Cal Poly

Caltech

SCEC
AN NSF+USGS CENTER

UC Irvine

UCLA

UC Santa Barbara

USC

Fault Displacement Hazard Initiative Database

Alexandra Sarmiento, P.E., C.E.G.
Danielle Madugo
Yousef Bozorgnia, Ph.D, P.E.
Andi Shen
Silvia Mazzoni, Ph.D.
University of California, Los Angeles

Grigorios Lavrentiadis, Ph.D.
University of California, Berkeley and Los Angeles

Timothy Dawson, C.E.G.
California Geological Survey

Christopher Madugo, Ph.D.
Albert Kottke, Ph.D., P.E.
Pacific Gas & Electric Company

Stephen Thompson, Ph.D., C.E.G.
Lettis Consultants International

Stéphane Baize, Ph.D.
Institut de Radioprotection et de Sûreté Nucléaire

Christopher Milliner, Ph.D.
California Institute of Technology

Fiia Nurminen
Paolo Boncio, Ph.D.
Università G. d'Annunzio di Chieti-Pescara

Francesco Visini, Ph.D.
Istituto Nazionale di Geofisica e Vulcanologia

A report on research supported by Pacific Gas & Electric Company, High-Speed Rail Authority, California Department of Transportation, Southern California Gas Company, Los Angeles Department of Water and Power, and California Energy Commission.

Report GIRS-2021-08
DOI: 10.34948/N36P48

University of California, Los Angeles (headquarters)

Natural Hazards Risk & Resiliency Research Center

B. John Garrick Institute for the Risk Sciences

Fault Displacement Hazard Initiative Database

Alexandra Sarmiento, P.E., C.E.G.
Danielle Madugo
Yousef Bozorgnia, Ph.D, P.E.
Andi Shen
Silvia Mazzoni, Ph.D
University of California, Los Angeles

Grigorios Lavrentiadis, Ph.D.
University of California, Berkeley and Los Angeles

Timothy Dawson, C.E.G.
California Geological Survey

Christopher Madugo, Ph.D.
Albert Kottke, Ph.D., P.E.
Pacific Gas & Electric Company

Stephen Thompson, Ph.D., C.E.G.
Lettis Consultants International

Stéphane Baize, Ph.D.
Institut de Radioprotection et de Sûreté Nucléaire (IRSN)

Christopher Milliner, Ph.D.
California Institute of Technology

Fiia Nurminen
Paolo Boncio, Ph.D.
Università G. d'Annunzio di Chieti-Pescara

Francesco Visini, Ph.D.
Istituto Nazionale di Geofisica e Vulcanologia

A report on research conducted with support from Pacific Gas & Electric Company, California High-Speed Rail Authority, California Department of Transportation, Southern California Gas Company, Los Angeles Department of Water and Power, and California Energy Commission.

Report GIRS-2021-08

Natural Hazards Risk and Resiliency Research Center
B. John Garrick Institute for the Risk Sciences
University of California, Los Angeles (Headquarters)

August 3, 2021 (Initial Release)
July 19, 2022 (Revision 3)
March 6, 2024 (Revision 3.1)
May 3, 2024 (Revision 3.2)

ABSTRACT

This report presents the development and results of a new surface rupture mapping and fault displacement database. The new database provides an updated and standardized collection of fault displacement measurements and surface rupture maps. The work was completed as part of the Fault Displacement Hazard Initiative (FDHI) Project, which is a multi-year and community-based research project coordinated by the University of California. Next-generation fault displacement models are being developed through the FDHI Project, and the new models will improve estimates of the probability, amplitude, and spatial distribution of principal and distributed displacements in surface-rupturing earthquakes. The FDHI Database provides a common set of inputs that can be used by model development teams, allowing a more systematic comparison of model performance. Our new database contains metadata and geospatially-controlled surface rupture and fault displacement data from 75 global historical earthquakes of **M** 4.9 to 8.0 and all styles of faulting. The data were collected collaboratively through a literature review and have been assessed in detail for completeness, accuracy, and consistency. Analysis and geologic interpretation of the raw data were performed to meet model development needs, including the development of an event-specific coordinate system for each earthquake, classifying ruptures and measurements as principal or distributed, and developing recommended net slip amplitudes from reported slip components. All information is contained in a structured relational database, and the contents have been aggregated into flatfiles for formal documentation and end-user convenience. The FDHI Database is anticipated to be used by multiple model development teams in the FDHI Project and will also support related research across the geoscience community. The database and its documentation are available through the Natural Hazards Risk and Resiliency Research Center (NHR3) web site (https://www.risksciences.ucla.edu/nhr3).

ACKNOWLEDGMENTS

Support for this project was provided by the Pacific Gas & Electric Company, California High-Speed Rail Authority, California Department of Transportation, Southern California Gas Company, Los Angeles Department of Water and Power, and California Energy Commission. The California Geological Survey, United States Geological Survey, Southern California Earthquake Center, California Institute of Technology, and Lettis Consultants International were partners in this project.

The support of these organizations is gratefully appreciated. The opinions, findings, conclusions, or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the study sponsors, the Natural Hazards Risk and Resiliency Research Center (NHR3), or the Regents of the University of California.

In reviewing and compiling datasets for this project, we reached out to many researchers with data requests and queries, and we greatly appreciate the timely and thorough responses we received. The authors also wish to thank Dr. Paolo Zimmaro for early discussions and advice on the relational database schema and for recommending the SchemaSpy software to document the schema in HTML format.

The FDHI database was developed through constructive interactions and teamwork with the FDHI Project team members. The authors are grateful to the model developers for their guidance, feedback, and technical review of the database contents.

ADDENDUM FOR REVISION 3

The initial version of this report was publicly released as Revision 2, dated August 3, 2021. The authors subsequently compiled nine additional historical surface-rupturing earthquake datasets and expanded the dataset for one earthquake. For convenience, the text and appendices of this report has been updated to reflect these additions, including the contents of the digital files in Appendix A. A summary of the changes is provided below. The updated version of this report is Revision 3, dated July 19, 2022.

The following earthquakes were added to the database in Revision 3:

- 2001 **M** 7.8 Kunlun (Kokoxili), Northern Tibet, EQ_ID = 67
- 2019 **M** 4.9 Le Teil, France, EQ_ID = 68
- 2016 **M** 6.2 Norcia (#1), Italy, EQ_ID = 69
- 1979 M_L 5.2 Homestead Valley, California, EQ_ID = 70
- 2018 **M** 7.5 Palu, Indonesia, EQ_ID = 71
- 2009 **M** 6.3 L'Aquila, Italy, EQ_ID = 72
- 1988 **M** 6.77 Spitak, Armenia, EQ_ID = 73
- 1993 **M** 6.2 Killari, India, EQ_ID = 74
- 1953 **M** 7.3 Yenice-Gonen, Turkey, EQ_ID = 75

The dataset for the San Fernando earthquake (EQ_ID = 25) was updated in Revision 3 as follows:

- A new composite rupture map was created from two datasets: CGS (which was previously the rupture map basis in Revision 2), and USGS. The additional USGS linework is generally in the Sylmar area, north of I-210. The geographic coordinates and rank classification were updated for several measurement sites to be consistent with the new linework.

ADDENDUM FOR REVISION 3.2

The database flatfiles and related digital products have been assigned to a separate Digital Object Identifier (DOI) and relocated to a permanent repository hosted by the UCLA Dataverse. These data can now be accessed at https://doi.org/10.25346/S6/Y4F9LJ and cited as follows:

> Sarmiento, A., Madugo, D., Bozorgnia, Y., Shen, A., Mazzoni, S., Lavrentiadis, G., Dawson, T., Madugo, C., Kottke, A., Thompson, S., Baize, S., Milliner, C., Nurminen, F., Boncio, P., and Visini, F. (2021). "Flatfiles and Related Digital Products for the Fault Displacement Hazard Initiative Database, Release 3 dated 19 July 2022." *The B. John Garrick Institute for the Risk Sciences at UCLA Engineering*. Dataset. https://doi.org/10.25346/S6/Y4F9LJ.

This report remains available at https://doi.org/10.34948/N36P48 and can be cited as:

> Sarmiento, A., Madugo, D., Bozorgnia, Y., Shen, A., Mazzoni, S., Lavrentiadis, G., Dawson, T., Madugo, C., Kottke, A., Thompson, S., Baize, S., Milliner, C., Nurminen, F., Boncio, P., and Visini, F. (2021). "Fault Displacement Hazard Initiative Database." Report No. GIRS-2021-08, Revision 3.1 dated 6 March 2024. Los Angeles, CA: *The B. John Garrick Institute for the Risk Sciences at UCLA Engineering*. https://doi.org/10.34948/N36P48.

Previously, the database flatfiles and related digital products were provided as electronic supplements to Appendix A of this report. For a short period of time, they were also provided under a separate DOI hosted by the Natural Hazards Risk and Resiliency Research Center (NHR3) at UCLA (which was Revision 3.1 of this report); however, that DOI is retired now and users should use the new DOI at the UCLA Dataverse (10.25346/S6/Y4F9LJ) to access the database flatfiles and related digital products.

The text throughout this report, and particularly in Chapter 6 and Appendix A, has been updated to reflect the new online location (i.e., the DOI) for the flatfiles and related digital products. Otherwise, the scope and content of this report is unchanged, and *the database contents are unchanged from the July 19, 2022 release*.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 Database Project Overview

The Fault Displacement Hazard Initiative (FDHI) Project is a multi-year and community-based research project coordinated by the University of California. The objective of the project is to develop a next-generation fault displacement database and models to estimate the amplitude and spatial distribution of principal and distributed displacements in surface-rupturing earthquakes. The new models will provide improved estimates of probabilistic and deterministic fault displacement hazard. To support the FDHI Project objective, we have developed a modern database of fault displacement measurements and surface rupture maps, incorporating earthquakes as recent as November 2019.

The FDHI Database was developed in collaboration with model developers, engineering community end-users, and project sponsors. The collaboration included monthly FDHI Project meetings beginning in June 2018, frequent Database Team meetings (nominally bi-weekly), and several topical working group meetings relating to model development. We also convened a one-day workshop in October 2019 to identify end-user needs and interface issues related to the new fault displacement models (Sarmiento et al., 2019a). The workshop was attended by over 40 professionals from industry, government, and academia specializing in seismic field geology, geodesy, model development, and simulations. Interim progress on the FDHI database and models was presented at the 2019 and 2021 Seismological Society of America Annual Meetings (Sarmiento et al., 2019b, 2021; Bozorgnia et al., 2021).

This Chapter presents an overview of the FDHI Database Project, including the motivation, goals, and intended use of the database and related products. We also provide a list of the key contributions of this database to the geoscience community, a summary of the database contents, and describe the report organization.

1.1 MOTIVATION AND GOALS

Surface-rupturing earthquakes produce permanent ground displacements along fault zones that can damage infrastructure (e.g., Proctor et al., 1972; Lee and Loh, 1999; Brandenberg et al., 2019). Surface rupture is generally defined as the instantaneous breaking of the ground surface along a fault in an earthquake. Not all earthquakes break the ground surface, but those that do are differentiated as "surface-rupturing earthquakes." The process and manifestation of surface rupture

are distinct from other earthquake-related phenomena (such as liquefaction, lateral spreading, and landsliding), in that surface rupture is the result of a focused earthquake energy release along a fault plane at depth, whereas liquefaction, lateral spreading, and landsliding are deformations triggered by ground shaking (California Geological Survey, 2018). Displacements across a surface rupture can be significant (e.g., 12 m on the Kekerengu Fault in the 2016 **M** 7.8 Kaikoura earthquake; Kearse et al., 2018) and can adversely impact infrastructure. However, site-specific engineering solutions can be developed to allow structures to accommodate fault displacements (e.g., the Trans-Alaska Pipeline; Cluff et al., 2003).

The FDHI Project was initiated to develop a new fault displacement database and models in response to an increasing need to improve fault rupture and displacement hazard estimates for a variety of engineered structures and systems (Baize and Scotti, 2017). Several fault displacement models are currently used in standard practice (e.g., Youngs et al., 2003; Petersen et al., 2011; Moss and Ross, 2011; Wesnousky, 2008; Wells and Coppersmith, 1994); however, these models have significant differences in their input datasets, estimated displacement metrics, modeling techniques, and treatment of uncertainty. The next-generation fault displacement models developed through the FDHI Project will help mitigate these issues by using a common database and producing multiple displacement models in a coordinated research program.

Our new database (the FDHI Database) provides a common set of inputs that have been assessed for data quality and relevant metrics and metadata for use in the development of the new fault displacement models. Similar community-based and coordinated model development projects for ground motions have demonstrated the benefits of using a common database in model development (Chiou et al., 2008; Ancheta et al., 2014; Bozorgnia et al., 2014; Goulet et al., 2014; Bozorgnia and Stewart, 2020; Bozorgnia et al., in press). While the key benefit is that model performance can be more systematically evaluated and compared, the development of a common database is also more efficient for the scientific community and can support other research projects.

The fundamental goal of the FDHI Database Project was to support the development of new fault displacement models by systematically collecting, reviewing, and organizing relevant data in a database. The minimum required content included geospatial control for fault displacement measurements and mapped ruptures, inclusion of measurements on distributed faults, and first-order analysis and interpretation of raw data for global earthquakes of all magnitudes and styles of faulting. The database development was content- and quality-driven, with an emphasis on longevity, and the process involved extensive collaboration with the model developers to ensure the content addressed model development needs.

We performed repeated quality assurance (QA) and quality control (QC) checks on the database, with the support of participatory review from the model developers, to produce a more reliable and stable product. For this project, our data quality evaluations of completeness, accuracy, and consistency were considered to address QA. Data content requests and reviews by the model development teams ensured the final product addressed model development needs, addressing QC.

To further support longevity, the database was constructed as a structured relational database[1] A relational database is readily expandable to new data and new types of data, and it inherently contributes to QA/QC by minimizing errors due to repetition, enforcing data entry constraints, and maintaining references between individual data entries.

1.2 INTENDED USE AND COMMUNITY PRODUCTS

The FDHI Database was developed primarily for model developers to use in developing models that estimate the probability of principal and distributed surface rupture occurrence, as well as the amplitude of principal and distributed net displacement. While the database contents, QA/QC efforts, and analysis and geologic interpretation of the raw data were geared toward model developer needs, other researchers and industry professionals may find this collection of datasets useful. We encourage users of the database and its products to review Chapter 2 of this report, which discusses surface rupture manifestation and data collection and documentation methods, to understand the strengths and limitations of the original datasets used in this database. As discussed in Chapter 3, this database includes only global historical surface-rupturing earthquakes with sufficient data to meet the project event and dataset criteria.

All data, metadata[2], and interpretations are contained across 37 tables and 365 columns in one relational database file. The database contents have been aggregated into flatfiles[3], in *.csv format, for convenience and user-friendly documentation (Appendix A). We recommend most users of the FDHI Database (including model developers, geoscience researchers, and industry professionals) use the flatfiles to access the contents of the database. We used our knowledge of the database schema to produce the flatfiles and check for errors and inconsistencies; therefore, the flatfiles are the formal documentation of the database contents.

As described in Chapter 6 and Appendix A, three flatfiles are required to represent the three distinct information types contained in the database:

1. Measurements flatfile
2. Ruptures flatfile
3. Event-specific coordinate system (ECS) flatfile

For further convenience, these flatfiles are also provided as ESRI shapefiles for use in various Geographic Information System (GIS) software. We also created individual Google Earth

[1] A relational database uses a defined schema to store different data types in individual tables, relate the data between tables using key fields, and hold the information and schema in a single file. See Chapter 5 and Appendix B of this report for discussion.
[2] The term "metadata" is used herein to refer to information supporting data. For example, a displacement measurement is considered data, and information on the measurement technique (e.g., tape measure, optical image correlation) is considered metadata.
[3] A flatfile is a table created from a relational database.

*.kmz files for each earthquake in the FDHI Database. These flatfiles and related digital products are publicly available at https://doi.org/10.25346/S6/Y4F9LJ.

1.3 PROJECT HIGHLIGHTS

The data quality review, analysis, and geologic interpretation efforts completed in this project are a unique feature of the FDHI Database and have resulted in a reliable and stable product that can be used by model development teams and the broader geoscience community. Significant advancements in this database, relative to other similar compilations, are summarized below.

- Our custom relational database was designed to systematically manage the project data and metadata while establishing a lasting framework that is expandable and extensible as additional earthquake data are available, new measurement techniques develop, and user needs evolve.
- The data were collected through an extensive literature review and were systematically assessed for completeness, accuracy, and consistency.
- Multiple data sources are included for the same event, where available, providing more complete spatial coverage of measurements and surface ruptures and allowing database users to make comparisons of alternative datasets.
- Terrain metrics are included for every measurement site, and geologic information is included for most sites (where available).
- A new event-specific coordinate system algorithm is developed to supplement geographic coordinates with strike-parallel and strike-normal ordinates for all measurement sites and surface rupture linework.
- All surface ruptures in the database are classified as principal or distributed rank based on detailed geologic evaluations.
- We introduce two new fault displacement measurement rank classifications (cumulative and total) to better distinguish measurements associated with multiple ruptures or wide measurement apertures. All fault displacement measurements in the database are classified as total, cumulative, principal, or distributed. Hanging wall and footwall flags are included for distributed measurements in reverse, normal, and oblique style earthquakes.
- Recommended net slip values (preferred and bounding maximum/minimum) are calculated from the reported slip components for each measurement. The basis for the calculations of each value is documented, and we assign a quality code with recommended usage in model development to each value.

1.4 DATABASE CONTENTS SUMMARY

We have assembled a geospatially-controlled relational database of surface rupture maps, measurements, and associated metadata for 75 global historical earthquakes of **M** 4.9 to 8.0 occurring between 1872 and 2019. Figure 1.1 shows the spatial distribution of the epicenters for 71 events in the database for which this information is available, and Figure 1.2 shows the same information for events in the conterminous United States and Mexico. The relative regional distribution of the 75 events is illustrated in the pie chart in Figure 1.3. Approximately 40% of the earthquakes in the database are from Western North America, which includes California, Nevada, Idaho, Montana, Alaska, and Mexico. One quarter of the events are from Japan, China, or Southeast Asia. There are also several events from Australia, which is a stable continental region.

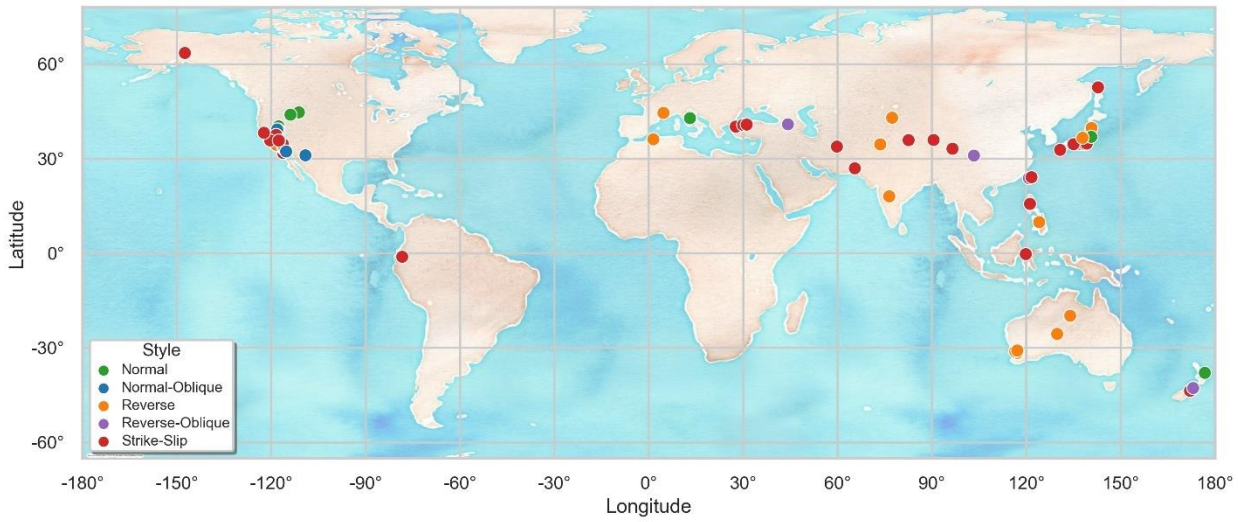

Figure 1.1. Epicentral locations of 71 of the 75 earthquakes in the FDHI Database (color-coded by style of faulting; see inset legend). Epicenters for the following events are not available: 1872 Owens Valley, California; 1912 Acambay, Mexico; 1986 Marryat Creek, Australia; and 2012 Pukatja, Australia.

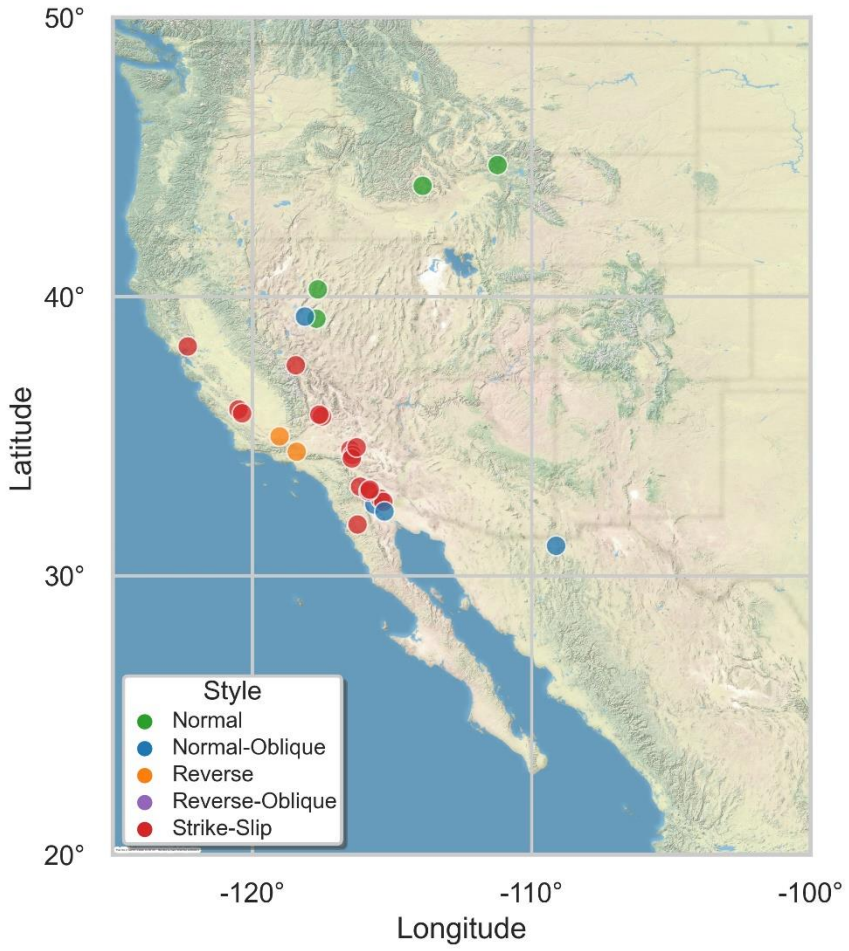

Figure 1.2. Epicentral locations of earthquakes in the FDHI Database in the conterminous United States and Mexico (color-coded by style of faulting; see inset legend). Epicenters for the 1872 Owens Valley, California and 1912 Acambay, Mexico earthquakes are not available.

Figure 1.4 shows the magnitude and style of faulting characteristics of the earthquakes in the database. All 66 events are ordered by date on the abscissa. Roughly 45% of the events are dominantly strike-slip, 20% are normal, and 35% are reverse. Overall, the events in the database span a magnitude range that corresponds to the hazard levels of interest for engineering design and analysis in active tectonic settings like California: more frequent smaller events (M ~6 to 6.5) that are more or less congruent with code-based hazard levels (e.g., ASCE 7-16 design response spectra), and larger (M ~7) events that are similar to the maximum considered earthquake hazard levels (e.g., ASCE 7-16 MCE_R level).

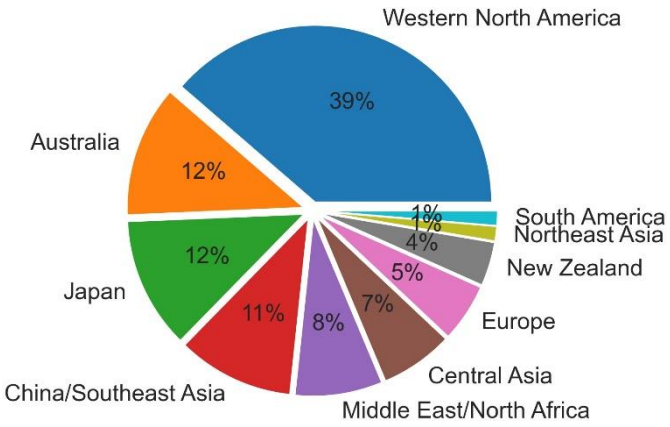

Figure 1.3. Regional distribution of earthquakes in the FDHI Database.

Event-specific metadata (including magnitude, magnitude type, hypocenter location, and style of faulting) are included for all earthquakes in the database. Dataset metadata (e.g., citation, mapping scale) are also included. The database includes over 87,000 individual point-in-space observations with geospatial control, including over 40,000 fault displacement measurements for a range of slip components (Figure 1.5). Surficial geologic unit classification (bedrock, young/old/undifferentiated alluvium; see Chapter 4.1 for definitions) is available for over 26,000 observation sites (Figure 1.5). The database also contains surface rupture maps for each earthquake with geospatial control on the rupture line vertices.

At the request of the model development teams, we also performed geologic interpretation of the rupture linework and displacement measurements to distinguish principal and distributed faulting (Chapter 2.5.3), aggregate the reported slip components into recommended net slip values for use in model development, and explicitly flag alternative measurements at the same location. Finally, as also requested by the model development teams, we developed an event-specific coordinate system (ECS) for each earthquake in the database. The ECS is a two-dimensional projection of the event data that accounts for curvature and discontinuities in the surface rupture trace. An example of the ECS is given in Figure 1.6, which also shows the mapped surface ruptures and interpreted principal/distributed classifications from the 1971 San Fernando (Sylmar) earthquake.

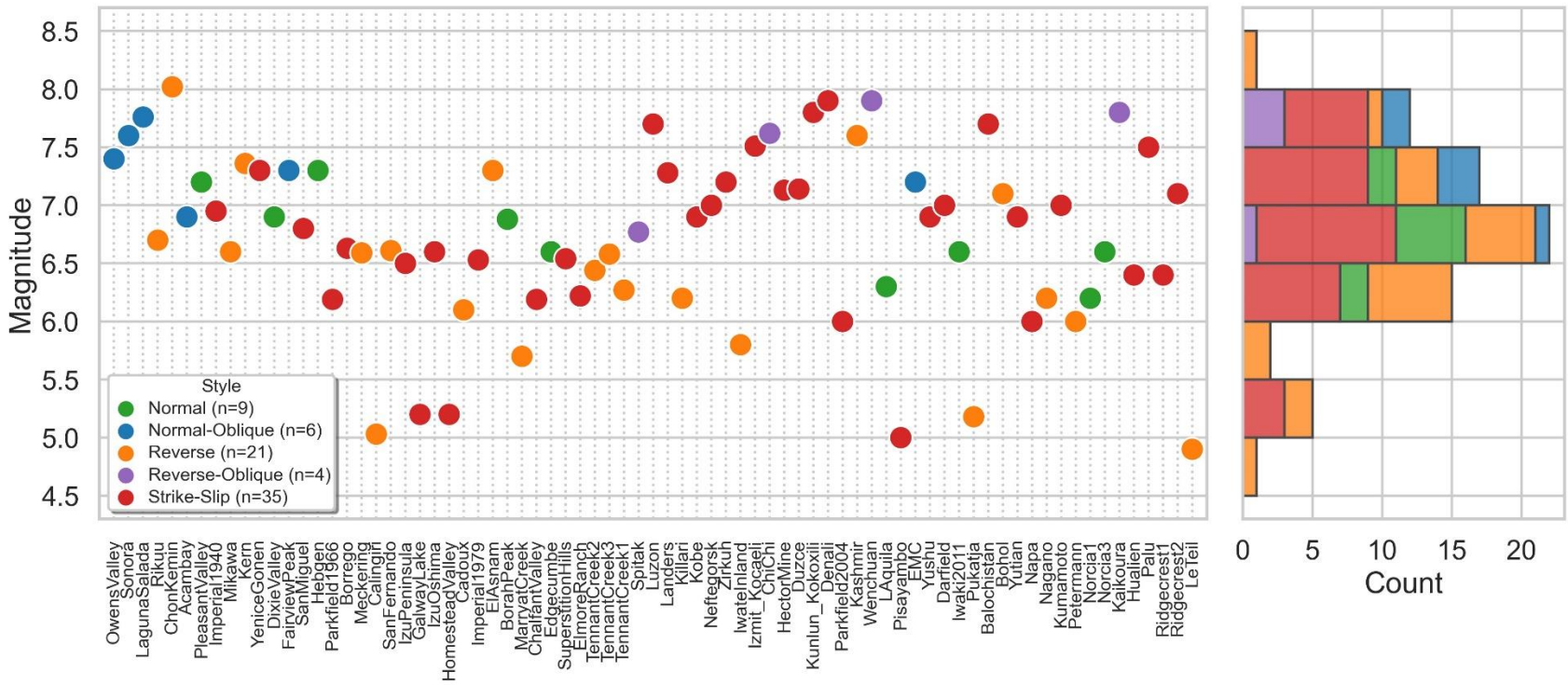

Figure 1.4. Magnitude and style of faulting of the 75 events in the FDHI Database.

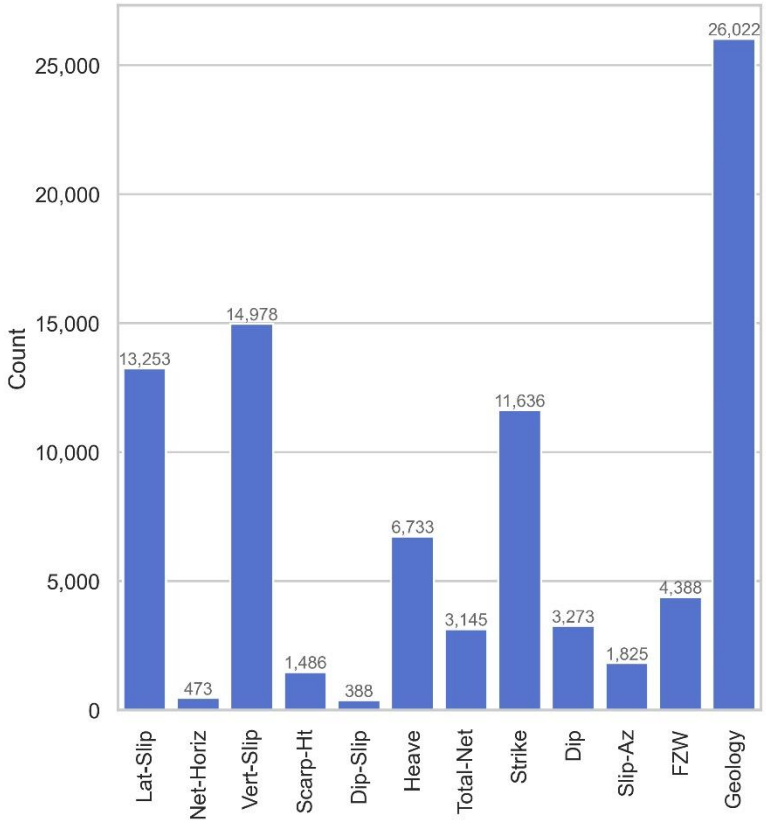

Figure 1.5. Number of measurements contained in the FDHI Database across all earthquakes and datasets. Slip components defined in Chapter 2.5.1 and in flatfiles described in Appendix A.

