PEER NGA-West2 Database:
A Database of Ground Motions Recorded in Shallow Crustal Earthquakes in Active Tectonic Regions

T.D. Ancheta & Y. Bozorgnia
Pacific Earthquake Engineering Research Center, Univeristy of California, Berkeley, CA, U.S.A.

R. Darragh & W.J. Silva
Pacific Engineering and Analysis, El Cerrito, CA, U.S.A.

B. Chiou
California Department of Transportation, Sacramento, CA, U.S.A.

J.P. Stewart
University of California, Los Angeles, CA, U.S.A.

D.M. Boore & R. Graves
U.S. Geological Survey, Menlo Park & Pasadena, CA, U.S.A.

N.A. Abrahamson
Pacific Gas & Electric Company, San Francisco, CA, U.S.A.

K.W. Campbell
EQECAT, Inc., Beaverton, OR, U.S.A.

I.M. Idriss
University of California, Davis, CA, U.S.A.

R.R. Youngs
AMEC Geomatrix, Oakland, CA, U.S.A.

G.M. Atkinson
University of Western Ontario, London, Ont. Canada

SUMMARY

The NGA-West 2 project database expands on the current PEER NGA ground-motion database to include worldwide ground-motion data recorded from shallow crustal earthquakes in active tectonic regimes post 2003. Since 2003, numerous well-recorded events have occurred worldwide, including the 2003 M6.6 Bam (Iran), 2004 M6 Parkfield (CA), 2008 M7.9 Wenchuan (China), 2009 M6.3 L'Aquila (Italy), 2010 M7 El Mayor-Cucupah (CA and Mexico), 2010 M7 Darfield (NZ), 2011 M6.1 Christchurch (NZ), and several well-recorded shallow crustal earthquakes in Japan, among other events. The NGA database has been extensively expanded to include the recorded ground-motion data, and metadata, in these and other recent events. The updated strong-motion database (NGA-West 2) currently includes 8611 three-component records from 334 shallow crustal events. The updated database has a magnitude range of 3.4 to 7.9, and a distance range of 0.05 to 1533 km. The estimated or measured time-averaged shear-wave velocity in the top 30 meters at the recording sites (i.e., V_{s30}) ranges from 94 to 2100 m/sec. The NGA-West 2 database more than doubles the size of the previous NGA database. The database includes uniformly processed time series as well as response spectral ordinates for 111 periods ranging from 0.01 to 20 seconds and 11 different damping ratios. Extensive metadata have also been collected and added to the database. The database is currently being utilized by NGA researchers to update the 2008 ground-motion prediction equations (GMPEs).

Keywords: earthquake, strong motion, database, metadata

1. INTRODUCTION

The importance of a common high-quality ground-motion database was recognized in the NGA project. Having project investigators use a common database fostered collaboration between ground motion prediction equation (GMPE) development groups and made model-to-model comparisons more meaningful. The scope of the NGA-West 2 database update are to add new ground-motion data and to improve metadata as well as new supporting information to the NGA database to aid the update of the NGA-West GMPEs. A major element of the improved database is the inclusion of significant shallow crustal events occurred post 2003, the cut-off date for events in the previous NGA database.

This paper begins with a brief description and overview of NGA-West 2 database. The new events included in the updated data set are highlighted. The four different metadata sources (source, path, site, time series and spectra) are described afterward focusing new features added during the NGA-West 2 project.

1.1. Database Overview

The NGA-West 2 database started with the PEER NGA database (http://peer.berkeley.edu/peer_ground_motion_database) of ground motion recording from shallow crustal events in active tectonic regions that was completed in 2006 (Chiou et al., 2008). The NGA database was at the time the largest set of uniformly processed set of ground motions available. The NGA-West 2 database continues the similar methodology of data collection as in the previous database to now include records prior to February 2011. The NGA database consists of a set of strong motion records (text files) and metadata tables. The entire set of records was uniformly processed with the PEER strong motion processing algorithm developed by Pacific Engineering and Analysis (PEA) which is detailed below (Darragh et al. 2004). The metadata tables were developed under direction of different working groups and final information entered into data tables went through a significant review process. Each working group contained a panel of experts within that field of interest (e.g. seismologists for the Earthquake Source Working Group). A summary flat file was created from the records and metadata tables and contains the key information used by the individual NGA-West 2 GMPE model developers. In addition to the three-component as-recorded spectra a new orientation-independent rotational spectra called RotDnn (Boore, 2010) is provided. The available spectra are now provided for periods 0.01 to 20 seconds and for 11 different damping ratios ranging from 0.1 to 30%.

