18TH WORLD CONFERENCE ON EARTHQUAKE ENGINEERING
WCEE2024
MILAN

MILAN, ITALY
30TH JUNE - 5TH JULY 2024
www.wcee2024.it

NEW DEVELOPMENTS AT THE CENTER FOR ENGINEERING STRONG-MOTION DATA (CESMD)

L. Hagos[1], H. Haddadi[1], L.S. Schleicher[2], J. Steidl[3], E. Thompson[4], H. Crume[1], M. Dhar[1], and N. Leue[1]

[1] California Geological Survey, Sacramento, California, USA, Lijam.Hagos@conservation.ca.gov

[2] U.S. Geological Survey, Earthquake Science Center, Moffett Field, California, USA

[3] Earth Research Institute, University of California Santa Barbara, California, USA

[4] U.S. Geological Survey, Geologic Hazards Science Center, Golden, Colorado, USA

Abstract: The Center for Engineering Strong-Motion Data (CESMD), an internationally utilized joint center of the U.S. Geological Survey (USGS) and the California Geological Survey (CGS), provides a single access point for earthquake strong-motion records and station metadata from the CGS California Strong-Motion Instrumentation Program (CSMIP), the USGS National Strong-Motion Project (NSMP), the USGS Advanced National Seismic System (ANSS), and other affiliates. The CESMD has been continuously improving its webtools to facilitate the access of strong-motion data and metadata for use in post-earthquake response and for scientific and engineering research applications. The Center provides raw and processed strong-motion data via the Engineering Data Center (EDC) and the Virtual Data Center (VDC) web portals. This paper focuses on the strong-motion products provided by the EDC where more than 48,000 records with peak ground accelerations greater than 0.1% g from over 2400 earthquakes are currently hosted and on the ongoing efforts to develop data access tools and applications. The new developments and ongoing efforts in the EDC include: 1) enhancements to the CESMD webservices to facilitate access to station metadata, earthquake information, and strong motion records, 2) new features to the interactive map interface, improving the visualization and access to earthquake, station, and record information, 3) efforts to develop a new web application tool for data format conversion from a number of data formats, 4) efforts to unify varying waveform data formats into a consistent format, 5) ongoing efforts to compile seismic station site geology, measured or inferred Vs30 values, shear-wave profiles, National Earthquake Hazards Reduction Program (NEHRP) site class, and available structural instrument deployment schematics, and 6) special studies pages for research topic-specific ground motion datasets that offer uniform processing of records from a variety of sources.

1. Introduction

The Center for Engineering Strong-Motion Data (CESMD) (Haddadi et al., 2008; strongmotioncenter.org) was established in 2007 as a cooperative effort between the U.S. Geological Survey (USGS) National Strong-Motion Project (NSMP) and the California Geological Survey (CGS) Strong Motion Instrumentation Program (CSMIP) in Moffett Field and Sacramento, California, respectively. The goal of the CESMD is to provide a single access point to quality-controlled strong-motion data from the United States (U.S.) and international earthquakes in a timely manner. The CESMD aims to provide earthquake data with a PGA threshold of 0.1% g and a magnitude threshold of M3.0 and above in California, and M4.0 within the rest of the conterminous United States, Hawaii, Puerto Rico, and Alaska. Datasets of interest to the engineering and scientific communities, such as event sequences in areas of induced seismicity and significant global events, are also processed and posted at the CESMD when available. In line with this general goal, the CESMD working group has been making efforts to: 1) develop standalone and web-based tools and applications to increase the utilization of strong-motion data products, 2) establish collaborative work with local and international strong-motion data centers for quick exchange of data, and 3) improve the delivery of waveform and parametric data in a unified data format.

2. The EDC and VDC Web Portals

The CESMD currently serves strong-motion data via either the Engineering Data Center (EDC) (https://www.strongmotioncenter.org/cgi-bin/CESMD/search1.pl) or the Virtual Data Center (VDC) (https://www.strongmotioncenter.org/vdc/scripts/default.plx) portals. Historically, strong-motion records in California are exchanged with the California Integrated Seismic Network (CISN) partners in near real-time for automatic processing and dissemination via the EDC portal of the CESMD. In recent years (~circa 2019), when the NSMP stood up its automated strong-motion data processing workflow, the CESMD expanded to include records automatically processed in near real-time from the U.S Advanced National Seismic System (ANSS) (USGS, 2017) networks located within the continental U.S., Hawaii, Alaska, and Puerto Rico (Schleicher et al., 2024, Steidl et al., 2022) at the EDC.

In parallel, the CESMD as part of its mission over the last several decades has endeavoured to populate strong motion records shared from a variety of international counterparts for significant global earthquakes. For international earthquakes occurring outside of the U.S. before 2019, generally most of these datasets can be found within the VDC. Since our current automated processing workflow and associated webservices have been developed for the EDC, more recent international earthquake datasets occurring outside the U.S. since 2019 (when the appropriate permissions are obtained) can be found within the EDC. For most international earthquakes hosted by the VDC, users are provided with virtual links to remote data repositories where the actual strong-motion data files reside.

One main difference between the EDC and VDC portals is that the EDC database has a station-based structure while the VDC is channel-based. To make user data access easier (via both webservices and through the GUI user interface), we are unifying both the EDC and VDC into one Merged Data Center (MDC) which we are executing in two phases. Phase 1 involves merging the EDC and VDC databases while keeping the public facing interfaces as is, and phase 2 will involve the complete merging of the two portals and web-interfaces.

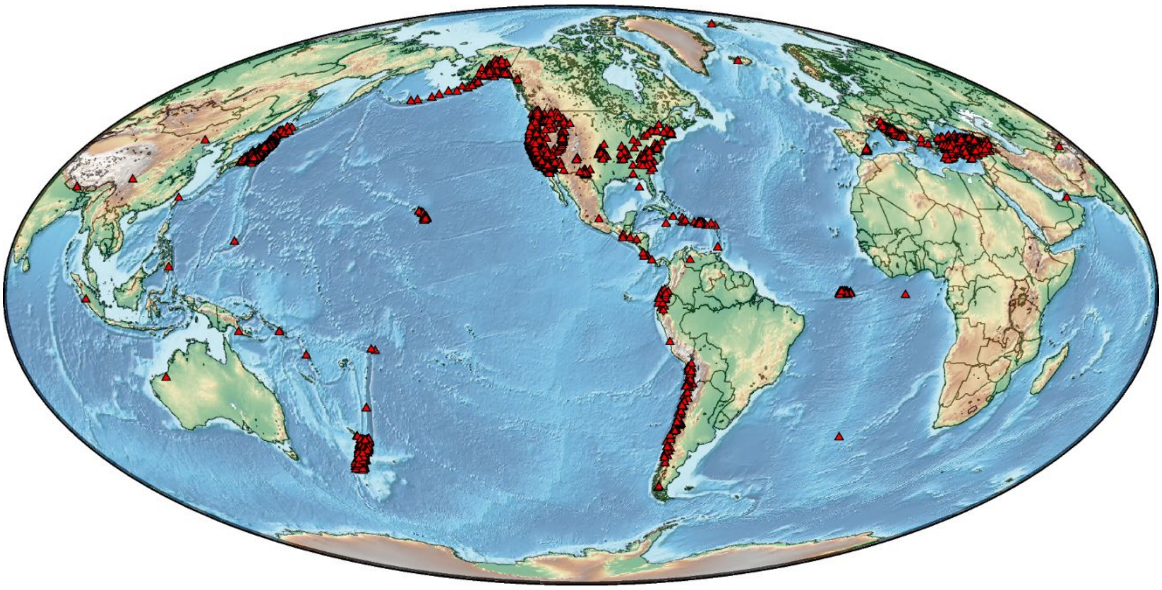


Figure 1. Map showing the seismic stations (red triangles) that contributed to the strong-motion data available at the Engineering Data Center (EDC) of the CESMD.

