Processed ground-motion records from induced earthquakes for use in engineering applications

Karen Assatourians[1] and Gail M. Atkinson[1]

[1]Dept. Earth Sciences, Western University, London, Canada N6A 5B7

Submitted to Can. J. Civil Eng., April 2018; Revised Dec. 12, 2018; Accepted March 2019

Corresponding author: Karen Assatourians

email addresses: (karenassatourians@yahoo.com, gmatkinson@aol.com)

word count equivalent: 9000 words

Abstract

We compile and process an electronic database of ground motions recorded on accelerometers and broadband seismographic instruments for induced earthquakes of **M**≥4 at distances <50 km in central and eastern North America. Most of the data are from Oklahoma, with some records from Alberta. Our focus is on the subset of available records that are of most interest for engineering analyses aimed at evaluation of the potential hazards from induced events, which is a pressing issue in western Canada and other regions experiencing induced seismicity. We considered all records to 50 km for events of **M**≥4.5. For events of **M**4 to 4.5, we select records at close distance (<10 km), having good signal strength (PGA >~3%g), in order to allow high-quality time histories to be obtained. These records have strong signal-to-noise ratio, making them suitable for engineering applications, such as dynamic analysis, after proper scaling. The selected records are windowed, filtered and instrument-corrected to compile a set of records having acceptable acceleration, velocity and displacement time histories. The records and their response spectra are provided as an electronic supplement at http://www.seismotoolbox.ca/IS_Strong_Motions/ . We note that the record set is not suitable as a response spectra database for development of ground-motion prediction equations, because for **M**<4.5 the record selection is biased to records with higher amplitudes. Rather, the intended use of the records is as seed records, which can be readily scaled in the time domain to approximately represent induced-event target scenarios for engineering applications.

Key words: earthquake time histories, induced earthquakes, dynamic analysis, critical infrastructure

Introduction

The rate of earthquake activity in central and eastern North America (CENA) has risen sharply due to induced events related to oil and gas development (Ellsworth 2013; Atkinson et al. 2015a; Atkinson et al. 2016), causing a notable increase in seismic hazard in many areas (Atkinson et al. 2015b; Petersen et al. 2015). There is a need for engineering analyses to evaluate the impact of induced events on structural response, which in turn leads to a requirement for suitable earthquake time histories for such analyses. This need is particularly pressing for the evaluation of critical infrastructure in regions such as western Alberta and eastern B.C., which are experiencing significant rates of induced seismic activity (Atkinson et al. 2016; Atkinson 2017). Suitable records are those of **M**~4 to 6 (where **M** is moment magnitude) recorded at close distances (<10 km from the hypocenter) because these are the scenarios that dominate the hazard (e.g. Bourne et al. 2015; Atkinson et al. 2015b), and which are most likely to produce strongly-felt motions of potential concern. Such records are scarce because the requirement for close distances greatly limits the number of available recordings. It is therefore necessary to consider a somewhat wider range of distances in the search for records, and to accept that significant scaling may be required to bring the amplitudes of the records to the levels expected at very close distances. Records from a number of similar regions should also be considered, to obtain the widest-possible database. Here we assemble and process a ground-motion database for induced-seismicity applications, considering events of **M**≥4 at distances <50 km.

The considered events have a shallow focus and occurred in areas where most of the seismicity is believed to be induced. However, we do not attempt to address whether every selected event was induced. As noted by Yenier and Atkinson (2015), there are no obvious differences in ground motions between natural and induced events in CENA for events of the same magnitude and focal depth. Therefore, in the development of a ground-motion database for induced events, it is not of primary importance to identify whether every event can be definitively classified as induced.

The purpose of the database compiled here is to provide spectral amplitudes and processed time histories for those records of most relevance in assessing the engineering response of structures to induced events. The database comprises publicly-available broadband and accelerometer records from events in Oklahoma and Alberta from 2010 to 2016 at distances <10 km for **M** 4 to **M** 4.5, and at distances to 50 km for **M**≥4.5. Selected records from large events (**M**>4.5) at close distances will be added to the database in future as they become available. It should be noted that this database, comprised of a few hundred records, is not intended to be a spectral-amplitude database for GMPE development for induced events, because we have focused our study on the available records having the strongest shaking – and the dataset is thus potentially-biased. For broader GMPE development purposes, a larger compilation project is required, involving thousands of records over a broad distance range, so that the effects of source, path and site can be separated empirically. This larger database compilation is in progress under a separate study. There is also a Next-Generation-Attenuation (NGA) project for Induced Seismicity underway at Pacific Engineering Research Center (PEER) that will compile a suitable database for GMPE development. While these larger projects are being conducted, we compile a targeted database of the records of most interest for engineering and provide their processed time histories. The focus on a subset of records allows for record-by-record inspection and processing, which may not be feasible for very large databases that need to rely on fully- automatic processing. We visually inspect every record to ensure that the resulting time histories are reasonably well-behaved in acceleration, velocity and displacement.

Database and Processing

There have been a significant number of recorded induced events of **M**≥4 in the last few years, with most of the records coming from Oklahoma or western Canada. The events in Oklahoma are primarily induced by wastewater disposal (Ellsworth 2013), whereas those in western Canada are often induced by hydraulic fracturing (Atkinson et al. 2015a; 2016). There is no compelling evidence to date that triggered earthquakes, whether related to disposal or hydraulic fracturing, are fundamentally different from shallow natural earthquakes. In terms of ground motions that are generated near the source, low-frequency ground motions are controlled by the seismic moment, which defines the moment magnitude. High-frequency ground motions are characterized by the stress parameter. The stress parameter scales with magnitude and focal depth and may vary regionally, but there appears to be no discernible difference based on whether an event is natural or induced (Yenier and Atkinson 2015). Shallow earthquakes, whether natural or induced, have lower stress parameters than deeper events, on average. This would tend to lower the spectral amplitudes at high frequencies. On the other hand, shallow events can be experienced at very close distances, due to the short distance from the hypocentre to the surface, and this tends to lead to increased spectral amplitudes. These factors offset each other, resulting in similar spectral amplitudes for natural and induced events. In terms of signal duration, this will scale with magnitude and distance. Induced events that are of small-to-moderate magnitude (e.g. **M**<5) will be relatively short in duration compared to large regional earthquakes. The maximum size of induced events is not known, but is likely similar to the maximum size for tectonic events (van der Elst et al. 2016).

The scaling trends of both natural and induced events were evaluated by Atkinson and Assatourians (2017) and used to identify candidate GMPEs that reflect current knowledge of ground-motion trends for induced events in CENA. That study suggests that we could potentially select either natural or induced events (or some combination) as proxies for induced events in development of a time history database. However, since we seek records of moderate events at very close distances, and therefore very shallow depth, induced-earthquake records are a natural choice, and are the most plentiful. We thus restrict our record selection to recent earthquakes in regions where induced seismicity is occurring frequently, and is being recorded on high-quality seismic networks. These regions are Oklahoma and western Canada. Most of the near-distance records from western Canada are not publicly available, and thus the focus of the time history database is on Oklahoma records. The stations recording the events are high-quality broadband seismographic stations or strong-motion instruments recording three-component waveforms sampled at 40, 100 and 200 samples/sec. We prefer records at higher sampling rates (100 or 200 samples/sec), but have considered 40 samples/sec records for the largest two events (**M**>5.5) when these are the only records available.

Ground-motion records are from a database compiled from publicly-available broadband and accelerometer recordings for events recorded in Oklahoma and Alberta, processed as described by Assatourians and Atkinson (2010). Most of the data were downloaded directly from IRIS (Incorporated Research Institutes for Seismology). To select records for the study, we examine earthquake catalogues and station lists from western Canada and Oklahoma to identify events of **M**≥4 that have one or more records at hypocentral distances of R_{hypo}<20 km, focusing on the closest available records. For the largest events (**M**≥4.5), we consider records to distances of 50 km because there are few very close records. The largest events considered are the 2011 Prague and 2016 Pawnee, Oklahoma events of **M**~5.6 to 5.8 (where 5.8 is the moment magnitude for Pawnee is that given by the U.S. Geological Survey; the Oklahoma Geological Survey lists its magnitude as **M**=5.6). **Figure 1** shows the geographical distribution of earthquakes and stations in and around Oklahoma from which most of data in this study are drawn.

Figure 2 shows the distribution of the initial database in magnitude, distance and horizontal-component peak ground acceleration (PGA). The records of most interest are those with significant spectral amplitudes at intermediate-to-high frequencies, as indicated in the figure by the 5%-damped, horizontal-component pseudo-acceleration spectral amplitude (PSA) at 5 Hz. We focus our more detailed time history development on records for **M**≥4.5 at distances to 50 km, plus records of **M**4 to 4.5, at hypocentral distances < 11 km, and having 5-Hz PSA >90 cm/s^2 (on at least one horizontal component). We emphasize that for GMPE development, we would be interested in an unbiased dataset (not just the stronger records) and we would therefore not make such a selection for GMPE development. In this study, by contrast, we wish to obtain records of good signal quality that will result in well-behaved records in the time domain, and we therefore focus on the subset of stronger signals.

All records shown in **Figure 2** (at left) are processed to obtain peak motions and response spectra. A selected subset of stronger records (at right) is chosen for more detailed analysis and processing. The selected records are those having strong signal; this includes all records of **M**≥4.5 to 50 km, plus records of **M**4-4.5 at Rhypo<11 km with 5-Hz PSA >90 cm/s^2 on at least one horizontal component. Moreover, we exclude any records of poor quality as based on visual inspection of the signals (Note: poor records are those with a quality index of 6, where the quality index is described in the Appendix.) Table 1 lists the events for which more detailed processing was performed, and from which the time history database is developed. The moment magnitude of each event (**M**) was determined as described by Novakovic and Atkinson (2015) for Western Canada earthquake ($\mathbf{M}_{NA15}$), or taken from Oklahoma Geological Survey (OK) and Advanced National Seismic System (ANSS) catalogues for Oklahoma earthquakes.