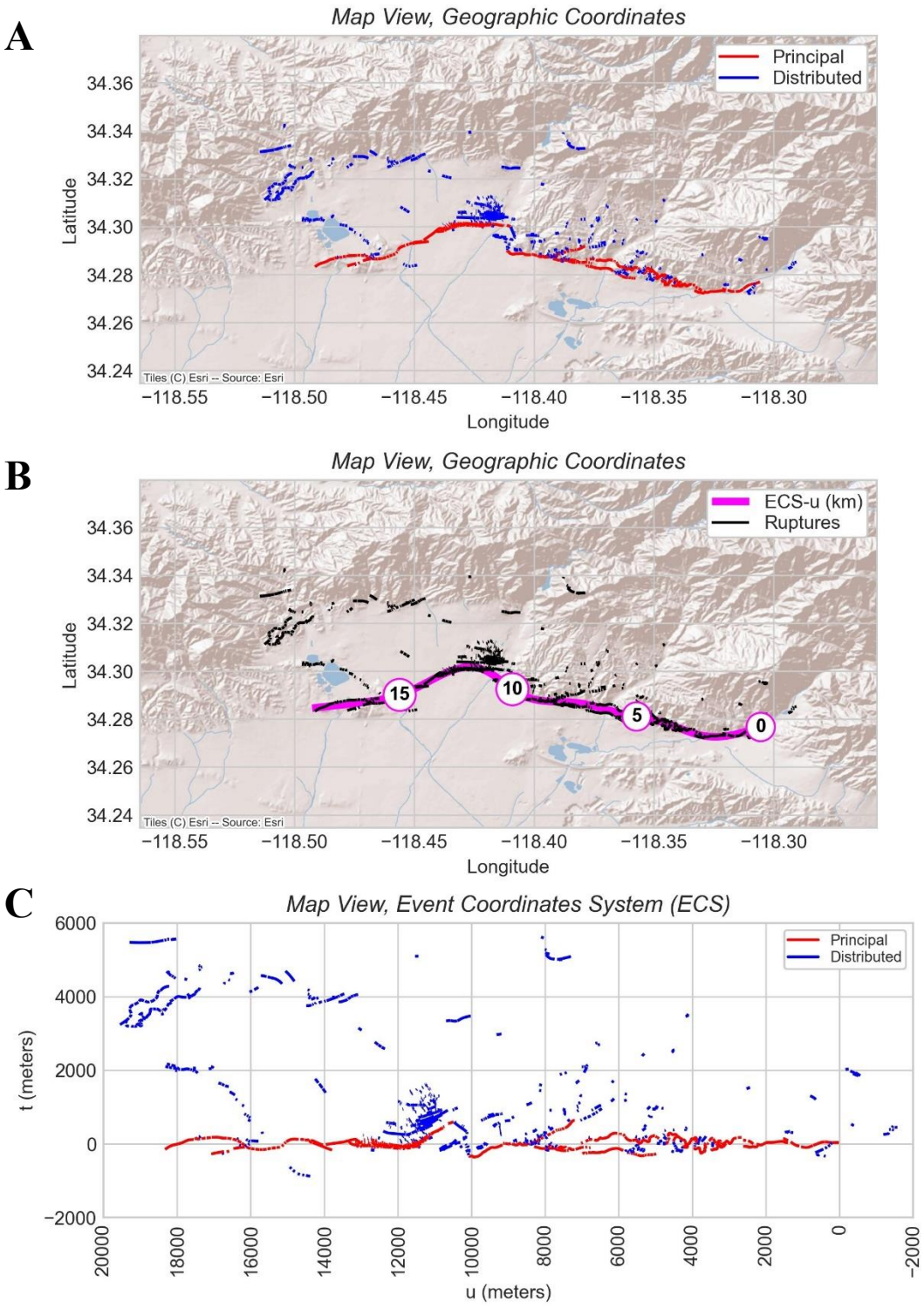

Figure 1.6. Surface rupture map from 1971 San Fernando (Sylmar) earthquake. EQ_ID = 25 in the FDHI Database. (A) Principal and distributed ruptures, geographic coordinates. (B) Ruptures and event-specific coordinate system (ECS) reference line, geographic coordinates; white circles are distance along reference line in kilometers. (C) Principal and distributed ruptures projected into ECS.

1.5 REPORT ORGANIZATON

The Chapters in this report document the development of the FDHI Database. (A separate report will document new models developed under the FDHI Project using this database). The process of building the database began with a systematic review of surface rupture characteristics, data collection tools, techniques, and reporting standards, and existing fault displacement and surface rupture compilations. We collaborated with the model developers to determine the initial database contents and then developed a custom relational database structure to accommodate the range of data types. We then reviewed the available published literature for geospatially-controlled measurements and rupture mapping from historical surface-rupturing earthquakes and performed first-order analysis and geologic interpretation of the raw datasets. All datasets were reviewed for quality and completeness before being imported into the database. Finally, we aggregated the database contents into flatfiles (*.csv format) for formal documentation and generated ESRI shapefiles (*.shp format) and Google Earth files (*.kmz format) for end-user convenience (Chapter 6 and Appendix A).

1.5.1 Definitions

Important terms are usually defined as they are introduced. In some cases, we defer detailed definitions to a specific chapter. For convenience, the following list summarizes some important terms and where they are defined in this report:

- Database-related terms (*relational database*, *flatfile*, and *metadata*) are defined in the footnotes in Chapter 1.2.
- *Principal* and *distributed* faulting are defined where they are first introduced in Chapter 2.5.3 are discussed in further detail in Chapter 4.3. The *total* and *cumulative* fault displacement measurement classifications are defined in Chapter 4.3.
- *Slip components* (e.g., vertical slip, fault-parallel slip) are defined in Chapter 2.5.1.

1.5.2 Chapters Overview

Chapter 2 of this report describes surface rupture characteristics with photographs of various types of complexities and reviews typical data sources and tools used to develop surface rupture maps and collect fault displacement measurements. Chapter 2 also provides a summary of key terms relating to surface rupture and fault displacements and their usage in this project.

Chapter 3 documents the data collection approach for this project. Chapter 3 includes event and dataset criteria, an overview of existing compilations, the standard workflow for developing data for each earthquake and dataset, and discussion of intentionally excluded data.

Chapter 4 describes the data analysis and interpretation applied to the collected data. The analyses include spatial analysis performed in GIS software to develop geologic information, elevation data and metrics, and the ECS. The interpretations include classifying or ranking ruptures and measurements generally as principal or distributed (additional classifications include cumulative and total, as described in Chapter 4.3), developing recommended net slip amplitudes for use in model development based on reported slip components, assigning recommended usage and quality codes to the measurement data, and explicitly identifying co-located alternative displacement measurements.

Chapter 5 describes technical aspects of the relational database development, including the database management system, database schema, and process of populating the database. More detail on the relational database is also documented in Appendix B.

Chapter 6 provides an overview of the flatfile documentation and contents. The flatfiles are described in more detail in Appendix A of this report and are publicly available at https://doi.org/10.25346/S6/Y4F9LJ.

Chapter 7 discusses the QA/QC process applied throughout the database development and documentation.

Chapter 8 presents the conclusions of this report.

1.6 REFERENCES

American Society of Civil Engineers (ASCE) (2017). Minimum Design Loads and Associated Criteria for Buildings and Other Structures. ASCE Standard ASCE/SEI 7-16. American Society of Civil Engineers.

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., and Donahue, J. L., (2014). NGA-West2 database, *Earthquake Spectra*, **30**, 989–1005.

Baize, S. and Scotti, O. (2017). Fault Displacement Hazard Analysis Workshop in Menlo Park (USGS, California), 8-9/12/2016, Synthesis and Perspectives, *Institut Radioprotection Sûreté Nucléaire*, 72 p.

Brandenberg, S. J., Wang, P., Nweke, C. C., Hudson, K., Mazzoni, S., Bozorgnia, Y., Goulet, C. A., Hudnut, K. W., Davis, C. A., Ahdi, S. K., Zareian, F., Fayaz, J., Koehler, R. D., Chupik, C., Pierce, I., Williams, A., Akciz, S., Hudson, M. B., Kishida, T., Brooks, B., Gold, R., Ponti, D., Scharer, K., McPhillips, D., DuRoss, C., Ericksen, T., Hernandez, J., Patton, J.,

Olson, B., Dawson, T., Treiman, J., Blake, K., Buchhuber, J., Madugo, C., Sun, J., Donnellan, A., Lyzenga, G., and Conway, E. (2019). Preliminary report on engineering and geological effects of the July 2019 Ridgecrest earthquake sequence. Report GEER-064, *Geotechnical Extreme Event Reconnaissance Association*.

Bozorgnia, Y., Abrahamson, N. A., Ahdi, S. K., Ancheta, T. D., Al Atik, L., Atchuleta, R. J., Atkinson, G.M., ... and Youngs, R. R. (in press). NGA-Subduction Research Program, *Earthquake Spectra*.

Bozorgnia, Y., Abrahamson, N., Arcos, B., Baize, S., Boncio, P., Chao, S.-H., Chen, R., Chiou, B., Dawson, T., Donahue, J., Goulet, C., Hanson, K., Kottke, A., Kuehn, N., Kuo, C.-H., Lavrentiadis, G., Madugo, D., Mazzoni, S., Milliner, C., Moss, R., Nurminen, F., Pace, B., Petersen, M., Sarmiento, A., Shen, A., Thomas, K., Thompson, S., Visini, F., Wang, Y., and Youngs, R. R. (2021). An Overview of the Fault Displacement Hazard Initiative Research Program. Presented at Seismological Society of America 2021 Annual Meeting, in *Seismological Research Letters*, **92**(2B), 1360-1361.

Bozorgnia, Y. and Stewart, J. P. (2020). Data Resources for NGA-Subduction Project. PEER Report 2020/02, *Pacific Earthquake Engineering Research Center*, University of California, Berkeley, CA.

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., and Youngs, R. (2014). NGA-West2 Research Project, *Earthquake Spectra*, **30**(3), 973-987.

California Geological Survey (2018). Earthquake Fault Zones, A Guide for Government Agencies, Property Owners / Developers, and Geoscience Practitioners for Assessing Fault Rupture Hazards in California, Revised 2018, California Geological Survey Special Publication 42.

Chiou, B., Darragh, R., Gregor, N., and Silva, W. (2008). NGA project strong-motion database, *Earthquake Spectra*, **24**(1), 23-44.

Cluff, L. S., Page, R. A., Slemmons, D. B., and Crouse, C. B. (2003). Seismic hazard exposure for the Trans-Alaska pipeline, *In* Advancing Mitigation Technologies and Disaster Response for Lifeline Systems (pp. 535-546).

Goulet, C. A., Kishida, T., Ancheta, T. D., Cramer, C. H., Darragh, R. B., Silva, W. J., Hashash, Y. M. A., Stewart, J. P., Woodell, K. E., and Youngs, R. R. (2014). PEER NGA-East Database, PEER Report 2014/17, *Pacific Earthquake Engineering Research Center*, University of California, Berkeley, CA.

Kearse, J., Little, T. A., Van Dissen, R. J., Barnes, P. M., Langridge, R., Mountjoy, J., Ries, W., Villamor, P., Clark, K. J., Benson, A., Lamarche, G., Hills, M., and Hemphill-Haley, M.

(2018). Onshore to offshore ground-surface and seabed rupture of the Jordan–Kekerengu–Needles fault network during the 2016 Mw 7.8 Kaikōura earthquake, New Zealand, *Bulletin of the Seismological Society of America*, **108**(3B), 1573-1595.

Lee, G. C. and Loh, C. H. (1999). Preliminary report from MCEER-NCREE workshop on the 921 Taiwan earthquake, *Multidisciplinary Center for Earthquake Engineering Research*.

Moss, R. E. S. and Ross, Z. E. (2011). Probabilistic fault displacement hazard analysis for reverse faults, *Bulletin of the Seismological Society of America*, **101**(4), 1542-1553.

Petersen, M. D., Dawson, T. E., Chen, R., Cao, T., Wills, C. J., Schwartz, D. P., and Frankel, A. D. (2011). Fault displacement hazard for strike-slip faults, *Bulletin of the Seismological Society of America*, **101**(2), 805-825.

Proctor, R. J., Crook Jr., R., McKeown, M. H., and Moresco, R. L. (1972). Relation of known faults to surface ruptures, 1971 San Fernando earthquake, southern California, *Geological Society of America Bulletin*, **83**(6), 1601-1618.

Sarmiento, A., Baize, S., Boncio, P., Bozorgnia, Y., Dawson, T., Lavrentiadis, G., Madugo, D., Nurminen, F., Shen, A., and Thompson, S. (2021). A Next-Generation Fault Displacement and Surface Rupture Database. Presented at Seismological Society of America 2021 Annual Meeting, in *Seismological Research Letters*, **92**(2B), 1360.

Sarmiento, A., Dawson, T., Lavrentiadis, G., and Bozorgnia, Y. (2019a). Summary Report on the Fault Displacement Workshop Held on October 1, 2019, *Southern California Earthquake Center*, SCEC Award #19201.

Sarmiento, A., Abrahamson, N., Baize, S., Bozorgnia, Y, Chen, R., Coppersmith, K. J., Dawson, T. E., Donahue, J. L., Jacob, V., Ketabdar, M., Kottke, A. R., Kuehn, N., Lavrentiadis, G., Madugo, C., Mazzoni, S., Milliner , C., Shamsaabadi, A., Shantz, T., Shen, A., Thompson, S. C., and Youngs, R. R. (2019b). A New Model Database for Next-Generation Fault Displacement Hazard Analysis. Presented at Seismological Society of America 2019 Annual Meeting, Seattle, Washington, in *Seismological Research Letters*, **90**(2B), 1058.

Wells, D. L. and Coppersmith, K. J. (1994). New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement, *Bulletin of the Seismological Society of America*, **84**(4), 974-1002.

Wesnousky, S. G. (2008). Displacement and geometrical characteristics of earthquake surface ruptures: Issues and implications for seismic-hazard analysis and the process of earthquake rupture, *Bulletin of the Seismological Society of America*, **98**(4), 1609-1632.

Youngs, R. R., Arabasz, W. J., Anderson, R. E., Ramelli, A. R., Ake, J. P., Slemmons, D. B., McCalpin, J. P., Doser, D. I., Fridrich, C. J., Swan III, F. H., Rogers, A. M., Yount, J. C., Anderson, L. W., Smith, K. D., Bruhn, R. L., Knuepfer, P. L. K., Smith, R. B., dePolo, C. M., O'Leary, D. W., Coppersmith, K. J., Pezzopane, S. K., Schwartz, D. P., Whitney, J.

W., Olig, S. S., and Toro, G. R. (2003). A methodology for probabilistic fault displacement hazard analysis (PFDHA), *Earthquake Spectra*, **19**(1), 191-219.

2 Surface Rupture Characteristics, Data Collection Methods, and Terminology

2.1 INTRODUCTION

The documentation of surface ruptures in an earthquake can vary significantly, owing to variations in mapping scale, rupture characteristics or expression, areas investigated, degree of preservation, level of effort, mapping standards, and the purpose of the rupture map. As a result, many scientists, engineers, and practitioners may be surprised by the variability of complexity and detail contained or absent in the database for individual earthquakes. Similarly, fault displacement measurements are also sometimes oversimplified, inconsistently documented, or incomplete for many reasons, and users of this database may also be surprised by the non-uniform spatial distribution of displacement measurements or the variability in alternative measurements at the same location. In this Chapter, we present examples through of photographs to show a range of surface rupture complexities and measurable, unmeasurable, or ambiguous features. We also provide an overview of the methods, tools, and techniques used in collecting surface rupture and fault displacement measurements after an earthquake. Finally, we define relevant geologic terms used in standard practice and the FDHI Database Project.

The goal of this Chapter is to give users of this database an understanding of what surface ruptures look like, how ruptures are represented on a map, and the strengths and limitations of the different methods, tools, and techniques that are used to document surface ruptures and fault displacements. An important conclusion of this Chapter is that variable documentation of the simplifications, assumptions, and terminology in rupture mapping and fault displacement measurement reporting is common in the original datasets we reviewed, and therefore the degree of simplification varies between events and sometimes within an event.

2.2 EXAMPLE MANIFESTATIONS OF SURFACE RUPTURE

Total surface rupture lengths are tens to hundreds of kilometers long, which presents many challenges in documenting the character and spatial distribution of the individual surface ruptures. Common challenges include inaccessible areas (e.g., private property, difficult terrain, offshore), deformation obscured by dense vegetation, snow, or landslides, post-earthquake ground surface

modification (e.g., infrastructure repair work, weather/storm events), and resource limitations (e.g., time, budget, workforce). Distributed ruptures away from the principal fault and ruptures with small displacements suffer more of these limitations, as they are more easily overlooked by field reconnaissance teams and more susceptible to modification by surface processes. Although the quick turnaround time of high-resolution satellite-based data mitigates many of these challenges (Morelan and Hernandez, 2020; Milliner and Donnellan, 2020) for very recent and future earthquakes, the documentation of nearly all historical earthquakes was limited by these challenges.

In this Chapter, we present maps and photographs of surface rupture to reveal (un)intentional simplifications inherent in original datasets and the challenges in quantitatively capturing all deformation associated with a surface rupture. The maps demonstrate important differences in mapping scale that can arise from the data collection method or the intended use of the map. The photographs in this Chapter are an inexhaustive set of examples of surface rupture complexity at site-specific scales. Together, the maps and photographs are a useful introduction to help users of the FDHI Database visualize surface rupture complexity and understand the types of data included, data collection methods, and strengths and limitations of the methods.

Measuring fault displacement requires, at minimum, knowledge or inference of the ground surface and geometry of the physical features before the earthquake. Specifically, we must be able to identify physical features that were adjacent, or in some other known spatial configuration, before the earthquake and then measure their current separation (i.e., piercing points). Common piercing points include but are not limited to: geologic or geomorphic features such as channel margins or thalwegs, terrace risers, and alluvial fans; and cultural features such as roadways, vehicle tracks and fences. The photographs in this Chapter also provide some examples of measurable, unmeasurable, and ambiguous features. Furthermore, we show examples where incomplete documentation of surface rupture complexity can also cause ambiguity in reported displacement measurements.

2.2.1 Mapping Scale

Thorough and consistent documentation of surface ruptures is complicated by the fractal (scale-invariant) nature of faults (Turcotte, 1990). Site-specific factors, such as near-surface soil/rock conditions, also affect how ruptures manifest on the ground surface (e.g., Bray et al., 1994; Moss et al., 2013). Surface rupture characteristics vary widely from simple, discrete planar faulting to complex or diffuse networks of fissures or minute cracking. Mapping scale is particularly important for complex ruptures: small-scale maps have limited resolution that reduce detail, whereas large-scale maps can retain a high level of detail[4]. For example, Figure 2.1 shows a portion of the rupture mapping from the 1966 **M** 6.19 Parkfield, California earthquake at two different

[4] For example, a 1:500-scale map has more detail (larger scale) than a 1:50,000-scale map (smaller scale) (Avery and Berlin, 1992).

mapping scales that underscore the fractal nature of surface rupture. The full rupture was documented on a 1:24,000-scale topographic base map as mostly continuous curvilinear rupture traces, but the geologists also mapped minute fractures (~15 cm in length) at a very large scale in several locations. The site-specific mapping in Figure 2.1B is along an 8-m-long portion of the rupture. The purpose of site-specific mapping is usually to document a specific structural complexity or illustrate a representative level of complexity for a portion of the rupture.

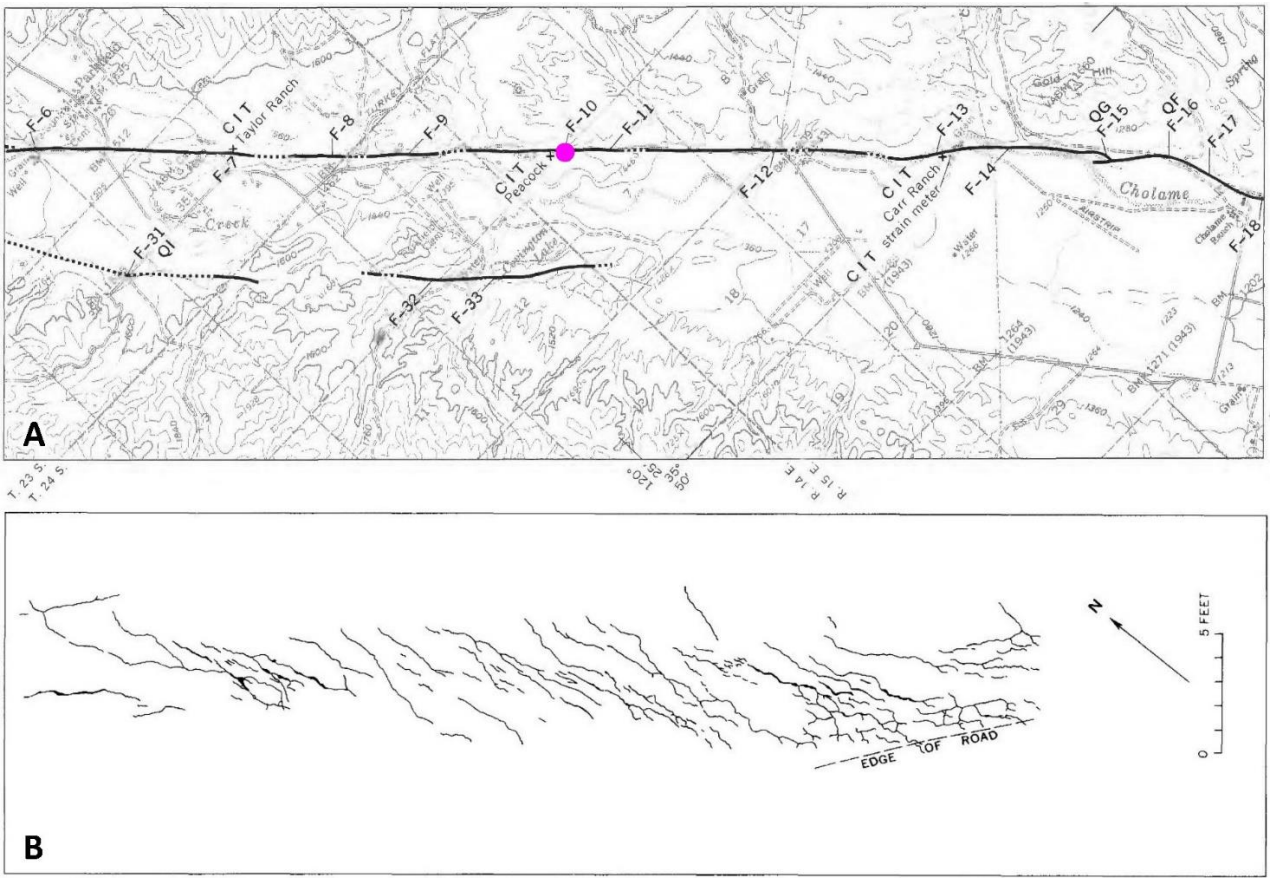

Figure 2.1. Surface rupture datasets from 1966 **M** 6.19 Parkfield, California earthquake. (A) Ruptures mapped on 1:24,000-scale 7.5' quadrangle topographic map; magenta dot is location F10. (B) Fractures in asphalt road mapped from vertical photographs at location F10. Source: Brown et al. (1967), USGS Professional Paper 579.

Figure 2.2 is a 4-km-long section from the 2010 **M** 7.2 El Mayor-Cucapah, Mexico earthquake surface rupture at two different mapping scales. The Teran et al. (2015) mapping was completed at a 1:500 scale, and the Fletcher et al. (2014) mapping is a regional simplification at an unreported scale (likely ~1:50,000 or smaller). The Fletcher et al. (2014) compilation manually simplified the surface rupture based on the geometry and kinematics to investigate the fault rupture process in the earthquake.

Figure 2.3 shows a portion of the 2019 **M** 7.2 Ridgecrest, California earthquake at three different mapping scales. The Ponti et al. (2020) dataset in this area was developed from unmanned aerial vehicle (UAV) imagery at a very large scale (1:100 or larger), capturing small fractures (~2 m in length) that might not be of engineering significance; however, the mapping scale is less than 1:10,000 in other areas. The CGS lidar mapping (unpublished work) is based on a 1:1,000 scale. The differences in the ruptures mapped from these datasets are related to both the resolution of the source data and geologists' interpretations. The DuRoss et al. (2020) dataset was manually simplified from the Ponti et al. (2020) dataset for the purpose of developing fault displacement profiles. The scale of the DuRoss et al. (2020) dataset is not reported but is likely on the order of 1:24,000.

Mapping scale is generally controlled by the data collection method (Figure 2.3), intended usage of the map (Figure 2.2), or external limiting factors (e.g., inaccessible areas, resource limitations). Therefore, the level of detail in surface rupture maps can vary significantly between earthquakes, in different areas of the same earthquake, or even within the same area of an earthquake. This is an important limitation in the database that is inherited from the original datasets. Moreover, most surface rupture maps are not compiled at scales that are appropriate for site-specific engineering applications.

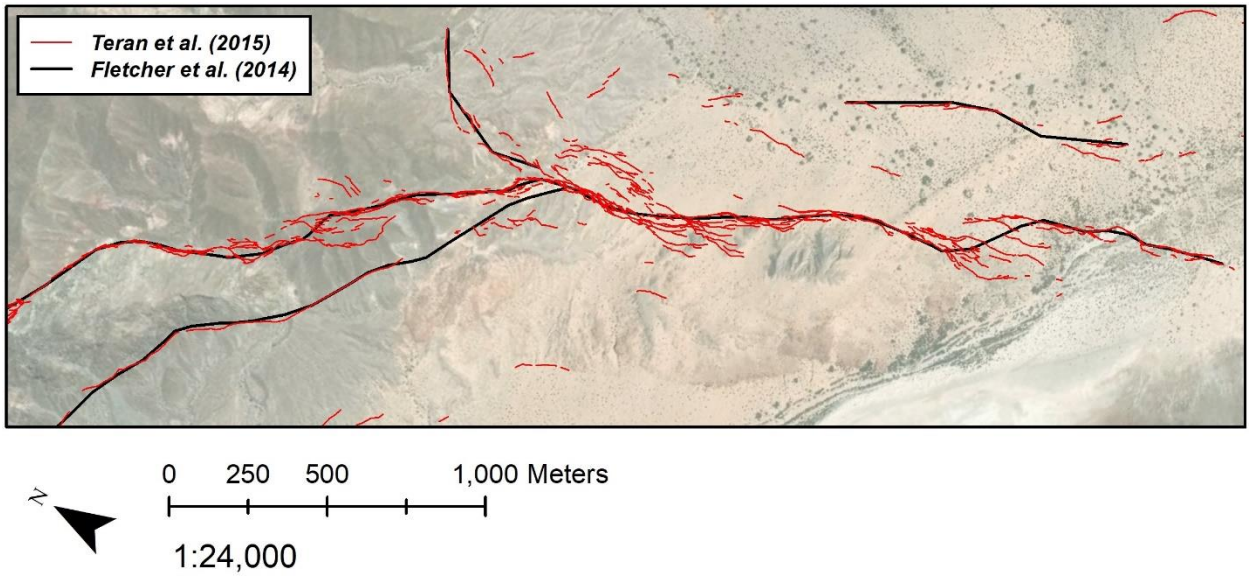

Figure 2.2. Surface rupture datasets from the 2010 **M** 7.2 El Mayor-Cucapah, Mexico (EMC) earthquake. Reported mapping scale from Teran et al. (2015) is 1:500; Fletcher et al. (2014) mapping scale not reported, but estimated to be ~1:50,000 or smaller. Area shown is near 32.548°N, 115.691°W.

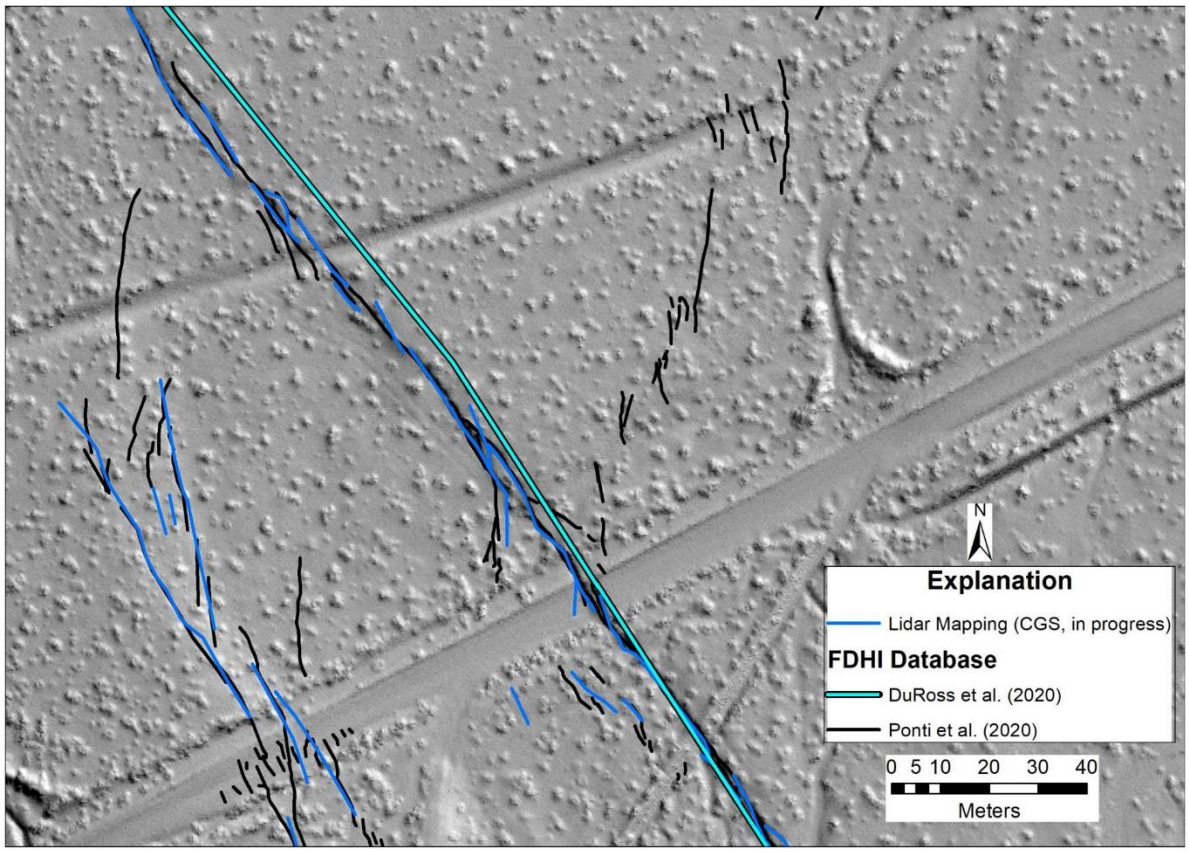

Figure 2.3. Surface rupture datasets from the 2019 **M** 7.2 Ridgecrest, California earthquake. Ponti et al. (2020) mapping scale is 1:100 in this area (varies elsewhere); unpublished lidar mapping by CGS (Dawson, T.), scale 1:1,000; and DuRoss et al. (2020) mapping scale not reported, but estimated to be ~1:24,000.