In addition to United States earthquakes, the CESMD currently hosts data from significant international earthquakes in Japan, New Zealand, Italy, Türkiye, Chile, and Mexico. As of December 2023, there are over 48,000 station records from 2,400 earthquakes available in the EDC contributed by more than 68 networks as shown in Figure 1. The strong-motion data added to the EDC has significantly increased in recent years (Figure 2) due to two main factors:- 1) Starting in 2018, the CESMD adopted a new policy to lower the PGA threshold for posting records from 0.5% g to 0.1% g, and 2) The USGS NSMP adopted an automated workflow that involves the acquisition and processing of strong-motion data from the U.S. ANSS regional networks and records from networks outside of the U.S. for international earthquakes (Steidl et al., 2022). Currently, there is an ongoing project to backfill significant global earthquakes at the CESMD where possible (Shao et al., 2024). In addition, we have a large dataset of small amplitude (PGA < 0.5% g) strong-motion records from earthquakes prior to 2018 that users can access through links to an FTP site found at the Internet Quick Report (IQR) pages (see Section 3). These small amplitude records are missing from the EDC database, and an effort is underway to add them to the database. As the traditional boundary between weak and strong ground motion records becomes blurred with modern seismic instrumentation and big data and machine learning technology, a further change to the PGA threshold may be considered to host available waveform data above appropriate signal-to-noise ratio (SNR) thresholds to support a variety of earthquake engineering and event response applications (Aagaard et al., 2022).

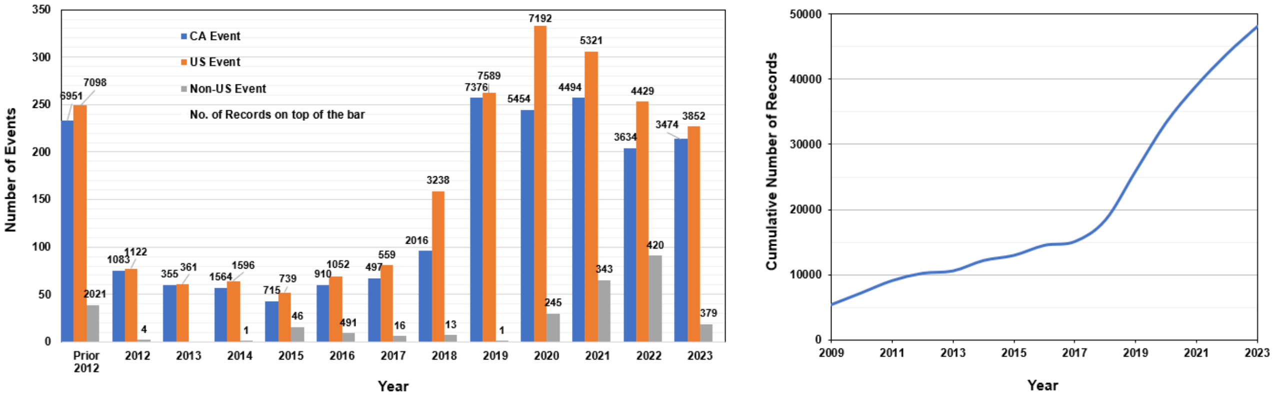


Figure 2. Left: Histogram showing the number of events and records at the EDC available at CESMD in California (blue), the broader U.S. (orange), and non-U.S. events (gray) and Right: Cumulative number of all records in the EDC available at the CESMD as of December 2023.

As part of the NSMP automated workflow, collaborative efforts to collect data from the U.S. and international data centers resulted in a sharp increase of the volume of data added to the EDC as shown in Figure 2. For non-U.S. data providers, the CESMD adds proper data attribution by adding the logos and URLs of the data providers to the event pages. In addition, the CESMD team is working to provide citation information (Shao et al., 2024) to users so that they can incorporate proper attribution to the data contributing networks and data centers when they utilize CESMD data for their scientific research and analysis.

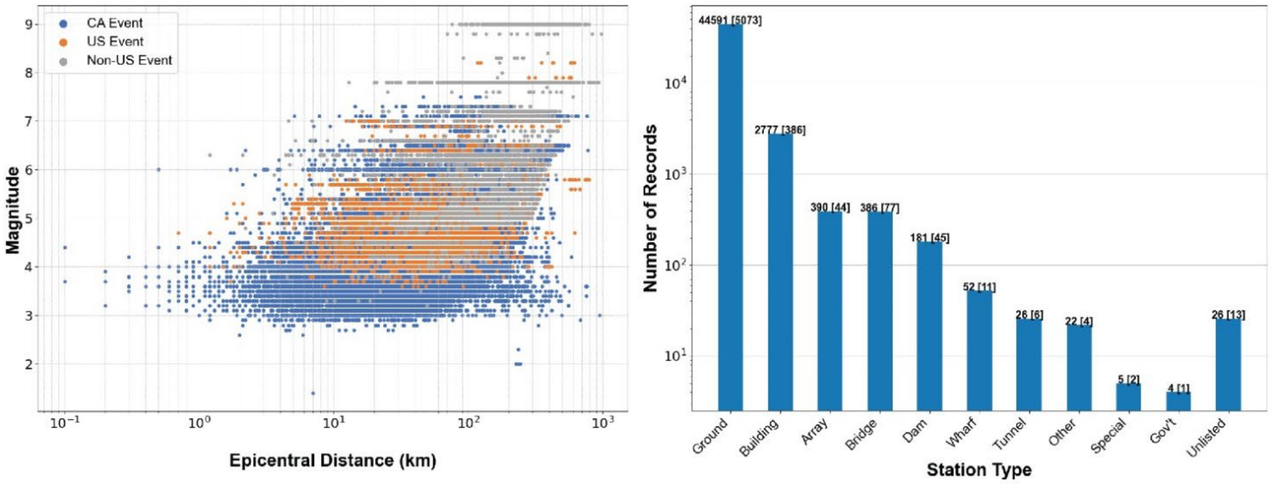


Figure 3. Left: Magnitude-distance distribution of records hosted at the EDC for California (blue), other U.S. states (orange), and other countries (gray). Right: Histogram showing a breakdown of contributions by station type indicating availability of records from ground response and a variety of man-made structural station sites. Bracketed values above bars indicate number of stations responsible for creating records.

The magnitude-distance distribution (Figure 3) of data at the EDC shows a good distribution of records in the lower magnitude and distance ranges but shows scarcity of data in the near-field for larger magnitude earthquakes. The dataset in Figure 3 includes records from ground response stations and man-made structures classified by the station types queried from the CESMD webservices for the EDC. Data download statistics (Figure 4) indicate users' interest in downloading both raw and processed records. However, CESMD users downloaded more processed data than raw data, suggesting the usefulness for the CESMD to distribute processed data by careful consideration of processing parameters such as filter corners.

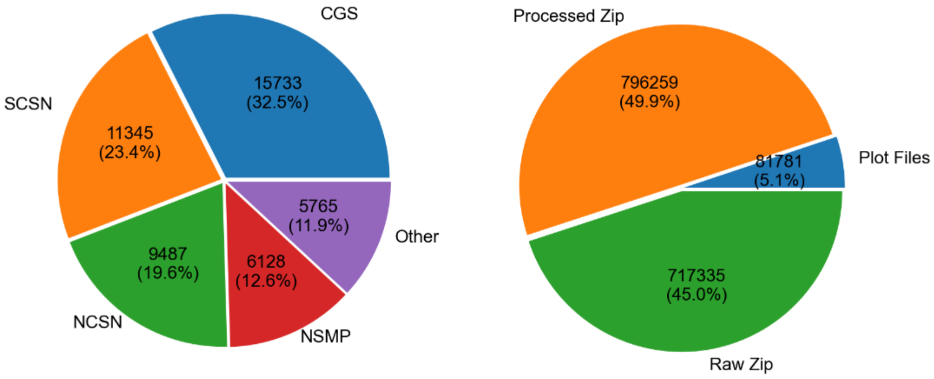


Figure 4. Left: Visualization of the networks that contributed strong-motion data to the CESMD. Right: Visualization of the type and number of downloaded waveform data files.

