Title

User-interaction with a web-served global ground motion relational database

Permalink

https://escholarship.org/uc/item/9tw6k14c

Authors

Buckreis, Tristan
Nweke, CC
Wang, Pangfei
et al.

Publication Date

2024-07-01

Peer reviewed

USER-INTERACTION WITH A WEB-SERVED GLOBAL GROUND MOTION RELATIONAL DATABASE

T.E. Buckreis[1], C.C. Nweke[2], P. Wang[3], S.J. Brandenberg[4], R. Shams[5], M.E. Ramos-Sepulveda[6], R. Pretell[7], S. Mazzoni[8], & J.P. Stewart[4]

[1] Postdoctoral Scholar, University of California, Los Angeles, Los Angeles, CA, USA, tristanbuckreis@ucla.edu

[2] Assistant Professor, University of Southern California, Los Angeles, CA, USA

[3] Assistant Professor, Old Dominion University, Norfolk, VA, USA

[4] Professor, University of California, Los Angeles, Los Angeles, CA, USA

[5] Ph.D. Student, University of Southern California, Los Angeles, CA, USA

[6] Ph.D. Candidate, University of California, Los Angeles, Los Angeles, CA, USA

[7] Assistant Professor, University of Nevada, Reno, Reno, NV, USA

[8] Project Scientist, University of California, Los Angeles, Los Angeles, CA, USA

Abstract: *We present an application programming interface (API) which facilitates public access to a global relational database of earthquake ground motion intensity measures, associated metadata, and time-series data. Next Generation Attenuation (NGA)-East and NGA-West2 project spreadsheets have been adapted into a relational database format composed of multiple tables through a series of primary and foreign keys. The combined dataset has been expanded to include contributions from earthquakes, generally with magnitudes greater than* ***M****3.9, that have occurred since the conclusion of the data synthesis component of both projects in 2011. Currently the database includes 62,449 ground motions recorded at 9,092 stations for 899 events. The database is accessible through an API, which allows users to interact with and query the database directly without detailed knowledge of structure query language (SQL). Simple queries are constructed by appending relatively straightforward query string parameters to the end of a uniform resource location (URL) that serves as an endpoint, which returns only data that satisfy the query constraints. The web-served nature of the database means that users have immediate access to ground motion data as soon as it is collected, reviewed, and uploaded. Furthermore, integrated end-to-end workflows – which do not require files to be downloaded and saved in local memory – are made possible through the API. The structure of the database has been designed to accommodate growth, with ongoing efforts to integrate global ground motion data in anticipation of the NGA-West3 project, and improve ease-of access through the API.*

1. Introduction

There have been many efforts to develop collections of ground motion data and metadata, which include the Next Generation Attenuation (NGA)-West2 (Bozorgnia et al. 2014),NGA-East (Goulet et al. 2021), and NGA-Subduction (Mazzoni et al. 2022) projects for shallow-crustal events in active-tectonic regions, crustal events in stable continental regions, and subduction events, respectively. The metadata compiled in these projects, while broadly similar, have differences that reflect considerations that may be unique for a given tectonic regime. Ground motion data are provided as intensity measures (e.g., peak metrics and pseudo-spectral accelerations at specified damping levels: PSA). The ground motion data and metadata from these and other similar projects are provided as downloadable files, and the time-series are also available for download in increments through publicly-accessible web sites (e.g., https://ngawest2.berkeley.edu/).

The nature of the data management and data releases has evolved over time to suit the needs of the individual projects. For example, NGA-West2 data were organized through the flatfile, where updates were entered. The flatfile size is about 45 MB, which is large but manageable to work with on most PC's. The NGA-Subduction database was managed as a relational database, with data disseminated from the database in a flatfile format and as separate tables that reflect the organizational structure of the database. Managing the data using a relational database greatly enhances data integrity because it eliminates the need for duplicate entry. For example, if an earthquake magnitude is updated, it is changed in a single entry in the appropriate table. By contrast, the same magnitude appears in many locations in the flatfile, raising the potential for conflicting entries. The flatfile released for the NGA-Subduction project is 130 MB, which is quite large and cumbersome to work with. This is a primary reason why data were released as separate tables.

The proposed work is an extension of the NGA efforts that allows users to create custom flatfiles and integrate ground motion data into end-to-end workflows. We recognize ground motion databases are likely to become too large to facilitate model development efforts from a single flatfile that contains all relevant fields. Furthermore, we recognize that many users do not need all of the data contained within the flatfile, and customized data requests could result in less memory demand and more efficient interactions. Our goal is to create a resource where users can query the data quantities they wish to retrieve in a customized output format for integration into their workflow.

To achieve this goal, we developed a relational ground motion database (GMDB) that has been populated to-date with the NGA-East and NGA-West2 datasets (but not the time-series), in addition to data from more recent events (including time-series) that was compiled prior to the onset of the NGA-West3 project. Details of the GMDB are provided in Buckreis et al. (202x), and a high-level overview of the data and organizational structure is presented following this introduction. The second issue led us to develop a representational state transfer (REST) application programming interface (API), which acts as an intermediary between the user and the database. Our API enables users to request a specific subset of the data they need from the GMDB through simple web-services, without necessarily downloading files or loading the entire database into local memory. This paper presents the API as of October 2023, and illustrates how it can be used to interact with data stored in the GMDB.