For each of the records shown in **Figure 2**, we download 3-component ground motions from IRIS for all stations within 20 km or within 50 km for $\mathbf{M}\geq4.5$. For each record, a window of 160 seconds duration is obtained around the signal containing 20 seconds of pre-event noise, the strong part of the signal (i.e. the signal window), and the significant coda. The records are processed using an updated version of the ICORRECT program developed by Assatourians and Atkinson (2010). This program was developed to follow the general processing guidelines for earthquake records as discussed by Boore and Bommer (2005), in a way that can be implemented efficiently in an automatic format that will be valid for both broadband and accelerometer records. The Appendix provides a summary of the processing procedures, and the steps taken to refine them for the selected subset of records, in order to ensure high-quality time series in acceleration, velocity and displacement, free from trends or other artifacts.

For each record we provide the instrument-corrected ground acceleration, velocity and displacement. The peak values of the processed records for each component are also calculated. Caution is required for records associated with quality 5 (minor clipping), as the peak motion values may not be reliable; however the response spectra are insensitive to such minor clipping. The response spectrum is calculated from the accelerograms using the algorithm of Nigam and Jennings (1969). We also calculate the Arias intensity (I_A):

$$I_A = \frac{\pi}{2g}\int_0^T a^2(t)dt$$

where *a(t)* is the acceleration time history in units of m/s^2, *g* is the acceleration of gravity in units of m/s^2, and *T* represents the complete duration of recording. A common measure of significant duration is the time interval between 5% and 75% of I_A, denoted $D_{5\text{-}75}$ (more details are provided in Kempton and Stewart 2006). The value of $D_{5\text{-}75}$ is also calculated.

Cumulative absolute velocity (CAV) is defined as the integral of the absolute value of an acceleration time series (Campbell and Bozorgnia 2010):

$$CAV = \int_0^{t_{\max}} |a(t)|dt$$

where $a(t)$ is the acceleration time series in m/s^2, t is time, and t_{max} is the total duration of the time series. Here, we calculate CAV over the whole signal duration for our 160-sec record-processing window. Users may calculate standardized CAV (Campbell and Bozorgnia 2010) if desired, after selecting a set of time series from the database and applying proper scaling factors. We didn't calculate standardized CAV for processed accelerograms because their values for scaled records will not be related to our CAV measure through simple scaling.

Some attributes of the processed database

As a guide to the amplitude levels and distance scaling of the ground motions, in **Figure 3** we compare the recorded horizontal-component (geomean) ground motions from events of **M** 4.0 to 4.5 in Oklahoma and the few available Alberta records to selected ground-motion prediction equations (GMPEs), at close distances . The plotted GMPEs are those identified by Atkinson and Assatourians (2017) as being appropriate for induced events in CENA. In order to produce an unbiased figure of amplitudes at close distances, all of the available data in the magnitude-distance range are included (i.e. all data on the left side of **Figure 1**). The site conditions of the recording stations are not yet classified in available databases. To make an approximate correction to the B/C reference condition of the GMPEs for **Figure 3**, it is assumed that all records are on NEHRP (National Earthquake Hazards Reduction Program) site class C, with V_{s30}=450 m/s (where V_{s30} is the time-averaged shear-wave velocity over the top 30 m). The site correction factors of Seyhan and Stewart (2014), assuming linear site response, are used to make a first-order correction from C to B/C; the Seyhan and Stewart (2014) site corrections were also used by Atkinson (2015) and Yenier and Atkinson (2015) to correct observations to B/C before developing their GMPEs. The assumption of Class C for Oklahoma stations is likely a reasonable average when taken over the database, but is not intended to represent a realistic site correction for any individual record. The use of an average site correction factor will map into increased variability of the ground-motion amplitudes. More detailed site corrections will require compilation of information on site conditions, and/or empirical regressions to determine site terms.

From **Figure 3** we conclude that, despite the larger scatter in the data, the observations at distances <15 km are generally consistent with the GMPEs, especially when one considers that few (or none) of these data were used in the GMPE derivations, and that the conversions of observations to B/C conditions were not site-specific. A noteworthy observation is that despite general consistency of observed amplitudes with GMPEs, the decay of amplitudes in the first 20 km appears to be quite steep, especially at high frequencies. The slope is steeper than the trend of $R_{hypo}^{-1.7}$ that applies to the Atkinson (2015; A15) GMPE (at 5 Hz), and much steeper than the decay of $R_{hypo}^{-1.0}$ that is often assumed in ground-motion modeling. We emphasize the steepness of the distance scaling by plotting a line of slope 1/R for reference, at an arbitrary amplitude level, on **Figure 3**. The steep amplitude scaling with distance is apparent only at small-to-moderate magnitudes because for large magnitudes this effect is counteracted by an increasing near-distance saturation effect (e.g. Yenier and Atkinson 2014). The steep amplitude decay is an important factor in scaling of records for induced-seismicity hazard studies. This steep decay is common to the selected GMPEs shown on **Figure 3**. We caution that many GMPEs assume a distance scaling of 1/R, or otherwise invoke a more pronounced near-distance saturation that may result in underestimation of expected near-distance ground motions, if the GMPE was developed from regression of data at distances beyond 10 km. Note that values of PGA in the range from 10% to 40%g are not unusual for events of **M** 4 to 4.5 at distances within 10 km of the hypocenter. The large variability of amplitudes is also noteworthy, although some of the variability comes from the range of unknown site conditions.

Organization of the time history database

The processed time histories are provided at the following URL location:
http://www.seismotoolbox.ca/IS_Strong_Motions/

We also provide an overall index table of response spectra for all selected records, along with their key attributes, to aid in selection of records for further evaluation against a set of desired criteria. The individual time history files are compressed in a number of zip files, along with associated PSAs for those records. Grouping a number of time histories and associated PSAs in zip files is done for easier access and download of the required files. The time history and PSA files have identical headers for the same record. Each PSA file contains two columns: frequency and corresponding PSA values. The body of time history files have seven columns: time, instrument-corrected time history values of acceleration, velocity, displacement, raw data count (uncorrected), Husid, and cumulative absolute values. The filenames follow a logical convention: each file name specifies the corresponding event date/time, recording station name, channel, and location code. For example "2010.10.11.13.33.40.ARK1.EHE.--.tra" carries time histories of an event on 2010/10/11 at 13:33:40 recorded at station ARK1 on component EHE. Note that the records are unscaled. As described in the next section, they may require scaling to be suitable for specific purposes.

Some Suggestions for Scaling of Time Histories

It can be concluded from **Figure 3** that the expected amplitudes of ground motion will depend strongly on hypocentral distance. It should also be recognized that amplitudes will vary with magnitude and with site condition. Moreover, some records will be stronger-than-average, while others will be weaker. For these reasons, a common practice in using time histories for evaluation of the response of structures is to first scale them to approximately match a target spectrum having the desired amplitudes and spectral content. In general, the target spectrum can be defined based on a probabilistic seismic hazard analysis, using either the uniform hazard spectrum or the conditional mean spectrum (e.g. Baker 2011). This is the approach usually taken for site-specific analysis considering natural seismicity. The same approach can be used to define the target for induced-seismicity applications, but with some modifications. For induced seismicity, the source zone to consider for the hazard includes specific oil and gas operations, or a collection of such operations. The likelihood of induced events needs to be assessed, and considered within the context of the assigned magnitude recurrence parameters for the analysis. Whilst conceptually straightforward, the probabilistic assessment of the target spectrum is fraught with difficulty due to the very large uncertainties affecting the key rate parameters and the processes that control them. Examples of use of the probabilistic approach to assign a target spectrum are provided by Atkinson et al. (2015b) and Atkinson (2017).

Alternatively, the target spectrum is sometimes based on a postulated scenario of interest. This approach is particularly applicable to induced-seismicity applications, in which we are especially interested in the effects of events at a close distance – and in which this distance may be known because it is tied to specific operations. The use of a scenario is also sometimes used for preliminary evaluations due to its conceptual simplicity and transparency. Atkinson (2017) used a combination of probabilistic analysis and the scenario approach to argue that hydraulic fracture operations should be kept a minimum distance of 5 km from critical infrastructure that might be vulnerable to strong ground motions from moderate events, such as older major dams built with minimal seismic resistance. In that context, the target spectrum considered was based on a scenario event of **M**4.5 at R_{hypo}= 5km, which may have a likelihood of the order of 1/10,000 for operations in areas prone to induced seismicity.

To illustrate the use of the time histories, we use the scenario approach to define a target spectrum for the event considered by Atkinson (2017) - an earthquake of **M**4.5 at R_{hypo}= 5km. The target is defined as the median-plus-sigma spectrum predicted by the three GMPEs shown on **Figure 3**, for this magnitude and distance (where sigma is the standard deviation, assumed here to be 0.3 $\log_{10}$ units; see Atkinson and Assatourians 2017). We assume that only linear amplitude scaling is to be applied. We note that it is also common practice to use spectral matching techniques to more close match a target spectrum (e.g. Hancock et al. 2006). However, it is not clear that such procedures offer any real advantage, other than to make the records appear more similar to the target (Bazzurro and Luco 2006; NIST 2012). Therefore, we restrict our focus in this demonstration to linear amplitude scaling procedures.

To find the records that best match the target spectra with only simple linear amplitude scaling being applied, we seek those records that have suitable spectral shape, without any dramatic site-response peaks in the spectra that cause them to deviate significantly from the target shapes. The actual site conditions are not known at most sites, so these shape checks are our best tool to select records with appropriate site characteristics.

To aid in identifying the most suitable records we determine, for each record, the mean value of $\log_{10}$(PSA(targ)/PSA(obs)), along with its standard deviation, over a selected frequency range, assumed here to be from 1 to 10 Hz. PSA(targ) is the target median-plus-sigma spectral values, and PSA(obs) are the corresponding values calculated from the instrument-corrected accelerograms. The metric being used in this computation is the geometric mean of the two horizontal components of ground motion.

The value of (PSA(targ)/PSA(obs)), averaged over the selected frequencies, gives the scaling factor that needs to be applied to the records so that the geomean of the horizontal components will match the target. The same scaling factor is applied to vertical component motions. The standard deviation is a measure of how closely the shape of the record matches that of the target, over the selected frequency range. We prioritize the records based on the signal quality and existence of noise, necessary scaling factor, and misfit RMS respectively, then inspect the response spectra of the observations against the target graphically to select the subset of most suitable records. Note that most records will require a significant scaling factor because they are at greater distances than the target distance.