2.2.2 Single Discrete Surface Rupture

Conceptually, the simplest surface rupture pattern is a single, continuous, linear, and discrete fault, which is an infinitely thin feature that can be represented as a line on a map (sometimes informally referred to as a "knife-edge" rupture). The photographs in Figure 2.4 show examples of continuous, linear, and discrete surface ruptures from the 2019 **M** 7.1 Ridgecrest and 1906 **M** 7.9 San Francisco earthquakes in California (USGS, 2021).

The surface rupture characteristics in both photographs in Figure 2.4 are similar (continuous, linear, and discrete) and would be represented similarly on most surface rupture maps. There are subtle differences in the surface rupture characteristics that might be reflected on a site-specific scale map. For example, the surface soils in Figure 2.4B are displaced across a zone of deformation that is at least one-meter-wide. The zone of deformation is wider in the foreground of the picture, partly due to perspective but mainly due to a localized area of extension that forms a fissure. The alternating fissures and push-ups are a common manifestation of surface rupture in strike-slip earthquakes referred to as a "moletrack." The rupture continuity is also more variable

in Figure 2.4B (such as above the woman's head in the photograph), and there are several short ruptures at a high angle to the main rupture with vertical displacement. A site-specific scale map of the area in Figure 2.4B might show the rupture as a zone (area in map view) and distinguish the fissure boundaries and high-angle splays.

The examples in Figure 2.4 are from right-lateral fault rupture in large earthquakes. It is unlikely that right-lateral displacements could be measured at either location, as distinctly displaced features (i.e., piercing points) are not visible in either photograph. (Note that the evaporite boundaries in Figure 2.4A are irregular and not reliable markers.) Alternatively, Figure 2.5 shows another example of a continuous, linear, and discrete surface rupture from the 2019 **M** 7.1 Ridgecrest, California earthquake that crosses a small channel margin, providing reliable piercing points for measuring right-lateral displacement.

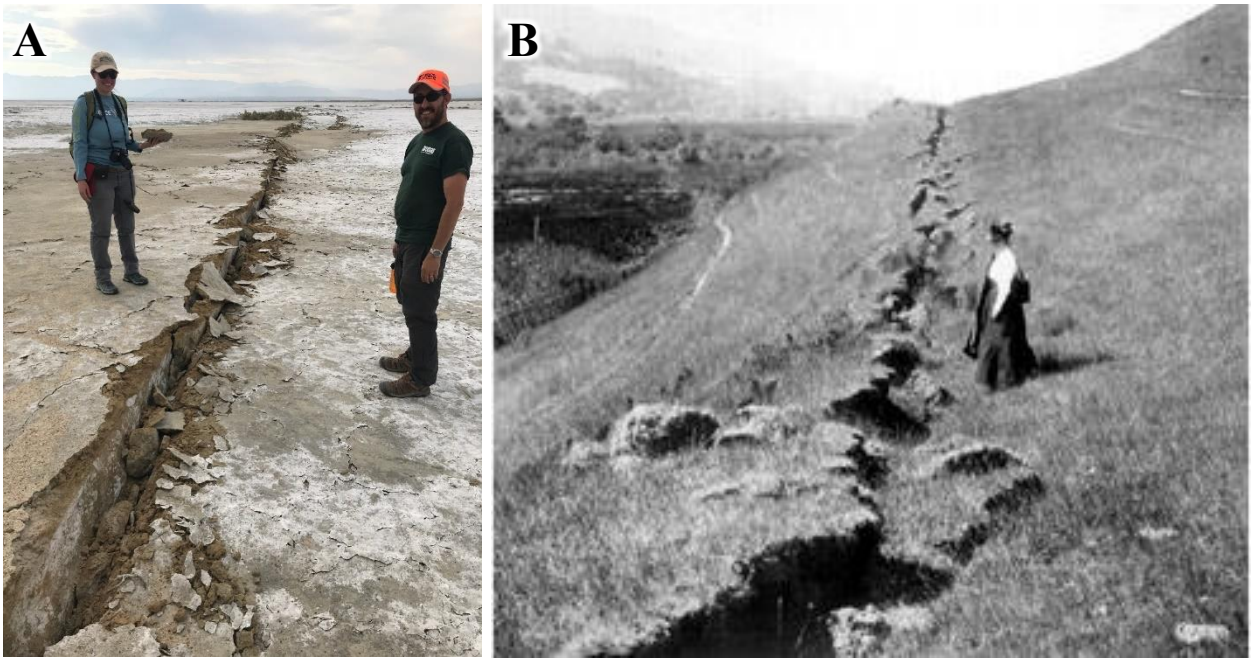

Figure 2.4. (A) Photograph along portion of 2019 Ridgecrest, California **M** 7.1 surface rupture, location not reported. (B) Photograph along portion of 1906 **M** 7.9 San Francisco, California surface rupture near 38.057312°N, 122.807878°W. Source: USGS Earthquake Photo Collections (USGS, 2021).

Figure 2.5. Photograph along portion of 2019 Ridgecrest, California **M** 7.1 surface rupture showing right-lateral displacement of a small channel margin (location not reported). Source: USGS Earthquake Photo Collections (USGS, 2021).

2.2.3 Distributed Ruptures and Deformation

Earthquakes can also produce multiple surface ruptures across zones hundreds of meters wide. Figure 2.6 is a photograph of the 1999 **M** 7.1 Hector Mine earthquake showing a fault zone roughly 25 meters in width (USGS, 2021). Several surface ruptures are visible in the photograph. Although many of the individual ruptures are continuous, linear, and discrete traces, the overall surface rupture at this location is relatively complex due to the amount and density of individual ruptures. Few, some, or all of the individual ruptures could be captured in a rupture map, depending on the scale or purpose of the map. Vertical slip can be measured along multiple locations of most of the ruptures visible in Figure 2.6, but piercing points for lateral slip measurements are not discernable in the photograph. Measurement reporting in rupture zones like this can vary, with individual measurements on separate ruptures (which may or may not be delineated on a rupture map) or slip measurements summed across closely-spaced ruptures. The latter case is not always distinguished in the original datasets.

Figure 2.6. Photograph along portion of 1999 Hector Mine, California **M** 7.1 surface rupture (location not reported). Source: USGS Earthquake Photo Collections (USGS, 2021).

In addition to discrete surface rupture traces, nonbrittle deformation in the form of warping, rotation, or tilting is also common in earthquakes. In the photograph in Figure 2.7, surface deformation is mainly accommodated by broad warping and multiple discontinuous fissures. The area across which warping or tilting of the ground surface occurs is typically poorly documented in rupture mapping: sometimes a trace is mapped at the base or center of the scarp, implying discrete rupture, or sometimes linework is omitted because a discrete rupture is absent. The discontinuous fissures are sometimes, but not always, included on rupture maps. Vertical slip is usually measured across the zone. Many engineered structures are sensitive to both the displacement amplitude and the width of the zone; however, the width of the zone is usually not reported, and measurement or rupture metadata might not distinguish between discrete surface rupture and broad warping.

Figure 2.7. Photograph along portion of 2016 Kaikoura, New Zealand **M** 7.8 surface rupture. Source: Madugo, C.

2.2.4 Site-Specific Complexity

Relatively simple rupture traces can transition into diffuse or discontinuous traces over short distances. Figure 2.8 is a photograph of surface rupture on the Kaikoura Fault northwest of Clarence, New Zealand from the 2016 Kaikoura **M** 7.6 earthquake. A robust and continuous subvertical rupture (fault scarp) is clear in the center of the photograph; however, the surface rupture expression suddenly changes from the distinct linear trace to a zone of deformation on the right side of the photograph. A rupture map might record individual ruptures in the zone of deformation or simply continue the linear trace from the left side of the picture. Several secondary ruptures are also visible in the hill above the main rupture. The secondary ruptures may or may not be included on a rupture map, depending on the scale or purpose of the map.

The range of surface rupture expressions in Figure 2.8 provides examples of simple and complicated locations for measuring fault displacement. The main rupture decays from a distinct linear trace (left) to a zone of distributed deformation (right). Multiple secondary ruptures are also evident in the hill above the main rupture. The uniformity of the vegetative grass highlights vertical displacements throughout the photograph area, and the topography provides some constraints on lateral displacements on the smaller, high-angle secondary faults (between the main rupture and the geologist in the photograph); therefore, reliable vertical displacements can be obtained on many of these faults. The offset fencing in the hills, near the geologist in the photograph, provides excellent piercing points for measurement of lateral displacement. Although it is difficult to see in the photograph, broken and offset fencing near the robust linear part of the main rupture can also be used to measure lateral and net displacements. Within the diffuse part of the main rupture, a good measurement of vertical displacement is probably possible while lateral displacement measurement opportunities might be limited. Most importantly, the relatively similar amplitude of vertical displacements across the main fault in this photograph span different fault zone widths: a subvertical discrete rupture on the left side of the picture, and a roughly two-meter-wide zone on the right side of the picture. Many engineered structures are sensitive to both the displacement amplitude and the width of the zone; however, width is usually not documented.

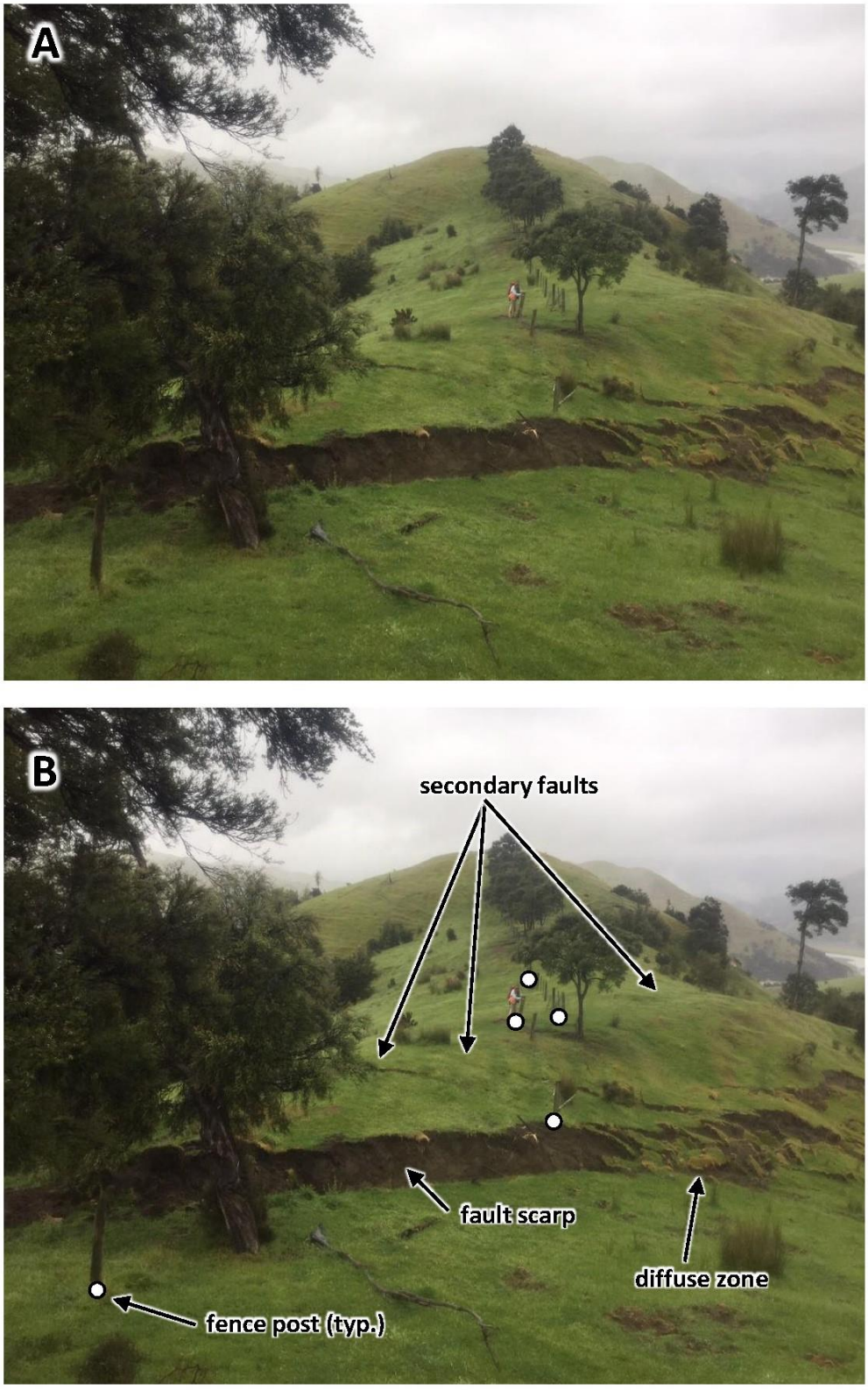

Figure 2.8. Photograph along portion of 2016 Kaikoura, New Zealand **M** 7.8 surface rupture. (A) Unannotated. (B) Annotated based on discussion in text. Source: Madugo, C.

The photograph in Figure 2.9 is also from the 2016 Kaikoura earthquake and is located within a few hundred meters of the picture in Figure 2.8. The rupture in Figure 2.9 manifests at the surface as two (sub)parallel principal faults forming a pop-up (or positive flower) structure with significant fissuring along the fault traces. The two rupture traces may or may not be shown on a smaller scale rupture map. Vertical slip can be measured on both sides of the structure; however, we have found that documentation of slip measurements on structures like this is sometimes ambiguous. Ideally, both rupture lines would be shown on a map, the measurement sites would correlate with the linework, and fault motion indicators (i.e., upside/downside for measurements or ruptures) would be reported. Due to positional location errors, measurement sites might not be precisely correlated with the linework; alternatively, the dataset originator might report two measurements at the center of the structure. In both cases, it can be challenging to discern the measured feature unless the dataset originator indicates the measurement is part of a pop-up structure. Further ambiguity can arise from the rupture mapping scale, as structures like this might be represented with one approximated rupture line. While there is significant fissuring on both sides of the pop-up structure, reporting standards vary: fissures may or may not be noted in measurement comments, and the width or depth of fissures is not always reported or identified as contributing to a reported net slip measurement.

Figure 2.9. Photograph along portion of 2016 Kaikoura, New Zealand **M** 7.8 surface rupture. Source: Madugo, C.

2.3 DATA COLLECTION METHODS

There are three general methods for mapping surface ruptures and measuring fault displacements: field-based observation, interpretation of remotely collected data, and automated or semi-automated analysis of pre- and post-earthquake digital data (Table 2.1). Typical data sources, tools and techniques, and advantages and disadvantages of the methods are discussed below.

Table 2.1. Generalized data source and analysis methods for surface rupture mapping and fault displacement measurements.

Type	Basis	Method
Field-based	Digital maps or imagery, printed maps, aerial photographs, GPS positioning	Ruptures and measurements are assessed on the ground or from low altitude aerial reconnaissance
Remote	Post-event aerial photographs, orthophotographs, satellite imagery, lidar	Ruptures and measurements are interpreted offsite
Automated/Semi-Automated	Geospatially-controlled and co-registered pre- and post-earthquake digital data	Rupture locations and measurements are calculated using differencing or change detection algorithms; results are interpreted for geologic consistency

2.3.1 Field-Based

Field-based methods include the conventional "boots on the ground" geologic mapping of surface ruptures on various media, such as digital or printed maps or aerial photographs. A key advantage of this method is the potential to document site-specific complexity if the mapping scale is sufficiently small. The related disadvantage is that artificial variability in fault trace complexity may be inadvertently documented if the scale of the base map changes along the length of the total surface rupture. Another key drawback of ground-based mapping is the potential to overlook surface ruptures outside the area surveyed. Reconnaissance air-based mapping is often used to guide ground teams to the spatial extent of surface ruptures. In areas with difficult terrain, restricted access, or when ground-based mapping is limited by personnel or time, air-based mapping might be the only method used.

A key advantage of field-based measurements is that it can be easier to confirm the measured displacements are only from the most recent earthquake in the field, as some offset geomorphic features may record displacements from multiple events. Fault displacements are measured with tools like measuring tapes, folding rulers, and hand levels or leveling staffs. These tools are usually better suited for discrete ruptures or distinct piercing points, but they are also used along warps and folds. Surveying methods and tools are also commonly used to construct profiles

across (sub)vertical scarps (e.g., warps and folds) and along laterally-offset linear features. Surveying instruments, such as pole-mounted Global Navigation Satellite System (GNSS) receivers or total station theodolites, record the location and elevation of points along the profile. The data are collected in the field and analyzed or interpreted in the office. Measurement locations are either manually determined and annotated on various media (e.g., digital or printed maps or imagery) or automatically determined using portable GNSS devices such as handheld units, mobile phones, or tablet computers.

2.3.2 Remote

Remote methods require office-based geologic interpretation of remotely collected post-earthquake data. Typical data types include aerial photographs, orthophotographs, satellite imagery, unmanned aerial vehicle (UAV) photographs, or lidar collected after the earthquake. Surface ruptures and offset features are identified by geologic interpretation of the imagery or hillshades created from lidar-based digital elevation models. A key advantage of remote methods is the ability to quickly collect data over a large area, increasing the likelihood observations are collected before post-earthquake ground surface modification (e.g., infrastructure repairs, weather/storm events) obscures the rupture.

The ability to recognize ruptures and measure offset features is limited by the resolution of the imagery or elevation model, and vertical components of displacement cannot be measured from imagery. Misinterpretations can occur when mapping exclusively from remote data without any field verification, and the propensity for error is related to the dataset resolution, vegetation, pre-existing faulting, and the complexity of the rupture. For example, surface ruptures may be unobservable in lower resolution data, dense vegetation, or earthquakes with complex and diffuse ruptures, and offset features can be difficult to measure in dense vegetation or areas with diffuse ruptures. Pre-existing faults or lineaments that did not rupture in the earthquake of interest could be misinterpreted as fresh surface ruptures, and it can be difficult to discern displacements due exclusively to the most recent earthquake.

While misinterpretations can occur when mapping exclusively from remote data without any field verification, some physical settings are well-suited to remote interpretation. For example, maintained agricultural fields are usually flat with aligned rows of plants, and manmade structures like fences, walls, and roads provide excellent piercing points. Ideally, remote and field-based methods are used in tandem or iteratively to develop a verified surface rupture map and dataset of fault displacements.

2.3.3 Automated or Semi-Automated

Automated or semi-automated methods use differencing or change detection analysis and require geospatially-controlled and co-registered pre- and post-earthquake digital data. These methods

include optical image correlation (also called pixel mapping) and differential lidar analysis. Mapping surface ruptures and measuring fault displacement are both possible, and a key advantage of these methods is the ability to collect and uniformly analyze data, with a high degree of spatial accuracy, over a large area. In particular, displacements can be measured across kilometer-scale apertures[5]. However, the results can be sensitive to the aperture over which displacement is measured (e.g., Gold et al., 2015; Zinke et al., 2014), and the threshold detection limits are a function of the spatial resolution of the pre- and post-earthquake datasets. Ideally, remote and field-based methods are used in conjunction with these methods to develop a verified surface rupture map and to understand how displacement measurements vary between different apertures or scales.

Surface rupture identification varies depending on the data and technique, but fully automated approaches generally extract points or cells of maximum pre- and post-earthquake positional differences for various gradients, such as displacement, strain, or rotation, (e.g., Howell et al., 2020; Milliner et al., 2016, 2021) and linearize or connect the cells using a greedy (spatial optimization) algorithm or similar approach (Milliner et al., in prep.). Semi-automated approaches, in which an analyst interprets and manually digitizes rupture linework from gradient maps, are more common (e.g., Milliner et al., 2015, 2016; Zinke et al., 2019; Nissen et al., 2014). This allows the analyst to apply geologic judgement to identify rupture ends, manually adjust areas with decorrelation artifacts, and manually differentiate areas of landsliding or lateral spreading. However, due to the data resolution and analysis procedures, closely-spaced parallel faults usually cannot be detected, and distinguishing between continuous linear ruptures, en-echelon rupture patterns, or non-brittle warping can be difficult.

Displacement measurement methods vary depending on the data and technique, but in general pre- and post-earthquake differences, such as displacement or strain, are calculated for each point or cell. Fault-normal profiles are constructed along the length of a defined rupture, and a functional form is fit across the profile to calculate the relative displacement on each side of the fault (e.g., Milliner et al., 2015, 2016; Gold et al., 2013). The length and width of the profiles varies depending on the data and technique. The key advantage to using these methods to measure fault displacement is that the total, or wide-aperture, fault displacement can be calculated, as well as the fault zone width and the accumulation of displacement across the zone. When sufficient elevation control is available, vertical displacement can also be calculated (Oskin et al., 2012; Nissen et al., 2014; Zinke et al., 2019). However, the results need to be reviewed to distinguish landsliding, lateral spreading, and natural or engineered landform changes from fault displacement (Nissen et al., 2014; Howell et al., 2020; Zinke et al., 2019).

[5] Measurement aperture is the length or area over which fault displacement is measured. Displacement outside the aperture window is not observed and therefore cannot be detected.

2.4 MEASUREMENT UNCERTAINTY

The epistemic uncertainty of fault displacement measurements is a function of several factors, including but not limited to degree of preservation, knowledge of pre-rupture geometry, the quality or reliability of the offset feature, measurement aperture, local variations in style of faulting, measurement tools, and time elapsed between the earthquake and the measurement. Brief examples of the factors are provided below. The accumulation of these factors inherently requires the use of more judgment in measuring displacements, and the effects of judgment and implicit biases (e.g., anchoring and confirmation biases) on measurement uncertainty are difficult to quantify (Arrowsmith and Rockwell, 2012).

- Degree of preservation is influenced by rupture characteristics (discrete vs. diffuse), site conditions (e.g., climate, vegetation, surface material properties), and amplitude of displacement (e.g., smaller displacements are more perishable).
- Offset features can be geologic, geomorphic, or cultural, and the quality of measurements across a feature is related to how the feature intersects the rupture (i.e., obliquity) and how confidently its pre-rupture configuration can be reconstructed. For example, cultural features such as fences, roads, and canals are often linear and more easily reconstructed, whereas sinuous channel or terrace margins can permit a range of pre-rupture configurations (e.g., Gold et al., 2011; Arrowsmith and Rockwell, 2012). The style of faulting can also affect how well offset features can be reconstructed; for example, fault-normal shortening (e.g., under-thrusting), can bury or obscure features. Confidence that the feature is offset by only one earthquake is also an important part of feature reliability.
- Measurement aperture can affect reported displacement measurements because some slip can be accommodated through continuous warping tens of meters beyond a discrete rupture trace (Rockwell et al., 2002).
- Localized changes in deformation mechanisms (e.g., displacement components or style of faulting) due to structural complexity or sudden changes in surface material properties can obscure or confuse displacement measurements. For example, rupture expression and displacements can be exaggerated or obfuscated by local refraction processes or when a rupture crosses from native soil onto hardscapes such as asphalt or concrete.
- Remote-based (Table 2.1) measurements are limited by the resolution of the dataset. Vegetation or atmospheric obstructions can limit imagery-based assessments.
- Measurements collected immediately after an earthquake are less likely to reflect afterslip or degradation.

In our experience compiling and analyzing the database, field- and remote-based (Table 2.1) displacement measurements are most often reported as a single preferred value. When

uncertainties are documented, they are more commonly symmetrical (i.e., +/- value) and only rarely asymmetrical. Asymmetrical uncertainties are generally more robust because they reflect explicit evaluation of both the range of pre-rupture reconstructions and the most likely value (e.g., Gold et al., 2011; Scharer et al., 2014). Documentation on the meaning and application of uncertainties is also rare, in part because the preferred displacement and uncertainties do not represent a mean and standard deviation; however, it is typically assumed that uncertainties represent a minimum and maximum (Scharer et al., 2014). Conversely, measurements calculated using differencing or change detection algorithms (e.g., differential lidar, optical image correlation) are usually reported as mean values with a standard deviation or confidence interval (Milliner et al., 2016; Gold et al., 2015).

In addition to measurement uncertainty (i.e., the epistemic uncertainty discussed above), measurement errors can also occur. Gold et al. (2013) conducted a systematic evaluation of measurement errors using field- and remote-based methods at three locations in the 2010 **M** 7.2 El Mayor-Cucapah, Mexico earthquake. They conclude an uncertainty (two standard deviations or 2σ) of 11% to 17% is necessary to capture measurement errors.

2.5 TERMINOLOGY

The terminology used to describe surface rupture patterns and fault displacement measurements varies in professional literature and standard practice. In this Chapter, we identify relevant terms used by geologists and define how they are used in the FDHI Database Project.

2.5.1 Fault Displacements

Several terms are used by geologists to describe the magnitude, amplitude, or amount of ground surface movement across a surface rupture. For example, the terms *displacement*, *slip*, *separation*, and *offset* are inconsistently and interchangeably applied in professional literature and standard practice. For this project, we define *slip* as the actual relative displacement and *separation* as the apparent relative displacement. We use *offset* as a verb or adjective when describing the geomorphic or anthropogenic features displaced across a rupture. The terms *style*, *sense*, and *direction* all describe the relative movement of the ground surface across a surface rupture.

Surface ruptures generate three-dimensional ground surface displacements. As a result, several components of relative displacement (slip) can be measured (Figure 2.10). The slip components are also sometimes called slip vectors; however, recognizing that the term *vector* has a specific meaning in math and engineering (i.e., magnitude and azimuthal direction), we use *component* instead, because the azimuthal direction (i.e., rake) is rarely reported. The measured slip component is usually clearly documented in the original datasets, and we retained this information in the database. The style of slip (e.g., normal, reverse, left-lateral, right-lateral) is also usually reported in the original datasets and retained in the database.

Based on the geometric relationships in Figure 2.10, the net (three-dimensional) slip can be calculated from the net horizontal slip and dip slip measurements if it is not directly measured (Equations 2.1 through 2.3). In a pure strike-slip offset, the fault-parallel slip (FPS) is equivalent to the net slip (TDS). Similarly, in a pure normal or pure reverse offset, the dip slip (ADS) is equivalent to the net slip (TDS).

$$TDS = \sqrt{NHS^2 + VS^2} \tag{2.1}$$

$$NHS = \sqrt{FPS^2 + FNS^2} \tag{2.2}$$

$$ADS = \sqrt{FNS^2 + VS^2} \tag{2.3}$$

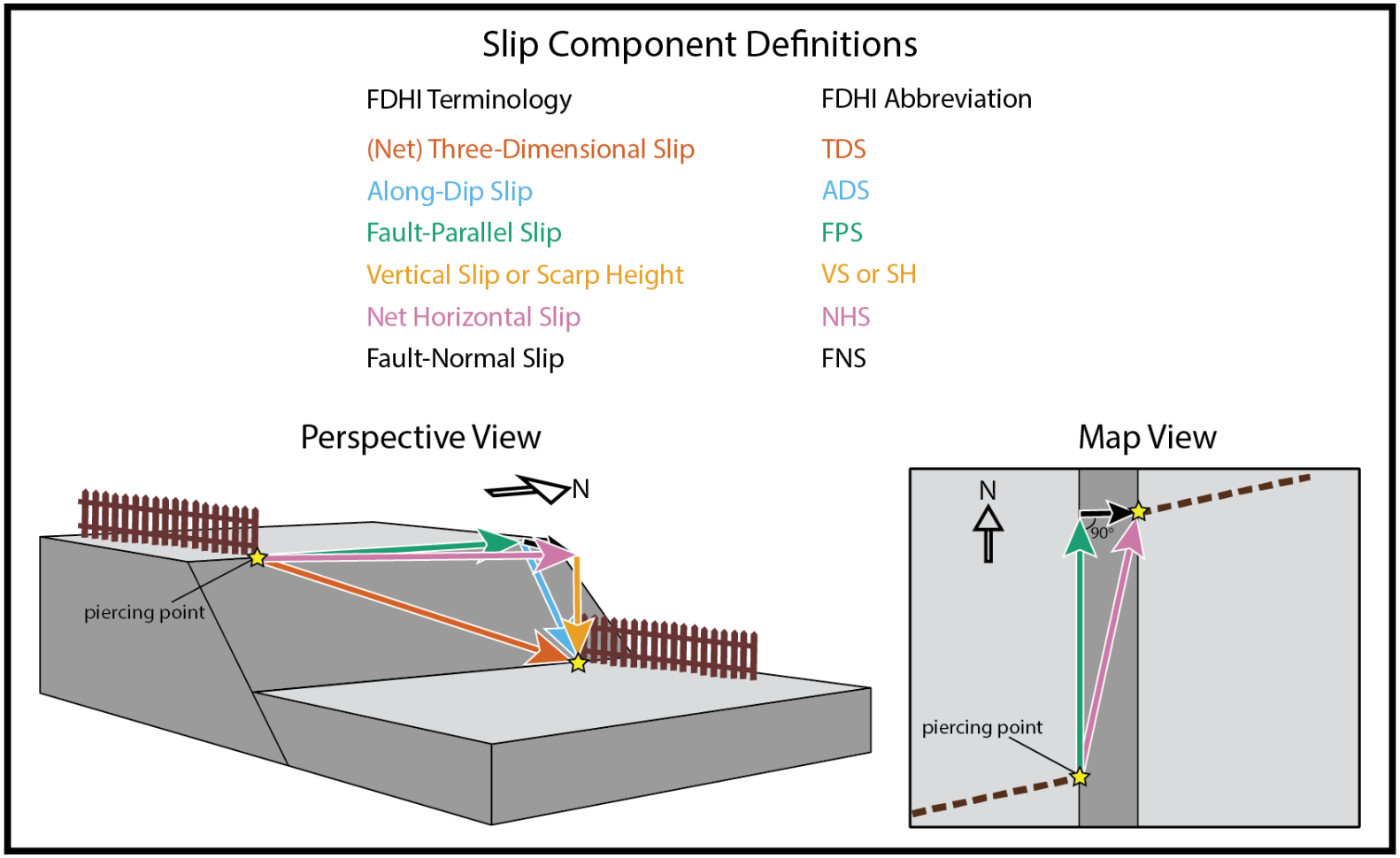

Figure 2.10. Fault displacement slip component definitions used in FDHI Database. Adapted from Ponti et al. (2020).

Vertical displacement measurements are sensitive to the slope of the ground surface before the earthquake (Caskey, 1995; Yang et al., 2015) (Figure 2.11). If the pre-earthquake ground surface was not flat, then vertical measurements that only consider the distance between the piercing points (orange arrow in Figure 2.10) are an apparent displacement (separation). The actual vertical component of the relative displacement (slip) is measured by projecting the original ground surface across the rupture (Figure 2.11). We refer to the actual vertical component of the relative displacement as *vertical slip* and the apparent vertical component of the relative displacement as *scarp height* (Figures 2.10 and 2.11). As shown in Figure 2.11, vertical slip and scarp height measurements can vary significantly depending on the slope and direction of the original ground surface. Fortunately, this distinction in vertical displacement measurements is usually documented in original datasets and retained in the database.