2. Ground Motion Data

The GMDB includes the global database for shallow crustal earthquakes in active tectonic regions from the NGA-West2 project (Bozorgnia et al. 2014) and the database for central and eastern North America for shallow crustal earthquakes in stable continental regions from the NGA-East project (Goulet et al. 2021). The authors have developed ground motions in a consistent manner to supplement the NGA project datasets for a number of events since completion of the NGA database projects around 2011. These studies include site response and path studies in southern and northern California (Wang 2020, Nweke et al. 2022, and Buckreis et al. 2023a), studies to assess the combined bias of NGA-East GMMs with site response models (Ramos-Sepulveda et al. 2023), studies to assess the usable-ranges of data recorded by the Community Seismic Network (Stewart et al. 2023), and ground motion analysis for the 2014 Napa earthquake, 2019 Ridgecrest earthquake sequence, and 2023 Kahramanmaraş earthquake sequence (Kishida et al. 2016, Ahdi et al. 2020, and Buckreis et al. 2023b, respectively). Table 1 presents a summary of the data compiled in each data collection effort. Data compiled as part of recent efforts are grouped into "Added-SCR" (SCR = stable continental region) and "Added-ACR" (ACR = active crustal region), to readily associate these new data to prior NGA database products.

Table 1. Summary of original sources of compiled ground motion data.

Data Collection	References	No. of Events	No. of Stations	No. of Records
NGA-East	Goulet et al. (2021)	82	1,271	9,376
Added-East	Ramos-Sepulveda et al. (2023)	100	1,482	6,892
NGA-West2	Bozorgnia et al. (2014)	599	4,149	21,513
Added-West	Ahdi et al. (2020)	3†	82	1,483
	Buckreis (2022)	70	1,313	9,879
	Buckreis et al. (2023b)	4	493	1,181
	Kishida et al. (2016)	1	419	419
	Stewart et al. (2023)	26†	1,071	5,545
	Wang (2020)	28†	972	6,211

†Some overlap in the number of events; Stewart et al. (2023) processed data for different stations.

Figure 1 presents epicenter locations for all earthquakes with data in the GMDB. Added events (blue and orange symbols) range in magnitude from **M**3.9 to **M**7.8, and occurred between 2011 and 2023. It is important to note that the added events in California do not include all **M**3.9+ events in and around California since 2011 – the work to date was motivated by specific station-related data needs rather than a desire to produce a comprehensive dataset. The number of shallow-crustal events in active tectonic regimes (i.e., ACR events) has grown from 599 to 717, and the number of SCR earthquakes has grown from 82 to 182, when compared to their corresponding NGA databases.

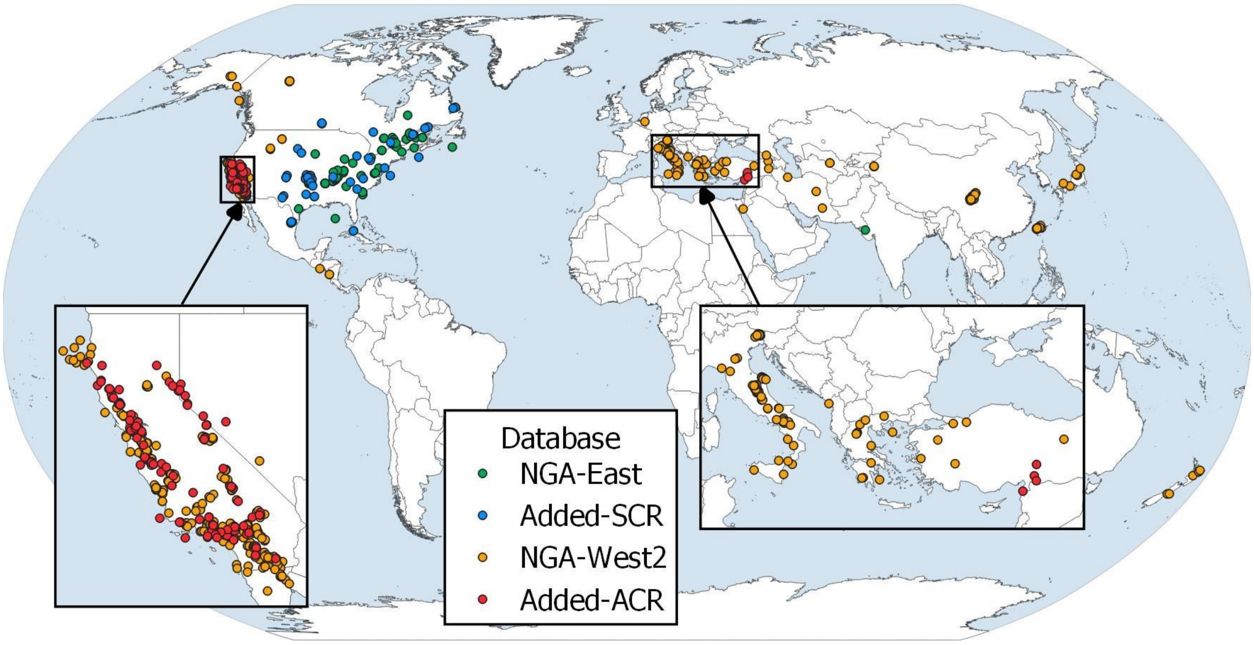


Figure 1. Map of earthquakes included in the GMDB. Symbol colors represent data collection effort: green = NGA-East, blue= Added-SCR, orange = NGA-West2, and red = Added-ACR.

Raw ground motions for all "Added" events were downloaded from the Incorporated Research Institutions for Seismology (IRIS; Trabant et al. 2012), the Center for Engineering Strong Motion data (CESMD; (https://www.strongmotioncenter.org), the Earthquake Data Center System of Türkiye (TDVMS; https://tdvms.afad.gov.tr), or directly from local networks (e.g., seismic network maintained by the California Department of Water Resources). The data were screened for duplicate recordings and for meaningful signals (above the noise threshold). Recent versions of standard PEER/NGA signal processing methods (e.g., Goulet et al. 2021) were used to provide uniformly processed time-series with high- and low-pass corner frequencies selected on a component-by-component basis using either fully-manual or semi-automated (e.g., Ramos-Sepulveda et al. 2023) procedures which involve human-inspection of each component of every record.