Figure 4 shows a sample of 11 selected records based on these criteria and compares their geomean horizontal-component spectra to that of the target spectrum, after scaling. The selected records range in magnitude from **M**4.2 to **M**4.9, and in hypocentral distance from 6.4 km to 38.7 km; because the hypocentral distances are greater than the target of 5 km, significant scaling is required. **Table** 2 lists the attributes of the selected records and the applied scaling factors. The time histories of acceleration, velocity and displacement for two record sets along with their response spectra after scaling, are provided in **Figures 5** to **12**. Note that the accelerations are significant, while the displacements are small, even after scaling.

Conclusion

We have compiled an electronic database of high-quality processed ground motions from induced earthquakes of **M**≥4 at distances <50 km in central and eastern North America. The records are suitable for engineering analyses aimed at evaluation of the potential hazards from induced events, which is a pressing issue in western Canada and other regions experiencing induced seismicity. The intended use is as seed records that can be scaled in the time domain to approximately represent induced-event target scenarios. The records and their response spectra are provided as an electronic supplement at http://www.seismotoolbox.ca/IS_Strong_Motions/ .

Acknowledgements

The financial support of the Natural Sciences and Engineering Research Council is gratefully acknowledged. Partial support from the U.S. Geological Survey, under Grant G16AP00107 is also acknowledged. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the opinions or policies of the U.S. Geological Survey.

References

Abrahamson, N. A., Silva W. J., and Kamai, R. 2014. Summary of the Abrahamson, Silva, and Kamai NGA-West2 ground-motion relations for active crustal regions, Earthquake Spectra, **30**, 1025-1056.

Ancheta, T., Darragh, R., Stewart, J., Seyhan, E., Silva, W., et al. 2014. PEER NGAWest2 database, Earthquake Spectra, **30**, 989-1006.

Assatourians, K., and Atkinson, G. 2010. Database of processed time series and response spectra for Canada: An example application to study of the 2005 MN5.4 Riviere du Loup, Quebec earthquake. Seism. Res. L., **81,** 1013-1031.

Atkinson, G. 2015. Ground-motion prediction equation for small-to-moderate events at short hypocentral distances, with application to induced seismicity hazards. Bull. Seism. Soc. Am., **105**, doi: 10.1785/0120140142.

Atkinson, G. 2017. Strategies to prevent damage to critical infrastructure from induced seismicity. FACETS, Science Application Forum, doi: 10.1139/facets-2017-0013.

Atkinson, G., and Assatourians, K. 2017. Are ground-motion models derived from natural events applicable to the estimation of expected motions for induced earthquakes? Seism. Res. L., **88**(2A), 1-12.

Atkinson, G., Eaton, D., Ghofrani, H., Walker, D., Cheadle, B., et al. 2016. Hydraulic fracturing and seismicity in the Western Canada Sedimentary Basin. Seism. Res. L., **87**, doi: 10.1785/0220150263.

Atkinson, G., Assatourians, K., Cheadle, B., and Greig. W. 2015a. Ground motions from three recent earthquakes in western Alberta and northeastern British Columbia and their implications for induced-seismicity hazard in eastern regions. Seism. Res. L.,86, 1022-1031.

Atkinson, G., Ghofrani, H., and Assatourians, K. 2015b. Impact of Induced Seismicity on the Evaluation of Seismic Hazard: Some Preliminary Considerations. Seism. Res. L., 86, 1009-1021.

Baker, J.W. 2011. Conditional Mean Spectrum: Tool for ground motion selection, Journal of Structural Engineering, **137**, 322-331.

Bazzurro, P., and Luco, N. 2006. Do scaled and spectrum-matched near-source records produce biased nonlinear structural responses?. Proc. 8th U.S. Natl Conf. Earthq. Eng., San Francisco, April 2006.

Boore, D. M. 2010. Orientation-Independent, Nongeometric-Mean Measures of Seismic Intensity from Two Horizontal Components of Motion. Bull. Seism. Soc. Am. **100**, 1830-1835.

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

Boore, D., and Kishida, T. 2017. Relations between some horizontal-component ground-motion intensity meaures used in practice. Bull. Seism. Soc. Am., **107**, 334-343.

Bourne, S., Oates, S., Bommer, J., Dost, B., Van Elk, J., et al. 2015. A Monte Carlo method for probabilistic hazard assessment of induced seismicity due to conventional natural gas production. Bull. Seism. Soc. Am., **105**, 1721-1738.

Campbell, K., and Bozorgnia, Y. 2010. Analysis of cumulative absolute velocity (CAV) and JMA instrumental seismic intensity (I_{JMA}) using the PEER-NGA strong motion database. PEER report 2010/102. www.peer.berkeley.edu.

Ellsworth, W. 2013. Injection-induced earthquakes. Science, 341 (6142), 12 July 2013. DOI:10.1126/science.1225942.

Hancock, J., Watson-Lamprey, J., Abrahamson, N., Bommer, J., Markatis, A., et al. 2006. An improved method of matching response spectra of recorded earthquake ground motion using wavelets. J. Earthq. Eng., **10**, Special Issue 1, 67-89.

Kempton, J., and Stewart, J. 2006. Prediction equations for significant duration of earthquake ground motions considering site and near-source effects. Earthquake Spectra, **22**, 985-1013.

Nigam, N. C., and Jennings, P. C. 1969. Calculation of response spectra from strong-motion earthquake records. Bull. Seism. Soc. Am., **59**, 909-922.

NIST 2012. Selecting and scaling earthquake ground motions for performing response-history analyses. NEHRP Consultants Joint Venture (A partnership of the Applied Technology Council and the Consortium of Universities for Research in Earthquake Engineering). NIST GCR 11-917-15. National Institute of Standards and Technology, U.S. Dept. Commerce, Wa. 256 pp.

Novakovic, M., and Atkinson, G. 2015. Preliminary evaluation of ground motions from earthquakes in Alberta. Seism. Res. L., **86**,doi 10.1785/0220150059.

Petersen, M., Mueller, C., Moschetti, M., Hoover, S., Rubinstein, J., et al. 2015. Incorporating induced seismicity in the 2014 United States National Seismic Hazard Model - Results of 2014 workshop and sensitivity studies. U.S. Geol. Surv. Open-file Rpt. 2015-1070.

Seyhan, E. and Stewart, J. 2014. Semi-empirical nonlinear site amplification from NGA-West2 data and simulations. Earthq. Spectra, **30**, 1241-1256.

Van der Elst, N., Page, M., Weiser, D., Goebel, T., and Hosseini, S. 2016. Induced earthquake magnitudes are as large as (statistically) expected. J. Geophys. Res. Solid Earth, 121, 4575-4590.

Yenier, E., and Atkinson, G. 2014. Point-source modeling of moderate-to-large magnitude earthquakes and associated ground-motion saturation effects. Bull. Seism. Soc. Am., **104**, 1458-1478.

Yenier, E., and Atkinson, G. 2015. A regionally-adjustable generic GMPE based on stochastic point-source simulations. Bull. Seism. Soc. Am., **105**, 1989-2009.

Figures

Figure 1 – Geographical distribution of earthquakes and stations (and network names) in and around Oklahoma. These earthquakes and records are the main source of data used in this study.

Figure 2 – Selection of the processed database in magnitude, distance and amplitude; the two horizontal components are plotted. Left: PSA at 5 Hz for available records of ***M****4-4.5 to 20 km, and* ***M****≥4.5 to 50 km. Right: the subset selected for more detailed processing to produce engineering time histories. In the selected subset, poor records and weak signals have been removed (see text for details).*

Figure 3 – Observed horizontal-component ground motions (symbols) for induced events of ***M****4.0 to 4.5 (converted to B/C) in Oklahoma (OK) and Alberta (AB), compared to Atkinson (2015) (A15 alternative-h model), Yenier and Atkinson (2015) (YA15 CENA; assumed depth=4km) and Abrahamson et al. (2014) (ASK14, unspecified depth) GMPEs (lines). ASK14 and YA15 are plotted versus rupture distance; A15 and observations are plotted versus* R_{hypo}*. Heavy dashed line at bottom of each panel shows 1/R trendline, plotted at an arbitrary amplitude level. Note that only the stronger records in this magnitude range are included in the time history database – see Figure 2 for details.*

*Figure 4 – Selected records scaled to the target median-plus-sigma spectrum for an event of **M**=4.5 at Rhypo=5km.*

*Figure 5 – Scaled accelerograms of **M**4.4 event at 8.5km distance (event number 17) for matching the **M**4.5 scenario event at 5 km for each of the three components.*

*Figure 6 – Scaled velocity time series of **M**4.4 event at 8.5km distance (event number 17) for matching the **M**4.5 scenario event at 5 km for each of the three components.*

*Figure 7 – Scaled displacement time series of **M**4.4 event at 8.5km distance (event number 17) for matching the **M**4.5 scenario event at 5 km for each of the three components.*

*Figure 8 – Response spectra of **M**4.4 event at 8.5km distance (event number 17) in acceleration, velocity and displacement, after scaling the records to the **M**4.5 scenario event at 5 km. HN1 and HN2 are the two orthogonal horizontal components; HNZ is the vertical component.*

*Figure 9 – Scaled accelerograms of **M**4.9 event at 18.4km distance (event number 9) for matching the **M**4.5 scenario event at 5 km for each of the three components.*

*Figure 10 – Scaled velocity time series of **M**4.9 event at 18.4km distance (event number 9) for matching the **M**4.5 scenario event at 5 km for each of the three components.*

*Figure 11- Scaled displacement time series of **M**4.9 event at 18.4km distance (event number 9) for matching the **M**4.5 scenario event at 5 km for each of the three components.*

*Figure 12 – Response spectra of **M**4.9 event at 18.4km distance (event number 9) in acceleration, velocity and displacement, after scaling the records to the **M**4.5 scenario event at 5 km. HN1 and HN2 are the two orthogonal horizontal components; HNZ is the vertical component.*

