0.083	0.312 $\pm$ 0.020	-0.045 $\pm$ 0.004	1.157
15	-3.333 $\pm$ 0.043	0.306 $\pm$ 0.010	-0.050 $\pm$ 0.003	1.153
20	-3.167 $\pm$ 0.057	0.305 $\pm$ 0.016	-0.053 $\pm$ 0.004	1.246
25	-3.256 $\pm$ 0.052	0.364 $\pm$ 0.014	-0.068 $\pm$ 0.003	1.073
30	-3.680 $\pm$ 0.057	0.466 $\pm$ 0.014	-0.081 $\pm$ 0.003	0.788
35	-3.662 $\pm$ 0.074	0.473 $\pm$ 0.017	-0.084 $\pm$ 0.003	0.740
40	-3.931 $\pm$ 0.066	0.557 $\pm$ 0.016	-0.100 $\pm$ 0.003	0.593
45	-4.005 $\pm$ 0.059	0.573 $\pm$ 0.013	-0.100 $\pm$ 0.002	0.592
50	-4.104 $\pm$ 0.057	0.602 $\pm$ 0.011	-0.106 $\pm$ 0.002	0.563
55	-4.138 $\pm$ 0.058	0.603 $\pm$ 0.015	-0.103 $\pm$ 0.004	0.540
60	-4.083 $\pm$ 0.055	0.601 $\pm$ 0.009	-0.104 $\pm$ 0.002	0.537
65	-4.182 $\pm$ 0.059	0.616 $\pm$ 0.009	-0.103 $\pm$ 0.003	0.517
70	-4.266 $\pm$ 0.030	0.617 $\pm$ 0.006	-0.098 $\pm$ 0.003	0.469
75	-4.565 $\pm$ 0.073	0.685 $\pm$ 0.020	-0.107 $\pm$ 0.005	0.422
80	-4.791 $\pm$ 0.071	0.728 $\pm$ 0.015	-0.111 $\pm$ 0.003	0.408
85	-4.777 $\pm$ 0.098	0.721 $\pm$ 0.020	-0.108 $\pm$ 0.003	0.406
90	-4.844 $\pm$ 0.031	0.733 $\pm$ 0.009	-0.109 $\pm$ 0.002	0.384
95	-4.943 $\pm$ 0.048	0.750 $\pm$ 0.010	-0.110 $\pm$ 0.002	0.383
100	-5.024 $\pm$ 0.048	0.764 $\pm$ 0.011	-0.111 $\pm$ 0.003	0.376
105	-5.061 $\pm$ 0.029	0.770 $\pm$ 0.006	-0.111 $\pm$ 0.002	0.367
110	-5.053 $\pm$ 0.022	0.768 $\pm$ 0.004	-0.110 $\pm$ 0.002	0.381
115	-5.051 $\pm$ 0.026	0.766 $\pm$ 0.006	-0.109 $\pm$ 0.002	0.370
120	-5.121 $\pm$ 0.060	0.763 $\pm$ 0.012	-0.104 $\pm$ 0.004	0.375
125	-5.158 $\pm$ 0.060	0.763 $\pm$ 0.013	-0.102 $\pm$ 0.003	0.387
130	-5.178 $\pm$ 0.057	0.770 $\pm$ 0.013	-0.103 $\pm$ 0.004	0.377
135	-5.175 $\pm$ 0.054	0.770 $\pm$ 0.013	-0.103 $\pm$ 0.004	0.379
140	-5.150 $\pm$ 0.059	0.763 $\pm$ 0.016	-0.102 $\pm$ 0.004	0.385
145	-5.129 $\pm$ 0.053	0.755 $\pm$ 0.015	-0.099 $\pm$ 0.004	0.395
150	-5.115 $\pm$ 0.049	0.751 $\pm$ 0.014	-0.099 $\pm$ 0.004	0.396
155	-5.102 $\pm$ 0.050	0.748 $\pm$ 0.015	-0.098 $\pm$ 0.004	0.397
160	-5.095 $\pm$ 0.056	0.747 $\pm$ 0.015	-0.098 $\pm$ 0.004	0.383
165	-5.069 $\pm$ 0.044	0.741 $\pm$ 0.012	-0.097 $\pm$ 0.003	0.389
170	-5.049 $\pm$ 0.047	0.733 $\pm$ 0.010	-0.094 $\pm$ 0.002	0.381

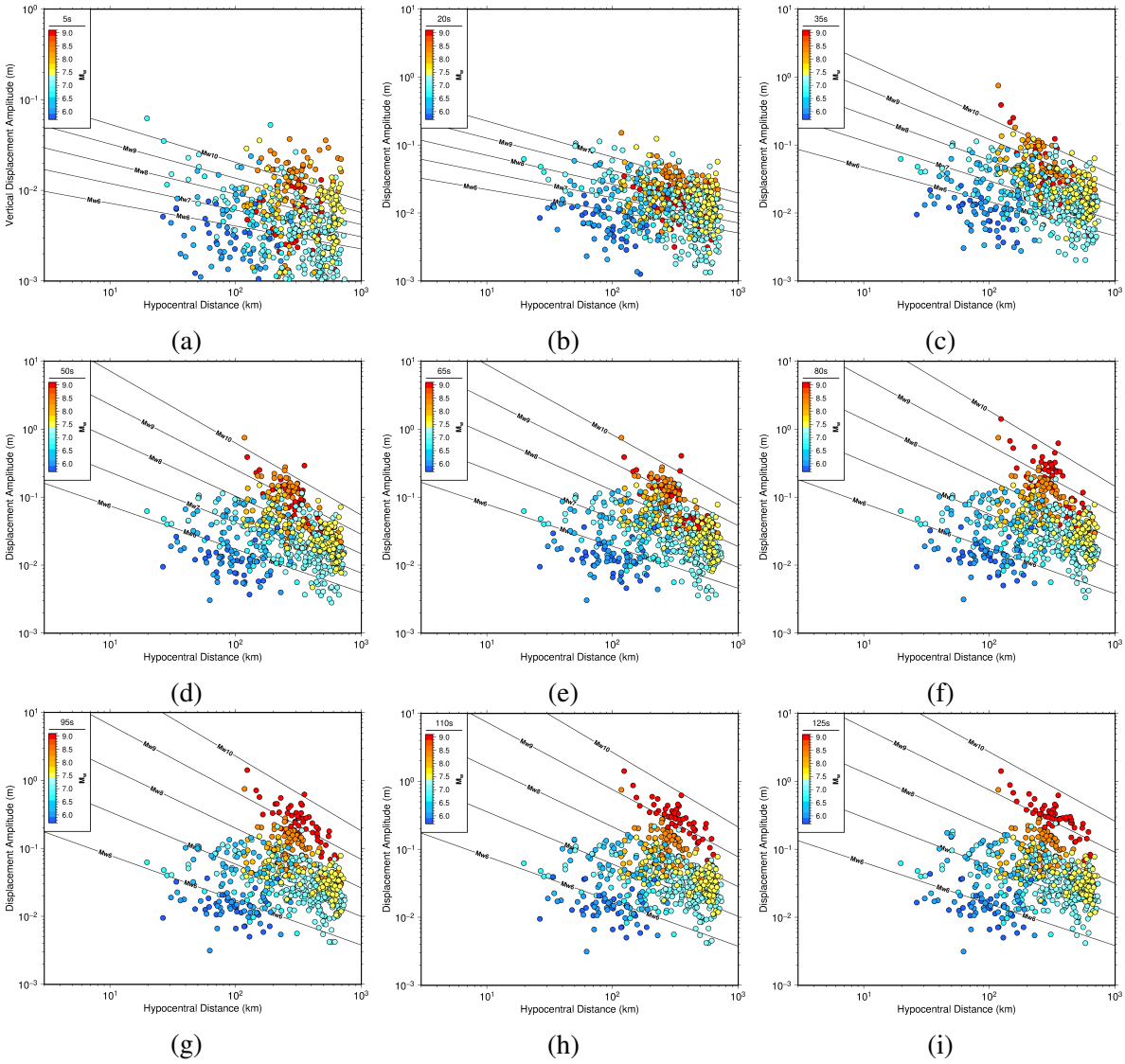


Figure S3.2: Seismogeodetic displacement evolution from vertical displacement amplitude. Best fitting scaling relations for the 14 considered events using Equation 3.5 (finite fault source assumption), presented at 15 s intervals. Amplitudes are derived from vertical-only component of motion. Each data point represents the observation at a single seismogeodetic station pair, where the color of the data point denotes the magnitude of the observed event. Black lines denote the best fitting magnitude scaling relation, from Equation 3.5.

3.8 Acknowledgements

This work is in preparation for publication as **Goldberg, D.E., Melgar, D. and Bock, Y. Evaluation of earthquake determinism with seismogeodetic displacements, *in prep*.** The dissertation author was the primary investigator and author of this material. We gratefully acknowledge funding support from the National Aeronautics and Space Administration grants NNX14AQ53G, NNX16AM04A, and NNX13AT33G. We thank the Geospatial Information Authority of Japan for the GNSS data and maintenance of the GNSS Earth Observation Network as well as the National Research Institute for Earth Science and Disaster Resilience for maintenance of the strong-motion networks, KiK-net and K-net. Figures were produced using Generic Mapping Tools (Wessel et al., 2013).

Chapter 4

Geodetic Observations of Rupture Evolution

4.1 Background

The temporal evolution of seismic moment release is a subject of fundamental interest in earthquake source physics and applied seismology, particularly for large and damaging events. Specifically, whether a large earthquake presents characteristics different from an earthquake of lesser magnitude at some point in the seismic rupture has been widely debated. Observations have suggested there is strong determinism, that is, it should be possible to estimate the final magnitude of an event in the first few seconds (the nucleation phase) of large rupture (e.g. Colombelli et al., 2014; Olson and Allen, 2005; Zollo et al., 2006). However, this hypothesis has been disputed. An observational basis has also been found for the contrasting view, that there is no determinism whatsoever and that nucleation is a magnitude-independent process (e.g. Rydelek and Horiuchi, 2006) such that the final magnitude cannot be determined from observations of only the first few seconds of rupture. In particular, Meier et al. (2016) make a strong case for this universal magnitude-independent nucleation by analyzing near-field strong motion records for moderate

events ($4.0 < M_w < 8.5$), finding no evidence that rupture onsets can be used to predict final magnitude. Recently, analyses of large databases of finite fault models obtained from teleseismic inversions have also been studied to shed light on these issues and have led to a more nuanced perspective. Melgar and Hayes (2017) find evidence that large events behave as self-similar slip pulses (Heaton, 1990) and propose a model of weak determinism for rupture evolution. In this view, sometime following nucleation (tens of seconds) rupture organizes into a self-similar pulse whose properties are diagnostic of final magnitude. Meier et al. (2017), analyzing the same dataset, suggest a similar model but argue that rupture can only be distinguished once peak moment rate occurs, roughly a third of the way into a large rupture. Underlying these recent studies is a shift in the hypothesis being tested away from whether rupture nucleation is deterministic and towards when, following nucleation, information of the rupture evolution can be used to infer final magnitude.

These theoretical and observational considerations, which have made the problem difficult to solve, are further compounded by measurement challenges. Observing large earthquakes at close distances is not without difficulties. Inertial seismometers with both low- and high-gain are the most commonly used tool for regional earthquake observations. High-gain, broadband instruments are more sensitive to small ground motions, but their dynamic range is exceeded during heavy shaking, rendering their recordings unusable in the near-field of large magnitude events. Low-gain accelerometers capture the strongest shaking, but are unable to distinguish between rotational and translational motions. As earthquakes get larger in magnitude, rotational motions become more important (Trifunac and Todorovska, 2001), leading to an inaccurate representation of ground motion, particularly at long periods. These errors in measurement, referred to as baseline offsets, are corrected in real-time with high-pass filtering (e.g. Boore et al., 2002). This correction dampens the influence of the baseline offsets, but leads to a band-limited signal that excludes the permanent displacement (the 0 Hz static offset), as well as the long-period band of the record. Furthermore, it reduces the observed amplitude of maximum displacement (e.g.

Melgar et al., 2015a). The consequence is that with traditional seismic observations, the magnitude of large, destructive earthquakes is often underestimated– even after rupture is complete– in a well-documented condition known as magnitude saturation (e.g. Colombelli et al., 2012; Hoshiba and Ozaki, 2014).

Saturation is of practical importance because rapid magnitude calculation has implications for seismic risk mitigation. Once evaluated, magnitude, coupled with a suitable ground motion prediction equation, is the main measure used by early warning systems to provide an estimate of expected shaking at a given location before it occurs (e.g. Kohler et al., 2017). Prompt and accurate assessments of an earthquake's size and expected ground motion are also useful for first responders as an initial estimate of the extent of damage. In subduction zone regions, current local tsunami warning systems are driven simply by location and magnitude (Hoshiba and Ozaki, 2014). Magnitude uncertainties can inhibit the effectiveness of an early warning, and particularly a tsunami evacuation order, because the predicted damage and affected areas will be underestimated. A poignant example was observed during the 2011 Mw9.1 Tohoku-oki earthquake offshore Japan (Yun and Hamada, 2014). The earthquake was estimated as M_w8.1, 122 seconds after origin and upgraded to M_w8.4 only after 74 minutes, at which point the tsunami had already inundated parts of the coast nearest to the source. When the earthquake was finally observed at teleseismic distances, the magnitude was upgraded to M_w9.1 (Hoshiba and Ozaki, 2014).

In an effort to overcome magnitude saturation and improve warnings, a considerable amount of attention has been devoted to other ground motion measurements. In particular, Global Navigation Satellite System (GNSS) observations can provide high-rate broadband displacements (Bock et al., 2011) devoid of baseline offsets and reliable down to the longest periods (Melgar et al., 2012). However, high-rate GNSS data have a lower sensitivity than seismic instrumentation: 1-2cm in the horizontal and $\sim$5-10cm in the vertical (Genrich and Bock, 2006). Therefore, GNSS data are not suitable for detection of the small amplitude P-wave arrivals, inhibiting the ability of utilizing such data to effectively pinpoint the timing of early magnitude evolution. Furthermore,

GNSS data are typically sampled at much lower sampling rates (1-10Hz), primarily due to the verbosity of phase and pseudorange observations to multiple satellites at multiple wavelengths for each epoch. Seismogeodesy, the optimal combination of collocated seismic and geodetic instrumentation, provides a favorable dataset for exploring the subtleties of early observations (Bock et al., 2011; Melgar et al., 2012; Crowell et al., 2013; Saunders et al., 2016; Goldberg and Bock, 2017). The combination dataset has the temporal resolution of the seismic instrumentation, and results in a displacement time series that is more accurate than that using integrated and filtered seismic instrumentation, and with reduced noise compared to GNSS-only, improving the sensitivity. Importantly, this approach enables the detection of P-wave arrivals in seismogeodetic velocities, improving the timeliness of a warning (Goldberg and Bock, 2017). Algorithms that leverage these broadband displacement data and use them for magnitude calculation have been developed as well. Crowell et al. (2013) first noted that peak ground displacement (PGD) measured with seismogeodetic instrumentation is reliable for a simple point-source magnitude scaling law and Melgar et al. (2015a) confirmed for a global suite of large events measured with geodetic instrumentation, that no saturation is observed in PGD with respect to either magnitude or source to station distance. Based on these observations, Crowell et al. (2016) designed and implemented a real-time PGD magnitude algorithm for the west coast of the United States. By their very nature, magnitude algorithms that rely on PGD scaling laws are limited in solution speed, because the peak displacement occurs sometime after S-waves arrive at a station, often many tens of seconds behind the P-wave. However, Melgar et al. (2015a) noted that, in particular, when stations are close to the source such as in the 2010 M_w8.8 Maule earthquake, PGD algorithms produced reliable magnitude estimates in $\sim$60 s– well before the source process is complete ($\sim$150 s). Melgar and Hayes (2017) later reasoned that this was consistent with the weakly deterministic self-similar slip-pulse model of rupture, and that PGD should occur at near-source sites as soon as the slip pulse has propagated close to it and before the source process is complete. The magnitude-dependent temporal evolution of displacement ground motion amplitude has

previously been investigated using strong-motion seismic instrumentation (e.g. Colombelli et al., 2014). However, an assessment of the time behavior of geodetically-derived PGD has not been carried out. High-rate GNSS and seismogeodetic networks are still evolving and thus, there are only limited data sets. The only other synthesis of PGD observations (Melgar et al., 2015a) includes ten earthquakes across multiple tectonic settings, with only a few recordings for some events.

In this work, we present a systematic assessment of the temporal evolution of geodetically-derived PGD, which we refer to as PGD(t), and discuss its implications for the weak determinism model of rupture evolution and for rapid magnitude calculation and early warning. We limit the geographic reach of our study to Japan, reducing major global variation, and utilize the very dense GNSS Earth Observation Network (GEONET) operated by the Geospatial Information Authority (GSI, www.gsi.go.jp). We present the results from 14 medium to large magnitude events, M_w5.7-9.1 (Figure 4.1), each observed by between 177 and 700 GNSS stations (Table 4.1). Note that for this study we have only used Global Positioning System (GPS) observations, but will continue to use the term "GNSS," as observations from any other satellite constellation could be implemented in the same fashion. To maximize the number of observations, we produce 1 Hz GNSS waveforms for each event and time-align them to P-wave arrival times by interpolating arrivals from the overlapping strong motion networks (KiK-net and K-net) operated by the National Research Institute for Earth Science and Disaster Resilience (NIED, www.kyoshin.bosai.go.jp). From these dense displacement observations we show that PGD(t) is consistent with the weak determinism model. While initially PGD(t) behaves the same for events of all magnitudes, in the first tens of seconds, before the source is complete, there is clear separation in PGD(t) as a function of final magnitude. Finally, we create synthetic kinematic rupture models for thrust faults to demonstrate the rupture characteristics that contribute to early identification of magnitude and provide recommendations for bolstering early warning efforts in light of our findings.

4.2 Datasets

In order to study the temporal evolution of PGD we acquired 1 Hz RINEX data from GSI GNSS stations in Japan for 14 earthquakes. The data acquired were for ∼1200 GEONET stations, shown in Figure 4.1 (purple triangles) along with the locations of the 14 earthquakes included in the study. Earthquake details are given in Table 4.1. The earthquakes occurred between 2003 and 2016, with magnitudes between M_w5.7 and M_w9.1. We processed the GNSS data using precise point positioning (PPP) (Zumberge et al., 1997; Geng et al., 2013). Because the onset P-wave amplitudes are usually below GNSS noise, we needed an alternative way of determining the start of the GNSS displacement record at each site. To that end, we relied on the very dense K-NET and KiK-net strong motion networks. For each event, the P-wave arrival times were manually picked from the vertical waveforms of all available strong-motion sites. The Japanese strong-motion networks are operated in triggered-mode, and unfortunately, in some cases records begin after the P-wave arrivals. Incomplete records like these were excluded. We use the P-wave arrival times at seismic sites to time-align the GNSS waveforms. In this analysis of PGD evolution, we considered GNSS stations contained within the geographic footprint described by the properly triggered strong motion sites, relaxing the constraint that GNSS stations be collocated with a strong motion accelerometer. There are many GNSS stations with clear displacement waveforms, but without nearby strong motion station coverage from which we could sufficiently determine the P-wave onsets at that GNSS location. These sites were not considered in our analysis (e.g. Figure 4.2). This reduces the number of available sites from ∼1200 in the entire GNSS network to anywhere between 177 and 700 for each event (Table 4.1, Figure 4.2, Figures S4.1-S4.13). The hypocentral distance, R, of each source-station pair was computed from the catalog hypocenter location determined by the NIED.

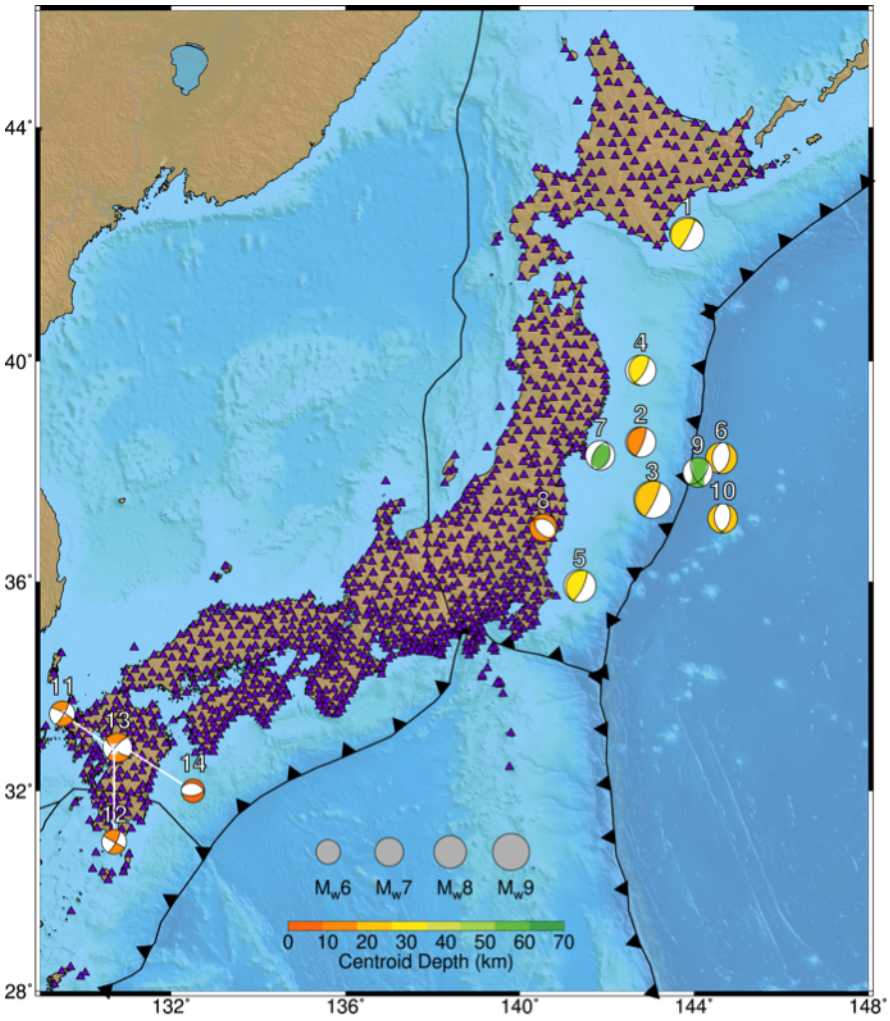


Figure 4.1: Map of GEONET stations (purple triangles) and the 14 earthquakes (focal mechanisms) considered in this study. Circle diameter corresponds to earthquake magnitude, and color refers to centroid depth. The number listed above each focal mechanism corresponds to the superscript in the event location column of 4.1.

Table 4.1: Earthquake event information for the 14 events included in the analysis. Superscript on location column corresponds to event number in Figure 4.1. Earthquake origin time and hypocenter location are from NIED. Origin time is given as Year-Month-Day Hour:Minute:Second. Magnitude is from Global CMT (globalCMT.org). The right-hand column denotes the number of GNSS stations recording the event within the area also covered by properly triggered strong-motion stations (for P-wave arrival interpolation).

Location	M_w	Origin Time (UTC)	Hypocenter			No.
			Lon (°E)	Lat (°N)	Depth (km)	Stations
Tokachi-oki[1]	8.3	2003-09-25 19:50:07.64	144.0785	41.7797	42.0	217
Miyagi[2]	7.3	2011-03-09 02:45:12.97	143.2798	38.3285	8.3	286
Tohoku-oki[3]	9.1	2011-03-11 05:46:18.12	142.3730	38.2970	29.0	493
Iwate[4]	7.4	2011-03-11 06:08:53.05	142.7815	39.8390	31.7	201
Ibaraki[5]	7.9	2011-03-11 06:15:34.46	141.2653	36.1083	43.2	340
N. Honshu[6]	7.6	2011-03-11 06:25:44.04	144.8940	37.8367	34.0	567
Miyagi[7]	7.1	2011-04-07 14:32:43.43	141.9237	38.2028	60.7	700
E. Fukushima[8]	6.7	2011-04-11 08:16:12.02	140.6727	36.9457	6.4	450
N. Honshu[9]	7.2	2012-12-07 08:18:20.28	144.3153	37.8158	46.0	575
N. Honshu[10]	7.1	2013-10-25 17:10:18.39	144.5687	37.1963	56.0	273
Kumamoto[11]	6.2	2016-04-14 12:26:34.43	130.8087	32.7414	11.4	226
Kumamoto[12]	6.0	2016-04-14 15:03:46.45	130.7777	32.7007	6.7	201
Kumamoto[13]	7.0	2016-04-15 16:25:05.47	130.7630	32.7545	12.5	254
Kumamoto[14]	5.7	2016-04-15 16:45:55.45	130.8990	32.8632	10.6	177

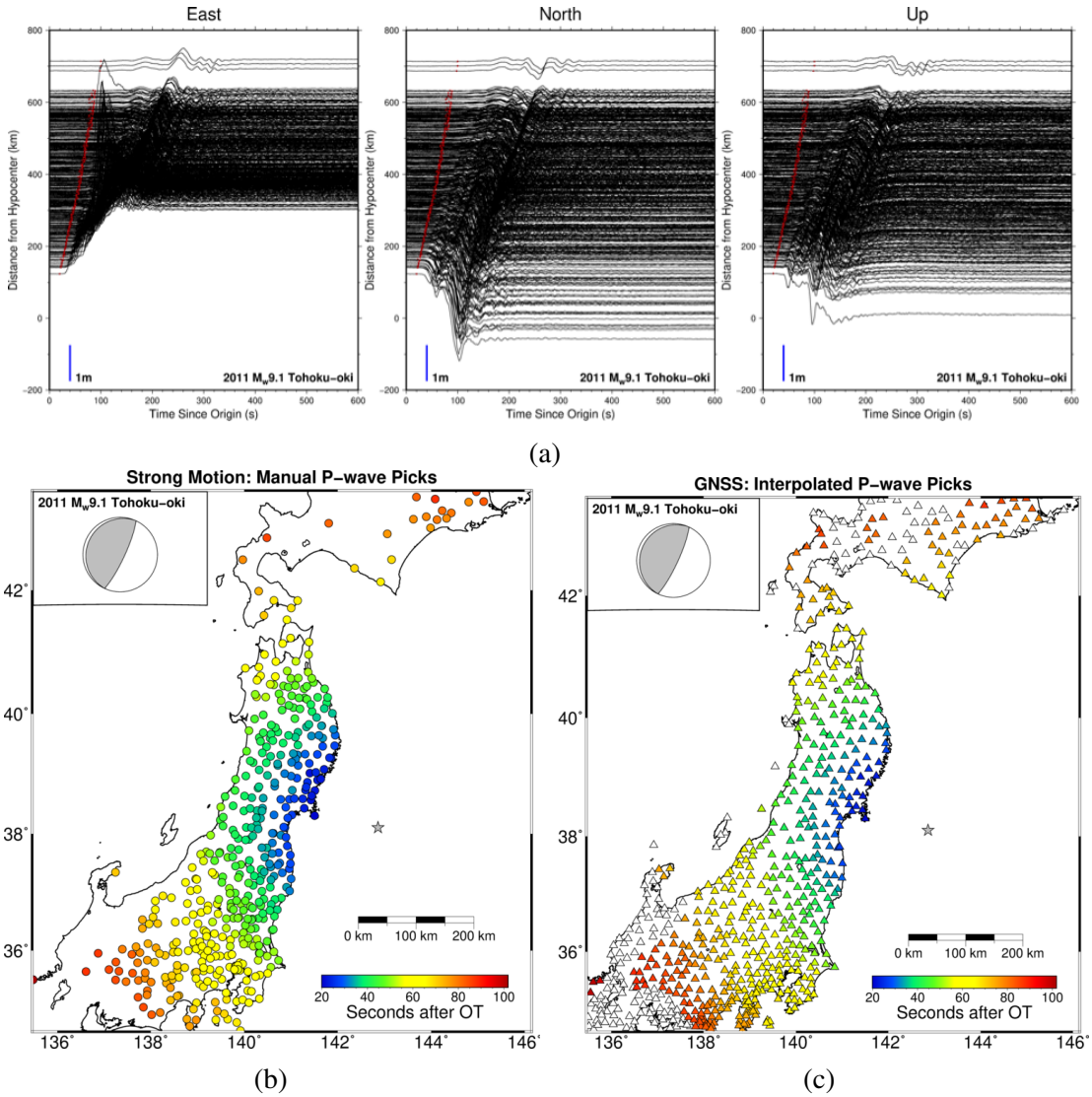


Figure 4.2: Example earthquake dataset, 2011 M_w9.1 Tohoku-oki earthquake. (a) PPP solutions for the three components of motion (left: East, center: North, right: Up) for all GNSS stations included in our analysis. Waveforms are offset by hypocentral distance. Red dots denote the P-wave arrival time assigned from (c). (b) Strong-motion seismic station locations (circles), colored by manually picked P-wave arrival times. (c) Available GNSS stations (triangles). Colored stations are within the footprint covered by strong-motion P-wave picks, and have been assigned a P-wave arrival time via interpolation from (b). White triangles are beyond the region covered by P-wave picks and were not included in subsequent analysis. Earthquake epicenter is denoted by grey star, with associated focal mechanism in top left corner of (b) and (c).

4.3 Methods

We now explore the evolution of peak displacement amplitude as a proxy for moment evolution. Equation 4.1 expresses the relation between maximum amplitude of displacement at the surface, *PGD*, hypocentral distance, *R*, and earthquake magnitude, M_{w} (e.g. Fowler, 2005). *A*, *B*, and *C* are constants to be estimated (Wu and Zhao, 2006; Crowell et al., 2013) given the known values of *PGD*, *R*, and M_{w}.

$$\log_{10}(PGD) = A + BM_w + C\log_{10}(R) \tag{4.1}$$

The distance attenuation term in Equation 4.1, $C\log_{10}(R)$, is sufficient for a point source approximation of rupture. While this assumption may hold for early observation times, later in rupture the interpretation of the rupture as a finite fault is required. To do so, we follow earlier studies, supplementing the distance term with a magnitude dependence (Crowell et al., 2013; Melgar et al., 2015a):

$$\log_{10}(PGD) = A + BM_w + CM_w\log_{10}(R) \tag{4.2}$$

It is unclear precisely when during rupture it becomes more appropriate to adjust from describing the earthquake as a point source to describing the earthquake as a finite fault, thus we present results using both source approximations. Throughout the text, emphasis will be made to identify results based on their source assumption: point source (Equation 4.1) or finite fault (Equation 4.2).

We have modified the scaling equations to include a weight matrix, *W*, that accounts for the different number of observations (data points) in different magnitude ranges. We divide the data into magnitude bins (M_{w} <7, 7<M_{w} <7.5, 7.5<M_{w} <8, 8<M_{w} <8.5, M_{w} >8.5) and weight by the inverse of the norm of PGDs in each bin. In this way, each bin becomes equally important in the resulting inversion and there is no bias towards preferentially fitting some part of

the magnitude range spanned by more events or by larger signals. We include only measurements of peak displacement that are above the expected GNSS noise. A typical value of GNSS noise is $\sim$1 cm in the horizontal direction and $\sim$5 cm in the vertical (Genrich and Bock, 2006). To avoid fitting stations whose observations do not exceed the noise, we remove all observations where peak displacement is smaller than 4 cm. Once a station's maximum observed displacement has exceeded 4 cm, it is introduced into the regression.

Coefficients A, B, and C in Equations 4.1 and 4.2 can be estimated from the observations at any point in time following P-wave onset. We use an L1-norm minimization scheme (Melgar et al., 2015a), to reduce sensitivity to outliers. We first construct the total displacement waveform as a function of time, $D(t)$, at each station such that

$$D(t) = \left(N(t)^2 + E(t)^2 + U(t)^2\right)^{1/2} \tag{4.3}$$

where $N(t)$, $E(t)$, and $U(t)$ are the north-south, east-west, and up-down displacement waveforms, respectively. Peak ground displacement as a function of time is then the maximum of $D(t)$, observed up to a given epoch

$$PGD(t) = \max[D(\tau)]; \qquad 0 < \tau \leq t \tag{4.4}$$

where 0 denotes the P-wave arrival time at a particular station. P-wave onsets at each GNSS station are interpolated from the arrival times observed at nearby strong motion sites (Figure 4.2). For the time dependent regression (Equations 4.1 and 4.2) we study the scaling properties of PGD observed over increasing windows following P-wave onset. We begin with a window of only 5 seconds (to investigate earthquake onset observations) and expand the window in 5-s intervals up to a final value of 170 s, when the final value of PGD(t) is achieved for all events and distances considered. For each 5 second window we carry out a regression for the best fitting set of coefficients (A, B, and C, Equations 4.1, 4.2). These coefficients at each 5-second interval

are calculated by randomly removing 10% of the stations from each earthquake, solving the regression, and repeating 100 times. In each iteration, we test the reliability of the relationship by invoking the removed 10% of stations to estimate each earthquake's magnitude for the relationship derived from the other 90% of data. We assess the amount of data subsequent to the P-wave arrival required for reliable magnitude estimation. This is different from previous approaches (Crowell et al., 2013, 2016; Melgar et al., 2015a), which have generated one set of coefficients with the final total value of PGD. In our formulation, A, B, and C vary as a function of observation time, hence PGD(t).

4.4 Results and Discussion

4.4.1 Evolution of Maximum Displacement

We performed the PGD(t) analysis for the 14 earthquakes in Japan (Table 4.1). For conciseness, we focus on the results using the finite fault assumption, which we suggest is the more appropriate assumption for the majority of the observation times. These results, discussed in detail below, are presented in Figure 4.3 and Table 4.2. The corresponding results for the point source solution can be found in Figure 4.4 and Table 4.3.

Figure 4.3 illustrates the best fitting relationships at several times after P-wave onset, beginning with only 5 s of data (Figure 4.3a) and increasing by 15 s in subsequent panels. Details of the best fitting coefficients at each time step are available in Table 4.2. At 5 s, only a small number of PGD observations exceed the noise threshold of 4 cm (Figure 4.3a), thus the best fitting PGD(t) relationships cannot be considered reliable and we are limited in drawing conclusions about these early stages. In each iteration, 10% of the data were removed prior to inversion for the scaling coefficients (A, B, and C). The estimated coefficients were then applied to the removed 10% to test reliability. The standard deviation of the residuals between known magnitude and calculated magnitude of the removed 10% is listed in the right hand column of Table 4.2.

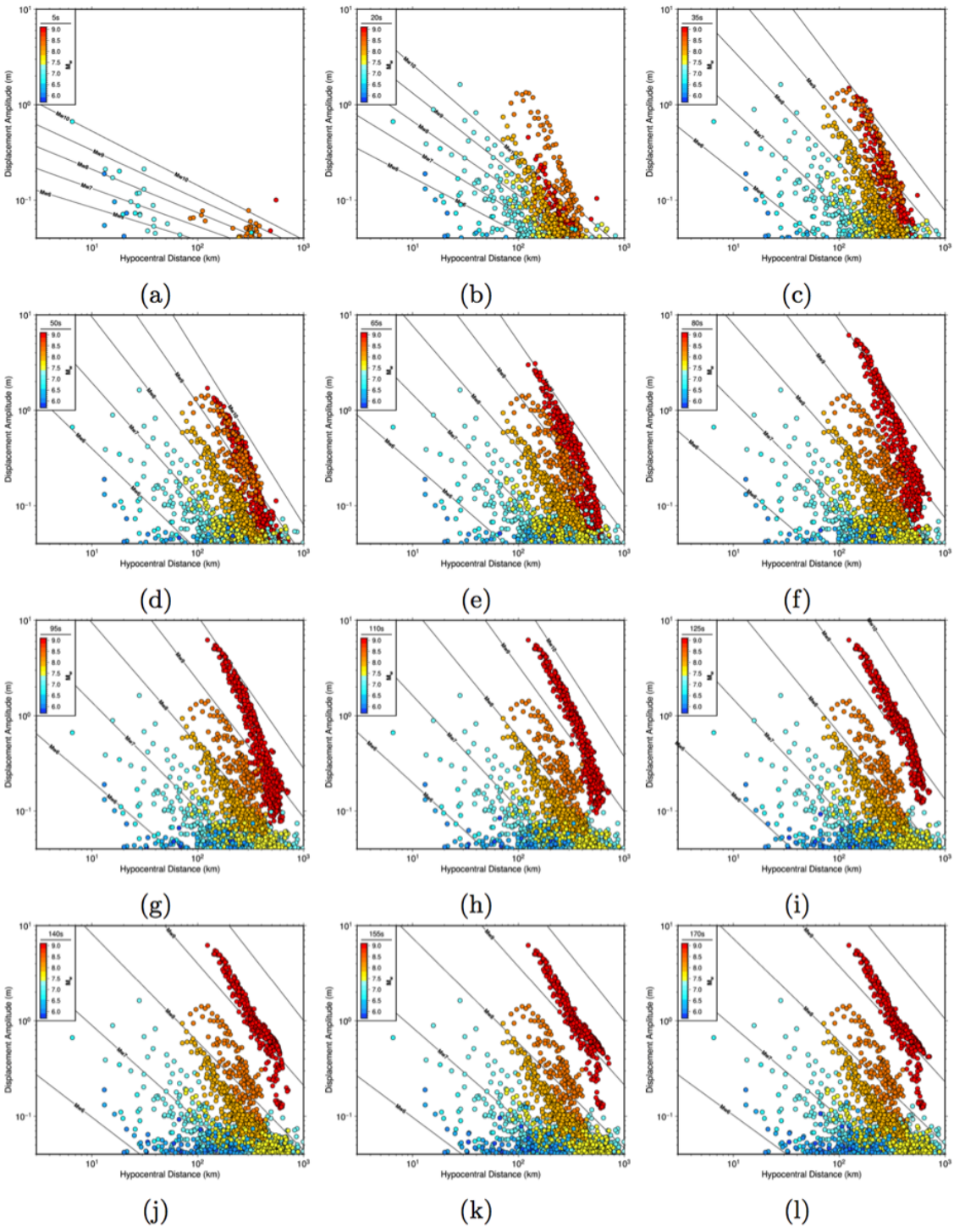


Figure 4.3: Best fitting magnitude scaling relations for the 14 considered events using Equation 4.2 (finite fault assumption), presented at 15-second intervals after P-wave arrival. Black lines denote the relation described by the best fitting coefficients A, B, and C at each time step. The lower limit of the y-axis is 4 cm, the chosen GNSS noise floor.