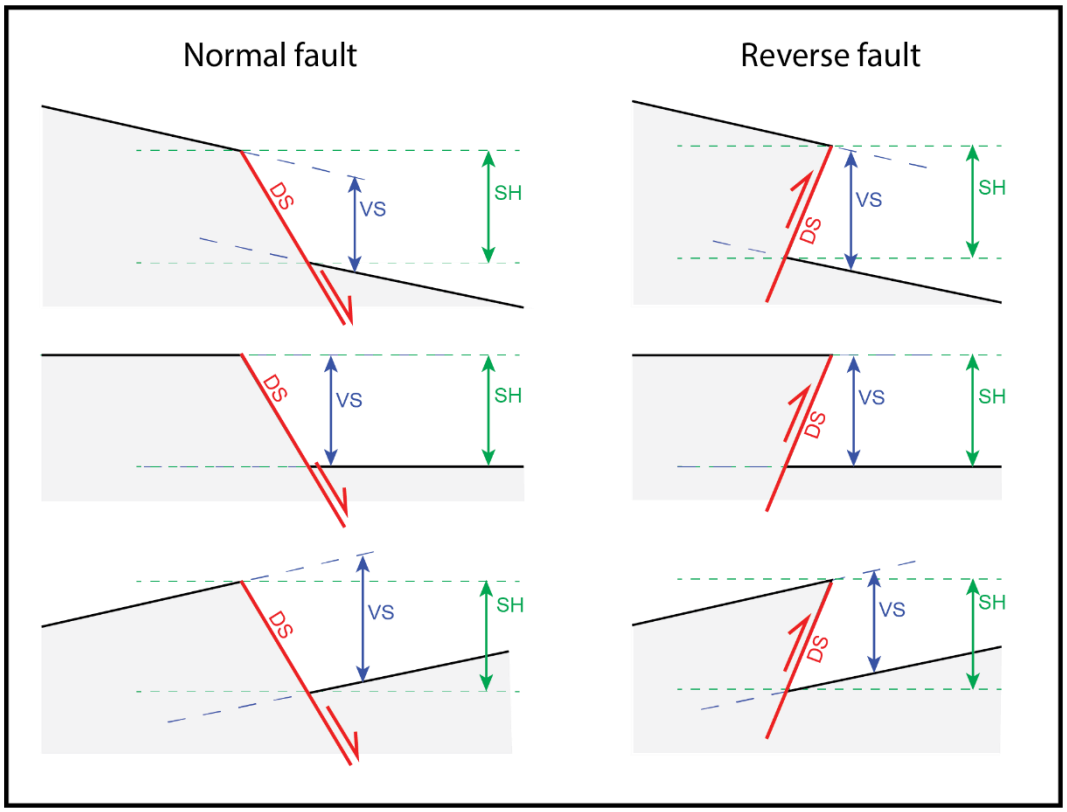

Figure 2.11. Schematic ground surface configurations and vertical fault displacement measurements for normal and reverse faults (profile view). DS = dip-slip; VS = vertical slip; SH = scarp height. Note that vertical slip and scarp height are vertical components of the slip vector in a Cartesian reference frame.

2.5.2 Discrete Slip and Continuous Deformation

Discrete expressions of surface rupture are clear locations of significant fault displacement (Figures 2.4 through 2.6). Measurements on discrete surface ruptures represent an infinitely narrow

deformation zone width. conversely, broad warps represent continuous deformation across a wider zone (Figure 2.7). Although continuous deformation is easier to visualize with vertical displacement (Figure 2.7), lateral fault displacement also produces shear zones of continuous deformation that are evident when a continuous linear feature is offset (Figure 2.12). Measurements across broad warps or shears correspond to a specific deformation zone width and are measured within a specific aperture. In the schematic in Figure 2.12, all of the displacement (discrete slip plus continuous deformation) occurs between marker numbers 3 and 12, which represents the deformation zone width, and the surveyed length extends from marker numbers 1 to 14, representing the measurement aperture. While profile-based measurements are relatively common, measurement aperture and deformation zone width are rarely reported in the original datasets, and we include this information in the database when it is available. Fault zone width, which we interpret to represent the zone encompassing a network of closely-spaced ruptures, is sometimes reported and included in the database; however, the terms fault zone width and deformation zone width are often used ambiguously in professional literature and standard practice.

2.5.3 Principal (Primary) and Distributed (Secondary) Faulting

Surface ruptures are sometimes categorized as principal or distributed to reflect causative sources or relative significance of individual ruptures. The terms *principal* and *distributed* are sometimes used interchangeably with *primary* and *distributed*, respectively, in professional literature and standard practice. For this project, we use *primary* and *distributed*¸ and we follow the definitions in Coppersmith and Youngs (2000) and Youngs et al. (2003):

- **Principal faulting** is slip on the primary faults or tectonic/seismogenic features responsible for the earthquake.
- **Distributed faulting** is the secondary slip that occurs on other faults, splays, fractures, or shears near the principal fault.

As discussed in detail in Chapter 4.3 of this report, the FDHI Database includes classifications of rupture linework as principal or distributed, based on information reported in the original data sources or interpreted by the Database Team.

Fault displacement measurements can also be categorized as principal or distributed, based on the classification of their associated surface rupture (Youngs et al., 2003; Petersen et al., 2011, Nurminen et al., 2020). We introduce two additional measurement categories (cumulative and total) in this project to better distinguish measurements associated with multiple ruptures or wider measurement apertures. Cumulative measurements represent slips summed across multiple known principal ruptures, one principal rupture and one or more distributed ruptures, or principal rupture(s) plus continuous deformation (e.g., Figure 2.12). Total measurements represent wide-aperture slips calculated from the differencing or image correlation methods discussed in Chapter 2.3.3. Because total measurements capture the total displacement across a wide aperture, the

measurement might not be associated with a single specific fault. The process of classifying fault displacement measurements as principal, distributed, cumulative, or total is discussed in Chapter 4.3 of this report.

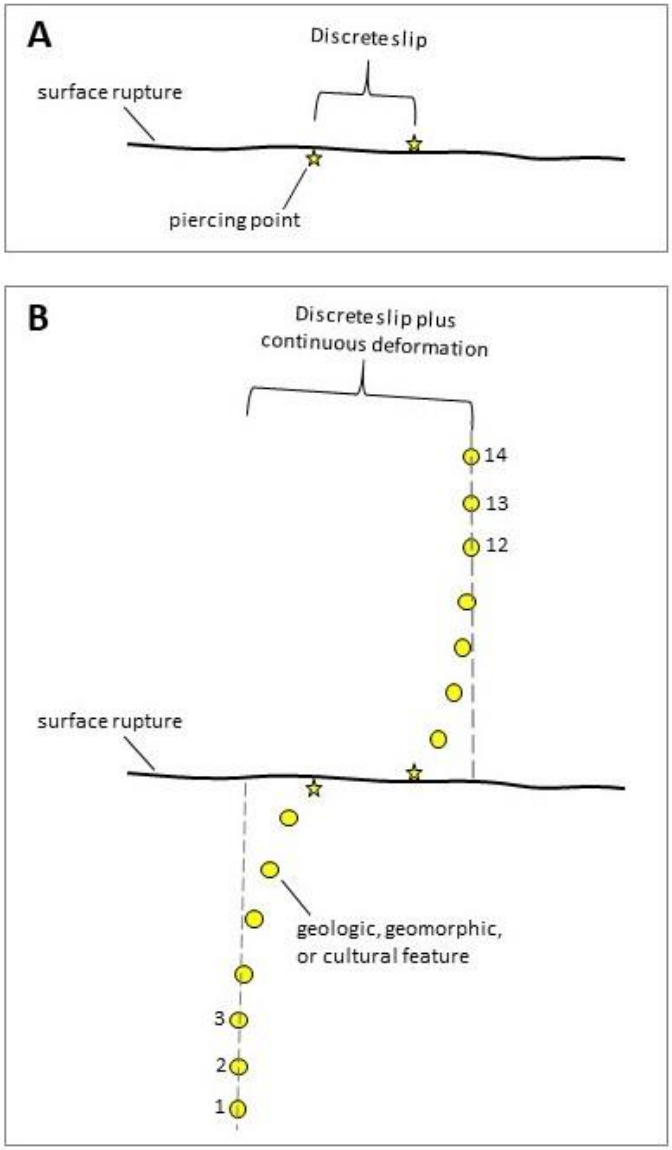

Figure 2.12. Plan-view schematics of right-laterally offset piercing points slip measurements (not to scale). (A) Narrow-aperture measurement captures discrete slip. (B) Wider aperture measurement discrete slip and continuous warping or deformation.

2.6 REFERENCES

Arrowsmith, J. R. and Rockwell, T. K. (2012). Repeatability, accuracy, and precision of surface slip measurements from high-resolution topographic data: Collaborative research with Arizona State University and San Diego State University, *Final Technical Report U.S. Geological Survey National Earthquake Hazards Reduction External Research Program*, USGS Award Number G11AP20029 (ASU) & G11AP20020 (SDSU).

Avery, T. E. and Berlin, G. L. (1992). Fundamentals of remote sensing and airphoto interpretation. Macmillan.

Bray, J. D., Seed, R. B., Cluff, L. S., and Seed, H. B. (1994). Earthquake fault rupture propagation through soil, *Journal of Geotechnical Engineering*, **120**(3), 543-561.

Brown, R. D. Jr. and Vedder, J. G. (1966). Surface tectonic fractures along the San Andreas Fault, The Parkfield-Cholame California earthquakes of June-August 1966, *U.S. Geological Survey Professional Paper 579.*

Caskey, S. J. (1995). Geometric relations of dip slip to a faulted ground surface: new nomograms for estimating components of fault displacement, *Journal of Structural Geology*, **17**(8), 1197-1202.

Coppersmith, K. J. and Youngs, R. R. (2000). Data needs for probabilistic fault displacement hazard analysis, *Journal of Geodynamics*, **29**(3-5), 329-343.

DuRoss, C. B., Gold, R. D., Dawson, T. E., Scharer, K. M., Kendrick, K. J., Akciz, S. O., Angster, S. J., Bachhuber, J., Bacon, S., Bennett, S. E. K., et al. (2020). Surface Displacement Distributions for the July 2019 Ridgecrest, California, Earthquake Ruptures, *Bulletin of the Seismological Society of America*, **110**(4), 1400-1418.

Fletcher, J. M., Teran, O. J., Rockwell, T. K., Oskin, M. E., Hudnut, K. W., Mueller, K. J., Spelz, R. M., Akciz, S. O., Masana, E., Faneros, G., Fielding, E. J., Leprince, S., Morelan, A. E., Stock, J., Lynch, D. K., Elliott, A. J., Gold, P., Liu-Zeng, J., González-Ortega, A., Hinojosa-Corona, A., and González-García, J. (2014). Assembly of a large earthquake from a complex fault system: Surface rupture kinematics of the 4 April 2010 El Mayor–Cucapah (Mexico) Mw 7.2 earthquake, *Geosphere*, **10**(4), 797-827.

Gold, R. D., Reitman, N. G., Briggs, R. W., Barnhart, W. D., Hayes, G. P., and Wilson, E. (2015). On-and off-fault deformation associated with the September 2013 Mw 7.7 Balochistan earthquake: Implications for geologic slip rate measurements, *Tectonophysics*, **660**, 65-78.

Gold, P. O., Oskin, M. E., Elliott, A. J., Hinojosa-Corona, A., Taylor, M. H., Kreylos, O., and Cowgill, E. (2013). Coseismic slip variation assessed from terrestrial LiDAR scans of the El Mayor–Cucapah surface rupture, *Earth and Planetary Science Letters*, **366**, 151-162.

Gold, R. D., Cowgill, E., Arrowsmith, J. R., Chen, X., Sharp, W. D., Cooper, K. M., and Wang, X. F. (2011). Faulted terrace risers place new constraints on the late Quaternary slip rate for the central Altyn Tagh fault, northwest Tibet, *Geological Society of America Bulletin*, **123**(5-6), 958-978.

Howell, A., Nissen, E., Stahl, T., Clark, K., Kearse, J., Van Dissen, R., Villamor, P., Langridge, R., and Jones, K. (2020). Three-dimensional surface displacements during the 2016 Mw 7.8 Kaikōura earthquake (New Zealand) from photogrammetry-derived point clouds, *Journal of Geophysical Research: Solid Earth*, **125**(1), e2019JB018739.

Milliner, C., and Donnellan, A. (2020). Using daily observations from Planet Labs satellite imagery to separate the surface deformation between the 4 July Mw 6.4 foreshock and 5 July Mw 7.1 mainshock during the 2019 Ridgecrest earthquake sequence, *Seismological Research Letters*, **91**(4), 1986-1997.

Milliner, C., Donnellan, A., Aati, S., Avouac, J. P., Zinke, R., Dolan, J. F., Wang, K., and Bürgmann, R. (2021). Bookshelf Kinematics and the Effect of Dilatation on Fault Zone Inelastic Deformation: Examples from Optical Image Correlation Measurements of the 2019 Ridgecrest Earthquake Sequence, *Journal of Geophysical Research: Solid Earth*, **126**(3), e2020JB020551.

Milliner, C. W. D., Dolan, J. F., Hollingsworth, J., Leprince, S., and Ayoub, F. (2016). Comparison of coseismic near-field and off-fault surface deformation patterns of the 1992 Mw 7.3 Landers and 1999 Mw 7.1 Hector Mine earthquakes: Implications for controls on the distribution of surface strain, *Geophysical Research Letters*, **43**(19), 10-115.

Milliner, C. W., Dolan, J. F., Hollingsworth, J., Leprince, S., Ayoub, F., and Sammis, C. G. (2015). Quantifying near-field and off-fault deformation patterns of the 1992 Mw 7.3 Landers earthquake, *Geochemistry, Geophysics, Geosystems*, **16**(5), 1577-1598.

Morelan, A. E. and Hernandez, J. L. (2020). Increasing Postearthquake Field Mapping Efficiency with Optical Image Correlation, *Bulletin of the Seismological Society of America*, **110**(4), 1419-1426.

Moss, R. E. S., Stanton, K. V., and Buelna, M. I. (2013). The impact of material stiffness on the likelihood of fault rupture propagating to the ground surface, *Seismological Research Letters*, **84**(3), 485-488.

Nissen, E., Maruyama, T., Arrowsmith, J. R., Elliott, J. R., Krishnan, A. K., Oskin, M. E., and Saripalli, S. (2014). Coseismic fault zone deformation revealed with differential lidar: Examples from Japanese Mw ~7 intraplate earthquakes, *Earth and Planetary Science Letters*, **405**, 244-256.

Nurminen, F., Boncio, P., Visini, F., Pace, B., Valentini, A., Baize, S., and Scotti, O. (2020). Probability of occurrence and displacement regression of distributed surface rupturing for reverse earthquakes, *Frontiers in Earth Science*, **8**(581605).

Oskin, M. E., Arrowsmith, J. R., Corona, A. H., Elliott, A. J., Fletcher, J. M., Fielding, E. J., Gold, P. O., Garcia, J. J. G., Hudnut, K. W., Liu-Zeng, J., and Teran, O. J. (2012). Near-field deformation from the El Mayor–Cucapah earthquake revealed by differential LIDAR, *Science*, **335**(6069), 702-705.

Petersen, M. D., Dawson, T. E., Chen, R., Cao, T., Wills, C. J., Schwartz, D. P., and Frankel, A. D. (2011). Fault displacement hazard for strike-slip faults, *Bulletin of the Seismological Society of America*, **101**(2), 805-825.

Ponti, D. J., Blair, J. L., Rosa, C. M., Thomas, K., Pickering, A. J., Akciz, S., Angster, S., Avouac, J.-P., Bachhuber, J., Bacon, S., et al. (2020). Documentation of Surface Fault Rupture and Ground-Deformation Features Produced by the 4 and 5 July 2019 Mw 6.4 and Mw 7.1 Ridgecrest Earthquake Sequence, *Seismological Research Letters*, **91**(5), 2942-2959.

Rockwell, T. K., Lindvall, S., Dawson, T., Langridge, R., Lettis, W., and Klinger, Y. (2002). Lateral offsets on surveyed cultural features resulting from the 1999 Izmit and Duzce earthquakes, Turkey, *Bulletin of the Seismological Society of America*, **92**(1), 79-94.

Scharer, K. M., Salisbury, J. B., Arrowsmith, J. R., and Rockwell, T. K. (2014). Southern San Andreas fault evaluation field activity: Approaches to measuring small geomorphic offsets—Challenges and recommendations for active fault studies, *Seismological Research Letters*, **85**(1), 68-76.

Teran, O. J., Fletcher, J. M., Oskin, M. E., Rockwell, T. K., Hudnut, K. W., Spelz, R. M., Akciz, S. O., Hernandez-Flores, A. P., and Morelan, A. E. (2015). Geologic and structural controls on rupture zone fabric: A field-based study of the 2010 Mw 7.2 El Mayor–Cucapah earthquake surface rupture, *Geosphere*, **11**(3), 899-920.

Turcotte, D. L. (1990). Implications of chaos, scale-invariance, and fractal statistics in geology, *Global and Planetary Change*, **3**(3), 301-308.

United States Geological Survey (USGS) (2021). Earthquake Photo Collections. URL: https://www.usgs.gov/natural-hazards/earthquake-hazards/science/earthquake-photo-collections?qt-science_center_objects=0#qt-science_center_objects.

Yang, X., Li, W., and Qin, Z. (2015). Calculation of reverse-fault-related parameters using topographic profiles and fault bedding, *Geodesy and Geodynamics*, **6**(2), 106-112.

Youngs, R. R., Arabasz, W. J., Anderson, R. E., Ramelli, A. R., Ake, J. P., Slemmons, D. B., McCalpin, J. P., Doser, D. I., Fridrich, C. J., Swan III, F. H., Rogers, A. M., Yount, J. C., Anderson, L. W., Smith, K. D., Bruhn, R. L., Knuepfer, P. L. K., Smith, R. B., dePolo, C. M., O'Leary, D. W., Coppersmith, K. J., Pezzopane, S. K., Schwartz, D. P., Whitney, J. W., Olig, S. S., and Toro, G. R. (2003). A methodology for probabilistic fault displacement hazard analysis (PFDHA), *Earthquake Spectra*, **19**(1), 191-219.

Zinke, R., Hollingsworth, J., Dolan, J. F., and Van Dissen, R. (2019). Three-dimensional surface deformation in the 2016 Mw 7.8 Kaikōura, New Zealand, earthquake from optical image correlation: Implications for strain localization and long-term evolution of the Pacific-Australian plate boundary, *Geochemistry, Geophysics, Geosystems*, **20**(3), 1609-1628.

Zinke, R., Hollingsworth, J., and Dolan, J. F. (2014). Surface slip and off-fault deformation patterns in the 2013 Mw 7.7 Balochistan, Pakistan earthquake: Implications for controls on the distribution of near-surface coseismic slip, *Geochemistry, Geophysics, Geosystems*, **15**(12), 5034-5050.

3 Data Collection

This Chapter documents the data collection approach for the FDHI Database Project. The data collection process for the FDHI Database followed a systematic approach that entailed defining event and dataset selection criteria based on model developer needs, reviewing existing compilations, and developing and implementing a standard workflow to review and process data. Each candidate dataset was carefully reviewed for quality and compliance with the event and dataset selection criteria.

3.1 SELECTION CRITERIA

The event and dataset selection criteria for the FDHI Database were guided by model development needs and project timelines. Specific criteria for events and dataset contents are discussed in this Chapter. In general, we found the dataset criteria (i.e., dataset contents) to be more limiting than the event criteria. To support the model development project schedule, we enforced a database entry cut-off date of October 2020 by ceasing our data search and compilation efforts.

3.1.1 Event Criteria

The event selection criteria for the FDHI Database Project were broadly constrained to historical surface-rupturing crustal earthquakes. We considered global, shallow crustal events of all styles of faulting, all magnitudes, and both active and stable tectonic regimes. We limited our criteria to historical earthquakes to ensure the data capture single-event ruptures and displacements only. By constraining our criteria to events that produced surface rupture, we do not include information on earthquakes that did not reach the ground surface. The database does not include an exhaustive collection of global historical surface-rupturing earthquakes because most events do not meet the dataset criteria (Chapter 3.1.2). As a result, several significant historical surface-rupturing earthquakes are not in the database (Chapter 3.4).

3.1.2 Dataset Criteria

The dataset selection criteria generally relate to data content and quality. Only historical, inferred single-event measurement and rupture data from tectonic faulting are included (i.e., shaking related

features and paleoseismic data were excluded). The minimum required content included geospatial control for fault displacement measurements and mapped ruptures. This is an important criterion because the new models will consider the spatial distribution for both displacement amplitudes and surface ruptures. We also specifically sought datasets that include measurements on distributed faults and datasets with detailed surface rupture mapping. Given the extensive areas involved in surface rupture of large earthquakes, some datasets are limited to specific portions of the rupture; therefore, we included multiple slip measurement datasets for the same event, when available, for more complete spatial coverage. In some cases, multiple slip measurements are available for the same site, and they are included in the database. All data were collected from published reports and journal papers and reviewed for quality (Chapter 3.3.2). We specifically excluded datasets with insufficient geospatial control, irrelevant content, or known errors or issues (Chapter 3.4).

As discussed in Chapter 3.2, the data for many older events are documented as displacement profiles, with limited or no information on measurement site location, distributed displacements, or mapped ruptures. Therefore, we found the geospatial control requirement to be the most limiting factor in including additional events. We also observed a general paucity of distributed rupture and displacement information reported in older events, which we infer to be a data completeness issue and not reflective of earthquake-specific characteristics.

3.2 EXISTING COMPILATIONS

Several previous studies have compiled observational data from historical surface-rupturing earthquakes (Table 3.1). The existing compilations vary in their content and completeness, generally due to limitations in available data and the intended usage of the compilation. There is also considerable event overlap between the compilations, owing to the limited number of historical surface-rupturing earthquakes. For example, the Baize et al. (2020) compilation includes the Takao et al. (2018) data, and the Nurminen et al. (2020) compilation includes the Boncio et al. (2018) data.

We reviewed the existing compilations to identify candidate events and datasets that met our selection criteria (Table 3.1). Our review was limited to published reports and journal papers, as we did not have access to proprietary compilations. In general, compilations with geospatial (two-dimensional) control for displacements and mapped ruptures were considered further for inclusion in the FDHI Database. Each candidate dataset was carefully reviewed for quality (i.e., completeness, accuracy, and consistency; Chapter 3.3.2), and a subset of events in the available compilations that did not meet our data quality requirements were not included in the database (Chapter 3.4).

Table 3.1. Review of existing compilations.

Compilers	No. of Events / Style[1]	Data Format	Reported Displacement	Displacement Spatial Control[3]	Surface Rupture Mapping	Rupture Spatial Control[3]	Meets Initial FDHI Database Selection Criteria
Wells & Coppersmith (1994)	244 / All	Tabulated	Max & mean	None	None	None	No
Pezzopane & Dawson (1996)	9 / NML & SS	Maps	Principal & distributed	2-D	Principal & distributed	2-D	Yes
Wesnousky (2008)	37 / All	Displacement profiles	Principal / main trace(s)	1-D	Principal	None	No
Petersen et al. (2011)	24 / SS	Displacement profiles, GIS	Principal & distributed	Varies, 1-D / 2-D	Varies	Varies, 2-D / none	Yes (subset)
Takao et al. (2018)	22 / SS & RV	GIS	Principal & distributed	2-D	Principal & distributed	2-D	Yes
Boncio et al. (2018)	11 / RV	Maps, GIS[2]	Distributed	2-D	Principal & distributed	2-D	Yes
Baize et al. (2020)	45 / All	GIS[2]	Principal & distributed	2-D	Varies	2-D	Yes
Nurminen et al. (2020)	15 / RV	GIS[2]	Distributed	2-D	Principal & distributed	2-D	Yes[4]

[1] Style of faulting abbreviations: SS = Strike-Slip; NML = Normal; RV = Reverse; OBL = Oblique.

[2] Includes tabulated data with geographic coordinates.

[3] One-dimensional (1-D) and two-dimensional (2-D).

[4] Compilation was released after database expansion cut-off date and is only partially included.

3.3 STANDARD WORKFLOW

We followed a standard workflow for events and datasets meeting the selection criteria to ensure the necessary data and metadata were collected, reviewed, and formatted for import into the relational database. In the following Chapters, we discuss the event metadata development and workflows for collecting the surface rupture, measurement, and geologic data and metadata.

3.3.1 Event Metadata

Earthquakes are identified in the FDHI Database by a unique integer called the earthquake identification number ("EQ_ID"). The earthquake identification number is based on the order the event was added to the database; therefore, the identification numbers are arbitrary and are not in event chronological order. Event metadata collected for the FDHI Database consisted of the common name, magnitude, magnitude type, style of faulting, region, and origin information (date in Coordinated Universal Time and hypocentral location). The seismic moment (in dyne-centimeters) was calculated based on the magnitude per Hanks and Kanamori (1979), except where it was directly reported. Earthquake metadata were assembled from the professional literature, including peer-reviewed journal publications and published reports. We preferentially selected metadata from the NGA-West2 database, which was subject to robust QA/QC process (Ancheta et al., 2013), when available; otherwise, we used an authoritative journal publication (preferred), the USGS database, or the local geoscience authority.

Table 3.2 lists the basic event metadata for all 75 earthquakes in the FDHI Database. This information, along with hypocentral locations and seismic moment, is also included in the flatfiles (Appendix A). The sources for the magnitude, style, and hypocenter metadata are also listed in Table 3.2.

Table 3.2. Event metadata.

EQ_ID	Name	Region	Date	Style[1]	Magnitude, Type[2]	Magnitude / Style Source	Hypocenter Source
1	Landers	California	6/28/1992	SS	7.28, Mw	NGA-West2	NGA-West2
2	HectorMine	California	10/16/1999	SS	7.13, Mw	NGA-West2	NGA-West2
3	EMC	Mexico	4/4/2010	NML-OBL	7.2, Mw	NGA-West2 / Teran et al. (2015)	NGA-West2
4	Balochistan	Pakistan	9/24/2013	SS	7.7, Mw	Gold et al. (2015)	USGS
5	Izmit_Kocaeli	Turkey	8/17/1999	SS	7.51, Mw	NGA-West2	NGA-West2
6	Borrego	California	4/9/1968	SS	6.63, Mw	NGA-West2	NGA-West2
7	Imperial1979	California	10/15/1979	SS	6.53, Mw	NGA-West2	NGA-West2
8	SuperstitionHills	California	11/24/1987	SS	6.54, Mw	NGA-West2	NGA-West2
9	Kobe	Japan	1/16/1995	SS	6.9, Mw	NGA-West2	NGA-West2
10	Denali	Alaska	11/3/2002	SS	7.9, Mw	NGA-West2	NGA-West2
11	Duzce	Turkey	11/12/1999	SS	7.14, Mw	NGA-West2	NGA-West2
12	Wenchuan	China	5/12/2008	RV-OBL	7.9, Mw	NGA-West2	NGA-West2
13	Napa	California	8/24/2014	SS	6.0, Mw	Ponti et al. (2019)	USGS
14	Yushu	China	4/13/2010	SS	6.9, Mwc	USGS	USGS
15	Hualien	Taiwan	2/6/2018	SS	6.4, Mw	Kuo et al. (2018)	USGS
16	ChiChi	Taiwan	9/20/1999	RV-OBL	7.62, Mw	NGA-West2	NGA-West2
17	Kumamoto	Japan	4/15/2016	SS	7, Mww	USGS	USGS
18	Nagano	Japan	11/22/2014	RV	6.2, Mww	USGS	USGS
19	Kashmir	Sub-Himalaya	10/8/2005	RV	7.6, Mw	Kaneda et al. (2008)	USGS
20	Kaikoura	New Zealand	11/13/2016	RV-OBL	7.8, Mw	Zinke et al. (2019)	USGS
21	Darfield	New Zealand	9/3/2010	SS	7.0, Mw	NGA-West2	NGA-West2
22	Parkfield2004	California	9/28/2004	SS	6.0, Mw	NGA-West2	NGA-West2
23	Norcia3	Italy	10/30/2016	NML	6.6, Mww	USGS	USGS
24	Hebgen	Montana	8/18/1959	NML	7.3, Mw	USGS	USGS

EQ_ID	Name	Region	Date	Style[1]	Magnitude, Type[2]	Magnitude / Style Source	Hypocenter Source
25	SanFernando	California	2/9/1971	RV	6.61, Mw	NGA-West2	NGA-West2
26	Bohol	Philippines	10/15/2013	RV	7.1, Mww	USGS	USGS
27	Acambay	Mexico	11/19/1912	NML-OBL	6.9, mB	Langridge et al. (2000)	n/a
28	Imperial1940	California	5/19/1940	SS	6.95, Mw	NGA-West2	NGA-West2
29	Parkfield1966	California	6/28/1966	SS	6.19, Mw	NGA-West2	NGA-West2
30	FairviewPeak	Nevada	12/16/1954	NML-OBL	7.3, Mw	USGS	USGS
31	DixieValley	Nevada	12/16/1954	NML	6.9, Mw	USGS	Baize et al. (2020)
32	GalwayLake	California	6/1/1975	SS	5.2, ML	Kanamori and Fuis (1976)	USGS
33	Sonora	Mexico	5/3/1887	NML-OBL	7.6, Mw	USGS	USGS
34	PleasantValley	Nevada	10/2/1915	NML	7.2, Mw	USGS	USGS
35	Kern	California	7/21/1952	RV	7.36, Mw	NGA-West2	NGA-West2
36	ChalfantValley	California	7/21/1986	SS	6.19, Mw	NGA-West2	NGA-West2
37	Zirkuh	Iran	5/10/1997	SS	7.2, Mw	Berberian et al. (1999)	Nemati (2015)
38	Petermann	Australia	5/20/2016	RV	6.0, Mw	Gold et al. (2019)	Geoscience Australia
39	OwensValley	California	3/26/1872	NML-OBL	7.4, Mw	USGS	n/a
40	LagunaSalada	Mexico	2/23/1892	NML-OBL	7.76, Mw	USGS	USGS
41	Iwaki2011	Japan	4/11/2011	NML	6.6, Mw	Toda and Tsutsumi (2013) / JMA	JMA
42	Ridgecrest1	California	7/4/2019	SS	6.4, Mw	USGS	USGS
43	Ridgecrest2	California	7/6/2019	SS	7.1, Mw	USGS	USGS
44	ElAsnam	Algeria	10/10/1980	RV	7.3, Mw	Hamdache et al. (2010) / Yielding et al. (1981)	Hamdache et al. (2010)
45	Cadoux	Australia	6/2/1979	RV	6.1, Mw	King et al. (2019)	King et al. (2019)
46	Calingiri	Australia	3/10/1970	RV	5.03, Mw	King et al. (2019)	King et al. (2019)
47	MarryatCreek	Australia	3/30/1986	RV	5.7, Mw	King et al. (2019)	King et al. (2019)
48	Meckering	Australia	10/14/1968	RV	6.59, Mw	King et al. (2019)	King et al. (2019)

EQ_ID	Name	Region	Date	Style[1]	Magnitude, Type[2]	Magnitude / Style Source	Hypocenter Source
49	Pukatja	Australia	3/23/2012	RV	5.18, Mw	King et al. (2019)	n/a
50	TennantCreek1	Australia	1/22/1988	RV	6.27, Mw	King et al. (2019)	King et al. (2019)
51	TennantCreek2	Australia	1/22/1988	RV	6.44, Mw	King et al. (2019)	King et al. (2019)
52	TennantCreek3	Australia	1/22/1988	RV	6.58, Mw	King et al. (2019)	King et al. (2019)
53	SanMiguel	Mexico	2/9/1956	SS	6.8, Ms	USGS / Doser (1992)	USGS
54	Yutian	China	2/12/2014	SS	6.9, Mw	Li et al. (2016)	USGS
55	Luzon	Philippines	7/16/1990	SS	7.7, Mwc	USGS	USGS
56	BorahPeak	Idaho	10/28/1983	NML	6.88, Mw	NGA-West2	NGA-West2
57	ElmoreRanch	California	11/24/1987	SS	6.22, Mw	NGA-West2	NGA-West2
58	Pisayambo	Ecuador	3/26/2010	SS	5.0, Mw	Champenois et al. (2017)	Champenois et al. (2017)
59	Rikuu	Japan	8/31/1896	RV	6.7, U	Baize et al. (2020)	Baize et al. (2020)
60	Mikawa	Japan	1/12/1945	RV	6.6, Mw	USGS / Baize et al. (2020)	USGS
61	IzuPeninsula	Japan	5/8/1974	SS	6.5, Ms	Baize et al. (2020)	Baize et al. (2020)
62	IzuOshima	Japan	1/14/1978	SS	6.6, Mwc	USGS / Baize et al. (2020)	Baize et al. (2020)
63	IwateInland	Japan	9/3/1998	RV	5.8, Mwc	USGS / Baize et al. (2020)	Baize et al. (2020)
64	Edgecumbe	New Zealand	3/2/1987	NML	6.6, Mw	NGA-West2	NGA-West2
65	Neftegorsk	Russia	5/27/1995	SS	7.0, Mwc	USGS	USGS
66	ChonKemin	Kyrgyzstan	1/3/1911	RV	8.02, Mw	Kulikova and Kruger (2015)	Kulikova and Kruger (2015)
67	Kunlun_Kokoxili	Northern Tibet	11/14/2001	SS	7.8, Mwc	USGS	USGS
68	LeTeil	France	11/11/2019	RV	4.9, Mww	Ritz et al. (2020) / USGS	Delouis et al. (2021)
69	Norcia1	Italy	8/24/2016	NML	6.2, Mww	USGS	USGS
70	HomesteadValley	California	3/15/1979	SS	5.2, ML	USGS	USGS
71	Palu	Indonesia	9/28/2018	SS	7.5, Mww	USGS	USGS
72	LAquila	Italy	4/6/2009	NML	6.3, Mw	NGA-West2	NGA-West2
73	Spitak	Armenia	12/7/1988	RV-OBL	6.77, Mw	NGA-West2	NGA-West2
74	Killari	India	9/29/1993	RV	6.2, Mwb	USGS	USGS
75	YeniceGonen	Turkey	3/18/1953	SS	7.3, Mw	USGS / Kürçer et al. (2019)	USGS

[1] Style of faulting abbreviations: SS = Strike-Slip; NML = Normal; RV = Reverse; OBL = Oblique

[2] Magnitude types from USGS (2021): Mw = moment magnitude, details not reported; Mwc = moment magnitude based on centroid moment tensor inversion of long-period surface waves; Mww = moment magnitude based on centroid moment tensor inversion of W-phase; mB = body-wave magnitude; ML = local magnitude; Ms = surface-wave magnitude; U = unspecified

3.3.2 Surface Rupture and Measurement Data and Metadata

The workflow for developing surface rupture and measurement information for each earthquake consisted of three steps: (1) literature review; (2) dataset processing; and (3) data quality evaluation. The first step served to identify available datasets that met the selection criteria, and the second step produced uniformly formatted datasets for import into the relational database. The third step documented our quality assessments of the data, which were ultimately used to provide data usage recommendation to the model development teams (Chapter 4.6). The goal of the workflow was to ensure all data were systematically collected and reviewed.