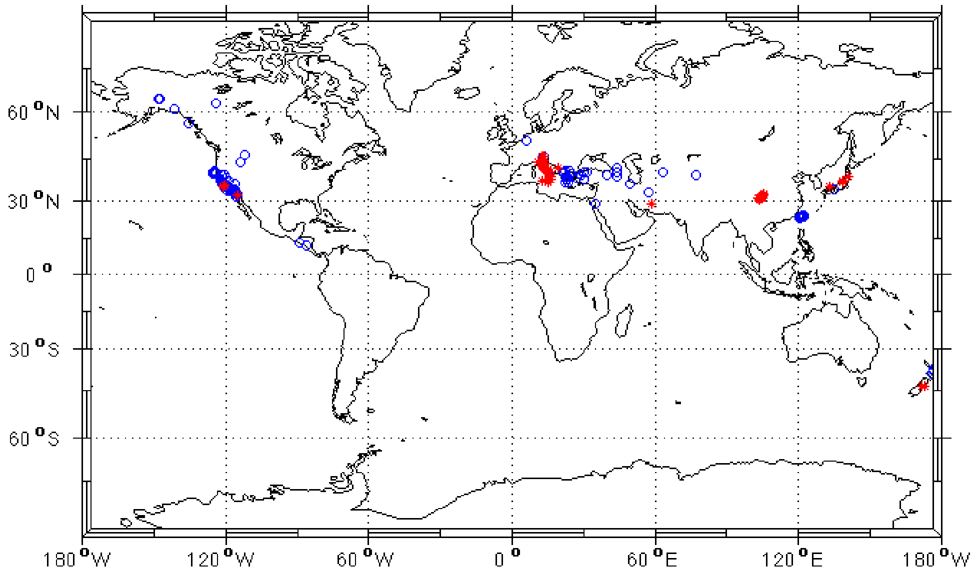

Figure 1. Map of the epicentre distribution of the 334 events. Open circles are events in the previous NGA database and solid stars are events added in the NGA-West 2 database.

Currently the NGA-West 2 database contains 8611 multi-component records from 334 events. Figure 1 shows the distribution of the hypocenter locations and highlights the 178 events that were added. Figure 2 shows a comparison of the magnitude-closest distance distribution between the records in the NGA and NGA-West 2. This new set contains roughly double the number of events and records from the previous database.