Table 4.2: Summary of best fitting magnitude scaling coefficients for a finite fault assumption (Equation 4.2) for windows of observation, in 5-second increments.

Window (s)	$A \pm \sigma_A$	$B \pm \sigma_B$	$C \pm \sigma_C$	σ^1_{err}
5	-2.276 ± 0.460	0.257 ± 0.086	-0.057 ± 0.013	1.602
10	-2.716 ± 0.175	0.381 ± 0.038	-0.085 ± 0.007	0.576
15	-1.903 ± 0.072	0.224 ± 0.015	-0.062 ± 0.003	1.463
20	-2.532 ± 0.137	0.392 ± 0.028	-0.097 ± 0.005	0.955
25	-3.065 ± 0.030	0.493 ± 0.006	-0.106 ± 0.001	0.720
30	-3.800 ± 0.051	0.626 ± 0.010	-0.121 ± 0.002	0.486
35	-4.472 ± 0.059	0.777 ± 0.015	-0.147 ± 0.004	0.433
40	-4.616 ± 0.053	0.826 ± 0.012	-0.159 ± 0.003	0.444
45	-4.446 ± 0.051	0.825 ± 0.012	-0.166 ± 0.003	0.493
50	-4.540 ± 0.069	0.868 ± 0.023	-0.178 ± 0.006	0.535
55	-4.662 ± 0.058	0.874 ± 0.016	-0.172 ± 0.005	0.467
60	-4.772 ± 0.073	0.864 ± 0.014	-0.160 ± 0.003	0.448
65	-4.881 ± 0.055	0.882 ± 0.009	-0.161 ± 0.002	0.414
70	-5.116 ± 0.060	0.913 ± 0.012	-0.160 ± 0.002	0.364
75	-5.397 ± 0.078	0.950 ± 0.017	-0.160 ± 0.003	0.366
80	-5.574 ± 0.061	0.968 ± 0.013	-0.158 ± 0.003	0.366
85	-5.635 ± 0.080	0.976 ± 0.018	-0.157 ± 0.004	0.354
90	-5.634 ± 0.060	0.982 ± 0.013	-0.159 ± 0.003	0.365
95	-5.695 ± 0.051	0.994 ± 0.010	-0.160 ± 0.002	0.376
100	-5.790 ± 0.062	1.015 ± 0.013	-0.163 ± 0.002	0.361
105	-5.933 ± 0.065	1.040 ± 0.013	-0.166 ± 0.002	0.368
110	-6.220 ± 0.079	1.089 ± 0.016	-0.170 ± 0.003	0.360
115	-6.407 ± 0.061	1.116 ± 0.013	-0.171 ± 0.003	0.359
120	-6.672 ± 0.052	1.155 ± 0.010	-0.172 ± 0.002	0.368
125	-6.847 ± 0.051	1.173 ± 0.011	-0.170 ± 0.002	0.365
130	-6.914 ± 0.033	1.160 ± 0.007	-0.159 ± 0.002	0.366
135	-6.870 ± 0.034	1.134 ± 0.007	-0.150 ± 0.001	0.363
140	-6.788 ± 0.028	1.105 ± 0.006	-0.142 ± 0.001	0.359
145	-6.746 ± 0.036	1.097 ± 0.007	-0.140 ± 0.001	0.366
150	-6.686 ± 0.025	1.085 ± 0.005	-0.139 ± 0.001	0.369
155	-6.657 ± 0.021	1.079 ± 0.004	-0.138 ± 0.001	0.367
160	-6.637 ± 0.020	1.075 ± 0.004	-0.137 ± 0.001	0.378
165	-6.598 ± 0.030	1.068 ± 0.007	-0.136 ± 0.002	0.362
170	-6.551 ± 0.029	1.062 ± 0.007	-0.135 ± 0.002	0.363

[1] Standard deviation of residuals between known magnitude and magnitude calculated with best fitting coefficients.

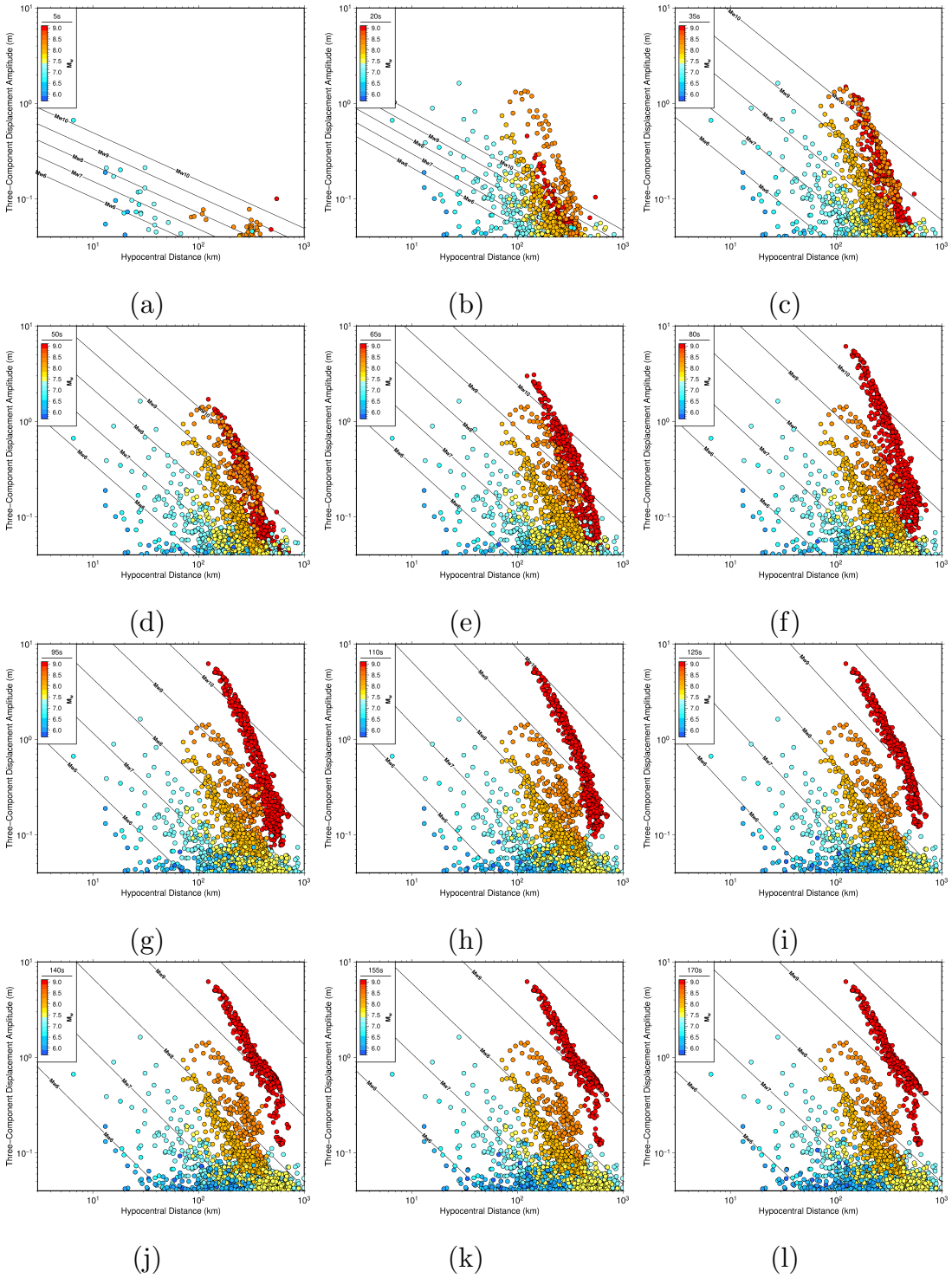


Figure 4.4: Best fitting magnitude scaling relations for the 14 considered events using Equation 4.1 (point source assumption), presented at 15-second intervals after P-wave arrival. Black lines denote the relation described by the best fitting coefficients A, B, and C at each time step. The lower limit of the y-axis is 4 cm, the chosen GNSS noise floor.

Table 4.3: Summary of best fitting magnitude scaling coefficients for a point source assumption (Equation 4.1) for windows of observation, in 5-second increments.

Window (s)	$A \pm \sigma_A$	$B \pm \sigma_B$	$C \pm \sigma_C$	σ_{err}
5	-1.505 ± 0.369	0.170 ± 0.078	-0.501 ± 0.115	0.948
10	-1.125 ± 0.095	0.135 ± 0.024	-0.527 ± 0.047	1.085
15	-0.926 ± 0.057	0.091 ± 0.010	-0.455 ± 0.023	1.430
20	-0.928 ± 0.063	0.146 ± 0.009	-0.620 ± 0.029	1.102
25	-1.365 ± 0.028	0.239 ± 0.005	-0.709 ± 0.010	0.688
30	-1.788 ± 0.053	0.323 ± 0.012	-0.791 ± 0.021	0.495
35	-2.081 ± 0.036	0.394 ± 0.006	-0.896 ± 0.015	0.412
40	-1.958 ± 0.075	0.406 ± 0.009	-0.976 ± 0.033	0.471
45	-1.762 ± 0.041	0.377 ± 0.008	-0.960 ± 0.025	0.511
50	-1.814 ± 0.059	0.390 ± 0.009	-0.968 ± 0.035	0.512
55	-2.072 ± 0.059	0.432 ± 0.007	-0.983 ± 0.037	0.429
60	-2.255 ± 0.058	0.453 ± 0.007	-0.961 ± 0.025	0.408
65	-2.203 ± 0.049	0.459 ± 0.006	-0.999 ± 0.027	0.392
70	-2.355 ± 0.072	0.488 ± 0.007	-1.023 ± 0.032	0.375
75	-2.538 ± 0.075	0.518 ± 0.008	-1.040 ± 0.036	0.354
80	-2.726 ± 0.060	0.538 ± 0.007	-1.024 ± 0.027	0.346
85	-2.775 ± 0.077	0.552 ± 0.010	-1.048 ± 0.025	0.364
90	-2.828 ± 0.065	0.562 ± 0.006	-1.053 ± 0.024	0.367
95	-2.834 ± 0.057	0.566 ± 0.008	-1.060 ± 0.031	0.351
100	-2.862 ± 0.056	0.577 ± 0.008	-1.081 ± 0.028	0.366
105	-3.014 ± 0.055	0.606 ± 0.008	-1.108 ± 0.027	0.363
110	-3.205 ± 0.039	0.647 ± 0.009	-1.157 ± 0.021	0.360
115	-3.341 ± 0.035	0.673 ± 0.009	-1.174 ± 0.027	0.345
120	-3.623 ± 0.050	0.714 ± 0.009	-1.186 ± 0.018	0.357
125	-3.790 ± 0.027	0.740 ± 0.006	-1.196 ± 0.019	0.371
130	-3.959 ± 0.025	0.756 ± 0.005	-1.174 ± 0.014	0.361
135	-4.045 ± 0.022	0.759 ± 0.005	-1.141 ± 0.018	0.362
140	-4.077 ± 0.018	0.750 ± 0.004	-1.096 ± 0.015	0.376
145	-4.085 ± 0.017	0.745 ± 0.005	-1.074 ± 0.014	0.380
150	-4.060 ± 0.015	0.740 ± 0.004	-1.065 ± 0.013	0.362
155	-4.055 ± 0.018	0.736 ± 0.004	-1.055 ± 0.013	0.371
160	-4.035 ± 0.021	0.731 ± 0.004	-1.044 ± 0.012	0.366
165	-4.001 ± 0.027	0.724 ± 0.004	-1.034 ± 0.011	0.380
170	-3.962 ± 0.027	0.719 ± 0.003	-1.033 ± 0.011	0.363

[1] Standard deviation of residuals between known magnitude and magnitude calculated with best fitting coefficients.

We require 20 s of data after the P-waves have arrived before the best fitting relation allows estimation within 1 magnitude unit (one-sigma), and 55 s of data before the residual consistently comes within 0.5 magnitude units. The final error is $\pm$0.36 magnitude units, consistent with previous studies (Melgar et al., 2015a). The size of the error implies that earthquake magnitude is indistinguishable from observations of displacement amplitude early in rupture, but becomes more reliable as stations record their final PGD value. As the evolution of peak displacement progresses, it is visually clear that magnitude is differentiated when displacement amplitude recorded from the smaller of the two earthquakes has achieved PGD, and the PGD(t) observations from the larger earthquake continue to increase. This is perhaps clearest in Figure 4.3c; all earthquakes $<$M$_w$8 appear separated, but the two largest events (2003 M$_w$8.3 Tokachi-oki and 2011 M$_w$9.1 Tohoku-oki) have overlapping distributions of data points. By Figure 4.3e, the GNSS stations nearest the M$_w$8.3 event (orange) have recorded peak displacement, and are now exceeded by the observations of the M$_w$9.1 (red) at the same hypocentral distances. As the later, highest amplitude seismic waves continue to propagate, this separation becomes apparent at increasing distances in subsequent panels. Similarly, because we do not see evidence of separation prior to final PGD, the best fitting scaling relations are poor at earlier stages. Evidence of the insufficient fit is notable in Figure 4.3e-g, where the M$_w$9.1 observations (red) approach M$_w$=10.

The timing of separation between events of different magnitudes is consistent with the amount of time it takes for the smaller of any two compared events to reach peak ground displacement. Our findings are therefore in agreement with the notion of universal initial (over the first few seconds) rupture behavior regardless of final magnitude. That is, these geodetic observations do not support the idea of strong determinism from earthquake onset parameters. However, we acknowledge limitations in the availability of near-fault data from the larger events in our study. Next, we consider the case for weak determinism at later rupture times.

4.4.2 Observational Timing of Peak Ground Displacement

We find that final PGD is the first reliable proxy for magnitude, therefore we must then address when, with respect to rupture initiation, we expect to observe final PGD. The answer places a lower limit on the timeliness of accurate magnitude estimation, and contributes to our understanding of earthquake development and determinism. Figure 4.5 shows the relationship between timing of PGD and hypocentral distance for a subset of the earthquakes considered (those with wide hypocentral distance coverage). The black dashed line denotes an estimate of the duration between P- and S-wave (S-P) arrival time at each distance for a 1D velocity model (Table 4.4). S-waves are responsible for larger ground motion, thus this line denotes the lower bound of the timing of PGD with respect to the P-wave arrival. Assuming the same Earth structure for each event, maximum ground displacement should follow shortly after this demarcation, regardless of earthquake size. From the observational data in Figure 4.5, two major features are apparent: first, the largest event (M_w9.1, red) takes considerably longer than the smaller events to reach PGD even at the same hypocentral distances, a feature previously noted in studies using strong motion observations (e.g. Colombelli et al., 2012); and second, for the remaining events, the observed time to PGD follows a trend with a shallower slope than the S-P line at short distances, but becomes similarly steep to that line at greater distances. Finally, the hypocentral distance at which the slope changes from shallower than the S-P line to similar to the S-P line is greater for larger magnitude events. The data points that appear below the S-P line may be due to a more complex velocity structure or low PGD amplitudes exceeded by GNSS noise earlier in the time series.

Considering that the largest earthquakes are offshore thrust events, they are primarily observed at long distances, with most observations further than 100 km from the hypocenter. However, understanding this relationship at smaller hypocentral distances is a crucial step toward discerning the minimum amount of time required to accurately measure PGD, and subsequently, estimate earthquake magnitude. We consider the physical basis for the observed relationship

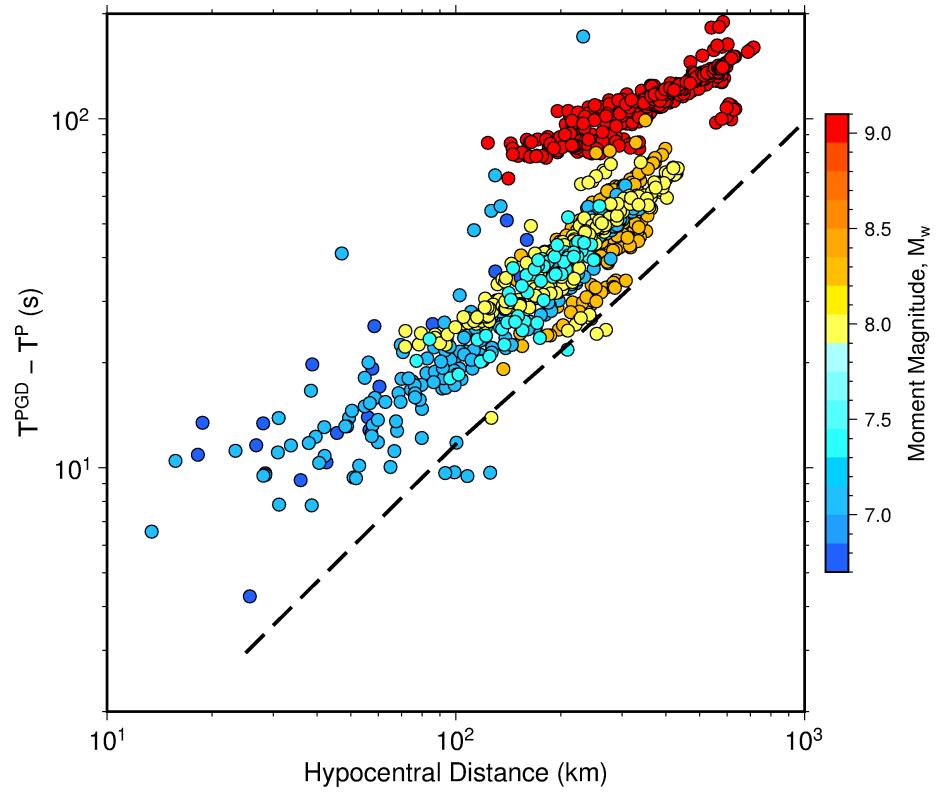


Figure 4.5: Observed time to peak ground displacement. Relationship between hypocentral distance and the time from P-wave arrival to PGD for a subset of the earthquakes considered; those with wide and overlapping hypocentral distance coverage. Black dashed line is the expected S-wave arrival minus P-wave arrival (S-P) time for a 1D velocity model (Table 4.4) and 25 km source depth. Each colored circle represents a single station, and is colored based on the magnitude of the observed earthquake.

between hypocentral distance and time to PGD. We hypothesize that at close distances, a station is most sensitive to slip on the portion of the finite fault closest to it, rather than the rupture surface as a whole. Displacement then should be related to the local duration of slip, or rise time. In turn, sites farther afield will be sensitive to the integrated signal from the fault as a whole and thus time to PGD will be related to the source duration. In other words, for distant stations, the fault can be approximated as a point source, while at closer stations, heterogeneity of the finite fault becomes important. This transition point from near- to far-field will depend on the frequency content of the radiated signal. The distance from the fault at which this change occurs may account for the observed change in slope in Figure 4.5.

Galetzka et al. (2015) demonstrated that near-field high-rate GNSS recordings of the 2015 M_w7.8 Nepal earthquake were most consistent with kinematic rupture characterized by a simple slip pulse. Similarly, Melgar and Hayes (2017) reported that rise time scales with earthquake magnitude. From these observations we further hypothesize that the observed pattern in Figure 4.5 is in part due to the different average rise times associated with increasing magnitude of these events. The notion of magnitude-dependent rise time is consistent with a weakly deterministic rupture model in which rupture organizes within the first tens of seconds into a slip pulse that has a width diagnostic of magnitude.

4.4.3 Synthetic Modeling of PGD Time

To examine the influence of rise time on the timing of peak displacement, we supplement the observations with synthetic modeling using the SW4 software (geodynamics.org/cig/software/sw4, Petersson and Sjögreen (2017a,b)). This allows us to specifically test the rupture characteristics that we hypothesize are responsible for our observations of PGD(t), and gain insight into behavior at distances not covered by our observational dataset. We test a simple model of a north-south striking, 20° west dipping thrust fault with homogeneous slip. Rupture begins at the center of the fault and propagates bilaterally toward the northern and southern ends. A grid of receivers is

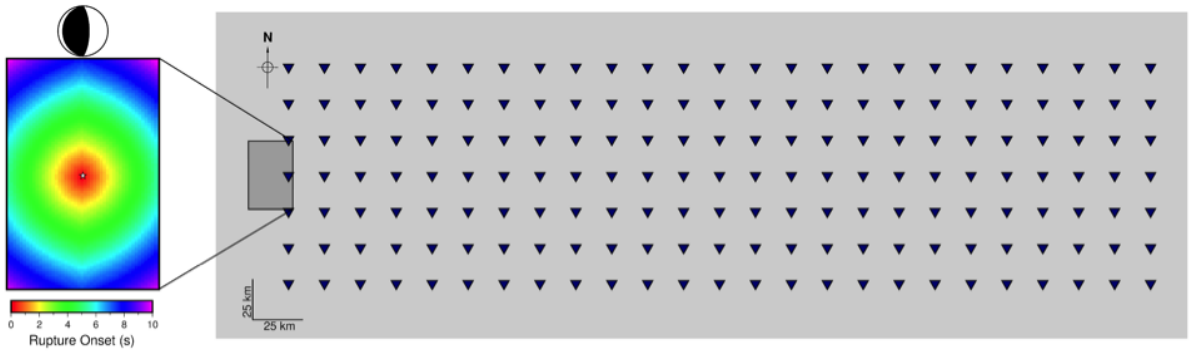


Figure 4.6: Schematic of subduction-zone event model design, shown here for the M_w7.0 test. The bilateral propagating fault strikes North-South and dips 20° west. Dip-slip (thrust) motion is uniform across the fault, with rupture onset determined by a rupture propagation velocity of 2.8 km/s. Fault is colored by rupture onset time, with white star representing the hypocenter location, at 25 km depth. Inverted triangles denote the receiver locations where synthetic observations are made. For larger magnitude input models, rows of receivers are added to the North and South of those shown, spanning at least 0.5 fault lengths to the North and South of the fault limits.

Table 4.4: One-dimensional Earth structure assumed for synthetic models. V_p and V_s are the P-wave and S-wave velocities, respectively. Q_p and Q_s are the P- and S-wave seismic attenuation.

Layer	Thickness (km)	V_p (km/s)	V_s (km/s)	Density (kg/m^3)	Q_p	Q_s
1	3	5.50	3.14	2.3	600	300
2	15	6.00	3.55	2.4	600	300
3	15	6.70	3.83	2.8	600	300
4	67	7.80	4.46	3.2	600	300
Half-space	∞	8.00	4.57	3.3	600	300

located along the overriding plate. Figure 4.6 depicts a schematic of the model set up. Our testing includes five models with varying magnitude (M_w6.5, M_w7.0, M_w7.5, M_w8.0, and M_w8.5). Each model differs from the others in two magnitude-dependent ways: fault dimensions (length and width) are determined from the subduction zone rupture scaling laws of Blaser et al. (2010), and the average rise time for each event is assigned from the scaling relation described in Melgar and Hayes (2017). Rupture speed is held fixed at 2.8 km/s and thus we assume that each of these models is a bilateral propagating slip pulse. We assume a 1D Earth structure, given in Table 4.4. Receivers measure displacement, and are located in an evenly spaced grid, between 25 and 500 km from the hypocenter.

Figure 4.7 shows the expected timing of PGD relative to the P-wave arrival for our suite of tests. At close distances, it is apparent that longer rise times and larger fault dimensions are associated with delayed PGDs. Indeed, this is the case on a station-by-station basis for all stations within ∼250 km of the hypocenter. Beyond that distance, surface waves become dominant, obscuring the pattern. Second, there appear to be two disparate slopes, one in the near-field and one in the far-field, similar to what is observed in the PGD measurements from GNSS displacement data (Figure 4.5). The location of the change between slopes appears to be related to magnitude as well, with this transition point occurring at greater hypocentral distances for larger magnitudes. Equation 4.1, which assumes a point source, is better equipped at distances beyond this change in slope, where the the finite fault dimensions are small compared to the distance between source and receiver. Modifications are likely required to properly assess observations in the near-field.

Figure 4.7 provides insight into how early it might be possible to determine magnitude from PGD measurements. The black dashed line indicates the S-P travel time for the assumed velocity model and a source depth of 25 km. The colored dashed lines in Figure 4.7 correspond to the sum of the S-P value and half-duration for the event of the same color. The M_w6.5 event has a rupture duration of only 5.5 s, and is complete prior to any PGD observations at the receiver stations. The M_w7.0 has a duration comparable to the time required to observe PGD, about 10 s. Above M_w7.0, there is a consistent pattern demonstrating that at close enough distances, the PGD metric is available prior to rupture completion. In the case of the M_w8.5 event, though rupture lasts ∼65 seconds, only about 35 seconds of observations following the P-wave are required to observe PGD at stations within ∼90 km of the hypocenter. Thus, our findings are consistent again with a weakly deterministic earthquake rupture model, wherein metrics which can differentiate the earthquake magnitude are available tens of seconds after first observations, and before completion of rupture. Our model is simplistic in its assumption of homogeneous slip, whereas a fault is more likely to experience heterogeneity in both slip and rise time, leading to additional variation in the

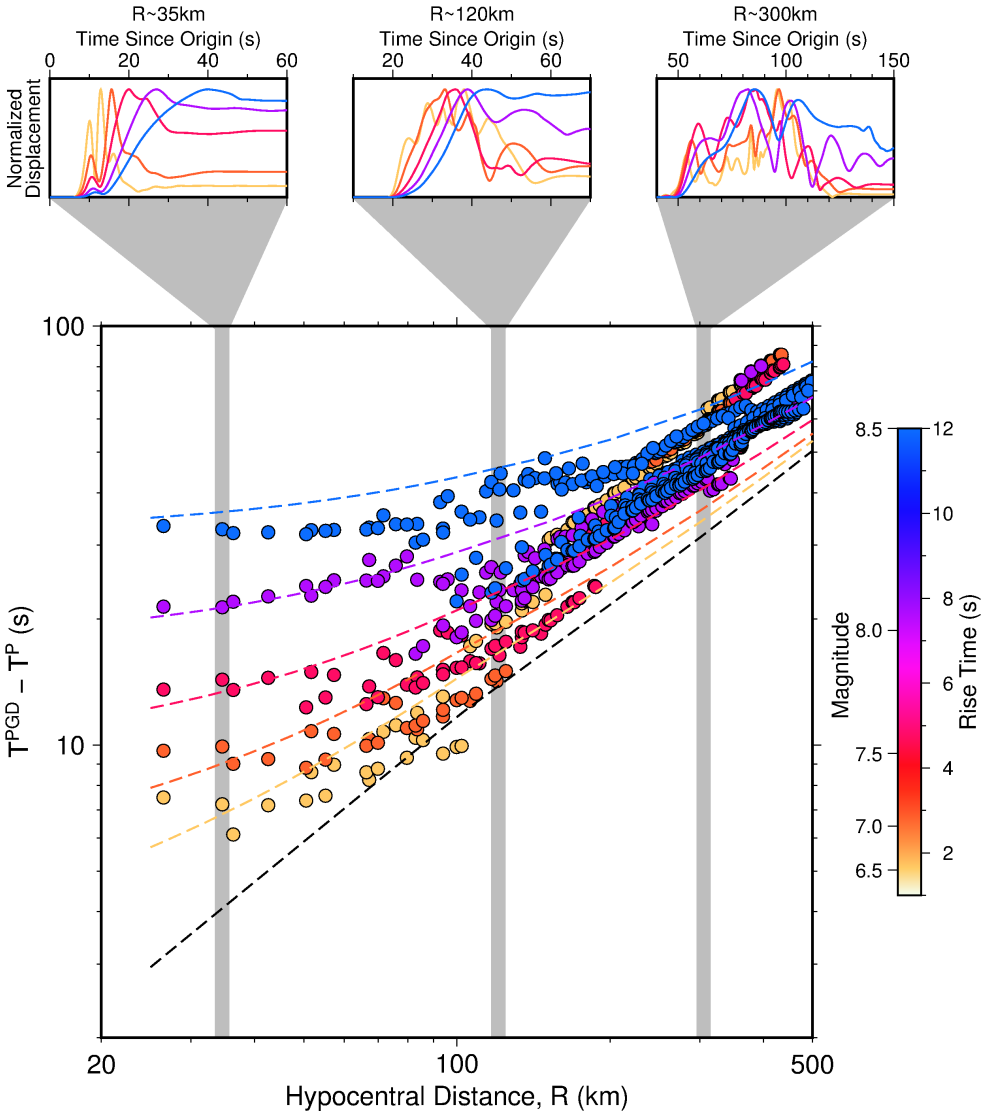


Figure 4.7: Modeled time to peak ground displacement. Timing of peak ground displacement relative to P-wave arrival time as a function of distance from the hypocenter for a bilateral rupture propagation. Earthquakes are modeled from M_w6.5 to M_w8.5, in 0.5 magnitude unit intervals, with associated fault dimensions (Blaser et al., 2010) and rise times (Melgar and Hayes, 2017). Each circle represents observations from a single receiver and is colored by the rise-time of the observed earthquake. The black dashed line is the S-P travel time for the 1D velocity model (Table 4.4) and 25 km source depth. Colored dashed lines represent the value of the half-duration of rupture added to the S-P travel time. Above, examples of waveforms from receivers at three representative hypocentral distances show the pattern of timing of peak ground displacement for the different rise time events. Each waveform is normalized to its corresponding PGD.

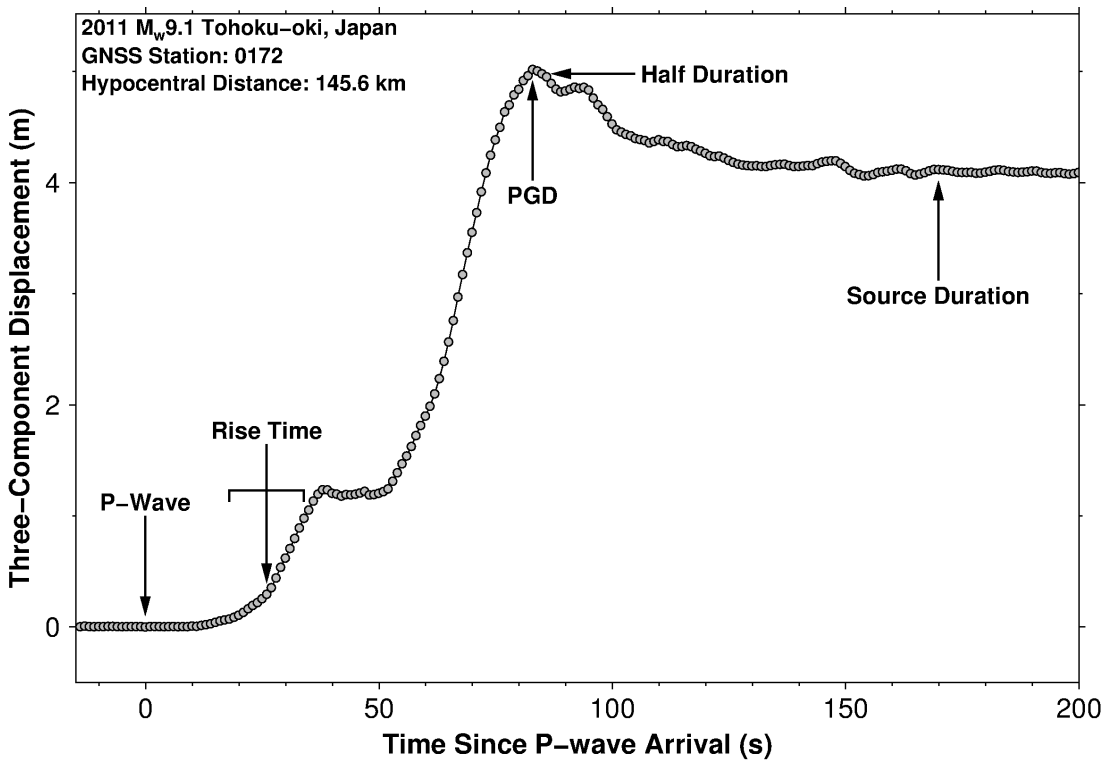


Figure 4.8: Time series of GNSS-derived displacements for GEONET station 0172, detailing the timing of earthquake parameters relative to the P-wave arrival during the 2011 M_w9.1 Tohoku-oki earthquake. Earthquake source duration is from the U.S. Geological Survey finite-fault source time function (Hayes, 2017) with average rise time (arrow) and standard deviation (bracket) from Melgar and Hayes (2017).

timing of PGD observation. Thus, our model represents an average of the expected behavior for earthquakes of the represented magnitudes.

We further demonstrate the timing of relevant parameters with an example displacement time series observed at GNSS station 0172, located 145.6 km from the 2011 M_w9.1 Tohoku-oki earthquake (Figure 4.8). We denote the timing of the observed PGD relative to the P-wave arrival, as well as the length of the earthquake-source properties including rise time, half duration, and source duration relative to the P-wave arrival. For the Tohoku-oki event, the full source duration is 170 s (Hayes, 2017) and the average rise time is 26 s with a standard deviation of 8 s (Melgar and Hayes, 2017). For this M_w9.1 event, the PGD metric is observed in slightly less time than the earthquake half-duration, even at a hypocentral distance of almost 150 km.

Large events most likely behave like slip pulses (Heaton, 1990; Melgar and Hayes, 2017) with rise time and width that scale with the eventual final magnitude of the event. Furthermore, GNSS stations will be most sensitive to the portion of the fault closest to them. As a result, we suggest that at close distances, stations are not required to observe the complete rupture in order to have magnitude-identifying qualities. Rather, they must only observe the passing of the slip-pulse, which is much shorter in duration than the full earthquake rupture. Thus, if stations are available close enough to the source, it will be possible to infer the magnitude of an earthquake prior to rupture completion. While our simulation was conducted for a thrust faulting environment, there is evidence that this pulse-slip behavior is exaggerated for long, narrow faults, such as continental strike-slip faults (Day, 1982; Zheng and Rice, 1998). Thus, our findings are especially relevant to the ongoing implementation of earthquake early warning in California, USA (Kohler et al., 2017). Evidence of weak determinism places enormous importance on the location of receivers capable of both timing the P-wave arrival (typically seismic instrumentation) and measuring accurate PGD (typically geodetic instrumentation) in near-fault locations and in real time. In the continental strike-slip regions, real-time networks have been designed to instrument the near-fault regions. For large subduction zone earthquakes that rupture mostly offshore and pose tsunami risk, such near-field measurements remain a challenge. It reinforces the need for real-time offshore seismic and geodetic instrumentation such as ocean bottom strong motion seismometers, absolute pressure, and GPS-acoustic to improve earthquake and tsunami early warning (e.g. Yoshioka and Matsuoka, 2013; Imano et al., 2015; Takahashi et al., 2015; Saito and Tsushima, 2016; Takahashi et al., 2015; Yokota et al., 2016).

4.5 Conclusions

Reliable magnitude estimation from earthquake early onset properties ($\sim$5 s) is not supported by our geodetic observations. However, earthquake magnitude is discernible prior to

rupture completion of the largest events using GNSS-derived peak displacements, indicating a weak determinism. Furthermore, the relationship between source-receiver distance and timing of maximum displacement amplitude demonstrate that our geodetic observations are consistent with a previously proposed source model describing rupture as a slip-pulse of magnitude-dependent width. Changes in slip-pulse width affect the timing of PGD observations. This timing provides a measure of the observation length required to compute an accurate magnitude estimate. Our findings suggest that high priority should be placed on installation of near-fault seismogeodetic instrumentation capable of both P-wave arrival detection and accurate displacement measurements, including ocean bottom seismometers, pressure sensors and GPS-acoustic seafloor instruments.

4.6 Supporting Information

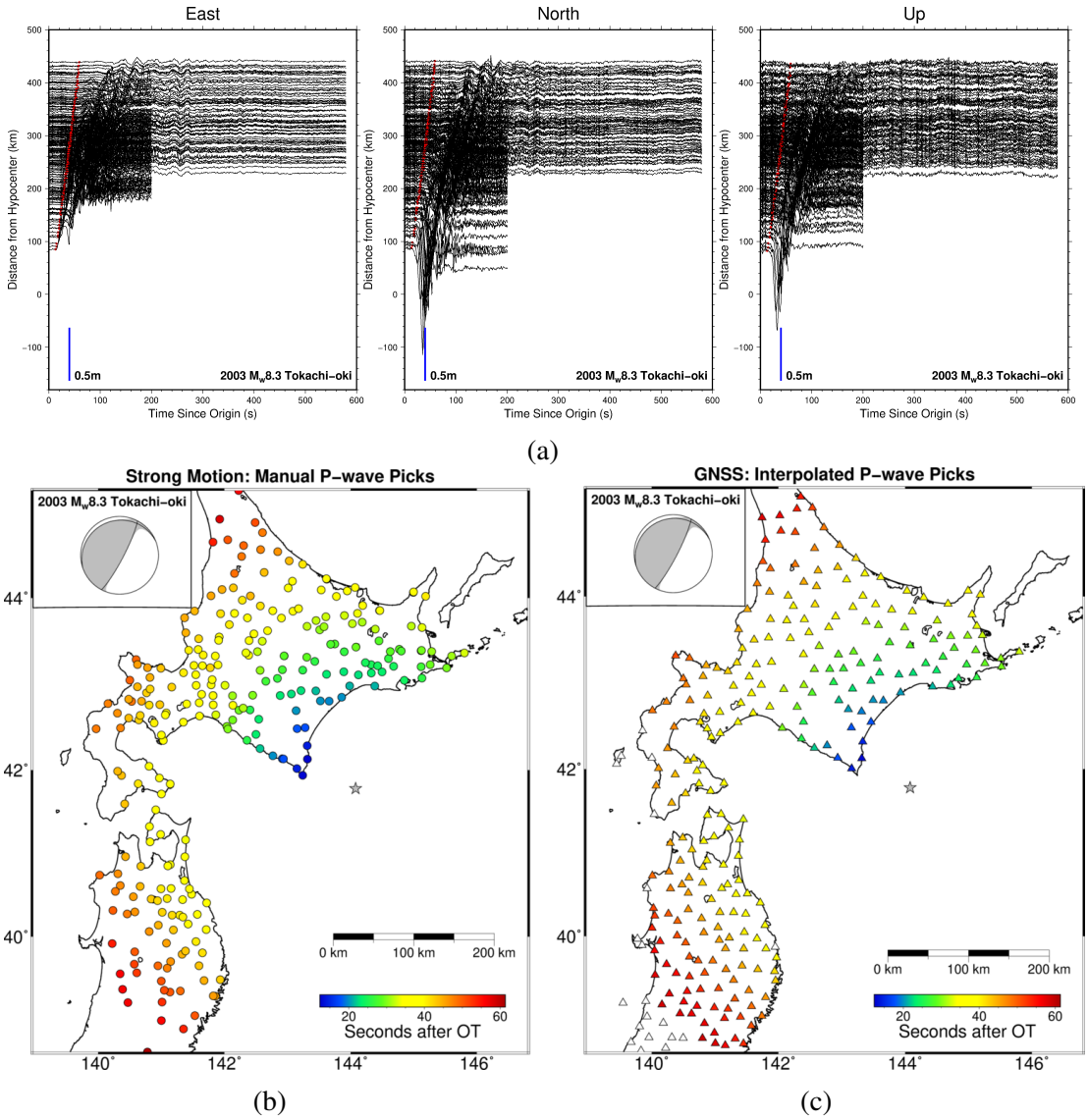


Figure S4.1: Earthquake dataset, 2003 M_w8.3 Tokachi-oki earthquake. (a) PPP solutions for the three components of motion (left: East, center: North, right: Up) for all GNSS stations included in our analysis. Waveforms are offset by hypocentral distance. Red dots denote the P-wave arrival time assigned from (c). (b) Strong-motion seismic station locations (circles), colored by manually picked P-wave arrival times. (c) Available GNSS stations (triangles). Colored stations are within the footprint covered by strong-motion P-wave picks, and have been assigned a P-wave arrival time via interpolation from (b). White triangles are beyond the region covered by P-wave picks and were not included in subsequent analysis. Earthquake epicenter is denoted by grey star, with associated focal mechanism in top left corner of (b) and (c).

