Focus Section: Historical Earthquake Data and Research

California Historical Intensity Mapping Project (CHIMP): A Consistently Reinterpreted Dataset of Seismic Intensities for the Past 162 Yr and Implications for Seismic Hazard Maps

Leah Salditch[*1,2], Molly M. Gallahue[1], Madeleine C. Lucas[1], James S. Neely[1,2], Susan E. Hough[3], and Seth Stein[1,2]

Abstract

Historical seismic intensity data are useful for myriad reasons, including assessment of the performance of probabilistic seismic hazard assessment (PSHA) models and corresponding hazard maps by comparing their predictions to a dataset of historically observed intensities in the region. To assess PSHA models for California, a long and consistently interpreted intensity record is needed. For this purpose, the California Historical Intensity Mapping Project (CHIMP) has compiled a dataset that combines and reinterprets intensity information that has been stored in disparate and sometimes hard-to-access locations. The CHIMP dataset also includes new observations of intensity from archival research and oral history collection. Version 1 of the dataset includes 46,502 intensity observations for 62 earthquakes with estimated magnitudes ranging from 4.7 to 7.9. The 162 yr of shaking data show observed shaking lower than expected from seismic hazard models. This discrepancy is reduced, but persists, if historical intensity data for the largest earthquakes are smoothed to reduce the effects of spatial undersampling. Possible reasons for this discrepancy include other limitations of the CHIMP dataset, the hazard models, and the possibility that California seismicity throughout the historical period has been lower than the long-term average. Some of these issues may also explain similar discrepancies observed for Italy and Japan.

Cite this article as Salditch, L., M. M. Gallahue, M. C. Lucas, J. S. Neely, S. E. Hough, and S. Stein (2020). California Historical Intensity Mapping Project (CHIMP): A Consistently Reinterpreted Dataset of Seismic Intensities for the Past 162 Yr and Implications for Seismic Hazard Maps, *Seismol. Res. Lett.* **XX**, 1–20, doi: 10.1785/0220200065.

Supplemental Material

Introduction

Because major earthquakes and the resulting strong shaking are rare events in any one area, it is difficult to assess how well earthquake hazard models and the corresponding maps describe the actual shaking that occurs. The problem is challenging both because of limitations in the available data and because of conceptual issues in how to assess the performance of probabilistic forecasts (Gneiting and Katzfuss, 2014; Marzocchi and Jordan, 2014; Wang, 2015; Vanneste *et al.*, 2018). Ideally, assessments should be prospective, that is, use only shaking that occurred after a model was made. For example, Brooks *et al.* (2018, 2019) compared intensities from (presumed) induced earthquakes in the central United States with predictions of a 1-yr hazard model. However, shaking data recorded since probabilistic seismic hazard assessment (PSHA) began typically span a time period that is short compared to the return period of a PSHA model, and hence rarely include data from the moderate and large earthquakes that control hazard. The methods referenced previously have therefore been developed to allow historically observed intensities to be compared numerically to PSHA model predictions, over the duration of the historical catalog.

Retrospective assessments, or hindcasting, using compilations of historical shaking data spanning hundreds of years, can reduce this problem (Stirling and Petersen, 2006; Stirling and Gerstenberger, 2010; Mak *et al.*, 2014; Nekrasova *et al.*, 2014; Stein *et al.*, 2015; Brooks *et al.*, 2016, 2017, 2018, 2019; Mak and Schorlemmer, 2016). Such assessments compare PSHA model predictions to historically observed shaking from earthquakes that occurred before the models were made.

1. Department of Earth and Planetary Sciences, Northwestern University, Evanston, Illinois, U.S.A.; 2. Northwestern Institute for Policy Research, Evanston, Illinois, U.S.A.; 3. U.S. Geological Survey, Pasadena, California, U.S.A.

*Corresponding author: leah@earth.northwestern.edu

We use the term "historical" to mean the time before the modern instrumental catalog, that is, pre-1970s and going back to the earliest available written accounts of shaking.

In this article, we consider the U.S. Geological Survey (USGS) National Seismic Hazard Model (NSHM) for California and surroundings that uses fault rupture forecasts from Uniform California Earthquake Rupture Forecast, Version 3 (UCERF-3), and ground-motion models (GMMs) from the Next Generation Attenuation-West2 Project (Bozorgnia *et al.*, 2014; Field *et al.*, 2014; Petersen *et al.*, 2014). These models include information about historical earthquakes, but do not explicitly use historically observed shaking intensities, some of which had not yet been compiled when the models were made. Because the hazard model parameters were not chosen to specifically match the past intensity data, comparing the model and historic shaking data can yield insight into the models' performance and potential improvements.

Historical accounts of shaking by witnesses of an earthquake provide the basis for assignment of seismic intensity, a measure based on the effects of shaking on man-made structures and objects within them. Intensity is therefore a good tool for characterizing the distribution of shaking at many locations, both near and far from the source. The U.S. government has been collecting first-hand accounts of shaking over the past century, and newspaper accounts go back even further (Byerly and Dyk, 1936; Toppozada and Branum, 2004). These reports have been collected in various government publications and assigned individual intensity levels using the Modified Mercalli intensity (MMI) scale (Wood and Neumann, 1931; Richter, 1958). The best practices for assigning MMI values have evolved over time (Ambraseys, 1971; Hough, 2014). Compiling a consistent record of intensity data requires consistent reinterpretation of intensity assignments, a need that motivated the development of the California Historical Intensity Mapping Project (CHIMP).

This article presents version 1 of the CHIMP (CHIMP-1) dataset of intensity values from large and moderate earthquakes expected to control the maximum shaking in California over the past 162 yr (Fig. 1; see the supplemental material available in this article). The dataset includes "Did You Feel It?" (DYFI) data for recent earthquakes (~1990 onward), intensities reinterpreted from felt reports since about 1924 when the U.S. government began collecting them systematically, intensities inferred from historical accounts, and for two earthquakes, intensities from oral history collected by the authors of this study. CHIMP contains consistently reinterpreted intensity assignments for each individual earthquake and a map of historically observed maximum shaking intensities. We also present CHIMP version 1A (CHIMP-1A), which includes smoothed data for the three largest and oldest earthquakes in the dataset.

We compare the CHIMP datasets to the 2018 USGS NSHM (Rukstales and Petersen, 2019) to explore various aspects of how the models perform and possible approaches to improving

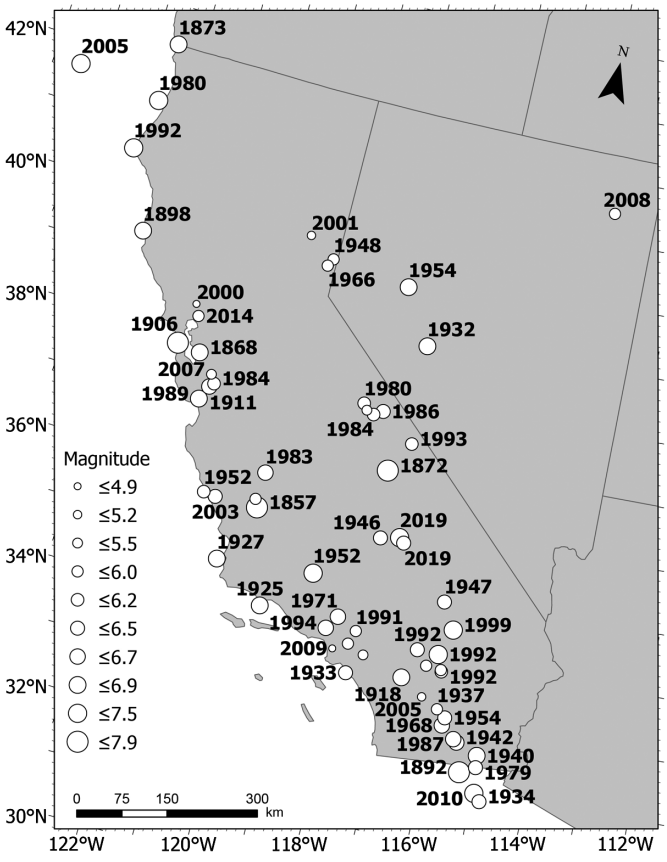

Figure 1. Map of earthquakes included in the maximum shaking dataset, labeled by year and scaled by moment magnitude.

them. We use different metrics to explore how model performance varies in space and time. The shaking dataset has been developed specifically for comparison with hazard models, so features intrinsic to the historical data are identified and addressed in the performance assessment.

Historical Macroseismic Intensities

CHIMP builds on past studies of historical and instrumental earthquakes in California, including seminal work by Toppozada *et al.* (1981) and Boatwright and Bundock (2005, 2008), long-running postcard questionnaire programs by government agencies (e.g., Byerly and Dyk, 1936; Dewey *et al.*, 1995), and, most recently, the USGS DYFI system. DYFI collects macroseismic information over the web and assigns community internet intensity (CII) values using an algorithm (Wald *et al.*, 1999) adapted from the community decimal intensity (CDI) algorithm developed by Dengler and Dewey (1998). DYFI initially reported average CII values within postal ZIP codes and increasingly reports intensities for geocoded cells. DYFI intensities have also been collected retrospectively for some large earthquakes. For example, although DYFI debuted in 1999, DYFI intensities from the 1992 Landers

and 1994 Northridge earthquakes are available from over 300 and over 700 ZIP codes, respectively. Efforts such as CHIMP and DYFI are part of a growing effort worldwide to collect, disseminate, and analyze regional and global seismic intensity information (Stucchi *et al.*, 2004; Locati and Cassera, 2010; LaMontagne and Burke, 2019), although methods and intensity scales still vary widely (De Rubeis *et al.*, 2019). Some intensity data for past earthquakes worldwide (1638–1985) are available via the National Oceanic and Atmospheric Administration (NOAA) earthquake intensity database (see Data and Resources). For U.S. earthquakes, this database includes intensities from the United States earthquakes report series, with some modification. The California Geological Survey's historic earthquake online database (see Data and Resources) contains felt report summaries and intensity assignments for earthquakes in California and adjoining regions, drawing from earlier work by Toppozada *et al.* (1981) and others.

Combining and consistently interpreting intensity data from different sources and times requires care, for several reasons:

1. A large body of work, starting from Ambraseys (1971, 1983), shows that, although environmental effects such as landslides and liquefaction potentially provide useful information about ground motions, without careful consideration they are not reliable indicators of peak ground acceleration (PGA), as was assumed in the initial formulation of the MMI scale. For this reason, we exclude reports that mention only environmental effects in CHIMP.
2. The assignment of intensity based on a written account is inherently subjective, and the degree of conservatism varies among researchers.
3. The NOAA database is especially problematic, with originally assigned intensities of 1–3 (or 1–4) generally listed as MMI 3 (or MMI 4).
4. Assessment of intensity values evolved significantly from the early twentieth century onward. Early assessments gave more weight to subjective human response, for example, if people reported being frightened by shaking. Moreover, accepted practice was to assign a higher intensity if a single indicator corresponded to intensity level N, even if overall effects suggested $N - 1$ (or even $N - 2$) (Dewey, personal comm., 2018).

The issue of subjectivity in (2) has been addressed in recent work showing that differences in intensity estimates decrease as the level of expertise increases, and that assigning intensities through numeric forms, such as DYFI and equivalent systems, reduces the difference in estimates from historical documents (Sira *et al.*, 2019). MMI assignments tend to vary by ± 1 units between researchers (Hough and Page, 2011; Salditch *et al.*, 2018). Reinterpretations in CHIMP are given as the mean of two independent assignments to help reduce uncertainty. For the 1857 Ft. Tejon earthquake, the oldest earthquake in CHIMP and hence the most uncertain, four independent intensity assignments were made and averaged. Although it remains difficult if not impossible to formally assign uncertainties to intensity assignments, ± 1 unit is a reasonable estimate of uncertainty for an individual assignment (see Hough and Page, 2011).