For each earthquake, we performed a literature review to collect candidate datasets with information bearing on surface rupture and fault displacement measurements. We began with existing compilations (Table 3.1) and the references therein and then supplemented event datasets with information from other sources when possible. Each candidate dataset was carefully reviewed for compliance with the selection criteria (particularly geospatial control, single-event data, and no known errors or issues). We reached out to dataset originators on several occasions to ask questions, confirm our intended usage of the data, or request digital source files for data presented in figures.

Multiple data sources were included for the same event in several cases, providing more complete spatial coverage of measurements and surface ruptures and/or technically defensible alternative measurements at the same location. Table 3.3 lists the measurement and surface rupture mapping sources included for all 75 events in the FDHI Database. To systematically track data sources, we assigned each dataset a unique identifier called the dataset identification number (DS_ID). The dataset identification number is based on the order the dataset was added to the database, and it is not related to a specific earthquake because some datasets present data for multiple earthquakes (e.g., existing compilations). In many cases, the most complete surface rupture maps and measurement datasets were generated by different researchers. In a few cases, multiple surface rupture datasets were available for the same earthquake. We manually combined supplementary surface rupture datasets to develop a single rupture dataset for more complete spatial coverage (Table 3.4). When alternative rupture datasets could not be combined, generally due to different mapping scales in overlapping areas, both rupture datasets are included in the database (e.g., 2010 El Mayor-Cucapah, Mexico; 2010 Darfield, New Zealand; and both 2019 Ridgecrest, California earthquakes). We provide recommendations to model development teams on alternative rupture datasets in Chapter 4.6.1. The information on all datasets used, based on dataset identification number (DS_ID) and citation, is included in the flatfile documentation in Appendix A.

Table 3.3. Measurement and surface rupture data sources included in the FDHI Database.

EQ_ID	Name	Measurements: [DS_ID][1] Source	Ruptures: [DS_ID][1] Source
1	Landers	[3] Milliner et al. (2016)	[6] Petersen et al. (2011)
		[6] Petersen et al. (2011)	
2	HectorMine	[2] Field et al. (2013)	[6] Petersen et al. (2011)
		[3] Milliner et al. (2016)	
		[6] Petersen et al. (2011)	
		[99] Chen et al. (2015)	
3	EMC	[18] Fletcher et al. (2014)	[17] Teran et al. (2015)
			[18] Fletcher et al. (2014)
4	Balochistan	[23] Gold et al. (2015)	[23] Gold et al. (2015)
		[75] Zinke et al. (2014)	
5	Izmit_Kocaeli	[6] Petersen et al. (2011)	[6] Petersen et al. (2011)
		[144] Rockwell et al. (2002)	
6	Borrego	[6] Petersen et al. (2011)	[6] Petersen et al. (2011)
7	Imperial1979	[6] Petersen et al. (2011)	[6] Petersen et al. (2011)
8	SuperstitionHills	[100] Sharp et al. (1989)	[100] Sharp et al. (1989)
9	Kobe	[6] Petersen et al. (2011)	[6] Petersen et al. (2011)
		[86] Baize et al. (2020)	
10	Denali	[39] Haeussler et al. (2004)	[24] Haeussler (2009)
		[40] Crone et al. (2004)	
		[90] Schwartz et al. (2012)	
11	Duzce	[37] Pucci et al. (2006)	[160] FDHI Manual Compilation[2] based on: Akyuz et al. (2002); pers. comm., Dawson, T.; and Duman et al. (2005)
		[38] Hartleb et al. (2002)	
		[43] Akyuz et al. (2002)	
		[144] Rockwell et al. (2002)	
12	Wenchuan	[44] Liu-Zeng et al. (2009)	[47] pers. comm., Liu-Zeng, J.
		[45] Liu-Zeng et al. (2010)	
		[46] Liu-Zeng et al. (2012)	
		[50] Xu et al. (2009)	
		[51] Tan et al. (2012)	
		[158] Nurminen et al. (2020)	
13	Napa	[56] Ponti et al. (2019)	[56] Ponti et al. (2019)
14	Yushu	[57] Li et al. (2012)	[57] Li et al. (2012)
		[58] Guo et al. (2012)	
15	Hualien	[61] Kuo et al. (2018)	[62] Huang et al. (2019)
		[62] Huang et al. (2019)	
16	ChiChi	[20] pers. comm., Kuo, Y.-T. and Yu, W.	[142] FDHI Manual Compilation[2] based on: Baize et al. (2020) & pers. comm., Kuo, Y.-T. and Yu, W.
		[65] Huang (1999)	
		[66] Lee et al. (2003)	
		[158] Nurminen et al. (2020)	
17	Kumamoto	[67] Shirahama et al. (2016)	[156] FDHI Manual Compilation[2] based on: Shirahama et al. (2016) and Goto et al. (2017)

EQ_ID	Name	Measurements: [DS_ID][1] Source	Ruptures: [DS_ID][1] Source
18	Nagano	[68] Okada et al. (2015)	[70] Ishimura et al. (2019)
		[69] Katsube et al. (2017)	
		[70] Ishimura et al. (2019)	
19	Kashmir	[71] Kaneda et al. (2008)	[71] Kaneda et al. (2008)
		[158] Nurminen et al. (2020)	
20	Kaikoura	[73] Zinke et al. (2019)	[107] FDHI Manual Compilation[2] based on: GNS Science and Zinke et al. (2019)
		[32] Kearse et al. (2018)	
		[33] Langridge et al. (2018)	
		[34] Williams et al. (2018)	
		[106] Howell et al. (2020)	
21	Darfield	[77] Litchfield et al. (2014)	[80] Villamor et al. (2012) & [103] Langridge et al. (2016)
		[78] Quigley et al. (2012)	
		[79] Elliott et al. (2012)	
22	Parkfield2004	[83] Rymer et al. (2006)	[83] Rymer et al. (2006)
23	Norcia3	[87] pers. comm., Boncio, P., based on: Brozzetti et al. (2019) and Villani et al. (2018)	[87] pers. comm., Boncio, P., based on: Civico et al. (2018); Brozzetti et al. (2019); and unpublished work
24	Hebgen	[84] Johnson et al. (2018)	[157] FDHI Manual Compilation[2] based on: Johnson et al. (2018) and USGS (1964)
		[86] Baize et al. (2020)	
25	SanFernando	[86] Baize et al. (2020)	[167] FDHI Manual Compilation[2] based on: California Geological Survey (2019) and USGS (1971)
		[2] Field et al. (2013)	
26	Bohol	[91] Rimando et al. (2019)	[91] Rimando et al. (2019)
27	Acambay	[93] Urbina and Camacho (1913)	[94] Langridge et al. (2000)
28	Imperial1940	[96] Rockwell and Klinger (2013)	[104] FDHI Manual Compilation[2] based on: California Geological Survey (2019); Rockwell and Klinger (2013); and Trifunac and Brune (1970)
		[162] pers. comm., Dawson, T.	
29	Parkfield1966	[108] Brown and Vedder (1966)	[108] Brown and Vedder (1966)
30	FairviewPeak	[98] Caskey et al. (1996)	[98] Caskey et al. (1996)
31	DixieValley	[98] Caskey et al. (1996)	[98] Caskey et al. (1996)
32	GalwayLake	[97] Hill and Beeby (1977)	[97] Hill and Beeby (1977)
33	Sonora	[110] Suter (2015)	[113] pers. comm., Suter, M.
		[111] Suter (2008a)	
		[112] Suter (2008b)	
34	PleasantValley	[117] Wallace et al. (1984)	[117] Wallace et al. (1984)
35	Kern	[122] Buwalda and St. Amand (1955)	[116] pers. comm., Thompson, S.
36	ChalfantValley	[125] Kahle et al. (1986)	[127] FDHI Manual Compilation[2] based on: Lienkaemper et al. (1987) and dePolo and Ramelli (1987)
		[126] Lienkaemper et al. (126)	

EQ_ID	Name	Measurements: [DS_ID][1] Source	Ruptures: [DS_ID][1] Source
37	Zirkuh	[124] Berberian et al. (1999)	[123] Francesca (2020)
38	Petermann	[120] Gold et al. (2019)	[120] Gold et al. (2019)
39	OwensValley	[129] Beanland and Clark (1994)	[129] Beanland and Clark (1994)
		[128] Haddon et al. (2016)	
		[2] Field et al. (2013)	
40	LagunaSalada	[130] Rockwell et al. (2015)	[130] Rockwell et al. (2015)
41	Iwaki2011	[131] Toda and Tsutsumi (2013)	[141] FDHI Manual Compilation[2] based on: Toda and Tsutsumi (2013) and Mizoguchi et al. (2012)
		[140] Mizoguchi et al. (2012)	
42	Ridgecrest1	[132] DuRoss et al. (2020)	[132] DuRoss et al. (2020)
			[145] Ponti et al. (2020)
43	Ridgecrest2	[132] DuRoss et al. (2020)	[132] DuRoss et al. (2020)
			[145] Ponti et al. (2020)
44	ElAsnam	[134] Philip and Meghraoui (1983)	[134] Philip and Meghraoui (1983)
		[135] Yielding et al. (1981)	
45	Cadoux	[136] King et al. (2019)	[136] King et al. (2019)
46	Calingiri	[136] King et al. (2019)	[136] King et al. (2019)
47	MarryatCreek	[136] King et al. (2019)	[136] King et al. (2019)
48	Meckering	[136] King et al. (2019)	[136] King et al. (2019)
49	Pukatja	[136] King et al. (2019)	[136] King et al. (2019)
50	TennantCreek1	[136] King et al. (2019)	[136] King et al. (2019)
51	TennantCreek2	[136] King et al. (2019)	[136] King et al. (2019)
52	TennantCreek3	[136] King et al. (2019)	[136] King et al. (2019)
53	SanMiguel	[139] Harvey (1985)	[139] Harvey (1985)
54	Yutian	[147] Li et al. (2016)	[146] pers. comm., Liu-Zeng, J.
		[146] pers. comm., Liu-Zeng, J.	
55	Luzon	[148] Nakata et al. (1996)	[148] Nakata et al. (1996)
56	BorahPeak	[119] Crone et al. (1987)	[150] FDHI Manual Compilation[2] based on: Crone et al. (1987) and Vincent (1995)
		[149] Vincent (1995)	
		[151] DuRoss et al. (2019)	
57	ElmoreRanch	[100] Sharp et al. (1989)	[100] Sharp et al. (1989)
58	Pisayambo	[86] Baize et al. (2020)	[86] Baize et al. (2020)
59	Rikuu	[86] Baize et al. (2020)	[86] Baize et al. (2020)
60	Mikawa	[86] Baize et al. (2020)	[86] Baize et al. (2020)
61	IzuPeninsula	[86] Baize et al. (2020)	[86] Baize et al. (2020)
62	IzuOshima	[86] Baize et al. (2020)	[86] Baize et al. (2020)
63	IwateInland	[86] Baize et al. (2020)	[86] Baize et al. (2020)
64	Edgecumbe	[86] Baize et al. (2020)	[86] Baize et al. (2020)
65	Neftegorsk	[154] pers. comm., Pinegina, T., Kozhurin, A., & Arcos, B.	[154] pers. comm., Pinegina, T., Kozhurin, A., & Arcos, B.
66	ChonKemin	[152] Arrowsmith et al. (2017)	[152] Arrowsmith et al. (2017)
67	Kunlun_Kokoxili	[52] Xu et al. (2002)	[165] Fu et al. (2005)
		[92] Klinger et al. (2005)	
68	LeTeil	[87] pers. comm., Baize, S.	[87] pers. comm., Baize, S.
69	Norcia1	[87] pers. comm., Boncio, P.	[87] pers. comm., Boncio, P.

EQ_ID	Name	Measurements: [DS_ID][1] Source	Ruptures: [DS_ID][1] Source
70	HomesteadValley	[168] Hill et al. (1980)	[168] Hill et al. (1980)
71	Palu	[169] Wu et al. (2021)	[171] Natawidjaja et al. (2021)
		[170] Jaya et al. (2019)	
		[171] Natawidjaja et al. (2021)	
72	LAquila	[87] pers. comm., Boncio, P.	[87] pers. comm., Boncio, P.
73	Spitak	[173] Nurminen et al. (2020)	[173] Nurminen et al. (2020)
74	Killari	[173] Nurminen et al. (2020)	[173] Nurminen et al. (2020)
		[174] Rajendran et al. (1996)	
75	YeniceGonen	[175] Kürçer et al. (2019)	[175] Kürçer et al. (2019)

[1] See Appendix A, Chapter 5 for full citations for each "DS_ID"

[2] See Table 3.4

Table 3.4. Events with surface rupture maps manually combined from multiple datasets.

EQ_ID	Name	Sources
11	Duzce	Akyuz et al. (2002); pers. comm., Akyuz, S. to Sarmiento, A., dated 28 Dec. 2018; pers. comm., Dawson, T. to Sarmiento, A., dated 18 Jul. 2018; Duman et al. (2005)
16	ChiChi	Baize et al. (2019); pers. comm., Kuo, Y.-T. & Yu, W. to Dawson, T., dated 29 Aug. 2018
17	Kumamoto	Shirahama et al. (2016); Baize et al. (2020); Goto et al. (2017)
20	Kaikoura	GNS Science (2018); Zinke et al. (2019)
24	Hebgen	Johnson et al. (2018); USGS (1964)
25	SanFernando	California Geological Survey (2019); USGS (1971)
28	Imperial1940	California Geological Survey (2019); Rockwell and Klinger (2013); Trifunac and Brune (1970)
36	ChalfantValley	Lienkaemper et al. (1987); dePolo et al. (1987)
41	Iwaki2011	Toda and Tsutsumi (2013); Mizoguchi et al. (2012)
56	BorahPeak	Crone et al. (1987); Vincent (1995)

Measurement and rupture data are provided in the professional literature in multiple formats. Measurement information is typically reported in tables that are embedded in the publication or attached as electronic supplements. The electronic supplements are usually data tabulated in *.csv format (or similar) or encoded for direct use in Geographic Information System (GIS) software as ESRI shapefiles or XML files (e.g., *.kml). Rupture linework is usually provided in GIS format (i.e., shapefiles or *.kml files). In some cases, measurement and rupture data were only provided as maps on figures or plates in the publication. For these datasets, we carefully georeferenced the maps using GIS software (Chapter 3.5), based on the map scale and projection, against appropriate base maps provided by ESRI (typically topographic maps). We then digitized the measurement locations or rupture linework and manually entered the relevant information in the shapefile attribute table.

Basic data cleaning and processing was performed on all original data to generate uniformly formatted ESRI shapefiles. We used the ESRI shapefile format because it was convenient and reliable for performing first-order geospatial analyses and geologic interpretation (Chapter 4) and standardizing data elements for importing into the relational database (Chapter 5). The shapefile format also allowed us to plot the data on base maps and visually inspect the data for quality assessments.

Processing of each measurement and rupture dataset ensured consistent formatting of reported data and metadata. For measurement data, the processing generally included the following: organizing reported measurement components (e.g., Figure 2.10) and uncertainties or ranges; identifying missing or duplicate data entries; adding measurement metadata; and adding event, dataset, site, and measurement identifier information (EQ_ID, DS_ID, PT_ID, and MEAS_ID, respectively). We used Microsoft Excel and Python ("pandas" library dataframes) to

format measurement datasets and cull duplicated data. The processing for surface rupture datasets generally included: adding rupture metadata; adding event, dataset, and rupture identifier information (EQ_ID, DS_ID, and RUP_ID, respectively); and performing topological checks in GIS to cull duplicated linework. Processed measurement and rupture datasets were converted into ESRI shapefiles, taking care to properly project the data based on the original coordinate system.

Our data quality review considered three data quality metrics: completeness, accuracy, and consistency. Completeness refers to the spatial extent of the rupture mapping and displacement measurement data in the FDHI Database, relative to the known spatial extent of the surface rupture. We reviewed multiple data sources for each earthquake to develop the most complete dataset possible for each event. Accuracy relates to the reliability of the original data and includes both the spatial accuracy of the measurement location and the accuracy of the reported measurement. We also carefully reviewed the data for each earthquake for internal consistency within an individual data source and between multiple data sources.

The results of our data quality review are included in the database and were used to develop recommendations for the model development teams (Chapter 4.6). The database contains text descriptions of geographic areas that are incomplete, where applicable, to document our assessment of completeness. Co-located or alternative measurements within the same data source and/or from different sources (for the same event) are explicitly identified in the database as part of our consistency evaluation. Finally, through our accuracy assessment, we identified some individual measurements that might be incomplete and/or erroneous (from data sources that are otherwise reliable); these are still included in the database for completeness, but they are explicitly flagged for potential accuracy issues. We developed a quality code system to methodically track our assessments of measurement accuracy and consistency, as described in Chapter 4.6.2.

3.3.3 Geologic Data and Metadata

Geologic datasets were developed from published digital geologic maps to allow the model development teams to investigate geologic controls on fault displacements. These maps were typically regional-scale and published by state or national geoscience authorities in ESRI shapefile format (i.e., georeferenced). Minimal data cleaning and processing was required for the geologic datasets. We retained the unit lithologic and age descriptions as reported, and we added a generalized geology category consisting of bedrock, young alluvium (Holocene), old alluvium, and undifferentiated alluvium. Digital geologic data were available for most events in the database; however, we were unable to acquire information for earthquakes in Africa and some parts of Asia. The geospatial analyses that relate measurement and rupture data to geologic data are described in Chapter 4.1.

3.4 EXCLUDED DATA

Our standard workflow for developing surface rupture and measurement information included a data quality evaluation through which each candidate dataset was carefully reviewed for completeness, accuracy, consistency, and compliance with the event and dataset selection criteria. In this process, we identified and intentionally excluded published datasets with known quality issues and subsets of existing compilations that did not meet the project's quality standards.

We evaluated the existing compilations of fault displacement and surface rupture data in Table 3.1 in detail. Most of the event datasets that met the initial selection screening criteria in Table 3.1 were included in the FDHI Database; however, some of events did not meet the dataset selection criteria or data quality standards (Table 3.5). A subset of events reported in the existing compilations were given lower priority due to difficulty accessing original data or relatively limited number of observations and were not resolved prior to the database entry cut-off date. This subset included the following four earthquakes: 1988 Spitak, Armenia; 1944 Gerede-Bolu, Turkey; 1976 Motagua, Guatamala; 1954 Rainbow Mountain, Nevada.

Table 3.5. Events in existing compilations not included in FDHI Database due to selection criteria or data quality.

Event	Compiler	Comments
1869 Olinghouse, Nevada	Pezzopane & Dawson (1996)	Incomplete rupture mapping, no slip measurements
1903 Wonder, Nevada	Pezzopane & Dawson (1996)	No slip measurements
1932 Cedar Mountain, Nevada	Pezzopane & Dawson (1996)	Incomplete rupture mapping
1934 Excelsior Mountains, Nevada	Pezzopane & Dawson (1996)	No slip measurements
1934 Hansel Valley, Utah	Pezzopane & Dawson (1996)	No slip measurements
1950 Fort Sage, California	Pezzopane & Dawson (1996)	No slip measurements
1954 Stillwater, Nevada	Pezzopane & Dawson (1996)	Incomplete rupture mapping, few slip measurements
1980 Mammoth Lakes, California	Pezzopane & Dawson (1996)	Multi-event earthquake sequence
1993 Eureka Valley, California	Pezzopane & Dawson (1996)	Few slip measurements
1857 Fort Tejon, California	Petersen et al. (2011)	Incomplete rupture mapping
1930 Kita-Izu, Japan	Petersen et al. (2011)	No/incomplete geospatial control
1939 Erzincan, Turkey	Petersen et al. (2011)	No/incomplete geospatial control
1942 Irba-Niksar, Turkey	Petersen et al. (2011)	No/incomplete geospatial control
1943 Tosya, Turkey	Petersen et al. (2011)	No/incomplete geospatial control
1967 Mudurnu, Turkey	Petersen et al. (2011)	No/incomplete geospatial control
1981 Sirch, Iran	Petersen et al. (2011)	No/incomplete geospatial control
1998 Fandoqa, Iran	Petersen et al. (2011)	No/incomplete geospatial control
1938 Kussharo, Japan	Baize et al. (2020)	No rupture mapping

Event	Compiler	Comments
1891 Nobi, Japan	Baize et al. (2020)	Data quality concerns: duplicate sites (coordinates) with conflicting measurements and several sites have no data (meaning is ambiguous)
1927 North Tango, Japan	Baize et al. (2020)	Data quality concerns: duplicate sites (coordinates) with conflicting measurements and several sites have no data (meaning is ambiguous)
1930 North Izu, Japan	Baize et al. (2020)	Data quality concerns: duplicate sites (coordinates) with conflicting measurements and several sites have no data (meaning is ambiguous)
1943 Tottori, Japan	Baize et al. (2020)	Data quality concerns: duplicate sites (coordinates) with conflicting measurements and several sites have no data (meaning is ambiguous)
2000 Tottori Pref. West., Japan	Baize et al. (2020)	Incomplete rupture mapping, data quality concerns: duplicate sites (coordinates) with conflicting measurements and several sites have no data (meaning is ambiguous)
1944 La Laja, Argentina	Baize et al. (2020)	Few slip measurements
1959 Deshibori, Japan	Baize et al. (2020)	Few slip measurements
1918 Omachi, Japan	Baize et al. (2020)	Few slip measurements
1939 Oga, Japan	Baize et al. (2020)	Few slip measurements
2004 Niigata Pref. Chuetsu, Japan	Baize et al. (2020)	Few slip measurements
1984 Nagano Prefecture West., Japan	Baize et al. (2020)	No rupture mapping
2008 Iwate-Miyagi Inland, Japan	Baize et al. (2020)	Data quality concerns: rupture mapping and slip measurements contaminated with non-tectonic deformation (landslides)
1983 Coalinga/Nunez, California	Boncio et al. (2018)	Multi-event earthquake sequence
2016 Kumamoto, Japan	Lin et al. (2016)	Data quality concerns: publication was retracted; see Stein (2019) and Lin et al. (2019)
2014 Nagano, Japan	Lin et al. (2015)	Data quality concerns: appears to mix landslide and fault scarps; see Ishimura et al. (2019)
2008 Wenchuan, China	Lin et al. (2009)	Data quality concerns: slip measurements appear to be erroneous; see Feng et al. (2017)

Event	Compiler	Comments
2001 Kunlun/Kokoxili, Tibet	Lin et al. (2002)	Data quality concerns: slip measurements may be multi-event; see Xu et al. (2002)

3.5 SOFTWARE

The following software was used to manage the collected datasets:

- ESRI ArcMap and ArcGIS Desktop software version 10.7, Advanced license.
- Global Mapper version 19.

3.6 REFERENCES

Akyuz, H. S., Hartleb, R., Barka, A., Altunel, E., Sunal, G., Meyer, B. and Armijo, V. R. (2002). Surface rupture and slip distribution of the 12 November 1999 Duzce earthquake (M 7.1), North Anatolian fault, Bolu, Turkey, *Bulletin of the Seismological Society of America*, **92**(1), 61-66.

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. and Donahue, J. L. (2013). PEER NGA-West2 Database, PEER Report 2013/03, *Pacific Earthquake Engineering Research Center*, University of California, Berkeley, CA.

Baize, S., Nurminen, F., Sarmiento, A., Dawson, T., Takao, M., Scotti, O., Azuma, T., Boncio, P., Champenois, J., Cinti, F. R., Civico, R., Costa, C., Guerrieri, L., Marti, E., McCalpin, J., Okumura, K., and Villamor, P. (2020). A worldwide and Unified Database of Surface Ruptures (SURE) for fault displacement hazard analyses, *Seismological Research Letters*, **91**(1), 499-520.

Berberian, M., Jackson, J. A., Qorashi, M., Khatib, M. M., Priestley, K., Talebian, M., and Ghafuri-Ashtiani, M. (1999). The 1997 May 10 Zirkuh (Qa'enat) earthquake (Mw 7.2): faulting along the Sistan suture zone of eastern Iran, *Geophysical Journal International*, **136**(3), 671-694.

Boncio, P., Liberi, F., Caldarella, M., and Nurminen, F. C. (2018). Width of surface rupture zone for thrust earthquakes: implications for earthquake fault zoning, *Natural Hazards and Earth System Sciences*, **18**(1), 241-256.

Brozzetti, F., Boncio, P., Cirillo, D., Ferrarini, F., Nardis, R., Testa, A., Liberi, F., and Lavecchia, G. (2019). High-resolution field mapping and analysis of the August–October 2016

coseismic surface faulting (central Italy earthquakes): Slip distribution, parameterization, and comparison with global earthquakes, *Tectonics*, **38**, 417–439.

California Geological Survey (2019). Alquist-Priolo (AP) Earthquake Fault Zones (EFZ) Geographic Information System (GIS) files, accessed December 2019.

Champenois, J., Baize, S., Vallée, M., Jomard, H., Alvarado, A., Espin, P., Ekström, G., and Audin, L. (2017). Evidences of surface rupture associated with a low-magnitude (M w 5.0) shallow earthquake in the Ecuadorian Andes, *Journal of Geophysical Research: Solid Earth*, **122**(10), 8446-8458.

Civico, R., Pucci, S., Villani, F., Pizzimenti, L., De Martini, P. M., Nappi, R., and Open EMERGEO Working group (2018). Surface ruptures following the 30 October 2016 Mw 6.5 Norcia earthquake, Central Italy, *Journal of Maps*, **14**(2), 151–160.

Crone, A. J., Machette, M. N., Bonilla, M. G., Lienkaemper, J. J., Pierce, K. L., Scott, W. E., and Bucknam, R. C. (1987). Surface faulting accompanying the Borah Peak earthquake and segmentation of the Lost River fault, central Idaho, *Bulletin of the Seismological Society of America*, **77**(3), 739-770.

Delouis, B., Oral, E., Menager, M., Ampuero, J. P., Trilla, A. G., Régnier, M., and Deschamps, A. (2021). Constraining the point source parameters of the 11 November 2019 Mw 4.9 Le Teil earthquake using multiple relocation approaches, first motion and full waveform inversions, *Comptes Rendus. Géoscience*, **353**(S1), 1-24.

dePolo, C. M. and Ramelli, A. R. (1987). Preliminary report on surface fractures along the White Mountains fault zone associated with the July 1986 Chalfant Valley earthquake sequence, *Bulletin of the Seismological Society of America*, **77**(1), 290-296.

Doser, D. I. (1992). Faulting processes of the 1956 San Miguel, Baja California, earthquake sequence, *Pure and Applied Geophysics*, **139**(1), 3-16.

Duman, T. Y., Emre, O., Dogan, A. and Ozalp, S. (2005). Step-over and bend structures along the 1999 Duzce earthquake surface rupture, North Anatolian fault, Turkey, *Bulletin of the Seismological Society of America*, **95**(4), 1250-1262.

Feng, G., Jónsson, S., and Klinger, Y. (2017). Which fault segments ruptured in the 2008 Wenchuan earthquake and which did not? New evidence from near-fault 3D surface displacements derived from SAR image offsets *Bulletin of the Seismological Society of America*, **107**(3), 1185-1200.

GNS Science (2018). "2016 Kaikoura Rupture" GIS download, accessed 9 October 2018.

Gold, R. D., Clark, D., Barnhart, W. D., King, T., Quigley, M., and Briggs, R. W. (2019). Surface rupture and distributed deformation revealed by optical satellite imagery: The intraplate 2016 Mw 6.0 Petermann Ranges earthquake, Australia, *Geophysical Research Letters*, **46**(17-18), 10394-10403.

Gold, R. D., Reitman, N. G., Briggs, R. W., Barnhart, W. D., Hayes, G. P., and Wilson, E. (2015). On-and off-fault deformation associated with the September 2013 Mw 7.7 Balochistan earthquake: Implications for geologic slip rate measurements, *Tectonophysics*, **660**, 65-78.

Goto, H., Tsutsumi, H., Toda, S., and Kumahara, Y. (2017). Geomorphic features of surface ruptures associated with the 2016 Kumamoto earthquake in and around the downtown of Kumamoto City, and implications on triggered slip along active faults, *Earth, Planets and Space*, **69**(1), 26.

Hamdache, M., Peláez, J. A., Talbi, A., and Casado, C. L. (2010). A unified catalog of main earthquakes for Northern Algeria from AD 856 to 2008, *Seismological Research Letters*, **81**(5), 732-739.

Hanks, T. C. and Kanamori, H. (1979). A moment magnitude scale, *Journal of Geophysical Research: Solid Earth*, **84**(B5), 2348-2350.