Figure 2 presents the data distributions of the NGA and Added contributions with respect to magnitude and rupture distance. Metadata for the Added datasets was assigned according to standard protocols developed within NGA projects, as described in Goulet et al. (2021), Ahdi et al. (2022), and Contreras et al. (2022), and as outlined for application in the GMDB in Buckreis et al. (202x). The datasets have significantly expanded for events with 4 < **M** < 5, and several significant earthquakes have also been added: **M**6.02 August 24, 2014 South Napa; **M**6.48 July 4, 2019 Searles Valley; **M**7.06 July 6, 2019 Ridgecrest; **M**6.49 May 15, 2020 Monte Cristo Range, Nevada; **M**7.8 February 6, 2023 Pazarcik, Türkiye; **M**6.8 February 6, 2023 Nurdağı, Türkiye; **M**7.7 February 6, 2023 Elbistan, Türkiye; and **M**6.3 February 20, 2023 Yayladağı, Türkiye. The majority of the data are recorded at far distances (≥ 100 km), with most of the near-source data coming from the NGA-West2 database, with the exception of the **M**7.8 February 6, 2023 Pazarcik earthquake.

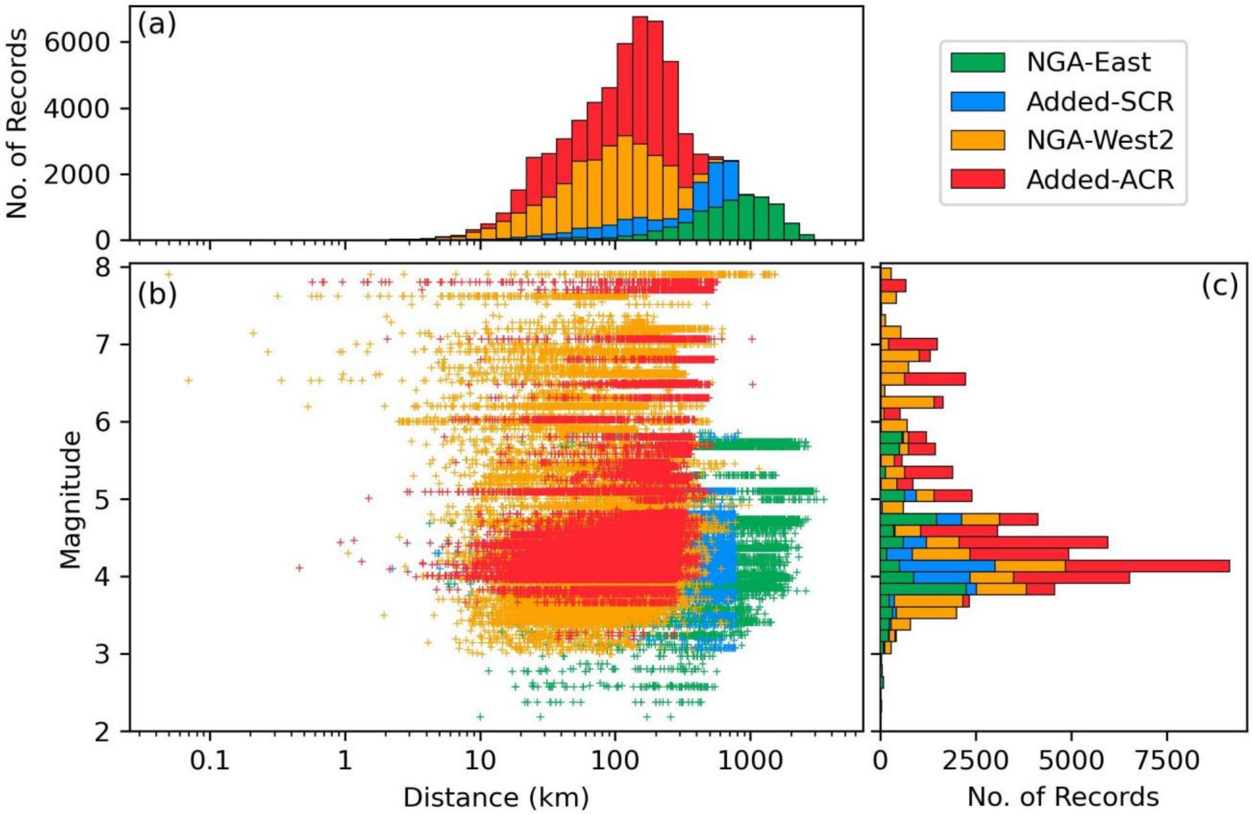


Figure 2. Distribution of data in the GMDB: (a) histogram of number of records per distance bin; (b) scatter-plot in magnitude vs distance space; and (c) histogram of number of records per magnitude bin. Colors represent data collection effort: green = NGA-East, blue= Added-SCR, orange = NGA-West2, and red = Added-ACR.

In total, the GMDB is comprises 62,499 three-component ground motions recorded at 9,092 stations for 899 events. A breakdown of the original-sources of data is provided in Table 1. Intensity measures such as peak ground acceleration (PGA), peak ground velocity (PGV), 5%-damped PSA, Arias Intensity (I_A; Arias 1970), times corresponding to select percentiles of I_A used to facilitate calculation of significant durations (I_A-times), cumulative absolute velocity (CAV; EPRI 1988), CAV computed from acceleration time-histories considering a threshold of 5 cm/s^2 (CAV$_5$; Kramer and Mitchell 2006), and the orientation-independent horizontal component Fourier amplitude spectra (effective amplitude spectra – EAS; Kottke et al. 2021) are computed for all motions. Only the “Added” time-series are stored in the GMDB.