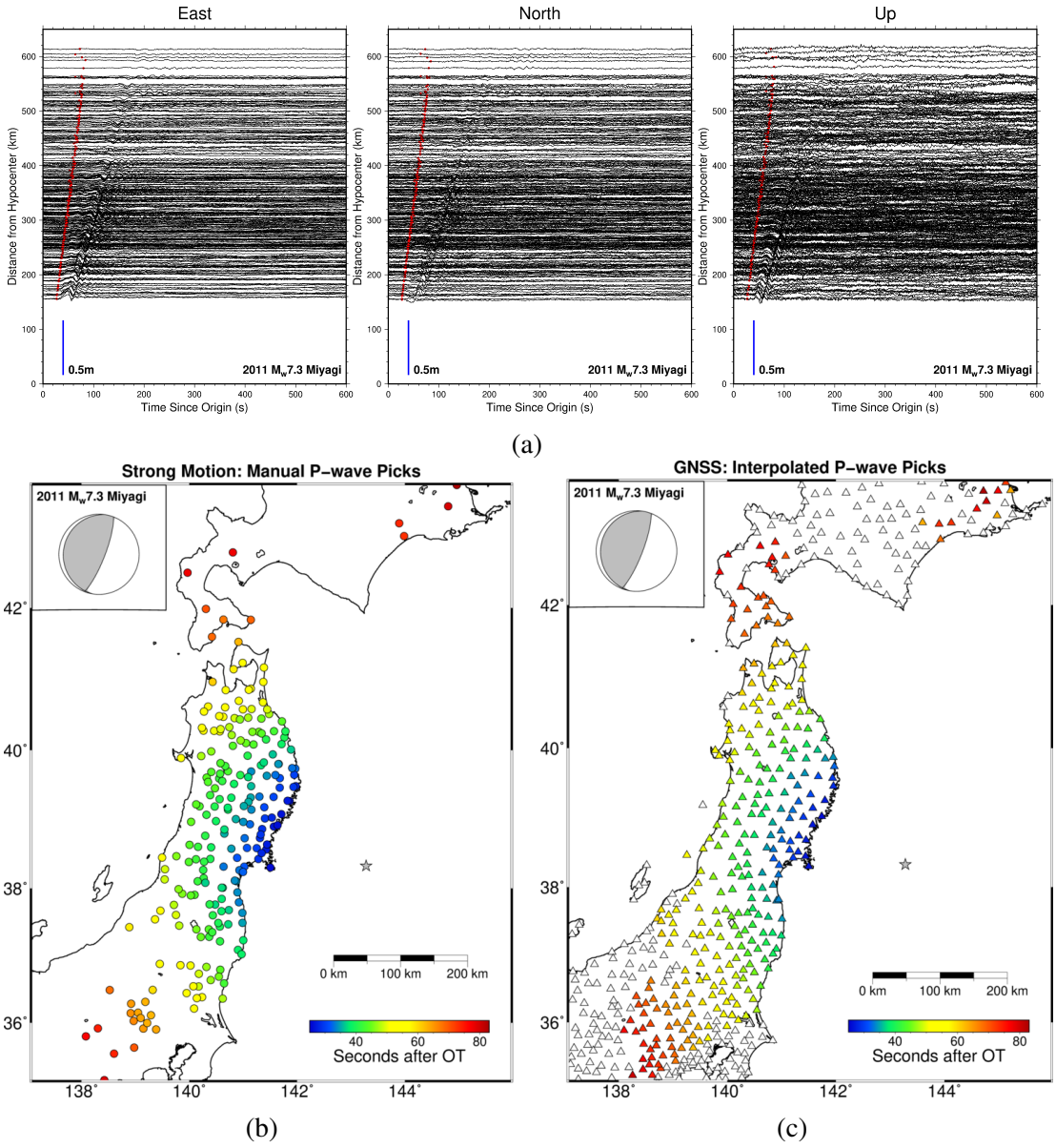


Figure S4.2: Earthquake dataset, 2011 M_w7.3 Miyagi earthquake. (a) PPP solutions for the three components of motion (left: East, center: North, right: Up) for all GNSS stations included in our analysis. Waveforms are offset by hypocentral distance. Red dots denote the P-wave arrival time assigned from (c). (b) Strong-motion seismic station locations (circles), colored by manually picked P-wave arrival times. (c) Available GNSS stations (triangles). Colored stations are within the footprint covered by strong-motion P-wave picks, and have been assigned a P-wave arrival time via interpolation from (b). White triangles are beyond the region covered by P-wave picks and were not included in subsequent analysis. Earthquake epicenter is denoted by grey star, with associated focal mechanism in top left corner of (b) and (c).

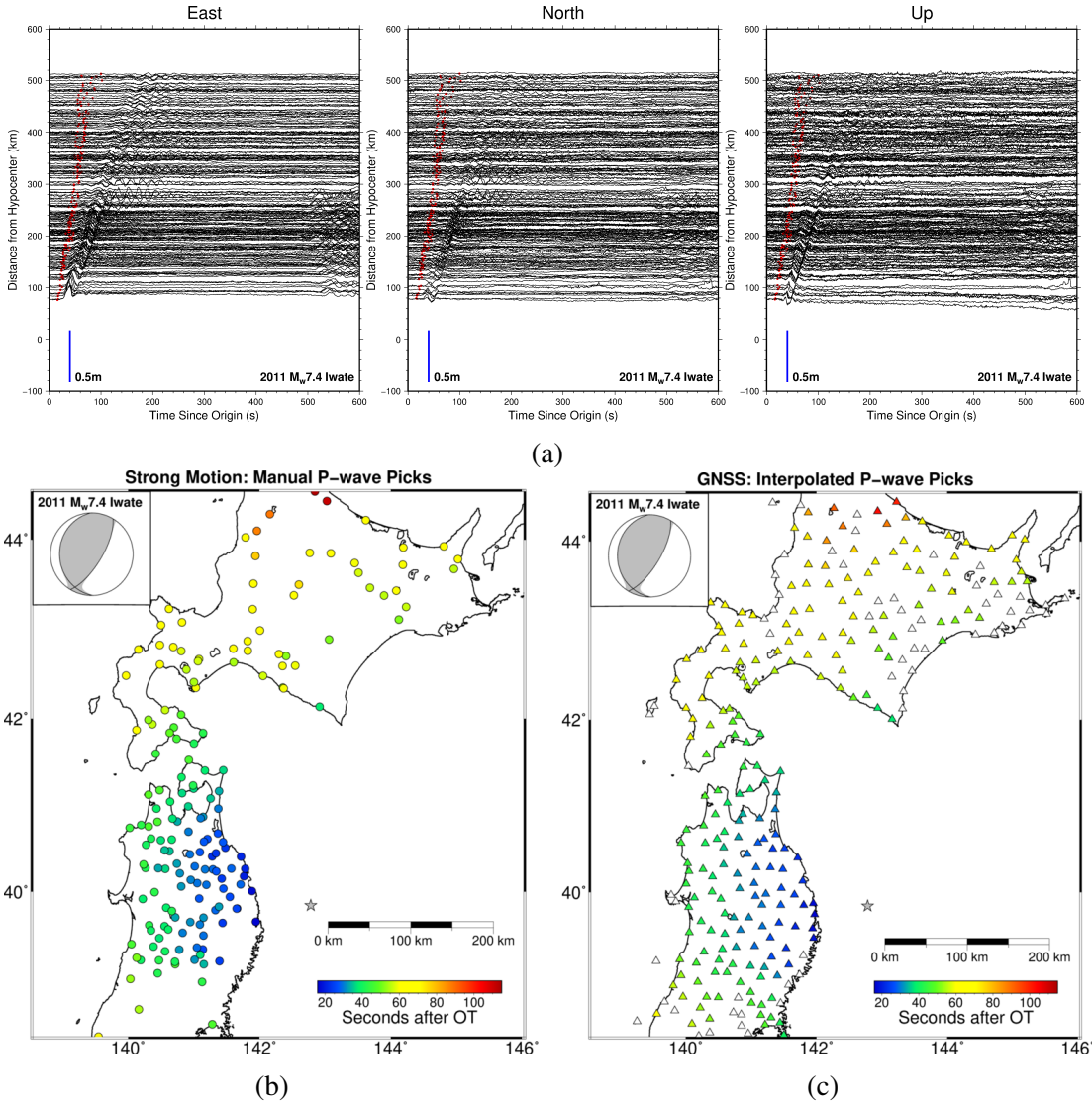


Figure S4.3: Earthquake dataset, 2011 M_w7.4 Iwate earthquake. (a) PPP solutions for the three components of motion (left: East, center: North, right: Up) for all GNSS stations included in our analysis. Waveforms are offset by hypocentral distance. Red dots denote the P-wave arrival time assigned from (c). (b) Strong-motion seismic station locations (circles), colored by manually picked P-wave arrival times. (c) Available GNSS stations (triangles). Colored stations are within the footprint covered by strong-motion P-wave picks, and have been assigned a P-wave arrival time via interpolation from (b). White triangles are beyond the region covered by P-wave picks and were not included in subsequent analysis. Earthquake epicenter is denoted by grey star, with associated focal mechanism in top left corner of (b) and (c).

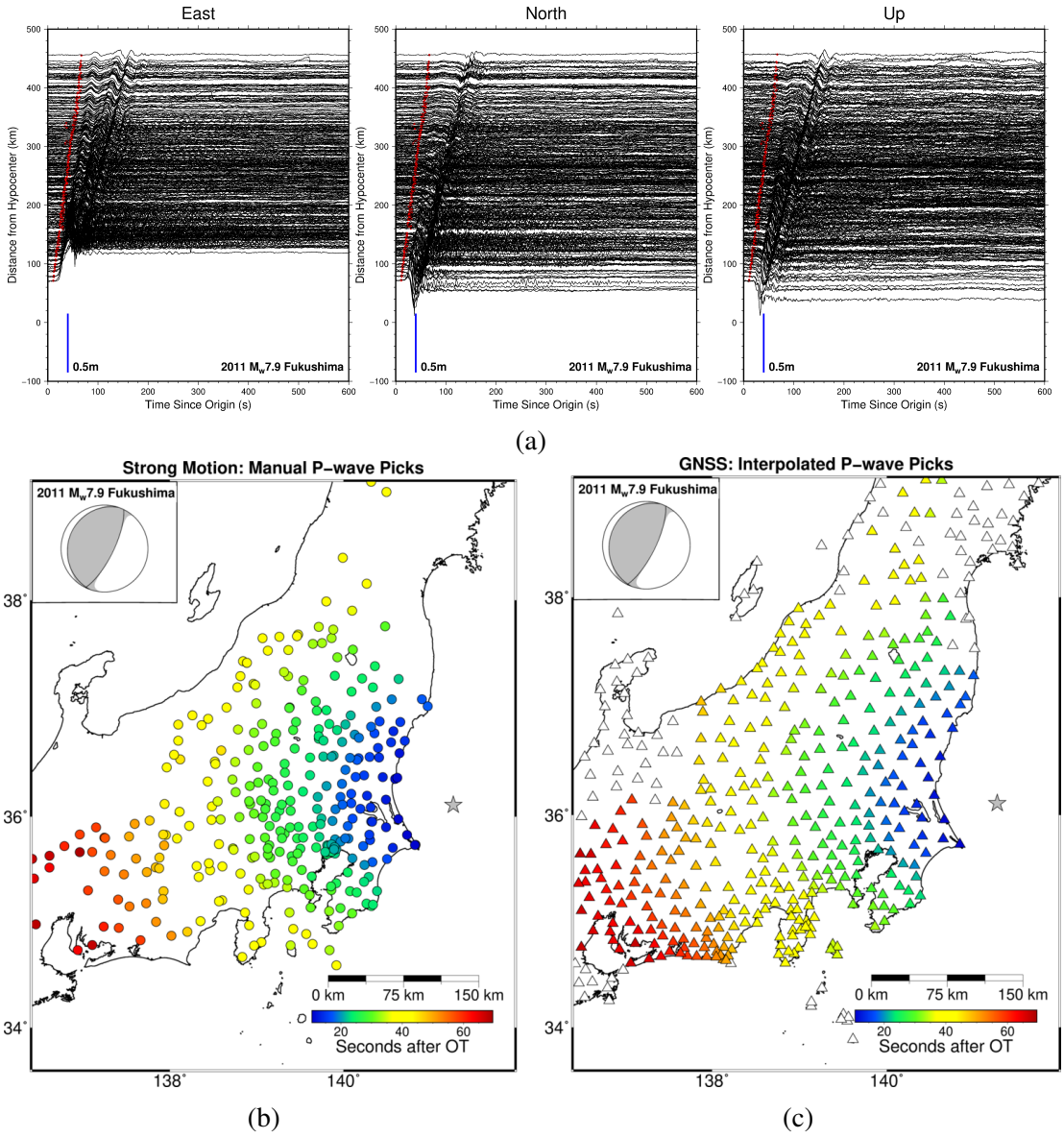


Figure S4.4: Earthquake dataset, 2011 M_w7.9 Fukushima earthquake. (a) PPP solutions for the three components of motion (left: East, center: North, right: Up) for all GNSS stations included in our analysis. Waveforms are offset by hypocentral distance. Red dots denote the P-wave arrival time assigned from (c). (b) Strong-motion seismic station locations (circles), colored by manually picked P-wave arrival times. (c) Available GNSS stations (triangles). Colored stations are within the footprint covered by strong-motion P-wave picks, and have been assigned a P-wave arrival time via interpolation from (b). White triangles are beyond the region covered by P-wave picks and were not included in subsequent analysis. Earthquake epicenter is denoted by grey star, with associated focal mechanism in top left corner of (b) and (c).

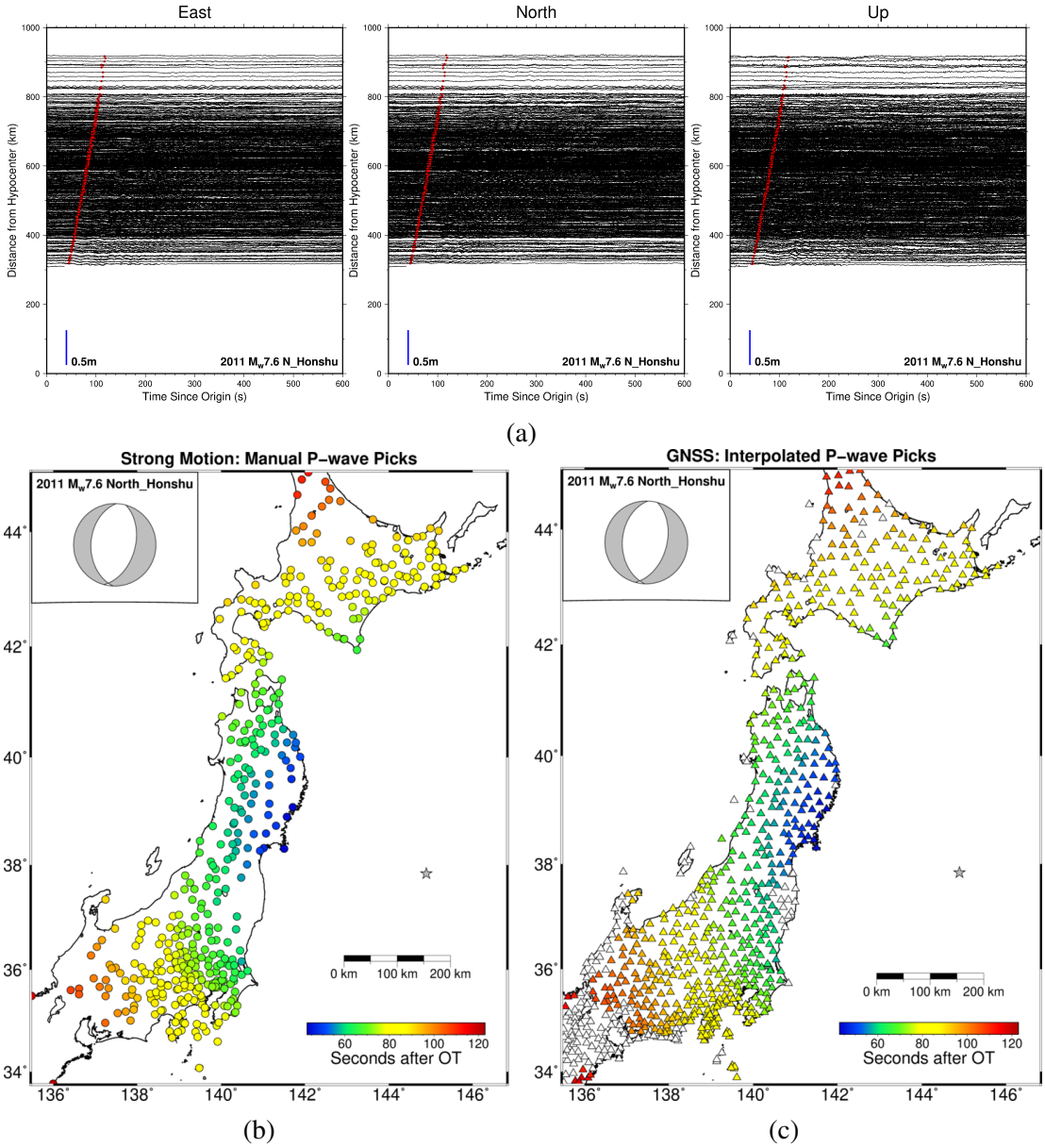


Figure S4.5: Earthquake dataset, 2011 M_w7.6 North Honshu earthquake. (a) PPP solutions for the three components of motion (left: East, center: North, right: Up) for all GNSS stations included in our analysis. Waveforms are offset by hypocentral distance. Red dots denote the P-wave arrival time assigned from (c). (b) Strong-motion seismic station locations (circles), colored by manually picked P-wave arrival times. (c) Available GNSS stations (triangles). Colored stations are within the footprint covered by strong-motion P-wave picks, and have been assigned a P-wave arrival time via interpolation from (b). White triangles are beyond the region covered by P-wave picks and were not included in subsequent analysis. Earthquake epicenter is denoted by grey star, with associated focal mechanism in top left corner of (b) and (c).

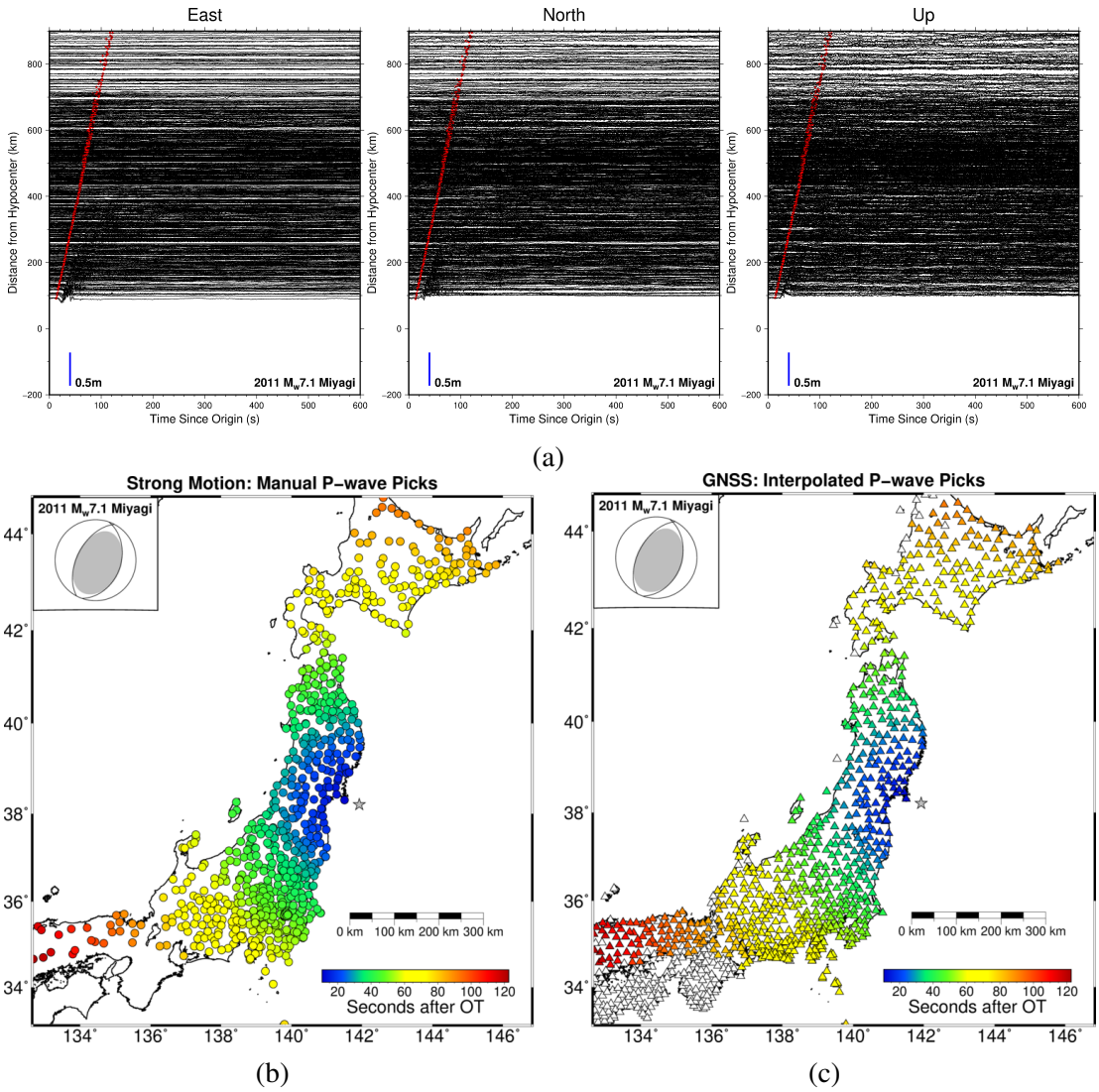


Figure S4.6: Earthquake dataset, 2011 M_w7.1 Miyagi earthquake. (a) PPP solutions for the three components of motion (left: East, center: North, right: Up) for all GNSS stations included in our analysis. Waveforms are offset by hypocentral distance. Red dots denote the P-wave arrival time assigned from (c). (b) Strong-motion seismic station locations (circles), colored by manually picked P-wave arrival times. (c) Available GNSS stations (triangles). Colored stations are within the footprint covered by strong-motion P-wave picks, and have been assigned a P-wave arrival time via interpolation from (b). White triangles are beyond the region covered by P-wave picks and were not included in subsequent analysis. Earthquake epicenter is denoted by grey star, with associated focal mechanism in top left corner of (b) and (c).

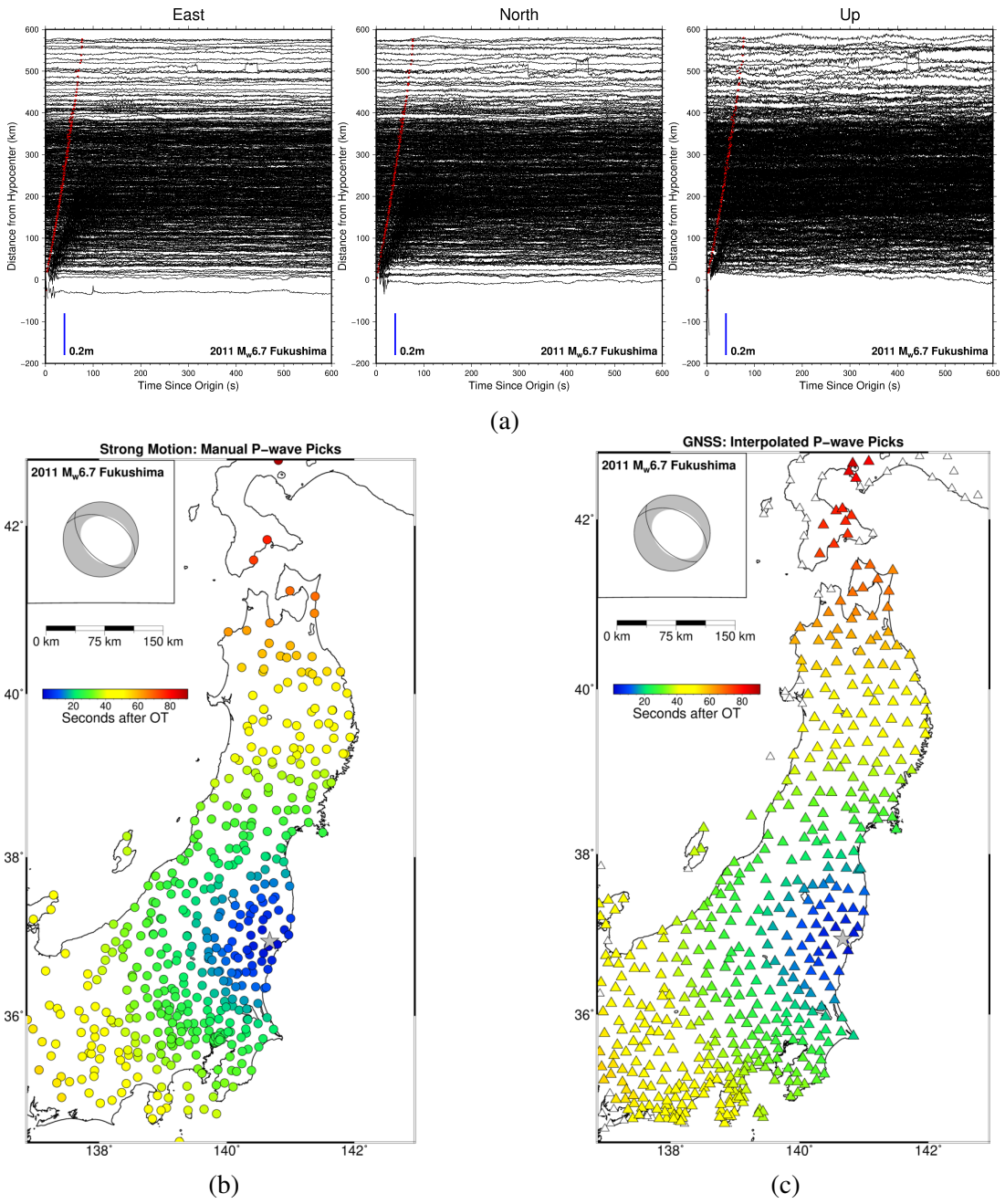


Figure S4.7: Earthquake dataset, 2011 M_w6.7 Fukushima earthquake. (a) PPP solutions for the three components of motion (left: East, center: North, right: Up) for all GNSS stations included in our analysis. Waveforms are offset by hypocentral distance. Red dots denote the P-wave arrival time assigned from (c). (b) Strong-motion seismic station locations (circles), colored by manually picked P-wave arrival times. (c) Available GNSS stations (triangles). Colored stations are within the footprint covered by strong-motion P-wave picks, and have been assigned a P-wave arrival time via interpolation from (b). White triangles are beyond the region covered by P-wave picks and were not included in subsequent analysis. Earthquake epicenter is denoted by grey star, with associated focal mechanism in top left corner of (b) and (c).

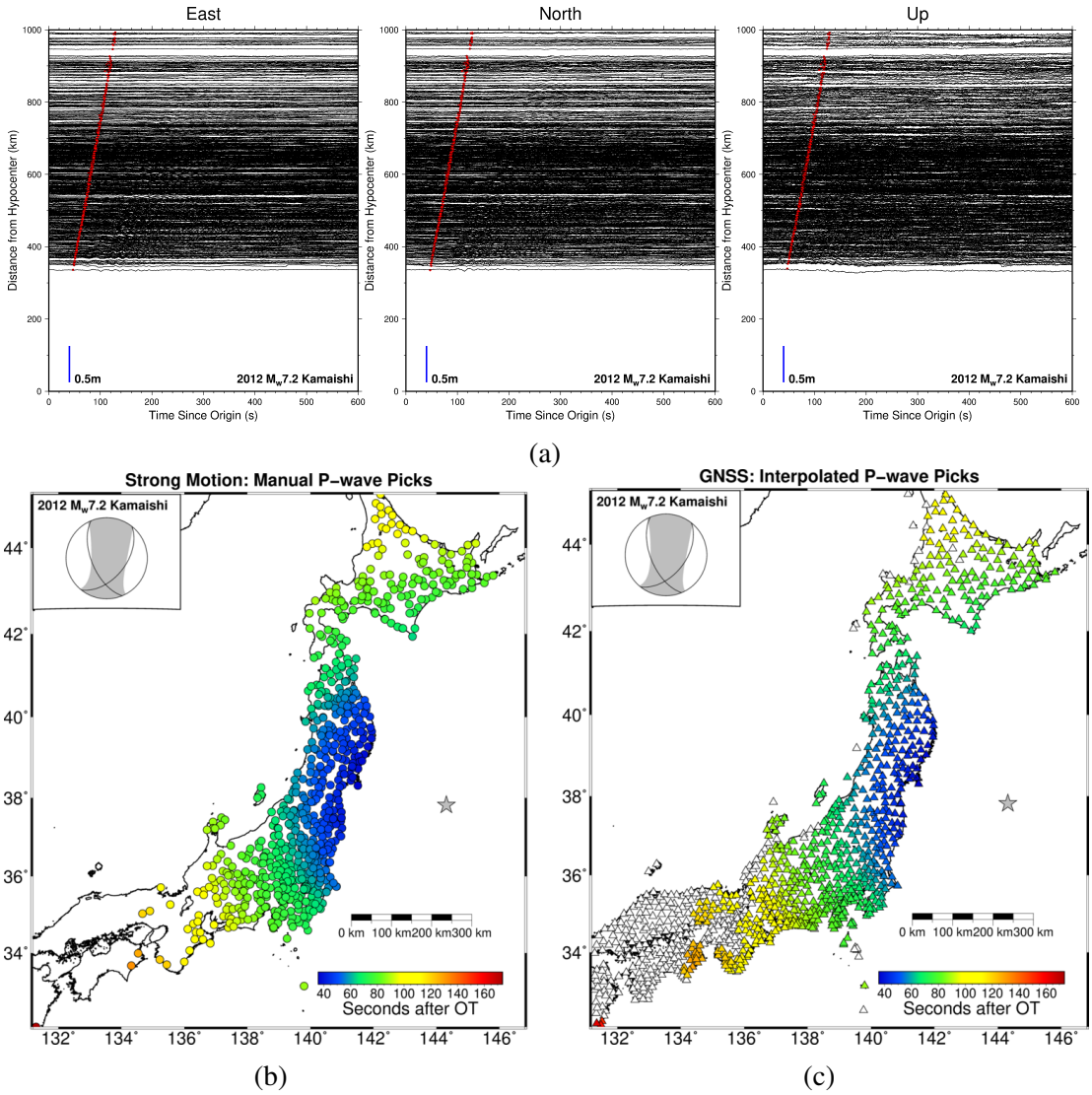


Figure S4.8: Earthquake dataset, 2012 M_w7.2 Kamaishi earthquake. (a) PPP solutions for the three components of motion (left: East, center: North, right: Up) for all GNSS stations included in our analysis. Waveforms are offset by hypocentral distance. Red dots denote the P-wave arrival time assigned from (c). (b) Strong-motion seismic station locations (circles), colored by manually picked P-wave arrival times. (c) Available GNSS stations (triangles). Colored stations are within the footprint covered by strong-motion P-wave picks, and have been assigned a P-wave arrival time via interpolation from (b). White triangles are beyond the region covered by P-wave picks and were not included in subsequent analysis. Earthquake epicenter is denoted by grey star, with associated focal mechanism in top left corner of (b) and (c).

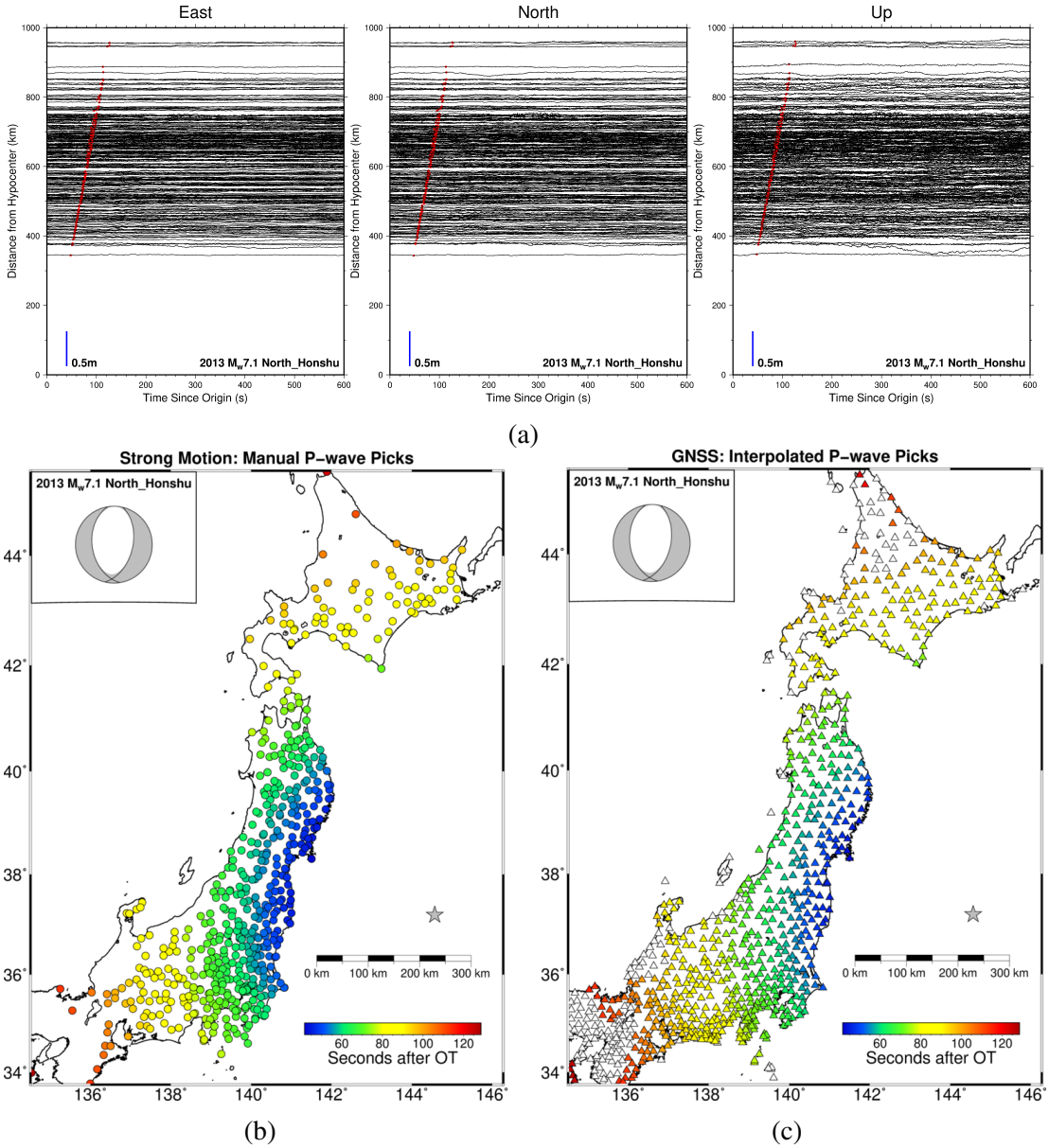


Figure S4.9: Earthquake dataset, 2013 M_w7.1 North Honshu earthquake. (a) PPP solutions for the three components of motion (left: East, center: North, right: Up) for all GNSS stations included in our analysis. Waveforms are offset by hypocentral distance. Red dots denote the P-wave arrival time assigned from (c). (b) Strong-motion seismic station locations (circles), colored by manually picked P-wave arrival times. (c) Available GNSS stations (triangles). Colored stations are within the footprint covered by strong-motion P-wave picks, and have been assigned a P-wave arrival time via interpolation from (b). White triangles are beyond the region covered by P-wave picks and were not included in subsequent analysis. Earthquake epicenter is denoted by grey star, with associated focal mechanism in top left corner of (b) and (c).