3. Strong-Motion Data Products

Strong-motion data collected for California earthquakes are exchanged between the CISN networks in near real-time and are automatically processed by the CSMIP's Strong-motion Automated Recovery and Analysis (SARA) system (Shakal et al., 2018). In the contiguous U.S. and its territories, the NSMP acquires, reviews, and processes strong-motion data using the Processing and Review Interface for Strong-Motion Data (PRISM) software (Jones et al., 2017) for ANSS accelerometers. All data products are disseminated via the Internet Quick Report (IQR) of the CESMD for quick access and utilization in post-earthquake evaluation and further analysis. The IQR is an event-specific page containing the data products generated for the event. When available, the IQR provides the following data products:

- Raw and processed waveform data.
- A flat file containing peak ground motions (PGA, PGV, PGD, and SA at 3 periods) in text and CSV format. Both CSMIP and NSMP currently compute spectral accelerations at 91 periods, and they are considering expanding the number of periods to 100 to accommodate requests from the engineering community.
- Files with plots of acceleration, velocity, and displacement timeseries and spectra.
- Link to an interactive map showing color-coded stations representing ranges of PGA values.
- Downloadable KML file for Google Earth/Geographic Information System (GIS) applications (Open Geospatial Consortium).
- Link to ShakeMap – a graphical representation of ground shaking (Wald et al., 2021; Worden et al., 2018; Worden et al., 2020).

3.1 Data Processing and Quality

While the CESMD strives to provide the above-mentioned data product types when available, the data received from the contributing networks may come with different sampling rates, processing approaches, and data formats. In California, most modern triggered systems record 200 samples per second (sps) and the real-time stations record at 100 sps. For some of the stations belonging to the CISN networks that record both 100 sps and 200 sps data, the highest sampling rate record will be posted at the CESMD and used in ShakeMap. In general, the 100 sps and 200 sps data are similar in the 0.2 - 40 Hz frequency band, with obvious differences at higher frequencies approaching and beyond the 50 Hz Nyquist frequency of the 100 sps data (Figure 5). Engineering applications requiring frequencies greater than 40 Hz will need to utilize the 200 sps data.

Decimating the 100 sps data by a factor of 2 results in significant differences in spectral acceleration (SA) (Figure 5) and Fourier amplitude (not shown) for frequencies greater than 10 Hz, demonstrating the high-frequency differences that can result from decimated data.

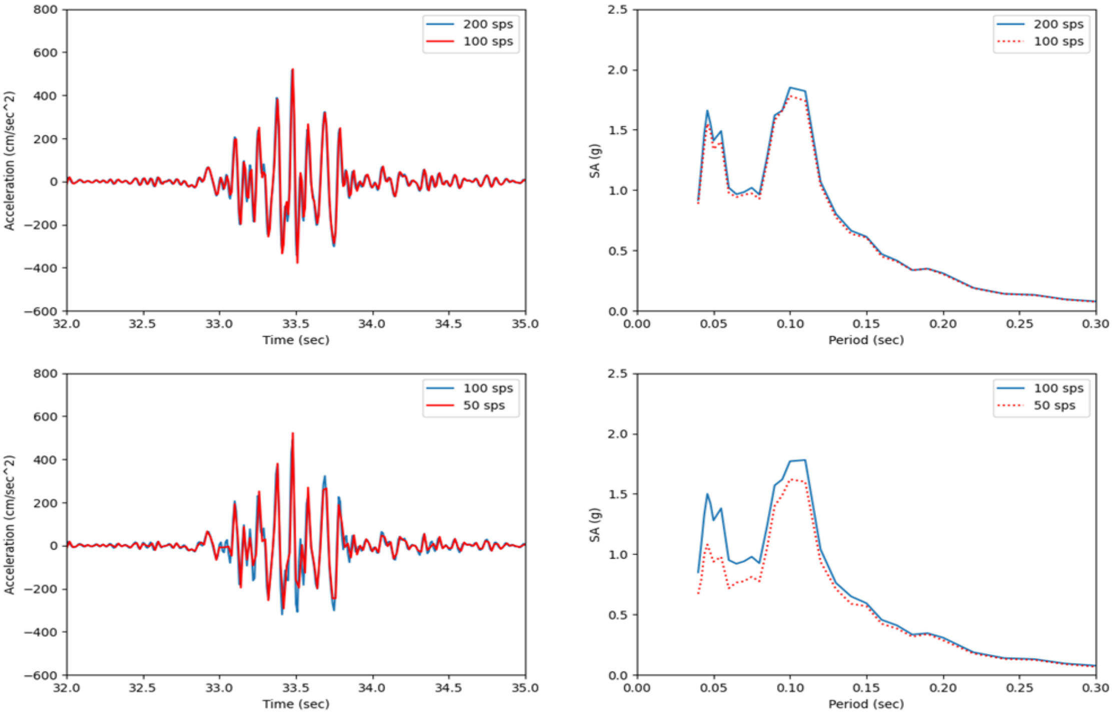


Figure 5. Comparison of processed 100 sps and 200 sps records (bandpass filter: 0.3 Hz – 23 Hz) and corresponding spectral acceleration (SA) at NSMP station NP5230 for the M4.9 Anza earthquake of April 3rd, 2020. Top: A direct comparison between the 100 sps and 200 sps data. Bottom: A comparison of the same records but with decimated samples by half. Large amplitude differences are observed for frequencies higher than about 10 Hz as shown in the spectral plot (Reproduced from Hagos et al., 2022).

Despite small differences in the data processing workflows, the SARA and PRISM automated systems largely result in similar processed waveforms and spectral amplitude values (Jones et al., 2017; Kalkan and Stephens, 2017; Shakal et al., 2018). Both systems currently use default filter corners for automatic processing which may later be adjusted upon review by a seismologist. One of the most important challenges in processing strong-motion acceleration data is the choice of the long-period filter corner, i.e., the low-cut, which is the longest period at which the data are reliable and is dependent on the SNR (Boore & Bommer, 2005). The example in Figure 6 shows a record from the 2019 M7.1 Ridgecrest earthquake where differences in the PGA, PGV, Fourier amplitude spectra (FAS), and SNRs are clearly observed when the long-period corner is adjusted from 3.3s to 15s. In this example, the 15s low-cut is most appropriate based on a SNR threshold of 3 (Figure 6). The differences between the 3.3s and 15s low-cut filtered signals are more pronounced in the PGV records than the PGA records, with higher PGV values recorded for signals filtered with a 15s low-cut (Figure 6). The choice of the long-period filter corner is clearly important to consider in any automated strong-motion data processing system.

The quality control and filter determination of strong-motion data are intertwined processes that can be addressed simultaneously. Therefore, it is imperative that automated data processing systems include an automatic calculation of the SNR for quality assessment and objectively determine the optimal filtering frequency band using a combination of methods which may consider the SNR, the FAS slope (which should decay proportionally to the reciprocal of the frequency squared at long-periods), and the integrated displacement signal. While there is a need for this type of automated and uniform processing to determine key processing parameters in real-time, manual reviews and adjustments cannot be completely avoided as a post-processing step and processing decisions should be standardized. There are already a few ongoing collaborative efforts in standardizing the strong-motion data processing and the generation of data products in a unified format. The CESMD team is currently working on a plan for benchmarking the SARA (Shakal et al., 2018), PRISM (Jones et al, 2017) and the recently developed gmprocess software (Hearne et al., 2019; Schleicher et al., 2024). Other efforts in this area include an initiative by the Consortium of Organizations for Strong Motion Observation Systems (COSMOS) (https://strongmotion.org) to organize a workshop that is aimed at developing standards for processing and quality control of strong-motion data.