3. Relational Database

The organizational structure of the database – the schema – defines the tables, fields, and relationships among the tables in the database. The current GMDB schema contains 32 tables that can be broadly grouped into five categories: event, site-station, ground motion, auxiliary, and junction. A diagram depicting the tables and their relationships is presented in Figure 3, where arrows indicate the relationships between tables. Relationships between tables are identified by shared fields called “keys”, with a “primary key” consisting of one or more fields that uniquely identify a table entry, and a “foreign key” consisting of one or more fields in a separate table that identify the entry in the table containing the corresponding primary key. For example, each event and station are assigned a unique “event_id” and “station_id” in the “event” and “station” tables (primary keys), respectively, which are used to identify the event and station metadata for a particular ground motion in the “motion” table (where they are foreign keys).

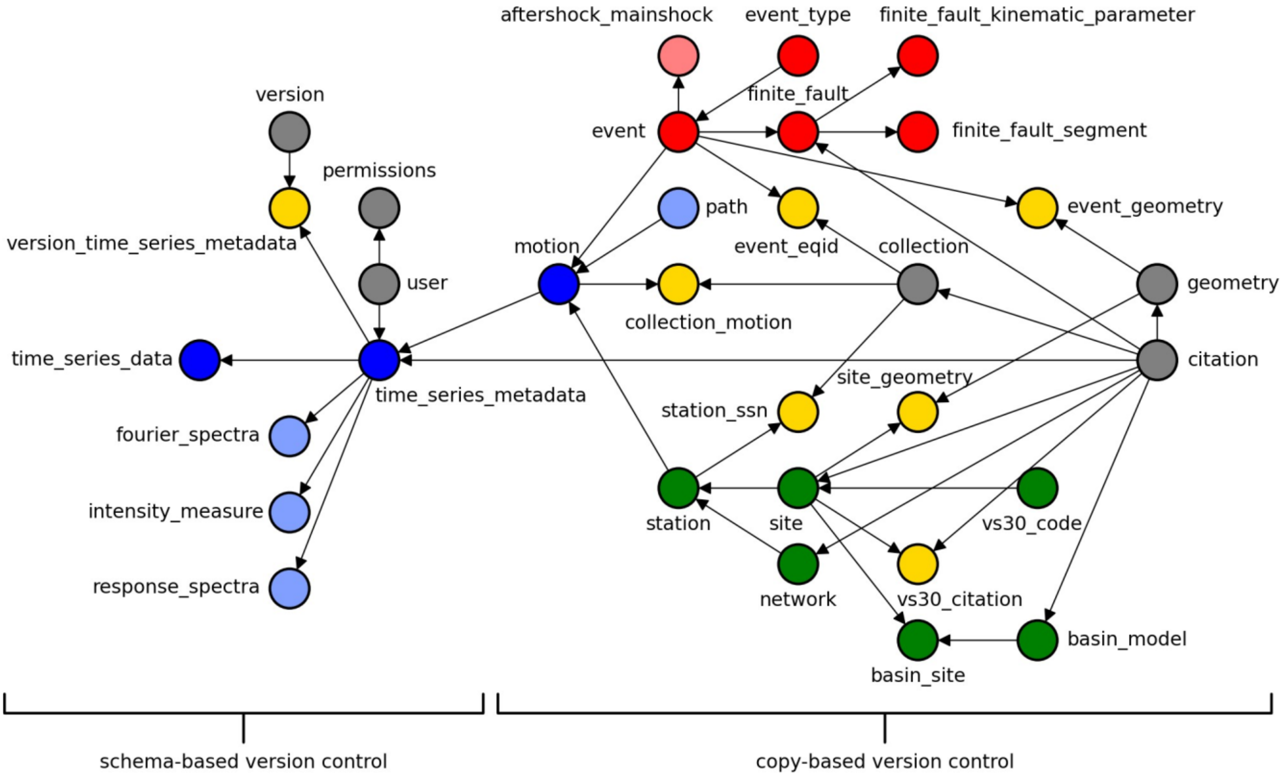


Figure 3. Diagram of simplified GMDB schema depicting primary and foreign key relationships; event tables colored red; station-site tables colored green; ground motion tables colored blue; auxiliary tables colored gray; and junction tables colored yellow. Lighter shades (e.g., path table) indicate tables whose contents are computed from values stored in other tables. Arrow direction indicates direction of foreign key dependency (e.g., event_id, path_id, and station_id are foreign keys in the motion table). [from Buckreis et al. 202x]

“Event” tables contain event-related metadata including finite-fault representations; “station-site” tables store information about the physical attributes and metadata related to the location of the recording instrument; “ground motion” tables store path metrics, computed intensity measures, and time-series data; “auxiliary” tables store information that is not inherently related to ground motion data (e.g., citations, user information, version control, etc.); and “junction” tables store many-to-many relationships between two tables. A comprehensive presentation of each table and the version control is given in Buckreis et al. (202x), and descriptions for all of the fields in each table are provided at https://doi.org/10.34948/G4RP4K.

4. Application Programming Interface

One of the primary objectives of the effort to create the GMDB was to make ground motion data and associated metadata publicly accessible and easy user-friendly. To accomplish this objective, we developed a representational state transfer (REST) application programming interface (API) and an online tool to assist users with accessing and interacting with the data. Our API enables users to request data from the GMDB using relatively straightforward query string parameters appended to the end of a uniform resource location (URL) that serves as an endpoint. Using the API to retrieve data requires only a basic understanding of the database structure (i.e., names of tables and fields) and no knowledge of structure query language (SQL) – which otherwise represent significant obstacles for most users. Key features of the API are explained in this document, and complete documentation is provided by Buckreis et al. (202x).