Factor (4) is the most potentially problematic because the assessment of numerical intensities evolved over time from assigning intensities based on the most dramatic reported effects, to assigning representative values for an area. The DYFI algorithm determines representative intensities by averaging responses to each individual question on the questionnaire and calculating a weighted average intensity within a ZIP code or geocoded cell (Wald *et al.*, 1999).

Intensity values have also been colored by a reporting bias, whereby the media focuses on the most dramatic damage rather than the overall level of damage (Richter, 1958; Hough, 2013, 2014). This bias tends to be strongest when written reports are short, so more extensive accounts typically provide a better sense of overall effects. Thus, the bias tends to be more significant for earlier earthquakes, for which newspaper and other accounts are often especially fragmentary, than for more recent events. A reporting bias still commonly exists in coverage of even recent earthquakes, such that care is needed if intensities are determined from available media accounts.

Hough (2014) proposed a correction curve approach to convert conventional intensities to DYFI intensities. This approach was developed based on older published conventional intensity values. Over time, as noted, the assignment of traditional intensities has evolved toward a more conservative approach, such that subjective assignments better align with DYFI values. The reinterpretations undertaken as part of the CHIMP effort were largely motivated by the need to address this issue for earthquakes for which a modern reinterpretation had not been done by studies such as Boatwright and Bundock (2005, 2008). The CHIMP dataset itself provides an argument against applying a correction curve approach; for example, the extensive set of intensities for the 1906 earthquake (Boatwright and Bundock, 2005) includes 23 MMI 1 s (not felt) and 93 2.0–2.5 values, out of a total of 684 locations. The correction curve would reduce these low values, in many cases to intensities below 2, which would be inconsistent with DYFI assignments. We recognize that subjectively determined intensities will never be entirely consistent with those determined by an algorithm such as that used by the DYFI system, a fundamental limitation of the CHIMP dataset. Nonetheless, the value of creating a complete dataset outweighs the unavoidable limitations associated with use of different intensity types.

The previous limitations of early intensity data, some of which persist with modern studies, caused some researchers to denigrate the value of such "unscientific" data (Hough, 2000). Characterization of shaking severity by a single number has limitations, because the character of ground motions depends on duration and frequency content as well as peak velocity or peak acceleration. Overwhelmingly, shaking intensities are controlled by frequencies between 1 and 8 Hz (Trifunac and Brady, 1975; Sokolov and Cherov, 1998). With rare exception, intensity data thus cannot constrain long-period (<1 Hz) shaking effects (Hanks and Johnston, 1992; Hough, 2014). Accordingly, intensity magnitudes based on macroseismic data alone provide an estimate of energy magnitude, but only limited constraint on moment magnitude (Hough, 2014). By the same token, historically observed intensities are not expected to provide much constraint on the levels of long-period ground motions that will potentially affect large modern structures.

Intensity data do, however, provide an integrated measure of shaking over the main frequency range of engineering concern, and are increasingly recognized to be of great value, if they are interpreted carefully with an appreciation of limitations. To cover the range of perceptible earthquake ground motions with a 10-step intensity scale, each step in intensity must correspond robustly to a factor of approximately 2 in PGA (Hough, 2000). DYFI and other internet-based systems demonstrate that consistently interpreted intensity data provide surprisingly reliable indicators of ground motions (e.g., Atkinson and Wald, 2007; Worden *et al.*, 2012) and can provide important insights about earthquake source parameters, site, and path effects. USGS ShakeMaps—which depict intensity values calculated from instrumental recordings—rely on DYFI data to flesh out shaking distributions, in particular, in areas where instrumentation is sparse. As we will discuss shortly, consistently interpreted intensity data can be used for assessing the performance of PSHA models, in particular, where PSHA models are cast in terms of intensity (e.g., Brooks *et al.*, 2016, 2017, 2018, 2019).

CHIMP Dataset

The CHIMP dataset encompasses 62 earthquakes occurring between 1857 and 2019 (Fig. 1). Because reconsideration of all macroseismic data for California earthquakes would be prohibitively time-consuming, we focus primarily on the set of $M_w > 6$ earthquakes since 1857 that we expect will control maximum observed shaking throughout the state (Table 1). We make exceptions on the magnitude cutoff for some recent, smaller events that may control the maximum shaking in their epicentral area. Hence the data go back 162 yr but are not uniformly complete—the coverage and number of contributing reports are heavily weighted toward the more recent DYFI end of the dataset. Earlier earthquakes included are fewer and larger in magnitude. Including moderate earthquakes is only possible for the recent DYFI-era events because of the better spatial coverage overall and sensitivity to smaller events. The 1987 Whittier Narrows earthquake is the earliest in the dataset with magnitude lower than M_w 6, followed by Sierra Madre in 1991. The final six earthquakes smaller than M_w 6 occur during 2000–2009. Our selection of events for the CHIMP dataset was subjective, so we may have missed other moderate events that controlled intensities in their local area.

CHIMP includes the 15 April 1898 Mendocino earthquake, the only pre-1900 earthquake in California for which an instrumental magnitude has been determined (Abe, 1994), and several large historical earthquakes: 1857 Fort Tejon, 1868 Hayward, 1872 Owens Valley, 1873 California-Oregon border, and 1892 Laguna Salada. Intensity values have been reinterpreted for these earthquakes (Hough and Elliot, 2004; Boatwright and Bundock, 2008; Hough and Hutton, 2008; Brocher, personal comm., 2019, unpublished USGS Open-File Report). Intensity distributions have been revisited for several of the largest twentieth century earthquakes, including 1906 San Francisco (Boatwright and Bundock, 2008), 1925 Santa Barbara (Hough and Martin, 2018), 1933 Long Beach (Hough and Graves, 2020), and 1952 Kern County (Salditch *et al.*, 2018). DYFI data are available for 25 earthquakes (Table 1), including the 1989 Loma Prieta earthquake.

We reconsidered the intensity distributions for 29 earthquakes. For most of these earthquakes, summaries of postcard questionnaires are available in reports published by the U.S. Weather Bureau, Coast and Geodetic Survey, and later by the USGS. The primary contributors are "United States Earthquakes" and "Abstracts of Earthquake Reports." The latter represents an intermediate phase between the primary sources, the original questionnaires and press reports, and the secondary summaries given in "United States Earthquakes." The abstracts document effects not commonly reported in the former such as effects observed in low-intensity communities for earthquakes that produced high intensities elsewhere (Dewey, personal comm., 2019). The abstracts from 1929 to 1973 are available as scanned copies online through resources such as the Hathi Trust Digital Library (see Data and Resources), and most exist as paper or digital copies in the archives of the National Earthquake Information Center in Golden, Colorado. The summaries in these reports required transcription and reinterpretation of the originally assigned intensities. These data sources were supplemented by newspaper accounts (e.g., Toppozada *et al.*, 1981), which, in some cases, were augmented with accounts gleaned from searchable online newspaper repositories and other sources, which again required transcription and intensity assignments.

Full characterization of the shaking distribution of the 29 historical and early instrumental earthquakes would be valuable for myriad reasons, but prohibitively time-consuming. Because our work focuses on maximum intensities observed throughout California, and the handful of largest earthquakes (1857,

TABLE 1
Earthquakes in the CHIMP-1 Dataset

CHIMP Label	Date (yyyy/mm/dd)	Location	Moment Magnitude	Epicenter Longitude	Epicenter Latitude	Number of IDP
1857	1857/01/09	Fort Tejon*	7.9	−120.300	35.700	71
1868	1868/10/21	Hayward	6.8	−122.100	37.700	162
1872	1872/03/26	Owens Valley	7.8	−118.100	36.700	147
1873	1873/11/23	California-Oregon border	6.9	−124.200	42.000	120
1892	1892/02/23	Laguna Salada	7.8	−115.630	32.550	37
1898	1898/04/15	Mendocino*	6.9	−123.800	39.200	33
1906	1906/04/18	San Francisco	7.9	−122.550	37.750	684
1911	1911/07/01	South Bay*	6.6	−121.750	37.250	38
1918	1918/04/21	San Jacinto*	6.8	−117.000	33.750	141
1925	1925/06/29	Santa Barbara	6.8	−119.800	34.300	237
1927a	1927/11/18	Bishop*	5.5	−118.750	37.500	8
1927b	1927/11/04	Lompoc*	6.9	−120.774	34.813	160
1932	1932/12/20	Nevada*	6.8	−117.910	38.631	3
1933	1933/03/10	Long Beach	6.4	−118.000	33.631	223
1934	1934/12/31	Colorado River*	6.4	−115.176	32.180	39
1937	1937/03/25	Buck Ridge*	6.0	−116.250	33.400	65
1940	1940/05/18	Imperial Valley*	6.9	−115.381	32.844	201
1942	1942/10/21	Fish Creek*	6.6	−115.785	32.975	143
1946	1946/03/15	Walker Pass*	6.3	−117.944	35.702	187
1947	1947/04/10	Manix*	6.5	−116.532	34.983	220
1948a	1948/12/04	Desert Hot Springs*	6.0	−116.331	33.983	275
1948b	1948/12/29	Northeast California*	6.0	−120.080	39.550	81
1952a	1952/07/21	Kern County	7.5	−118.998	34.958	1,062
1952b	1952/11/22	Bryson*	6.2	−121.328	35.723	86
1954a	1954/03/19	San Jacinto	6.4	−116.081	33.299	149
1954b	1954/07/06	Nevada*	6.8	−118.530	39.420	170
1966	1966/09/12	Northern California-Truckee*	5.9	−120.160	39.438	62
1968	1968/04/08	Borrego*	6.6	−116.103	33.180	306
1971	1971/02/09	Sylmar	6.6	−118.370	34.416	115
1979	1979/10/15	Imperial Valley	6.4	−115.359	32.667	19
1980a	1980/05/25	Mammoth*	6.1	−118.831	37.590	23
1980b	1980/11/08	Eureka*	7.2	−124.253	41.117	18
1983	1983/05/02	Coalinga	6.7	−120.312	36.232	50

Number of intensity data points (IDPs) for "Did You Feel It?" (DYFI) are the total number of gridded data points in the file provided, which is not the same as the number of individual reports—those can be found on the respective U.S. Geological Survey (USGS) event pages whose URLs are in the Appendix. Magnitudes and epicenters are also from the respective USGS event page. CHIMP, California Historical Intensity Mapping Project.
*Events reinterpreted in this study.
(Continued next page.)