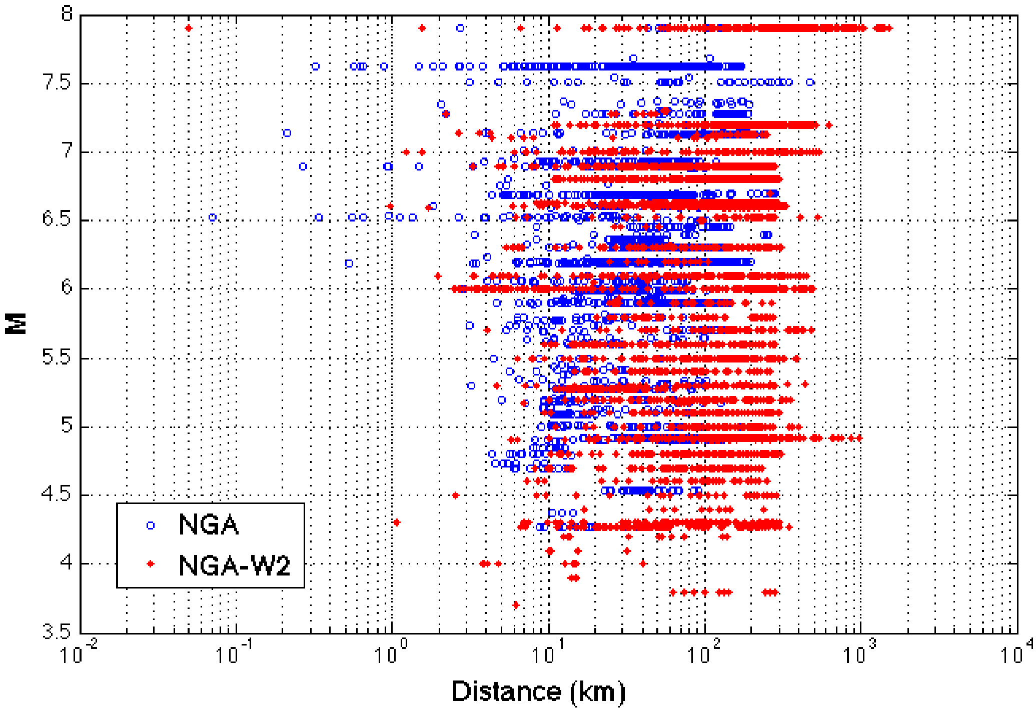

Figure 2. Magnitude-distance distribution of strong-motion records in the NGA-West 2 database.

1.2. Metadata Tables

As in the previous database, four metadata tables were created by different working groups: record catalog, finite source table, a site table, and propagation path table. The data tables used in the NGA project (Chiou et al. 2008) were used as the start of the current tables. In addition to adding information to the databases from additional station/events there was a significant review of existing metadata. New metadata (i.e. R_x) added to each data table are described below. Additionally, improvements have been made toward uniformity and transparency in metadata collection and estimation. A subset of the pertinent information in all the tables are summarized in the database flat file which is used by the various working groups in the NGA-West 2 project.

The record catalog is a list of the strong motion recordings included in the database. The record catalog also contains the spectra, intensity measures (PGA, PGV and PGD), and selected filter corner frequencies. Each record is given a unique record sequence number (RSN) as it is added. The earthquake source table contains earthquake source information for the selected events such as magnitude, hypocenter location, finite fault dimensions (as available), and seismic moment. Events are given a unique earthquake ID number (EQID) as they are added. The site table is a collection of site information for each recording station such as station location, V_{s30}, codes indicating how V_{s30} was established, and various proxies used for V_{s30} estimation. A station is given a unique station sequence number (SSN) as it is added. The three ID numbers (RSN, SSN, and EQID) facilitate a linkage between the tables and aided in the creation of the final flat file.

The following sections give an overview of added metadata, improvement on the metadata and record processing, and review of the new database.

1.3. NGA-West 2 Flat File

The NGA-West 2 flat file serves as a single file of key metadata and ground-motion parameters to be used by NGA developers in regression analysis. It is formed by merging a subset of information within all four tables. The flat file now contains more than 130 columns of metadata, PGA, PGV, PGD and pseudo absolute spectral acceleration at 111 periods. A total of 55 flat files are created for spectral acceleration for the three as-recorded components, RotDnn (to be explained later), and GMRotI50 at 11 different damping ratios.

2. TIME SERIES AND SPECTRA

As in the previous NGA database, Pacific Engineering and Analysis (PEA) collected digitized but otherwise unprocessed accelerograms from various agencies around the world and uniformly processed the raw accelerograms. The PEER processing procedure, including instrument correction, bandpass filtering (removal of unwanted noise), and baseline correction, is described in Darragh et al. (2004). A major change from that procedure is the systematic use of acausal Butterworth filter, whereas previously causal filter was the preferred type of filter. Records were also re-evaluated to extend their usable frequency (and hence reprocessing), for identifying late triggers, for alignment in absolute time, and for potential co-located instruments. The processed time series were used to calculate various types of response spectra as described in Section 2.2.

2.1. Time Series Processing and Station Flags

To improve the quality of the accelergrams and intensity measures included in the database a significant effort was made toward uniformity in the data processing and the evaluation of the completeness of the recorded event. More than 1600 records that were originally passed-through without correction for the original NGA database had their raw accelerograms collected and processed using the PEER methodology. Stations in the previous database that were co-located with another station were removed from the NGA-West 2 database. The entire record set was evaluated for late *p*-/*s*-triggers (e.g. missing portions of the *p*-wave/*s*-wave). Records with portions of the *p*-wave or *s*-wave lost due to a late trigger were flagged.