The URL used to request data through the API contains the following three elements, as described in Figure 4: (1) the “base URL” utilized by the API, (2) a “service” indicator which describes the nature of the query, and (3) a “query string” to specify the query constraints and data format. Supported services are “schema” to obtain information about the organizational structure and field descriptions, “flatfile” to extract flatfile-style tables of data, and “timeseries” to download time-series data (currently under development).

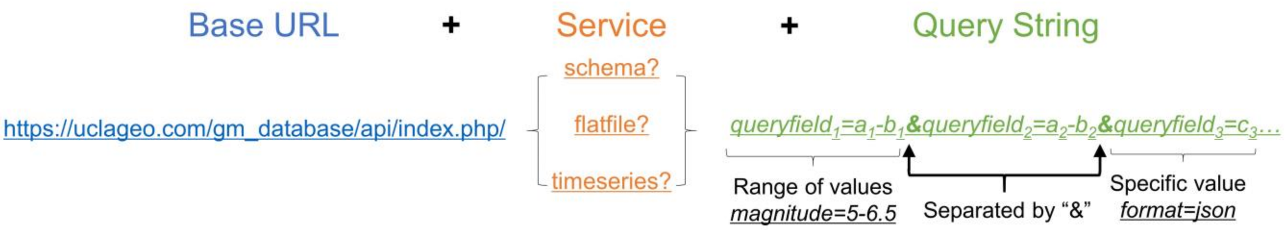


Figure 4. Schematic of how to format a valid uniform resource location (URL) for the GMDB API: (blue) base URL included with all API-requests; (orange) service indicator; and (green) optional query string.

When the user enters a URL, it is passed through the API to the serve where the service and query string parameters are parsed, and a SQL query is subsequently constructed and executed. The SQL queries are highly optimized and can be quite complicated. For example, the query optimizers in MySQL perform early row lookups, which means that if users request 50 records beginning at row 10,000, the first 10,050 records will be retrieved and the first 10,000 will then be discarded. This can cause significant performance problems, particularly for tables that contain many fields and/or fields that occupy significant memory (e.g., time-series data). To overcome this problem, we first create a temporary table that stores only the primary keys from the tables we wish to query. This table is very light, involving only indexed fields, and the early row lookup issue does not cause a significant performance problem. We then join this temporary table to the other tables to retrieve only the data we need. These optimizations render complicated queries that make the API much more efficient than would generally be achieved if users were permitted to directly craft their own SQL queries because most users, even those who have a working understanding of SQL, do not understand these limitations of the query optimizers.

Query strings are service-dependent. The flatfile service has the greatest number of valid query string parameters, because it is the primary means to query data through the API. Users can condition queries on individual fields included in the GMDB. For example, if a user wishes to query a flatfile for records from large magnitude earthquakes with short rupture distances, they can append “&magnitude=6-8&rrup=0-15” to the query string. Additional query string parameters can be used to control the output format, limit the number of rows, order the data by specific fields, and other functions. A complete list of the valid query string parameters for each service documentation can be found at https://doi.org/10.34948/G4RP4K.