TABLE 1 (continued)
Earthquakes in the CHIMP-1 Dataset

CHIMP Label	Date (yyyy/mm/dd)	Location	Moment Magnitude	Epicenter Longitude	Epicenter Latitude	Number of IDP
1984a	1984/04/24	Morgan Hill*	6.2	−121.679	37.310	194
1984b	1984/11/23	Round Valley*	6.1	−118.607	37.460	33
1986a	1986/07/08	North Palm Springs	6.0	−116.608	33.999	40
1986b	1986/07/21	Chalfant Valley*	6.4	−118.443	37.538	32
1987a	1987/10/01	Whittier Narrows*	5.9	−118.079	34.061	127
1987b	1987/11/24	Superstition Hills*	6.6	−115.852	33.015	14
1989	1989/10/17	Loma Prieta	6.9	−121.880	37.036	27
1991	1991/06/28	Sierra Madre	5.8	−117.993	34.270	31
1992a	1992/04/22	Joshua Tree*	6.1	−116.317	33.960	59
1992b	1992/04/25	Rio Dell	7.2	−124.449	40.335	1
1992c	1992/06/28	Landers	7.3	−116.437	34.200	157
1992d	1992/06/28	Big Bear	6.3	−116.827	34.203	310
1993	1993/05/17	Big Pine*	6.1	−117.774	37.165	5
1994	1994/01/17	Northridge	6.7	−118.537	34.213	351
1999	1999/10/16	Hector Mine	7.1	−116.265	34.603	721
2000	2000/09/03	Northern California	4.9	−122.413	38.379	213
2001	2001/08/10	Northern California	5.2	−120.617	39.811	295
2003	2003/12/22	San Simeon	6.5	−121.100	35.700	1,070
2004	2004/09/28	Shandon	6.0	−120.366	35.818	565
2005a	2005/06/12	Anza	5.2	−116.567	33.532	489
2005b	2005/06/14	Mendocino	7.2	−125.953	41.292	388
2007	2007/10/30	San Francisco Bay	5.5	−121.774	37.434	5,775
2008a	2008/02/21	Wells, Nevada	5.9	−114.872	41.144	997
2008b	2008/07/29	Chino Hills	5.4	−117.766	33.949	7,289
2009	2009/05/17	Lennox	4.7	−118.336	33.938	589
2010	2010/04/04	Baja	7.2	−115.295	32.286	4,170
2014	2014/08/24	Napa	6.0	−122.312	38.215	3,938
2019a	2019/07/04	Ridgecrest	6.4	−117.504	35.705	1,242
2019b	2019/07/06	Ridgecrest	7.1	−117.599	35.770	12,045

Number of intensity data points (IDPs) for "Did You Feel It?" (DYFI) are the total number of gridded data points in the file provided, which is not the same as the number of individual reports—those can be found on the respective U.S. Geological Survey (USGS) event pages whose URLs are in the Appendix. Magnitudes and epicenters are also from the respective USGS event page. CHIMP, California Historical Intensity Mapping Project.
*Events reinterpreted in this study.

1872, 1906, etc.) effectively establishes MMI 4 as a lower bound on maximum observed intensities anywhere in the state, we reviewed available sources to compile information for earthquakes for which MMI 5 or greater shaking was observed, although still including lower intensities values were described by full felt reports. If a report contained just a location and MMI assignment without a report, we did not include it in CHIMP. For some earthquakes, the few reports of moderately high intensities are straightforward to compile. For more widely felt earthquakes, however, such as the 1918 San Jacinto and 1927 Lompoc

earthquakes, macroseismic information is more plentiful and required more time to review.

Reinterpretation methods for CHIMP differ from the original MMI classification by Wood and Neumann (1931), updated by Richter (1958), in the following ways:

1. Taking the conservative approach of Ambraseys (1983) to secondary geologic indicators.
2. Using quartile decimal (e.g., 5.25) rather than integer values of MMI when some but not all indicators for a certain shaking level are present.
3. Giving more weight to the disturbance of objects than to subjective human reactions. Originally, the disturbance of objects and personal reactions was given about equal weight for the lower intensities (U.S. Department of Commerce [USDOC], Environmental Services Administration, Coast and Geodetic Survey, 1966).
4. Assigning MMI 5 only when accounts describe toppling of small objects, a key objective indicator for this intensity level (Richter, 1958).
5. Differentiating between MMI 2 (felt or felt by few) and 3 (felt by many or hanging objects swing), rather than characterizing weakly felt intensities as MMI 1–3, as was the earlier practice in the U.S. Earthquake report series.
6. Assigning MMI 1 for sites "reported not felt", following the recommendation of Ambraseys (1983) and generally in line with current practice (including DYFI).

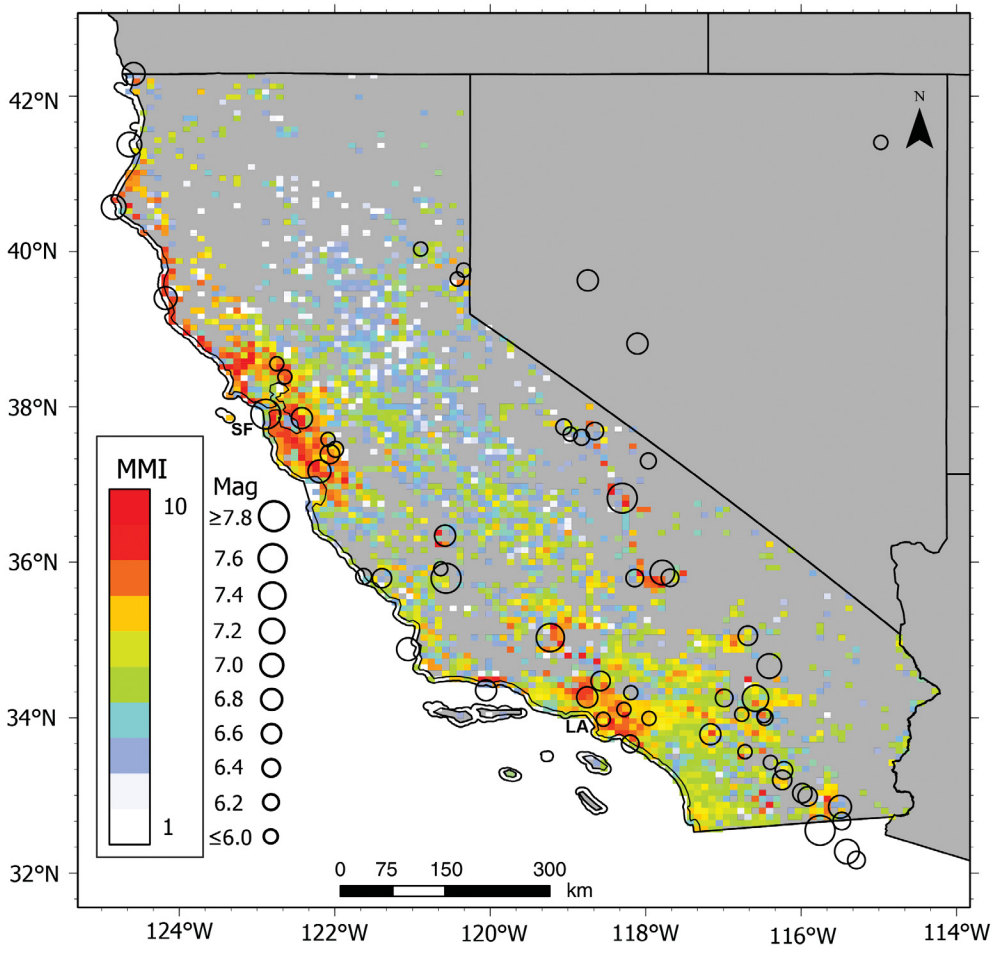

Figure 2. California Historical Intensity Mapping Project-1 (CHIMP-1) map of maximum observed intensities in California. Open circles represent the causative epicenters, scaled by moment magnitude. LA, Los Angeles; SF, San Francisco.

CHIMP-1 includes 46,502 intensity data points (IDPs). The final dataset reports MMI to one decimal place, rounded up, in keeping with DYFI standards. The vast majority (88% or 40,887 IDPs) of those points are from DYFI, although they account for just 12% of the time period covered and 48% of the earthquakes. Historical IDPs, of which 5615 were included in this study, represent the remaining 12%. Of those 5615 individual historical felt reports, 2794 were reinterpreted by two authors of this study. We give the mean value of the two assignments in CHIMP-1. These assignments matched 38% of the time, 72% were within ±0.5 MMI units, and 94% were within ±1 unit. Figure 2 shows the maximum intensities observed in 10 × 10-km grid cells for which one or more intensity value is included in the CHIMP-1 dataset.

CHIMP can give different levels of detail depending on user needs. So that this information can be reinterpreted later, we provide in the repository (see Data and Resources) the full felt report for each reinterpreted intensity assignment and give the two assignments that go into the mean. The dataset provides files of longitude, latitude, and MMI or CDI (see supplemental material).

Oral History Collection

In the 1990s, during the transition from postcards to the DYFI system, which went online at the turn of the millennium, the

USGS collected intensity information for only the most damaging events, for example, the M_w 7+ 1992 Landers and 1994 Northridge earthquakes. As a result, two minimally damaging M_w 6.1 earthquakes, those of 1992 in Joshua Tree and 1993 near Big Pine, did not have intensity data compiled in the usual systematic way. Because these earthquakes predate the onset of the DYFI system but postdate the collection of postal questionnaires, there is a lack of publicly available intensity information for them.

To correct for this data gap, two authors of this study, Salditch and Gallahue, gathered intensity information by collecting oral history reports of shaking by local witnesses and newspaper accounts. This field work resulted in 13 in-person interviews with residents, several more questionnaire responses (designed after DYFI surveys), and dozens of newspaper reports giving 61 new IDPs for Joshua Tree and five for Big Pine. The latter location is much more sparsely populated and so provided fewer new IDPs. The new data are especially useful for the Owens Valley portion of the CHIMP map, which has areas without reports in the Sierra Nevada and Basin and Range regions.

Most interviewees learned of the study from advertisements. Local media, including radio station Z 107.7 and The Hi-Desert Star newspaper, covering the high desert around Joshua Tree, as well as The Inyo Register and 100.7 KIBS, covering the Owens Valley and Eastern Sierra Nevada region around Big Pine, ran announcements and articles about the project. The Yucca Valley, Palm Springs, and Joshua Tree Public Libraries generously hosted drop-in interviews. Newspaper accounts came from the archives of the Desert Sun on microfilm at the Palm Springs Public Library, and bound volumes of the Hi-Desert Star available at their offices in Yucca Valley.

This oral history collection could be replicated for other moderate earthquakes occurring during the lifetime of current generations. Memories are the most evanescent source of intensity data available—and therefore most at risk of being lost.

Retrospective Assessment of Hazard Maps with Historical Data

We assessed the performance of two of the 2018 USGS NSHM models following approaches we have used elsewhere (Brooks *et al.*, 2016, 2017, 2018, 2019). These involve comparing the hazard models' forecasts to the historic shaking data using different performance metrics to assess various aspects of how the models performed. Assessing a model's performance can be used to improve future models and the models used to generate them, in that the factors contributing to the model's performance (source locations, maximum magnitude, GMM, etc.) can be evaluated. These comparisons can give insight into current models and possible approaches to improving them.

At any point on a PSHA map, the probability p that during t years of observations shaking will exceed the value shown in a model with a T year return period is assumed to be described by a negative exponential distribution, $p = 1 - \exp(-t/T)$ (Cornell, 1968; Field, 2010). We consider two of the 2018 USGS models, both with $t = 50$ yr, but differing in return periods. For one, $p = 0.1$ or 10%, giving $T = 475$ yr. For the other, $p = 0.02$ or 2%, giving $T = 2475$ yr. Equivalently, during t years on average 10% and 2% of the sites should experience shaking greater than shown on the map with return period $T = 475$ and 2475 yr, respectively. Utilizing the ergodic assumption, which states that a system has the same behavior over time as it does over space, we assume that p also reflects the fraction of sites where observed shaking exceeds the modeled value. This approach, introduced by Ward (1995) and used in many subsequent analyses (e.g., Albarello and D'Amico, 2008; Fujiwara *et al.*, 2009; Miyazawa and Mori, 2009; Stirling and Gerstenberger, 2010; Nekrasova *et al.*, 2014; Tasan *et al.*, 2014; Mak and Schorlemmer, 2016) considers many sites to avoid the difficulty that large motions at any given site are infrequent.