2.2. RotDnn Spectra

The NGA-West 2 project has selected an orientation independent spectra that is not calculated from the geometric mean of the two horizontal components called RotDnn (Boore, 2010). The 'D' stands for a period dependent rotation angle and 'nn' is the fractile of the rotated spectra sorted by amplitude. Therefore, the rotated spectra over a range of periods will have a non-uniform set of rotation angles. RotDnn gives the spectra from the smallest to largest for any rotation angle. The median amplitude over all non-redundant angles, RotD50, will be used in the development of the updated ground motion prediction equations (GMPE). A comparison of RotD50 to GMRotI50 by Boore (2010) indicates a slight increase in the overall variation in RotD50.

The period dependent minimum and maximum rotated spectral amplitude, RotD00 and RotD100, will be utilized by the Directionality and Directivity working groups to develop equations to convert the RotD50 to the RotD100 (maximum spectral amplitude) and new directivity equations.

The RotDnn values were calculated using the program nga2psa_rot_gmrot.for (personal comm, Boore). Reasons for not providing rotated spectra were: missing horizontal component azimuth, missing a horizontal component, misalignment in absolute time of the two horizontal components.

3. FINITE SOURCE MODELS

The 173 added earthquakes were given source parameters based on the available studies on each event. For large well-recorded events, multiple finite fault models were available and reviewed. As in the previous database, the method of selecting the preferred geometry was based on the type of data used in the inversion (i.e. GPS, geodedic, teleseismic, and strong motion). For the events with multiple solutions, the fault model that used strong motion data in the inversion was preferred. The preferred finite fault model went through several iterations of review by the earthquake source working group, which in some cases involved the review of reports that were in press.

Trimming the areal extent of the finite fault plane of newly added earthquakes is performed using the methodology used in the NGA project. An example application of the finite fault model selection and fault trimming is provided in Stewart et al. (2012a) for one of the events in the database (L'Aquila, Italy).

4. PATH METRICS

Metadata entered into the propagation path table included various distance measures, hanging wall indicator, radiation pattern coefficients, and directivity parameters. Two new distance measures have been added, R_x and R_y, which are related to the generalized coordinate T and U defined in Appendix A of Spudich and Chiou (2008). Furthermore, directivity parameters have been expanded to include multiple models of directivity that are being developed within the NGA-West 2 project. The set of directivity models considered is intended to aid the direct inclusion of directivity effects into the updated NGA GMPEs.

In the previous NGA database, distance metrics such as the Joyner-Boore distance and the distance to the closest point on fault rupture are missing when the finite fault model is not available. A decision is made by the NGA-West 2 developers to adopt the method to simulate finite fault planes for events without a finite fault model but with minimal information of hypocenter, magnitude, and fault plane solution (or style of faulting). The goal of the simulation routine is to obtain an approximate fault geometry that may be used to compute distance metrics and a few other path metadata that require knowledge of finite fault geometry. The methodology is briefly described here but detailed in Youngs (2006), which was reproduced as Appendix B of Chiou and Youngs (2008). In the simulation routine, the missing fault plane information is simulated by random sampling of pertinent probabilistic distributions of fault area, fault aspect ratio, and hypocenter position on the fault plane. With the random sampling, the routine simulates a set of 101 random fault planes that are spun and slid in space but locked to the given hypocenter location. For the set of simulated fault planes the median distance from site to the rupture is estimated. The selected fault plane is the simulated plane that best fits the set of median closest distances to the rupture.

5. SITE DATABASE

The site database is a collection of station information for the stations included in the NGA-West 2 database. The metadata collected for each station included station identification, station location, instrument type and housing (GMX letters), recommended values of V_{s30} (time-averaged shear-wave velocity to 30-meter depth), codes indicating the basis for the recommended V_{s30}, various proxies used for V_{s30} estimation (slope, terrain, geology, geotechnical categories; details in Stewart et al. 2012b), and isoseismal depths (i.e., depths to shear-wave velocity horizons from various 3D velocity models). The site database started with the site table from the NGA project and a major effort was made to update and improve the completeness of the information, especially with respect to V_{s30} and isoseismal depth information.