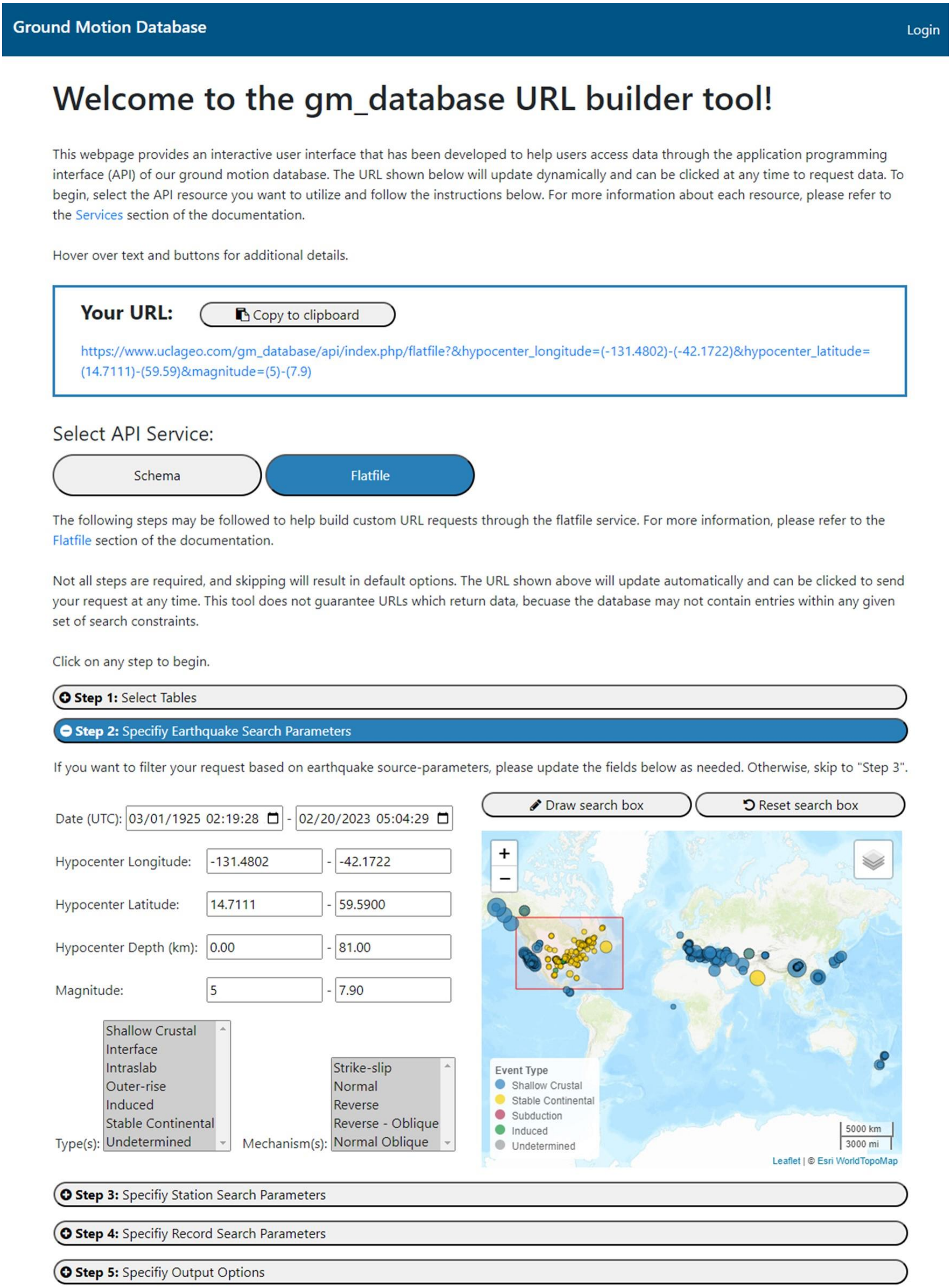
5. Online Tool

The API is relatively straightforward, but requires users to have basic knowledge of the tables and fields in the database, and a working knowledge of the API’s query string parameters. To help users construct URL’s we developed a URL Builder Tool available at https://doi.org/10.34948/G4RP4K, which currently supports the “schema” and “flatfile” services. The advantage of this format is that users do not need any pre-installed software to use the tool, as long as they have a web browser and an internet connection. The webpage provides an interactive user interface to construct URL’s, as illustrated in Figure 5. Steps to using the tool are presented below:

1. Users first select the tables they want to query. If the user does not specify any tables, then a default set of tables is used.

2. If the request requires screening based on earthquake related metadata (e.g., origin time, location, magnitude), the user can specify constraints. This form includes an interactive map where the user can preview event-specific metadata, and dynamically draw a coordinate bounding box, as shown in Figure 5.
3. If the request requires screening based on station-site related metadata (e.g., location, time-averaged shear wave velocity in the upper 30 m – V_{S30}), the user can specify constraints. This form also includes an interactive map of the stations.
4. If the request requires screening based on information about the ground motion records themselves (e.g., distance, peak ground acceleration – PGA, components), the user can specify constraints.
5. The user can customize the returned data structure (i.e., format, sorting, limit).

As illustrated in Figure 5, the URL is updated dynamically as the forms are populated. Individual fields within each form of the tool are automatically populated with the parametric ranges of data currently stored in the GMDB, and include quality checks to ensure that the constraints are valid (e.g., longitude values must be between -180º and 180º; minimum bounds are less than maximum bounds). Users are not required to complete each step, and if any steps are skipped the API will use default parameters. The tool does not guarantee URL’s that return data, because the database may not contain entries within any given set of search constraints (e.g., M=10-11 would return no data).



Ground Motion Database

Welcome to the gm_database URL builder tool!

This webpage provides an interactive user interface that has been developed to help users access data through the application programming interface (API) of our ground motion database. The URL shown below will update dynamically and can be clicked at any time to request data. To begin, select the API resource you want to utilize and follow the instructions below. For more information about each resource, please refer to the Services section of the documentation.

Hover over text and buttons for additional details.

Your URL:

https://www.uclageo.com/gm_database/api/index.php/flatfile?&hypocenter_longitude=(-131.4802)-(-42.1722)&hypocenter_latitude=(14.7111)-(59.59)&magnitude=(5)-(7.9)

Select API Service:

The following steps may be followed to help build custom URL requests through the flatfile service. For more information, please refer to the Flatfile section of the documentation.

Not all steps are required, and skipping will result in default options. The URL shown above will update automatically and can be clicked to send your request at any time. This tool does not guarantee URLs which return data, becuase the database may not contain entries within any given set of search constraints.

Click on any step to begin.

Step 1: Select Tables

Step 2: Specifiy Earthquake Search Parameters

If you want to filter your request based on earthquake source-parameters, please update the fields below as needed. Otherwise, skip to "Step 3".

Date (UTC): 03/01/1925 02:19:28 - 02/20/2023 05:04:29

Hypocenter Longitude: -131.4802 - -42.1722

Hypocenter Latitude: 14.7111 - 59.5900

Hypocenter Depth (km): 0.00 - 81.00

Magnitude: 5 - 7.90

Type(s): Shallow Crustal, Interface, Intraslab, Outer-rise, Induced, Stable Continental, Undetermined

Mechanism(s): Strike-slip, Normal, Reverse, Reverse - Oblique, Normal Oblique

Step 3: Specifiy Station Search Parameters

Step 4: Specifiy Record Search Parameters

Step 5: Specifiy Output Options

Figure 5. Screenshot of GMDB API URL Builder Tool flatfile service options. The tool is available at https://doi.org/10.34948/G4RP4K.

6. Interacting with the GMDB

The default data format returned by the API is an HTML table that is viewable in a web browser, which is useful for viewing the data, but not for working with it. To facilitate interaction with the data, URL's can be configured to download files in a comma-separated value (CSV) format (i.e., "format=csv"), or return the data in JavaScript Object Notation (JSON) format (i.e., "format=json"). CSV files are useful for users who wish to save and work with the data locally, however this necessitates that the file be read and parsed, even if only a subset of the data is actually used, and may result in files becoming out-of-date as new data are added to the database. However, the CSV format is useful for reproducibility since the database is constantly evolving. The JSON format is useful for users who wish to directly work with the returned data in an integrated end-to-end computing workflow without saving files to local memory, and/or perform many different queries within their workflow.