Comparison between the predicted and observed maximum shaking allows the calculation of performance metrics (Stein *et al.*, 2015). Assuming that the frequency sample estimates correspond to the probabilities p, the fractional exceedance metric $M0 = |f - p|$ is the absolute value of the difference between the observed fraction of points f above the diagonal line in Figure 3a—sites where the largest observed shaking exceeds prediction—to the fraction p expected. The remaining sites plot below the line, because the model predicted shaking higher than observed. As the ratio of the observation time to the return period (t/T) increases, p should also increase because an increasing fraction of the area will have experienced larger earthquakes and thus higher shaking.

Hence the shaking shown in a model with $T = 500$ yr should be exceeded at 22% of the sites in 125 yr $(t/T = 0.25)$, 39% of the sites in 250 yr $(t/T = 0.5)$, and 63% of the sites in 500 yr $(t/T = 1)$ (Fig. 3b). M0, which is based on the definition of PSHA, has the limitation that it counts exceedances as binary—shaking at a site either exceeded the modeled value or did not. Hence a model with exceedances at exactly as many sites as predicted (M0 = 0) could significantly overpredict or underpredict the magnitude of shaking (Stein *et al.*, 2015). Because hazard model assessment is a relatively new enterprise, and only a few cases have so far been assessed, there is currently no threshold defined for a "good" score on the M0 metric.

The related question of how well a model could realistically be expected to describe observations, given the uncertainties in model parameters and variability in earthquake occurrence, has been investigated via numerical simulations of ground motion by assuming earthquakes occur randomly within a study area (Vanneste *et al.*, 2018). As shown by Vanneste *et al.* (2018) (Fig. 3c), some places experience shaking higher than on the hazard map, whereas others experience shaking lower

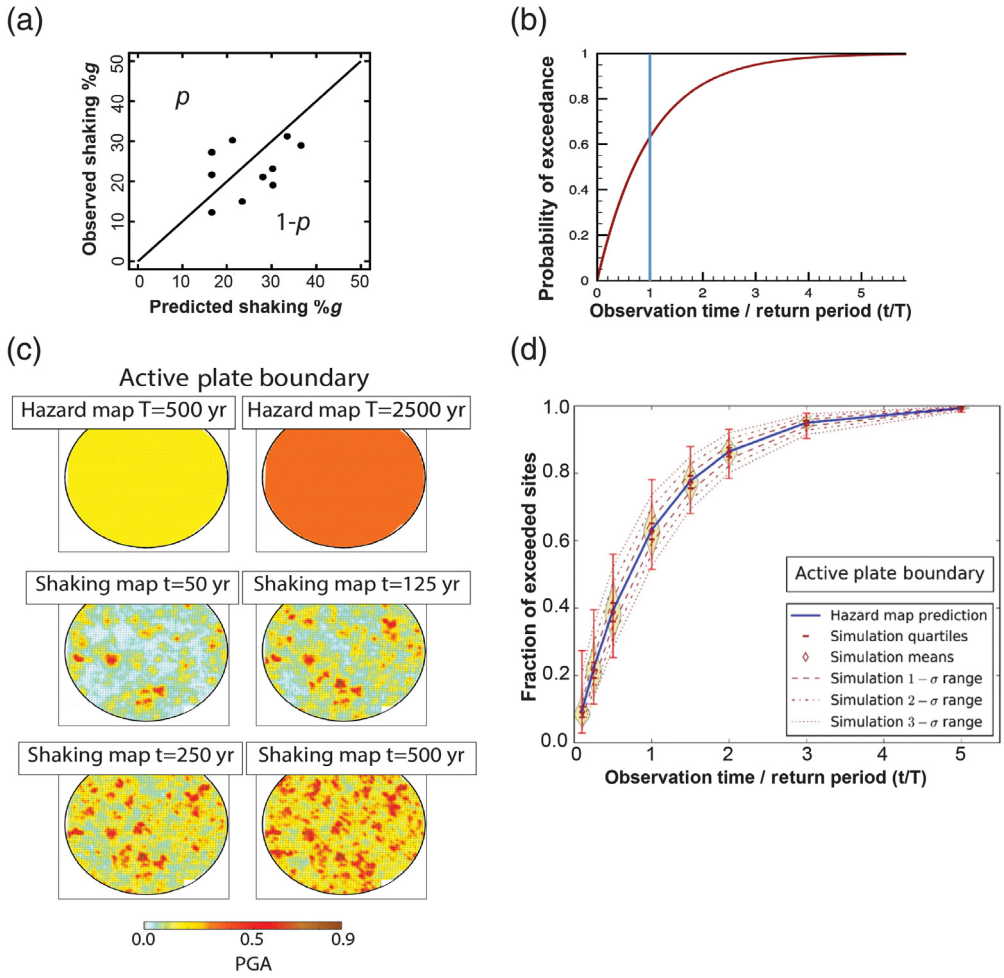

Figure 3. (a) Schematic observed versus predicted plot showing p, the predicted fraction of sites at which the largest shaking exceeded the mapped values. (b) p should increase as the ratio of the observation time t to the model's return period T increases. (c) Numerical simulation of maximum shaking over time assuming earthquakes occur randomly (with a spatially uniform distribution) within the model area with seismicity comparable to active plate boundaries. Top row: Hazard models for return periods of 500 and 2500 yr. The models and associated maps are uniform across the area, because the expected level of shaking is the same. Middle and bottom rows: Maps of maximum shaking at each point after observation times of 50, 125, 250, and 500 yr, for one simulation. (d) Comparison between predicted and "observed" fractions of sites where maximum shaking exceeds that predicted by the 500 yr hazard model as a function of observation time. Scatter decreases for longer simulations. (c,d) Modified from Vanneste *et al.* (2018).

than shown on the map. For example, after only 50 yr some sites experienced shaking stronger than shown on the $T = 2500$-yr map. When large earthquakes happen, shaking often exceeds that shown on the hazard maps. As the observation time increases, the fraction of sites exceeding the mapped value increases. An ensemble of simulations yields shaking distributions whose mean is consistent with the model, but individual shaking histories show large scatter (Fig. 3d). The scatter decreases for longer simulations (increasing t/T), because as observation time increases, the largest earthquakes and resulting shaking are increasingly likely to have occurred.

We also define a squared-misfit metric $M1 = (1/N)\Sigma(s_i - x_i)^2$, in which x_i and s_i are the maximum observed shaking and predicted shaking at each of the N sites (Stein *et al.*, 2015). Graphically, M0 reflects the fraction of sites plotting above the diagonal line in Figure 3a, whereas M1 reflects how close to the line sites plot. M0 is a metric based on the definition of PSHA, describing how well a PSHA model predicts the fractional exceedance that occurs. M1, a metric which is not based on the definition of PSHA, describes how spatially similar the observed shaking and hazard model are. M1 quantifies the comparison of maps of predicted and observed shaking, similar to a visual comparison.

The two metrics characterize different aspects of map performance. Hence, together they give a fuller picture of map performance than one measure could (Stein *et al.*, 2015; Brooks *et al.*, 2017, 2018, 2019). PSHA models do not predict specific shaking levels but rather probabilities of shaking exceedance. Thus a lower M1 score does not necessarily mean that the map has performed better—as measured by M0—than one with a higher M1 score. Situations may arise for which decreasing M1 may produce larger M0 scores. When comparing metric scores, it is important to be aware of potential trade-offs in these metrics.

Because hazard model assessment is a relatively new enterprise and only a few cases have so far been assessed, there is currently no threshold defined for a "good" score on either the M0 or M1 metrics. These metrics, and others that can be used (Stein *et al.*, 2015), are thus most useful as a tool to compare maps to observations and assess the performance of different maps. Future consideration of many maps and numerical simulations should provide improved understanding of the meaning of high- and low-metric scores (Vanneste *et al.*, 2018; Brooks *et al.*, 2019).

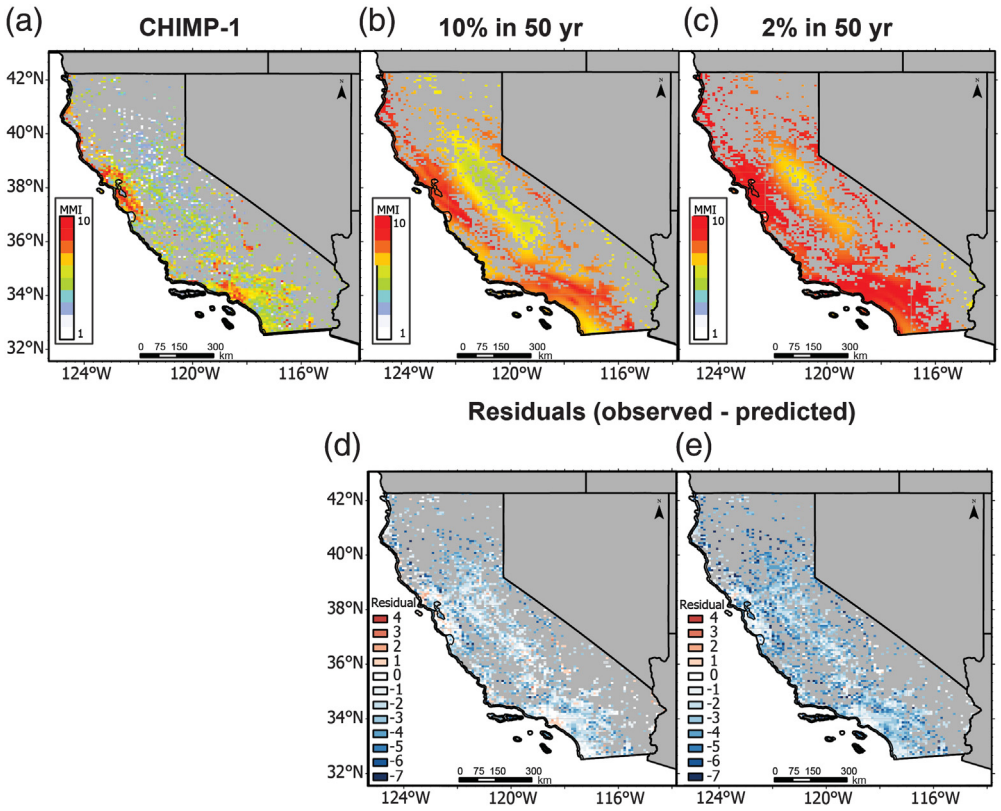

Figure 4. Comparison of (a) CHIMP-1 maximum shaking dataset to the 2018 U.S. Geological Survey (USGS) hazard models, for (b) 475-yr return period model with 10% chance of exceedance in 50 yr and (c) 2475-yr return period model with a 2% chance of exceedance in 50 yr. (d,e) Residuals (dataset—model values) relative to both models.

Comparison of Maps and Dataset for California

Using the performance metrics described previously, we compare the CHIMP-1 dataset with the 2018 USGS time-independent seismic hazard models for California (Rukstales and Petersen, 2019). We limit our assessment to the state of California because the rupture forecast model used there (Field *et al.*, 2014) differs from the models used in surrounding states such as Nevada and Arizona (Working Group on California Earthquake Probabilities [WGCEP], 2013). Figure 2 shows the maximum shaking values in CHIMP-1, which are sorted into 10 × 10-km grid cells, giving 38% spatial coverage of California. For the comparison, we use two of the 2018 hazard models that assume a reference site condition to be National Earthquake Hazards Reduction Program site class boundary B/C, corresponding to firm rock, very dense soil, and soft rock, and a $V_{S30} = 760$ m/s (Petersen *et al.*, 2020). One has $T = 475$ yr and hence 10% probability of exceedance in 50 yr, and one has $T = 2475$ yr corresponding to a 2% probability of exceedance in 50 yr. USGS PSHA models include exceedance forecasts for a number of ground-motion parameters. Models of hazard in terms of MMI are available, calculated using ground-motion intensity conversion equations (GMICEs) of Worden *et al.* (2012) to convert PGA to MMI intensity. Mapped hazard values (Fig. 4) are shown for points at which CHIMP-1 has values.