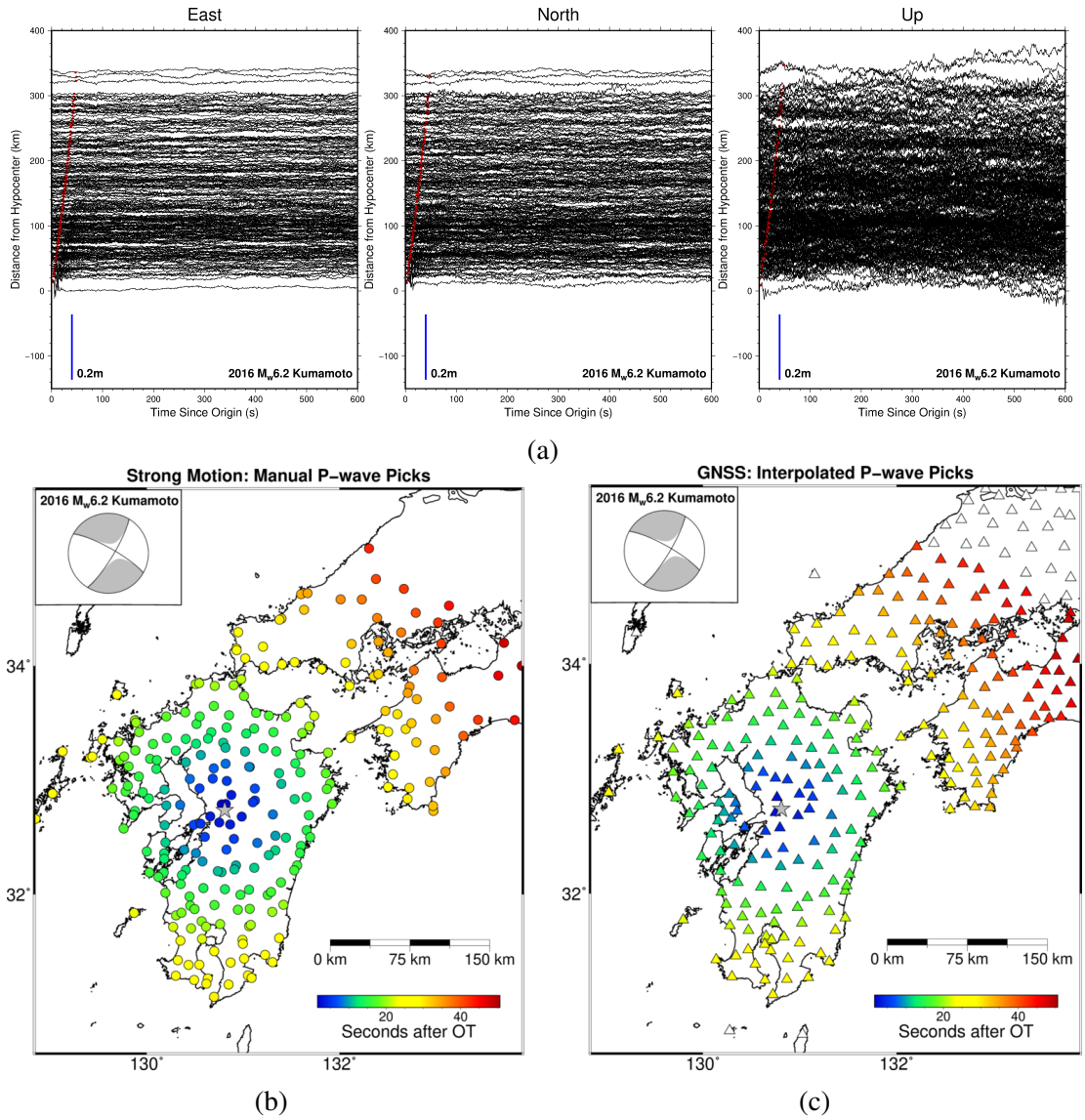


Figure S4.10: Earthquake dataset, 2016 M_w6.2 Kumamoto earthquake. (a) PPP solutions for the three components of motion (left: East, center: North, right: Up) for all GNSS stations included in our analysis. Waveforms are offset by hypocentral distance. Red dots denote the P-wave arrival time assigned from (c). (b) Strong-motion seismic station locations (circles), colored by manually picked P-wave arrival times. (c) Available GNSS stations (triangles). Colored stations are within the footprint covered by strong-motion P-wave picks, and have been assigned a P-wave arrival time via interpolation from (b). White triangles are beyond the region covered by P-wave picks and were not included in subsequent analysis. Earthquake epicenter is denoted by grey star, with associated focal mechanism in top left corner of (b) and (c).

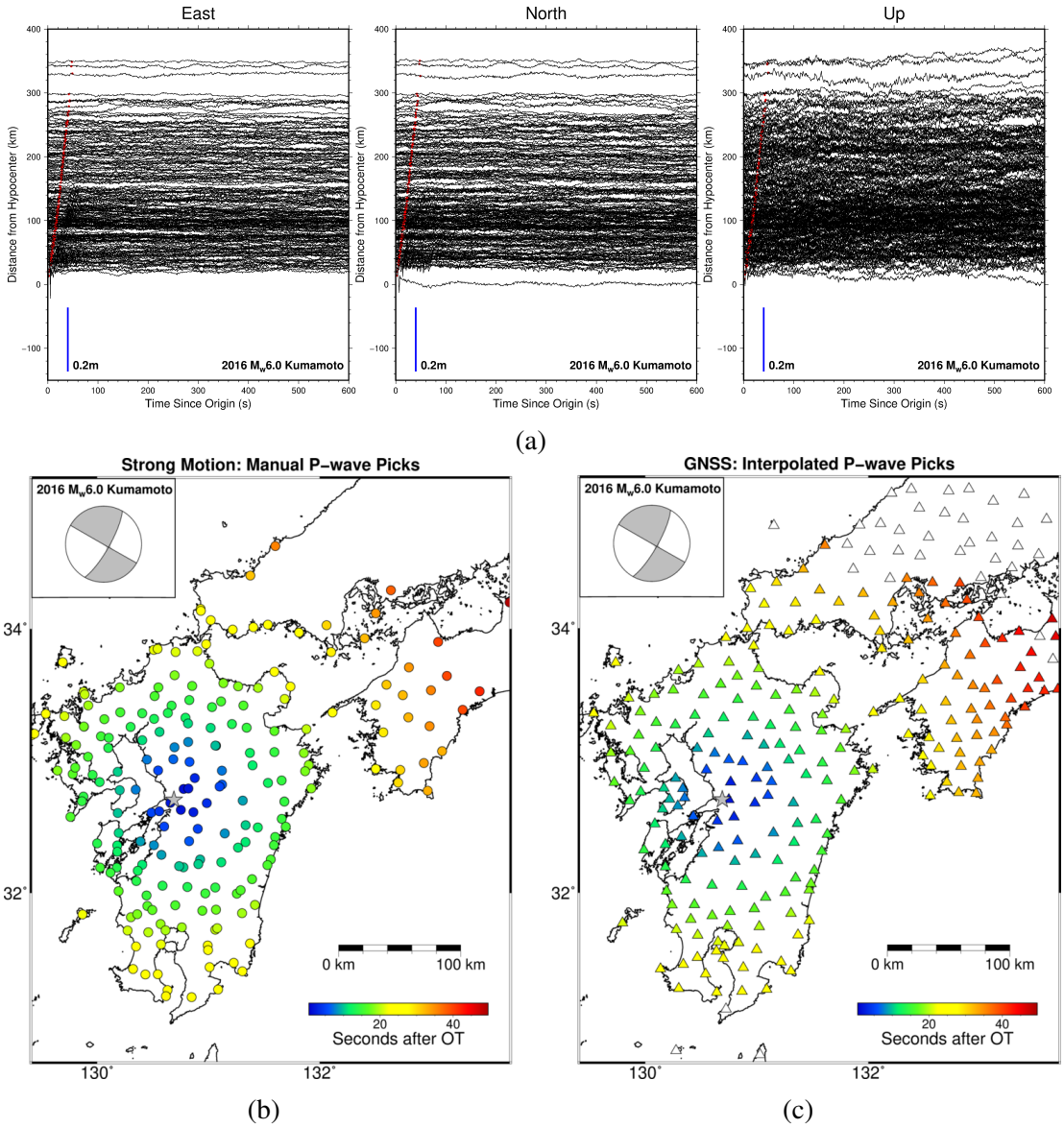


Figure S4.11: Earthquake dataset, 2016 M_w6.0 Kumamoto earthquake. (a) PPP solutions for the three components of motion (left: East, center: North, right: Up) for all GNSS stations included in our analysis. Waveforms are offset by hypocentral distance. Red dots denote the P-wave arrival time assigned from (c). (b) Strong-motion seismic station locations (circles), colored by manually picked P-wave arrival times. (c) Available GNSS stations (triangles). Colored stations are within the footprint covered by strong-motion P-wave picks, and have been assigned a P-wave arrival time via interpolation from (b). White triangles are beyond the region covered by P-wave picks and were not included in subsequent analysis. Earthquake epicenter is denoted by grey star, with associated focal mechanism in top left corner of (b) and (c).

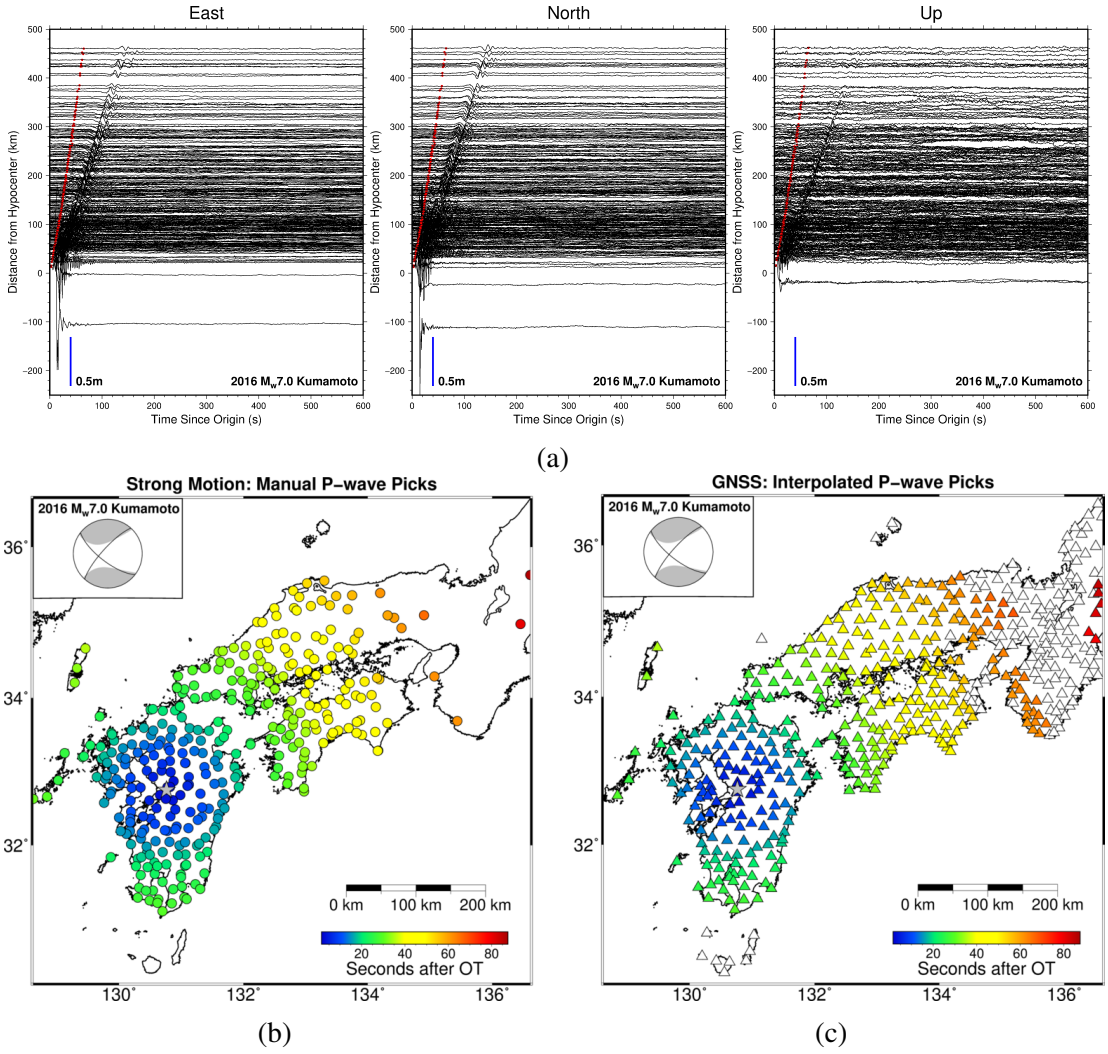


Figure S4.12: Earthquake dataset, 2016 M_w7.0 Kumamoto earthquake. (a) PPP solutions for the three components of motion (left: East, center: North, right: Up) for all GNSS stations included in our analysis. Waveforms are offset by hypocentral distance. Red dots denote the P-wave arrival time assigned from (c). (b) Strong-motion seismic station locations (circles), colored by manually picked P-wave arrival times. (c) Available GNSS stations (triangles). Colored stations are within the footprint covered by strong-motion P-wave picks, and have been assigned a P-wave arrival time via interpolation from (b). White triangles are beyond the region covered by P-wave picks and were not included in subsequent analysis. Earthquake epicenter is denoted by grey star, with associated focal mechanism in top left corner of (b) and (c).

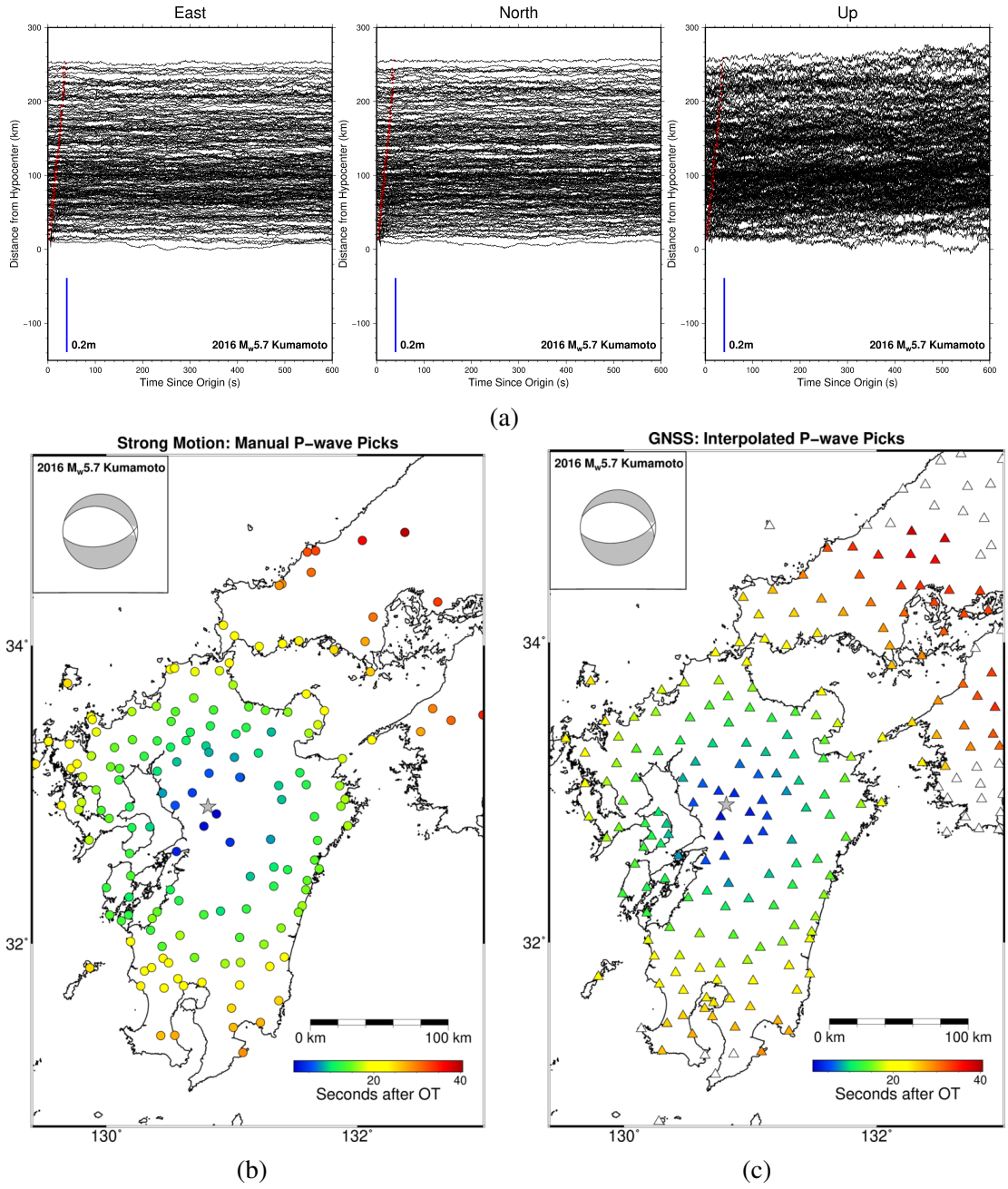


Figure S4.13: Earthquake dataset, 2016 M_w5.7 Kumamoto earthquake. (a) PPP solutions for the three components of motion (left: East, center: North, right: Up) for all GNSS stations included in our analysis. Waveforms are offset by hypocentral distance. Red dots denote the P-wave arrival time assigned from (c). (b) Strong-motion seismic station locations (circles), colored by manually picked P-wave arrival times. (c) Available GNSS stations (triangles). Colored stations are within the footprint covered by strong-motion P-wave picks, and have been assigned a P-wave arrival time via interpolation from (b). White triangles are beyond the region covered by P-wave picks and were not included in subsequent analysis. Earthquake epicenter is denoted by grey star, with associated focal mechanism in top left corner of (b) and (c).

4.7 Acknowledgements

This work has been submitted for publication as **Goldberg, D.E., Melgar, D., Bock, Y., and Allen, R.M (2018). Geodetic Observations of Rupture Evolution.** ***Journal of Geophysical Research: Solid Earth, in revision.*** The dissertation author was the primary investigator and author of this material. We are thankful for the contributions of Aldo Zollo, Thomas Hanks, an anonymous reviewer, and the Associate Editor of the Journal of Geophysical Research to this work. We gratefully acknowledge funding support from the National Aeronautics and Space Administration grants NNX14AQ53G, NNX16AM04A, and NNX13AT33G to UC San Diego, from the Gordon and Betty Moore Foundation grant GBMF3024 to UC Berkeley, and from the U.S. Geological Survey grants G16AC00348 and G17AC00346 to UC Berkeley. We thank the Geospatial Information Authority of Japan for the GNSS data and maintenance of the GNSS Earth Observation Network as well as the National Research Institute for Earth Science and Disaster Resilience for maintenance of the strong-motion networks, KiK-net and K-net. Figures were produced using Generic Mapping Tools (Wessel et al., 2013).

Chapter 5

Multisensor Seismic Structural Monitoring

5.1 Background

A familiar adage among seismologists is "earthquakes don't kill people, buildings kill people." Indeed, an earthquake in the absence of the built environment is unlikely to cause loss of life (tsunamigenesis is an important exception), whereas man-made infrastructure has amplified the destruction caused by a seismic event. Broken gas lines resulting from an earthquake cause fires and inadequate or nonexistent seismic structural codes result in fallen buildings. Even structures well designed for seismic threat may experience degradation over time in response to repeated seismic activity or natural building aging. Following an earthquake or other natural hazard, it is standard to conduct visual inspection of buildings in the affected area to determine whether they are safe for reentry, or must be "red tagged" until they are properly reinforced. Currently, the University of California San Diego (UCSD) Facilities Design and Construction Engineering Services (FD&C) is responsible for performing damage assessment following a natural hazard, to determine whether any university buildings have sustained damage compromising to their structural integrity. The inspections, which are primarily visual, lack objective, numerical, data driven metrics from which to establish those determinations.

(a) (b)

Figure 5.1: Geisel Library baseline UAV imagery, July 28, 2017. (a) View of library rooftop from UAV imagery. North direction is approximately to the right of the image. (b) Additional view of the Geisel Library from baseline UAV survey. The snake head walkway (bottom right of image) is approximately due west of the library structure.

We devise a multi-sensor approach to structural health monitoring, to better inform the decision making of the FD&C following an earthquake. We describe testing and development underway at the UCSD Geisel Library, an eight-story reinforced concrete structure 100 feet high and 200 feet in diameter. In collaboration with researchers at the UCSD Qualcomm Institute and KDM Meridian, we employ exterior imaging of the structure from unmanned aerial vehicle (UAV) photography and Light Detection and Ranging (LIDAR) surveys to create a digital surrogate (cybermodel) of the building, which can be used as a basis of comparison following any natural hazard (Figure 5.1). The digital surrogate is georeferenced using continuous Global Positioning System (GPS) and and conventional land surveying methods. The UCSD Geisel Library monitoring project is part of an initial proof-of-concept phase of testing. A successful culmination of this work will result in the expansion of this methodology to other vulnerable buildings across the UCSD campus and beyond.

The Scripps Orbit and Permanent Array Center (SOPAC) has developed hardware to upgrade continuous GPS stations to seismogeodetic capability. The Scripps Institution of Oceanography (SIO) Geodetic Module (GM) is a low-power, microprocessor-based, internet-capable device that controls real-time data streams from any geodetic GPS receiver and the SIO micro-electro-mechanical systems (MEMS) accelerometer package. The SIO MEMS accelerometer is a

(a) (b) (c)

Figure 5.2: Seismogeodetic instrumentation at Geisel Library, UC San Diego (a) Geisel library reference station, LBRF, consisting of deep drilled braced GPS monument and SIO MEMS accelerometer, affixed to the vertical leg of the GPS monument, located behind the Biology Field Station on the UCSD campus. (b) GPS and SIO MEMS accelerometer installed on the rooftop of the Geisel Library. (c) Top of GPS radome, with temporary target applied to facilitate georeferencing of UAV imagery to precisely known location at center of radome.

+/-2g, 16-bit, low power, waterproof, three-axis, undamped acceleration sensor (Saunders et al., 2016). Monitoring of the Geisel Library includes five such seismogeodetic stations, described below.

The library reference station, LBRF, is located on the UCSD campus near the Biology Field Station. LBRF is a continuous GPS station, with an SIO MEMS accelerometer affixed to the vertical brace of the drill-braced monument and includes a GM for real-time data streaming (Figure 5.2a). The Geisel library itself is instrumented with three GM and Accelerometer Package (GAP) stations on the roof (Figure 5.2b): Library North (LBNO), Library Southeast (LBSE) and Library Southwest (LBSW). An additional SIO MEMS accelerometer is located in the basement (Library Basement, LBBA). Station LBBA is time-tagged and telemetered via connection to the GM tied to rooftop station LBNO. The rooftop stations have multiple utilities for natural hazard structural monitoring: (1) the GPS radome locations, which are known with millimeter-level

Figure 5.3: Geisel Library LIDAR and UAV survey. LIDAR (main image) survey conducted by KDM Meridian and UAV (inset, top right) survey conducted by the Qualcomm Institute at the UCSD Geisel Library on July 28, 2017.

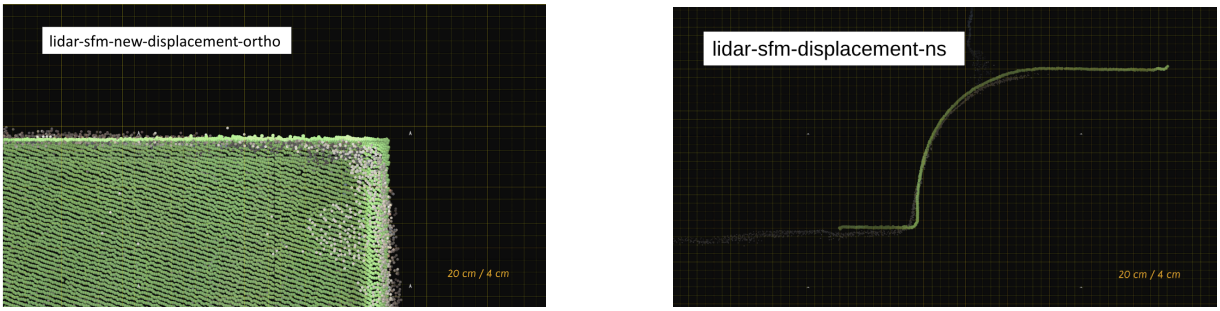
precision, can be used to georeference the UAV and LIDAR surveys, (2) the accelerometer data can be used to identify any significant deviations from the building's normal frequency response, and (3) the high-rate, real-time rooftop data stream can be used to monitor any significant changes or distortions to the building's position during or after a seismic event or other natural hazard. There have been no considerable seismic events in the vicinity since data collection began in early 2017, thus the the latter utility will not be considered here. Instead, we describe the role of geodetic instrumentation in cybermodel georeferencing (Section 5.2) and evaluate the potential of the accelerometer instrumentation to analyze the structure's frequency response (Section 5.3).

5.2 GPS Georeferencing of Digital Structural Surrogate

The first priority of the multisensor structural monitoring is repeat surveys, conducted each year and following any natural hazard, to create a digital surrogate of the external building structure. Repeat surveys allow systematic comparison of a pre- and post-event structure, creating a reliable, data-driven approach to damage identification, without relying solely on visual inspection. The three rooftop GPS stations and additional survey points on the rooftop and library base are used to georeference the digital surrogate, and align the UAV and LIDAR surveys. An initial baseline UAV and LIDAR survey of the Geisel Library was conducted on July 28, 2017 (Figure 5.3), with a second repeat survey one year later on July 23, 2018. Here, we describe the role of the GPS instrumentation within these repeat surveys.

5.2.1 GPS Data Processing

The GPS observations are processed in real-time with the RTD Pro software, wherein the positions of the three rooftop stations are estimated relative to the base station, LBRF. The underlying assumption of relative positioning is that the base station is fixed in space, and other stations in the network move in relation to it. This method is in contrast to precise point positioning (PPP), wherein the GPS location is solved relative to an external, globally defined reference frame. Generally, PPP is the preferred processing method when considering ground motions from moderate-to-large earthquakes, because unless the reference station is sufficiently far away, it will also be subject to seismic motions, and will prevent accurate evaluation of the ground motions experienced at the remaining stations in the network. In the case of building monitoring, however, relative positioning is an appropriate method because we are concerned not with the seismic ground motions themselves, but with the motion of the structure relative to the ground. In other words, because the reference site is in close proximity to the structure, we expect that the ground motions at the reference site and at the building's base will be similar. Thus, the



(a) (b)

Figure 5.4: Point cloud alignment of LIDAR (green) and UAV (grey) surveys. (a) Top down cross-sectional view, (b) Side cross-sectional view, south to north. Images courtesy of Eric Lo, Qualcomm Institute.

results of this processing method will show the displacement at the building roof relative to the ground level, such that any change in position of the building relative to the reference station station following a seismic event or other natural hazard is indicative of distortion in the building structure.

5.2.2 Precise GPS Location for Imagery Alignment

The initial baseline UAV, LIDAR, and ground survey was completed on July 28, 2017. A second survey, conducted on July 23, 2018 was used to confirm repeatability. The 2018 survey included three important improvements: (1) the resolution of the camera aboard the UAV was increased to allow clear identification of the GPS radomes in the resulting images, (2) temporary targets were affixed to the GPS radomes to assist with UAV survey alignment to the precisely known GPS positions (Figure 5.2c), and (3) the UAV was equipped with a high-rate (1 Hz), precision GPS, such that the UAV flight path is known and can be used to facilitate alignment of the imagery.

Figure 5.4 shows the alignment of UAV (grey) and LIDAR (green) point clouds for the 2018 survey, completed by collaborators at the Qualcomm Institute. Once aligned with one another, these two surveys are then tied into the reference frame defined by the continuous rooftop

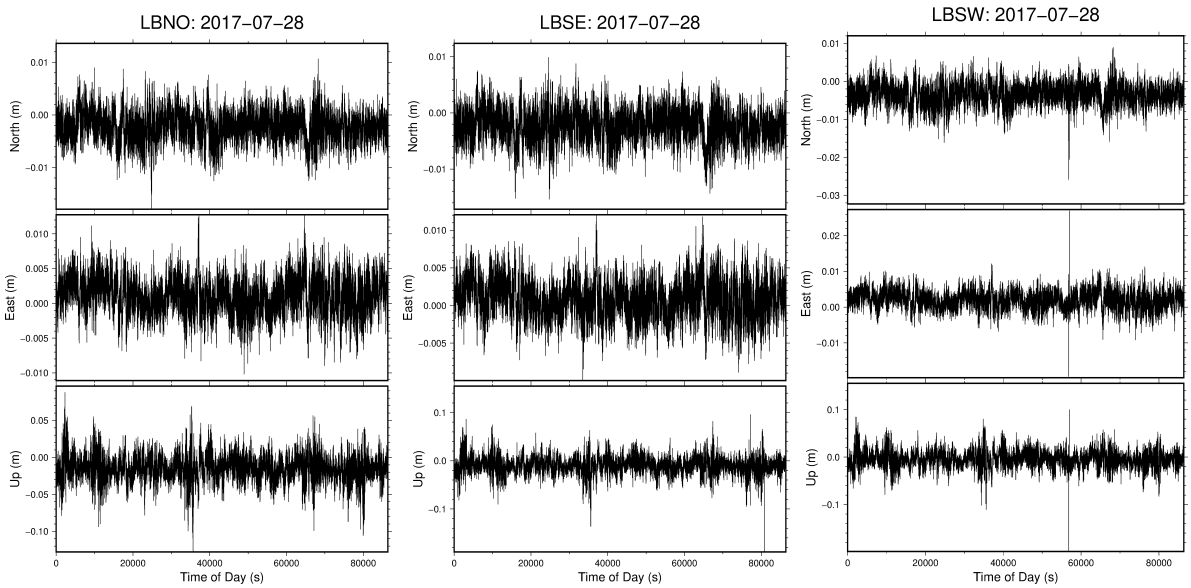


Figure 5.5: Geisel Library rooftop GPS observations (a) Library North (LBNO), (b) Library Southeast (LBSE), and (c) Library Southwest (LBSW) for July 28, 2017, the date of the original baseline UAV and LIDAR survey.

Table 5.1: Average daily position of Geisel Library GPS Stations, July 28, 2017.

Station	Longitude (°E)	Latitude(°N)	Elevation (m)
LBRF	-117.230295890	32.886120270	56.4548
LBSE	-117.237508504	32.881122059	115.1448
LBSW	-117.237648061	32.881121104	115.1455
LBNO	-117.237545742	32.881212088	115.1495

GPS locations and ground survey points. Constraining certain markers in the imagery results in misfit elsewhere in the image, relative to known survey points. For example, in the recent 2018 survey, optimal alignment to the reference frame defined by the ground survey markers at the base of the structure resulted in a $\sim$4 cm horizontal offset between the precisely known continuous rooftop GPS locations, and their corresponding locations derived from the aligned imagery. This misfit exceeds the typical noise in GPS location estimate, implying that the alignment must still be improved. Additional work is now required to account for this offset, improve the mathematical transformations for imagery alignment, and optimize the complementary use of sensors.

The 1 Hz GPS positions at the rooftop stations for the full day of the initial baseline

Table 5.2: Average daily position of Geisel Library GPS Stations, July 23, 2018.

Station	Longitude (°E)	Latitude(°N)	Elevation (m)
LBRF	-117.230295890	32.886120270	56.4548
LBSE	-117.237508488	32.881122072	115.1319
LBSW	-117.237648046	32.881121120	115.1334
LBNO	-117.237545732	32.881212102	115.1376

survey are shown in Figure 5.5. Average daily positions of the three rooftop GPS stations during the 2017 baseline survey and 2018 repeat survey are given in Tables 5.1 and 5.2, respectively. There have been no significant seismic events between July 2017 and July 2018, thus a successful repeat study should show no considerable change with respect to the baseline survey. However, the 2018 repeat survey showed minor changes to the building position relative to the 2017 survey. On average, the three rooftop stations moved 1.587 mm north, 1.277 mm east, and 1.23 cm down, relative to their July 2017 positions.

Horizontal GPS-derived displacements are precise to less than 1 cm in the horizontal, and sustained trends can be discerned on smaller scales, with a sufficiently long time series. Consider the small northeastward motion implied by the difference between the average positions of the GPS stations on the two survey dates one year apart. When placed in the more complete time series (Figure 5.6), it becomes clear that this observed change in location is within natural variability of the healthy structure. The average daily GPS-derived positions at the three rooftop stations are shown over a period of 13 months, beginning July 1, 2017 and ending July 31, 2018. The dates of the two surveys, July 28, 2017 and July 23, 2018, are highlighted in red. In the context of the longer time series, the variation between the two surveys clearly falls within the typical scatter. Long period trends in the time series appear to show yearly, seasonal variations, however such speculation cannot be confirmed until our time series is extended through additional annual cycles. The results of the GPS positions one year apart emphasize the importance of consideration of the continuous GPS record to assess changes to the building locations, rather than relying only on discrete, single-day sampling. Conducting once yearly surveys, without regard

for the longer time series, may provide a misleading view of variation in the building position. Discrete sampling during approximately the same time each year means that seasonal variations may not be well understood, but seasonal trends are undoubtedly an important consideration should a natural hazard occur at a different time of year, prompting a resurvey. Longer, continuous time series make clear the typical day-to-day and month-to-month variation in the GPS positions. In this continuous context, the repeat surveys become more useful for understanding structural changes and long-term health.

5.2.3 Goals of Subsequent Repeat Survey

The repeat surveys described here are the first in a long-term monitoring project. The eventual goal will be to demonstrate the ability to operate in an automated fashion following a natural hazard, where accessibility to target structures may be limited. Reducing the need for operator interaction by automating the UAV flight path, for example, is a necessary step to prove feasibility of a repeat survey within rapid response procedures. In order to meet this goal, we suggest that subsequent UAV surveys include higher-rate GPS positioning or an inertial sensor to improve the spatial and temporal resolution of the recorded UAV flight path. The continuation of this work will include development of permanent targets for the rooftop GPS radomes to aid imagery alignment, and strategies to deploy additional targets at ground survey points within a disaster scenario. Finally, the optimization of survey alignment between sensors, and an analysis of the timeliness of each survey component is required to advance eventual inclusion in rapid response procedures.

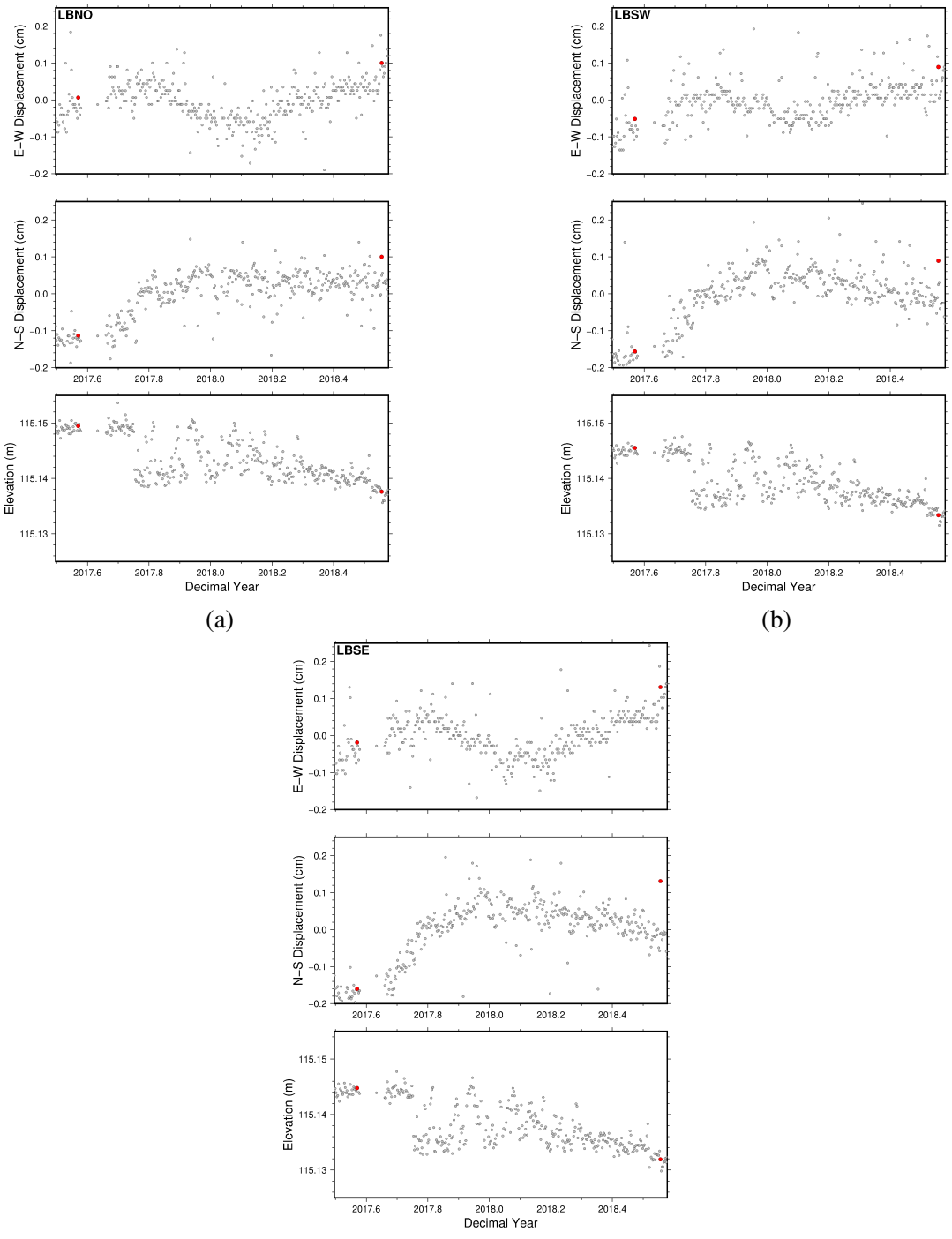


Figure 5.6: Average daily GPS-derived positions at rooftop stations (a) LBNO, (b) LBSW, and (c) LBSE over a 13 month (July 2017 - July 2018) period of observation. Survey dates (July 28, 2017 and July 23, 2018) are shown in red.