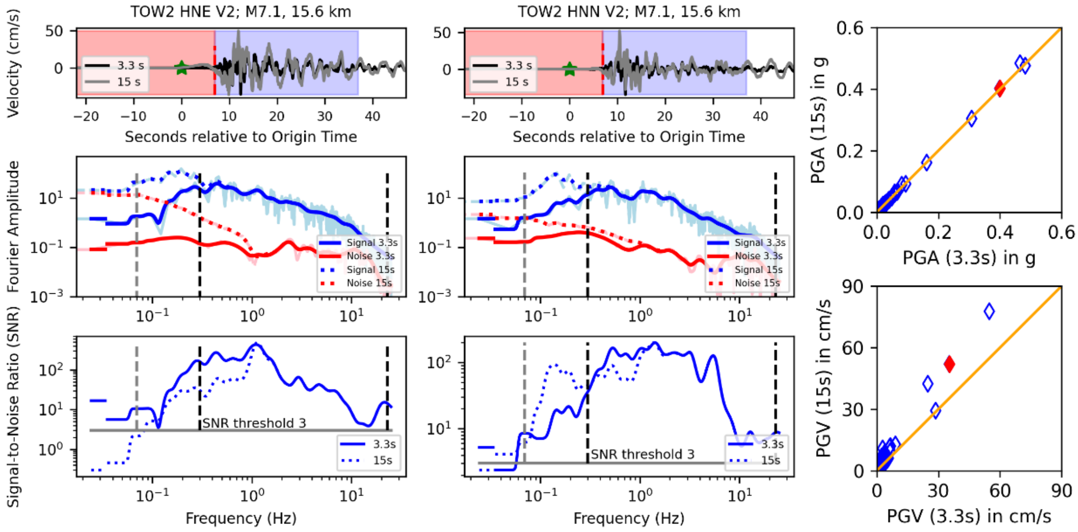


Figure 6. Left two panels: Example of two horizontal component records from the 2019 M7.1 Ridgecrest earthquake showing how changes in long-period filter corners from 15s (grey dashed lines) to 3.3s (left-most black dashed lines) result in amplitude differences mainly shown on the velocity and Fourier amplitude spectra (FAS), as well as changes in the ratio of the signal (shown in blue) to the noise (shown in red) with a signal-to-noise ratio (SNR) threshold of 3 indicated by the horizontal grey lines on the bottom panels. Right: Comparisons of PGA and PGV between data filtered at 3.3s and 15s for the long-period filter corner/low-cut. Red diamond is the PGV at station TOW2; blue diamonds are other stations that recorded the earthquake. These records are available at the EDC.

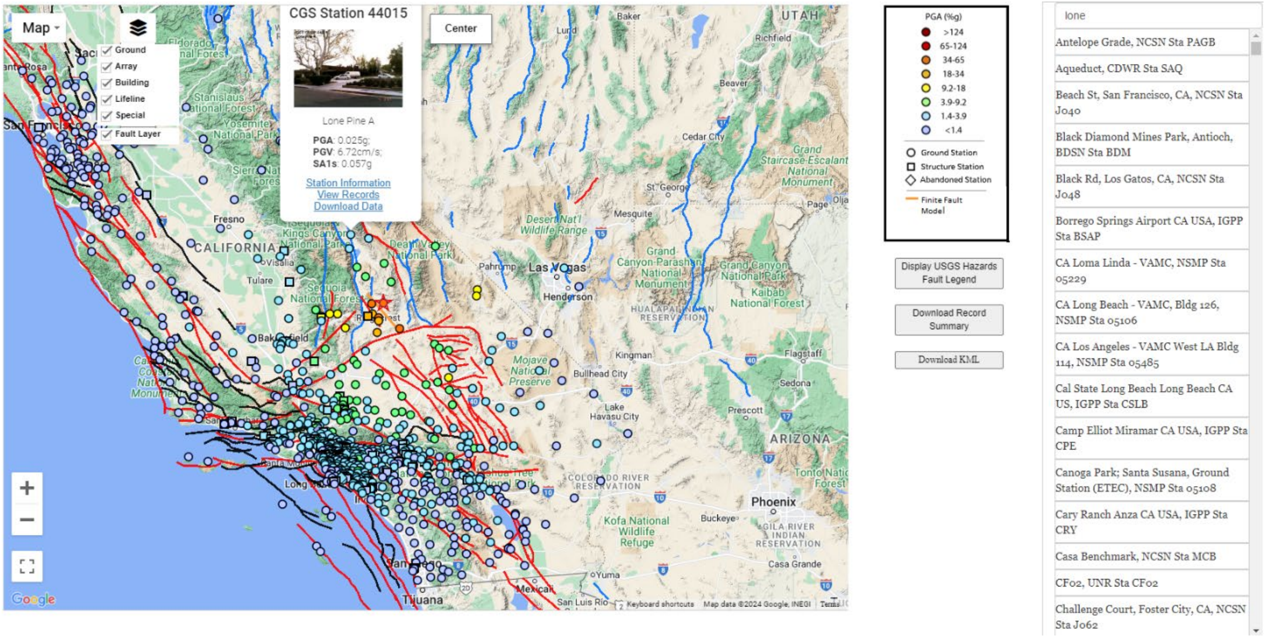
Strong-Motion Data Format

Waveform data that are processed by the SARA and PRISM systems are distributed in CGS or COSMOS formats. Data contributions shared by networks outside of the U.S. for some significant earthquakes are also available in formats adopted by the data providers. Recognizing the challenge that disseminating data in different formats poses to data users, the CESMD has made few data conversion tools available to users under “About CESMD” on the CESMD website. Further, the CESMD is currently in a process of developing an enhanced web-based data conversion application to increase the accessibility and utilization of data. The web application under development is being designed to convert data at the CESMD from various institutions not in COSMOS format to a variety of formats requested by the earthquake engineering community including MATLAB, Extensible Markup Language (XML), and tagged single column American Standard Code for Information Interchange (ASCII).

Most waveform data at the CESMD usually contain event origin information in the data file headers. For data exchanged in near real-time, the earthquake origin information could still be the initial reporting by the authoritative network. User verification and cross-checking of the earthquake origin information at the USGS event pages for updated information is useful. In the future, an event URL will be included in the comment lines of the waveform data in COSMOS format, which will give users quick access to the USGS event page where updated earthquake information can be retrieved.

3.2. The Interactive Map

The Interactive Map feature at the CESMD is an event-specific product that is accessible by clicking the icon at the top of event IQR pages. The Interactive Map provides a visual of the spatial distribution of the stations that recorded an event (Figure 7). The current version of the map has added several important features including a fault overlay (Hatem et al., 2022) and options for the user to select or deselect station types (e.g., Ground, Geotech array, Building, Lifeline) to display. The stations (ground and structure) are color-coded according to PGA values and are consistent with the intensity scale used in ShakeMap (Wald et al., 2021; Worden et al., 2018; Worden et al., 2020). Further, by clicking on the individual station icons a pop-up window appears providing users with quick links to the station page, the record plot, and the downloadable data files.