An integrated workflow can be used to circumnavigate computations on memory deficient machines (i.e., if there is insufficient allotable memory to save CSV or Excel files). Historically, this has not been an issue for most earthquake ground motion related studies, however as database sizes grow flatfiles become less efficient. Furthermore, if users require access to recent data, integrated workflows do not require additional effort, whereas new CSV files must be downloaded each time the database is updated. Pre-built packages, such as "requests" in Python and "httr" in R, can be used to query data within a programming environment, thus eliminating the need to store data locally (provided there is sufficient Random-Access Memory – RAM).

To highlight the advantage of an integrated workflow relative to traditional file-read-and-write workflow, we compare the time needed to query the first 1,000 rows of data from **M**5.0+ events in the NGA-West2 and GMDB databases. In the traditional workflow (Figure 6a), it takes approximately 58 sec to read and screen the data, whereas the integrated workflow takes only 5 sec to query the same amount of data (Figure 6b). The large difference in time is caused by the traditional requiring the entire flatfile (~45 MB) to be read and parsed, before screening and returning the first 1,000 rows, whereas the integrated workflow queries only the data that satisfies the search criteria (i.e., first 1,000 entries). It is important to note that the server-to-client communication comprises the majority of the time in the integrated workflow. Of course if the file size is small, then the traditional approach may perform just as well as the integrated approach through the API.


a)

```
%%time
import pandas as pd
DataFrameFull = pd.read_excel(r"G:\My Drive\NGA\Updated_NGA_West2_Flatfile_RotD50_d050_public_version.xlsx")

DataFrameLargeMag = DataFrameFull[DataFrameFull['Earthquake Magnitude'] >= 5.0] # Screen for records from events with M > 5
DataFrame1000 = DataFrameLargeMag.reset_index().loc[0:999]                      # Return 1,000 rows
```

```
CPU times: total: 17.8 s
Wall time: 58.6 s
```

b)

```
%%time
# Construct Uniform Resource Locator (URL):
URL = "https://www.uclageo.com/gm_database/api/index.php/flatfile?" # Base URL for flatfile service
URL += "tables=motion,event,path,station,site,response_spectra"     # List tables with desired data/metadata
URL += "&component=rotd50"                                          # Only interested in RotD50 intensity measures and PSA
URL += "&magnitude=5-10"                                            # Only interested in large magniutude (>= 5.0) data
URL += "&limit=1000&offset=0"                                       # Return 1,000 rows, starting with the first (offset=0)
URL += "&format=json"                                               # Return data as a JSON string

# Retrieve data through web-services request via API:
import requests, json
r= requests.get(URL, headers={"User-Agent":"XY"})                   # "Get" web-services request (querying through the API)
JsonObject = json.loads(r.text)                                     # JSON dictionary-like object

# Convert JsonObject to a Pandas DataFrame Object:
import pandas as pd
DataFrame = pd.DataFrame.from_dict(JsonObject)                      # Convert JSON dicrtionary-like object to DataFrame

DataFrame.head(5)
```

```
CPU times: total: 31.2 ms
Wall time: 4.76 s
```

*Figure 6. Screenshot of Python codes to query the first 1,000 rows of **M**5.0+ data using a) traditional approach of loading data from the NGA-West2 flatfile in local memory – total elapsed time = 58.6 s; b) integrated approach through the API – total elapsed time = 4.76 s.*

7. Summary and Conclusions

The GMDB marks a significant evolution from past practices for disseminating ground motion data, because data are available through a web-served database via an API. Our API enables users to request data from the GMDB in HTML, CSV, or JSON formats using relatively straightforward query string parameters appended to the end of a URL. Ground motion data are actively being curated and added to the GMDB, which currently contains 62,499 ground motions from 899 events assembled by NGA-East, NGA-West2, and more recent efforts including studies of the 2023 Kahramanmaraş earthquake sequence in Türkiye (Buckreis et al. 2023b). Users who wish to work with the most up-to-date data can do so through an integrated end-to-end workflow made possible by the API. In addition to repeatedly providing the most up-to-date data, this workflow does not require large amounts of allocable memory (needed to save large files to local machines), and is often more efficient than reading and parsing data from downloaded flatfiles. Current services supported by the API include "schema" for information about the organizational structure and field descriptions, and "flatfile" to query NGA-style flatfiles of ground motion data and related metadata. We are working on developing a "timeseries" service to return time-series data, and also continue to make improvements to the "flatfile" service to increase efficiency and reduce query times.

8. References

Ahdi S.K., Mazzoni S., Kishida T., Wang P., Nweke C.C., Kuehn N.M., Contreras V., Rowshandel B., Stewart J.P., Bozorgnia Y. (2020). Engineering characteristics of ground motions recorded in the 2019 Ridgecrest Earthquake Sequence, *Bulletin of the Seismological Society of America*, 110: 1474-1494.

Ahdi S.K., Kwak D.Y., Ancheta T.D., Contreras V., Kishida T., Kwok A.O.L., Mazzoni S., Ruz F., Stewart J.P. (2022). Site parameters applied in NGA-Sub database, *Earthquake Spectra*, 38(1), 494-520.

Arias A. (1970). A measure of earthquake intensity, in Hansen, R., ed., *Seismic design for nuclear power plants*: Cambridge, Mass., Massachusetts Institute of Technology Press, 438-483.