Tables

Table 1 – List of events for which records were analyzed

No	Date	Time [UTC]	Latitude [∘]	Longitude [∘]	Depth [km]	$\mathbf{M}^*$	# of records	Source
1	2011/02/28	05:00:51.51	35.269	-92.355	3.2	4.7[2]	6	USGS
2	2011/11/06	03:53:11.11	35.522	-96.780	3.1	5.7	3	OK
3	2011/11/08	02:46:58.58	35.518	-96.786	2.5	4.8	3	OK
4	2013/12/07	18:10:24.24	35.607	-97.385	8.4	4.5	9	OK
5	2014/06/16	10:47:35.35	35.592	-97.399	5.0	4.3	6	OK
6	2014/06/18	10:53:02.2	35.593	-97.396	5.0	4.1	6	OK
7	2014/08/19	12:41:35.35	35.773	-97.468	4.9	4.4	3	OK
8	2014/10/02	18:01:24.24	37.245	-97.955	5.0	4.3	15	USGS
9	2014/11/12	21:40:00.0	37.271	-97.621	4.0	4.9	30	USGS
10	2015/04/08	20:52:00.0	35.818	-97.420	2.5	4.3	3	OK
11	2015/07/20	20:19:03.3	36.843	-98.257	4.1	4.4	6	OK
12	2015/07/27	18:12:15.15	35.989	-97.572	5.0	4.5	3	OK
13	2015/09/18	12:35:17.17	35.987	-96.795	0.2	4.1	6	OK
14	2015/09/25	01:16:37.37	35.987	-96.787	2.9	4.0	12	OK
15	2015/10/10	22:03:05.5	35.986	-96.803	3.3	4.4	15	OK
16	2015/11/19	07:42:12.12	36.661	-98.458	5.9	4.7	6	OK
17	2015/11/23	21:17:46.46	36.838	-98.275	5.0	4.4	6	OK
18	2015/11/30	09:49:13.13	36.761	-98.056	2.3	4.7	12	OK
19	2015/12/29	11:39:19.19	35.665	-97.405	6.5	4.3	3	OK
20	2016/01/01	11:39:39.39	35.669	-97.406	5.8	4.2	3	OK
21	2016/01/12	18:27:23.23	54.411	-117.290	5.0	4.2[1]	6	NA15
22	2016/02/13	17:07:07.7	36.483	-98.735	3.2	5.1	6	OK
23	2016/07/08	21:31:58.58	36.477	-98.739	7.3	4.2	24	OK
24	2016/07/09	02:04:27.27	36.465	-98.756	7.2	4.4	9	OK
25	2016/09/03	12:02:44.44	36.426	-96.929	5.6	5.8	3	USGS

* 1 is $\mathbf{M}_{NA15}$, calculated using the Novakovic and Atkinson (2015) magnitude formulation (M. Novakovic, 2016, pers. comm.); 2 is duration magnitude; all others are moment magnitude as given by Oklahoma Geological Survey (OK) or the U.S. Geological Survey (USGS). Note: negative longitudes used for western hemisphere.

*Table 2 – Characteristics of the Selected and Scaled Records: **M**4.5 at Rhypo= 5km. (Rhypo in km; PGA, PGV, and PGD in cgs units; AI and CAV in m/s, and Duration (Dur) in seconds).*

Event #	M	Sta	Comp	R_{hypo} [km]	PGA [cm/s^2]	PGV [cm/s]	PGD [cm]	AI [m/s]	CAV [m/s]	Dur [sec.]	Factor
4	4.5	OK001	HNE	13.1	1.65E+02	6.02E+00	5.43E-01	2.21E-01	4.55E+00	3.78	8.419
4	4.5	OK001	HNN	13.1	4.91E+02	8.69E+00	1.05E+00	6.45E-01	5.43E+00	1.055	8.419
4	4.5	OK001	HNZ	13.1	9.74E+01	1.70E+00	1.27E-01	8.65E-02	2.71E+00	3.155	8.419
8	4.3	KAN12	HHE	8.6	2.31E+02	5.13E+00	2.97E-01	3.77E-01	6.13E+00	2.62	4.487
8	4.3	KAN12	HHN	8.6	3.40E+02	8.26E+00	8.50E-01	7.21E-01	7.19E+00	2.285	4.487
8	4.3	KAN12	HHZ	8.6	3.50E+02	6.81E+00	3.26E-01	4.44E-01	4.72E+00	1.825	4.487
9	4.9	KAN01	HNE	18.4	2.31E+02	7.86E+00	9.51E-01	2.65E-01	4.56E+00	3.015	3.100
9	4.9	KAN01	HNN	18.4	2.90E+02	1.07E+01	8.25E-01	3.48E-01	4.98E+00	2.535	3.100
9	4.9	KAN01	HNZ	18.4	1.01E+02	2.66E+00	3.17E-01	7.41E-02	2.55E+00	3.935	3.100
11	4.4	OK032	HN1	7.3	2.33E+02	1.05E+01	7.75E-01	1.67E-01	2.24E+00	0.42	1.067
11	4.4	OK032	HN2	7.3	1.85E+02	5.91E+00	2.72E-01	1.27E-01	2.25E+00	0.67	1.067
11	4.4	OK032	HNZ	7.3	1.19E+02	1.86E+00	1.27E-01	4.45E-02	1.24E+00	1.23	1.067
16	4.7	OK032	HH1	27.9	2.68E+02	4.41E+00	2.86E-01	6.41E-01	1.07E+01	11.84	5.440
16	4.7	OK032	HH2	27.9	6.79E+02	9.55E+00	5.58E-01	9.66E-01	1.13E+01	6.68	5.440
16	4.7	OK032	HHZ	27.9	1.07E+02	1.28E+00	9.85E-02	1.73E-01	5.80E+00	15.21	5.440
17	4.4	OK032	HN1	8.5	3.69E+02	9.92E+00	7.43E-01	4.13E-01	4.28E+00	1.35	2.751
17	4.4	OK032	HN2	8.5	2.40E+02	4.78E+00	3.60E-01	2.29E-01	3.75E+00	1.49	2.751
17	4.4	OK032	HNZ	8.5	1.37E+02	1.76E+00	2.03E-01	6.56E-02	2.03E+00	1.935	2.751
18	4.7	STN19	HH1	16.4	3.40E+02	5.68E+00	8.59E-01	3.96E-01	4.08E+00	1.96	6.554
18	4.7	STN19	HH2	16.4	2.27E+02	5.22E+00	5.77E-01	2.26E-01	3.53E+00	1.92	6.554
18	4.7	STN19	HHZ	16.4	2.21E+02	4.48E+00	5.27E-01	1.96E-01	3.11E+00	1.52	6.554
18	4.7	STN20	HH1	38.7	4.62E+02	8.99E+00	9.54E-01	3.31E-01	4.60E+00	2.56	12.229
18	4.7	STN20	HH2	38.7	2.68E+02	5.48E+00	6.90E-01	1.74E-01	3.90E+00	4.24	12.229

18	4.7	STN20	HHZ	38.7	2.03E+02	3.04E+00	5.42E-01	1.01E-01	2.84E+00	2.37	12.229
21	4.2	WSK01	HH1	9.4	3.10E+02	7.28E+00	4.51E-01	1.54E-01	2.21E+00	0.45	9.257
21	4.2	WSK01	HH2	9.4	2.23E+02	9.69E+00	6.51E-01	1.29E-01	2.05E+00	0.82	9.257
21	4.2	WSK01	HHZ	9.4	8.02E+01	1.94E+00	1.99E-01	2.68E-02	1.13E+00	2.07	9.257
24	4.2	OK038	HN1	6.4	2.39E+02	8.32E+00	4.20E-01	2.03E-01	3.43E+00	1.83	1.499
24	4.2	OK038	HN2	6.4	2.58E+02	7.30E+00	3.22E-01	2.31E-01	3.40E+00	1.47	1.499
24	4.2	OK038	HNZ	6.4	2.91E+02	3.41E+00	6.29E-02	2.96E-01	3.63E+00	1.485	1.499
24	4.2	OK043	HN1	7.9	3.45E+02	1.35E+01	6.53E-01	3.78E-01	3.34E+00	0.47	2.460
24	4.2	OK043	HN2	7.9	3.09E+02	6.13E+00	1.92E-01	1.51E-01	2.68E+00	0.97	2.460
24	4.2	OK043	HNZ	7.9	1.95E+02	3.31E+00	1.46E-01	1.11E-01	2.09E+00	1.18	2.460

Appendix – Time Series Processing Procedures

The ICORRECT algorithm upon which the time series processing was based, and its validation, are described in Assatourians and Atkinson (2010). The processing includes deglitching, detrending, windowing and initial filtering (4^{th} order Butterworth with high-pass at 0.1 Hz), along with removal of the complex instrument response in the frequency domain. Differentiation and integration of broadband seismometer signals is carried out by multiplying and dividing the instrument-corrected velocity spectrum by iω (ω is angular frequency in radians/sec and i is $\sqrt{-1}$) in the complex frequency domain, to obtain acceleration and displacement, respectively. The processing produces an initial set of instrument-corrected acceleration, velocity and displacement records, having a useable frequency range from ~0.2 to 40 Hz. The flowchart of this processing stage is shown in **Figure A-1** (where the time histories in correct physical units are the result of the processing step shaded in grey and are input to the second (manual) stage of processing).

In a second stage of processing, applied to the selected subset of records (**Table 1**), every component is inspected and processed manually, refining the filtering as necessary to supress the effects of low-frequency noise and any other artifacts of automatic processing from appearing in displacement time series. The details of the secondary processing steps are illustrated in the flowchart of **Figure A-2**. We visually inspected the acceleration, velocity and displacement traces produced by the initial processing. If all traces were acceptable, no additional processing was performed, and a quality flag of 1 was assigned to these records (where quality=1 denotes a high-quality record over the entire frequency range from 0.2 to 40 Hz, as determined by visual inspection of all traces). If we noted that some of the velocity or raw records were clipped (due to limited dynamic range of the broadband instruments), we assessed the relative degree of clipping. The processed time series were assigned a quality flag of 5 if there is very minor clipping (e.g. clipping of a few points only, not appearing to significantly cut signal amplitudes) that would have minimal impact on the response spectra. A quality flag of 6 was assigned to strongly-clipped traces, which are judged to be useless as engineering time histories.