5.3 Geisel Library Frequency Response

In addition to monitoring major changes to the Geisel Library by aligning to the precise location of the rooftop instruments, the rooftop instrumentation has utility for monitoring structural response during periods of quiescence, which can in turn provide metrics for a more comprehensive understanding of damage following an extreme event. That is, by characterizing the healthy structure, we become better able to monitor for any future changes that might be indicative of structural instability (Trifunac et al., 2001; Prieto et al., 2010; Kohler et al., 2013, 2016; Massari et al., 2017). This can be accomplished by monitoring the fundamental period of the structure. The fundamental period (also called the natural period, fundamental frequency, or natural frequency) is the longest-period mode of vibration. In the most simplistic estimation, the fundamental period of a building (in seconds) is roughly equal to one tenth the number of stories (i.e. a 10-story building has a natural period of $\sim$1 s, a 20-story building has a natural period of $\sim$2 s, etc.). The frequency response of a structure in reality is more complex, depending on the building design and construction. The natural period has been shown to change with time due to various factors such as building aging, soil quality changes, and meteorological events (e.g. Clinton et al., 2006). The Geisel Library has a unique design with a significantly smaller ground-level footprint than that of the upper stories, just one of the many additional factors that may affect its frequency response.

The 1 Hz GPS (Nyquist frequency of 0.5 Hz) is not capable of investigating the frequencies of interest. Furthermore, due to the relatively high noise floor of the GPS compared to traditional seismic instrumentation, we do not expect the GPS to be sensitive to the minor, everyday motions of the structure. As a result, we pursue characterization of the Geisel Library frequency response with the seismic dataset alone. We compute a three-component sum of squares for each rooftop accelerometer station, and apply an adaptive sine multitaper power spectral density (PSD) estimation using the Python wrapper MtSpec (Riedel and Sidorenko, 1995; Prieto et al., 2009;

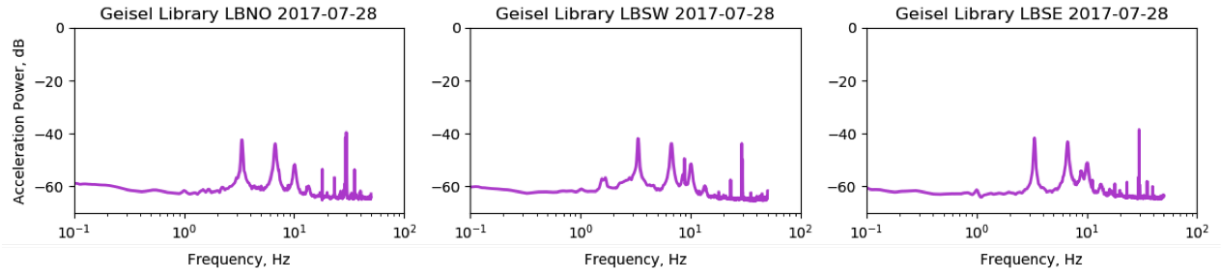


Figure 5.7: Example power spectral density of Geisel Library rooftop accelerometers during a period of quiescence on July 28, 2017. Left: Geisel roof station LBNO, Center: Geisel roof station LBSW, Right: Geisel roof station LBSE.

Krischer, 2016). From Figure 5.7, we see that the longest period (or lowest frequency) mode in the acceleration PSD is at a frequency greater than 3 Hz.

This frequency response observed in the SIO MEMS accelerometer is in contrast with typical analytical approximations of natural period. As an example, Patel et al. (2011) estimates the natural period of vibration, T, for a reinforced concrete structure of height, h, in meters, as

$$T \approx 0.075 h^{0.75} \tag{5.1}$$

which would estimate the natural period of the Geisel structure to be approximately 1 Hz, however the accelerometer data shows little power at that frequency. The same study determined a second equation that considers the dimension of the building base:

$$T \approx \frac{0.09h}{\sqrt{d}} \tag{5.2}$$

where d is the length of the building in meters at the base level. The Geisel Library's smaller diameter base is undoubtedly an important factor in its frequency response, however Equation 5.2 also underestimates the natural frequency observed by the MEMS accelerometers. While it is possible that the unique design of the Geisel library may be too complex for such simple rules of thumb, the unexpectedly high value of the observed natural period suggests the possibility that

the instrument is not precise enough to properly observe the structural response, and what we are observing is instead some product of the SIO MEMS or GAP instrumentation itself. In the following sections, we will evaluate the accelerometer observations in the frequency domain over a prolonged period of quiescence, to elucidate the origin of the observed response, and determine future directions for structural monitoring with this low-cost instrumentation.

5.3.1 Observed Frequency Response, Geisel Library Rooftop

The frequency response of a building is expected to change over time based on a variety of factors including meteorological impacts, seismological impacts, and building aging. The dry conditions and lack of local earthquakes during our period of study render it impossible to draw any conclusions about the impact of rainfall or seismicity on the Geisel Library frequency response. Instead, we concentrate on the effect of temperature and wind.

Figures 5.8, 5.9, and 5.10 show spectrograms of the Geisel Library during the 2-3PM hour (local time) for the first 160 days of 2018 at rooftop stations LBSE, LBSW, and LBNO, respectively. The color bar denotes the power at each frequency, where warmer colors indicate greater power. The frequency response over this study period is fairly consistent. There are three major frequency bands of higher power: $\sim$3.2 Hz, $\sim$6.5 Hz, and $\sim$30 Hz. Station LBSW appears to observe an additional, weaker peak between 9-10 Hz. Station LBSE appears to have the lowest noise floor (power remains in the cool colors at low frequency), which may be due to slight instrumentation differences, or other factors on the roof itself, including proximity to any machinery, heating ventilation and air conditioning system, etc. The observed natural period of the building fluctuates slightly from day to day, rising and falling in response to some combination of external factors. Figures 5.23, 5.21, and 5.22 show the PSD for each day in the 160-day study period (grey lines) and the average PSD over the 160-day period (black line) at rooftop stations LBSE, LBSW, and LBNO, respectively.

Figures 5.11 and 5.12 describe the relationship of fundamental frequency during the

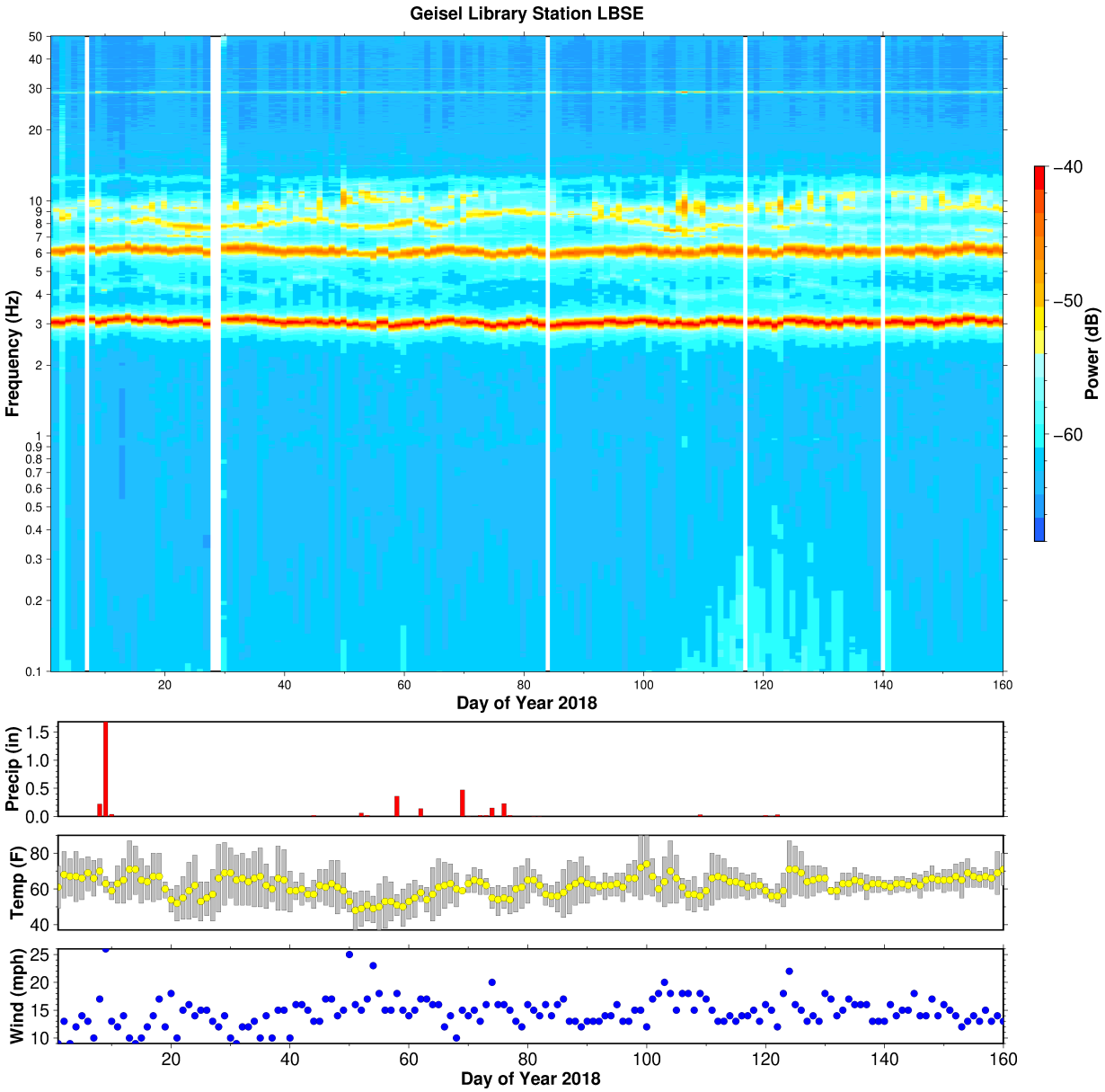


Figure 5.8: Top: Spectrogram at Geisel library rooftop station LBSE. Plot shows days 001-160 in 2018 at frequencies between 0.1 and 50 Hz. White areas denote gaps in data acquisition. The lowest persistent peak appears at ∼3.2Hz, with a second peak at twice the frequency of the first. Bottom: Weather metrics from the period of study: daily precipitation totals (top), mean daily temperature (yellow circles) with grey bar extending to daily low and high temperature (middle) and average daily wind speed (bottom).

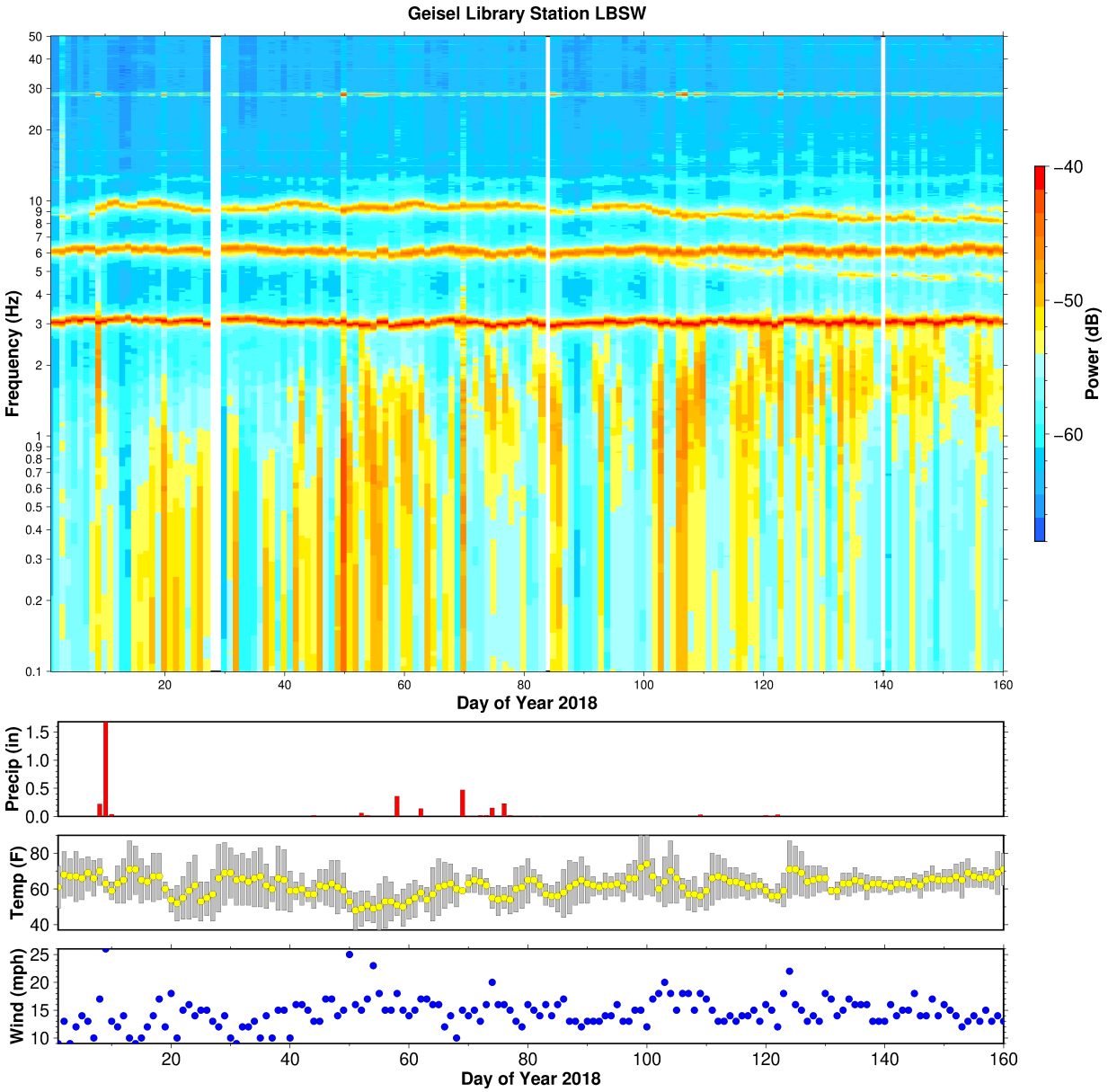


Figure 5.9: Top: Spectrogram at Geisel library rooftop station LBSW. Plot shows days 001-160 in 2018 at frequencies between 0.1 and 50 Hz. White areas denote gaps in data acquisition. The lowest persistent peak appears at ∼3.2Hz, with a second peak at twice the frequency of the first. Bottom: Weather metrics from the period of study: daily precipitation totals (top), mean daily temperature (yellow circles) with grey bar extending to daily low and high temperature (middle) and average daily wind speed (bottom).

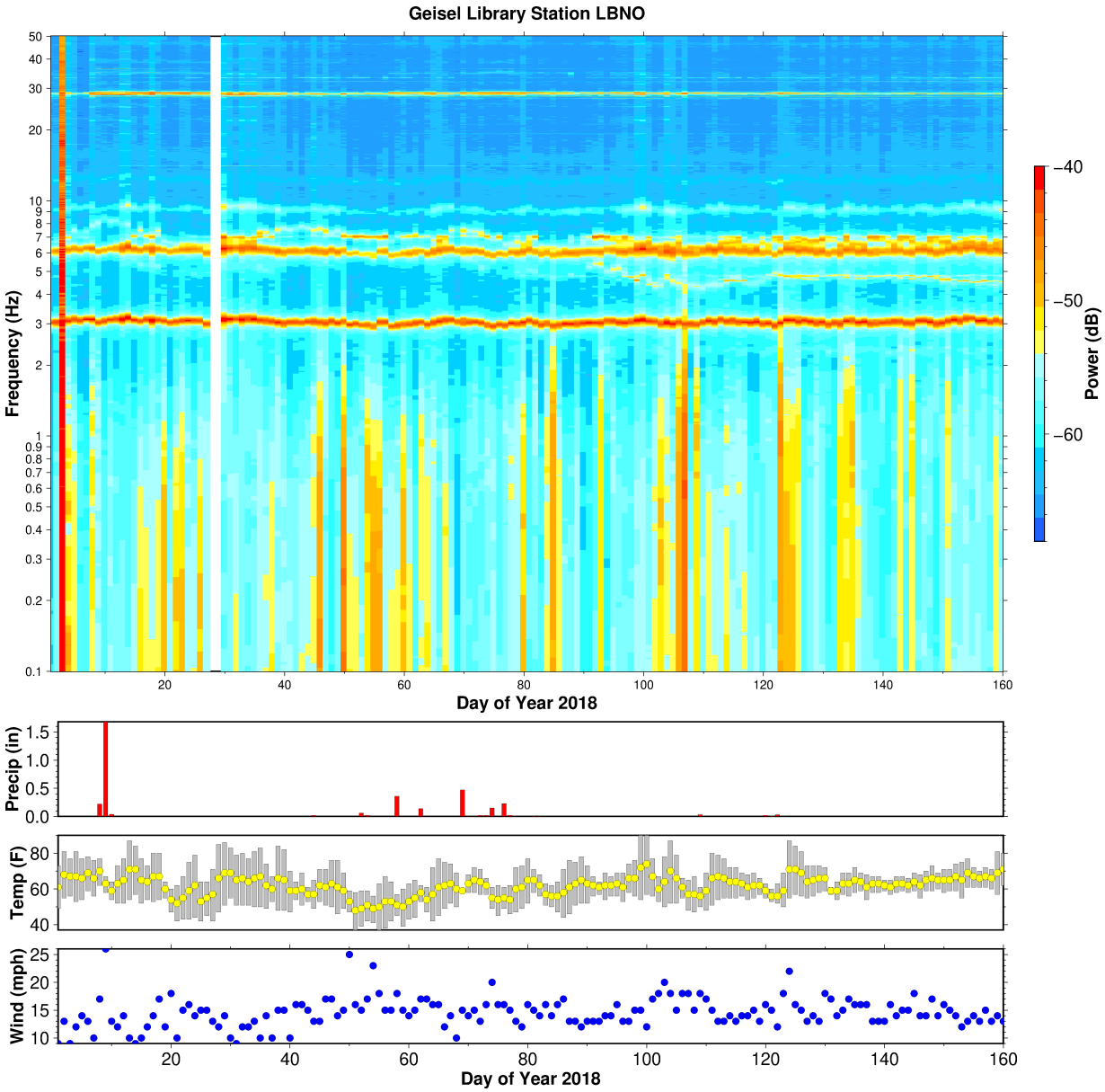


Figure 5.10: Top: Spectrogram at Geisel library rooftop station LBNO. Plot shows days 001-160 in 2018 at frequencies between 0.1 and 50 Hz. White areas denote gaps in data acquisition. The lowest frequency peak appears at ∼3.2Hz, with a second peak at twice the frequency of the first. Bottom: Weather metrics from the period of study: daily precipitation totals (top), mean daily temperature (yellow circles) with grey bar extending to daily low and high temperature (middle) and average daily wind speed (bottom).

2-3PM hour (local time) with maximum sustained wind speed and average daily temperature, respectively. The natural frequency of a structure is expected to fall during periods of heavy winds due to the increase of forced vibrations. Conversely, natural frequency is expected to rise with an increase in temperature because thermal expansion of the concrete structure creates a stiffening effect (Clinton et al., 2006). Our observations from the SIO MEMS appear to confirm these relationships. There is a weak negative correlation (correlation coefficient, $R \approx$-0.3) between observed natural frequency and maximum sustained winds (Figure 5.11), and a stronger positive correlation ($R \approx$0.6) between natural frequency and average daily temperature (Figure 5.12).

In order to better address the relationship between temperature and natural frequency, we calculate hourly PSDs to see smaller temporal-scale changes. For brevity, we focus the analysis on rooftop station LBNO, and consider a shorter duration within the month of February 2018 (Figure 5.13). Note that the vertical axis of Figure 5.13 extends from 1 to 10 Hz, which is different from previous figures. We compare these fluctuations in natural frequency to the hourly dry bulb temperature recorded at a National Oceanic and Atmospheric Administration (NOAA) weather station in Miramar, San Diego, roughly 9 km from the Geisel Library (Figure 5.14). The rise and fall of the temperature and the natural frequency appear to occur with the same period. Temperature and natural frequency are expected to be positively correlated– warming causes thermal expansion of the building, which stiffens the structure and raises the natural frequency– such that natural frequency should follow temperature with some lag. We find that the alignment between the two time series occurs with a shift in natural frequency by -5.5 hours (Figure 5.15). In other words, while temperature peaks just after midday, natural frequency peaks more than 5 hours later, during the evening. While it is not possible to prove causation, this simple analysis suggests it is possible that changing temperature may have a direct effect on the observed natural period. However, because both signals have a period of a day, it is also possible that some unknown daily signal is responsible for the periodic nature of the observed natural frequency.

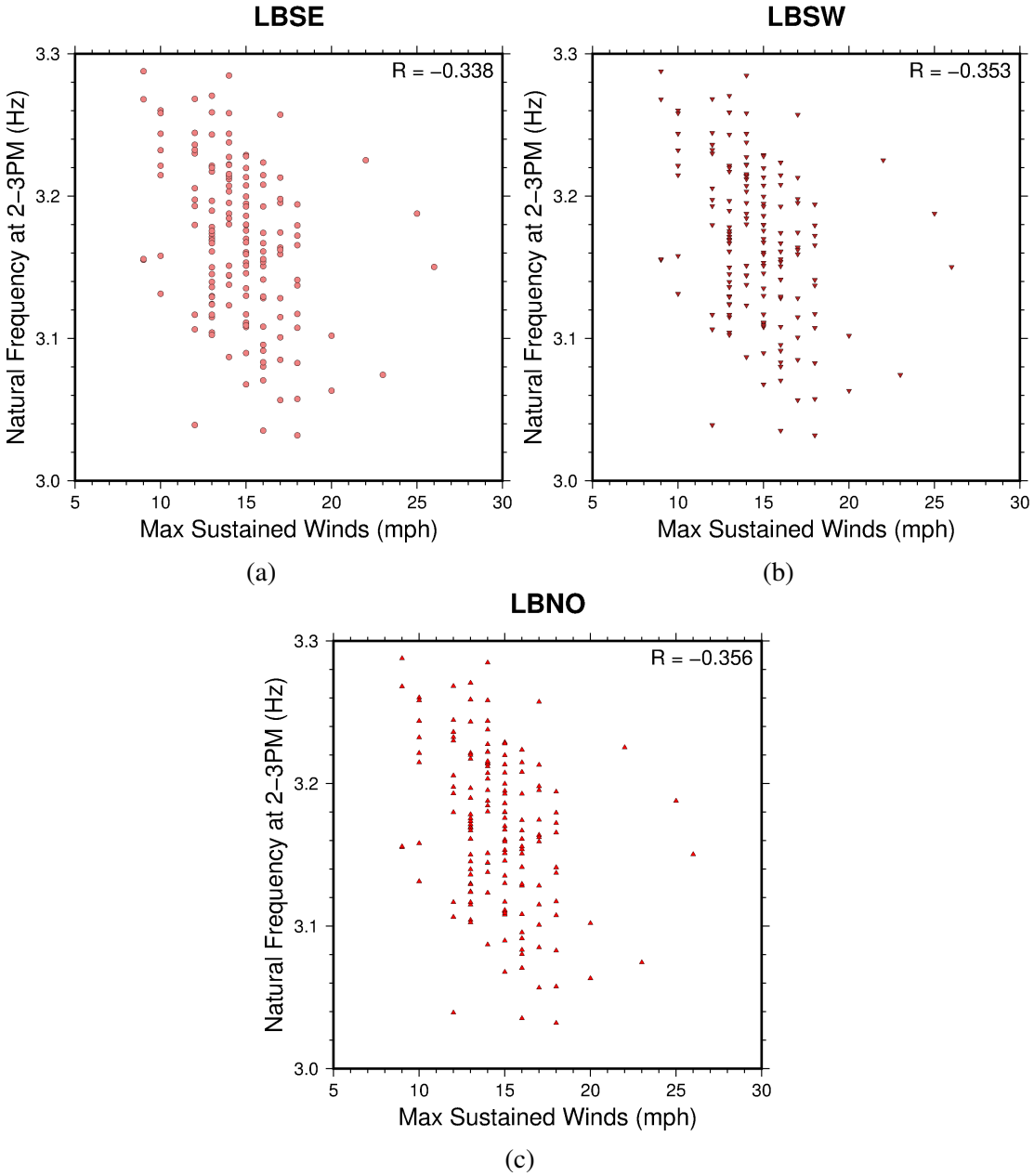


Figure 5.11: Correlation between natural frequency and maximum sustained wind speed at Geisel Library rooftop.

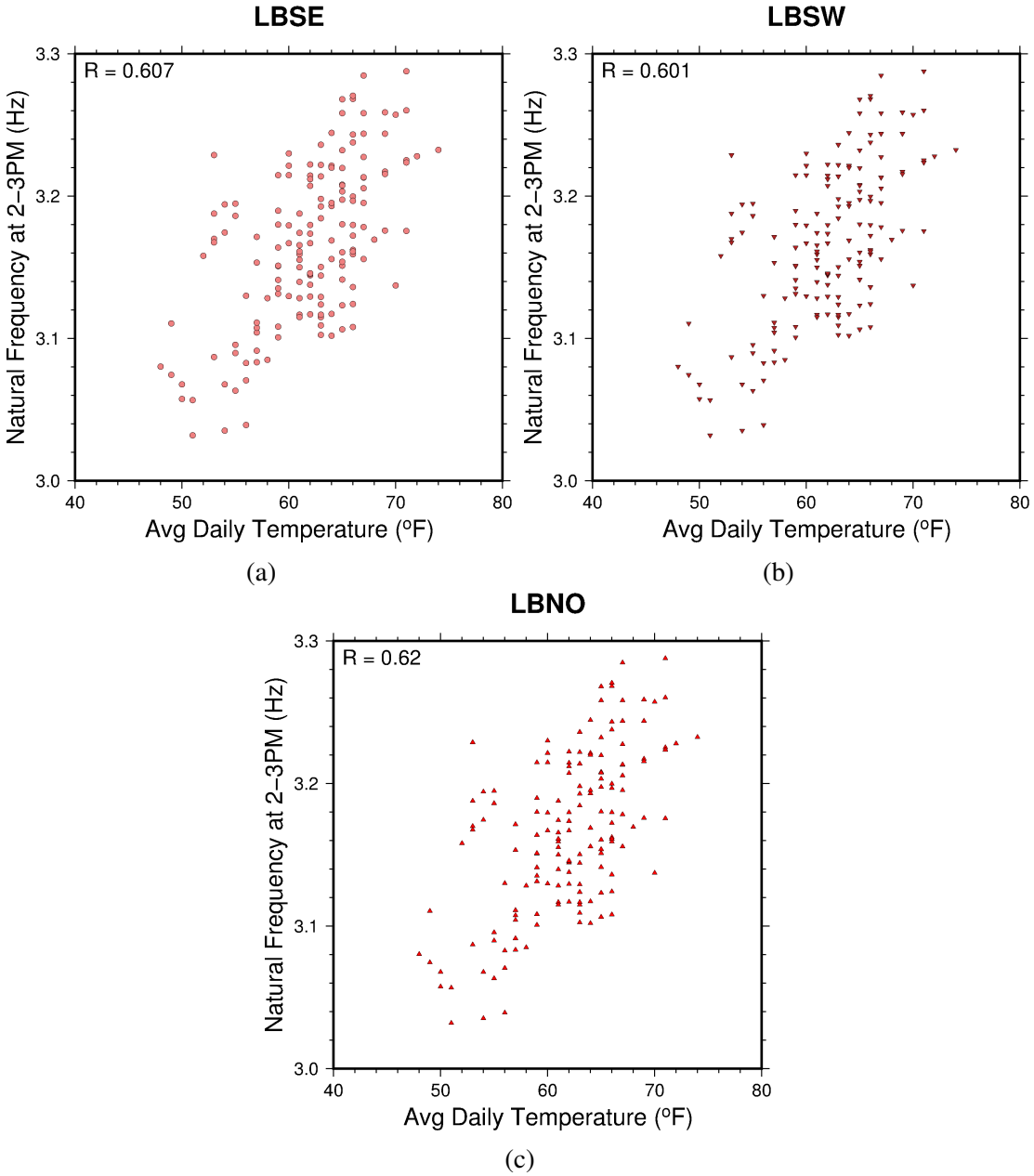


Figure 5.12: Correlation between natural frequency and average daily temperature at Geisel Library rooftop.

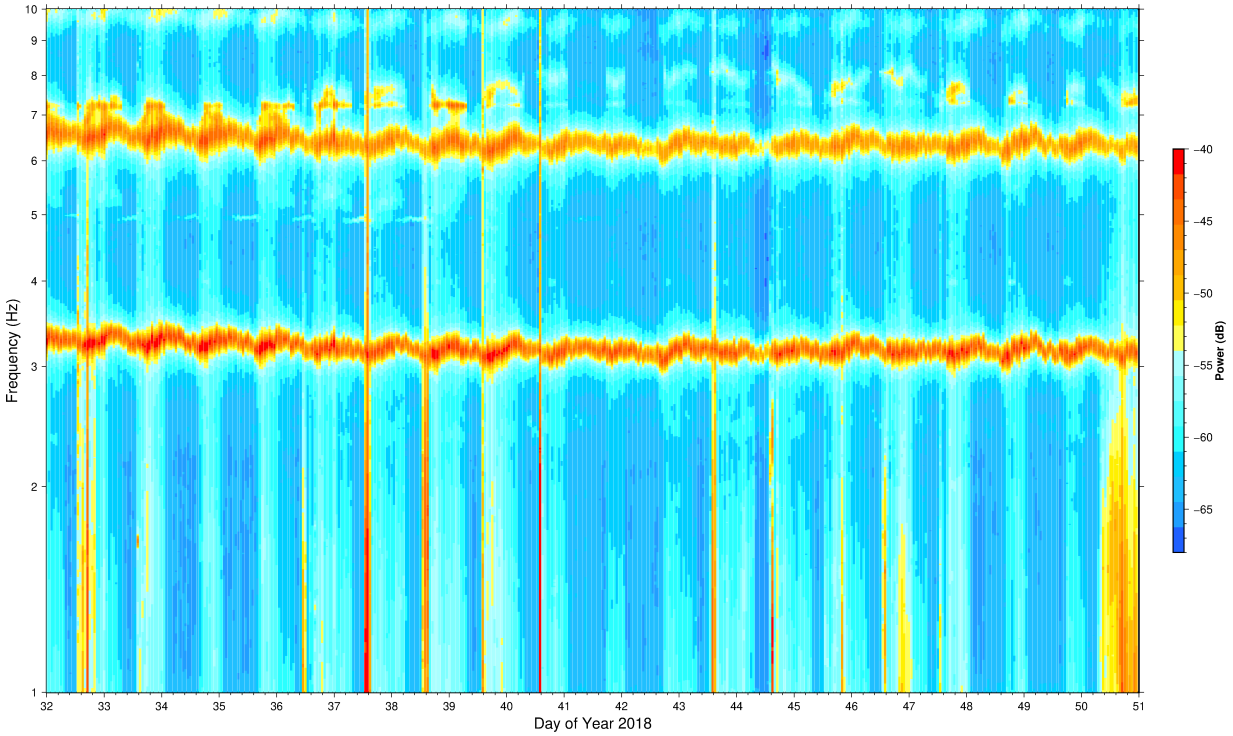


Figure 5.13: Hourly spectrogram at Geisel Library rooftop station LBNO for days 032-050 in 2018. Note that we show frequencies between 1 and 10 Hz (vertical axis), which is different from previous figures.

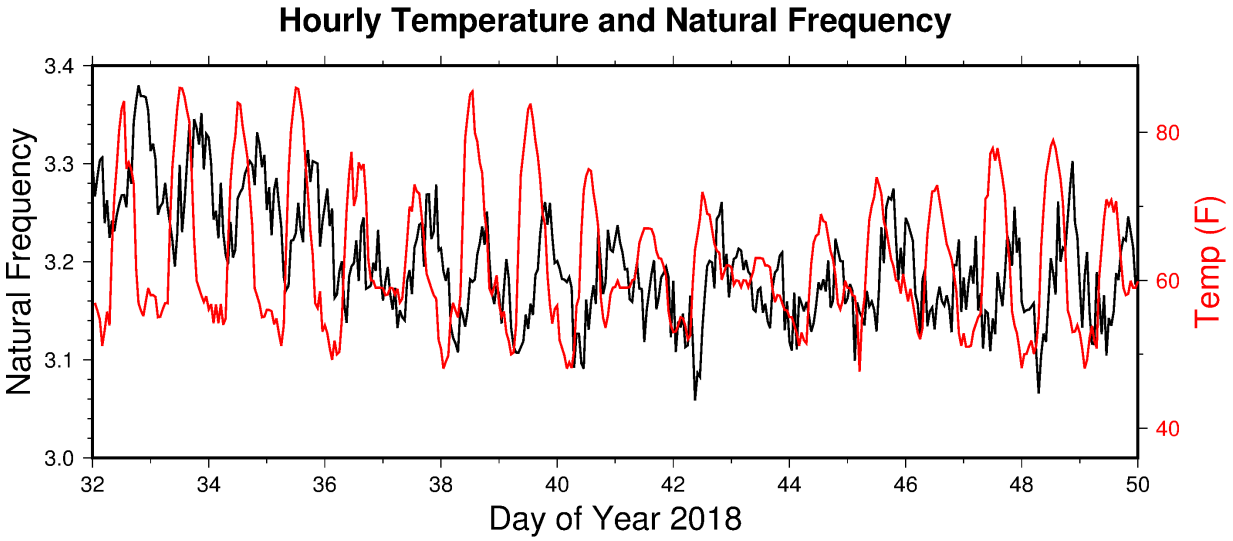


Figure 5.14: Hourly natural frequency (black) at Geisel Library rooftop station LBNO, and hourly dry bulb temperature (red) at NOAA weather station in Miramar.

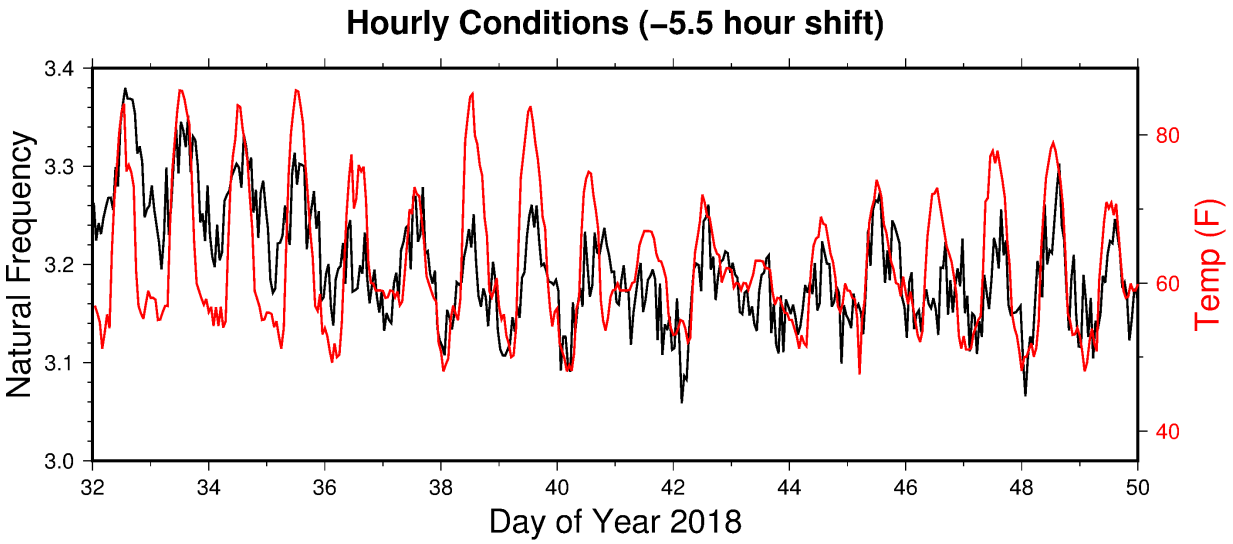


Figure 5.15: Hourly natural frequency (black) at Geisel Library rooftop station LBNO, and hourly dry bulb temperature (red) at NOAA weather station in Miramar. The natural frequency time series is shifted by -5.5 hours, indicating that the peak frequency follows temperature with a 5.5 hour lag.

5.3.2 Observed Frequency Response, Geisel Library Basement

In comparison to the rooftop stations, a similar spectrogram is observed for station LBBA, an SIO MEMS accelerometer located in the basement of the Geisel Library (Figure 5.16). Note that Figure 5.16 shows a consistent peak at 1 Hz (and a harmonic at 2 Hz), however this is likely due to filling of 1-second data gaps to facilitate the calculation of the PSD and therefore should not be interpreted as a true signal. Data are telemetered in 1-second packets (100 samples per second), thus missing data always comes in 1-second increments. Figure 5.24 shows the PSD for each day in the 160-day study period (grey lines) and the average daily PSD (black line) at basement station LBBA. This basement-level station is not expected to see the same building response as rooftop stations since it should be unaffected by the natural sway of the upper floors. That the response at station LBBA is so similar to the rooftop stations introduces some skepticism, and suggests the possibility that the observed frequency response is not in fact a representation of the building motion, but of some other factor, perhaps inherent to the instrument itself.

Again similar to its rooftop counterparts, the natural period observed at accelerometer

LBBA during the 2-3PM local hour is positively correlated ($R \approx 0.6$) with temperature (Figure 5.17b). That station LBBA is located in the temperature-controlled basement appears to discount the possibility that that we are observing a temperature effect on the instrument, rather than on the building, supporting the theory that we are properly imaging the building's frequency response. However, complicating this assessment is the fact that the basement station is not entirely unaffected by the outside world. Station LBBA is connected to a geodetic module on the library roof (the same as rooftop station LBNO), where the accelerometer signal is time tagged to the rooftop GPS instrument. The observed frequency content could therefore be affected by any instrumental issues resulting from the geodetic module as well.