The assignment of V_{s30} to a site follows the following general protocols:

1. When a V_s profile is available to a profile depth (z_p) > 30 m, V_{s30} is computed from the profile.
2. When a V_s profile is available to a profile depth of $10 < z_p < 30$ m, V_{sz} is computed to depth z_p and V_{s30} is computed from V_{sz} using available correlations (Boore, 2004; Boore et al., 2011).
3. When no profile is available but regionally-calibrated proxy-based relationship between surface geology and V_{s30} is available, that relationship is used. The utilized relationships apply to California and Italy, as described in Stewart et al. (2012b).
4. Other proxies are used when the above conditions are not met, including slope-, terrain-, and geotechnical category-based proxies. In Taiwan, elevation-based proxies are also considered.

The values of V_{s30} are assigned uncertainties in the site database. For the case of proxy-based estimates, the uncertainties are derived from data analysis as described in Stewart et al. (2012b), and tend to be relatively high for rock sites as compared to soil sites.

Depths to shear-wave velocity horizons (1.0, 1.5, and 2.5 km/s) were collected for 3D velocity models and boring measurements for basins in northern and southern California, Japan, and Taiwan. Depths were extracted for sites within the CVM-S4 and the CVM-H11.9 models southern California and for sites within an updated version of the model described by Boatwright et al. (2004) for northern California sites. For sites in Japan, depths were extracted from the Japan Seismic Hazard Information Station.

ACKNOWLEDGEMENTS

This study was sponsored by the Pacific Earthquake Engineering Research Center (PEER) and funded by the California Earthquake Authority, the California Department of Transportation, and the Pacific Gas & Electric Company. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect those of the sponsoring agencies.

REFERENCES

Boatwright, J, L Blair, R Catchings, M Goldman, F Perosi, and C Steedman (2004). *Using Twelve Years of USGS Refraction Lines to Calibrate the Brocher and Others (1997) 3D Velocity Model of the Bay Area*, U.S. Geological Survey Open File Report 2004–1282.

Boore, D.M. (2004). “Estimating Vs30 (or NEHRP site classes) from shallow velocity models (depth < 30 m), ” *Bull. Seism. Soc. Am*., **94**, 591 – 597.

Boore, D. M. (2010). Orientation-independent, nongeometric-mean measures of seismic intensity from two horizontal components of motion. *Bull. Seismol. Soc. Am.* **100**, 1830-1835.

Boore, D.M., Thompson, E.M. and Cadet, H. (2011). Regional correlations of V_{s30} and velocities averaged over depths less than and greater than 30 m. *Bull. Seism. S*oc. Am., **101**, 3046-3059.

Boore, D. M., Watson-Lamprey, J., Abrahamson, N. A. (2006). Orientation-independent measures of ground motion. *Bull. Seismol. Soc. Am.* **96**, 1502-1511.

Chiou, B., Darragh, R., Gregor, N. (2008). NGA project strong-motion database. *Earthquake Spectra* **24**, 23-44.

Chiou, B., and Youngs, R.R. (2008). NGA model for average horizontal component of peak ground motion and response spectra. *PEER report.*

Darragh, R., Silva, W.J., Gregor, N. (2004). Strong motion record processing procedures for the PEER center. *Proceedings of COSMOS Workshop on Strong-Motion Processing, ull. Seismol. Soc. Am.* 1-12.

Spudich, P., Chiou, B. (2008). Directivity in NGA earthquake ground motions: analysis using isochrone theory. *Earthquake Spectra.* **24**, 279-298.

Stewart, J.P., Lanzo,G., Pagliaroli, A., Scasserra, G., DiCapua, G., Peppolini, S., Darragh, R. and Gregor, N. (2012a). Ground motion recordings from the Mw 6.3 2009 L’Aquila earthquake in Italy and their engineering implications, *Earthquake Spectra*, **28:1**, 317-345.

Stewart, J.P., Seyhan, E., Boore, D.M.,, Campbell, K.W., Erdik, M., Silva, W.J., Di Alessandro, C., and Bozorgnia, Y. (2012b). Site effects in parametric ground motion models for the GEM-PEER global GMPEs project, 15th WCEE, Portugal. Submitted.