If the displacement records showed unrealistic pre-event artifacts such as ringings or bumps, which can result from the application of zero-phase bandpass filtering in automatic processing, then additional (zero-phase) bandpass filters with narrower pass bands were applied to the filtered data, and all motions were re-checked visually in both the time and frequency domains. Such artifacts can be problematic particularly for doubly-integrated accelerograms from relatively low magnitude earthquakes, which have weak signal at longer periods. In cases where this additional filtering repaired the artifacts to a negligible level, further processing was not performed, and time series were assigned the quality flag of 2 (good quality record, but over a more restrictive bandpass than the initial range of 0.2 to 40 Hz).

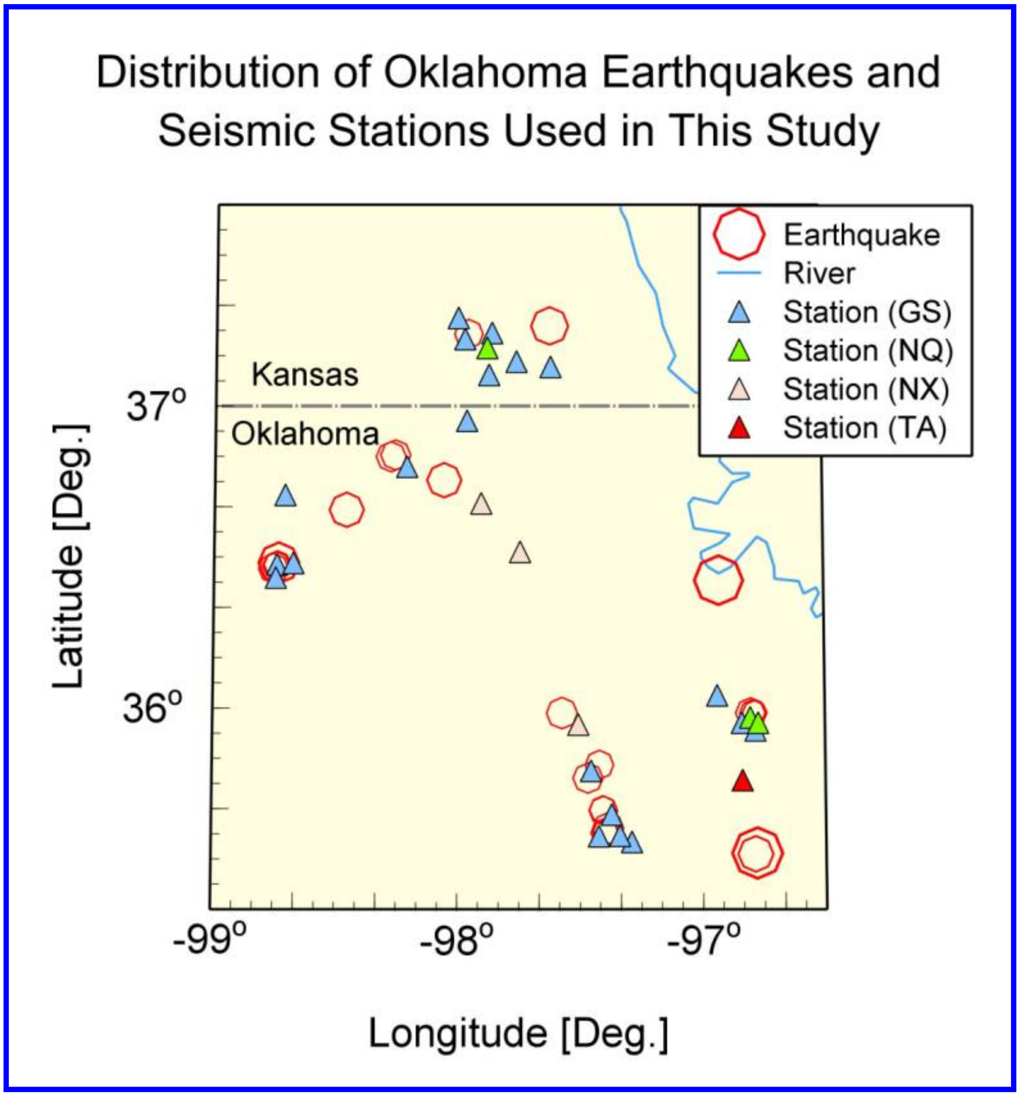
For records for which more restrictive band pass filtering did not reduce pre-event signal processing artifacts in the displacement records to negligible levels, the velocity records were windowed in three segments, which often resulted in improved signal recovery. The three segments are: 1- from the record start to the beginning of the signal window; 2- the main signal window; and 3- from the end of the main signal window to the end of record. The main signal window starts with the arrival of the P-wave and includes the entire strong part of the waveform; its selection is done by visual inspection of the velocity time series after application of bandpass filters and amplitude scaling, to allow for better visualization of the most significant part of the record. In such cases, the first segment (pre-event) was high-pass filtered with a corner frequency of 10Hz while the other two segments were baseline and trend corrected. Differentiation and integration for calculating acceleration and displacement time series for such records is carried out in the time domain using the entire time series. The signal portions are already filtered and detrended, so any residual trends after rejoining the segments are marginal and have little effect. Thus no additional filtering or tapering need be applied before merging sub-windows. Finally, it is confirmed visually that there are no unrealistic artifacts appearing in the Fourier or response spectra or time series of signals after going through this process. This process was generally successful in removing pre-event artifacts on displacement records, while having minimal impact on either Fourier or response spectra (except for some minor changes at very low frequencies, below the bandwidth of interest). The records processed by this multi-window processing approach are assigned a quality flag of 3.

As an additional check on our processed record amplitudes, we compared our automatically-processed response spectra for Oklahoma to records in common from a similar database compiled by the U.S. Geological Survey (Morgan Moschetti, personal communication 2017); their processing was based on the routines used by PEER (Pacific Earthquake Engineering Research Center) for the NGA (Next Generation Attenuation) projects (Ancheta et al. 2014). The values are not completely comparable because the USGS tabulated the orientation-independent horizontal measures, RotD50 and RotD100 (Boore 2010), whilst we have tabulated the as-recorded component measures (east component, north component). There are also other more minor differences such as the selected window for processing, and the filter parameter procedures. Nevertheless, our values of the geometric mean of the two horizontal components should be similar to the USGS values of RotD50 in most cases (e.g. see Boore and Kishida 2017). This expectation was realized, over a comparison of 56 records in common to the two databases. The average difference between our computed geomeans and the USGS values of RotD50 (for records in common to the two databases) was 5% to 10% over most frequencies. This is slightly larger than the difference found by Boore and Kishida (2017) between the geomean and RotD50 for the NGA-West2 database (2% to 8%), but acceptable given the small size of the sample and the automated processing procedures employed.

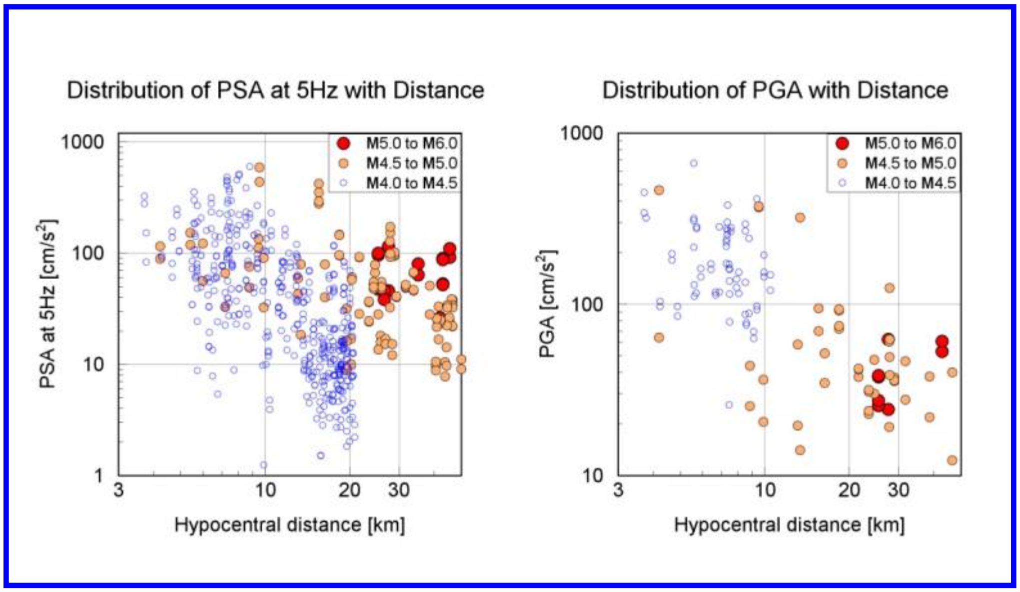
Appendix Figures

Figure A1 – Flowchart of signal processing steps in first (automatic) stage by program Qcorrect. Outputs from the step shaded in grey are time histories in correct units.