5.3.3 Observed Frequency Response, Ground-Based SIO MEMS

To further investigate the source of our observations of natural frequency, we consider the SIO MEMS affixed to the vertical leg of the deep drilled braced GPS monument at station LBRF. This instrument is identical to those installed on the Geisel Library, but of course will have no contribution from building response. Figure 5.18 shows the spectrogram for station LBRF and Figure 5.25 shows the PSD for each day in the 160-day study period (grey lines) and the average daily PSD (black line). The observed power of the reference station is generally below -60 dB at all frequencies, with a first spectral peak around 3.7 Hz. The rest of the spectrogram is described by harmonics of this first peak. While it is unclear what is causing the $\sim$3.7 Hz peak, the fluctuations from day to day also appear to be correlated with temperature (R$\approx$0.7, Figure 5.19a), similar to that of the Geisel Library rooftop and basement stations. The frequency fluctuations are not correlated to wind (Figure 5.19b).

It seems likely then, that this perceived temperature-related pattern is a feature not of the Geisel Library, but rather of the instrument being used to observe the library: the SIO GAP. Luckily, a newer, updated generation of instruments has been manufactured and is ready for installation. Testing of this upgraded instrumentation was conducted in October 2017 at nearby

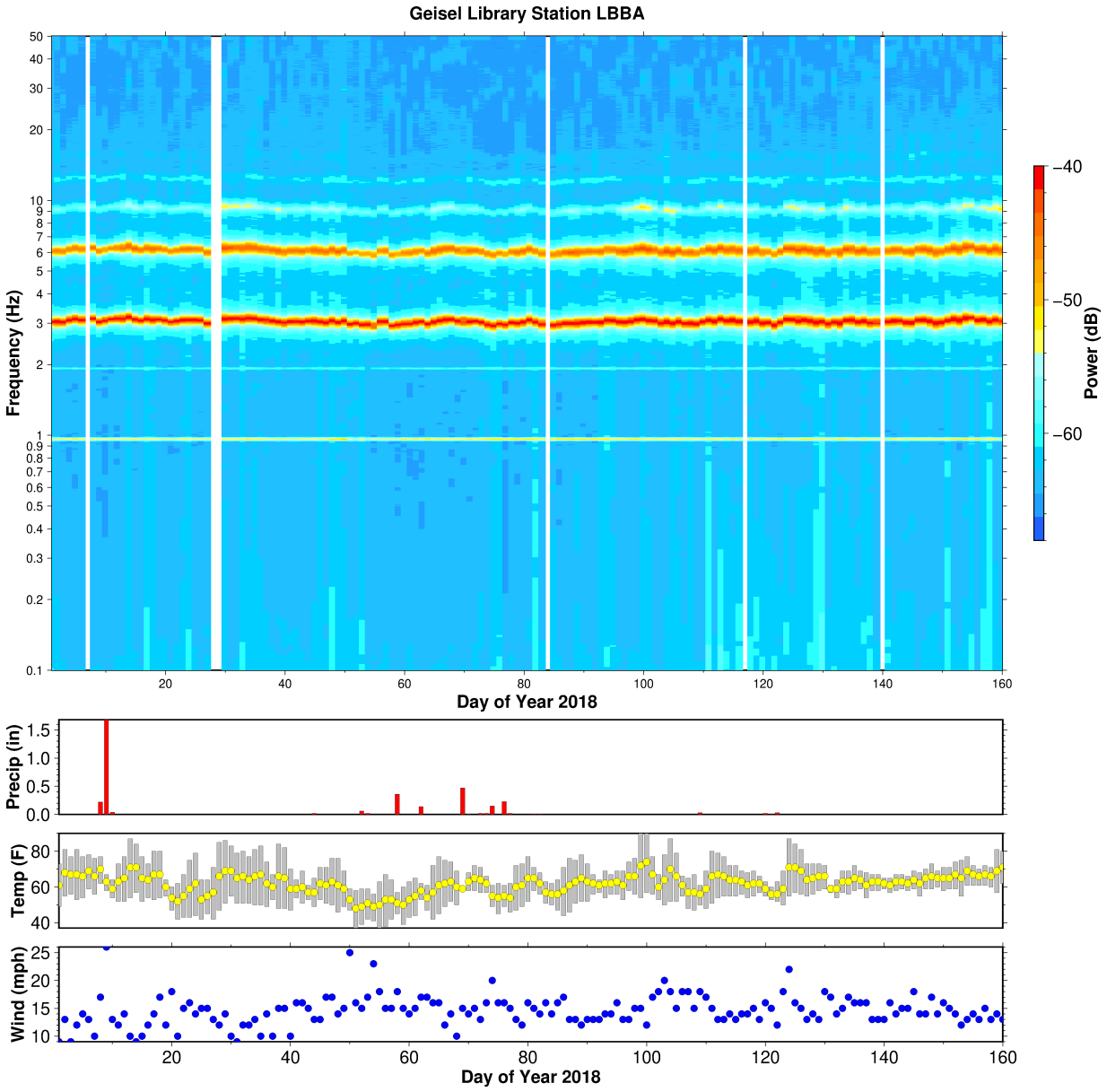


Figure 5.16: Top: Spectrogram at Geisel Library basement station LBBA. Plot shows days 001-160 in 2018 at frequencies between 0.1 and 50 Hz. White areas denote gaps in data acquisition. The lowest persistent peak appears at ∼3.2Hz, with a second peak at twice the frequency of the first. A consistent signal at 1Hz (and harmonic at 2Hz) is most likely due to filling of 1-s data gaps to allow PSD calculation, and is not interpreted as a true signal. Bottom: Weather metrics from the period of study: daily precipitation totals (top), mean daily temperature (yellow circles) with grey bar extending to daily low and high temperature (middle) and average daily wind speed (bottom).

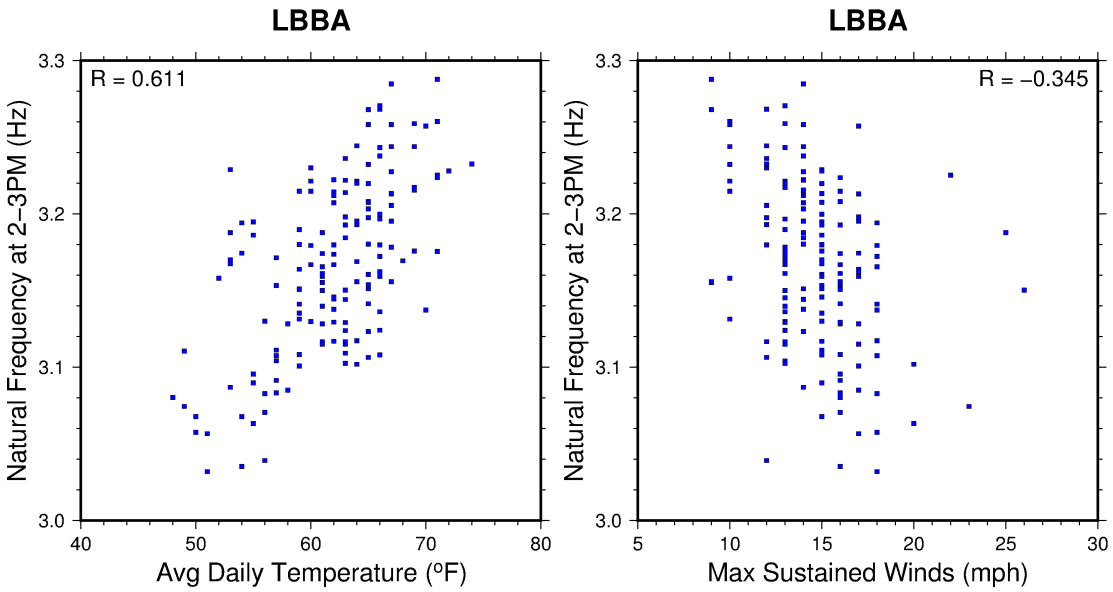


Figure 5.17: Correlation between natural frequency and weather metrics, basement station LBBA, (a) Correlation between natural frequency and average daily temperature. (b) Correlation between natural frequency and maximum sustained winds.

deep drilled braced GPS monument SIO5, located on Mt. Soledad in La Jolla, California. We calculate the PSD for each hour during the testing period. The resulting spectrogram is shown in Figure 5.20. Note that the color scale has changed, demonstrating the improved sensitivity of the newer instrumentation. Unlike station LBRF, there are no discernible peaks in this spectrogram. Instead, the observed power simply decreases with increasing frequency, and is lower in magnitude in comparison to the older generation instruments. We therefore suggest that the newer generation of instrument is more suitable for the task of monitoring of a structure's natural period and recommend installation of these updated sensors at the Geisel Library.

5.4 Discussion and Future Directions

Using a suite of instruments optimizes monitoring of structural health over time. LIDAR and UAV imagery show changes to the outside structure, as well as any changes in the overall

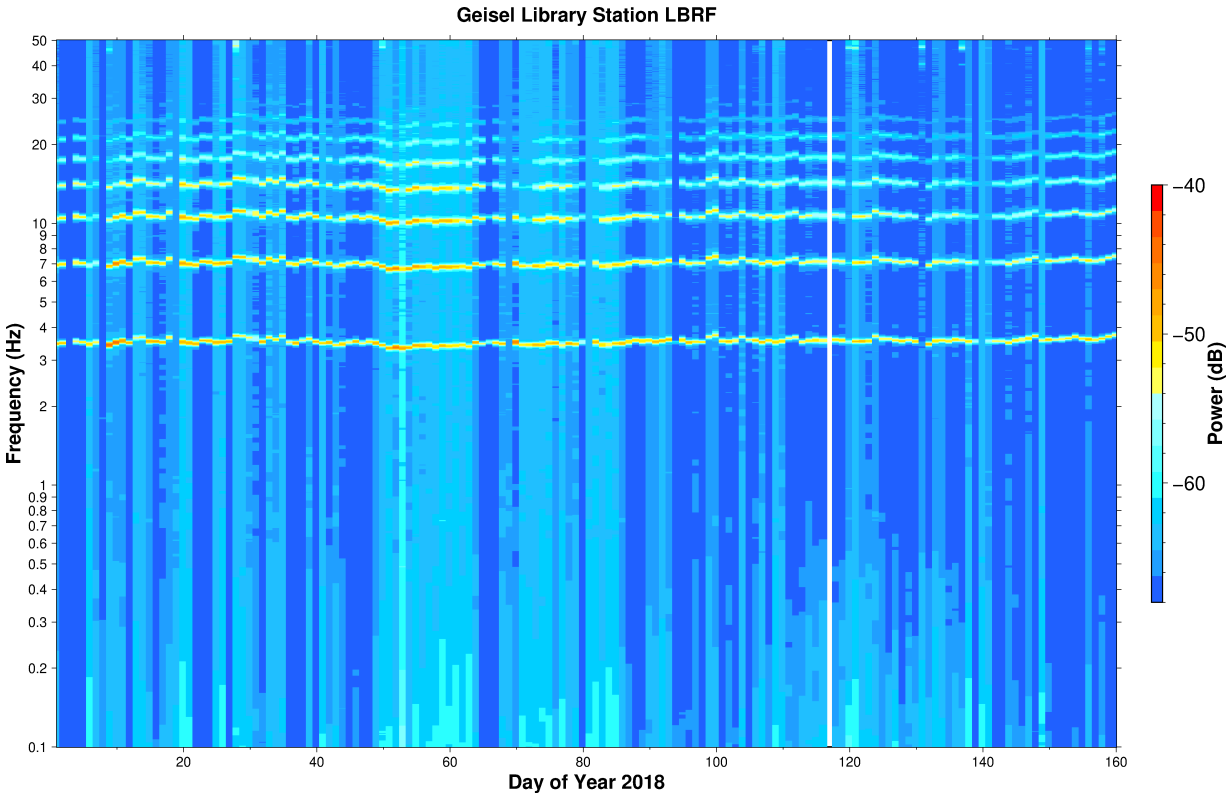
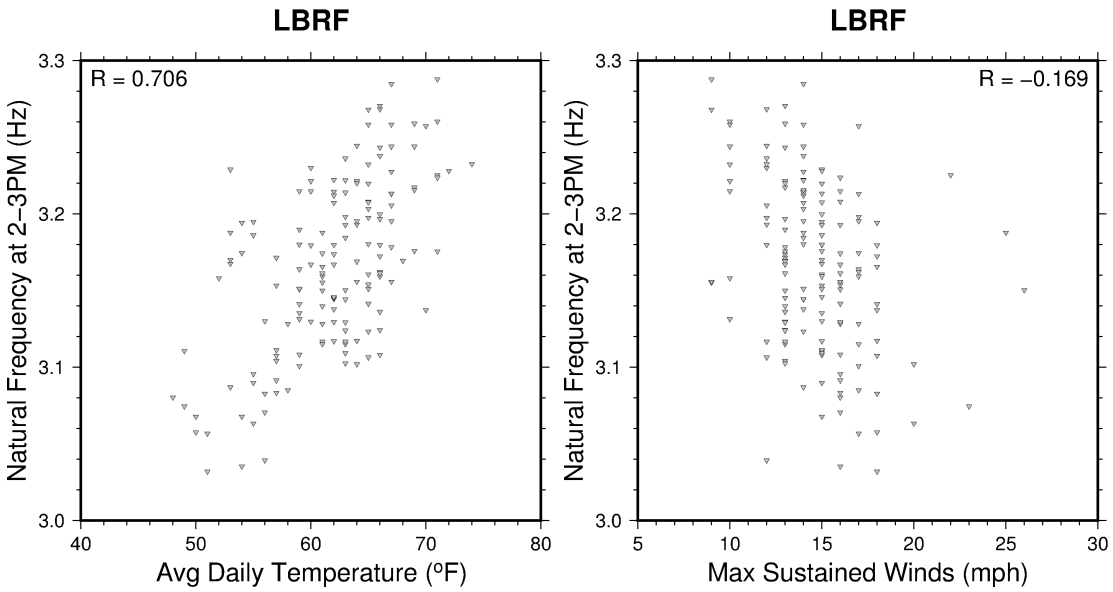


Figure 5.18: Spectrogram at Geisel library reference station LBRF. Plot shows days 001-160 in 2018 at frequencies between 0.1 and 50 Hz. White areas denote gaps in data acquisition. The lowest persistent frequency peak appears at ∼3.7Hz, with a second peak at twice the frequency of the first.



(a) (b)

Figure 5.19: Correlation between natural frequency and weather metrics, reference station LBRF, (a) Correlation between natural frequency and average daily temperature. (b) Correlation between natural frequency and maximum sustained winds.

shape of the building that could be indicative of damage. Alignment of that imagery to the precisely known locations of the rooftop GPS instrumentation and other survey points provides a reliable, repeatable method to track any significant motions of the building. We have demonstrated repeatability via two surveys roughly a year apart. While small differences in the average daily position of the building were observed between the two survey dates, continuous GPS monitoring over the course of the intervening year suggest that this is within typical variation, and likely due to normal variations rather than any tectonic or structural changes.

Despite discovery of an instrument deficiency which prevents accurate quantification of the building's natural period, our results suggest that the newest generation of SIO MEMS, which has a lower noise floor, and is devoid of a strong harmonic signal observed in the older generation instrument, will be suitable for structural monitoring applications. As a direct result of this investigation, the four SIO MEMS accelerometers located at Geisel Library, and the off-site reference station, will be replaced with the newer generation instruments.

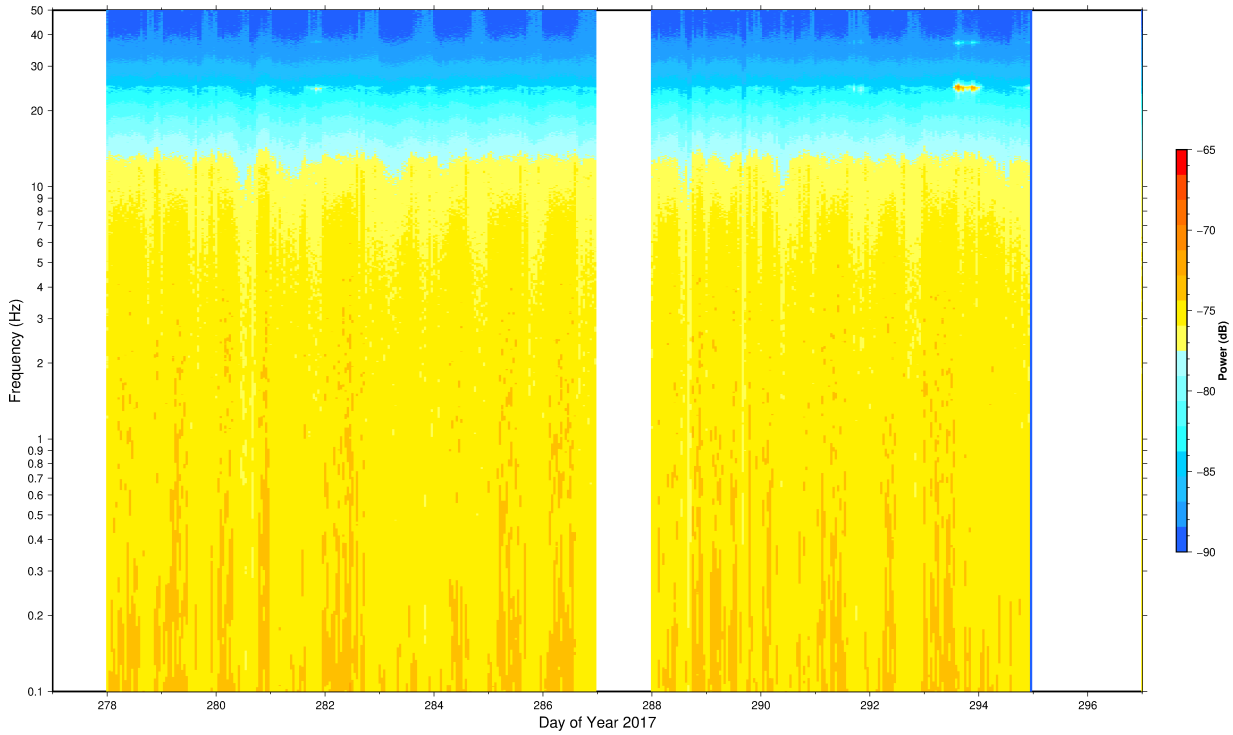


Figure 5.20: Spectrogram at deep drilled braced GPS station SIO5 on Mt. Soledad. Plot shows days 001-160 in 2018 at frequencies between 0.1 and 50 Hz. White areas denote gaps in data acquisition. There are no discernible peaks in the power spectral density calculated at this station.

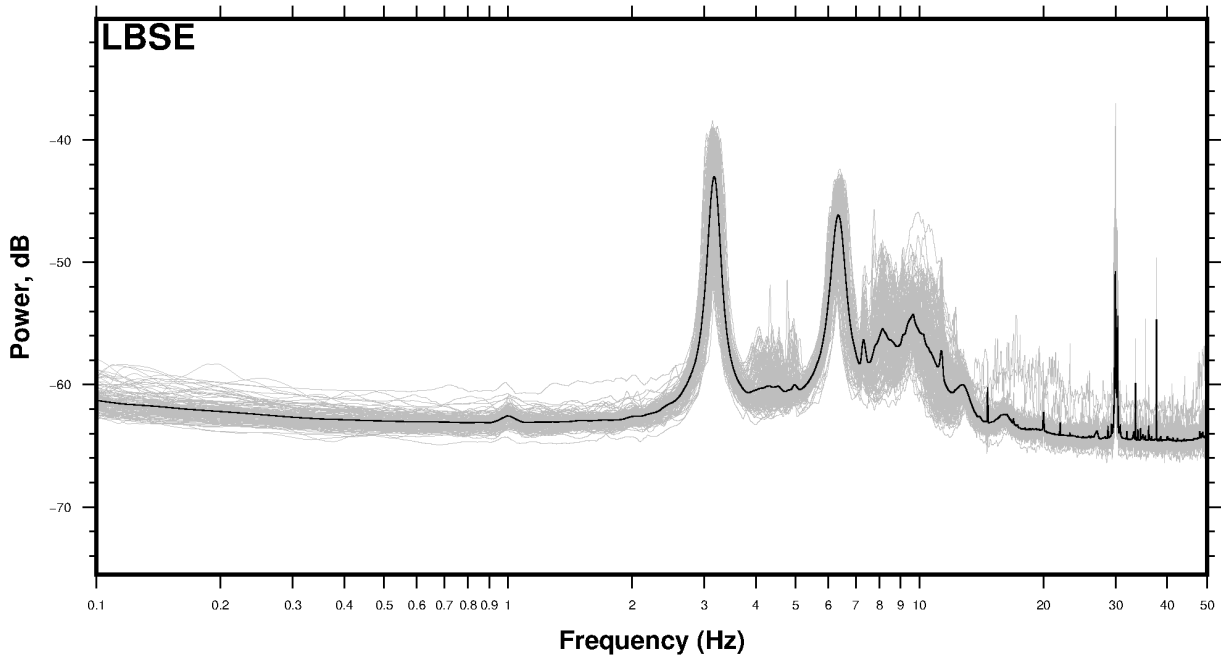


Figure 5.21: PSDs over the study period at rooftop station LBSE. Individual PSDs (grey lines) calculated for each day of study (2018-001 to 2018-160). Average PSD (black line) over the full time period.

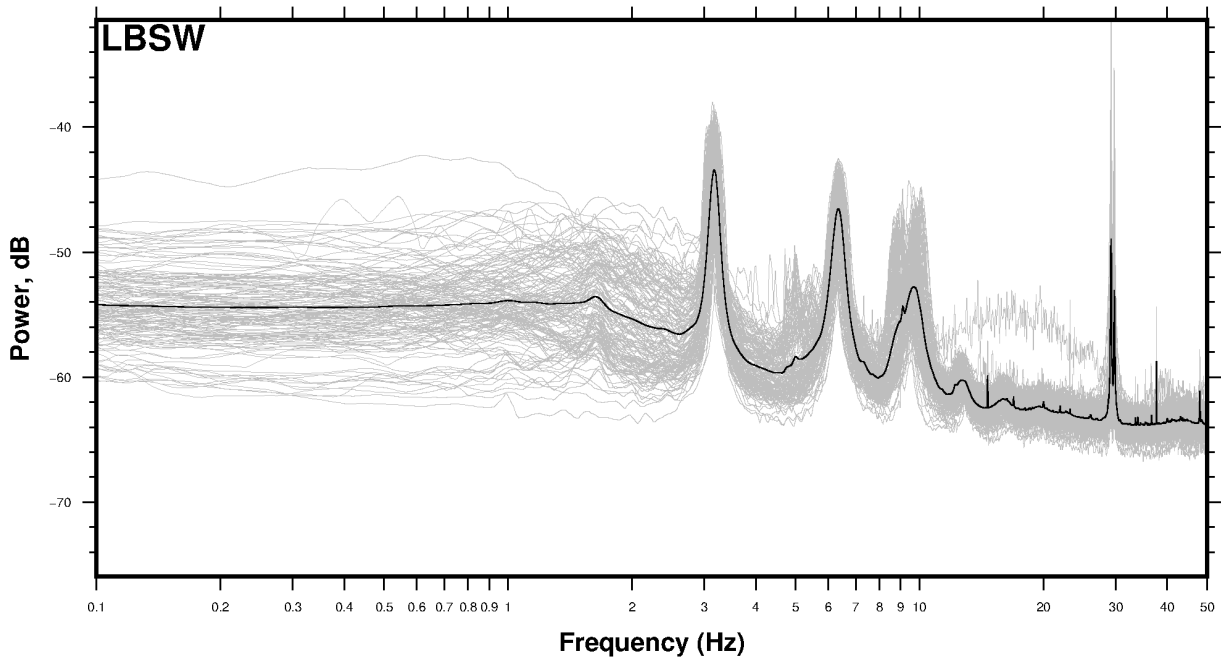


Figure 5.22: PSDs over the study period at rooftop station LBSE. Individual PSDs (grey lines) calculated for each day of study (2018-001 to 2018-160). Average PSD (black line) over the full time period.

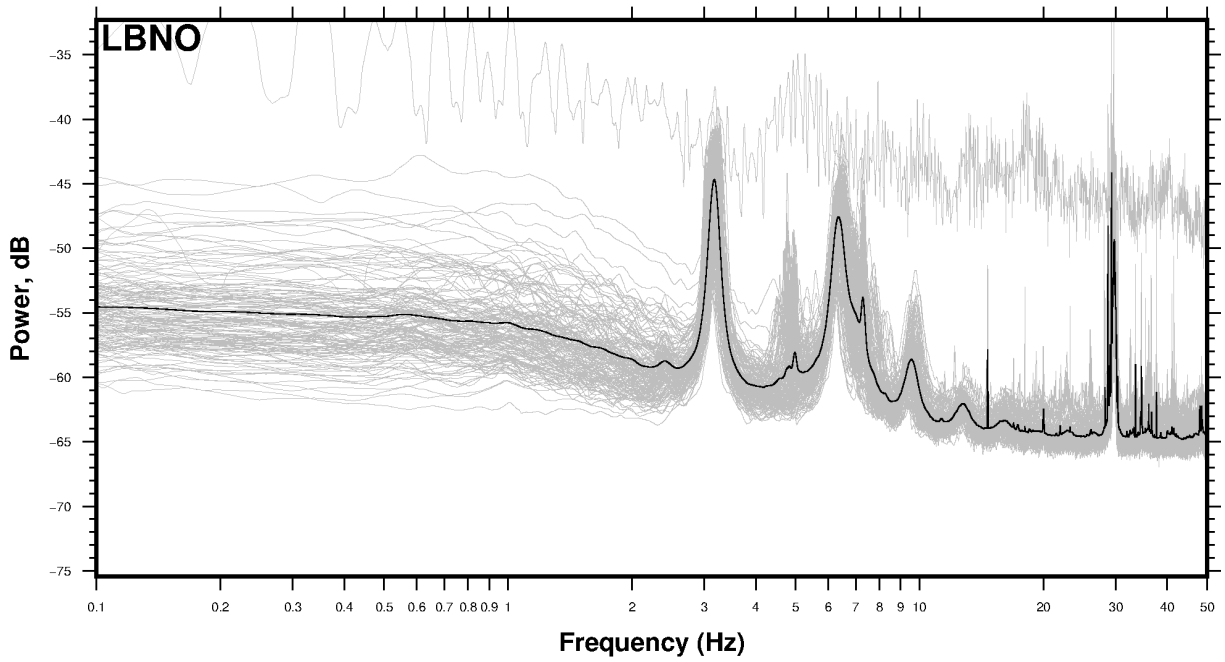


Figure 5.23: PSDs over the study period at rooftop station LBNO. Individual PSDs (grey lines) calculated for each day of study (2018-001 to 2018-160). Average PSD (black line) over the full time period.

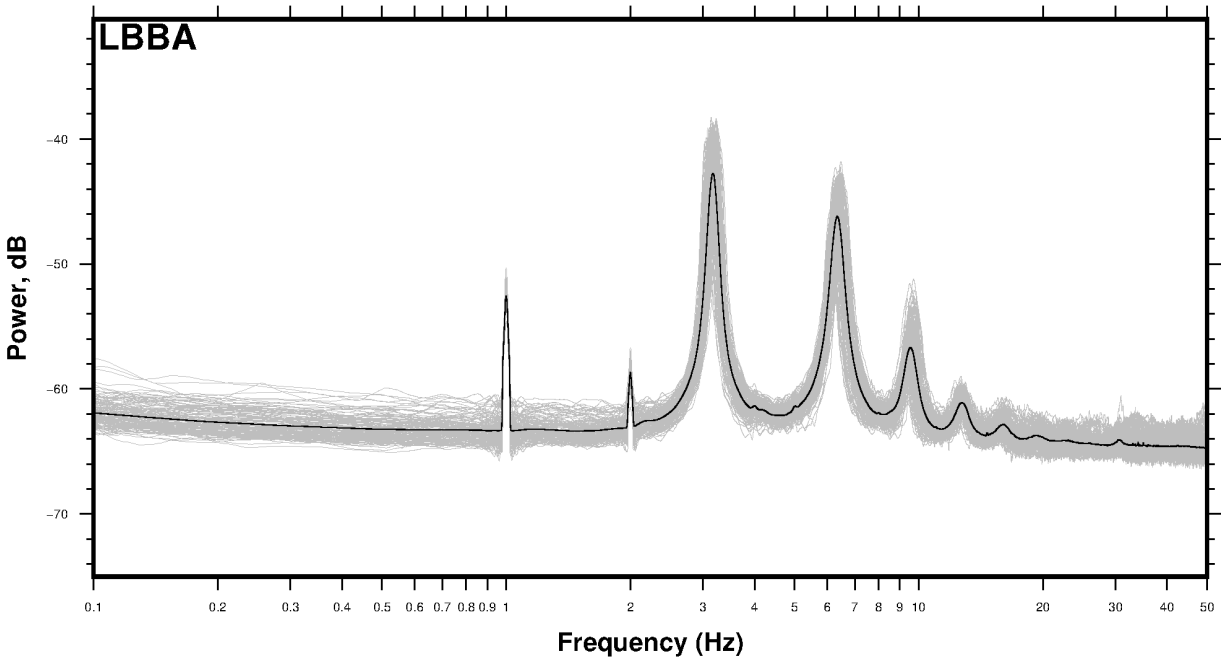


Figure 5.24: PSDs over the study period at basement station LBBA. Individual PSDs (grey lines) calculated for each day of study (2018-001 to 2018-160). Average PSD (black line) over the full time period.

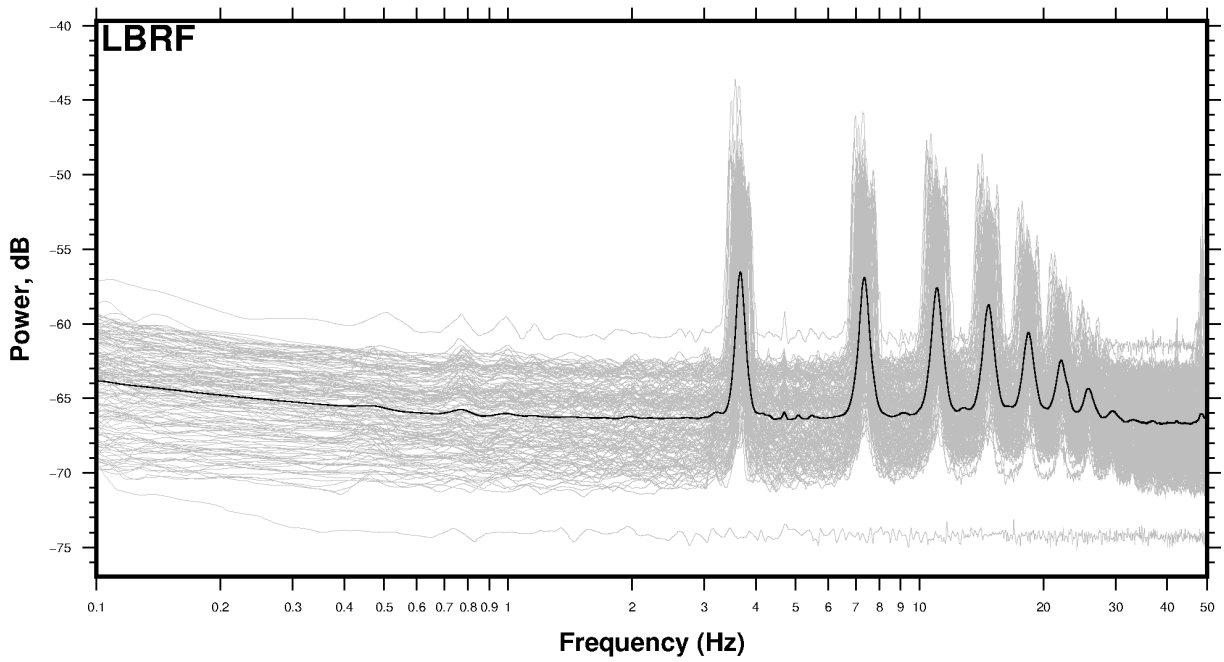


Figure 5.25: PSDs over the study period at reference station LBRF. Individual PSDs (grey lines) calculated for each day of study (2018-001 to 2018-160). Average PSD (black line) over the full time period.

This work represents the beginning stages of a larger effort to remove dependence on visual inspection and move toward a data-driven approach to structural monitoring. Following installation of the new generation of instruments, prolonged study of the wander of the structure's natural frequency will inform threshold settings to identify abnormal building behavior. As a continuation of this study, the building frequency response should be monitored on a near-real-time basis to study behavior with respect to weather, seasonal conditions, and building use. While the repeat studies are designed as individual snapshots from which structural changes can be identified, the seismogeodetic instrumentation provides a continuous record of structural health which helps to put the repeat surveys into longer-term context. Continuation of the structural monitoring procedures outlined here will involve the testing of automated data collection and analysis suitable for implementation in a natural hazard rapid response framework and eventual expansion to additional structures on the UCSD campus.

5.5 Acknowledgements

We thank our collaborators at the UCSD Qualcomm Institute, Falko Kuester and Eric Lo, as well as Rich Maher from KDM Meridian. This work is coauthored with Yehuda Bock, Falko Kuester, and Eric Lo, and is in preparation for publication. The dissertation author is the primary investigator and author of this material. We gratefully acknowledge funding support from the United States Army Corps of Engineers. Figures were produced using Generic Mapping Tools (Wessel et al., 2013).

Chapter 6

Conclusions and Future Outlook

This dissertation has shown the utility of merging geophysical instruments for understanding and responding to seismic hazards. Using traditional seismic datasets in combination with geodetic (GNSS) observations, we have addressed our objectives to devise and implement a fully seismogeodetic approach to earthquake and tsunami early warning, to investigate the source physics and rupture evolution of large, damaging earthquakes, to examine the limitations of rapid earthquake magnitude estimation, and to assess the utility of the low-cost seismic instrumentation for early warning and long-term structural monitoring.

6.1 Summary of the Dissertation

In Chapter 2, we outlined the framework for implementation of a prototype, self-contained seismogeodetic earthquake early warning (EEW) system. Using earthquakes in the western United States, Mexico, and Chile, we demonstrated that the seismogeodetic combination dataset can be used for accurate detection of seismic P-wave arrivals and assessment of earthquake location, while measuring the broad spectrum of seismic motion, including any permanent offsets. We described the lower limits of P-wave detection capabilities using low-cost micro-electro-mechanical systems (MEMS) accelerometers merged with observatory-grade GNSS and small to medium

sized events (2014 M_w4.0 Piedmont, CA and 2016 M_w5.2 Borrego Springs, CA earthquakes). We also suggested statistical methods by which to determine whether P-wave detections are consistent between stations, allowing us to remove poor quality detections and avoid propagating these errors to higher-order warning products. We analyzed our ability to resolve earthquake locations in areas of interest using the existing seismogeodetic network. From this analysis, we suggested existing GNSS stations to be prioritized for upgrade to seismogeodetic capability based on their ability to improve resolution of earthquake location in sparsely instrumented, offshore, or high-risk locations.

Chapter 3 addressed the objective of improving rapid earthquake source size estimations through seismogeodetic observation of P-wave amplitude. An open question in seismology surrounds the existence of earthquake determinism, the notion that properties observable early in rupture (e.g. the P-wave) are indicative of final rupture properties. If confirmed, deterministic rupture properties could be exploited to evaluate the eventual size of an earthquake prior to rupture completion, improving the timeliness of an early warning. Studies which rely on near-field seismic instrumentation display a limitation at large magnitudes wherein early onset parameters appear to scale with magnitude for small to moderate sized events, but scaling saturates for large magnitude events, making proper evaluation of earthquake size impossible above a certain threshold ($\sim$M7.5). Previous work (Crowell et al., 2013; Melgar et al., 2015a) has suggested that magnitude saturation is not inherent to the physics of large earthquakes, but rather due to the limitations of seismic instrumentation to properly quantify the long-period portion of the spectrum. P-wave amplitudes measured using seismic instrumentation alone become inaccurate at large magnitudes due to accelerometer biases, amplification of small errors upon integration from acceleration to displacement, or the filtering applied to ameliorate these issues. The seismogeodetic displacement solution is not plagued by such errors at long-periods, thus we investigated evidence for earthquake determinism in medium to large magnitude earthquakes observed by collocated seismic and GNSS stations. Despite previous work suggesting that P-

wave amplitude is predictive of final magnitude, we found that for a dataset of medium to large earthquakes regionally confined to Japan and each observed by 24-84 seismogeodetic stations, there is no evidence that P-wave amplitude is indicative of the final magnitude. We then tested at what point the maximum observed ground displacement diverges for large events, finding that in practice, the ability to distinguish earthquake size from near-field observations (within a few hundred km, as opposed to teleseismic distances) requires roughly one minute of data collection up to $\sim$M9, suggesting that magnitude is not discernible from the first few seconds of observation and that earthquake rupture is not strongly deterministic.

In Chapter 4, we investigate the possibility of weak determinism, that final earthquake properties can be estimated some time after rupture initiation, but prior to rupture completion. The available seismogeodetic dataset is limited in Japan, because the accelerometers are operated in triggered mode, and thus do not always begin collecting data prior to the P-wave arrival. Therefore, we relaxed the constraint that seismic and geodetic instruments be truly collocated, expanding the dataset to any GNSS station located within the geographical footprint of strong-motion stations that began recording prior to the P-wave arrival. In this way, we estimated the P-wave arrival time at each GNSS site via interpolation from the surrounding strong-motion instruments. With this larger dataset, we find that the final peak ground displacement is the first reliable indicator of earthquake magnitude. However, if stations are close enough to the source, we find that the observation of peak ground displacement occurs prior to rupture completion, implying a weak determinism. Through synthetic modeling of a simple, bilateral, pulse-slip rupture-propagation on a shallowly dipping thrust fault, we demonstrate that the final peak ground displacement is expected to be observed prior to rupture completion in the near field. We show similarity between the timing of peak ground displacement due to the modeled slip-pulse rupture and our observations, which suggests that large earthquakes behave as self-similar slip pulses of magnitude-dependent width. This rupture model describes a weakly deterministic rupture process such that the ability to determine slip-pulse width during rupture would allow reliable magnitude

estimation prior to rupture completion.