Ishimura, D., Toda, S., Mukoyama, S., Homma, S. I., Yamaguchi, K., and Takahashi, N. (2019). 3D Surface Displacement and Surface Ruptures Associated with the 2014 Mw 6.2 Nagano Earthquake Using Differential Lidar3D Surface Displacement and Surface Ruptures, *Bulletin of the Seismological Society of America*, **109**(2), 780-796.

Johnson, K. L., Nissen, E., and Lajoie, L. (2018). Surface Rupture Morphology and Vertical Slip Distribution of the 1959 Mw 7.2 Hebgen Lake (Montana) Earthquake From Airborne Lidar Topography, *Journal of Geophysical Research: Solid Earth*, **123**(9), 8229-8248.

Kanamori, H. and Fuis, G. (1976). Variation of P-wave velocity before and after the Galway Lake earthquake (ML= 5.2) and the Goat Mountain earthquakes (ML= 4.7, 4.7), 1975, in the Mojave desert, California, *Bulletin of the Seismological Society of America*, **66**(6), 2017-2037.

Kaneda, H., Nakata, T., Tsutsumi, H., Kondo, H., Sugito, N., Awata, Y., Akhtar, S. S., Majid, A., Khattak, W., Awan, A. A., Yeats, R. S., Hussain, A., Ashraf, M., Wesnousky, S. G., and Kausar, A. B. (2008). Surface rupture of the 2005 Kashmir, Pakistan, earthquake and its active tectonic implications, *Bulletin of the Seismological Society of America*, **98**(2), 521-557.

King, T. R., Quigley, M., and Clark, D. (2019). Surface-rupturing historical earthquakes in Australia and their environmental effects: new insights from re-analyses of observational data, *Geosciences*, **9**(10), 408.

Kulikova, G. and Krüger, F. (2015). Source process of the 1911 M 8.0 Chon-Kemin earthquake: Investigation results by analogue seismic records, *Geophysical Journal International*, **201**(3), 1891-1911.

Kürçer, A., Özalp, S., Özdemir, E., Uygun Güldoğan, Ç., and Duman, T. Y. (2019). Active tectonic and paleoseismologic characteristics of the Yenice-Gönen fault, NW Turkey, in light of

the 18 March 1953 Yenice-Gönen Earthquake (Ms=7.2), *Bulletin of the Mineral Research and Exploration*, **159**(159), 29-62.

Kuo, Y. T., Wang, Y., Hollingsworth, J., Huang, S.-Y., Chuang, R. Y., Lu, C.-H., Hsu, Y.-C., Tung, H., Yen, J.-Y., and Chang, C.-P. (2018). Shallow fault rupture of the Milun fault in the 2018 M w 6.4 Hualien earthquake: A high-resolution approach from optical correlation of Pléiades satellite imagery, *Seismological Research Letters*, **90**(1), 97-107.

Langridge, R. M., Weldon, R. J., Moya, J. C., and Suárez, G. (2000). Paleoseismology of the 1912 Acambay earthquake and the Acambay-Tixmadejé fault, Trans-Mexican volcanic belt, *Journal of Geophysical Research: Solid Earth*, **105**(B2), 3019-3037.

Li, H., Pan, J., Lin, A., Sun, Z., Liu, D., Zhang, J., Li, C., Liu, K., Chevalier, M.-L., Yun, K., and Gong, Z. (2016). Coseismic Surface Ruptures Associated with the 2014 Mw 6.9 Yutian Earthquake on the Altyn Tagh Fault, Tibetan Plateau, *Bulletin of the Seismological Society of America*, **106**(2), 595-608.

Lienkaemper, J. J., Pezzopane, S. K., Clark, M. M., and Rymer, M. J. (1987). Fault fractures formed in association with the 1986 Chalfant Valley, California, earthquake sequence: preliminary report, *Bulletin of the Seismological Society of America*, **77**(1), 297-305.

Lin, A., Satsukawa, T., Wang, M., Mohammadi Asl, Z., and Fueta, R. (2019). Retraction, *Science*, **364**(6439), 444.

Lin, A., Satsukawa, T., Wang, M., Mohammadi Asl, Z., Fueta, R., and Nakajima, F. (2016). Coseismic rupturing stopped by Aso volcano during the 2016 Mw 7.1 Kumamoto earthquake, Japan, *Science*, **354**(6314), 869-874.

Lin, A., Sano, M., Yan, B., and Wang, M. (2015). Co-seismic surface ruptures produced by the 2014 Mw 6.2 Nagano earthquake, along the Itoigawa–Shizuoka tectonic line, central Japan, *Tectonophysics*, **656**, 142-153.

Lin, A., Ren, Z., Jia, D., and Wu, X. (2009). Co-seismic thrusting rupture and slip distribution produced by the 2008 Mw 7.9 Wenchuan earthquake, China, *Tectonophysics*, **471**(3-4), 203-215.

Lin, A., Fu, B., Guo, J., Zeng, Q., Dang, G., He, W., and Zhao, Y. (2002). Co-seismic strike-slip and rupture length produced by the 2001 Ms 8.1 Central Kunlun earthquake, *Science*, **296**(5575), 2015-2017.

Mizoguchi, K., Uehara, S. I., and Ueta, K. (2012). Surface Fault Ruptures and Slip Distributions of the Mw 6.6 11 April 2011 Hamadoori, Fukushima Prefecture, Northeast Japan, Earthquake, *Bulletin of the Seismological Society of America*, **102**(5), 1949-1956.

Nemati, M. (2015). Slip distribution, aftershocks and co-seismic fault of some large Persian earthquakes, *Environmental Earth Sciences*, **73**(11), 7165-7181.

Nurminen, F., Boncio, P., Visini, F., Pace, B., Valentini, A., Baize, S., and Scotti, O. (2020). Probability of occurrence and displacement regression of distributed surface rupturing for reverse earthquakes, *Frontiers in Earth Science*, **8**(581605).

Petersen, M. D., Dawson, T. E., Chen, R., Cao, T., Wills, C. J., Schwartz, D. P., and Frankel, A. D. (2011). Fault displacement hazard for strike-slip faults, *Bulletin of the Seismological Society of America*, **101**(2), 805-825.

Pezzopane, S. K. and Dawson, T. E. (1996). Fault displacement hazard: A summary of issues and information. In *Seismotectonic Framework and Characterization of Faulting at Yucca Mountain, Nevada*, U.S. Geological Survey Administrative Report prepared for the U.S. Department of Energy, Chapter 9, 160 pp.

Ponti, D. J., Rosa, C. M., and Blair, J. L. (2019). The Mw 6.0 South Napa earthquake of August 24, 2014 – Observations of surface faulting and ground deformation, with recommendations for improving post-earthquake field investigations (No. 2019-1018). US Geological Survey.

Ritz, J. F., Baize, S., Ferry, M., Larroque, C., Audin, L., Delouis, B., and Mathot, E. (2020). Surface rupture and shallow fault reactivation during the 2019 Mw 4.9 Le Teil earthquake, France, *Communications Earth & Environment*, **1**(1), 1-11.

Rockwell, T. K. and Klinger, Y. (2013). Surface rupture and slip distribution of the 1940 Imperial Valley earthquake, Imperial fault, southern California: Implications for rupture segmentation and dynamics, *Bulletin of the Seismological Society of America*, **103**(2A), 629-640.

Shirahama, Y., Yoshimi, M., Awata, Y., Maruyama, T., Azuma, T., Miyashita, Y., Mori, H., Imanishi, K., Takeda, N., Ochi, T., Otsubo, M., Asahina, D., and Miyakawa, A. (2016). Characteristics of the surface ruptures associated with the 2016 Kumamoto earthquake sequence, central Kyushu, *Japan, Earth, Planets and Space*, **68**(1), 191.

Stein, R. S. (2019), Scientific fraud announced in two studies of the 2016 M=7.0 Kumamoto, Japan, earthquake, *Temblor*, http://doi.org/10.32858/temblor.018.

Takao, M., Kaneto, T., and Kurita, T. (2018). Outline of the PFDHA Method and Recent Studies on PFDHA in Japan. In *Best Practices in Physics-based Fault Rupture Models for Seismic Hazard Assessment of Nuclear Installations: Issues and Challenges Towards Full Seismic Risk Analysis*, 14 pp.

Teran, O. J., Fletcher, J. M., Oskin, M. E., Rockwell, T. K., Hudnut, K. W., Spelz, R. M., Akciz, S. O., Hernandez-Flores, A. P., and Morelan, A. E. (2015). Geologic and structural controls on rupture zone fabric: A field-based study of the 2010 Mw 7.2 El Mayor–Cucapah earthquake surface rupture, *Geosphere*, **11**(3), 899-920.

Trifunac, M. D. and Brune, J. N. (1970). Complexity of energy release during the Imperial Valley, California, earthquake of 1940, *Bulletin of the Seismological Society of America*, **60**(1), 137-160.

Toda, S. and Tsutsumi, H. (2013). Simultaneous reactivation of two, subparallel, inland normal faults during the Mw 6.6 11 April 2011 Iwaki earthquake triggered by the M w 9.0 Tohoku-oki, Japan, earthquake, *Bulletin of the Seismological Society of America*, **103**(2B), 1584-1602.

United States Geological Survey (USGS) (2021). Magnitude Types. URL: https://www.usgs.gov/natural-hazards/earthquake-hazards/science/magnitude-types?qt-science_center_objects=0#qt-science_center_objects.

USGS (1971). The San Fernando, California, earthquake of February 9, 1971: a preliminary report published jointly by the U.S. Geological Survey and the National Oceanic and Atmospheric Administration, *U.S. Geological Survey Professional Paper 733*.

USGS (1964). The Hebgen Lake, Montana, earthquake of August 17, 1959, *U.S. Geological Survey Professional Paper 435*.

Villani, F., Civico, R., Pucci, S., Pizzimenti, L., Nappi, R., De Martini, P. M., and Open EMERGEO Working Group (2018). A database of the coseismic effects following the 30 October 2016 Norcia earthquake in central Italy, *Scientific Data*, **5**, 180049.

Vincent, K.R. (1995). Implications for models of fault behavior from earthquake surface displacement along adjacent segments of the Lost River fault, Idaho, PhD. Dissertation, *University of Arizona*, 152 p.

Wells, D. L. and Coppersmith, K. J. (1994). New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement, *Bulletin of the Seismological Society of America*, **84**(4), 974-1002.

Wesnousky, S. G. (2008). Displacement and geometrical characteristics of earthquake surface ruptures: Issues and implications for seismic-hazard analysis and the process of earthquake rupture, *Bulletin of the Seismological Society of America*, **98**(4), 1609-1632.

Xu, X., Chen, W., Ma, W., Yu, G., and Chen, G. (2002). Surface rupture of the Kunlunshan earthquake (Ms 8.1), northern Tibetan plateau, China, *Seismological Research Letters*, **73**(6), 884-892.

Yielding, G., Jackson, J. A., King, G. C. P., Sinvhal, H., Vita-Finzi, C., and Wood, R. M. (1981). Relations between surface deformation, fault geometry, seismicity, and rupture characteristics during the El Asnam (Algeria) earthquake of 10 October 1980, *Earth and Planetary Science Letters*, **56**, 287-304.

Zinke, R., Hollingsworth, J., Dolan, J. F., and Van Dissen, R. (2019). Three-dimensional surface deformation in the 2016 Mw 7.8 Kaikōura, New Zealand, earthquake from optical image

correlation: Implications for strain localization and long-term evolution of the Pacific-Australian plate boundary, *Geochemistry, Geophysics, Geosystems*, **20**(3), 1609-1628.

4 Data Analysis

Analysis and geologic interpretation of the measurement and rupture data assembled from the professional literature was performed to meet model development needs. This Chapter documents the purposes and procedures of the analyses and interpretations. In brief, we completed geospatial analysis to develop geologic information, elevation data and metrics, and the event-specific coordinate systems (ECS) for each earthquake in the database. We also performed geologic evaluations of each dataset to classify (rank) measurements and ruptures and develop recommended net slip values and usage in model development. Incompatible measurement datasets and technically defensible alternative measurements are also explicitly flagged in the database.

4.1 GEOLOGY

Geologic data were included, when available, for each measurement site to allow the model development teams to investigate geologic controls on fault displacements. The geologic data included lithologic unit descriptions, unit age, general geology description, and distance to the closest mapped bedrock outcrop. The general geology description is six simple categories: bedrock, young alluvium (Holocene), old alluvium, undifferentiated alluvium, water, and glacier. The broad categories for alluvium include a range of sediments, such as fan, fluvial, colluvium, glacial, fluvio-glacial, lacustrine, and marine deposits. Young and old alluvial deposits are separated, when possible, to capture relative degrees of consolidation. The undifferentiated alluvium category is used when the source dataset does not provide age control. The bedrock distance parameter is the distance to the closest surface outcropping of bedrock, not the depth to bedrock. This parameter was included as a proxy for sediment thickness or basin depth (Milliner et al., 2015). The same geologic data are also reported for each rupture line vertex.

The geologic information for each measurement and rupture line vertex was calculated using built-in geospatial analysis tools in ArcGIS (see Chapter 4.7 for software versioning). We used our uniformly processed geologic datasets (Chapter 3.3.3) and measurement/rupture datasets (Chapter 3.3.2), in ESRI shapefile format, as inputs. The input datasets were projected from geographic coordinates into a projected coordinate system (in linear units) appropriate for the event location. The ArcGIS Identity Analysis Tool was used to calculate the geologic unit (including

lithology, general geology, and unit age) of each measurement site or rupture line vertex. The ArcGIS Select Analysis and Near Analysis tools were used to calculate the distance between each measurement site (or rupture line vertex) and the closest mapped bedrock outcrop.

4.2 ELEVATION DATA AND METRICS

Elevation data and related metrics were computed for each measurement site to allow the model development teams to investigate topographic effects on fault displacements. We used the 1 arc-second (30 meter) resolution digital elevation model derived from the Shuttle Radar Topography Mission (STRM) data (Farr et al., 2007) for all events except Denali (EQ_ID = 10), which was located outside the SRTM data coverage. For Denali, we down-sampled a 5-meter resolution digital surface model derived from InSAR data (Alaska Geospatial Council, 2021) to 30-meter resolution. The data extraction, down-sampling, and analyses were performed in Python using the Geospatial Data Abstraction Library (GDAL) and Scientific Python library (Chapter 4.7).

The computed metrics quantify ground slope and surface irregularities or terrain texture (i.e., density of topographic peaks and troughs) in the vicinity of the measurement site. The following elevation data and metrics are included in the database:

- Elevation (meters)
- Ground slope (percent)
- Terrain class per Iwahashi et al. (2018): geomorphic terrain class based on ground slope and texture (Table 4.1)
- Prominence or relative elevation per Rai et al. (2017): difference between pixel elevation at site and mean elevation of all pixels in N-meter radius (where N = 125, 250, 500, and 1000)
- Terrain roughness: largest difference between pixel elevation at site and elevation of all adjacent pixels
- Topographic Position Index (TPI): difference between pixel elevation at site and mean elevation of all adjacent pixels
- Terrain Ruggedness Index (TRI) per Riley et al. (1999): total elevation change between pixel elevation at site and all adjacent pixel elevations

Table 4.1. Terrain classification code after Iwahashi et al. (2018).

Code	Geomorphic terrain description
1	steep mountain, rough
2	steep mountain, smooth
3	moderate mountain, rough
4	moderate mountain, smooth
5	hills, rough in small and large scales
6	hills, smooth in small scale, rough in large scale
7	upper large slope
8	middle large slope
9	dissected terrace, moderate plateau
10	slope in and around terrace or plateau
11	terrace, smooth plateau
12	alluvial fan, pediment, bajada, pediplain
13	alluvial plain, pediplain
14	alluvial or coastal plain, pediplain
15	alluvial or coastal plain (gentlest), lake plain, playa

4.3 RANK CLASSIFICATION

At the request of the model development teams, we interpreted the rupture linework and displacement measurements to distinguish principal and distributed faulting. Previous fault displacement models have treated principal and distributed faulting separately (Youngs et al., 2003; Petersen et al., 2011), and some model development teams are anticipated to continue this approach. The model developers recognized value in using expert geologic interpretation and judgment to distinguish principal and distributed ruptures and displacements. Including this information in the FDHI Database allows the model development teams to use the same interpretations of the data.

We follow Coppersmith and Youngs (2000) and Youngs et al. (2003) in defining principal surface ruptures as the primary faults or tectonic/seismogenic features responsible for the earthquake and distributed surface ruptures as the secondary faults, splays, fractures, or shears near the principal fault (Table 4.2). These criteria served as the basis for the Youngs et al. (2003) and Petersen et al. (2011) fault displacement models. An alternative ranking classification system by Baize et al. (2020) and Nurminen et al. (2020), which further subdivides principal and distributed ruptures, was not implemented for this project. In general, our principal surface ruptures typically correspond to Nurminen et al. (2020) Rank 1 and Rank 1.5, and our distributed ruptures correspond to their Ranks 2, 21, and 22.

Fault displacement measurements can also be categorized as principal or distributed, based on the classification of their associated surface rupture (Youngs et al., 2003; Petersen et al., 2011). To better distinguish measurements summed across multiple ruptures or measured across very

wide apertures, we introduce two additional measurement rank classifications, cumulative and total, to respectively differentiate these measurements. Specifically, we use the cumulative classification for slip measurements summed across either (1) multiple principal ruptures, or (2) principal rupture(s) and one or more distributed ruptures. Total measurements represent wide-aperture slips calculated from the differencing or image correlation methods discussed in Chapter 2.3.3. The rank descriptions are summarized in Table 4.2. For distributed measurements in reverse, normal, and oblique style earthquakes, we also indicate if the site is located on the hanging wall or footwall.

Table 4.2. Rank classifications used in the FDHI Database.

Feature	Rank	Description
Rupture Line	Principal	Primary fault or tectonic/seismogenic feature responsible for the earthquake
	Distributed	Secondary features near the principal fault, such as other faults, splays, fractures, or shears
Measurement	Total	Wide-aperture displacements calculated from differencing or image correlation methods
	Cumulative	Displacement summed across multiple adjacent principal ruptures; or displacement summed across principal rupture(s) and adjacent distributed rupture(s); or displacement summed across principal rupture(s) and zone of continuous deformation
	Principal	Displacement on principal rupture
	Distributed	Displacement on distributed rupture

Surface rupture and measurement data reported in the professional literature often do not explicitly identify principal/primary or distributed/secondary faulting. As a result, we developed a workflow to manually assign principal, distributed, cumulative, and total rankings to all data in the FDHI Database (Figure 4.1). While we considered event characteristics such as style of faulting, dataset quality and completeness (e.g., mapping scale, known limitations such as inaccessible areas), and the original authors' interpretations in developing the rankings, the workflow relies largely on iterative application of geologic judgment. Figure 4.2 shows an example application of the workflow for part of the 1968 **M** 6.63 Borrego Mountain, California (EQ_ID = 6) earthquake.

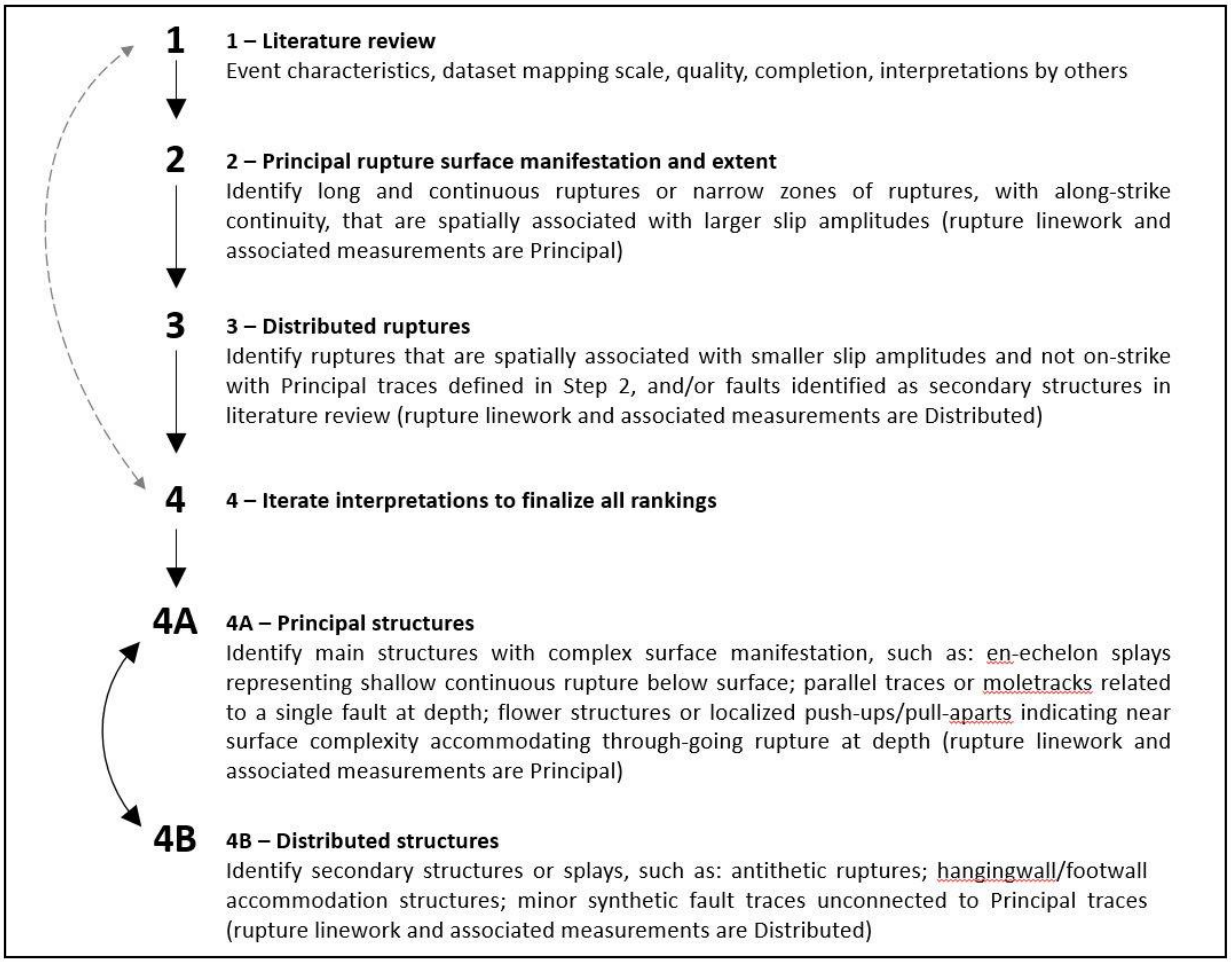

Figure 4.1. Flowchart for developing rank classifications based on geologic interpretation.

Principal surface rupture expression at the ground surface can vary significantly, as discussed in Chapter 2. Common patterns include the following: simple linear or curvilinear traces; segmented zones with en-echelon, anastomosing, or branching traces; moletrack zones; overlapping step-overs; flower or other slip-partitioning structures; and monoclinal warping or tilting. Examples of rank classifications for some of these patterns are shown in Figure 4.3 from the 1968 **M** 6.63 Borrego Mountain, California (EQ_ID = 6) and 1992 **M** 7.28 Landers, California (EQ_ID = 1) earthquakes. Although the classifications may be non-unique, they have been applied as consistently as possible across the contents of the database.

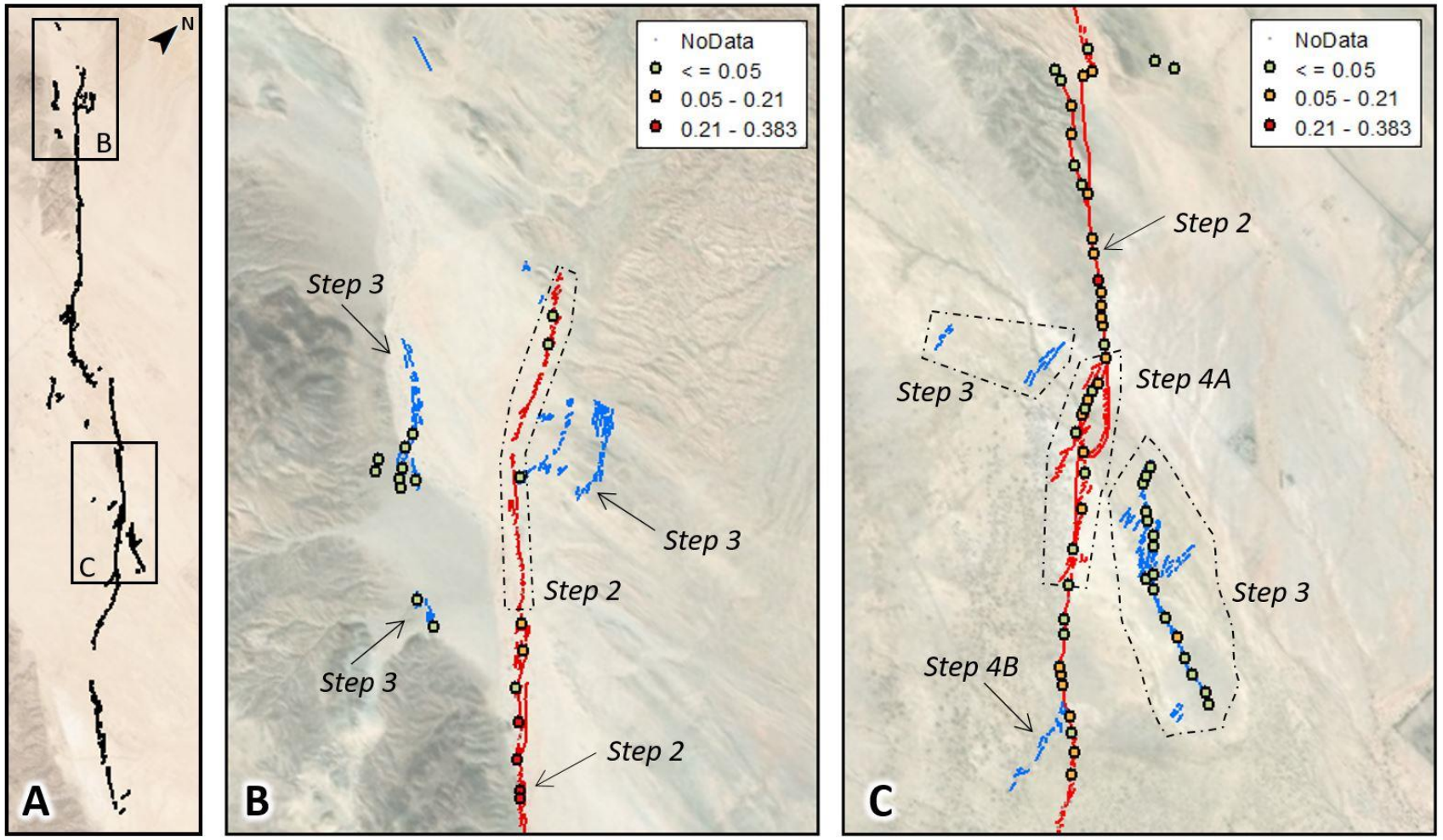

Figure 4.2. Example application of rank classification workflow applied to a portion of the 1968 **M** 6.63 Borrego Mountain, California earthquake; see Figure 4.1 for workflow steps and description. (A) Map of full surface rupture (black lines) at 1:350,000 scale. (B) and (C) Maps of rupture traces ranked in various steps as labeled; Red lines = principal ruptures; blue lines = distributed ruptures; filled circles are slip measurement sites, color-coding as shown in legend for recommended net displacement in meters. Arrows and black dashed polygons identify area considered in labeled workflow step. Map scale is 1:60,000.

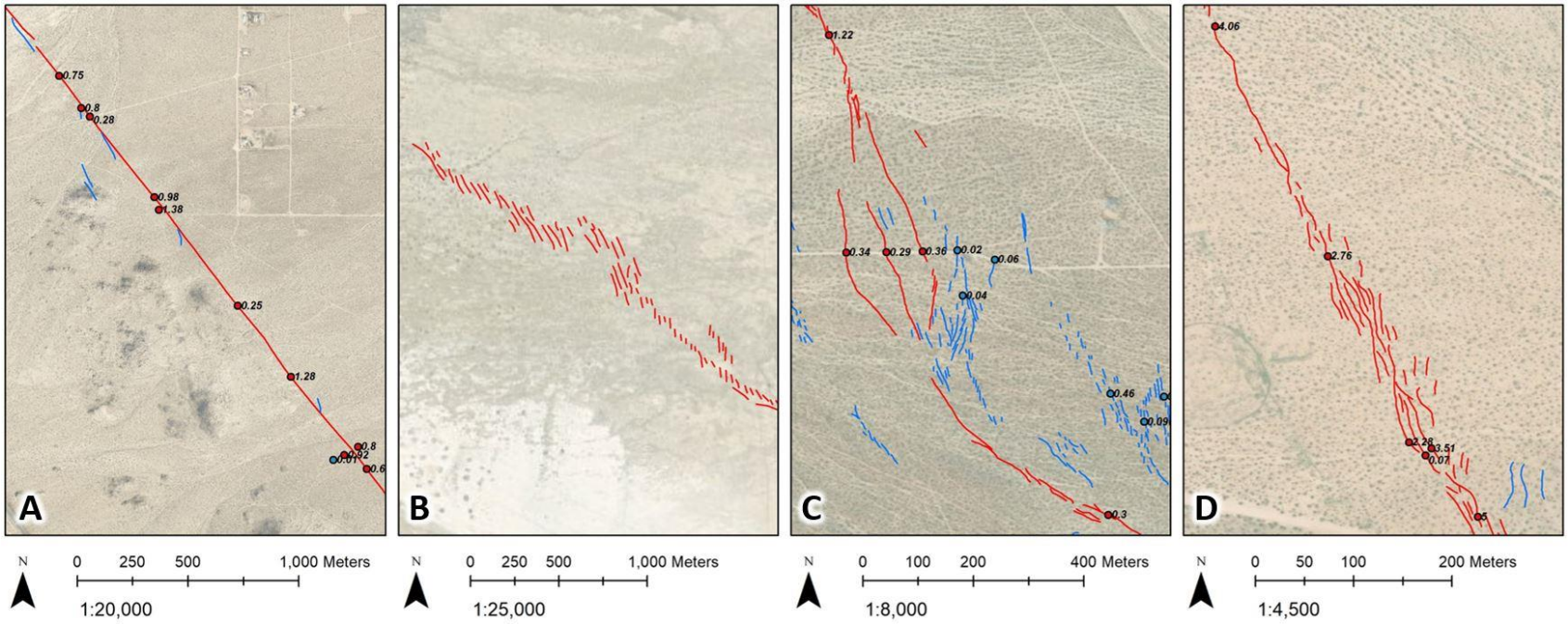

Figure 4.3. Example rank classifications for various surface rupture patterns. Red lines and circles are principal rank; blue lines and circles are distributed rank. Filled circles are slip measurement sites with recommended net displacement in meters. (A) Simple curvilinear principal fault trace from 1992 Landers, California **M** 7.28 earthquake. (B) Principal faulting as en-echelon overstepping array (R Riedel shears) from 1968 **M** 6.63 Borrego Mountain, California earthquake. (C) Tri-furcated/branching principal fault traces from Landers earthquake. (D) Anastomosing zone of principal faulting from Landers earthquake.