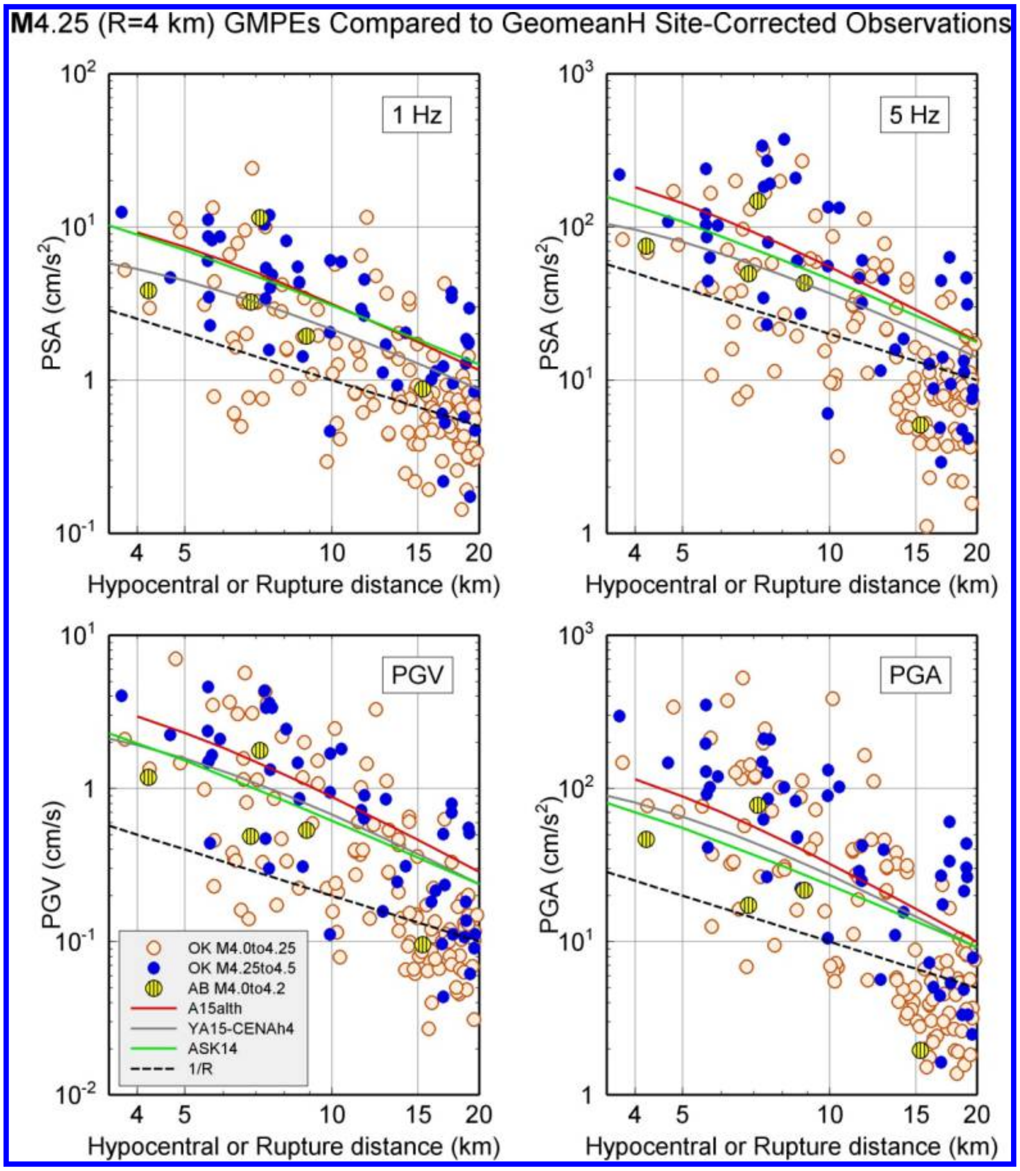
Figure A-2 – Flowchart of manual signal processing steps, illustrating assigned quality flags. The flags are an ordinal representation of the quality of final products with 1 being the highest quality and 6 being useless. Quality 1 signals don't need manual processing; Quality 2 signals are further bandpass filtered; Quality 3 signals are band pass filtered and multi-window processed; Quality 4 is reserved for weaker signals, not appearing in this study; Quality 5 signals carry minor clipping; Quality 6 signals are damaged and/or useless.



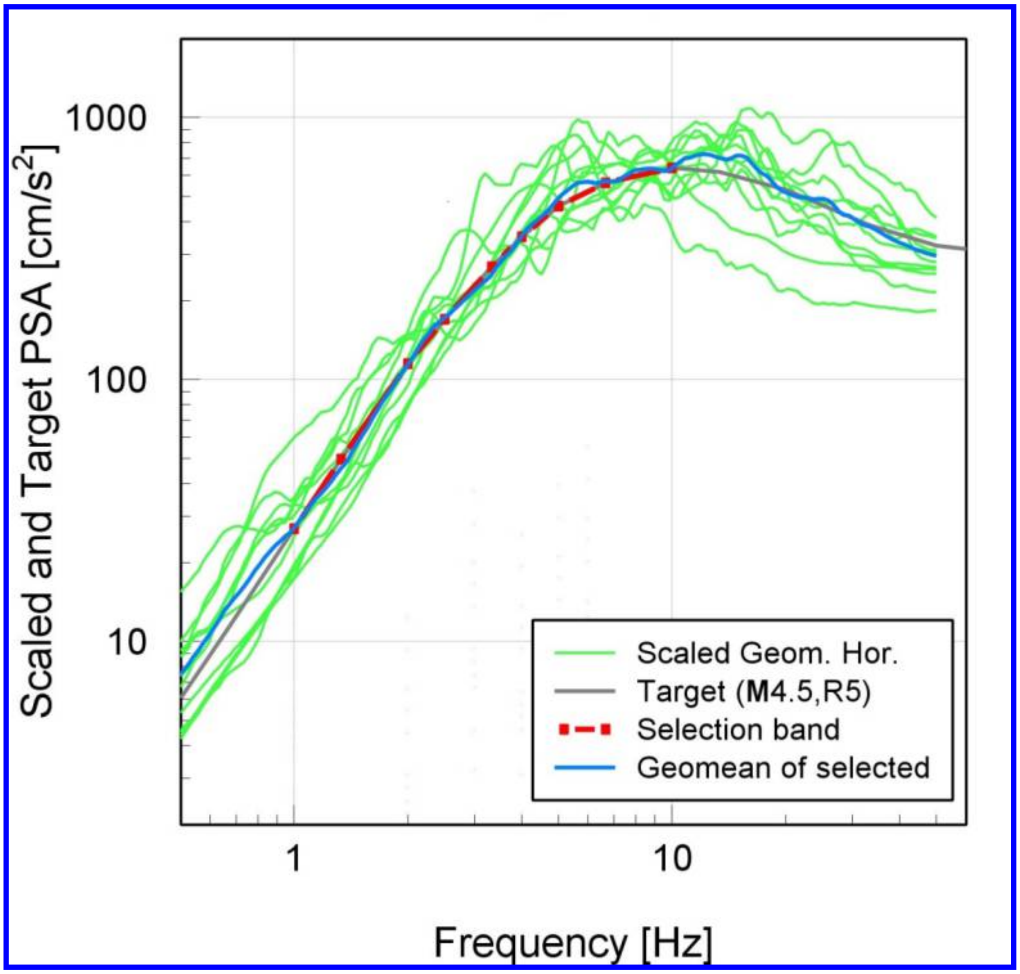
191x206mm (300 x 300 DPI)

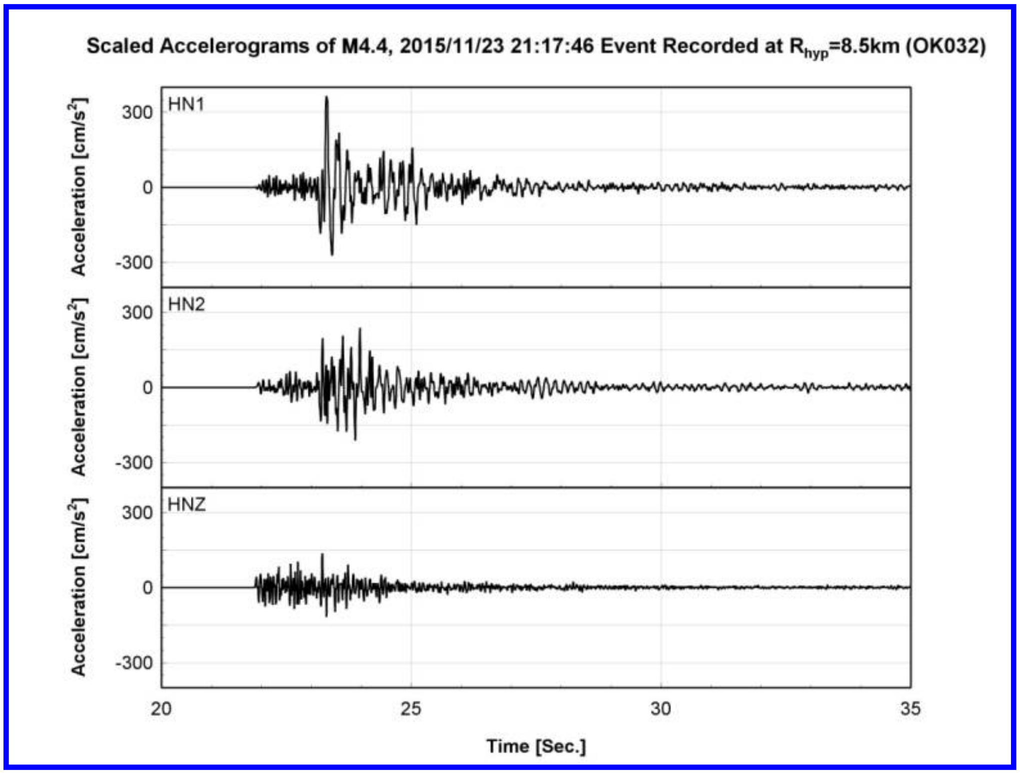


182x104mm (300 x 300 DPI)