In general, the shaking data are similar spatially in trends to the models. Some fault segments have not experienced a large earthquake since ~1857, notably the southern San Andreas. Such effects should ideally be accounted for in the model via the PSHA algorithm. For 162 yr of observations, $p = 1 - \exp(-162/475) = 29\%$ of the sites should have experienced maximum shaking greater than that shown in the 475-yr model. A smaller fraction, $p = 1 - \exp(-162/2475) = 6\%$, should have experienced maximum shaking greater than the values in the 2475-yr model, because 162 yr is a much smaller fraction (6% vs. 34%) of the model return period. The longer return period model predicts higher shaking, because the largest earthquakes and shaking are more likely to occur during the longer return period. However, comparison of the largest observed shaking at sites to predictions of the hazard models shows observed maximum shaking (Fig. 4) and fractional exceedances (Fig. 5) less than predicted. The residual plots of (observed–predicted) intensity in Figure 4 show the distribution of discrepancies. The 475-yr return period model has a mean residual of −2.3 MMI, and the 2475-yr return period model has a mean residual of −3.4 MMI. Thus, on average, the observations are approximately 3 MMI units lower than the hazard model values. The M0 score for the 475-yr model, in which f is more than four times smaller than p, is 0.2248 (Table 2). M0 for the 2475-yr model, in which f is an order of magnitude smaller than p, is 0.057 (again assuming that frequency sample estimates correspond to probabilities p). The M1 score for the 475-yr model is 8.800 and 15.048 for the 2475-yr model.

Possible causes of discrepancy

This apparent overprediction of intensities by the PSHA model may arise due to biases in the dataset, the hazard model, or chance. We discuss each of these possible biases in this section.

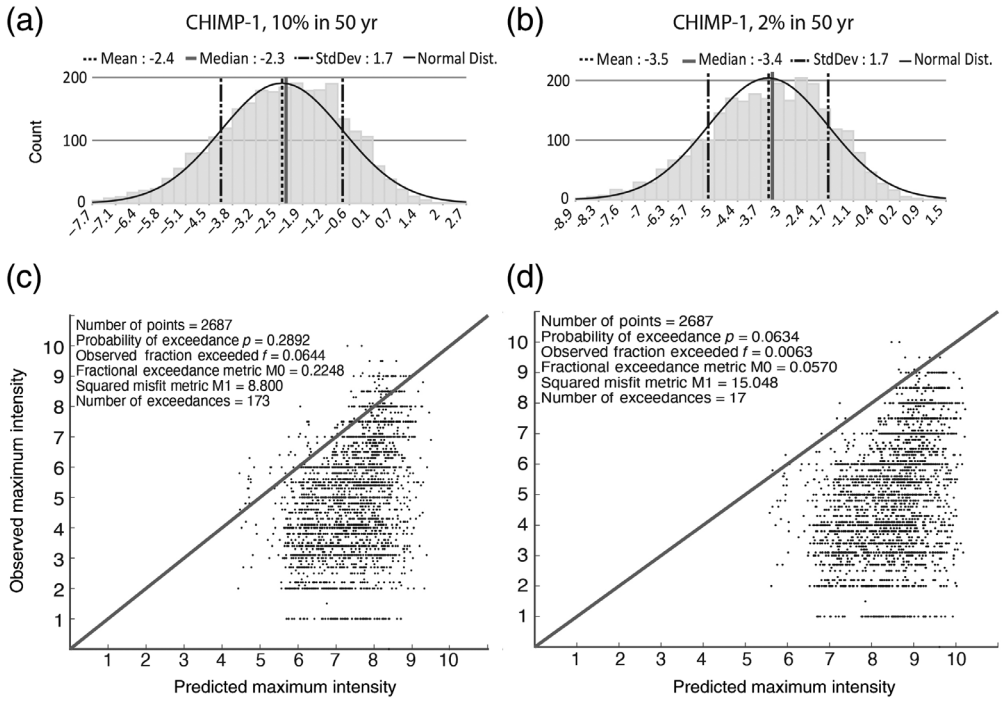

Figure 5. Histograms of residuals for CHIMP-1 points in California relative to (a) 2018 USGS hazard model for 10% in 50 yr and (b) 2018 USGS hazard model for 2% in 50 yr. Corresponding predicted versus observed plots for CHIMP-1 relative to (c) 10% in 50 yr model and (d) 2% in 50 yr model.

Data underestimation

The intensity data may be biased low. In particular, the shaking data do not capture the full extent of shaking from large historic earthquakes in the dataset, due to lack of population. Historical accounts, and hence historical intensity assignments, are biased by the locations and growth of population centers through time. Other potential issues may result from the nonuniform spatial sampling that becomes sparser the further back in time one goes: DYFI data are denser than traditional intensity data estimated from written or archival accounts, and values inferred from historical accounts are sparser still. Hence, some cells do not have a reported intensity value from the large historical events. Thus, if shaking from a more recent smaller earthquake is reported there, the maximum observed intensity in the CHIMP-1 dataset will be too low.

To address this incompleteness due to historical population distribution, we developed CHIMP-1A by adding intensity values interpolated from the existing historical IDPs for the large 1857, 1872, and 1906 events. We used the natural neighbor interpolation method in ArcGIS to produce a smoothed intensity map for each earthquake, which we then used to augment the CHIMP-1 maximum shaking dataset. This algorithm assigns a value to a query point by finding the closest subset of input samples and weighting the value proportionally to the area of overlapping Voronoi polygons (Sibson, 1981). This method produces an objective estimate, because it locally interpolates values depending only upon known data points and their spatial distribution. In contrast to traditional hand-drawn isoseismal maps, it is simple and reproducible (Sirovich *et al.*, 2002). However, because hand-drawn isoseismal maps were created using expert judgement and knowledge of local geology, we compared our smoothing results to them (Stover and Coffman, 1993) to confirm that results are reasonable. We added the smoothed data for each of the three historical earthquakes to the CHIMP-1 maximum observed dataset to study their effects on performance metrics (Fig. 6). The smoothing provided values almost everywhere in California, and more than doubles the total number of grid cells with observations and increases the total number of exceedances by a factor of 2–3 (Fig. 7). The residual plots now show slight underprediction in an area reaching from the Owens Valley to the Transverse Ranges, and slight overprediction in northernmost California.

The overall overprediction remains for CHIMP-1A but is slightly reduced. The mean residual for the $T = 475$-yr model improves from CHIMP-1 to CHIMP-1A by 0.4 MMI units to −1.9 MMI, and by 0.5 units to −2.9 MMI for the 2475-yr model. Interestingly, M0 remains largely unchanged, improving by 0.03 for the 475-yr model and by just 0.001 for the 2475-yr model (Table 2). However, M1 improves for both models, becoming 5.656 for the 475-yr model and 11.152 for the 2475 yr model.

A further possible data bias stems from the incompleteness of the CHIMP-1A dataset, via our assumption about which events will control maximum observed intensities. Earthquakes smaller than M_w 6 can generate locally high intensities. Because they are more common than larger earthquakes, they may control the maximum historically observed intensity in any one location. The DYFI database itself illustrates this point. Since the introduction of the DYFI system in 1999, 18 earthquakes, 11 of which are between M_w 4.7 and 5.9, generated CDI values of 6.0 or higher. (We do not consider earthquakes smaller than M_w 4.7). Of 48 earthquakes since 1999 that have generated CDI values of 5.0 or higher, 39 are smaller than M_w 6. It would be possible to assess the extent

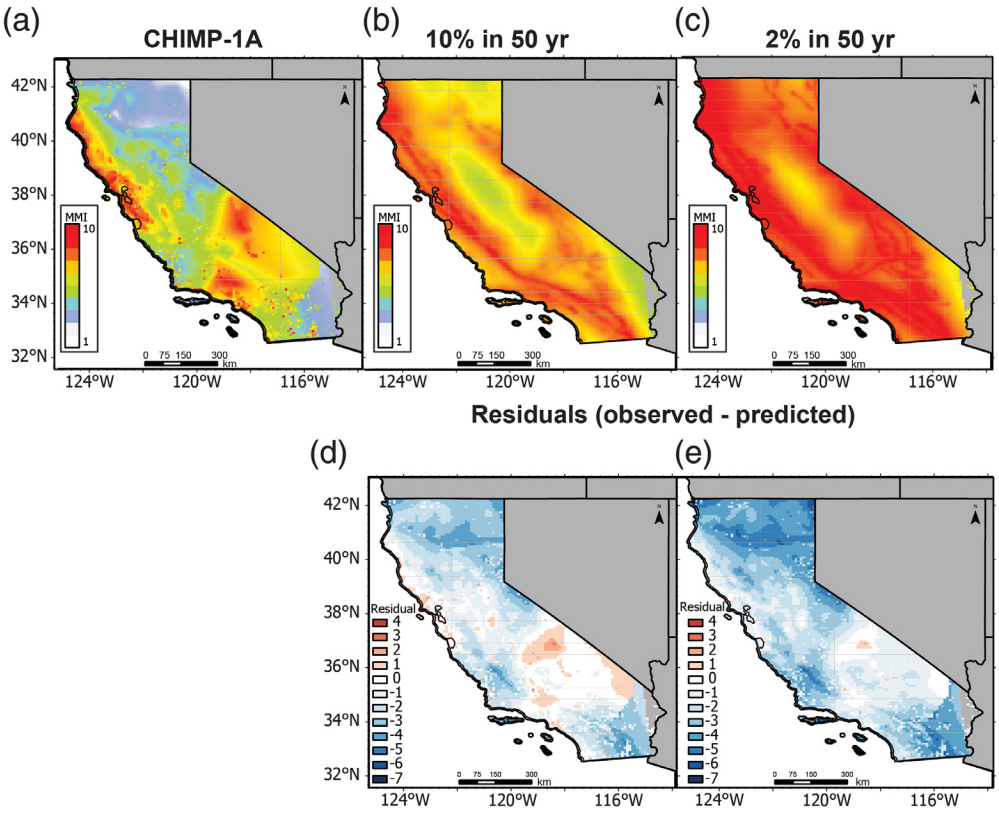

Figure 6. Comparison of (a) CHIMP-1A maximum shaking dataset, which includes smoothed data for 1857, 1872, and 1906 earthquakes, to the 2018 USGS hazard models with (b) 475-yr return period with 10% chance of exceedance in 50 yr and (c) 2475-yr return period with a 2% chance of exceedance in 50 yr. (d,e) Residuals (dataset—model values) relative to both models.

to which earthquakes smaller than M_w 6.0 contribute to the PSHA map by recalculating the map using only $M_w \geq 6.0$ sources.

Model overprediction

The model may be biased toward higher levels of shaking by various effects. One possible cause is that California seismicity during the historical period may have been lower than the long-term average, either due to random variability of earthquake occurrence or a stress shadow due to heightened activity from 1800 to 1918 that released much of the accumulated stress, though the latter is debated (Harris and Simpson, 1998; Felzer and Brodsky, 2005; Biasi and Scharer, 2019). There is compelling evidence that California has been in a state-wide lull in seismic moment release since 1910. Biasi and Scharer (2019) show that, effectively, the UCERF-3 map overpredicts the rate of surface-rupturing earthquakes on the San Andreas, San Jacinto, and Hayward fault systems. The most recent earthquake rupture forecast model for California (Field *et al.*, 2014), in fact, assumes that long-term state-wide earthquake rates are slightly higher than rates through the entire historic period. A number of other factors could bias the models to too-high values, including the normal tendency toward conservatism in engineering design (Vick, 2002).