4.4 PAIRING MEASUREMENT SITES TO MAPPED RUPTURES

For end-user convenience, the closest mapped rupture to each measurement is explicitly identified in the database. Specifically, we report the rupture identifier ("RUP_ID") and distance to the rupture for each measurement, considering the classification of the measurement and rupture. The closest mapped principal rupture trace is reported for measurements classified as principal, cumulative, or total. For measurements classified as distributed, the closest mapped rupture, regardless of classification, is returned. The calculations were performed using built-in geospatial analysis tools in ArcGIS. The uniformly processed measurement and rupture datasets (Chapter 3.3.2), in ESRI shapefile format, were used as inputs, and the input datasets were projected from geographic coordinates into a projected coordinate system (in linear units) appropriate for the event location. The ArcGIS Select Analysis and Near Analysis tools were used to calculate the distance between each measurement site and the closest mapped rupture.

Measurement sites commonly are not perfectly co-located on a mapped rupture. In our experience compiling and analyzing the database, we found that spatial discrepancies were mainly related to the format in which the original data were provided in the professional literature. Data from older events were more likely to be documented on topographic maps with hand-drawn rupture linework and measurement sites, and the measurement sites and mapped ruptures are generally co-located. Conversely, in many modern datasets, measurement locations are recorded by handheld GPS devices, and rupture linework is collected on various media (e.g., printed maps, aerial photographs, digital base maps) at variable scales. Our experience is that dataset originators do not consistently check for spatial compatibility between the measurement site coordinates and mapped rupture linework. Furthermore, inconsistencies between measurement locations and mapped ruptures are also common when the rupture and measurement datasets were generated by different researchers (Chapter 3.3.2).

4.5 EVENT-SPECIFIC COORDINATE SYSTEM (ECS)

In the FDHI database, the locations of displacement points and rupture-line vertices are defined in terms of the latitude and longitude coordinates. However, in fault displacement hazard analyses, the along-strike and perpendicular-to-strike distance metrics are used to describe the location of fault displacements.

The objective of the event coordinate system (ECS) is to provide a unique value of the along-strike and perpendicular-to-strike distance metrics for every data point for the events in the database. A key challenge is that some of the ruptures have multiple parallel strands, complicating the selection of a single value for each distance metric. In the proposed approach, the along-strike and perpendicular-to-strike distance metrics are defined based on a reference axis for each event. This reference axis is not intended to match individual rupture strands, but instead provides a local

coordinate system for the entire rupture profile. For instance, in the case of multiple sub-parallel ruptures strands, the reference axis will pass through the middle of the ruptures.

The location of the reference axis is estimated based on the location and amplitude of slip at the displacement measurement sites and the location of rupture-line vertices. With the reference axis determined, the u and t local axes are defined with respect to it. The u axis corresponds to the along-strike distance as measured from one arbitrary end of the rupture, and the t axis corresponds to the perpendicular-to-strike distance as measured from the reference axis.

An iterative process is used to estimate the location of the reference axis. At the start of each iteration, the location of the reference axis is expressed as a function of u:

$$f_{ref} = \begin{cases} x_{ref}(u) \\ y_{ref}(u) \end{cases} \tag{4.1}$$

where, $x_{ref}(u)$ and $y_{ref}(u)$ are the UTM coordinates. The starting solution for the reference axis corresponds to the first component of a principal component analysis of the displacement points and rupture-line vertices. In the subsequent iterations, the location of the reference axis is updated by minimizing the objective function g:

$$g = \sum_{i=1}^{n} wt_i \left[\left(x_{pt,i} - x_{ref}(u) \right)^2 + \left(y_{pt,i} - y_{ref}(u) \right)^2 \right] + \lambda \int_{u=0}^{L} \frac{\partial^2 x_{ref}}{\partial u} + \frac{\partial^2 y_{ref}}{\partial u^2} du \tag{4.2}$$

where x_{pt} and y_{pt} are ordinates of the displacement measurement sites and rupture-line vertices projected into Universal Transverse Mercator (UTM) coordinates. The first part of Equation 4.2 measures the weighted distance between the reference line and displacement measurement sites and rupture-line vertices, and the second part of the equation measures the curvature of the reference line scaled by the penalty factor, λ. The distance to the displacement measurement sites is weighted by the mean value of recommended net displacement, while the weights for the distance to the rupture-line vertices are equal to 0.01. Both the displacement measurement sites and rupture-line vertices are used in the calculation of the reference axis because the rupture lines commonly extend beyond the displacement measurement sites. With this weighting scheme, the reference axis is guided by the displacement measurement sites in areas of the surface rupture that are mapped by both displacement measurement sites and rupture vertices, whereas the reference axis is guided by the rupture-line vertices in parts of the surface rupture that are mapped only by rupture lines. Based on experts' review of preliminary results, the penalty factor, λ, is set to 0.1. Equation 4.2 ensures that the reference axis will pass close to the displacement points with the largest displacement values, as those points are assumed to be part of the main rupture, but also that the reference line remain smooth. The iterative procedure is terminated once the maximum

distance between the current and the previous reference axis is less than 10 m. Once the reference axis is determined, the second version of generalized coordinate system (GC2; Spudich and Chiou, 2015) is applied to calculate the u and t coordinates of all the displacement measurement sites and rupture-line vertices in the event.

As an example of an ECS calculation, Figure 4.4 shows the reference axis and u and t local coordinate axes for the surface rupture of the 1992 **M** 7.28 Landers earthquake. Overall, the reference line maintains smoothness and passes through the middle of the displacement measurement sites. Furthermore, the reference axis is consistent with the mapped ends of the fault rupture.

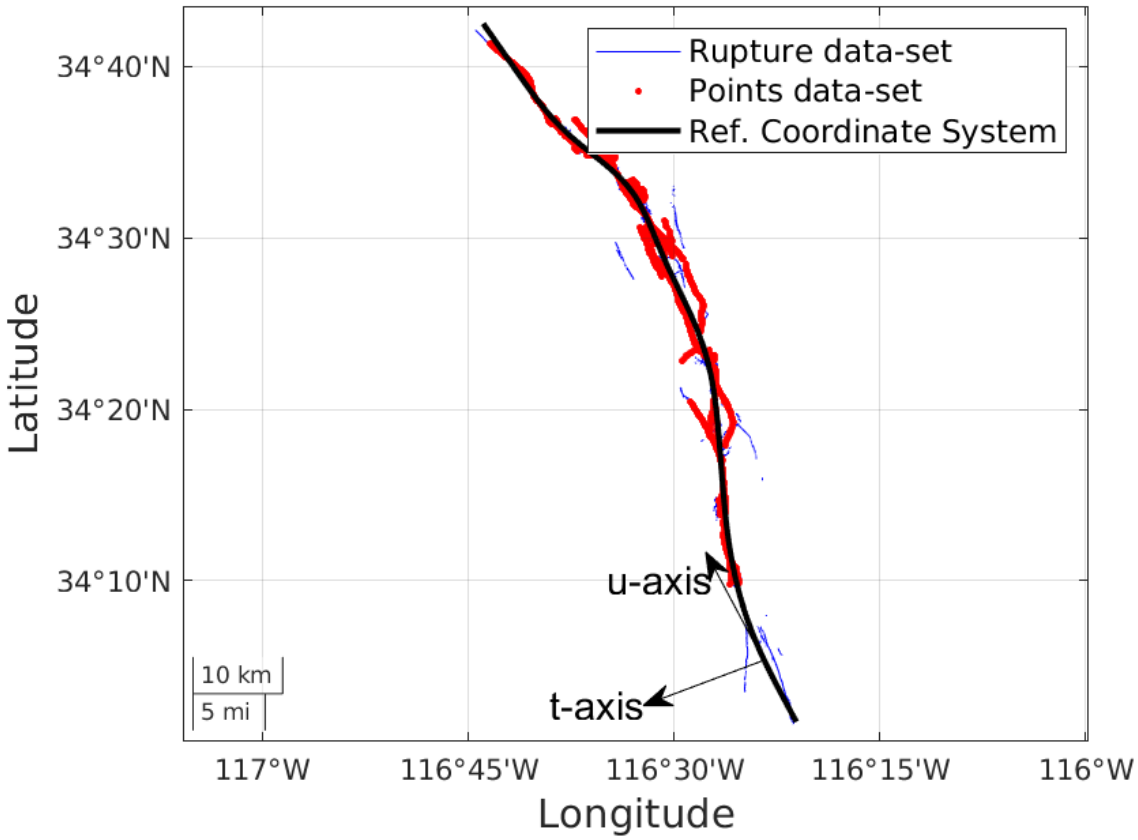

Figure 4.4. Event coordinate system for surface rupture of 1992 Landers, California **M** 7.28 earthquake (EQ_ID = 1).

4.6 RECOMMENDATIONS FOR MODEL DEVELOPERS

The FDHI Database was developed in collaboration with the model developers to ensure the content addressed model development needs. Below, we document differences in alternative rupture datasets that could impact the new models (Chapter 4.6.1), the development of

recommended net slip values and usage flags for each measurement in the database (Chapter 4.6.2), specific events with potential foreshock or aftershock contamination (Chapter 4.6.3), and specific events with spatial completeness limitations (Chapter 4.6.4).

4.6.1 Surface Rupture Data

We collected the highest quality surface rupture data available for the earthquakes in the database through our literature review (Chapter 3). When multiple supplementary surface rupture datasets were available for the same event, we manually combined the datasets to develop a single composite rupture dataset (Chapter 3.3.2 and Table 3.4). In a few cases, the available rupture datasets are alternatives (not supplements) and could not be combined due to different mapping scales in areas of overlap. The alternative rupture datasets are included in the database for completeness, and we do not identify a preferred dataset as any preference would depend on specific modeling needs. Table 4.3 lists the events that have alternative surface rupture datasets and characteristics of the individual datasets.

Table 4.3. Events with alternative surface rupture mapping datasets in the FDHI Database.

EQ_ID	Name	DS_ID	Scale[1]	Completeness	Source
3	EMC	17	Larger (1:500) (more detail), uniform throughout	Incomplete in liquefaction area (i.e., southeast of 32.268°N, 115.324°W)	Teran et al. (2015)
		18	Smaller (less detail), uniform throughout	Complete	Fletcher et al. (2014)
21	Darfield	80	Larger (more detail), uniform throughout	Complete	Villamor et al. (2012)
		103	Smaller (1:250,000) (less detail), uniform throughout	Complete	Langridge et al. (2016)
42	Ridgecrest1	132	Smaller (less detail), uniform throughout	Complete	DuRoss et al. (2020)
		145	Larger (more detail), varies throughout	Complete	Ponti et al. (2020)
43	Ridgecrest2	132	Smaller (less detail), uniform throughout	Complete	DuRoss et al. (2020)
		145	Larger (more detail), varies throughout	Complete	Ponti et al. (2020)

[1] Actual scale listed if reported in original source. Larger/smaller convention per Avery and Berlin (1992).

4.6.2 Fault Displacement Measurement Data

At the request of the model development teams, we provided recommended net slip values and usage flags for each measurement in the database. To complete this effort, we used custom measurement quality codes, a measurement technique compatibility identifier, and a measurement co-location identifier to guide our recommendations. Every measurement in the FDHI Database was evaluated in detail through this process.

We developed recommended net slip values for each measurement in the database. Fault displacement measurements are usually reported in the literature as a specific slip component, such as lateral slip or scarp height (Chapter 2.4.1). The individual slip component measurements as reported by the dataset originators are in the FDHI Database; however, to support model development, we also aggregated the reported slip components into recommended net slip values. Including this information in the FDHI Database allows the model development teams to develop displacement models for the same displacement metric, based on the same input data.

When dataset originators directly report a net (three-dimensional) slip component (TDS in Figure 2.10), we use this as the recommended net slip value; otherwise, the recommended net slip was calculated from the reported slip components for each measurement using basic trigonometric relationships (Figure 2.10 and Equations 2.1 through 2.3). In our experience, the dip angle and dip-slip component (ADS in Figure 2.10) were rarely reported in the source data, and the fault-normal component (FNS in Figure 2.10) was only occasionally reported; therefore, most of the calculated recommended net slip values imply a vertical fault (i.e., 90° dip). To systematically track the basis for the recommended net slip values, we list the reported slip components used in the calculation in a field called "recommended_net_vector_basis" in the database (cf. flatfile documentation in Appendix A). We also calculated upper and lower bounds of recommended net slips based on the bounding range calculated from the reported slip components.

We also created measurement quality codes to methodically document our assessments of the accuracy and consistency of every recommended net slip value in the database as part of our data quality review. The quality codes identify measurements with location errors or unreported slip components (relating to accuracy, as defined in Chapter 3.3.2) and sites that have alternative measurements (relating to consistency, as defined in Chapter 3.3.2). To track our consistency assessments, we also created a unique location identifier (“location_id”) for each earthquake and a compatibility or grouping identifier ("group_id"). Alternative measurements known or inferred to be at the same location have the same “location_id” (which is unique for each earthquake). The "group_id" field is used to explicitly separate the data in each earthquake into recommended sets that are internally compatible. The most common example is differentiating between wide-aperture measurements (e.g., based on optical image correlation) and field measurements collected on a discrete rupture. Other examples of incompatible measurements include events with datasets that mix vertical slip and scarp height, and measurement techniques that might unintentionally include slip from multiple events. Table 4.4 lists the groupings for each earthquake in the database.

Although all the information used to define the groupings is included in the database (e.g., individual slip components, measurement technique), we found that aggregating the relevant information into one field was a useful step towards developing recommended data usage flags for the model developers.

Table 4.4. Measurement technique groupings ("group_id" column) in the FDHI Database.

EQ_ID	Name	group_id	Measurement Technique
1	Landers	1_01	field-based measurements
1	Landers	1_02	optical image correlation
2	HectorMine	2_01	field-based measurements
2	HectorMine	2_02	optical image correlation
2	HectorMine	2_03	post-event lidar measurements (acquired ~10 yrs after earthquake)
3	EMC	3_01	field-based measurements
4	Balochistan	4_01	post-event high-resolution satellite imagery measurements
4	Balochistan	4_02	optical image correlation, densely spaced (~0.5 km average spacing)
4	Balochistan	4_03	optical image correlation, broadly spaced (~5.5 km average spacing)
5	Izmit_Kocaeli	5_01	field-based measurements
6	Borrego	6_01	field-based measurements
7	Imperial1979	7_01	field-based measurements
8	SuperstitionHills	8_01	field-based measurements
9	Kobe	9_01	field-based measurements
10	Denali	10_01	field-based measurements
11	Duzce	11_01	field-based measurements
12	Wenchuan	12_01	field-based measurements, based on vertical offset
12	Wenchuan	12_02	field-based measurements, based on scarp height
13	Napa	13_01	field-based measurements
14	Yushu	14_01	field-based measurements
15	Hualien	15_01	field-based measurements
15	Hualien	15_02	optical image correlation
16	ChiChi	16_01	field-based measurements
17	Kumamoto	17_01	field-based measurements
18	Nagano	18_01	field-based measurements
19	Kashmir	19_01	field-based measurements
20	Kaikoura	20_01	field-based measurements
20	Kaikoura	20_02	optical image correlation
21	Darfield	21_01	field-based measurements; post-event lidar measurements; post-event high-resolution satellite imagery measurements
22	Parkfield2004	22_01	field-based measurements

EQ_ID	Name	group_id	Measurement Technique
23	Norcia3	23_01	field-based measurements
24	Hebgen	24_01	post-event lidar measurements (acquired ~50 yrs after earthquake), based on vertical offset
24	Hebgen	24_02	field-based measurements, based on scarp height
25	SanFernando	25_01	field-based measurements
26	Bohol	26_01	field-based measurements
27	Acambay	27_01	field-based measurements (acquired ~125 yrs after earthquake)
28	Imperial1940	28_01	field-based measurements
29	Parkfield1966	29_01	field-based measurements
30	FairviewPeak	30_01	field-based measurements
31	DixieValley	31_01	field-based measurements
32	GalwayLake	32_01	field-based measurements
33	Sonora	33_01	field-based measurements (acquired ~125 yrs after earthquake)
34	PleasantValley	34_01	field-based measurements
35	Kern	35_01	field-based measurements
36	ChalfantValley	36_01	field-based measurements
37	Zirkuh	37_01	field-based measurements
38	Petermann	38_01	field-based measurements
38	Petermann	38_02	optical image correlation
39	OwensValley	39_01	field-based measurements (acquired ~100 yrs after earthquake)
39	OwensValley	39_02	post-event lidar measurements (acquired ~125 yrs after earthquake, with some field verification)
40	LagunaSalada	40_01	field-based measurements (acquired ~125 yrs after earthquake)
41	Iwaki2011	41_01	field-based measurements
42	Ridgecrest1	42_01	field-based measurements
43	Ridgecrest2	43_01	field-based measurements
44	ElAsnam	44_01	field-based measurements
45	Cadoux	45_01	field-based measurements
46	Calingiri	46_01	field-based measurements
47	MarryatCreek	47_01	field-based measurements
48	Meckering	48_01	field-based measurements
49	Pukatja	49_01	field-based measurements
50	TennantCreek1	50_01	field-based measurements
51	TennantCreek2	51_01	field-based measurements
52	TennantCreek3	52_01	field-based measurements
53	SanMiguel	53_01	field-based measurements
54	Yutian	54_01	field-based measurements
55	Luzon	55_01	field-based measurements
56	BorahPeak	56_01	field-based measurements
56	BorahPeak	56_02	post-event lidar measurements (acquired ~40 yrs after earthquake)

EQ_ID	Name	group_id	Measurement Technique
57	ElmoreRanch	57_01	field-based measurements
58	Pisayambo	58_01	field-based measurements
58	Pisayambo	58_02	InSAR slip inversion
59	Rikuu	59_01	field-based measurements
60	Mikawa	60_01	field-based measurements
61	IzuPeninsula	61_01	field-based measurements
62	IzuOshima	62_01	field-based measurements
63	IwateInland	63_01	field-based measurements
64	Edgecumbe	64_01	field-based measurements
65	Neftegorsk	65_01	field-based measurements
66	ChonKemin	66_01	field-based measurements (acquired ~100 yrs after earthquake)
67	Kunlun_Kokoxili	67_01	field-based measurements
67	Kunlun_Kokoxili	67_02	post-event high-resolution satellite imagery measurements
68	LeTeil	68_01	field-based measurements
68	LeTeil	68_02	InSAR slip inversion
69	Norcia1	69_01	field-based measurements
70	HomesteadValley	70_01	field-based measurements
71	Palu	71_01	field-based measurements
72	LAquila	72_01	field-based measurements
73	Spitak	73_01	field-based measurements
74	Killari	74_01	field-based measurements
75	YeniceGonen	75_01	field-based measurements

Recommended usage flags are included in the FDHI Database for each recommended net slip value. The flags are based on the quality codes and therefore are based on our assessment of the accuracy and consistency of the measurement. Table 4.5 lists the quality codes and the associated usage flag. We use three recommended usage flags: Keep, Check, and Toss. Recommended net slip values labeled as "Keep" are high quality data and can be used with confidence, provided that the model developer considers the rank (Chapter 4.3) and "group_id" associated with the recommended net slip value. Values labeled as "Toss" are low quality data that are erroneous or incomplete and should not be used for recommended net slip values; however, these measurement sites have other useful information (e.g., strike, dip), so they are preserved in the database. Finally, values labeled as "Check" might have quality issues related to consistency (i.e., alternative measurements) or accuracy (i.e., location errors or incomplete measurements), as documented in the quality code (Table 4.5). Model developers can use the quality codes to decide if values labeled as "Check" are appropriate for their models, again considering the rank and grouping associated with the recommended net slip value.

Table 4.5. Recommended net slip value quality codes used in the FDHI Database.

Quality Code	Explanation	Recommendation[1]	Model Development Usage Flag[1]
1	No known errors or issues (can be any rank or group_id)	Reliable data	Keep
2000	Multiple measurements (same rank and same group_id) available at same location_id (confident[2])	Review available alternative data	Check
2001	Multiple measurements (same rank and same group_id) available at same location_id (inferred[2])	Review available alternative data	Check
3000	Incomplete measurement, lateral slip component might be missing	Use with caution	Check
3001	Incomplete measurement, vertical slip component might be missing	Use with caution	Check
3002	Measurement might be minimum	Use with caution	Check
3003	Measurement might be maximum	Use with caution	Check
3004	Dataset originator quality is low	Use with caution	Check
3005	Deformation might not be tectonic	Use with caution	Check
3006	Incomplete measurement, extensional slip component might be missing	Use with caution	Check
4000	Location might be erroneous	Use with caution	Check
4001	Measurement might be erroneous	Use with caution	Check
5000	Measurement technique might mis-estimate vertical slip component	Use with caution	Check
9000	Other measurement at location_id is more complete	Unreliable data	Toss
9001	No measurement data	Unreliable data	Toss
9002	Incomplete measurement, significant lateral slip unaccounted for	Unreliable data	Toss
9003	Incomplete measurement, significant vertical slip unaccounted for	Unreliable data	Toss
9004	Measurement likely erroneous	Unreliable data	Toss
9005	Location likely erroneous	Unreliable data	Toss
9006	Deformation likely not tectonic	Unreliable data	Toss

[1] Applies to recommended net slip value; included in database for model developers.
[2] Measurements identified as co-located based on documentation from dataset originators (confident) or our evaluation of the reported slip components and site locations (inferred).

4.6.3 Foreshocks and Aftershocks

Spatiotemporal clustering of surface-rupturing earthquakes can cause difficulty in differentiating ruptures and displacements between events. We explicitly identify surface rupture and/or fault displacement data in the FDHI Database that might reflect deformation from an earthquake sequence ("multi_event_flag") or from an aftershock ("aftershock_flag"), where such information

is available. Two events in the database have areas that might have ruptured in an aftershock (1992 Landers, California and 2010 Yushu, China), and one event has data that captures both foreshocks and the mainshock (2016 Norcia, Italy). These events are listed in Table 4.6. Maps differentiating areas that might have ruptured in the mainshock and aftershock for the Landers, Yushu, and Kumamoto events are shown on Figure 4.5.

Table 4.6. Events in FDHI Database with potential foreshock or aftershock deformation.

EQ_ID	Name	Foreshock/Aftershock Notes
1	Landers	Southern-most portion (south of Pinto Mountain Fault) may have ruptured in aftershock; see Figure 4.5A (Hough et al., 1993)
14	Yushu	Northwestern portion may have ruptured in aftershock; see Figure 4.5B (Li et al., 2012)
23	Norcia3	Some measurements reflect unknown displacement from foreshocks (pers. comm., Boncio. P., based on: Brozzetti et al., 2019 and Villani et al., 2018b)

The 2016 Norcia, Italy **M** 6.6 earthquake (EQ_ID = 23) has more fault displacement measurements (n=5,718) than any other event in the FDHI Database. However, measurements from this event include an unspecified amount of displacement produced by foreshocks in areas that re-ruptured in the mainshock. The mainshock occurred on October 30, 2016 and was preceded by two surface-rupturing foreshocks on August 24, 2016 and October 26, 2016. The first foreshock (**M** 6.0 August 24, 2016) ruptured the southern portion of the mainshock rupture area, and the October 26, 2016 **M** 5.9 foreshock ruptured the northern portion. While some studies document fault displacements or displacement profiles of the August 24, 2016 foreshock (e.g., Villani et al., 2018a; Brozzetti et al., 2019), the contribution from foreshocks is not separated in the curated dataset used in the FDHI Database (Boncio, P., pers. comm.). The curated dataset was developed from extensive data quality reviews and was recommended by the model developers and SURE project colleagues as the authoritative dataset for this event. Model development teams and end users should be aware that the Norcia earthquake data in the FDHI Database is not strictly single-event, but rather includes an undetermined amount of deformation from **M** 6.0 and **M** 5.9 foreshocks.

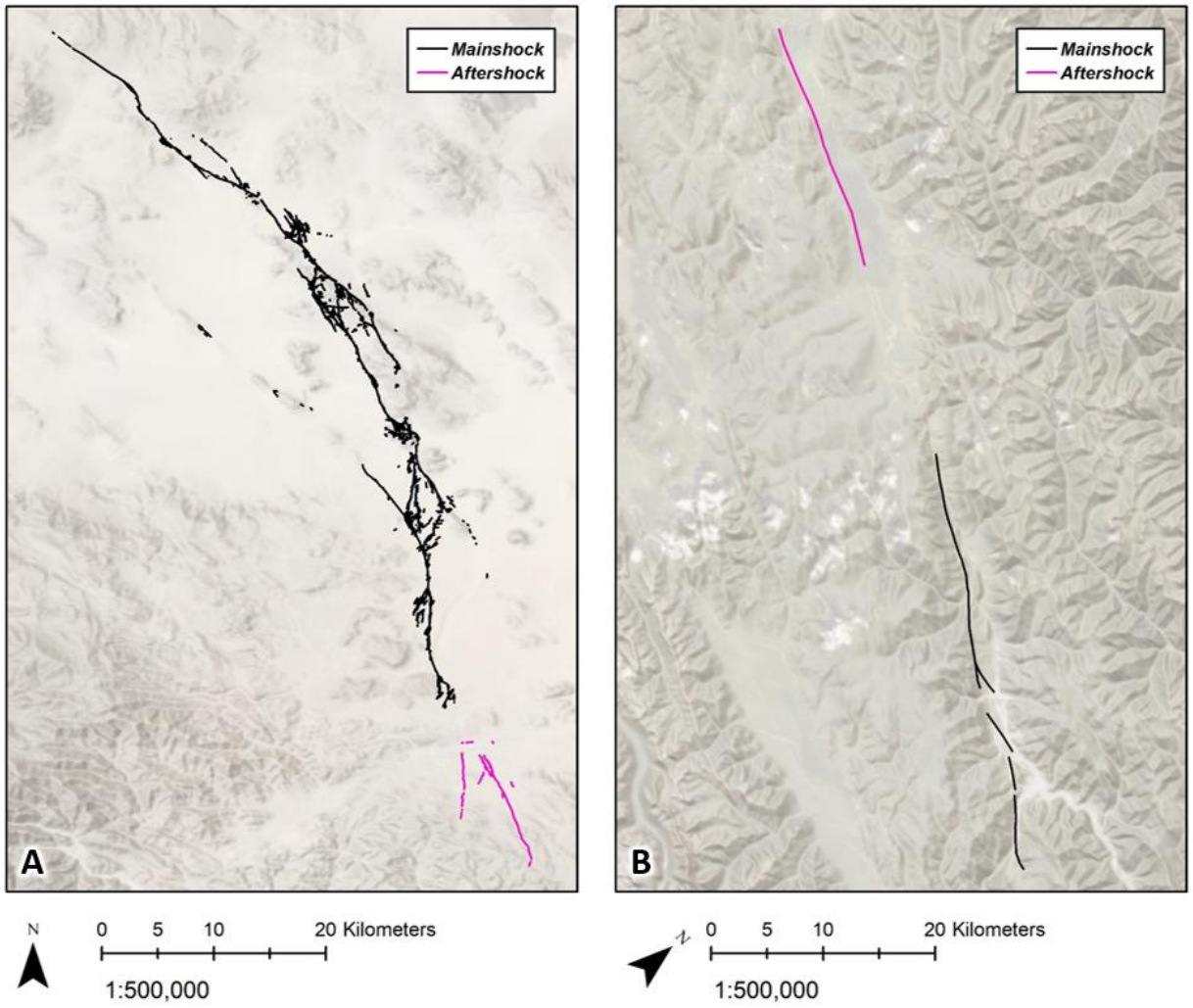

Figure 4.5. Spatial distribution of mainshock (black) and aftershock (magenta) surface ruptures in (A) 1992 **M** 7.28 Landers, California earthquake (EQ_ID = 1) and (B) 2010 **M** 6.9 Yushu, China earthquake (EQ_ID = 14).

Finally, we note two earthquake sequences in California (1987 Superstition Hills-Elmore Ranch and 2019 Ridgecrest) where the surface rupture and fault displacement data were successfully separated into individual events. The 1987 **M** 6.22 Elmore Ranch (EQ_ID = 57) earthquake ruptured several southwest-trending left-lateral faults and was shortly followed by the **M** 6.54 Superstition Hills earthquake (EQ_ID = 8), which ruptured a southeast-trending right-lateral fault system. The first event occurred at approximately six o'clock in the evening local time, and field investigation teams were not able to evaluate surface rutpures before the second event occurred roughly 12 hours later. Surface ruptures from the 1987 Superstition Hills-Elmore Ranch sequence are commonly differentiated based on fault strike and style of faulting (Sharp et al., 1989) (Figure 4.6.). Similarly, the 2019 Ridgrecrest earthquake sequence included two surface-rutpuring earthquakes that occurred 34 hours apart. Rapid response by field investiagtion and geodesy teams allowed surface ruptures and fault displacements from the **M** 6.4 foreshock (EQ_ID = 42) to be

documented prior to the **M** 7.1 mainshock (EQ_ID = 43), allowing the data from the sequence to be reliably separated into individual events (DuRoss et al., 2020; Milliner and Donnellan, 2020) (Figure 4.7).

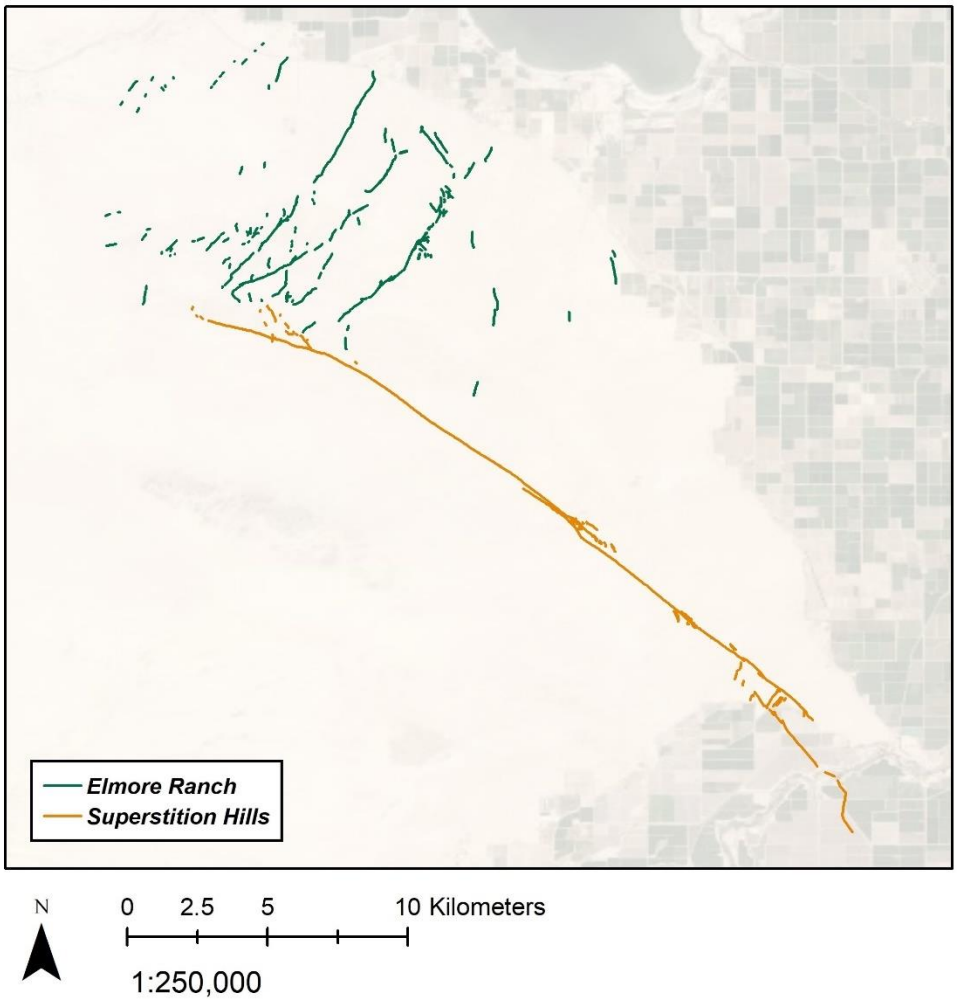

Figure 4.6. Surface ruptures from 1987 Superstition Hills-Elmore Ranch, California earthquake sequence. Green lines: **M** 6.22 Elmore Ranch earthquake (EQ_ID = 57). Orange lines: **M** 6.54 Superstition Hills earthquake (EQ_ID = 8).