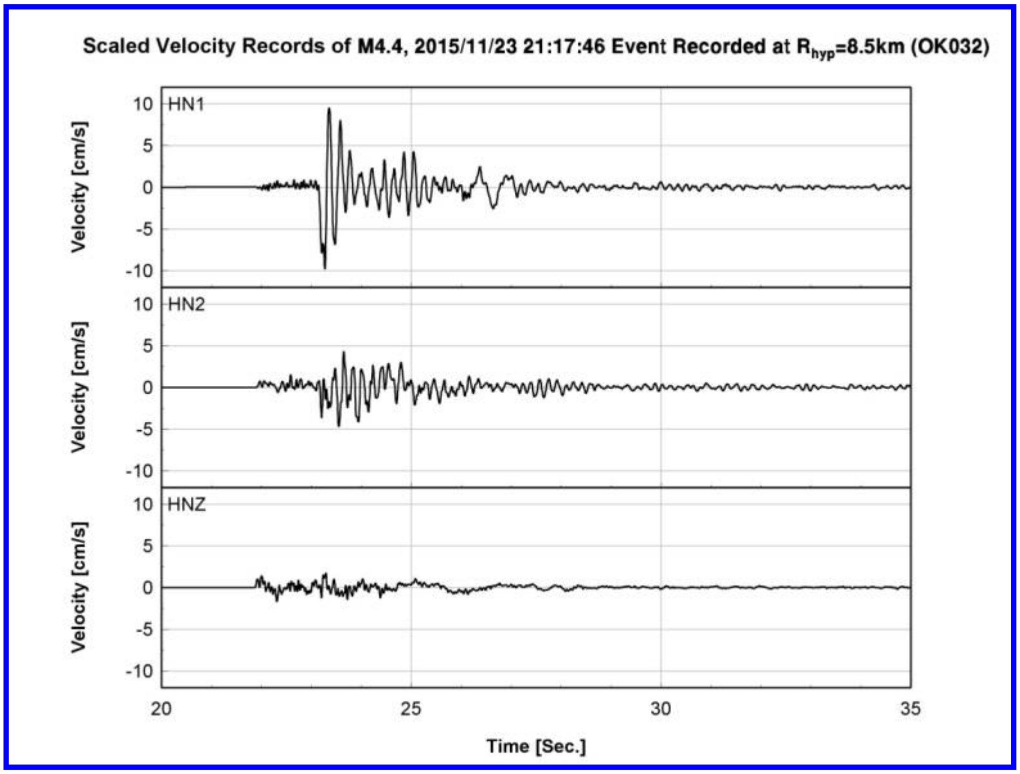


M4.25 (R=4 km) GMPEs Compared to GeomeanH Site-Corrected Observations
1 Hz
5 Hz
PGV
PGA
PSA (cm/s^2)
PGV (cm/s)
PGA (cm/s^2)
Hypocentral or Rupture distance (km)
OK M4.0to4.25
OK M4.25to4.5
AB M4.0to4.2
A15alth
YA15-CENAh4
ASK14
1/R

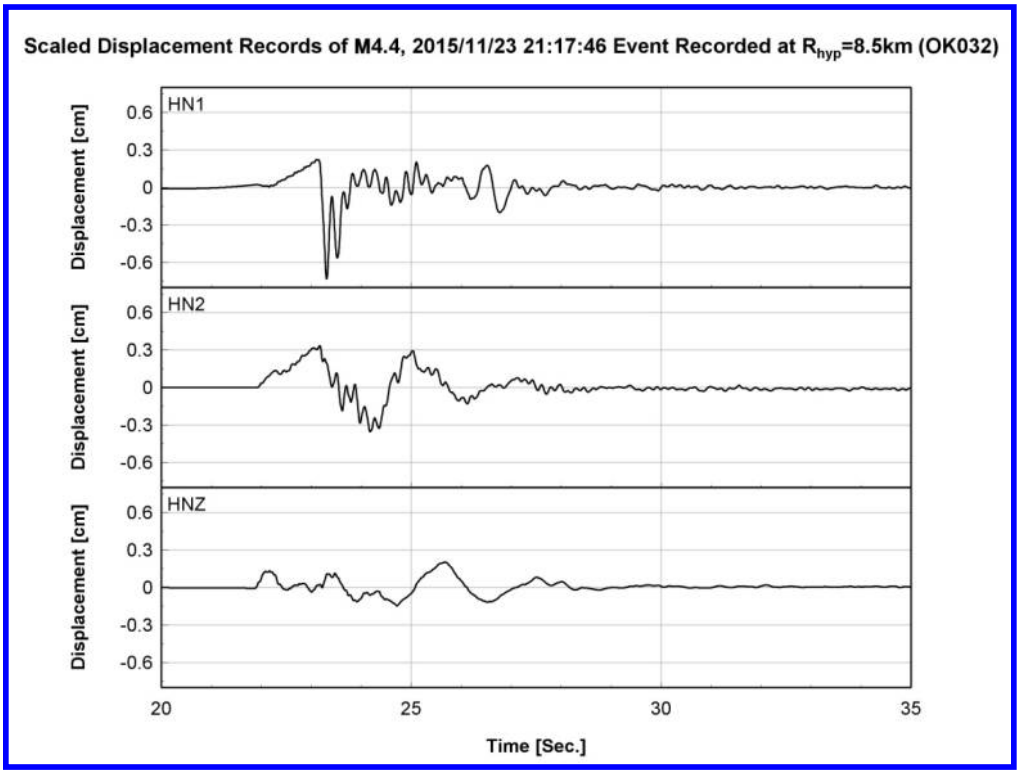

Scaled and Target PSA [cm/s²]
1000
100
10
1
10
Frequency [Hz]
Scaled Geom. Hor.
Target (M4.5,R5)
Selection band
Geomean of selected



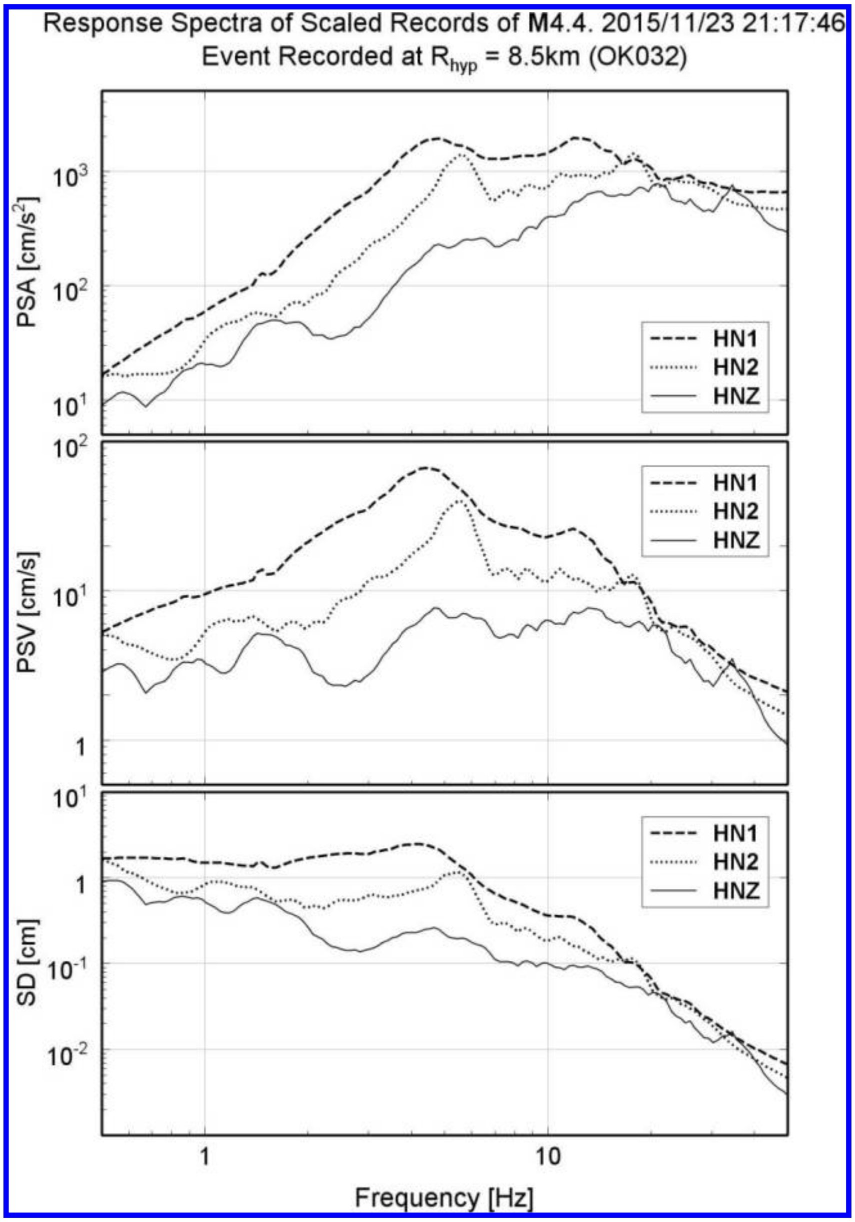
203x152mm (300 x 300 DPI)

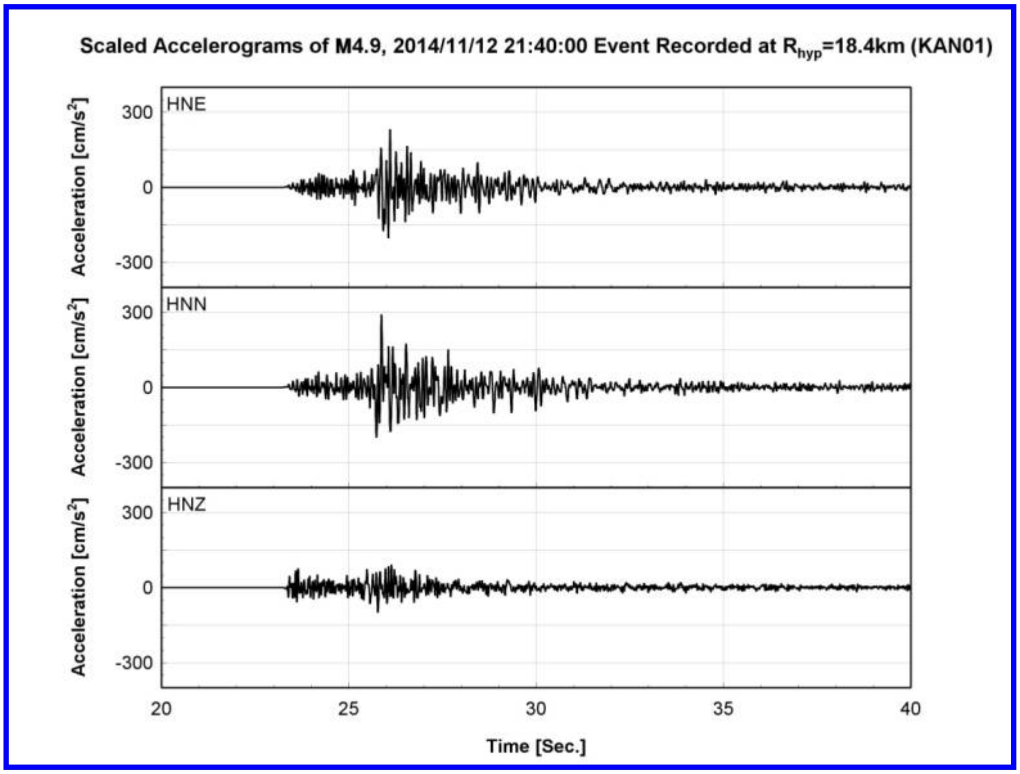


203x152mm (300 x 300 DPI)