Chance

Some of the misfit may arise purely by chance. Figure 3 illustrates this effect for an ideal hazard model. Even if all parameters are perfectly known, the variability of earthquake recurrence can give rise to a range of values. In a real case, for which parameters are not known a priori and are unlikely to be exactly estimated, the variability could be larger. If the ideal case (Vanneste *et al.*, 2018) is representative, then these simulation studies indicate that the misfit is large enough that it is unlikely to have arisen purely by chance due to variability in earthquake recurrence. Hence it likely represents, at least in part, biases in the hazard model, data, or both. Increasing the observed shaking at all sites in CHIMP-1 and CHIMP-1A by a constant shift (Fig. 8), which is possible given the uncertainty in historical intensity assignments, improves M0 and M1. Figure 8 shows the uncertainty range of the metrics for each dataset-model pair, given the inherent MMI uncertainty of ±1. The minima of M0 are within the uncertainty range of the intensities for all but the CHIMP-1 475-yr model, whereas the minima of M1 are outside that range for all models. A similar change could result from decreasing the predicted shaking or a combination of both effects.

Reconsideration of Significant Earthquakes

Although retrospective assessment of PSHA models motivated our creation of CHIMP-1, a dataset of observed intensities is potentially useful for many other purposes. Intensity datasets can be used to revisit magnitudes and locations of historic and early instrumental earthquakes, and to further explore the distribution of ground motions generated by recent as well as historic events. As an example, Figure 9 presents a shaking intensity map for the 1987 M_w 5.9 Whittier Narrows earthquake. To generate this distribution, we augmented the

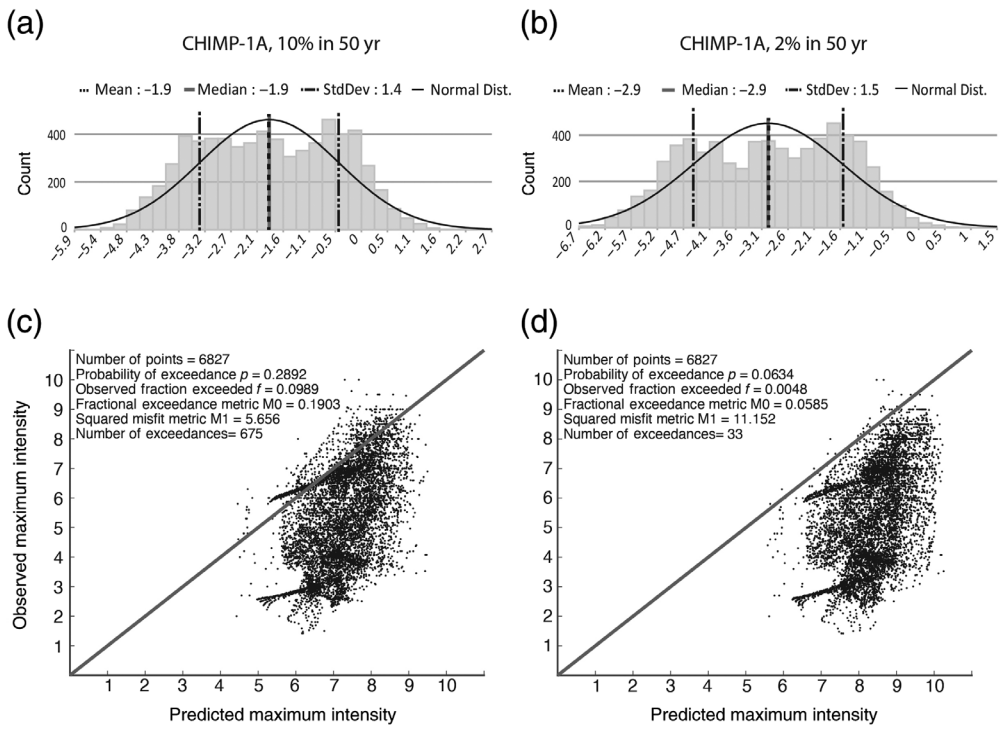

Figure 7. Comparison of CHIMP-1A, which includes smoothed data for the 1906, 1857, and 1872 earthquakes, to hazard maps. Histograms of residuals for CHIMP-1A points in California relative to (a) 2018 USGS hazard model for 10% in 50 yr and (b) 2018 USGS hazard model for 2% in 50 yr. Corresponding predicted versus observed plots for CHIMP-1A relative to (c) 10% in 50 yr model and (d) 2% in 50 yr model.

CHIMP-1 values with retroactively contributed DYFI (reported well after the time of the earthquake) intensities at all shaking levels (1-km geocoded values) and intensities estimated from available PGA data (see Data and Resources) using the Worden *et al.* (2012) GMICE. The intensity distribution is constrained by intensity values at 816 locations, revealing variability of shaking across much of the greater Los Angeles metropolitan region. There is a suggestion, for example, of elevated intensities inland of the Newport–Inglewood fault, where the Los Angeles basin deepens considerably. Similar amplification was observed in the 2008 M_w 5.4 Chino Hills earthquake (Hauksson *et al.*, 2008). The CHIMP-1 dataset will provide an opportunity to explore key questions regarding ground motions, including the variability of site response across geologically complex regions.

Discussion and Future Work

CHIMP illustrates the value of developing consistently interpreted shaking datasets and comparing them to hazard models. We see similar discrepancies, with historical intensity data much lower than hazard model predictions, for Italy and Japan (Stein *et al.*, 2015; Brooks *et al.*, 2016). This could be a coincidence or could indicate a common bias. For example, the hazard models' assumed seismicity and fault-slip rates may be too high, the GMMs may predict too-high shaking, or site effects could result in localized deamplification. We plan to investigate how well the GMMs fit the historical intensity data, as we did for the 1952 Kern County earthquake (Salditch *et al.*, 2018). Alternatively, the data may be biased low due to spatial sampling bias.

Even after including smoothed data for the largest events, CHIMP-1A may biased low by locally high intensities from moderate (M_w 4.5–6) earthquakes that are not in the CHIMP dataset. A statistical modeling approach might address this by estimating how many M_w 5s and M_w 6s are missing and calculating their expected MMI distribution using the intensity prediction equation of Atkinson *et al.* (2014). Another way

TABLE 2
Comparison of Performance Metrics between CHIMP-1 (Maximum Observations Only) and CHIMP-1A (Maximum Observations Plus Smoothed Data for 1906, 1857, and 1872)

Return Period	Model Probability Exceedance	Dataset Version	*p*	*f*	M0	M1
475 yr	10% in 50 yr	CHIMP-1	0.2892	0.0644	0.2248	8.800
		CHIMP-1A	0.2892	0.0989	0.1903	5.656
2,475 yr	2% in 50 yr	CHIMP-1	0.0634	0.0063	0.0570	15.048
		CHIMP-1A	0.0634	0.0048	0.0585	11.152

CHIMP, California Historical Intensity Mapping Project.

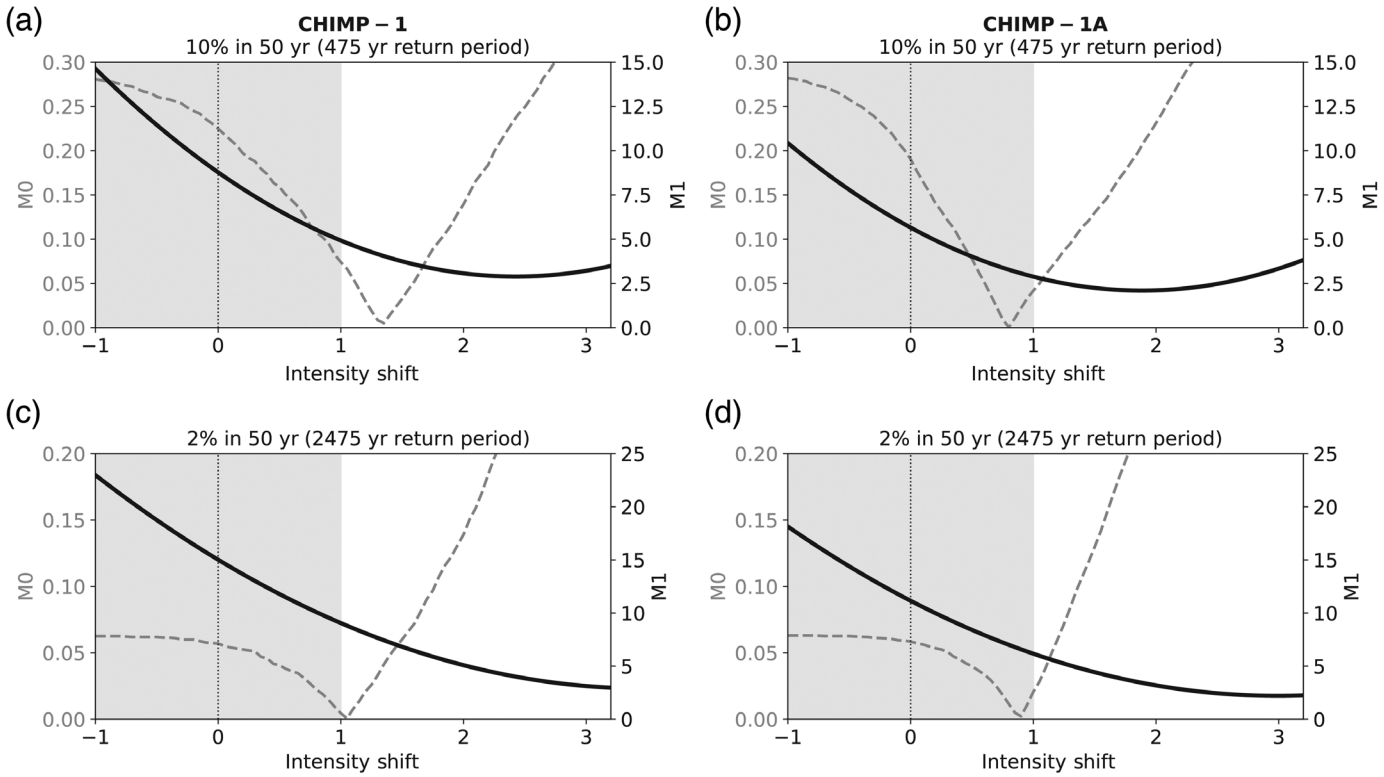

Figure 8. Effect of applying a uniform shift to either the map's predictions or the CHIMP-1 observations on performance metrics M0 (dashed gray line) and M1 (solid black line). The uncertainty for intensity values typically ranges ±1 unit (gray shaded region). Lines indicate how the metrics would change given a bulk shift. Positive intensity shift values correspond to an increase in CHIMP observations or a decrease in the map's predictions. Negative intensity shifts reflect a decrease in CHIMP observations or an increase in the map's predictions. (a) CHIMP-1 intensities compared to the 2018 USGS 10% in 50 yr model. (b) CHIMP-1A intensities compared to the 2018 USGS 10% in 50 yr model. (c) CHIMP-1 intensities compared to the 2018 USGS 2% in 50 yr model. (d) CHIMP-1A intensities compared to the 2018 USGS 2% in 50 yr model.

might be to start with the residual data, hypothesize that all of the misfit was due to small events, infer how many are needed, and see if that makes sense relative to the *b*-value curve. The DYFI dataset collected to date suggests that moderate earthquakes, that is, M_w 4.7–5.9, may be important in controlling hazard due to their prevalence.