CGS Station 44015

Lone Pine A

PGA: 0.025g;
PGV: 6.72cm/s;
SA1s: 0.057g

Station Information
View Records
Download Data

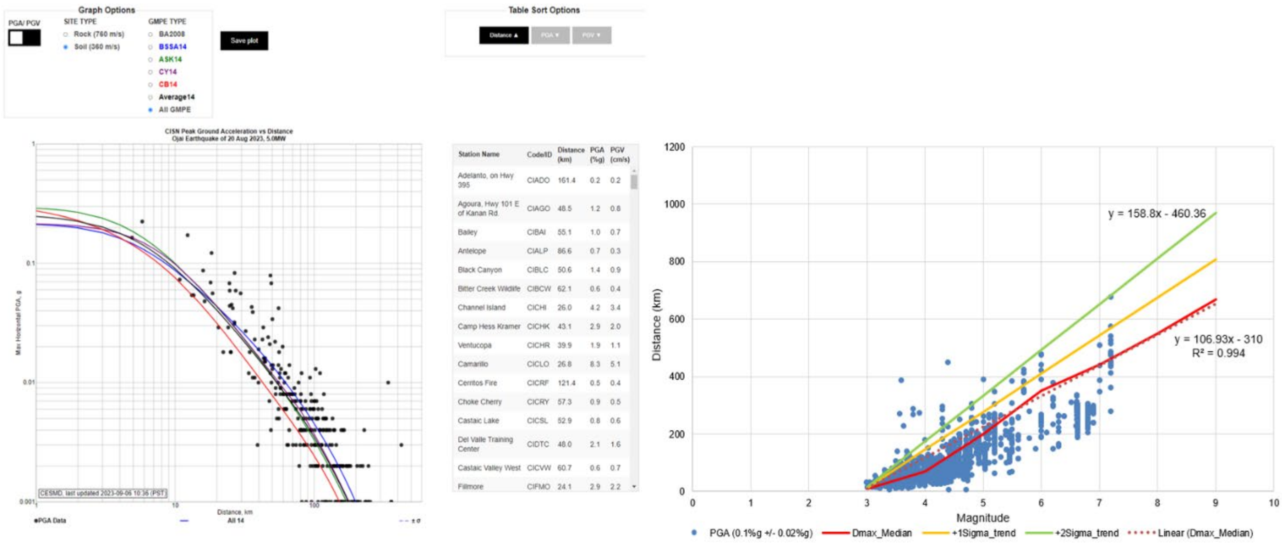
Antelope Grade, NCSN Sta PAGB
Aqueduct, CDWR Sta SAQ
Beach St, San Francisco, CA, NCSN Sta J040
Black Diamond Mines Park, Antioch, BDSN Sta BDM
Black Rd, Los Gatos, CA, NCSN Sta J048
Borrego Springs Airport CA USA, IGPP Sta BSAP
CA Loma Linda - VAMC, NSMP Sta 05229
CA Long Beach - VAMC, Bldg 126, NSMP Sta 05106
CA Los Angeles - VAMC West LA Bldg 114, NSMP Sta 05485
Cal State Long Beach Long Beach CA US, IGPP Sta CSLB
Camp Elliot Miramar CA USA, IGPP Sta CPE
Canoga Park; Santa Susana, Ground Station (ETEC), NSMP Sta 05108
Cary Ranch Anza CA USA, IGPP Sta CRY
Casa Benchmark, NCSN Sta MCB
CF02, UNR Sta CF02
Challenge Court, Foster City, CA, NCSN Sta J062

Figure 7. An example of the Interactive Map for the M7.1 Ridgecrest earthquake of July 2019, zoomed-in to show the distribution of the stations in the Los Angeles basin. An example of the pop-up window for a specific station (44015) with quick links to the station page, record plot, and the downloadable data files is also shown. The stations are color-coded according to peak shaking levels as shown in the map legend and consistent with ShakeMap.. This Interactive Map example is available at the CESMD website here: https://www.strongmotioncenter.org/stationmap/StationMapD1.php?ID=ci38457511.

3.3. Ground Motion Plots

Ground-motion plots on the IQR pages are dynamically generated for earthquakes with M4 and above (Figure 8). This product provides PGA and PGV vs distance plots. Recent developments to the ground motion plot product includes the 2014 four NGA-West2 ground-motion models (GMM) (Ancheta et al. 2014) and the Boore and Atkinson (2008) GMM, with an option to toggle between rock and soil sites for comparing observations with prediction. It also adds an equally weighted average curve of the four models consistent with the active crustal GMM used in ShakeMap. Currently, the GMMs are only displayed for earthquakes in the western U.S. For earthquakes outside of the western U.S. for which the NGA-West2 GMMs are developed, only the observations are displayed on the ground motion plot. The EDC plans to add other region-specific GMMs from across the world by utilizing available algorithms in OpenQuake (Pagani et al., 2014).

In real-time processing, automated quality control algorithms may not catch noisy records or rarely, some records can be associated with wrong events when multiple events occur close in time. In such cases, the ground motion plot can serve as another layer of post-processing quality control tools to identify outliers. Moreover, seismic networks can use the GMMs as a quick guide to determine the maximum distance for the harvesting and distribution of ground motion waveform and parametric data in real-time. For a PGA-based curve, the maximum distance for associating ground motions can be estimated from points at which the curve falls below the defined threshold (CISN criterion for exchange of waveform data is a PGA threshold of 0.1% g) with consideration of the standard deviation as shown in Figure 8 for earthquakes in the Western U.S. The distance shown on the ground motion plot is rupture distance when a finite fault model is available, or epicentral distance otherwise. The maximum distance for harvesting ground motion data in real-time processing is magnitude and region dependent.



Graph Options

PGA/ PGV

SITE TYPE
○ Rock (760 m/s)
● Soil (360 m/s)

GMPE TYPE
○ BA2008
○ BSSA14
○ ASK14
○ CY14
○ CB14
○ Average14
● All GMPE

Table Sort Options

Station Name	Code/ID	Distance (km)	PGA (%g)	PGV (cm/s)
Adelanto, on Hwy 395	CIADO	161.4	0.2	0.2
Agoura, Hwy 101 E of Kanan Rd.	CIAGO	48.5	1.2	0.8
Bailey	CIBAI	55.1	1.0	0.7
Antelope	CIALP	86.6	0.7	0.3
Black Canyon	CIBLC	50.6	1.4	0.9
Bitter Creek Wildlife	CIBCW	62.1	0.6	0.4
Channel Island	CICHI	26.0	4.2	3.4
Camp Hess Kramer	CICHK	43.1	2.9	2.0
Ventucopa	CICHR	39.9	1.9	1.1
Camarillo	CICLO	26.8	8.3	5.1
Cerritos Fire	CICRF	121.4	0.5	0.4
Choke Cherry	CICRY	57.3	0.9	0.5
Castaic Lake	CICSL	52.9	0.8	0.6
Del Valle Training Center	CIDTC	48.0	2.1	1.6
Castaic Valley West	CICVW	60.7	0.6	0.7
Fillmore	CIFMO	24.1	2.9	2.2

Figure 8. Left: Example Ground Motion Plot for the 2023 M5.0 Ojai earthquake showing the maximum horizontal ground motion versus distance using the NGA-West2 Ground Motion Prediction Equations (Ancheta et al. 2014). Right: A magnitude-based estimate of the maximum distance below which the threshold PGA falls to 0.1% g. These records are available at the EDC.