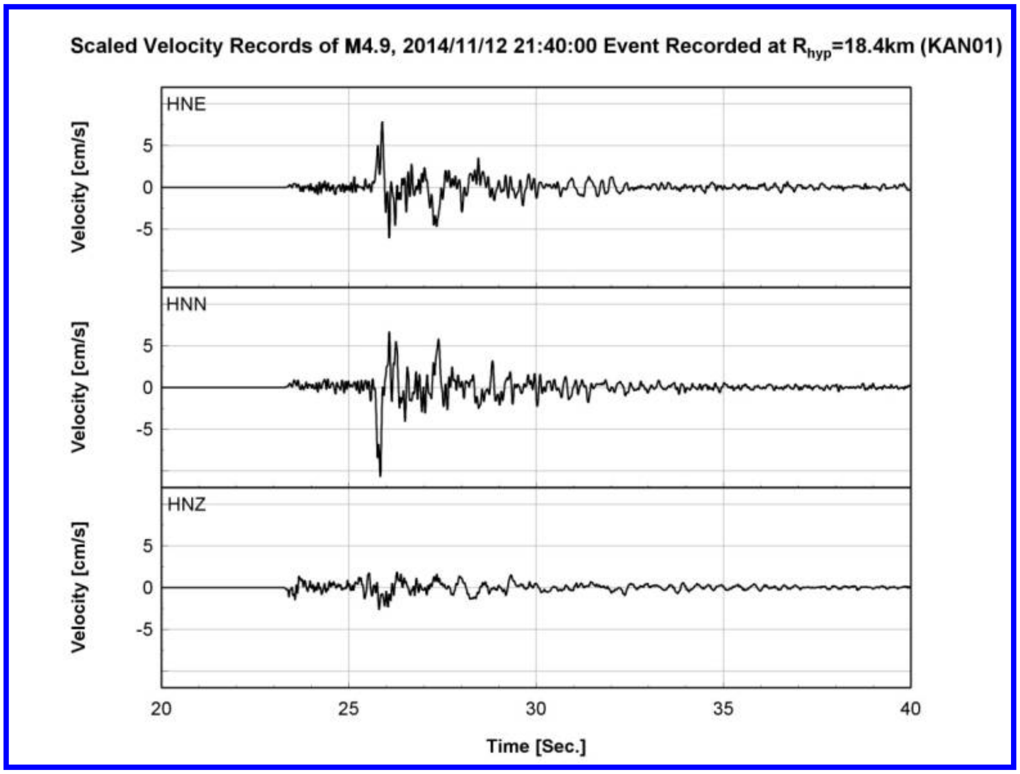


203x152mm (300 x 300 DPI)

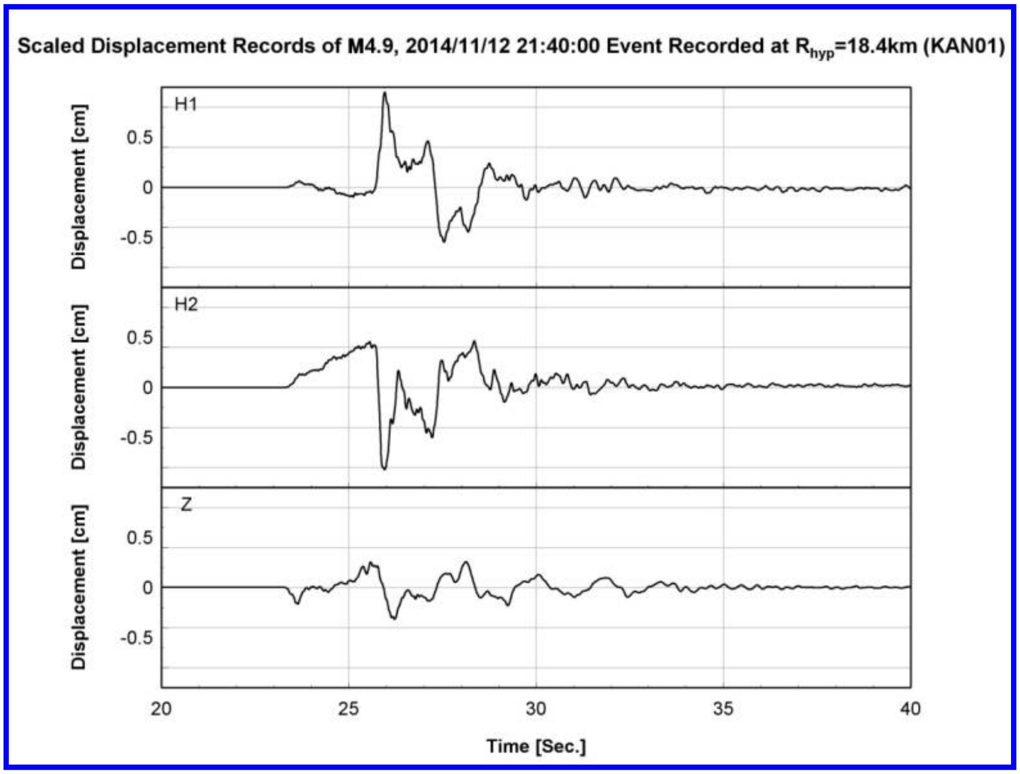

Response Spectra of Scaled Records of M4.4. 2015/11/23 21:17:46
Event Recorded at R_hyp = 8.5km (OK032)
PSA [cm/s²]
PSV [cm/s]
SD [cm]
HN1
HN2
HNZ
Frequency [Hz]



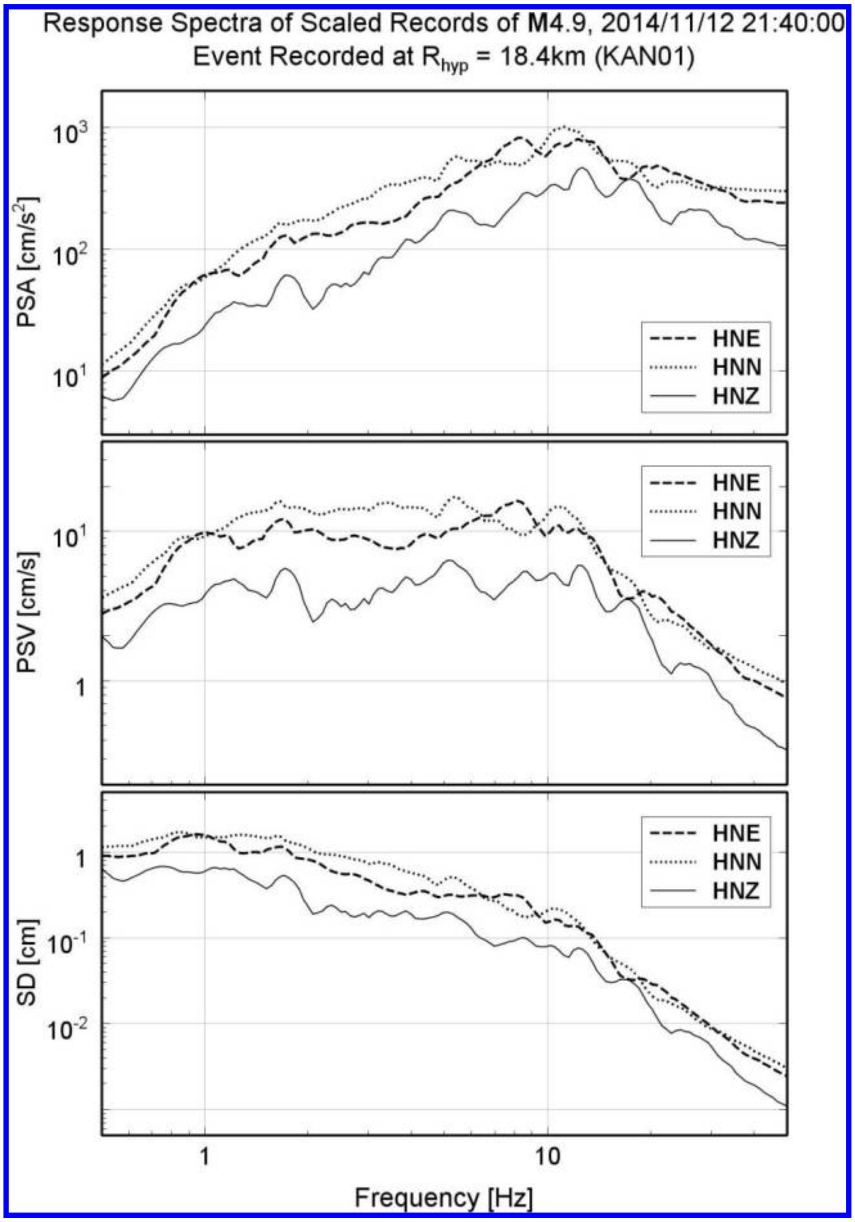
203x152mm (300 x 300 DPI)



203x152mm (300 x 300 DPI)



203x152mm (300 x 300 DPI)


Response Spectra of Scaled Records of M4.9, 2014/11/12 21:40:00
Event Recorded at R_hyp = 18.4km (KAN01)
PSA [cm/s²]
PSV [cm/s]
SD [cm]
HNE
HNN
HNZ
Frequency [Hz]