To address possible bias introduced by the different intensity data sources, detailed archival accounts of earthquake effects could be entered into the DYFI numerically coded questionnaire and then averaged according to DYFI procedures. This would provide a direct comparison of the consistency of subjective CHIMP intensities and DYFI values. Preliminary experiments with the 1987 Whittier earthquake data do demonstrate consistency between retroactively reported DYFI, historically assigned MMI, and instrumental data (PGA recordings converted to MMI via the method of Worden *et al.*, 2012). Converting CHIMP MMI to DYFI CDI, may not, as noted, be warranted, but if further analysis suggests discrepancies, the task would generally be tractable in the United States, given the relatively small size of the historical catalog. The conversion would be less tractable in Europe, for example, where the historical record is thousands of years longer than the electronic.

Discussion of the maximum observed shaking for an area leads to questions of the second highest shaking, the third highest shaking, and so forth. This question usually stems from desire to investigate the effects of removing a single earthquake from the record. Does the performance of the model change with the removal of a single event? By how much? There is also a benefit in seeing how model performance changes over time. Our current assessments involve static datasets, using the maximum shaking over an interval. It will be useful to explore the performance of the model as time progresses and compare this to that expected as the ratio of observation time to model return period increases. We have a number of sites at which the modeled shaking has been exceeded several times. These will let us explore how well examining model exceedances at many sites compares to examining multiple exceedances at individual sites over time. In addition, we plan to conduct simulation studies to explore the variability in the expected shaking likely to have arisen due to variability in earthquake recurrence, and its consequences for how well the models' predictions should match the observed shaking.

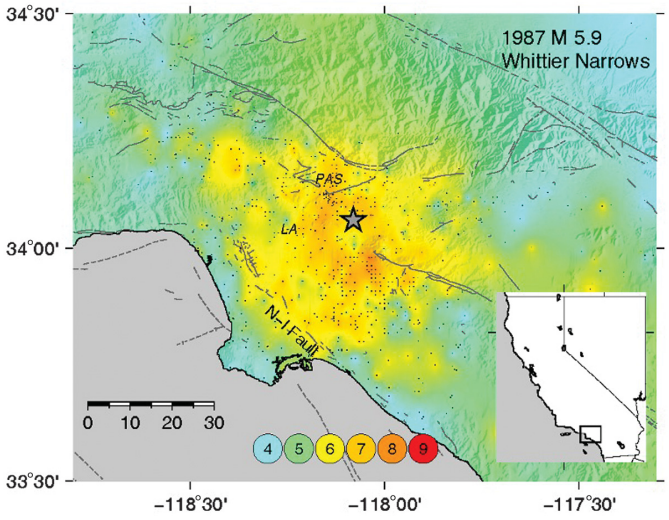

Figure 9. Shaking intensity map for the 1987 M_w 5.9 Whittier Narrows earthquake, generated using data from the CHIMP-1 dataset, augmented by retroactively contributed "Did You Feel It?" (DYFI) data at all distances and intensity values calculated from instrumentally recorded peak ground acceleration (PGA) data, converted to intensity using the ground-motion intensity conversion equation (GMICE) of Worden *et al.* (2012). Locations of Pasadena (PAS), central Los Angeles (LA), and the Newport–Inglewood fault zone (N-I fault) are indicated. Inset shows location of map within California.

Data and Resources

The California Historical Intensity Mapping Project (CHIMP) dataset is available as a supplemental material to this article. The supplement is divided into two parts: the dataset and the repository. Historic earthquakes in the dataset are .csv files formatted as LON, LAT, Modified Mercalli Intensity (MMI). "Did You Feel It?" (DYFI) earthquakes are .txt files in their original format. The repository contains full felt reports and MMI assignments for earthquakes that were reinterpreted for this study. Intensity data sources are listed in the Appendix. 2018 U.S. Geological Survey (USGS) hazard model data are available for download from Science Base at https://www.sciencebase.gov/catalog/item/5cbf47c4e4b0c3b00664fdef (last accessed December 2019). The National Oceanic and Atmospheric Administration (NOAA) intensity dataset can be accessed at https://www.ngdc.noaa.gov/hazard/intintro.shtml (last accessed February 2020). Most of the Abstracts of Earthquakes reports can be accessed online at http://www.hathitrust.org (last accessed February 2020). Strong-motion data, including recorded peak ground acceleration (PGA) values, for the 1987 Whittier Narrows earthquake are available at https://strongmotioncenter.org (last accessed February 2020).

Acknowledgments

The authors thank Arthur Zachary for help creating the California Historical Intensity Mapping Project (CHIMP) acronym. For helpful discussions and data, the authors thank Jim Dewey, David Wald, Vince Quitoriano, Tom Brocher, and Mark Petersen. The authors thank Morgan Page, Sarah Minson, Norm Abrahamson, Wolgang Bruestle, Justin Rubinstein, and an anonymous reviewer for their constructive reviews of earlier versions of this article, and Alison Bent for her stewardship of the journal. This work was supported by Northwestern University's Institute for Policy Research. Data collection for the 1992 Joshua Tree and 1993 Big Pine earthquakes was supported by an American Geophysical Union Centennial Celebrate 100 Grant.

References

Abe, K. (1994). Instrumental magnitudes of historical earthquakes, 1892 to 1898, *Bull. Seismol. Soc. Am.* **84,** 415–425.

Albarello, D., and V. D'Amico (2008). Testing probabilistic seismic hazard estimates by comparison with observations: An example in Italy, *Geophys. J. Int.* **175,** no. 3, 1088–1094.

Ambraseys, N. N. (1971). Value of historical records of earthquakes, *Nature* **232,** 375–379.

Ambraseys, N. N. (1983). Notes on historical seismicity, *Bull. Seismol. Soc. Am.* **73,** 1917–1920.

Atkinson, G. M., and D. J. Wald (2007). "Did You Feel It?" intensity data: A surprisingly good measure of earthquake ground motion, *Seismol. Res. Lett.* **78,** no. 3, 362–368.

Atkinson, G. M., C. B. Worden, and D. J. Wald (2014). Intensity prediction equations for North America, *Bull. Seismol. Soc. Am.* **104,** no. 6, 3084–3093, doi: 10.1785/0120140178.

Biasi, G. P., and K. M. Scharer (2019). The current unlikely earthquake hiatus at California's transform boundary paleoseismic sites, *Seismol. Res. Lett.* **90,** no. 3, 1168–1176.

Boatwright, J., and H. Bundock (2005). Modified Mercalli intensity maps for the 1906 San Francisco earthquake plotted in ShakeMap format, *U.S. Geol. Surv. Open-File Rept. 2005-1135, Version 1.0*, https://pubs.usgs.gov/of/2005/1135/ (last accessed February 2020).

Boatwright, J., and H. Bundock (2008). Modified Mercalli intensity maps for the 1868 Hayward earthquake plotted in ShakeMap format: Spreadsheet of 1868 intensity sites, *U.S. Geol. Surv. Open-File Rept. 2008-1121*, https://pubs.usgs.gov/of/2008/1121/ (last accessed February 2020).

Bozorgnia, Y., N. A. Abrahamson, L. A. Atik, T. D. Ancheta, G. M. Atkinson, J. W. Baker, A. Baltay, D. M. Boore, K. W. Campbell, B. S. J. Chiou, *et al.* (2014). NGA-West2 research project, *Earthq. Spectra* **30,** no. 3, 973–987.

Brooks, E., J. Neely, S. Stein, B. Spencer, and L. Salditch (2019). Assessments of the performance of the 2017 one-year seismic hazard forecast for the Central and Eastern United States via simulated earthquake shaking data, *Seismol. Res. Lett.* **90,** 1155–1167.

Brooks, E., S. Stein, and B. D. Spencer (2016). Comparing the performance of Japan's earthquake hazard maps to uniform and randomized maps, *Seismol. Res. Lett.* **87,** no. 1, 90–102.

Brooks, E., S. Stein, and B. Spencer (2017). Investigating the effects of smoothing on the performance of earthquake hazard maps, *Int. J. Earthq. Impact Eng.* **2,** no. 2, 121–134.

Brooks, E., S. Stein, B. D. Spencer, L. Salditch, M. D. Petersen, and D. E. McNamara (2018). Assessing earthquake hazard map performance for natural and induced seismicity in the central and eastern United States, *Seismol. Res. Lett.* **89,** 118–126, doi: 10.1785/0220170124.

Byerly, P., and H. Dyk (1936). The questionnaire program for collecting earthquake data, in *Earthquake Investigations in California, 1934–1935*, N. H. Heck (Editor), Special Publication No. 201,

U.S. Department of Commerce Coast and Geodetic Survey, Washington, D.C., 43–48.

Cornell, C. A. (1968). Engineering seismic risk analysis, *Bull. Seismol. Soc. Am.* **58,** no. 5, 1583–1606.

Dengler, L. A., and J. W. Dewey (1998). An intensity survey of households affected by the Northridge, California, earthquake of 17 January 1994, *Bull. Seismol. Soc. Am.* **88,** no. 2, 441–462.

De Rubeis, V., K. Van Noten, T. Patrizia, and P. Sbarra (2019). 2019 survey report on the internet macroseismic practice in Europe, *7th International Colloquium on Historical Earthquakes and Paleoseismicity Studies*, Barcelona, Spain, 4–6 November 2019.

Dewey, J. W., B. G. Reagor, L. Dengler, and K. Moley (1995). Intensity distribution and isoseismal maps for the Northridge, California earthquake of January 17, 1994, *U.S. Geol. Surv. Open-File Rept. 95-92*, https://pubs.er.usgs.gov/publication/ofr9592 (last accessed February 2020).

Felzer, K. R., and E. E. Brodsky (2005). Testing the stress shadow hypothesis, *J. Geophys. Res.* **110,** no. 5, B003277, doi: 10.1029/2004JB003277.

Field, E. (2010). Probabilistic seismic hazard analysis: A primer, available at http://www.opensha.org/ (last accessed January 2020).

Field, E. H., R. J. Arrowsmith, G. P. Biasi, P. Bird, T. E. Dawson, K. R. Felzer, D. D. Jackson, K. M. Johnson, T. H. Jordan, C. Madden, *et al.* (2014). Uniform California earthquake rupture forecast, version 3 (UCERF3)—The time-independent model, *Bull. Seismol. Soc. Am.* **104,** no. 3, 1122–1180.

Fujiwara, H., N. Morikawa, Y. Ishikawa, T. Okumura, J. Miyakoshi, N. Nojima, and Y. Fukushima (2009). Statistical comparison of national probabilistic seismic hazard maps and frequency of recorded JMA seismic intensities from the K-NET strong-motion observation network in Japan during 1997–2006, *Seismol. Res. Lett.* **80,** no. 3, 458–464.

Gneiting, T., and M. Katzfuss (2014). Probabilistic forecasting, *Annu. Rev. Stat. Appl.* **1,** 125–151.

Hanks, T. C., and A. Johnston (1992). Common features of the excitation and propagation of strong ground motion for North American earthquakes, *Bull. Seismol. Soc. Am.* **82,** 1–23.

Harris, R. A., and R. W. Simpson (1998). Suppression of large earthquakes by stress shadows: A comparison of Coulomb and rate-and-state failure, *J. Geophys. Res.* **103,** no. B10, 24,439–24,451.

Hauksson, E., K. Felzer, D. Given, M. Giveon, S. Hough, K. Hutton, H. Kanamori, V. Sevilgen, S. Wei, and A. Yong (2008). Preliminary report on the 29 July 2008 Chino Hills, eastern Los Angeles Basin, California, earthquake sequence, *Seismol. Res. Lett.* **79,** no. 6, 855–866, doi: 10.1785/gssrl.79.6.855.

Hough, S. E. (2000). On the scientific value of "unscientific" data, *Seismol. Res. Lett.* **71,** 483–485.