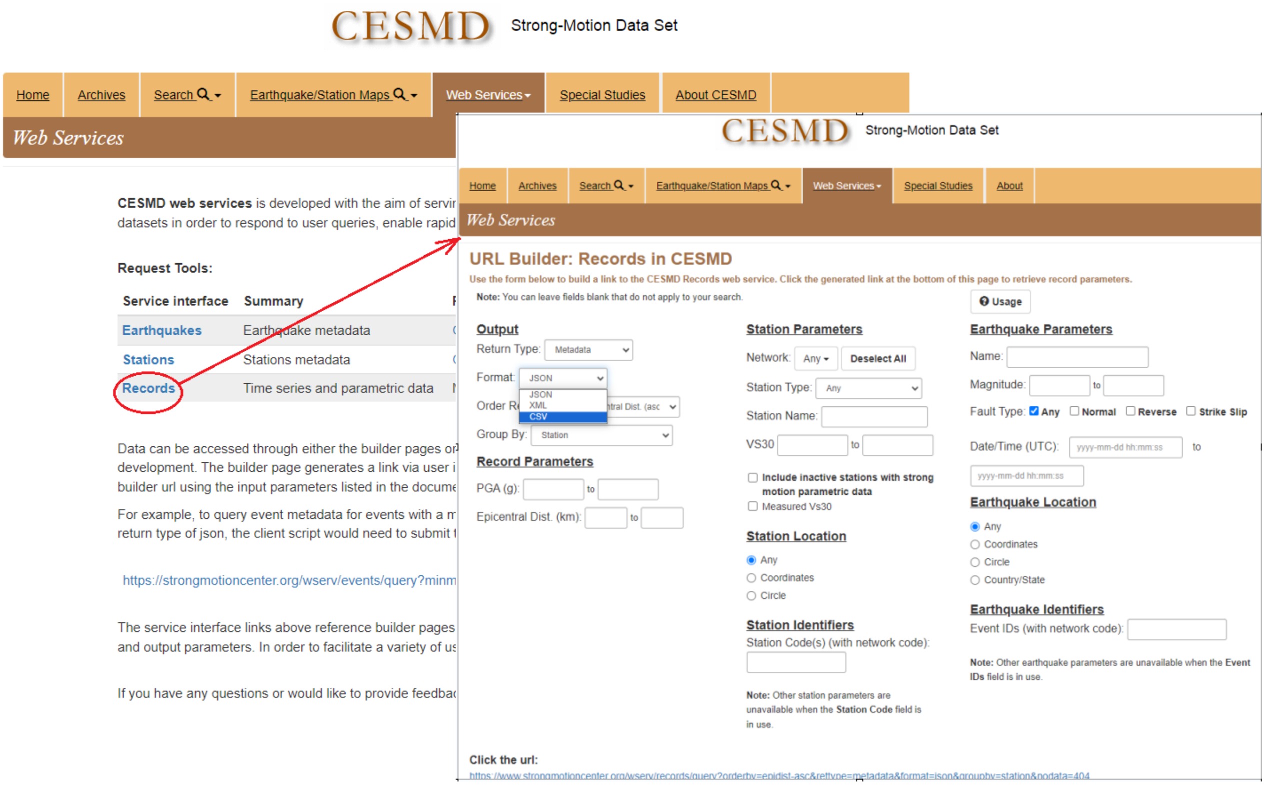
4. Data Access Tools

Users can query the CESMD for either event-specific data or from multiple earthquakes based on search parameters. In general, there are four ways to access data from the EDC:

1. Direct access to data via the event specific IQR pages
2. Search engine
3. Web services
4. Station pages and maps

Given the increase in the volume of the uploaded and utilized/downloaded parametric and waveform data, CESMD staff have been working to develop new and enhanced data access tools. In addition to the search engine, the CESMD has developed web services tools to facilitate the access and utilization of data. The CESMD web services is a RESTful Application Program Interface (API) that provides users with variety of functionalities to access event, station, and record data (Figure 9). Data users can query the EDC database by using the CESMD URL builder, or by building URLs programmatically using their own scripts, or by using the standalone client script that can be found at a GitHub repository (https://github.com/SMIP-CGS/cesmd-webservices-client).

While the event and station web service allow users to query earthquake and station metadata information, respectively, the record web-services provides a comprehensive record-based search that returns parametric outputs in CSV, XML and JSON formats, waveform data, and associated plot files. The web services allow users to download records and metadata grouped either by event or by station. Currently, there is a 200 record-per-download limit at a time to manage network bandwidth. The CESMD is working to improve the download of large datasets through changes to permission levels via registration pages to let users indicate their agreement to the CESMD bulk download policy.



CESMD web services is developed with the aim of servir
datasets in order to respond to user queries, enable rapid

Request Tools:

Service interface	Summary
Earthquakes	Earthquake metadata
Stations	Stations metadata
Records	Time series and parametric data

Data can be accessed through either the builder pages or
development. The builder page generates a link via user i
builder url using the input parameters listed in the docume

For example, to query event metadata for events with a m
return type of json, the client script would need to submit t

https://strongmotioncenter.org/wserv/events/query?minm

The service interface links above reference builder pages
and output parameters. In order to facilitate a variety of us

If you have any questions or would like to provide feedbac

URL Builder: Records in CESMD

Use the form below to build a link to the CESMD Records web service. Click the generated link at the bottom of this page to retrieve record parameters.

Note: You can leave fields blank that do not apply to your search.

Output

Return Type: Metadata

Format: JSON (JSON, XML, CSV)

Order By: Epicentral Dist. (asc)

Group By: Station

Record Parameters

PGA (g): to

Epicentral Dist. (km): to

Station Parameters

Network: Any Deselect All

Station Type: Any

Station Name:

VS30 to

☐ Include inactive stations with strong motion parametric data

☐ Measured Vs30

Station Location

● Any
○ Coordinates
○ Circle

Station Identifiers

Station Code(s) (with network code):

Note: Other station parameters are unavailable when the **Station Code** field is in use.

Usage

Earthquake Parameters

Name:

Magnitude: to

Fault Type: ☑ Any ☐ Normal ☐ Reverse ☐ Strike Slip

Date/Time (UTC): yyyy-mm-dd hh:mm:ss to yyyy-mm-dd hh:mm:ss

Earthquake Location

● Any
○ Coordinates
○ Circle
○ Country/State

Earthquake Identifiers

Event IDs (with network code):

Note: Other earthquake parameters are unavailable when the **Event IDs** field is in use.

Click the url:

https://www.strongmotioncenter.org/wserv/records/query?orderby=epidist-asc&rettype=metadata&format=json&groupby=station&nodata=404

Figure 9. A snapshot of the current CESMD record web-services URL builder showing the sections for user query parameters (Direct Link to CESMD Webservices Interface: https://www.strongmotioncenter.org/wserv/).

5. Station Site Characterization

Site characterization provides information about site geology, shear-wave velocity (Vs) profiles, time-averaged shear wave velocity in the top 30 m (Vs30), and the National Earthquake Hazards Reduction Program (NEHRP) site class. In addition to waveform data and related products, site-specific information is made available to users via the CESMD station page where station photos and schematics for structural response stations are displayed. Recognizing the importance of site characteristic information, the CSMIP and NSMP have been compiling Vs30 information for station sites from different sources. The CSMIP has an ongoing site characterization project to make direct Vs30 measurements at CGS ground station sites and has so far completed measurements at 291 sites (GeoVision, 2016; Petralogix, 2017; Yong et al., 2013). For stations where measured Vs30 is not available, CSMIP uses inferred Vs30 from proxy methods extracted from the Vs30 grid map for California by Thompson et al., 2014. While inferred Vs30 values from published maps and other proxies are used, they cannot replace measurements as demonstrated in Figure 10. So far, there are 530 stations with measured Vs30 values at the CESMD.

Expanding the compilation effort, current efforts include working to integrate a recent compilation of site geology and seismic velocity characteristics at the U.S. ANSS strong motion station sites at the CESMD (Huddleston et al., 2021). As shown in Figure 11, this compilation aggregates measurements of Vs30, three different Vs30 proxies (based on geology, terrain, and a hybrid approach), Vs profile data from compilations, statistics on Vs30 uncertainty, Geotechnical reports and Vs30 values determined from site databases of large-scale ground motion model development efforts (i.e., the most recent regional Next Generation Attenuation [NGA] Model projects (Bozorgnia et al., 2020; Goulet et al., 2018; Seyhan et al., 2014).