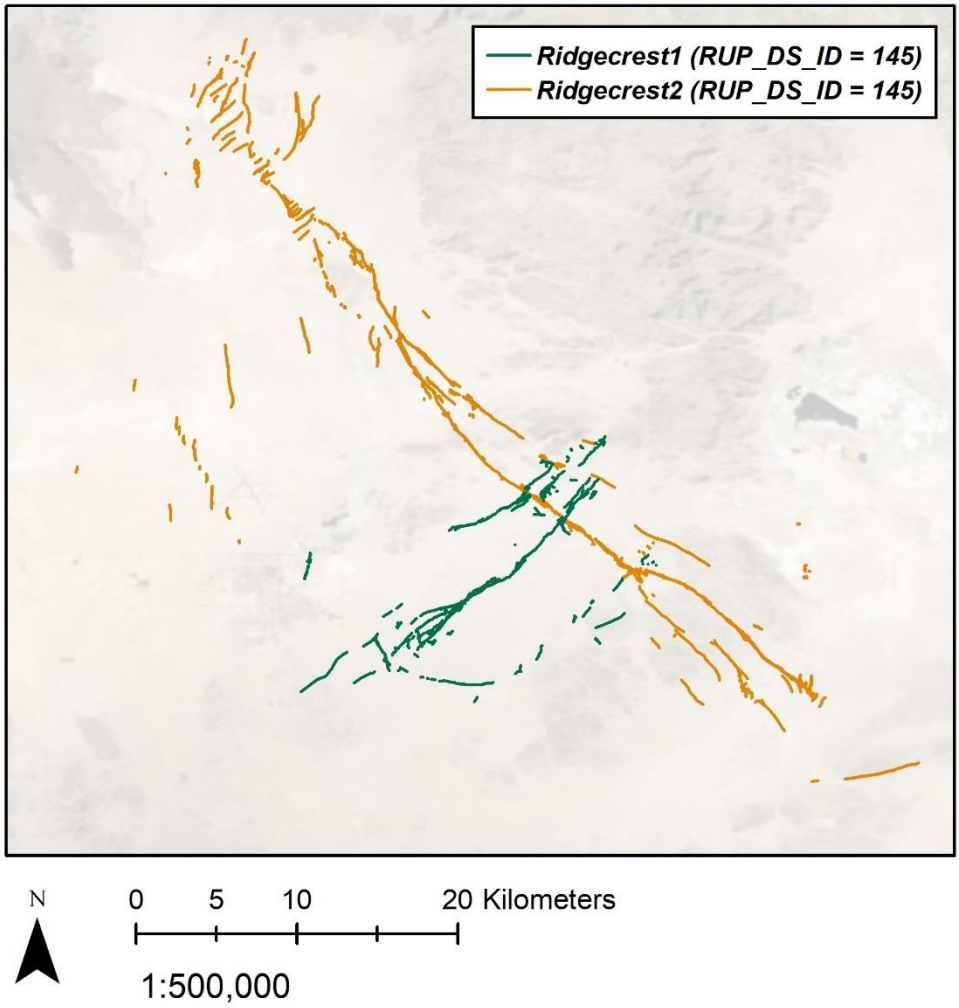

Figure 4.7. Surface ruptures from 2019 Ridgecrest, California earthquake sequence. Green lines: **M** 6.4 Ridgecrest1 earthquake (EQ_ID = 42). Orange lines: **M** 7.1 Ridgecrest2 earthquake (EQ_ID = 43). See Chapter 4.6.1 and Table 4.3 for discussion on alternative surface rupture datasets for the Ridgecrest earthquakes.

4.6.4 Spatial Completeness Limitations

As discussed in Chapter 2, logistical constraints can preclude full documentation of surface ruptures and fault displacement measurements in an earthquake. The level of detail in rupture mapping can vary in different areas of the rupture, and the spatial distribution of measurement sites is nonuniform. As part of our data quality review, we evaluated the completeness of the data for each earthquake in the FDHI Database relative to the known spatial extent of the surface rupture. While most of the events in the FDHI Database generally have complete spatial coverage of surface ruptures and measurements (notwithstanding variations in mapping scale and nonuniform spacing of measurement sites), a subset of events listed in Table 4.7 have incomplete

data in specific areas. Two earthquakes (1992 Landers, California and 2010 Yushu, China) have known slip gaps near areas that may have ruptured in aftershocks (Figure 4.5 and Table 4.6).

Table 4.7. Events in FDHI Database with known spatial completeness limitations.

EQ_ID	Name	Spatial Completeness Notes
1	Landers	Surface rupture mapping and measurements are complete; known slip gap near 34.147°N, 116.416°W
3	EMC	Extensive liquefaction southeast of 32.268°N, 115.324°W; no measurements in liquefaction area; no surface rupture mapping in "DS_ID = 17" in liquefaction area, but "DS_ID = 18" surface rupture mapping is complete[1]
5	Izmit_Kocaeli	Surface rupture mapping is complete, but no measurements in Sea of Marmara and Lake Sapanca
9	Kobe	Surface rupture mapping and measurements on Awaji Island are complete; possible undocumented rupture offshore (to northwest in Akashi Strait)
12	Wenchuan	Possible undocumented rupture to southwest
13	Napa	Surface rupture mapping is complete, but no measurements south of 38.225°N, 122.311°W (in Napa River estuary)
14	Yushu	Surface rupture mapping and measurements are complete; known slip gap near 33.135°N, 96.667°E
15	Hualien	Surface rupture mapping and measurements on island of Taiwan are complete; undocumented rupture offshore to northeast, and possible undocumented offshore to southeast
20	Kaikoura	Surface rupture mapping is complete; measurements in "group_id = 20_01" are concentrated in onshore Northern Domain and on Hundalee Fault[2]
26	Bohol	Surface rupture mapping on Bohol Island is complete; possible undocumented rupture offshore (to southwest in Cebu Strait); measurements are concentrated at northeastern area of rupture
27	Acambay	Surface rupture mapping is complete; measurements are concentrated at southeastern area of rupture
56	BorahPeak	Surface rupture mapping is complete; measurements in "group_id = 56_02" are concentrated in Thousand Springs-Mackay Fault area[2]
61	IzuPeninsula	Surface rupture mapping and measurements on Honshu Island are complete; possible undocumented rupture offshore (to southeast)
62	IzuOshima	Surface rupture mapping and measurements on Honshu Island are complete; possible undocumented rupture offshore (to southeast)
71	Palu	Surface rupture mapping is complete in onshore and offshore portion south of Tanimbaya Peninsula; possible undocumented rupture to north-northwest; no measurements offshore

[1] The two alternative surface rupture maps for the 2010 El Mayor-Cucapah (EMC) earthquake are differentiated by the dataset identification number (DS_ID); see Chapter 4.6.1 for discussion.
[2] The grouping identifier ("group_id") is used to explicitly separate the data in each earthquake into internally compatible sets based on measurement technique and aperture; see Chapter 4.6.2 for discussion.

4.7 SOFTWARE

The following software was used in the data analysis:

- ESRI ArcMap and ArcGIS Desktop software version 10.7, Advanced license
- Geospatial Data Abstraction Library (GDAL) version 3.2.1
- Scientific Python (SciPy) version 1.6.1

4.8 REFERENCES

Alaska Geospatial Council (2021). Alaska IFSAR Data and Status Web App. URL: http://agc.dnr.alaska.gov/elevation.html.

Avery, T. E. and Berlin, G. L. (1992). Fundamentals of remote sensing and airphoto interpretation. Macmillan.

Baize, S., Nurminen, F., Sarmiento, A., Dawson, T., Takao, M., Scotti, O., Azuma, T., Boncio, P., Champenois, J., Cinti, F. R., Civico, R., Costa, C., Guerrieri, L., Marti, E., McCalpin, J., Okumura, K., and Villamor, P. (2020). A worldwide and Unified Database of Surface Ruptures (SURE) for fault displacement hazard analyses, *Seismological Research Letters*, **91**(1), 499-520.

Brozzetti, F., Boncio, P., Cirillo, D., Ferrarini, F., Nardis, R., Testa, A., Liberi, F., and Lavecchia, G. (2019). High-resolution field mapping and analysis of the August–October 2016 coseismic surface faulting (central Italy earthquakes): Slip distribution, parameterization, and comparison with global earthquakes, *Tectonics*, **38**, 417–439.

Coppersmith, K. J. and Youngs, R. R. (2000). Data needs for probabilistic fault displacement hazard analysis, *Journal of Geodynamics*, **29**(3-5), 329-343.

DuRoss, C. B., Gold, R. D., Dawson, T. E., Scharer, K. M., Kendrick, K. J., Akciz, S. O., Angster, S. J., Bachhuber, J., Bacon, S., Bennett, S. E. K., et al. (2020). Surface Displacement Distributions for the July 2019 Ridgecrest, California, Earthquake Ruptures, *Bulletin of the Seismological Society of America*, **110**(4), 1400-1418.

Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D., and Alsdorf, D. (2007). The shuttle radar topography mission, *Reviews of Geophysics*, **45**(2).

Fletcher, J. M., Teran, O. J., Rockwell, T. K., Oskin, M. E., Hudnut, K. W., Mueller, K. J., Spelz, R. M., Akciz, S. O., Masana, E., Faneros, G., Fielding, E. J., Leprince, S., Morelan, A. E., Stock, J., Lynch, D. K., Elliott, A. J., Gold, P., Liu-Zeng, J., González-Ortega, A., Hinojosa-Corona, A., and González-García, J. (2014). Assembly of a large earthquake

from a complex fault system: Surface rupture kinematics of the 4 April 2010 El Mayor–Cucapah (Mexico) Mw 7.2 earthquake, *Geosphere*, **10**(4), 797-827.

Hough, S. E., Mori, J., Sembera, E., Glassmoyer, G., Mueller, C., and Lydeen, S. (1993). Southern surface rupture associated with the 1992 M7.4 Landers Earthquake: Did it all happen during the mainshock?, *Geophysical Research Letters*, **20**(23), 2615-2618.

Iwahashi, J., Kamiya, I., Matsuoka, M., and Yamazaki, D. (2018). Global terrain classification using 280 m DEMs: segmentation, clustering, and reclassification, *Progress in Earth and Planetary Science*, **5**(1), 1-31.

Langridge, R. M., Ries, W. F., Litchfield, N. J., Villamor, P., Van Dissen, R. J., Barrell, D. J. A., Rattenbury, M. S., Heron, D. W., Haubrock, S., Townsend, D. B., Lee, J. M., Berryman, K. R., Nicol, A., Cox, S.C., and Stirling, M. W. (2016). The New Zealand Active Faults Database, *New Zealand Journal of Geology and Geophysics*, **59**(1), 86-96.

Li, C. Y., Pang, J. Z., and Zhang, Z. Q. (2012). Characteristics, geometry, and segmentation of the surface rupture associated with the 14 April 2010 Yushu earthquake, eastern Tibet, China, *Bulletin of the Seismological Society of America*, **102**(4), 1618-1638.

Milliner, C. and Donnellan, A. (2020). Using Daily Observations from Planet Labs Satellite Imagery to Separate the Surface Deformation between the 4 July Mw 6.4 Foreshock and 5 July Mw 7.1 Mainshock during the 2019 Ridgecrest Earthquake Sequence, *Seismological Research Letters*, **91**(4), 1986-1997.

Milliner, C. W., Dolan, J. F., Hollingsworth, J., Leprince, S., Ayoub, F., and Sammis, C. G. (2015). Quantifying near-field and off-fault deformation patterns of the 1992 Mw 7.3 Landers earthquake, *Geochemistry, Geophysics, Geosystems*, **16**(5), 1577-1598.

Nurminen, F., Boncio, P., Visini, F., Pace, B., Valentini, A., Baize, S., and Scotti, O. (2020). Probability of occurrence and displacement regression of distributed surface rupturing for reverse earthquakes, *Frontiers in Earth Science*, **8**(581605).

Petersen, M. D., Dawson, T. E., Chen, R., Cao, T., Wills, C. J., Schwartz, D. P., and Frankel, A. D. (2011). Fault displacement hazard for strike-slip faults, *Bulletin of the Seismological Society of America*, **101**(2), 805-825.

Ponti, D. J., Blair, J. L., Rosa, C. M., Thomas, K., Pickering, A. J., Akciz, S., Angster, S., Avouac, J.-P., Bachhuber, J., Bacon, S., et al. (2020). Documentation of Surface Fault Rupture and Ground-Deformation Features Produced by the 4 and 5 July 2019 Mw 6.4 and Mw 7.1 Ridgecrest Earthquake Sequence, *Seismological Research Letters*, **91**(5), 2942-2959.

Rai, M., Rodriguez-Marek, A., and Chiou, B. S. (2017). Empirical terrain-based topographic modification factors for use in ground motion prediction, *Earthquake Spectra*, **33**(1), 157-177.

Riley, S. J., DeGloria, S. D., and Elliot, R. (1999). A terrain ruggedness index that quantifies topographic heterogeneity, *Intermountain Journal of Sciences*, **5**(1-4), 23-27.

Sharp, R. V., Budding, K. E., Boatwright, J., Ader, M. J., Bonilla, M. G., Clark, M. M., Fumal, T. E., Harms, K. K., Lienkaemper, J. J., Morton, D. M., O'Neill, B. J., Ostergren, C. L., Ponti, D. J., Rymer, M. J., Saxton, J. L., and Sims, J. D. (1989). Surface faulting along the Superstition Hills fault zone and nearby faults associated with the earthquakes of 24 November 1987, *Bulletin of the Seismological Society of America*, **79**(2), 252-281.

Spudich, P., and Chiou, B. (2015). Strike-parallel and strike-normal coordinate system around geometrically complicated rupture traces: Use by NGA-West2 and further improvements. US Department of the Interior, US Geological Survey.

Teran, O. J., Fletcher, J. M., Oskin, M. E., Rockwell, T. K., Hudnut, K. W., Spelz, R. M., Akciz, S. O., Hernandez-Flores, A. P., and Morelan, A. E. (2015). Geologic and structural controls on rupture zone fabric: A field-based study of the 2010 Mw 7.2 El Mayor–Cucapah earthquake surface rupture, *Geosphere*, **11**(3), 899-920.

Villamor, P., Litchfield, N., Barrell, D., Van Dissen, R., Hornblow, S., Quigley, M., Levick, S., Ries, W., Duffy, B., Begg, J., Townsend, D., Stahl, T., Bilderback, E., Noble, D., Furlong, K., and Grant, H. (2012). Map of the 2010 Greendale Fault surface rupture, Canterbury, New Zealand: application to land use planning, *New Zealand Journal of Geology and Geophysics*, **55**(3), 223-230.

Villani, F., Pucci, S., Civico, R., De Martini, P. M., Cinti, F. R., and Pantosti, D. (2018a). Surface faulting of the 30 October 2016 Mw 6.5 central Italy earthquake: Detailed analysis of a complex coseismic rupture, *Tectonics*, **37**, 3378–3410.

Villani, F., Civico, R., Pucci, S., Pizzimenti, L., Nappi, R., De Martini, P. M., and Open EMERGEO Working Group (2018b). A database of the coseismic effects following the 30 October 2016 Norcia earthquake in central Italy, *Scientific Data*, **5**, 180049.

Youngs, R. R., Arabasz, W. J., Anderson, R. E., Ramelli, A. R., Ake, J. P., Slemmons, D. B., McCalpin, J. P., Doser, D. I., Fridrich, C. J., Swan III, F. H., Rogers, A. M., Yount, J. C., Anderson, L. W., Smith, K. D., Bruhn, R. L., Knuepfer, P. L. K., Smith, R. B., dePolo, C. M., O'Leary, D. W., Coppersmith, K. J., Pezzopane, S. K., Schwartz, D. P., Whitney, J. W., Olig, S. S., and Toro, G. R. (2003). A methodology for probabilistic fault displacement hazard analysis (PFDHA), *Earthquake Spectra*, **19**(1), 191-219.

5 Relational Database Development

5.1 INTRODUCTION

A custom relational database was created to systematically manage the event, measurement, and rupture data and related metadata that were assembled and developed for this project. Relational databases use a defined schema to store different data types in individual tables, relate the data between tables using key fields, and hold the information and schema in a single file. Alternative data repository formats are typically collections of separate spreadsheets with limited or no cross-referencing. The relational database structure improves efficiency, quality control, and expandability, relative to spreadsheet formats, by minimizing errors due to repetition, enforcing data entry constraints, and enforcing defined relationships between tables. Relational databases are relatively new to the geotechnical engineering community, but they are well-established in the information technology and petroleum industries (Hoffman, 2003; Brandenberg et al., 2018; Mazzoni et al., 2020).

Relational database management systems can be server-based (client-server model) or embedded (serverless). For this project, we sought an open-source management system with a wide range of programming language support (e.g., Matlab toolbox, Python or R libraries) that was compatible with multiple computer operating systems (Windows, Macintosh, and Linux). We also decided a serverless management system was more appropriate because the database would not require multiple users to simultaneously update or query data entries. Based on these criteria, the SQLite database engine (Hipp, 2020) was selected as the relational database management system. Specific software versioning is reported in Chapter 5.3 and Appendix B.

This Chapter provides an overview of the relational database structure (or schema) and contents. More details bearing on the individual tables and schema are provided in Appendix B. The contents of the database have been aggregated into flatfiles for formal documentation and end-user convenience (Chapter 6 and Appendix A). We recommend most users of the FDHI Database use the flatfiles.

5.2 DATABASE STRUCTURE AND CONTENTS

The process of designing the relational database began with a systematic review of surface rupture characteristics, data collection tools, techniques, and reporting standards (Chapter 2), and existing fault displacement and surface rupture compilations (Chapter 3). We collaborated with the model developers to determine the initial database contents and then developed a custom schema to accommodate the range of data types. As the project progressed, additional data and interpretations requested by the model development teams were readily accommodated by the custom and flexible database schema.

Several different types of data are available to document historical surface-rupturing earthquakes. For this project, we broadly grouped the data types into four categories: earthquake information, rupture information, measurement information, and the event-specific coordinate system (ECS) model (Figure 5.1). Each category contains information such as metadata, geospatial data, direct observations, analysis outputs, or interpretations, as described in Chapters 3 and 4 of this report.

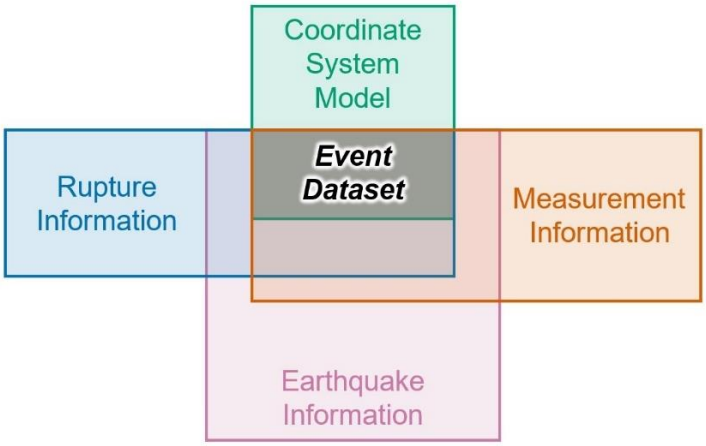

Figure 5.1. Schematic showing four data-type categories that collectively describe an earthquake (event) dataset.

The core database structure is shown in the relational schema diagram in Figure 5.2. Four database tables are emphasized by the yellow shading in the diagram, corresponding to the four data type categories from Figure 5.1. Placeholder table names ("RUP_otherTables" and "PT_otherTables") are shown in Figure 5.2 to illustrate the general relationship between the core database structure and the other individual observation or interpretation database. Table 5.1 summarizes the relationship between the data type categories and the core database schema. The entire database contains 37 individual tables. Appendix B contains additional documentation on the database schema, including lists of every table and column in the database and access to a digital version of the full schema.

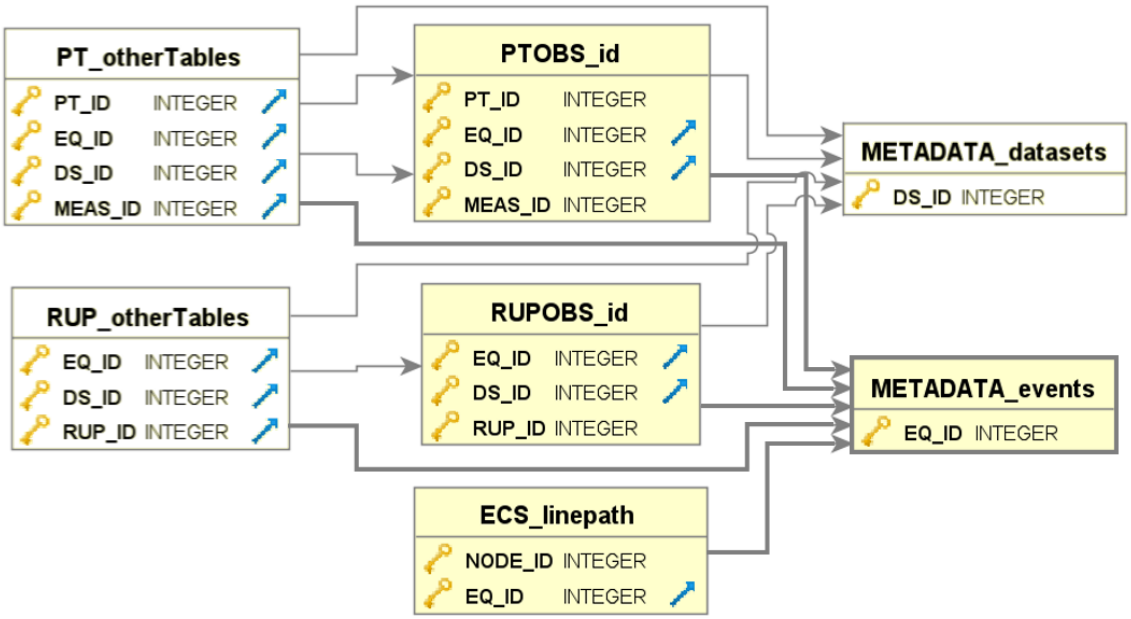

Figure 5.2. Relational schema diagram showing the core FDHI Database structure. Gold key symbol and blue arrow symbol represent primary and foreign keys, respectively.

Table 5.1. Parent tables in FDHI Database.

Data Category	Database Table Name	Table Type	Table Purpose/Contents
Earthquake Information	METADATA_events	Parent	Assign event identifier (EQ_ID); store event metadata
Measurement Information	PTOBS_id	Child & Parent	Assign measurement identifiers (PT_ID & MEAS_ID)
Rupture Information	RUPOBS_id	Child & Parent	Assign rupture line identifier (RUP_ID)
Coordinate System Model	ECS_linepath	Child	Store geographic coordinates for ECS reference line

The database contains metadata and geospatially-controlled surface rupture and fault displacement data from 75 global historical earthquakes. The process of developing the event metadata and surface rupture/fault displacement data and metadata is described in Chapter 3.3. Similarly, Chapter 4 documents the data analyses and interpretations performed to support the model development. Table 5.2 is a general summary of the database contents.

Table 5.2. Generalized list database contents.

Data Category	Contents[1]
Earthquake Information	EQ_ID, name, region, date, style, magnitude, magnitude type, seismic moment, hypocenter
Measurement Information	PT_ID, MEAS_ID, location_id, group_id, geographic coordinates, elevation data and metrics, slip measurements, site geology, classification/rank, recommended net slip values, recommended net slip quality code and suggested usage
Rupture Information	RUP_ID, NODE_ID, geographic coordinates, site geology, mapping accuracy/confidence, classification/rank
Coordinate System Model	reference line geographic coordinates, ECS ordinates for measurement sites, ECS ordinates for rupture line vertices

[1] Simplified listing of contents

5.3 SOFTWARE

Figure 5.2 was made using "DbVisualizer" (https://www.dbvis.com/). The following software versions were used to build, populate, and query the FDHI Database:

- SQLite database engine version 3.14.2 (Hipp, 2020; https://www.sqlite.org)
- Python version 2.7.15 (https://www.python.org)
- Python "sqlite3" module version 2.6.0 (https://docs.python.org)
- Python "pandas" library version 0.18.1 (https://pandas.pydata.org/)

5.4 REFERENCES

Brandenberg, S. J., Kwak, D. Y., Zimmaro, P., Bozorgnia, Y., Kramer, S. L., and Stewart, J. P. (2018). Next-Generation Liquefaction (NGL) case history database structure. In *Geotechnical Earthquake Engineering and Soil Dynamics V: Liquefaction Triggering, Consequences, and Mitigation*, American Society of Civil Engineers, 426-433.

Hipp, R. D. (2020). SQLite. Retrieved from https://www.sqlite.org/index.html.

Hoffman, D. R. (2003). *Effective Database Design for Geoscience Professionals*. PennWell Books.

Mazzoni S., Kishida T., Ahdi S.K., Contreras V., Darragh R.B., Chiou B.S.-J., Kuehn N., Bozorgnia Y., Stewart J.P. (2020). Chapter 2: Relational Database, in Data Resources for NGA-Subduction Project, PEER Report No. 2020/02, *Pacific Earthquake Engineering Research Cente*r, University of California, Berkeley, CA.

6 Flatfile Documentation

The relational database contents have been aggregated into flatfiles for documentation and usability. The database contains fault displacement measurements, surface rupture maps, and associated metadata (including event information, analysis results, and geologic interpretations), and this information is contained across 37 tables and 365 columns in one relational database file. As described in Chapter 5, we broadly grouped the content into four categories: earthquake information, measurement information, rupture information, and the event-specific coordinate system (ECS) model. We provide separate flatfiles for the three latter information categories: (1) a measurements flatfile; (2) a ruptures flatfile; and (3) an ECS model flatfile. Earthquake information is contained in all three flatfiles.

The flatfiles are the formal documentation of the database contents. We used our knowledge of the database schema to produce the flatfiles and check for errors and inconsistencies. We recommend most users of the FDHI Database (including model developers, geoscience researchers, and industry professionals) use the flatfiles to access the contents of the database. Appendix A provides more information on the flatfiles.

The FDHI Database flatfiles and related digital products are publicly available at https://doi.org/10.25346/S6/Y4F9LJ. The flatfiles are provided in comma-separated (CSV) file format, and the related digital products provide the flatfiles in map-based formats (i.e., Esri shapefiles, and Google Earth KMZ files.)

7 Quality Assurance and Quality Control

The FDHI Database development included a robust quality assurance (QA) and quality control (QC) effort to ensure the database contents were carefully assessed for quality and content for use in the development of the new fault displacement models. For this project, we consider evaluations of data quality (i.e., completeness, accuracy, and consistency; Chapter 3.3.2) to address QA, and data requests and reviews by the model development teams as relating to QC. The key components of the QA/QC effort included collaborating with the model developers to define the database contents, using a structured relational database, developing standard workflows to review and process datasets, performing analysis and interpretation of the data, and engaging the model developers in a participatory peer review of interim database versions. The QA/QC measures applied to the FDHI Database have resulted in a more reliable, stable, and useful product.

Our standard workflow for developing surface rupture, measurement, and event information for each earthquake (Chapter 3.3) was designed to support QA. Each candidate dataset was carefully reviewed for data quality and compliance with the event and dataset selection criteria (Chapter 3.1). The Database Team met regularly (approximately bi-weekly for two years) to review and discuss individual datasets and earthquake characteristics. The work developing the rank classifications (Chapter 4.3) and recommended net slip values and quality codes (Chapter 4.6.2) resulted in a comprehensive QA effort in which every entry of each dataset was evaluated in detail. We assigned quality codes to each measurement to identify good/reliable data, alternative data, potentially incomplete or erroneous data, and unreliable data (Table 4.2). In developing the rank classifications, different members of the Database Team independently developed rankings for the same event or subsets of the same event, and the results were compared and discussed. In general, there was high reliability and repeatability of the rankings; in some complex ruptures, the variations captured technically defensible alternative interpretations, and we coordinated to develop a preferred interpretation.

The database was created primarily for model developers to use in developing new fault displacement models; therefore, our QC efforts focused on ensuring the database content addressed model development needs. We held monthly meetings with the model developers for almost two years and attended several model development working group meetings in that time. The event and dataset criteria (Chapter 3.1) were established based on the modeling needs, and specific analysis and geologic interpretation (Chapter 4) of the raw data were also performed to support

model development. Participatory reviews of interim internal database versions by the modeling teams helped identify content important to the model development (QC) and data elements that needed further review (QA). Finally, using a relational database structure also supports the QA/QC effort by minimizing errors due to repetition, enforcing data entry constraints, and maintaining important references between data elements.

8 Conclusions

We have assembled a geospatially-controlled relational database of surface rupture maps, measurements, and associated metadata for 75 historical earthquakes of **M** 4.9 to 8.0 for all styles of faulting. All information is contained in a structured relational database, and the contents have been aggregated into flatfiles (*.csv format), ESRI shapefiles, and Google Earth files for formal documentation and end-user convenience. The work was completed as part of the Fault Displacement Hazard Initiative (FDHI) Project to support the development of next-generation fault displacement models, and the FDHI Database was developed in collaboration with the model developers. The new fault displacement models are anticipated to provide improved estimates of the amplitude and spatial distribution of principal and distributed displacements for future surface-rupturing earthquakes. While several fault displacement models are currently used in standard practice, there are significant differences in their input datasets, estimated displacement metrics, modeling techniques, and treatment of uncertainty. The FDHI Project will help mitigate these critical issues by using a common database and producing independent models in a coordinated research program. The new models will be useful for engineering design and analysis of critical infrastructure located on or near active fault zones and will be applicable for both deterministic and probabilistic fault displacement hazard analysis.

The data quality review, analysis, and geologic interpretation efforts completed in this project are a unique feature of the FDHI Database and have resulted in a reliable and stable product for model development teams and the geoscience community. The data were collected through an extensive literature review and were systematically assessed for completeness, accuracy, and consistency. Multiple source datasets are included for the same earthquake, where available, allowing database users to make comparisons in a common framework. The database also includes geologic data and terrain metrics, which have not been included in previous databases, allowing model developers to investigate geologic and topographic controls on fault displacements. The development and application of a new event-specific coordinate system (ECS) algorithm herein supplements geographic coordinates with strike-parallel and strike-normal ordinates for all surface rupture linework and measurement locations. All surface ruptures in the database are classified as principal or distributed rank based on detailed geologic evaluations. We introduce two additional measurement rank classifications (cumulative and total) in this project to better distinguish measurements associated with multiple ruptures or wide measurement apertures. While the

classification scheme may be non-unique, it has been applied as consistently as possible across the contents of the database. Hanging wall and footwall flags are included for distributed measurements in reverse, normal, and oblique style earthquakes. We also provide preferred and bounding (e.g., maximum and minimum) recommended net slip values calculated from the reported slip components. The basis for the calculations is tracked in the database, and each value is assigned a quality code. Finally, the structured relational database created for this project was designed to be expandable and extensible as additional earthquake data become available, new measurement techniques develop, and user needs evolve.