In Chapter 5, we apply the seismogeodetic approach to the task of structural monitoring, investigating the building response of the University of California San Diego Geisel Library. Repeat imaging of the structure provides a baseline of comparison for identifying any changes or damage following an earthquake or major weather event. The continuous GNSS position estimates between repeat surveys provide confidence in survey repeatability and are not indicative of any significant changes to the structure. Simultaneously, we analyzed continuous data collection from MEMS accelerometers located on the structure's roof and basement level during a period of quiescence with the goal of identifying any anomalous building behavior in the future. We chart the observed wander of the natural period of the building over several months, and correlate to local weather patterns. Ultimately, we identify a characteristic daily frequency fluctuation that appears to be related to the instrumentation itself, rather than the Geisel Library, preventing accurate understanding of the building behavior. Instead, we evaluate the behavior of the newest version of the Scripps Institution of Oceanography (SIO) MEMS accelerometer, and find that it does not share the erroneous frequency pattern of the older batch of instruments installed for monitoring of the Geisel Library. As a result of this work, the SIO MEMS accelerometers located on the Geisel Library were replaced with the updated instrument in September 2018. We expect this upgrade will allow proper observation of the natural frequency, and wander thereof, of the library structure. The continuation of this work, and application of these analyses to the new instrumentation, will contribute to an ongoing United States Army Corps of Engineers collaboration, the goal of which is to develop an automated, objective, data-driven approach to identify structural changes over time and provide metrics that can be used to evaluate any compromises to structural integrity.

6.2 Future Directions

There is considerable work left to do to improve early warning efforts, in light of recent research. Importantly, GNSS is still underutilized in operational systems. The G-FAST system (Crowell et al., 2016), demonstrated in the Pacific Northwest, United States is a notable exception. While the prototype seismogeodetic system for the western U.S. runs at the Scripps Orbit and Permanent Array Center, seismogeodetic capabilities such as those described in this dissertation require additional validation and testing to compel EEW operators to ingest GNSS and seismogeodetic data streams. GNSS is best utilized where we expect large ground motions. In that vein, subduction zone interfaces accommodate the majority of large (M7+) and all of the great (M9+) earthquakes. This means that GNSS is well-designed to elucidate the conditions at these plate boundaries, and rapidly identify the source properties of large seismic events. We are currently providing GNSS and seismogeodetic data to the National and Pacific Tsunami Warning Centers (TWCs) of the National Oceanic and Atmospheric Administration (NOAA), which are responsible for basin-wide and local tsunami warning (LTW). The seismogeodetic approach described within previous chapters is undergoing testing for implementation within the TWC operations. Importantly, the real-time seismogeodetic data stream is expected to improve rapid magnitude estimates, which will enhance the accuracy of LTW, and contribute to hazard mitigation efforts in the near-coastal regions of tsunamigenic earthquakes.

Rapid, accurate estimation of earthquake magnitude, a main focus of this dissertation, is still the primary metric required for early warning applications. Rapid confirmation of the magnitude estimate through calculation of the centroid moment tensor (CMT) solution will improve confidence and reliability of the warning. The CMT solution is a concise description of the size of the earthquake and orientation of faulting, and can also be used as a first-order estimate of the affected areas, bolstering rapid response efforts. As an example, Figure 6.1 shows an estimate of the region of permanent deformation resulting from the 2016 M_w7.0 Kumamoto,

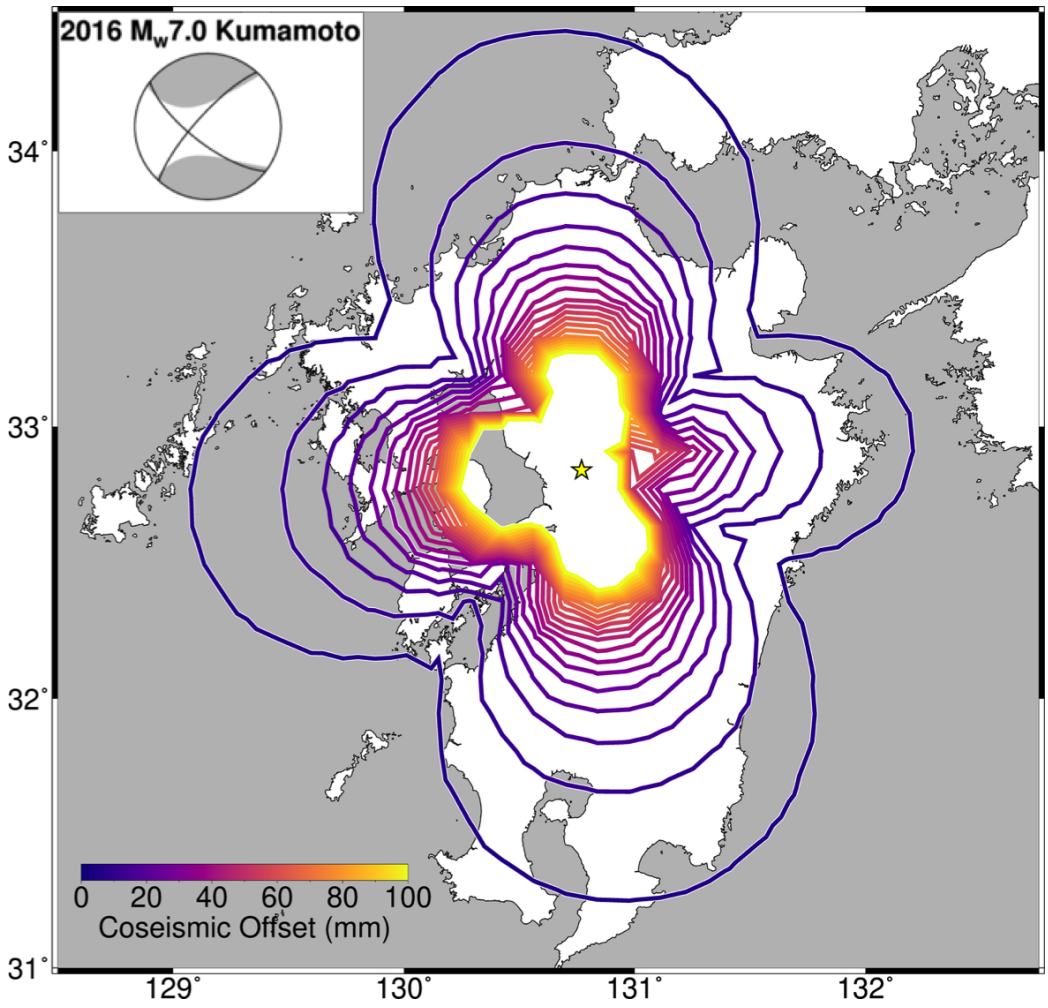


Figure 6.1: Estimated region of permanent deformation for the 2016 M_w7.0 Kumamoto, Japan earthquake, from a simple point source model of the centroid moment tensor solution, adapted from the open-source MudPy code (Melgar and Bock, 2015).

Japan strike-slip earthquake, derived from a point-source CMT solution. For offshore events, the CMT solution goes further to estimate the degree of vertical deformation, which directly affects the tsunamigenic potential of an earthquake. In ongoing work, and in collaboration with the NOAA TWCs, we continue to validate and implement rapid CMT estimation procedures from GNSS and seismogeodetic instrumentation (Melgar et al., 2012), and develop products (e.g. Figure 6.1) expected to benefit EEW and LTW efforts. Currently implemented rapid W-phase CMT solutions (Hayes et al., 2009; Dupatel et al., 2011) are limited in solution speed because they require that observations used to estimate the CMT be located at far enough distances such that the point-source assumption is valid. Including near-field geodetic and seismogeodetic data and expanding the analysis to a line-source CMT solution (finite fault assumption) as a continuation of this collaboration with the TWCs will further improve timeliness for LTW.

For offshore seismicity, it is becoming increasingly clear that land-based monitoring reveals only a limited perspective of processes at the subduction interface. The study of best practices for earthquake and tsunami early warning will certainly benefit from the integration of seafloor seismic and geodetic instrumentation. Tsunamigenic earthquakes have been modeled using a variety of instrument types, including land-based seismic instruments (e.g Suzuki et al., 2011), GNSS (e.g. Ozawa et al., 2011; Ohta et al., 2012), ocean bottom seismometers (OBS) (e.g. Eguchi et al., 1998), Global Positioning System Acoustics (GPS-A) (e.g. Spiess et al., 1998), ocean bottom pressure (OBP) sensors (e.g. Kubota et al., 2017), tide gauges (e.g. Merrifield et al., 2005), real-time kinematic GPS buoys (e.g. Kato et al., 2014), or some combination thereof (e.g. Pollitz et al., 2011; Melgar and Bock, 2013; Bletery et al., 2014). OBS instruments, designed as high-gain, broadband sensors suffer from the same clipping issues as land-based broadband seismometers in the near-field of large earthquakes. OBP and GPS-A instrumentation, however, are better-suited for the observation of large, near-field seismic events. Like land-based GPS, GPS-A provides a direct measurement of displacement, while OBPs can be used to interpret the vertical acceleration record, much like a strong-motion accelerometer on land, in the vertical component

of motion. Seafloor seismogeodesy is therefore an exciting frontier of rapidly expanding research in this field.

Cabled arrays, allowing for real-time observation of the seafloor, have the potential to dramatically improve the timeliness and accuracy of warnings for off-shore earthquakes. Japan's S-net system (e.g. Maeda et al., 2015), the VENUS and NEPTUNE arrays offshore Vancouver Island (Tunnicliffe et al., 2008), and the Ocean Observatories Initiative Cabled Array in Cascadia (Cowles et al., 2010) will inevitably capture events of interest to earthquake and tsunami early warning, thus an important stage of research will be the validation and implementation of real-time algorithms applicable to these seafloor instruments. Cabled arrays will improve near-field observation density, azimuthal coverage, timeliness of event detection, and reliability of tsunami forecasts.

In more general terms, subduction zones processes are not fully understood, in large part because of the difficulty of instrumenting these regions. Therefore, consideration of all available sensor types– land-based, seafloor, seismic, or geodetic– for these purposes will contribute not only to the advancement of early warning capabilities, but also to longer-term understanding of these plate boundaries in a way that will benefit general knowledge of tectonic processes. It is clear that because each instrument has its own strengths and weaknesses, our understanding of geophysical processes and the resulting hazards will be improved by incorporating a wide variety of observation tools.

6.3 Acknowledgements

The contents of this chapter are unpublished. The centroid moment tensor solution used in Figure 6.1 is from the Global Centroid Moment Tensor Project catalog, www.globalcmt.org. Figures were produced using Generic Mapping Tools (Wessel et al., 2013).

Appendix A

Merging Instrument Types: Seismogeodetic Kalman Filter Formulations

A.1 Background

The complex nature of earthquakes makes it difficult to accurately describe the properties of an event with a single instrumentation type. The three main earthquake monitoring instruments are broadband seismometers, strong-motion accelerometers, and Global Navigation Satellite Systems (GNSS). By design, each instrument has a frequency and amplitude of ground motion that it is best able to observe.

Broadband seismometers measure ground velocities at high sampling rates (typically 100-200 Hz), and have noise floors low enough to properly observe teleseismic earthquakes, despite attenuation through the Earth. The drawback of broadband seismometers is that they are limited in the intensity of ground motions that they can properly observe. Heavy shaking and

large ground motions in the nearfield of large seismic events exceeds the dynamic range of these instruments, in a problem known as *clipping*. Furthermore, as records approach the limits of the instrument's dynamic range, the instrument response becomes nonlinear, making it difficult to interpret the true ground motions.

Strong-motion accelerometers measure acceleration at sampling rates similar to broadband seismometers, but are better suited for large ground motions. Accelerometers are designed with lower gains, which allows the on-scale recording of shaking, even for the largest earthquakes. However, these instruments are unable to identify the difference between translational and rotational motions, meaning that tilts, rotations, and translations cannot be separated, and the true ground motion can be difficult to resolve. Furthermore, errors due to improper inclusion of tilts and rotations are amplified upon integration to velocity or double integration to displacement. As a result, integrated velocities include a linear drift, and doubly integrated displacements include a quadratic drift. Of course, this drift is unphysical and must be removed. There have been many methods put forth to remove the drift; the simplest method in real-time is a high-pass filter (Boore and Bommer, 2005). Removal of the long-period information makes it impossible to resolve a large earthquake's size, because the 0 Hz static offset is eliminated, and the amplitude of total displacement is underestimated (e.g. Melgar et al., 2015a).

GNSS provides a direct measurement of displacement in a defined reference frame, which allows proper estimation of the static offset and peak ground displacement caused by an earthquake. GNSS instrumentation is typically sampled at much lower sampling rates than seismic instrumentation (1-10 Hz for GNSS as opposed to 100-200 Hz for seismic). With current technologies, 10 Hz is the maximum sampling rate of real-time GNSS, limited by the verbosity of the phase and pseudorange observations to multiple satellites at multiple wavelengths for each epoch that must be accommodated for telemetry. Therefore, while GNSS is better suited

to measuring the long-period earthquake information, the high-frequency information is not adequately collected.

Bock et al. (2011) demonstrated a solution to this problem, in the form of a real-time Kalman filter combination of collocated GNSS observations and strong-motion accelerations. The Kalman filter is an algorithm that uses different measurements of the same quantity, each containing inaccuracies (biases, noise, etc.) to estimate the desired quantity with better accuracy than either measurement alone. The algorithm is divided into two main parts: the prediction step, and the measurement step. In the former, the current system state variables are predicted from their past behavior. In the latter, the prediction is updated by a newly acquired measurement, which itself is subject to some form of error or uncertainty. The output estimate is some weighted average of the predicted value and measured value, based upon the estimated uncertainties in the two.

The merging of the strong-motion accelerometer and GNSS results in a combination dataset that has the sampling rate of the accelerometer, is informed by the long-period information of the GNSS instrument, and has reduced noise compared to GNSS-only displacements. See Bock et al. (2011) and Melgar Moctezuma (2014) for a detailed description of the original Kalman filter seismogeodetic formulation. One drawback to the method is the relatively limited availability of collocated stations, because the GNSS and seismic networks have, until recently, been developed independently. In order to take advantage of the available instrumentation, this appendix includes derivations of Kalman filter formulations for additional combinations of instrumentation (GNSS, broadband seismometer, strong motion accelerometer).

In the following derivations, $d(t)$ is displacement, $v(t)$ is velocity, $a(t)$ is acceleration, and $\boldsymbol{x}$ is the system state, and $\boldsymbol{u}$ is the deterministic input. The first derivation will go through each step in detail, while the remaining instrument combinations will simply define the matrices

required for implementation.

A.2 Kalman Filter: GNSS and Broadband Velocity

It is reasonable to merge high-rate GNSS with broadband velocity data in regions where high amplitude shaking is not expected (i.e. no seismometer clipping). We follow Lewis et al. (2008) to write the continuous difference equation for the system:

The system state, $\boldsymbol{x}$, in this case is a 1x1 matrix of displacement, d.

$$\frac{d}{dt}\left[\ d(t)\ \right] = \frac{d}{dt}\boldsymbol{x}(t) = \boldsymbol{A}(t)\boldsymbol{x}(t) + \boldsymbol{B}(t)u(t) + \varepsilon(t) \tag{A.1}$$

where

$$x(t) = d(t) \qquad A = 0 \qquad B = 1 \qquad u(t) = v(t) \qquad \varepsilon = \varepsilon_v \quad . \tag{A.2}$$

ε_v is the noise in the seismometer at each time step. We assume Gaussian noise and the noise vector is distributed like $\varepsilon \sim (0, \boldsymbol{Q})$ where the covariance $\boldsymbol{Q}$ depends on the velocity noise, σ_v:

$$Q = \sigma_v \tag{A.3}$$

We can combine equations A.1 and A.2 to verify the system:

$$\begin{aligned} \frac{d}{dt}d(t) &= 0 \cdot d(t) + 1 \cdot v(t) + \varepsilon_v \\ \dot{d} &= v + \varepsilon_v \end{aligned} \tag{A.4}$$

We discretize the continuous definition such that,

$$\boldsymbol{x}_{k+1} = \boldsymbol{A}^s \boldsymbol{x}_k + \boldsymbol{B}^s v_k + \varepsilon_k \tag{A.5}$$

where the superscript s denotes the discretized version of the state transition matrices and the subscript k denotes a discrete time step. The discretized noise vector is assumed to be distributed as $\varepsilon_k \sim (0, \boldsymbol{Q}^s)$. We sample the continuous system at the sampling rate of the broadband seismometer, τ_v. In the discretized format, the state transition matrices are obtained by the first few terms of the infinite series expansion of the integral forms (Lewis et al., 2008):

$$\boldsymbol{A}^s = \boldsymbol{I} + \boldsymbol{A}\tau_v + \frac{\boldsymbol{A}^2\tau_v^2}{2} = 1$$

$$\boldsymbol{B}^s = \boldsymbol{B}\tau_v + \frac{\boldsymbol{A}\boldsymbol{B}\tau_v^2}{2} = \tau_v \tag{A.6}$$

$$\boldsymbol{Q}^{\boldsymbol{s}} = \boldsymbol{Q}\tau_v + \frac{1}{2}(\boldsymbol{A}\boldsymbol{Q} + \boldsymbol{Q}\boldsymbol{A}^T)\tau_v^2 + \frac{1}{3}\boldsymbol{A}\boldsymbol{Q}\boldsymbol{A}^T\tau_v^3 = \sigma_v\tau_v \quad .$$

The Kalman Filter formulation begins by introducing a measurement of the system states. We measure only displacement, such that:

$$z_k = d_k^{obs} = \boldsymbol{H}^{\boldsymbol{s}}\boldsymbol{x}_k + \eta_d \tag{A.7}$$

where d_k^{obs} is the GNSS measurement of displacement of epoch k with white Gaussian noise such that $\eta_d \sim (0, R^s)$. The sampling rate of the GNSS is τ_d, making the discretized matrices:

$$\boldsymbol{H}^{\boldsymbol{s}} = \boldsymbol{H} = 1 \tag{A.8}$$

$$R = \sigma_d, \qquad R^s = \frac{\sigma_d}{\tau_d} \tag{A.9}$$

We initialize the system states, x_0, and covariance, $\boldsymbol{P}_0$, as zero and the identity matrix,

respectively. Next, we estimate the *a priori* covariance as

$$\boldsymbol{P}_{k+1}^{-} = \boldsymbol{A}_k \boldsymbol{P}_k \boldsymbol{A}_k^T + \boldsymbol{Q}_k \tag{A.10}$$

and the a priori state estimate as

$$\boldsymbol{x}_{k+1}^{-} = \boldsymbol{A}_k \hat{\boldsymbol{x}}_k + \boldsymbol{B}_k u_k \tag{A.11}$$

where $u_k = v_k^{\text{obs}}$, the seismometer observation at each time step, from Equation A.2. The $\hat{\boldsymbol{x}}$ notation indicates an estimated quantity, and the $^{-}$ indicates a quantity obtained prior to a measurement update. These denote the *time update* stage of the filter, where system states are estimated in the absence of additional measurements. In the subsequent step, we include the *measurement update*:

$$\boldsymbol{P}_{k+1} = \left[(\boldsymbol{P}_{k+1}^{-})^{-1} + \boldsymbol{H}_{k+1}^T \boldsymbol{R}_{k+1}^{-1} \boldsymbol{H}_{k+1} \right]^{-1} \tag{A.12}$$

and the system state estimate is updated to:

$$\hat{\boldsymbol{x}}_{k+1} = \hat{\boldsymbol{x}}_k^{-} + \boldsymbol{K}_{k+1}(z_{k+1} - \boldsymbol{H}_{k+1} \boldsymbol{x}_{k+1}^{-}) \tag{A.13}$$

where $\boldsymbol{K}$, the *Kalman gain* is defined as:

$$\boldsymbol{K}_{k+1} = \boldsymbol{P}_{k+1} \boldsymbol{H}_{k+1}^T \boldsymbol{R}_{k+1}^{-1} \tag{A.14}$$

In summary, the Kalman gain, $\boldsymbol{K}$, weights the adjustment to the a priori estimate $\hat{\boldsymbol{x}}_k^{-}$ whenever a measurement of the system states z_k is available. $z_{k+1} - \boldsymbol{H}_{k+1} \boldsymbol{x}_{k+1}^{-}$ is the difference between the a priori state estimate and the observed measurement. Thus, the Kalman gain determines how much to trust the a priori estimate of the system state based on the seismometer, versus how much to correct it based on the GNSS measurement of displacement, since the Kalman gain is dependent on the covariance matrix $\boldsymbol{Q}$ of the seismometer and the covariance $\boldsymbol{R}$

of the GNSS measurements. The sampling rates of these two instruments are usually different (seismometer at higher sampling rates than GNSS). The output of the Kalman filter can have the frequency of the higher-sampled instrument, with a Kalman gain of zero when no measurement update is available. Equations for the forward filter and optional smoother are listed in tables A.1 and A.2, respectively.

Optional smoothing can be implemented in the same way on each of the following derivations, using the Rauch, Tung and Striebel smoother (RTS) (Rauch et al., 1965), and therefore is listed only in Section A.2 to avoid repetition. As in Melgar Moctezuma (2014), we use the term "N-sample Smoother" to denote that for near-real-time computation, the smoother can be applied using an N-sample lag behind the real-time data stream, rather than applying to the complete time series.

Table A.1: Summary of Kalman filter with GPS (measurement input) and broadband seismometer (deterministic input).

Forward Filter	
Initialize States	$\hat{\boldsymbol{x}}_0 = 0$
Initialize Covariance	$\boldsymbol{P}_0 = \boldsymbol{I}$
Time Update Covariance	$\boldsymbol{P}_{k+1}^{-} = \boldsymbol{A}_k \boldsymbol{P}_k \boldsymbol{A}_k^T + \boldsymbol{Q}_k$
Time Update System States	$\hat{\boldsymbol{x}}_{k+1}^{-} = \boldsymbol{A}_k \hat{\boldsymbol{x}}_k + \boldsymbol{B}_k v_k$
Introduce Measurement	$z_k = d_k$
Compute Kalman Gain	$\boldsymbol{K}_{k+1} = \boldsymbol{P}_{k+1} \boldsymbol{H}_{k+1}^T \boldsymbol{R}_{k+1}^{-1}$
Measurement Update of Covariance	$\boldsymbol{P}_{k+1} = \left[(\boldsymbol{P}_{k+1}^{-})^{-1} + \boldsymbol{H}_{k+1}^T \boldsymbol{R}_{k+1}^{-1} \boldsymbol{H}_{k+1} \right]^{-1}$
Measurement Update of System States	$\hat{\boldsymbol{x}}_{k+1} = \hat{\boldsymbol{x}}_k^{-} + \boldsymbol{K}_{k+1}(z_{k+1} - \boldsymbol{H}_{k+1} \boldsymbol{x}_{k+1}^{-})$

Table A.2: RTS smoother formulation.

RTS N-sample Smoother	
Initialize States	$\boldsymbol{x}_N = \boldsymbol{x}_k$
Initialize Covariance	$\boldsymbol{P}_N = \boldsymbol{P}_k^f$
Compute Smoother Gain	$\boldsymbol{F}_k = \boldsymbol{P}_k^f \boldsymbol{A}_k (\boldsymbol{P}_{k+1}^{f-})^{-1}$
Update Covariance	$\boldsymbol{P}_k = \boldsymbol{P}_k^f - \boldsymbol{F}_k (\boldsymbol{P}_{k+1}^{f-} - \boldsymbol{P}_{k+1}) \boldsymbol{F}_k^T$
Update System States	$\hat{\boldsymbol{x}}_k = \hat{\boldsymbol{x}}_k^f + \boldsymbol{F}_k (\hat{\boldsymbol{x}}_{k+1} - \hat{\boldsymbol{x}}_{k+1}^{f-})$

A.3 Real-Time Kalman Filter, GNSS and Accelerometer

Real-time data may experience gaps or poor telemetry from one of the instruments, thus both the GNSS displacements and accelerations must be used as measurements rather than using acceleration as a system input, as in Bock et al. (2011).

The system states are therefore the displacement, d, the velocity, v, the acceleration, a, and the accelerometer DC offset, Ω. The DC level of the strong-motion instrument is typically non-zero, due to small tilts from installation or changes in the site conditions over time, or because of normal instrument aging. If the DC level bias is unaccounted for, a spurious constant acceleration is introduced into the filter. A real-time implementation of the Kalman filter requires that the DC is estimated continually and corrected, reducing bias in the calculation. An abbreviated description of the relevant matrices are described below:

$$\frac{d}{dt}\boldsymbol{x}(t) = \boldsymbol{A}(t)\boldsymbol{x}(t) + \boldsymbol{B}(t)u(t) + \varepsilon(t) \tag{A.15}$$

There is no deterministic input, $u(t)$, thus Equation A.15 is reduced to

$$\frac{d}{dt}\boldsymbol{x}(t) = \boldsymbol{A}(t)\boldsymbol{x}(t) + \varepsilon(t) \tag{A.16}$$

where

$$x(t) = \begin{bmatrix} d(t) \\ v(t) \\ a(t) \\ \Omega(t) \end{bmatrix}, \qquad A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \qquad \varepsilon = \begin{bmatrix} 0 \\ 0 \\ \varepsilon_a \\ \varepsilon_\Omega \end{bmatrix} \tag{A.17}$$

where ε_a is the accelerometer noise and ε_Ω is the DC offset noise. Here, we are assuming that Ω has noise, and can change in small amounts through time (i.e. non-zero variance). But, there should also be process noise in the accelerations due to rotations. This means that ε_a would be small during periods of quiescence, and larger during heavy shaking, allowing the accelerations in the output to fluctuate as necessary.

We assume that $\varepsilon \sim (0, \boldsymbol{Q})$, where

$$\boldsymbol{Q} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_a & 0 \\ 0 & 0 & 0 & \sigma_\Omega \end{bmatrix} \tag{A.18}$$

We discretize the continuous system at the sampling rate we wish to have as output. This will be equal to the accelerometer sampling rate, τ_a. However, we restate this sampling rate as the Kalman output rate, τ_k, thus allowing τ_a to change in time (data gaps/ drop outs), but keeping the sampling rate of the output constant.

$$\boldsymbol{A}^s = \boldsymbol{I} + \boldsymbol{A}\tau_k + \frac{\boldsymbol{A}^2\tau_k^2}{2} = \begin{bmatrix} 1 & \tau_k & \frac{\tau_k^2}{2} & 0 \\ 0 & 1 & \tau_k & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{A.19}$$

$$\boldsymbol{Q^s} = \boldsymbol{Q}\tau_k + \frac{1}{2}(\boldsymbol{AQ} + \boldsymbol{QA}^T)\tau_k^2 + \frac{1}{3}\boldsymbol{AQA}^T\tau_k^3 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{\sigma_a\tau_k^3}{3} & \frac{\sigma_a\tau_k^2}{2} & 0 \\ 0 & \frac{\sigma_a\tau_k^2}{2} & \sigma_a\tau_k & 0 \\ 0 & 0 & 0 & \sigma_\Omega\tau_k \end{bmatrix} \tag{A.20}$$

The formulation begins by introducing a measurement of the system states. We measure displacement and acceleration, such that:

$$\boldsymbol{z}_k = \begin{bmatrix} d_k^{\text{obs}} \\ a_k^{\text{obs}} \end{bmatrix} = \boldsymbol{H}\boldsymbol{x}_k + \eta \tag{A.21}$$

$$\boldsymbol{H^s} = \boldsymbol{H} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \tag{A.22}$$

$$\eta \sim (0, \boldsymbol{R}^s) \qquad \boldsymbol{R} = \begin{bmatrix} \sigma_d & 0 \\ 0 & \sigma_a \end{bmatrix} \qquad \boldsymbol{R^s} = \frac{\boldsymbol{R}}{\tau} = \begin{bmatrix} \frac{\sigma_d}{\tau_d} & 0 \\ 0 & \frac{\sigma_a}{\tau_a} \end{bmatrix} \tag{A.23}$$

These are then applied in the time and measurement updates as in the original Kalman filter. For time steps that include only a measurement update from the accelerometer, the bottom rows of $\boldsymbol{H}$ and R^s are used. For those including only a GPS update, the top rows are used.

Equations for the forward filter are listed in Table A.3.

Table A.3: Summary of Kalman filter formulation with GPS and strong-motion accelerometer (measurement inputs), designed for implementation in real time.

Forward Filter	
Initialize States	$\hat{\boldsymbol{x}}_0 = [0,0,0,0]^T$
Initialize Covariance	$\boldsymbol{P}_0 = \boldsymbol{I}$
Time Update Covariance	$\boldsymbol{P}^-_{k+1} = \boldsymbol{A}_k \boldsymbol{P}_k \boldsymbol{A}_k^T + \boldsymbol{Q}_k$
Time Update System States	$\hat{\boldsymbol{x}}^-_{k+1} = \boldsymbol{A}_k \hat{\boldsymbol{x}}_k$
Introduce Measurement	$z_k = [d_k, a_k]^T$
Compute Kalman Gain	$\boldsymbol{K}_{k+1} = \boldsymbol{P}_{k+1} \boldsymbol{H}^T_{k+1} \boldsymbol{R}^{-1}_{k+1}$
Measurement Update of Covariance	$\boldsymbol{P}_{k+1} = \left[(\boldsymbol{P}^-_{k+1})^{-1} + \boldsymbol{H}^T_{k+1} \boldsymbol{R}^{-1}_{k+1} \boldsymbol{H}_{k+1} \right]^{-1}$
Measurement Update of System States	$\hat{\boldsymbol{x}}_{k+1} = \hat{\boldsymbol{x}}^-_k + \boldsymbol{K}_{k+1}(z_k + 1) - \boldsymbol{H}_{k+1} \boldsymbol{x}^-_{k+1}$

A.4 3 Instrument Input: GPS, Broadband, Accelerometer

A.4.1 Broadband Seismometer as Deterministic Input

Assume that we have a three-instrument input of strong-motion accelerations, $a(t)$, broadband seismic velocities, $v(t)$, and GPS positions, $d(t)$, such that $\tau_a = \tau_v$ and $\tau_a > \tau_d$. We assume that the strong motion and broadband seismic instruments are deterministic input, and GPS positions are the measurements.

$$\frac{d}{dt}\boldsymbol{x}(t) = \boldsymbol{A}(t)\boldsymbol{x}(t) + \boldsymbol{B}(t)u(t) + \varepsilon(t) \tag{A.24}$$

where

$$x(t) = \begin{bmatrix} d(t) \\ v(t) \end{bmatrix} \quad A = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{A.25}$$

$$u(t) = \begin{bmatrix} v(t) \\ a(t) \end{bmatrix} \quad \varepsilon = \begin{bmatrix} \varepsilon_v \\ \varepsilon_a \end{bmatrix}$$

We assume $\varepsilon \sim (0, \boldsymbol{Q})$, where:

$$Q = \begin{bmatrix} \sigma_v & 0 \\ 0 & \sigma_a \end{bmatrix} \tag{A.26}$$

Discretized,

$$\boldsymbol{A}^s = \boldsymbol{I} + \boldsymbol{A}\tau_a + \frac{\boldsymbol{A}^2\tau_a^2}{2} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\boldsymbol{B}^s = \boldsymbol{B}\tau_a + \frac{\boldsymbol{AB}\tau_a^2}{2} = \begin{bmatrix} \tau_a & 0 \\ 0 & \tau_a \end{bmatrix} \tag{A.27}$$

$$\boldsymbol{Q^s} = \boldsymbol{Q}\tau_a + \frac{1}{2}(\boldsymbol{AQ} + \boldsymbol{QA}^T)\tau_a^2 + \frac{1}{3}\boldsymbol{AQA}^T\tau_a^3 = \begin{bmatrix} \sigma_v\tau_a & 0 \\ 0 & \sigma_a\tau_a \end{bmatrix}$$

The formulation begins by introducing a measurement of the system states. We assume our measurement is only displacement, such that:

$$z_k = d_k^{obs} = \boldsymbol{H^s x}_k + \eta_d \tag{A.28}$$

$$\boldsymbol{H^s} = \boldsymbol{H} = \begin{bmatrix} 1 & 0 \end{bmatrix} \tag{A.29}$$

$$\eta_d \sim (0, R^s) \qquad R = \sigma_d \qquad R^s = \frac{\sigma_d}{\tau_d} \tag{A.30}$$

A summary is given in Table A.4.

Table A.4: Kalman filter formulation with GPS (measurement input), broadband seismometer (deterministic input), and strong-motion accelerometer (deterministic input).

Forward Filter	
Initialize States	$\hat{\boldsymbol{x}}_0 = [0,0]^T$
Initialize Covariance	$\boldsymbol{P}_0 = \boldsymbol{I}$
Time Update Covariance	$\boldsymbol{P}_{k+1}^- = \boldsymbol{A}_k \boldsymbol{P}_k \boldsymbol{A}_k^T + \boldsymbol{Q}_k$
Time Update System States	$\hat{\boldsymbol{x}}_{k+1}^- = \boldsymbol{A}_k \hat{\boldsymbol{x}}_k + \boldsymbol{B}_k [v_k, a_k]^T$
Introduce Measurement	$z_k = d_k$
Compute Kalman Gain	$\boldsymbol{K}_{k+1} = \boldsymbol{P}_{k+1} \boldsymbol{H}_{k+1}^T \boldsymbol{R}_{k+1}^{-1}$
Measurement Update of Covariance	$\boldsymbol{P}_{k+1} = \left[(\boldsymbol{P}_{k+1}^-)^{-1} + \boldsymbol{H}_{k+1}^T \boldsymbol{R}_{k+1}^{-1} \boldsymbol{H}_{k+1} \right]^{-1}$
Measurement Update of System States	$\hat{\boldsymbol{x}}_{k+1} = \hat{\boldsymbol{x}}_k^- + \boldsymbol{K}_{k+1}(z_{k+1}) - \boldsymbol{H}_{k+1} \boldsymbol{x}_{k+1}^-$

A.4.2 Broadband Seismometer as Deterministic Input with Accelerometer Biases

Again, in a real-time operational framework, the strong-motion accelerometer biases must be estimated and accounted for in real time. So, we add estimation of the DC offset, Ω, to Section A.4.1.

$$\frac{d}{dt}\boldsymbol{x}(t) = \boldsymbol{A}(t)\boldsymbol{x}(t) + \boldsymbol{B}(t)u(t) + \varepsilon(t) \tag{A.31}$$

where

$$x(t) = \begin{bmatrix} d(t) \\ v(t) \\ \Omega(t) \end{bmatrix} \quad A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \tag{A.32}$$

$$u(t) = \begin{bmatrix} v(t) \\ a(t) \end{bmatrix} \quad \varepsilon = \begin{bmatrix} \varepsilon_v \\ \varepsilon_a \\ \varepsilon_\Omega \end{bmatrix}$$

We assume $\varepsilon \sim (0, \boldsymbol{Q})$, where:

$$Q = \begin{bmatrix} \sigma_v & 0 & 0 \\ 0 & \sigma_a & 0 \\ 0 & 0 & \sigma_\Omega \end{bmatrix} \tag{A.33}$$

Discretized,

$$\boldsymbol{A}^s = \boldsymbol{I} + \boldsymbol{A}\tau_a + \frac{\boldsymbol{A}^2\tau_a^2}{2} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -\tau_a \\ 0 & 0 & 1 \end{bmatrix}$$

$$\boldsymbol{B}^s = \boldsymbol{B}\tau_a + \frac{\boldsymbol{A}\boldsymbol{B}\tau_a^2}{2} = \begin{bmatrix} \tau_a & 0 \\ 0 & \tau_a \\ 0 & 0 \end{bmatrix} \tag{A.34}$$

$$\boldsymbol{Q}^s = \boldsymbol{Q}\tau_a + \frac{1}{2}(\boldsymbol{A}\boldsymbol{Q} + \boldsymbol{Q}\boldsymbol{A}^T)\tau_a^2 + \frac{1}{3}\boldsymbol{A}\boldsymbol{Q}\boldsymbol{A}^T\tau_a^3 = \begin{bmatrix} \sigma_v\tau_a & 0 & 0 \\ 0 & \sigma_a\tau_a + \frac{\sigma_\Omega\tau_a^3}{3} & -\frac{\sigma_\Omega\tau_a^2}{2} \\ 0 & -\frac{\sigma_\Omega\tau_a^2}{2} & \sigma_\Omega\tau_a \end{bmatrix}$$

The formulation begins by introducing a measurement of the system states. We assume our measurement is only displacement, such that:

$$z_k = d_k^{obs} = \boldsymbol{H}^s\boldsymbol{x}_k + \eta_d \tag{A.35}$$

$$\boldsymbol{H}^s = \boldsymbol{H} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \tag{A.36}$$

$$\eta_d \sim (0, R^s) \qquad R = \sigma_d \qquad R^s = \frac{\sigma_d}{\tau_d} \tag{A.37}$$

A summary is given in Table A.5.

Table A.5: Kalman filter formulation with GPS (measurement input), broadband seismometer (deterministic input), and strong-motion accelerometer (deterministic input), accounting for accelerometer biases.