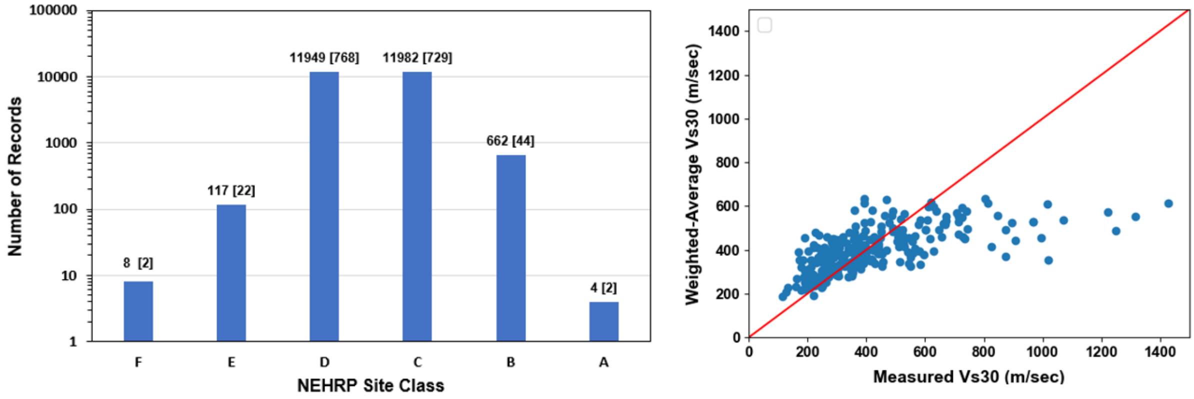
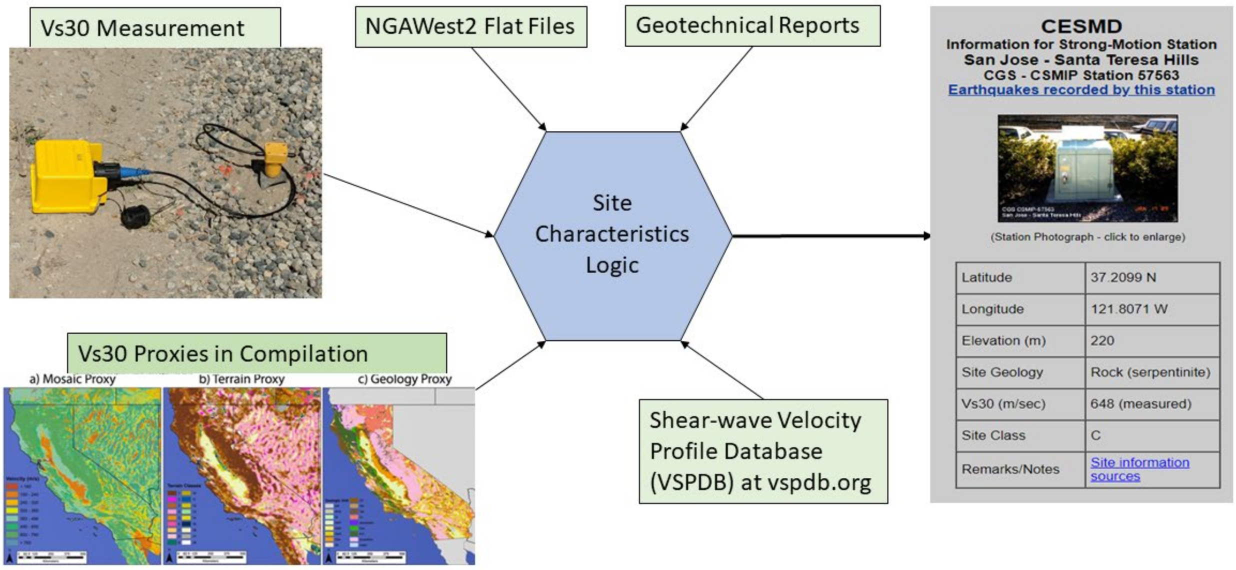


Figure 10. Left: Histogram showing the number of records at the CESMD EDC for each NEHRP site class from the number of stations shown in brackets. Right: Comparison between stations at the CESMD with measured Vs30 and the weighted average of 4 Proxy methods in NGA-West2.

Currently, the CESMD database contains representative information of site geology, Vs30 and NEHRP site class. The recent compilation includes extensive information per site compiled from multiple sources posing an interesting challenge to integrate and assign preferred site characteristic information. Criteria and an algorithm to automate this process are being developed. In addition to the preferred site information that goes to CESMD, users will also be provided with links to the full compilation found in Huddleston et al., 2021.



CESMD
Information for Strong-Motion Station
San Jose - Santa Teresa Hills
CGS - CSMIP Station 57563
Earthquakes recorded by this station

Latitude	37.2099 N
Longitude	121.8071 W
Elevation (m)	220
Site Geology	Rock (serpentinite)
Vs30 (m/sec)	648 (measured)
Site Class	C
Remarks/Notes	Site information sources

Figure 11. Diagram showing the sources of the collected data to be synthesized by a logic (in development) that assigns a representative Vs30 and Vs profile at station sites.

6. Special Studies

The availability of large amounts of recorded strong-motion data has allowed many researchers to perform studies addressing numerous specific topics of scientific interest. Results and conclusions derived from the same datasets may vary depending on the methodologies employed in the processing and interpretation of data. The CESMD has recently added “Special Studies” pages tab to the web portal (available at https://www.strongmotioncenter.org/specialstudies/) to serve as a repository for research-specific data releases which may in the future include synthetic data. The aim of the Special Studies page is to make datasets compiled and used in research projects that studied one or more earthquakes easily available to researchers who are interested in reproducing the results and/or in conducting their own independent study.

The CESMD Special Studies page currently has datasets from nine studies of earthquakes in Alaska, California, Hawaii, the Pacific Northwest, Utah, Great Valley, Portland, Reno, and Central and Eastern U.S.

7. Acknowledgements

All earthquake data stored at the Center for Engineering Strong-Motion Data (CESMD) are available at https://www.strongmotioncenter.org. We acknowledge the contribution of strong-motion data to the CESMD from multiple U.S. and international seismic networks enabling a larger geographic coverage of data collection. We acknowledge Han Shao for her complementary work on the CESMD data gaps and site statistics tools. We also thank the peer-reviewers for their valuable inputs and comments to improve this manuscript. Disclaimer: Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

8. References

Ancheta, T. D., Darragh, R. B., Stewart, J. P., Seyhan, E., Silva, W. J., Chiou, B.S.J., Wooddell, K. E., Graves, R. W., Kottke, A. R., Boore, D. M., Kishida, T., & Donahue J. L. (2014). 'NGA-West2 database', *Earthquake Spectra,* 30(3), pp. 989–1005. doi: https://doi.org/10.1193/070913EQS197.

Aagaard, B.T., Wald, D.J., Thompson, E.M., Hearne, M., & Schleicher, L.S. (2022). 'Improving the Development Pipelines for USGS Earthquake Hazards Program Real-Time and Scenario Products', *Proceedings of the 12th National Conference in Earthquake Engineering.* Earthquake Engineering Research Institute, Salt Lake City, Utah.

Boore, D.M., & Atkinson, G. (2008). 'Ground-motion prediction equations for the average horizontal component of PGA, PGV, and 5%-damped PSA at spectral periods between 0.01 s and 10.0 s', *Earthquake Spectra*, 24(1), pp. 99-138. doi: 10.1193/1.2830434.