Bozorgnia Y., Abrahamson N.A., Al Atik L., Ancheta T.D., Atkinson G.M., Baker J.W., Baltay A., Boore D.M., Campbell K.W., Chiou B.S.-J., Darragh R., Day S., Donahue J., Graves R.W., Gregor N., Hanks T., Idriss I.M., Kamai R., Kishida T., Kottke A., Mahin S.A., Rezaeian S., Rowshandel B., Seyhan E., Shahi S., Shantz T., Silva W., Spudich P., Stewart J.P., Watson-Lamprey J., Wooddell K., Youngs R. (2014). NGA-West2 research project, *Earthquake Spectra*, 30: 973-987.

Buckreis T.E. (2022). Customization of path and site response components of global ground motion models for application in Sacramento-San Joaquin Delta region of California, *Ph.D. Thesis*, UC Los Angeles.

Buckreis T.E., Nweke C.C., Wang P., Brandenberg S.J., Ramos-Sepulveda M.E., Shams R., Pretell R., Mazzoni S., Zimmaro P., Stewart J.P. (202x). A global application programming interface (API)-enabled earthquake ground motion relational database, *Earthquake Spectra*, (submitted).

Buckreis T.E., Stewart J.P., Brandenberg S.J., Wang P. (2023a). Subregional anelastic attenuation model for California, *Bulletin of the Seismological Society of America*, DOI: 10.1785/0120220173.

Buckreis T., Güryuva B., İçen A., Okcu O., Altindal A., Aydin M., Pretell R., Sandikkaya A., Kale Ö., Askan A., Brandenberg S., Kishida T., Akkar S., Bozorgnia Y., Stewart J. (2023b). Ground Motion Data from the 2023 Türkiye-Syria Earthquake Sequence, *DesignSafe-CI*., DOI: 10.17603/ds2-t115-bk16.

Contreras V., Stewart J.P., Kishida T., Darragh R.B., Chiou B.S.J., Mazzoni S., Youngs R.R., Kuehn N.M., Ahdi S.K., Wooddell K., Boroschek R., Rojas F., Ordenes J. (2022). NGA-Sub source and path database, *Earthquake Spectra*, 38(2), 799-840.

Electrical Power Research Institute (EPRI) (1988). A Criterion for Determining Exceedance of the Operating Basis Earthquake, *EPRI NP-5930*, Palo Alto, CA.

Goulet C.A., Kishida T., Ancheta T.D., Cramer C.H., Darragh R.B., Silva W.J., Hashash Y.M.A., Harmon J., Parker G.A., Stewart J.P., Youngs R.R. (2021). PEER NGA-East database, *Earthquake Spectra*, 37: 1331-1353.

Kishida T., Mazzoni S., Bozorgnia Y., Chiou B., Darragh R., Haddadi H.R., Kayen R., Markham C., Muin S. Silva W. (2016). Investigation of ground motions recorded during the 2014 South Napa earthquake, *SMIP16 Seminar*, October 2016.

Kottke A.R., Abrahamson N.A., Boore D.M., Bozorgnia Y., Goulet C.A., Hollenback J., Kishida T., Ktenidou O.J., Rathje E.M., Silva W.J., Thompson E.M., Wang X. (2021). Selection of random vibration theory procedures for the NGA-East project and ground motion modelling, *Earthquake Spectra*, 37: 1420-1439.

Kramer S.L., Mitchell R.A. (2006). Ground motion intensity measures for liquefaction hazard evaluation, *Earthquake Spectra*, 22(2): 413-438.

Mazzoni S., Kishida T., Stewart J.P., Contreras V., Darragh R.B., Ancheta T.D., Chiou B.-S.J., Sliva W.J., Bozorgnia Y. (2022). Relational databased used for ground-motion model development in the NGA-sub project, *Earthquake Spectra*, 38(2): 1529-1548.

Nweke C.C., Stewart J.P., Wang P., and Brandenberg S.J. (2022). Site response of sedimentary basins and other geomorphic provinces in southern California, *Earthquake Spectra*, 38(4): 2341-2370.

Ramos-Sepulveda M.E., Parker G.A., Buckreis T.E., Brandenberg S.J., Thompson E.M., Li M., Ilhan O., Rathje E.M., Hashash Y.M.A., Stewart J.P. (2023). Update of NGA-East database to include Central and Eastern North America events since November 2011, *SSA 2023 Annual Meeting*, San Juan, Puerto Rico, April 17-20, 2023.

Ramos-Sepulveda M.E., Parker G.A., Thompson E.M., Brandenberg S.J., Li M., Ilhan O., Hashash Y.M.A., Rathje E.M., Stewart J.P. (2023). High-pass corner frequency selection for implementation in the USGS automated ground motion processing tool, *Geo-Congress 2023: Geotechnical Data Analysis and Computation*, Los Angeles, CA, March 2023, Geotechnical Special Publication No. 342, EM Rathje, B Montoya, and MH Wayne (Eds.), 327-335, ASCE Geo-Institute.

Stewart J.P., Mohammad S., Nweke C.C., Shams R., Buckreis T.E., Kohler M.D., Bozorgnia Y. (2023). Usability of ground motions recorded by community seismic network, *Report GIRS-2023-08*, B. John Garrick Risk Institute, Natural Hazards Risk and Resiliency Research Center, UCLA (Center Headquarters), 63 pages.

Trabant C., Hutko A.R., Bahavar M., Karstens R., Ahern T., Aster R. (2012). Data Products at the IRIS DMC: Stepping Stones for Research and Other Applications, *Seismological Research Letters*, 83(5): 846–854.

Wang P. (2020). Predictability and Repeatability of Non-Ergodic Site Response for Diverse Geological Conditions, *Ph.D. Thesis*, UC Los Angeles.