Forward Filter	
Initialize States	$\hat{\boldsymbol{x}}_0 = [0,0,0]^T$
Initialize Covariance	$\boldsymbol{P}_0 = \boldsymbol{I}$
Time Update Covariance	$\boldsymbol{P}_{k+1}^- = \boldsymbol{A}_k \boldsymbol{P}_k \boldsymbol{A}_k^T + \boldsymbol{Q}_k$
Time Update System States	$\hat{\boldsymbol{x}}_{k+1}^- = \boldsymbol{A}_k \hat{\boldsymbol{x}}_k + \boldsymbol{B}_k [v_k, a_k]^T$
Introduce Measurement	$z_k = d_k$
Compute Kalman Gain	$\boldsymbol{K}_{k+1} = \boldsymbol{P}_{k+1} \boldsymbol{H}_{k+1}^T \boldsymbol{R}_{k+1}^{-1}$
Measurement Update of Covariance	$\boldsymbol{P}_{k+1} = \left[(\boldsymbol{P}_{k+1}^-)^{-1} + \boldsymbol{H}_{k+1}^T \boldsymbol{R}_{k+1}^{-1} \boldsymbol{H}_{k+1} \right]^{-1}$
Measurement Update of System States	$\hat{\boldsymbol{x}}_{k+1} = \hat{\boldsymbol{x}}_k^- + \boldsymbol{K}_{k+1}(z_{k+1}) - \boldsymbol{H}_{k+1} \boldsymbol{x}_{k+1}^-$

A.4.3 Broadband Seismometer as Measurement Input

Assume instead, that we have a three-instrument input of strong-motion accelerations, $a(t)$, broadband seismic velocities, $v(t)$, and GPS positions, $d(t)$, such that $\tau_v = \tau_d$ and $\tau_a > \tau_d$. We assume that the strong motion accelerations are deterministic input, and broadband velocities and GPS positions are the measurements. This formulation is similar to that derived in Melgar Moctezuma (2014), but includes the additional broadband velocity measurement.

$$\frac{d}{dt}\boldsymbol{x}(t) = \boldsymbol{A}(t)\boldsymbol{x}(t) + \boldsymbol{B}(t)u(t) + \varepsilon(t) \tag{A.38}$$

where

$$x(t) = \begin{bmatrix} d(t) \\ v(t) \end{bmatrix} \quad A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{A.39}$$

$$u(t) = a(t) \qquad \varepsilon = \begin{bmatrix} 0 \\ \varepsilon_a \end{bmatrix}$$

We assume $\varepsilon \sim (0, \boldsymbol{Q})$, where:

$$Q = \begin{bmatrix} 0 & 0 \\ 0 & \sigma_a \end{bmatrix} \tag{A.40}$$

Discretized,

$$\boldsymbol{A}^s = \boldsymbol{I} + \boldsymbol{A}\tau_a + \frac{\boldsymbol{A}^2\tau_a^2}{2} = \begin{bmatrix} 1 & \tau_a \\ 0 & 1 \end{bmatrix}$$

$$\boldsymbol{B}^s = \boldsymbol{B}\tau_a + \frac{\boldsymbol{A}\boldsymbol{B}\tau_a^2}{2} = \begin{bmatrix} \frac{\tau_a^2}{2} \\ \tau_a \end{bmatrix} \tag{A.41}$$

$$\boldsymbol{Q^s} = \boldsymbol{Q}\tau_a + \frac{1}{2}(\boldsymbol{AQ} + \boldsymbol{QA}^T)\tau_a^2 + \frac{1}{3}\boldsymbol{AQA}^T\tau_a^3 = \begin{bmatrix} \frac{\sigma_a\tau_a^3}{3} & \frac{\sigma_a\tau_a^2}{2} \\ \frac{\sigma_a\tau_a^2}{2} & \sigma_a\tau_a \end{bmatrix}$$

The formulation begins by introducing a measurement of the system states. We assume our measurement is only displacement, such that:

$$z_k = \begin{bmatrix} d_k^{obs} \\ v_k^{obs} \end{bmatrix} = \boldsymbol{H^s}\boldsymbol{x}_k + \eta \tag{A.42}$$

$$\boldsymbol{H^s} = \boldsymbol{H} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{A.43}$$

$$\eta \sim (0, \boldsymbol{R}^s) \qquad \boldsymbol{R} = \begin{bmatrix} \sigma_d & 0 \\ 0 & \sigma_v \end{bmatrix} \qquad \boldsymbol{R}^s = \frac{\boldsymbol{R}}{\tau} = \begin{bmatrix} \frac{\sigma_d}{\tau_d} & 0 \\ 0 & \frac{\sigma_v}{\tau_v} \end{bmatrix} \tag{A.44}$$

A summary is given in Table A.6.

Table A.6: Kalman filter formulation with GPS (measurement input), broadband seismometer (measurement input), and strong-motion accelerometer (deterministic input).

Forward Filter	
Initialize States	$\hat{\boldsymbol{x}}_0 = [0,0]^T$
Initialize Covariance	$\boldsymbol{P}_0 = \boldsymbol{I}$
Time Update Covariance	$\boldsymbol{P}_{k+1}^- = \boldsymbol{A}_k\boldsymbol{P}_k\boldsymbol{A}_k^T + \boldsymbol{Q}_k$
Time Update System States	$\hat{\boldsymbol{x}}_{k+1}^- = \boldsymbol{A}_k\hat{\boldsymbol{x}}_k + \boldsymbol{B}_k a_k$
Introduce Measurement	$z_k = [d_k, v_k]^T$
Compute Kalman Gain	$\boldsymbol{K}_{k+1} = \boldsymbol{P}_{k+1}\boldsymbol{H}_{k+1}^T\boldsymbol{R}_{k+1}^{-1}$
Measurement Update of Covariance	$\boldsymbol{P}_{k+1} = \left[(\boldsymbol{P}_{k+1}^-)^{-1} + \boldsymbol{H}_{k+1}^T\boldsymbol{R}_{k+1}^{-1}\boldsymbol{H}_{k+1}\right]^{-1}$
Measurement Update of System States	$\hat{\boldsymbol{x}}_{k+1} = \hat{\boldsymbol{x}}_k^- + \boldsymbol{K}_{k+1}(z_{k+1}) - \boldsymbol{H}_{k+1}\boldsymbol{x}_{k+1}^-$

A.4.4 Broadband Seismometer as Measurement Input with Accelerometer Biases

Assume once again, that we have a three-instrument input of strong-motion accelerations, $a(t)$, broadband seismic velocities, $v(t)$, and GPS positions, $d(t)$, such that $\tau_v = \tau_d$ and $\tau_a > \tau_d$. We assume that the strong motion accelerations are deterministic input, and broadband velocities and GPS positions are the measurements. We account for accelerometer biases, allowing real-time implementation, rather than post-processing alone. Once again, this formulation is similar to that derived in Melgar Moctezuma (2014), but includes the additional broadband velocity measurement.

$$\frac{d}{dt}\boldsymbol{x}(t) = \boldsymbol{A}(t)\boldsymbol{x}(t) + \boldsymbol{B}(t)u(t) + \varepsilon(t) \tag{A.45}$$

where

$$x(t) = \begin{bmatrix} d(t) \\ v(t) \\ \Omega(t) \end{bmatrix} \quad A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \tag{A.46}$$

$$u(t) = a(t) \qquad \varepsilon = \begin{bmatrix} 0 \\ \varepsilon_a \\ \varepsilon_\Omega \end{bmatrix}$$

We assume $\varepsilon \sim (0, \boldsymbol{Q})$, where:

$$Q = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \sigma_a & 0 \\ 0 & 0 & \sigma_\Omega \end{bmatrix} \tag{A.47}$$

Discretized,

$$\boldsymbol{A}^s = \boldsymbol{I} + \boldsymbol{A}\tau_a + \frac{\boldsymbol{A}^2\tau_a^2}{2} = \begin{bmatrix} 1 & \tau_a & -\frac{\tau_a^2}{2} \\ 0 & 1 & -\tau_a \\ 0 & 0 & 1 \end{bmatrix}$$

$$\boldsymbol{B}^s = \boldsymbol{B}\tau_a + \frac{\boldsymbol{A}\boldsymbol{B}\tau_a^2}{2} = \begin{bmatrix} \frac{\tau_a^2}{2} \\ \tau_a \\ 0 \end{bmatrix} \tag{A.48}$$

$$\boldsymbol{Q}^s = \boldsymbol{Q}\tau_a + \frac{1}{2}(\boldsymbol{A}\boldsymbol{Q} + \boldsymbol{Q}\boldsymbol{A}^T)\tau_a^2 + \frac{1}{3}\boldsymbol{A}\boldsymbol{Q}\boldsymbol{A}^T\tau_a^3 = \begin{bmatrix} \frac{\sigma_a\tau_a^3}{3} & \frac{\sigma_a\tau_a^2}{2} & 0 \\ \frac{\sigma_a\tau_a^2}{2} & \sigma_a\tau_a + \frac{\sigma_\Omega\tau_a^3}{3} & -\frac{\sigma_\Omega\tau_a^2}{2} \\ 0 & -\frac{\sigma_\Omega\tau_a^2}{2} & \sigma_\Omega\tau_a \end{bmatrix}$$

The formulation begins by introducing a measurement of the system states. We assume our measurement is displacement and broadband velocities, such that:

$$z_k = \begin{bmatrix} d_k^{obs} \\ v_k^{obs} \end{bmatrix} = \boldsymbol{H}^s\boldsymbol{x}_k + \eta \tag{A.49}$$

$$\boldsymbol{H}^s = \boldsymbol{H} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \tag{A.50}$$

$$\eta \sim (0, \boldsymbol{R}^s) \qquad \boldsymbol{R} = \begin{bmatrix} \sigma_d & 0 \\ 0 & \sigma_v \end{bmatrix} \qquad \boldsymbol{R}^s = \begin{bmatrix} \frac{\sigma_d}{\tau_d} & 0 \\ 0 & \frac{\sigma_v}{\tau_v} \end{bmatrix} \tag{A.51}$$

A summary is given in Table A.7.

Table A.7: Kalman filter formulation with GPS (measurement input), broadband seismometer (measurement input), and strong-motion accelerometer (deterministic input), accounting for accelerometer biases.

Forward Filter	
Initialize States	$\hat{\boldsymbol{x}}_0 = [0,0,0]^T$
Initialize Covariance	$\boldsymbol{P}_0 = \boldsymbol{I}$
Time Update Covariance	$\boldsymbol{P}^-_{k+1} = \boldsymbol{A}_k \boldsymbol{P}_k \boldsymbol{A}_k^T + \boldsymbol{Q}_k$
Time Update System States	$\hat{\boldsymbol{x}}^-_{k+1} = \boldsymbol{A}_k \hat{\boldsymbol{x}}_k + \boldsymbol{B}_k a_k$
Introduce Measurement	$z_k = [d_k, v_k]^T$
Compute Kalman Gain	$\boldsymbol{K}_{k+1} = \boldsymbol{P}_{k+1} \boldsymbol{H}_{k+1}^T \boldsymbol{R}_{k+1}^{-1}$
Measurement Update of Covariance	$\boldsymbol{P}_{k+1} = \left[(\boldsymbol{P}^-_{k+1})^{-1} + \boldsymbol{H}_{k+1}^T \boldsymbol{R}_{k+1}^{-1} \boldsymbol{H}_{k+1} \right]^{-1}$
Measurement Update of System States	$\hat{\boldsymbol{x}}_{k+1} = \hat{\boldsymbol{x}}^-_k + \boldsymbol{K}_{k+1}(z_{k+1}) - \boldsymbol{H}_{k+1} \boldsymbol{x}^-_{k+1}$

A.5 Acknowledgements

This appendix contains unpublished work meant to be used as a roadmap for future implementations of the seismogeodetic Kalman filter for various instrument availabilities and requirements.

Bibliography

Agnew, D. C. and Larson, K. M. (2007). Finding the repeat times of the GPS constellation. *GPS Solutions*, 11(1):71–76.

Allen, R. (1982). Automatic phase pickers: Their present use and future prospects. *Bulletin of the Seismological Society of America*, 72(6B):S225–S242.

Allen, R. M., Gasparini, P., Kamigaichi, O., and Bose, M. (2009). The Status of Earthquake Early Warning around the World: An Introductory Overview. *Seismological Research Letters*, 80(5):682.

Ammon, C. J., Ji, C., Thio, H.-K., Robinson, D., Ni, S., Hjörleifsdóttir, V., Kanamori, H., Lay, T., Das, S., Helmberger, D., Ichinose, G., Polet, J., and Wald, D. (2005). Rupture Process of the 2004 Sumatra-Andaman Earthquake. *Science*, 308(5725):1133 LP – 1139.

Billings, S. D., Sambridge, M. S., and Kennett, B. L. N. (1994). Errors in hypocenter location: Picking, model, and magnitude dependence. *Bulletin of the Seismological Society of America*, 84(6):1978–1990.

Blaser, L., Krüger, F., Ohrnberger, M., and Scherbaum, F. (2010). Scaling Relations of Earthquake Source Parameter Estimates with Special Focus on Subduction Environment. *Bulletin of the Seismological Society of America*, 100(6):2914–2926.

Bletery, Q., Sladen, A., Delouis, B., Vallée, M., Nocquet, J.-M., Rolland, L., and Jiang, J. (2014). A detailed source model for the Mw9.0 Tohoku-Oki earthquake reconciling geodesy, seismology, and tsunami records. *Journal of Geophysical Research: Solid Earth*, 119(10):7636–7653.

Bock, Y. (1991). Continuous Monitoring of Crustal Deformation. *GPS World*, 2(6):40–47.

Bock, Y. and Melgar, D. (2016). Physical applications of GPS geodesy: a review. *Reports on Progress in Physics*, 79(10):106801.

Bock, Y., Melgar, D., and Crowell, B. W. (2011). Real-Time Strong-Motion Broadband Displacements from Collocated GPS and Accelerometers. *Bulletin of the Seismological Society of America*, 101(6):2904–2925.

Bock, Y., Prawirodirdjo, L., and Melbourne, T. I. (2004). Detection of arbitrarily large dynamic ground motions with a dense high-rate GPS network. *Geophysical Research Letters*, 31(6).

Boore, D. M. (1999). Effect of baseline corrections on response spectra for two recordings of the 1999 Chi-Chi, Taiwan, earthquake. Technical report.

Boore, D. M. and Bommer, J. J. (2005). Processing of strong-motion accelerograms: needs, options and consequences. *Soil Dynamics and Earthquake Engineering*, 25(2):93–115.

Boore, D. M., Stephens, C. D., and Joyner, W. B. (2002). Comments on Baseline Correction of Digital Strong-Motion Data: Examples from the 1999 Hector Mine, California, Earthquake. *Bulletin of the Seismological Society of America*, 92(4):1543–1560.

Boore, D. M., Stewart, J. P., Seyhan, E., and Atkinson, G. M. (2014). NGA-West2 Equations for Predicting PGA, PGV, and 5% Damped PSA for Shallow Crustal Earthquakes. *Earthquake Spectra*, 30(3):1057–1085.

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

Chao, W.-A., Wu, Y.-M., and Zhao, L. (2010). An automatic scheme for baseline correction of strong-motion records in coseismic deformation determination. *Journal of Seismology*, 14(3):495–504.

Choi, K., Bilich, A., Larson, K. M., and Axelrad, P. (2004). Modified sidereal filtering: Implications for high-rate GPS positioning. *Geophysical Research Letters*, 31(22).

Clayton, R. W., Heaton, T. H., Chandy, M., Krause, A., Kohler, M. D., Bunn, J., Guy, R., Olson, M., Faulkner, M., Cheng, M., Strand, L., Chandy, R., Obenshain, D., Liu, A., and Aivazis, M. (2012). Community Seismic Network. *Annals of Geophysics; Vol 54, No 6 (2011)*.

Clinton, J. F., Bradford, S. C., Heaton, T. H., and Favela, J. (2006). The Observed Wander of the Natural Frequencies in a Structure. *Bulletin of the Seismological Society of America*, 96(1):237–257.

Cochran, E. S., Lawrence, J. F., Christensen, C., and Jakka, R. S. (2009). The Quake-Catcher Network: Citizen Science Expanding Seismic Horizons. *Seismological Research Letters*, 80(1):26–30.

Colombelli, S., Zollo, A., Festa, G., and Kanamori, H. (2012). Early magnitude and potential damage zone estimates for the great Mw 9 Tohoku-Oki earthquake. *Geophysical Research Letters*, 39(22).

Colombelli, S., Zollo, A., Festa, G., and Picozzi, M. (2014). Evidence for a difference in rupture initiation between small and large earthquakes. *Nature Communications*, 5:3958.

Cowles, T., Delaney, J., Orcutt, J., and Weller, R. (2010). The ocean observatories initiative: Sustained ocean observing across a range of spatial scales. *Marine Technology Society Journal*, 44(6):54–64.

Crowell, B. W., Bock, Y., and Melgar, D. (2012). Real-time inversion of GPS data for finite fault modeling and rapid hazard assessment. *Geophysical Research Letters*, 39(9).

Crowell, B. W., Bock, Y., and Squibb, M. B. (2009). Demonstration of Earthquake Early Warning Using Total Displacement Waveforms from Real-time GPS Networks. *Seismological Research Letters*, 80(5):772–782.

Crowell, B. W., Melgar, D., Bock, Y., Haase, J. S., and Geng, J. (2013). Earthquake magnitude scaling using seismogeodetic data. *Geophysical Research Letters*, 40(23):6089–6094.

Crowell, B. W., Schmidt, D. A., Bodin, P., Vidale, J. E., Gomberg, J., Renate Hartog, J., Kress, V. C., Melbourne, T. I., Santillan, M., Minson, S. E., and Jamison, D. G. (2016). Demonstration of the Cascadia GFAST Geodetic Earthquake Early Warning System for the Nisqually, Washington, Earthquake. *Seismological Research Letters*, 87(4):930–943.

Day, S. M. (1982). Three-dimensional simulation of spontaneous rupture: The effect of nonuniform prestress. *Bulletin of the Seismological Society of America*, 72(6A):1881–1902.

Dupatel, Z., Rivera, L., Kanamori, H., Hayes, G. P., Hirshorn, B., and Weinstein, S. (2011). Real-time W phase inversion during the 2011 off the Pacific coast of Tohoku Earthquake. *Earth, Planets and Space*, 63(7):5.

Earle, P. S. and Shearer, P. M. (1994). Characterization of global seismograms using an automatic-picking algorithm. *Bulletin of the Seismological Society of America*, 84(2):366–376.

Eguchi, T., Fujinawa, Y., Fujita, E., Iwasaki, S.-I., Watabe, I., and Fujiwara, H. (1998). A real-time observation network of ocean-bottom-seismometers deployed at the Sagami trough subduction zone, central Japan. *Marine Geophysical Researches*, 20(2):73–94.

Ellsworth, W. L. and Beroza, G. C. (1995). Seismic Evidence for an Earthquake Nucleation Phase. *Science*, 268(5212):851 – 855.

Emore, G. L., Haase, J. S., Choi, K., Larson, K. M., and Yamagiwa, A. (2007). Recovering Seismic Displacements through Combined Use of 1-Hz GPS and Strong-Motion Accelerometers. *Bulletin of the Seismological Society of America*, 97(2):357–378.

Fowler, C. (2005). *The Solid Earth: An introduction to Global Geophysics*. Cambridge University Press, Cambridge, 2nd edition.

Galetzka, J., Melgar, D., Genrich, J. F., Geng, J., Owen, S., Lindsey, E. O., Xu, X., Bock, Y., Avouac, J.-P., Adhikari, L. B., Upreti, B. N., Pratt-Sitaula, B., Bhattarai, T. N., Sitaula, B. P., Moore, A., Hudnut, K. W., Szeliga, W., Normandeau, J., Fend, M., Flouzat, M., Bollinger, L., Shrestha, P., Koirala, B., Gautam, U., Bhatterai, M., Gupta, R., Kandel, T., Timsina,

C., Sapkota, S. N., Rajaure, S., and Maharjan, N. (2015). Slip pulse and resonance of the Kathmandu basin during the 2015 Gorkha earthquake, Nepal. *Science*, 349(6252):1091 LP – 1095.

Geiger, L. (1912). Probability method for the determination of earthquake epicenters from the arrival time only. *Bulletin of St. Louis University*, 8(1):56–71.

Geng, J., Bock, Y., Melgar, D., Crowell, B. W., and 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(7):2124–2142.

Genrich, J. F. and Bock, Y. (1992). Rapid resolution of crustal motion at short ranges with the global positioning system. *Journal of Geophysical Research: Solid Earth*, 97(B3):3261–3269.

Genrich, J. F. and Bock, Y. (2006). Instantaneous geodetic positioning with 10–50 Hz GPS measurements: Noise characteristics and implications for monitoring networks. *Journal of Geophysical Research: Solid Earth*, 111(B3).

Goldberg, D. E. and Bock, Y. (2017). Self-contained local broadband seismogeodetic early warning system: Detection and location. *Journal of Geophysical Research: Solid Earth*, 122(4):3197–3220.

Grapenthin, R., Johanson, I. A., and Allen, R. M. (2014). Operational real-time GPS-enhanced earthquake early warning. *Journal of Geophysical Research: Solid Earth*, 119(10):7944–7965.

Hamilton, W. C. (1964). *Statistics in Physical Science*. Ronald Press Co.

Hauksson, E., Stock, J., Hutton, K., Yang, W., Vidal-Villegas, J. A., and Kanamori, H. (2011). The 2010 Mw 7.2 El Mayor-Cucapah Earthquake Sequence, Baja California, Mexico and Southernmost California, USA: Active Seismotectonics along the Mexican Pacific Margin. *Pure and Applied Geophysics*, 168(8-9):1255–1277.

Havskov, J. and Ottemöller, L. (2010). *Routine Data Processing in Earthquake Seismology*. Springer, New York.

Hayes, G. P. (2017). The finite, kinematic rupture properties of great-sized earthquakes since 1990. *Earth and Planetary Science Letters*, 468:94–100.

Hayes, G. P., Rivera, L., and Kanamori, H. (2009). Source Inversion of the W-Phase: Real-time Implementation and Extension to Low Magnitudes. *Seismological Research Letters*, 80(5):817–822.

Heaton, T. H. (1990). Evidence for and implications of self-healing pulses of slip in earthquake rupture. *Physics of the Earth and Planetary Interiors*, 64(1):1–20.

Hoshiba, M. and 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 Wenzel, F. and Zschau, J., editors, *Early Warning for Geological Disasters*, pages 1–28. Springer, Berlin, Heidelberg.

Imano, M., Kido, M., Ohta, Y., Fukuda, T., Ochi, H., Takahashi, N., and Hino, R. (2015). Improvement in the Accuracy of Real-Time GPS/Acoustic Measurements Using a Multi-Purpose Moored Buoy System by Removal of Acoustic Multipath. In Hashimoto, M., editor, *International Symposium on Geodesy for Earthquake and Natural Hazards (GENAH)*, pages 105–114. Springer, Cham.

Iwan, W. D., Moser, M. A., and Peng, C.-Y. (1985). Some observations on strong-motion earthquake measurement using a digital accelerograph. *Bulletin of the Seismological Society of America*, 75(5):1225–1246.

Kato, T., Yukihiro, T., Kinoshita, M., Kakimoto, H., Isshiki, H., Matsuishi, M., Yokoyama, A., and Tanno, T. (2014). Real-time observation of tsunami by RTK-GPS. *Earth, Planets and Space*, 52(10):8410845.

Kohler, M. D., Cochran, E. S., Given, D., Guiwits, S., Neuhauser, D., Henson, I., Hartog, R., Bodin, P., Kress, V. C., Thompson, S., Felizardo, C., Brody, J., Bhadha, R., and Schwarz, S. (2017). Earthquake Early Warning ShakeAlert System: West Coast Wide Production Prototype. *Seismological Research Letters*, 89(1):99–107.

Kohler, M. D., Heaton, T. H., and Cheng, M.-H. (2013). The community seismic network and quake-catcher network: enabling structural health monitoring through instrumentation by community participants. volume 8692, pages 86923X–8692–8.

Kohler, M. D., Massari, A., Heaton, T. H., Kanamori, H., Hauksson, E., Guy, R., Clayton, R. W., Bunn, J., and Chandy, K. M. (2016). Downtown Los Angeles 52-Story High-Rise and Free-Field Response to an Oil Refinery Explosion. *Earthquake Spectra*, 32(3):1793–1820.

Kong, Q., Allen, R. M., Schreier, L., and Kwon, Y.-W. (2016). MyShake: A smartphone seismic network for earthquake early warning and beyond. *Science Advances*, 2(2).

Krischer, L. (2016). mtspec Python wrappers 0.3.2 [Data set]. Zenodo.

Kubota, T., Saito, T., Suzuki, W., and Hino, R. (2017). Estimation of Seismic Centroid Moment Tensor Using Ocean Bottom Pressure Gauges as Seismometers. *Geophysical Research Letters*, 44(21):10,907–910,915.

Kuyuk, H. S., Allen, R. M., Brown, H., Hellweg, M., Henson, I., and Neuhauser, D. (2013). Designing a NetworkBased Earthquake Early Warning Algorithm for California: ElarmS2. *Bulletin of the Seismological Society of America*, 104(1):162–173.

Langbein, J. and Bock, Y. (2004). High-rate real-time GPS network at Parkfield: Utility for detecting fault slip and seismic displacements. *Geophysical Research Letters*, 31(15).

Larson, K. M. (2009). GPS Seismology. *Journal of Geodesy*, 83(3-4):227–233.

Larson, K. M., Bodin, P., and Gomberg, J. (2003). Using 1-Hz GPS Data to Measure Deformations Caused by the Denali Fault Earthquake. *Science*, 300(5624):1421 LP – 1424.

Laske, G., Masters, G., Ma, Z., and Psyanos, M. (2013). Update on CRUST1.0: A 1-degree global model of Earth's crust. In *EGU General Assembly*, page 2658, Vienna, Austria.

Lee, W. H. K. and Stewart, S. W. (2981). *Principles and Applications of Microearthquake Networks*. Academic Press, New York.

Lewis, F. L., Xie, L., and Popa, D. (2008). *Optimal and Robust Estimation*. CRC Press, Boca Raton, 2nd edition.

Lienert, B. R., Berg, E., and Frazer, L. N. (1986). HYPOCENTER: An earthquake location method using centered, scaled, and adaptively damped least squares. *Bulletin of the Seismological Society of America*, 76(3):771–783.

Lomax, A., Michelini, A., and Curtis, A. (2009). Earthquake location, direct, global-search methods. In *Encyclopedia of Complexity and Systems Science*, pages 2449–2473. Springer.

Maeda, T., Obara, K., Shinohara, M., Kanazawa, T., and Uehira, K. (2015). Successive estimation of a tsunami wavefield without earthquake source data: A data assimilation approach toward real-time tsunami forecasting. *Geophysical Research Letters*, 42(19):7923–7932.

Massari, A., Kohler, M. D., Clayton, R. W., Guy, R., Heaton, T. H., Bunn, J., Chandy, K. M., and Demetri, D. (2017). Dense Building Instrumentation Application for City-Wide Structural Health Monitoring. In *16th World Conference on Earthquake Engineering (16WCEE)*, Santiago, Chile.

Meier, M.-A., Ampuero, J. P., and Heaton, T. H. (2017). The hidden simplicity of subduction megathrust earthquakes. *Science*, 357(6357):1277 LP – 1281.

Meier, M.-A., Heaton, T. H., and Clinton, J. (2016). Evidence for universal earthquake rupture initiation behavior. *Geophysical Research Letters*, 43(15):7991–7996.

Melgar, D. and Bock, Y. (2013). Near-field tsunami models with rapid earthquake source inversions from land- and ocean-based observations: The potential for forecast and warning. *Journal of Geophysical Research: Solid Earth*, 118(11):5939–5955.

Melgar, D. and Bock, Y. (2015). Kinematic earthquake source inversion and tsunami runup prediction with regional geophysical data. *Journal of Geophysical Research: Solid Earth*, 120(5):3324–3349.

Melgar, D., Bock, Y., and Crowell, B. W. (2012). Real-time centroid moment tensor determination for large earthquakes from local and regional displacement records. *Geophysical Journal International*, 188(2):703–718.

Melgar, D., Crowell, B. W., Bock, Y., and Haase, J. S. (2013). Rapid modeling of the 2011 Mw 9.0 Tohoku-oki earthquake with seismogeodesy. *Geophysical Research Letters*, 40(12):2963–2968.

Melgar, D., Crowell, B. W., Geng, J., Allen, R. M., Bock, Y., Riquelme, S., Hill, E. M., Protti, M., and Ganas, A. (2015a). Earthquake magnitude calculation without saturation from the scaling of peak ground displacement. *Geophysical Research Letters*, 42(13):5197–5205.

Melgar, D., Geng, J., Crowell, B. W., Haase, J. S., Bock, Y., Hammond, W. C., and Allen, R. M. (2015b). Seismogeodesy of the 2014 Mw6.1 Napa earthquake, California: Rapid response and modeling of fast rupture on a dipping strike-slip fault. *Journal of Geophysical Research: Solid Earth*, 120(7):5013–5033.

Melgar, D. and Hayes, G. P. (2017). Systematic Observations of the Slip Pulse Properties of Large Earthquake Ruptures. *Geophysical Research Letters*, 44(19):9691–9698.

Melgar Moctezuma, D. (2014). *Seismogeodesy and Rapid Earthquake and Tsunami Source Assessment*. PhD thesis, Universty of California, San Diego.

Merrifield, M. A., Firing, Y. L., Aarup, T., Agricole, W., Brundrit, G., Chang-Seng, D., Farre, R., Kilonsky, B., Knight, W., Kong, L., Magori, C., Manurung, P., McCreery, C., Mitchell, W., Pillay, S., Schindele, F., Shillington, F., Testut, L., Wijeratne, E. M. S., Caldwell, P., Jardin, J., Nakahara, S., Porter, F.-Y., and Turetsky, N. (2005). Tide gauge observations of the Indian Ocean tsunami, December 26, 2004. *Geophysical Research Letters*, 32(9).

Nakamura, Y. (1988). On the urgent earthquake detection and alarm system (UrEDAS). In *Proc. of the 9th World Conference on Earthquake Engineering*, volume 7, pages 673–678. Tokyo-Kyoto Japan.

Nikolaidis, R. M., Bock, Y., Jonge, P. J., Shearer, P. M., Agnew, D. C., and Van Domselaar, M. (2001). Seismic wave observations with the Global Positioning System. *Journal of Geophysical Research: Solid Earth*, 106(B10):21897–21916.

Ohta, Y., Kobayashi, T., Tsushima, H., Miura, S., Hino, R., Takasu, T., Fujimoto, H., Iinuma, T., Tachibana, K., Demachi, T., Sato, T., Ohzono, M., and Umino, N. (2012). Quasi real-time fault model estimation for near-field tsunami forecasting based on RTK-GPS analysis: Application to the 2011 Tohoku-Oki earthquake (Mw 9.0). *Journal of Geophysical Research: Solid Earth*, 117(B2).

Olson, E. L. and Allen, R. M. (2005). The deterministic nature of earthquake rupture. *Nature*, 438:212.

Ozaki, T. (2011). Outline of the 2011 off the Pacific coast of Tohoku Earthquake (Mw 9.0). *Earth, Planets and Space*, 63(7):57.

Ozawa, S., Nishimura, T., Suito, H., Kobayashi, T., Tobita, M., and Imakiire, T. (2011). Coseismic and postseismic slip of the 2011 magnitude-9 Tohoku-Oki earthquake. *Nature*, 475:373.

Patel, S. K., Desai, A. N., and Patel, V. B. (2011). Effect of Number of Storeys To Natural Time Period of Building. In *National Conference on Recent Trends in Engineering & Technology*.

PérezCampos, X., Melgar, D., Singh, S. K., CruzAtienza, V., Iglesias, A., and Hjörleifsdóttir, V. (2013). Rapid Estimation of Fault Parameters for Tsunami Warning along the Mexican Subduction Zone: A Scenario Earthquake in the Guerrero Seismic Gap. *Seismological Research Letters*, 84(3):474–484.

Pesyna Jr, K. M., Heath Jr, R. W., and Humphreys, T. E. (2014). Centimeter positioning with a smartphone-quality GNSS antenna. In *Radionavigation Laboratory Conference Proceedings*, Tampa, FL.

Petersson, N. and Sjögreen, B. (2017a). Installing SW4, version 2.0. Technical Report LLNL-SM-741310. Technical report, Lawrence Livermore National Laboratory.

Petersson, N. and Sjögreen, B. (2017b). User's guide to SW4, version 2.0. Technical Report LLNL-SM-741439. Technical report, Lawrence Livermore National Laboratory.

Pollitz, F. F., Bürgmann, R., and Banerjee, P. (2011). Geodetic slip model of the 2011 M9.0 Tohoku earthquake. *Geophysical Research Letters*, 38(7).

Prieto, G. A., Lawrence, J. F., Chung, A. I., and Kohler, M. D. (2010). Impulse Response of Civil Structures from Ambient Noise Analysis. *Bulletin of the Seismological Society of America*, 100(5A):2322–2328.

Prieto, G. A., Parker, R. L., and Vernon, F. L. (2009). A Fortran 90 library for multitaper spectrum analysis. *Computers and Geosciences*, 35(8):1701–1710.

Rauch, H. E., Striebel, C. T., and Tung, F. (1965). Maximum likelihood estimates of linear dynamic systems. *AIAA Journal*, 3(8):1445–1450.

Riedel, K. S. and Sidorenko, A. (1995). Minimum bias multiple taper spectral estimation. *IEEE Transactions on Signal Processing*, 43(1):188–195.

Rydelek, P. and Horiuchi, S. (2006). Is earthquake rupture deterministic? *Nature*, 442:E5.

Saito, T. and Tsushima, H. (2016). Synthesizing ocean bottom pressure records including seismic wave and tsunami contributions: Toward realistic tests of monitoring systems. *Journal of Geophysical Research: Solid Earth*, 121(11):8175–8195.

Satake, K., Shimazaki, K., Tsuji, Y., and Ueda, K. (1996). Time and size of a giant earthquake in Cascadia inferred from Japanese tsunami records of January 1700. *Nature*, 379:246.

Saunders, J. K., Goldberg, D. E., Haase, J. S., Bock, Y., Offield, D. G., Melgar, D., Restrepo, J., Fleischman, R. B., Nema, A., Geng, J., Walls, C., Mann, D., and Mattioli, G. S. (2016). Seismogeodesy Using GPS and LowCost MEMS Accelerometers: Perspectives for Earthquake Early Warning and Rapid Response. *Bulletin of the Seismological Society of America*, 106(6):2469–2489.

Shearer, P. M. (1997). Improving local earthquake locations using the L1 norm and waveform cross correlation: Application to the Whittier Narrows, California, aftershock sequence. *Journal of Geophysical Research: Solid Earth*, 102(B4):8269–8283.

Smyth, A. and Wu, M. (2007). Multi-rate Kalman filtering for the data fusion of displacement and acceleration response measurements in dynamic system monitoring. *Mechanical Systems and Signal Processing*, 21(2):706–723.

Spiess, F. N., Chadwell, C., Hildebrand, J. A., Young, L. E., Purcell, G. H., and Dragert, H. (1998). Precise GPS/Acoustic positioning of seafloor reference points for tectonic studies. *Physics of the Earth and Planetary Interiors*, 108(2):101–112.

Suzuki, W., Aoi, S., Sekiguchi, H., and Kunugi, T. (2011). Rupture process of the 2011 Tohoku-Oki mega-thrust earthquake (M9.0) inverted from strong-motion data. *Geophysical Research Letters*, 38(7).

Takahashi, N., Ishihara, Y., Fukuda, T., Ochi, H., Tahara, J., Mori, T., Deguchi, M., Kido, M., Ohta, Y., Hino, R., Mutoh, K., Hashimoto, G., Motohashi, O., and Kaneda, Y. (2015). Buoy Platform Development for Observation of Tsunami and Crustal Deformation BT - International Symposium on Geodesy for Earthquake and Natural Hazards (GENAH). pages 97–103, Cham. Springer International Publishing.

Tatehata, H. (1997). The New Tsunami Warning System of the Japan Meteorological Agency BT - Perspectives on Tsunami Hazard Reduction: Observations, Theory and Planning. pages 175–188. Springer Netherlands, Dordrecht.

Trifunac, M. D., Ivanović, S. S., and Todorovska, M. I. (2001). Apparent Periods of a Building. II: Time-Frequency Analysis. *Journal of Structural Engineering*, 127(5):527–537.

Trifunac, M. D. and Todorovska, M. I. (2001). A note on the useable dynamic range of accelerographs recording translation. *Soil Dynamics and Earthquake Engineering*, 21(4):275–286.

Trnkoczy, A. (2012). *Parameter setting of STA/LTA trigger algorithm*. Potsdam: Deutsches GeoForschungsZentrum GFZ, is 8.1 edition.

Tunnicliffe, V., Barnes, C. R., and Dewey, R. (2008). Major advances in cabled ocean observatories (VENUS and NEPTUNE Canada) in coastal and deep sea settings. In *2008 IEEE/OES US/EU-Baltic International Symposium*, pages 1–7.

USGS-FAQs (2015). Why do so many earthquakes occur at a depth of 10km?

Wald, L. A., Hutton, L. K., and Given, D. D. (1995). The Southern California Network Bulletin: 1990-1993 Summary. *Seismological Research Letters*, 66(1):9–19.

Wessel, P., Smith, W. H. F., Scharroo, R., Luis, J., and Wobbe, F. (2013). Generic Mapping Tools: Improved Version Released. *Eos, Transactions American Geophysical Union*, 94(45):409–410.

Wu, Y.-M. and Kanamori, H. (2005). Experiment on an Onsite Early Warning Method for the Taiwan Early Warning System. *Bulletin of the Seismological Society of America*, 95(1):347–353.

Wu, Y.-M. and Wu, C.-F. (2007). Approximate recovery of coseismic deformation from Taiwan strong-motion records. *Journal of Seismology*, 11(2):159–170.

Wu, Y.-M. and Zhao, L. (2006). Magnitude estimation using the first three seconds P-wave amplitude in earthquake early warning. *Geophysical Research Letters*, 33(16).

Yokota, Y., Ishikawa, T., Watanabe, S., Tashiro, T., and Asada, A. (2016). Seafloor geodetic constraints on interplate coupling of the Nankai Trough megathrust zone. *Nature*, 534:374.

Yoshioka, S. and Matsuoka, Y. (2013). Interplate coupling along the Nankai Trough, southwest Japan, inferred from inversion analyses of GPS data: Effects of subducting plate geometry and spacing of hypothetical ocean-bottom GPS stations. *Tectonophysics*, 600:165–174.

Yun, N. Y. and Hamada, M. (2014). Evacuation Behavior and Fatality Rate during the 2011 Tohoku-Oki Earthquake and Tsunami. *Earthquake Spectra*, 31(3):1237–1265.

Zheng, G. and Rice, J. R. (1998). Conditions under which velocity-weakening friction allows a self-healing versus a cracklike mode of rupture. *Bulletin of the Seismological Society of America*, 88(6):1466–1483.

Zollo, A., Lancieri, M., and Nielsen, S. (2006). Earthquake magnitude estimation from peak amplitudes of very early seismic signals on strong motion records. *Geophysical Research Letters*, 33(23).

Zumberge, J. F., Heflin, M. B., Jefferson, D. C., Watkins, M. M., and Webb, F. H. (1997). Precise point positioning for the efficient and robust analysis of GPS data from large networks. *Journal of Geophysical Research: Solid Earth*, 102(B3):5005–5017.