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

Bozorgnia, Y., Stewart, J.P., Abrahamson, N.A., Ahdi, S.K., Ancheta, T.D., Archuleta, R.J., Atkinson, G.M., Boore, D.M, Boroschek, R., Campbell, K.W., Chiou, B.S.J., Contreras, V., Darragh R.B., Gregor, N., Gulerce, Z., Idriss, I.M., Ji, C., Kamai, R., Kishida, T., Kuehn, N., Kwak, D.Y., Kwok, A.O., Lin, P.S., Magistrale, H., Mazzoni, S., Muin, S., Midorikawa, S., Parker, G.A., Si, H., Silva, W.J., Walling, M., Wooddell, K.E., & Youngs, R.R. (2020). *Chapter 1: Introduction, in Data Resources for NGA-Subduction Project*. PEER Report No. 2020/02. https://peer.berkeley.edu/sites/default/files/2020_02_final_3.17.2020.pdf.

GEOVision Final Report (2016). *Surface Wave Measurements, California Strong Motion Instrumentation Program Stations, Riverside County, California.* GEOVision Project No. 16192.

Haddadi H., Shakal A., Stephens C., Savage W., Huang M., Leith W., Parrish J., & Borcherdt R. (2008). 'Center for Engineering Strong Motion Data (CESMD)'. *Proceedings of the 14th World Conference of Earthquake Engineering*, Beijing, China.

Hagos L., Haddadi H., Schleicher L., Steidl J., Gee L., & Dhar, M. (2022). 'Updates on the Center for Engineering Strong-Motion Data (CESMD)', *Proceedings of the 12th National Conference in Earthquake Engineering*, Earthquake Engineering Research Institute, Salt Lake City, UT. 2022.

Hatem, A., Reitman, N., Briggs, R., Gold, R., Thompson Jobe, J., and Burgette, R. (2022). 'Western U.S. Geologic Deformation Model for Use in the U.S. National Seismic Hazard Model 2023' *Seismological Research Letters*, 93 (6), pp. 3053–3067. Doi: https://doi.org/10.1785/0220220154.

Hearne, M., Thompson, E.M., Schovanec, H., Rekoske, J., Aagaard, B.T., & Worden, C.B. (2019). 'USGS automated ground motion processing software', USGS Software Release, doi: 10.5066/P9ANQXN3.

Huddleston, G.J., Schleicher, L., Knudsen, K., Gee, L., & Steidl, J. (2021). 'Compilation of Geologic and Seismic Velocity Characteristics at Advanced National Seismic System Strong Motion Accelerometer Sites', U.S. Geological Survey Data Release. doi: https://doi.org/10.5066/P9YKIUW2.

Huddleston, G., Schleicher, L., Blair, L., Knudsen, K., Gee, L, & Steidl, J. (2022). 'Interfacing the National Strong Motion Project (NSMP) Near-surface Site Characteristic Data Compilation at Advanced National

Seismic System Strong Motion Accelerometer Sites with Strong-Motion Waveform Datasets', *Proceedings of the 12th National Conference in Earthquake Engineering*, Earthquake Engineering Research Institute, Salt Lake City, Utah. 2022.

Jones, J., Kalkan, E., & Stephens, C. (2017). *Processing and Review Interface for Strong-Motion Data (PRISM)—Methodology and Automated Processing, Version 1.0.0*, U.S. Geological Survey Open-File Report No. 2017–1008. Doi: https://doi.org/10.3133/ofr20171008.

Kalkan, E., & Stephens, C. (2017) *Systematic comparisons between PRISM version 1.0.0, BAP, and CSMIP ground motion processing*, U.S. Geological Survey Open-File Report 2017-1020.

Open Geospatial Consortium. http://www.opengeospatial.org/standards/kml/.

Pagani, M., Monelli, D., Weatherill, G., Danciu, L., Crowley, H., Silva, V., Henshaw, P., Butler, L., Nastasi, M., Panzeri, L., Simionato, M., & Vigano, D. (2014). 'OpenQuake Engine: An Open Hazard (and Risk) Software for the Global Earthquake Model'. *Seismological Research Letters*, 85(3), pp. 692–702. Doi: https://doi.org/10.1785/0220130087.

Petralogix Engineering (2017). *Vs30 Site Characterization Report Los Angeles, Orange, Ventura, San Bernardino and Riverside Counties*. Project No. 2017-00006

Schleicher, L., Steidl, J., Thompson, E., Yong, A., Brody, J., Blair, L., Ahdi, S.K., Hearne, M., Aagaard, B., Hough, S.E., Shao, H., Huddleston, G., Heilpern, K., Marano, K., Gerragut, G., Worden, B., Wald, D., De Cristofaro, J., McClain, A., Dunham, B., Nget, D., Aragon, J., Gomez, J., Amador, V., Carrasco Rodriquez, V., Luna, E., Cembalski, D., Childs, D., Smith, J., Croker, D., & Gee, L. (2024). 'A Journey to the Center of the USGS National Strong-Motion Project Processing and Beyond', *Proceedings of the 18th World Conference on Earthquake Engineering*, Milan, Italy.

Seyhan, E., Stewart, J.P., Ancheta, T.D., Darragh, R.B., & Graves, R.W. (2014). 'NGA-West2 Site Database', *Earthquake Spectra*, 30(3), pp. 1007–1024. doi: https://doi.org/10.1193/062913EQS180M.

Shakal A., Haddadi H., & Reitz T. (2018). 'Developments in the CSMIP Strong-motion Automated Recovery and Analysis (SARA) System', *Proceedings of the 11th National Conference on Earthquake Engineering*, Earthquake Engineering Research Institute, Los Angeles, CA.

Shao, H., Brody, J., Schleicher, L., Marano, K., Hagos, L., Haddadi, H., Steidl, J., Thompson, E., & Hearne, M. (2024). 'International Data Gaps at the Center for Engineering Strong Motion Data', *World Conference on Earthquake Engineering 2024, Session - Exploring Earthquake Data Centers around the World*.

Steidl, J., Hegarty, P., Schleicher, L., Brody, J., & Gee, L. (2022). 'Modernization of Data Processing and Review at the U.S. Geological Survey's National Strong Motion Project (NSMP)', *Proceedings of the 12th National Conference in Earthquake Engineering*, Earthquake Engineering Research Institute, Salt Lake City, Utah. 2022.

Thompson, E, Wald, D. & Worden, C. A (2014). 'Vs30 Map for California with Geologic and Topographic Constraints', *Bulletin of the Seismological Society of America*, 104, pp. 2313-2321. https://doi.org/10.1785/0120130312, 2014.

U.S. Geological Survey, Earthquake Hazards Program (2017). Advanced National Seismic System (ANSS) Comprehensive Catalog of Earthquake Events and Products: Various, https://doi.org/10.5066/F7MS3QZH.

Wald D.J., Worden C.B., Thompson E.M., & Hearne M. (2021). 'ShakeMap operations, policies, and procedures'. *Earthquake Spectra*. 38(1), pp. 756-777. doi: https://doi.org/10.1177/87552930211030298.

Worden, C. B., Hearne, M., & Thompson, E. M. (2018). 'ShakeMap v4 Software', USGS Software Release. doi:10.5066/P97FHE0I.

Worden C.B., Thompson, E.M., Hearne, M. & Wald, D.J. (2020). 'ShakeMap Manual Online: Technical Manual, User's Guide, and Software Guide'. doi: https://doi.org/10.5066/F7D21VPQ.

Yong, A., Martin, A., Stokoe, K., & Diehl, J. (2013). *ARRA-Funded Vs30 Measurements Using Multi-Technique at Strong-Motion Stations in California and Central-Eastern United States*. USGS Open-File Report No. 2013-1102.