Hough, S. E. (2013). Spatial variability of "Did You Feel It?" intensity data: Insights into sampling biases in historical earthquake intensity distributions, *Bull. Seismol. Soc. Am.* **103,** 2767–2781.

Hough, S. E. (2014). Earthquake intensity distributions: A new view, *Bull. Earthq. Eng.* **12,** 135–155, doi: 10.1007/s10518-013-9573-x.

Hough, S. E., and A. Elliot (2004). Revisiting the 23 February 1892 Laguna Salada earthquake, *Bull. Seismol. Soc. Am.* **94,** no. 4, 1571–1578.

Hough, S. E., and R. W. Graves (2020). The1933 Long Beach, California, earthquake: Ground motions and rupture scenario, *Sci. Rep.*, doi: 10.1038/s41598-020-66299-w.

Hough, S. E., and K. Hutton (2008). Revisiting the 1872 Owens Valley, California, earthquake, *Bull. Seismol. Soc. Am.* **98,** no. 2, 931–949.

Hough, S. E., and S. S. Martin (2018). A proposed rupture scenario for the 1925 Mw 6.6 Santa Barbara, California, earthquake, *Tectonophysics* **747,** 211–224.

Hough, S. E., and M. Page (2011). Toward a consistent model for strain accrual and release for the New Madrid Seismic Zone, central United States, *J. Geophys. Res.* **116,** no. B03311, doi: 10.1029/2010JB007783.

LaMontagne, M., and K. B. S. Burke (2019). Status and future of macro seismic information in Canada, *Seismol. Res. Lett.* **90,** no. 2B, 1014.

Locati, M., and A. Cassera (2010). MIDOP: Macro seismic intensity data online publisher, *Rapporti Tecnici INGV* **123,** 92 pp.

Mak, S., R. A. Clements, and D. Schorlemmer (2014). The statistical power of testing probabilistic seismic-hazard assessments, *Seismol. Res. Lett.* **85,** 781–783.

Mak, S., and D. Schorlemmer (2016). A comparison between the forecast by the U.S. National Seismic Hazard Maps with recent ground-motion records, *Bull. Seismol. Soc. Am.* **106,** no. 4, 1817–1831.

Marzocchi, W., and T. H. Jordan (2014). Testing for ontological errors in probabilistic forecasting models of natural systems, *Proc. Natl. Acad. Sci. Unit. States Am.* **111,** no. 33, 11,973–11,978.

Miyazawa, M., and J. Mori (2009). Test of seismic hazard map from 500 years of recorded intensity data in Japan, *Bull. Seismol. Soc. Am.* **99,** 3140–3149.

Nekrasova, A., V. Kossobokov, A. Peresanand, and A. Magrin (2014). Comparison of NDSHA, PSHA seismic hazard maps and real seismicity for the Italian territory, *Nat. Hazards* **70,** no. 1, 629–641, .

Petersen, M. D., M. P. Moschetti, P. M. Powers, C. S. Mueller, K. M. Haller, A. D. Frankel, Y. Zeng, S. Rezaeian, S. C. Harmsen, O. S. Boyd, *et al.* (2014). Documentation for the 2014 Update of the United States National Seismic Hazard Maps, *U.S. Geol. Surv. Open-File Rept. 2014-1091*, 243 pp., https://pubs.usgs.gov/of/2014/1091/ (last accessed February 2020).

Petersen, M. D., A. M. Shumway, P. M. Powers, C. S. Mueller, M. P. Moschetti, A. D. Frankel, S. Rezaeian, D. E. McNamara, N. Luco, O. S. Boyd, *et al.* (2020). The 2018 update of the US National Seismic Hazard Model: Overview of model and implications, *Earthq. Spectra* **36,** no. 1, 5–41, doi: 10.1177/8755293019878199.

Richter, C. F. (1958). *Elementary Seismology*, W.F. Freemanand Company, San Francisco, California.

Rukstales, K. S., and M. D. Petersen (2019). Data release for 2018 Update of the U.S. National Seismic Hazard Model: U.S. Geological Survey data release, available at https://www.sciencebase.gov/catalog/item/5cbf47c4e4b0c3b00664fdef (last accessed December 2019).

Salditch, L., S. E. Hough, S. Stein, B. D. Spencer, E. M. Brooks, J. S. Neely, and M. C. Lucas (2018). The 1952 Kern County, California earthquake: A case study of issues in the analysis of historical intensity data for estimation of source parameters, *Phys. Earth Planet. In.* **283,** 140–151, doi: 10.1016/j.pepi.2018.08.007.

Sibson, R. (1981). A brief description of natural neighbor interpolation, Chapter 2, in *Interpolating Multivariate Data*, V. Barnett (Editor), John Wiley & Sons, New York, New York, 21–36.

Sira, C., S. Auclair, I. Bouroullec, C. Cornou, I. Douste-Bacque, O. Guilhot, H. Jomard, E. Koufoudi, K. Manchuel, A. Montabert, *et al.*

(2019). Exploratory workshop to study the effect of time on macroseismic analyses, *7th International Colloquium on Historical Earthquakes and Paleoseismicity Studies*, Barcelona, Spain, 4–6 November 2019.

Sirovich, L., F. Pettenati, F. Cavallini, and M. Bobbio (2002). Natural-neighbor isoseismals, *Bull. Seismol. Soc. Am.* **92,** 1933–1940.

Sokolov, V. Y., and Y. K. Cherov (1998). On the correlation of seismic intensity with Fourier amplitude spectra, *Earthq. Spectra* **14,** 679–694, doi: 10.1193/1.1586022.

Stein, S., B. D. Spencer, and E. M. Brooks (2015). Metrics for assessing earthquake hazard map performance, *Bull. Seismol. Soc. Am.* **105,** no. 4, 2160–2173.

Stirling, M., and M. Gerstenberger (2010). Ground motion-based testing of seismic hazard models in New Zealand, *Bull. Seismol. Soc. Am.* **100,** no. 4, 1407–1414.

Stirling, M. W., and M. Petersen (2006). Comparison of the historical record of earthquake hazard with seismic-hazard models for New Zealand and the continental United States, *Bull. Seismol. Soc. Am.* **96,** 1978–1994.

Stover, C. W. (1988). *United States Earthquakes, 1984*, Bulletin 1862, U.S. Government Publishing Office, 26 pp., doi: 10.3133/b1862.

Stover, C. W., and J. L. Coffman (1993). Seismicity of the United States, 1568–1989, *U.S. Geol. Surv. Profess. Pap. 1527*, United States Government Printing Office, Washington, D.C.

Stucchi, M., P. Albini, C. Mirto, and A. Rebez (2004). Assessing the completeness of historical earthquake data, *Ann. Geophys.* **47,** nos. 2/3, 659–673.

Tasan, H., C. Beauval, A. Helmstetter, A. Sandikkaya, and P. Guéguen (2014). Testing probabilistic seismic hazard estimates against accelerometric data in two countries: France and Turkey, *Geophys. J. Int.* **198,** 1554–1571.

Toppozada, T. R. (1984). *History of Earthquake Damage in Santa Clara County and Comparison of 1911 and 1984 Earthquakes*, California Division of Mines and Geology, special publications.

Toppozada, T. R,, and D. Branum (2004). California earthquake history, *Ann. Geophys.* **47,** nos. 2/3, doi: 10.4401/ag-3317.

Toppozada, T. R., C. R. Real, S. P. Bezore, and D. L. Parke (1981). Preparation of isoseismal maps and summaries of reported effects for pre-1900 California earthquakes, *U.S. Geol. Surv. Open-File Rept. 81-262*, https://pubs.usgs.gov/of/1981/0262/report.pdf (last accessed February 2020).

Trifunac, M. D., and A. G. Brady (1975). On the correlation of seismic intensity scales with peaks of recorded ground motion, *Bull. Seismol. Soc. Am.* **65,** no. 1, 139–162.

U.S. Department of Commerce (USDOC), Environmental Services Administration, Coast and Geodetic Survey (1966). Abstracts of earthquake reports for the Pacific Coast and the Western Mountain Region, *MSA-74 April, June, July 1952*, 2–115, https://hdl.handle.net/2027/uc1.31210003187075?urlappend=%3Bseq=70 (last accessed June 2020).

Vanneste, K., S. Stein, T. Camelbeeck, and B. Vleminckx (2018). Insights into earthquake hazard map performance from shaking history simulations, *Sci. Rep.* **8,** 1855.

Vick, S. G. (2002). *Degrees of Belief: Subjective Probability and Engineering Judgment*, ASCE Publications, Reston, Virginia.

von Hake, C. A., and M. K. Cloud (1984). *United States Earthquakes, 1966*, Department of Interior, U.S. Geological Survey, OFR 84–966.

Wald, D. J., V. Quitoriano, L. Dengler, and J. W. Dewey (1999). Utilization of the internet for rapid community intensity maps, *Seismol. Res. Lett.* **70,** no. 6, 680–693.

Wang, Z. (2015). Predicting or forecasting earthquakes and the resulting ground motion hazards: A dilemma for earth scientists, *Seismol. Res. Lett.* **86,** 1–5.

Ward, S. N. (1995). Area-based tests of long-term seismic hazard predictions, *Bull. Seismol. Soc. Am.* **85,** 1285–1298.

Wood, H. O., and F. Neumann (1931). Modified Mercalli intensity scale of 1931, *Bull. Seismol. Soc. Am.* **21,** no. 4, 277–283.

Worden, C. B., M. C. Gerstenberger, D. A. Rhoades, and D. J. Wald (2012). Probabilistic relationships between ground-motion parameters and modified Mercalli intensity in California, *Bull. Seismol. Soc. Am.* **102,** 204–221, doi: 10.1785/0120110156.

Working Group on California Earthquake Probabilities (WGCEP) (2013). Uniform California earthquake rupture forecast, version3 (UCERF3)—The time-independent model, *U.S. Geol. Surv. Open-File Rept. 2013-1165, California Geol. Surv. Spec. Rept. 228, and Southern California Earthquake Center Publication 1792*, available at http://pubs.usgs.gov/of/2013/1165/ (last accessed February 2020).

Appendix

Table A1 contains references for the sources of intensity data in the CHIMP dataset.

TABLE A1
References for Earthquake Data Sources

CHIMP Label	Date (yyyy/mm/dd)	Location	Reference
1857	1857/01/09	Fort Tejon*	Hough, S. E., L. Salditch, M. G. Gallahue (2020). This study.
1868	1868/10/21	Hayward	Boatwright and Bundock (2008)
1872	1872/03/26	Owens Valley	Hough and Hutton (2008)

CHIMP, California Historical Intensity Mapping Project; DYFI, "Did You Feel It?".
*Events reinterpreted in this study.
(Continued next page.)

TABLE A1 (continued)
References for Earthquake Data Sources

CHIMP Label	Date (yyyy/mm/dd)	Location	Reference
1873	1873/11/23	California-Oregon border	Tom Brocher, unpublished OFR, personal comm. (2019)
1892	1892/02/23	Laguna Salada	Hough and Elliot (2004)
1898	1898/04/15	Mendocino*	Toppozada *et al.* (1981, their Appendix D)
1906	1906/04/18	San Francisco	Boatwright and Bundock (2005)
1911	1911/07/01	South Bay*	Toppozada (1984)
1918	1918/04/21	San Jacinto*	Lucas, M. C. (2019). This study.
1925	1925/06/29	Santa Barbara	Hough and Martin (2018)
1927a	1927/11/18	Bishop*	Frank Nuemann, U.S. Coast and Geodetic Survey, Seismological report of July, August, September 1927. Serial number 495. Washington: G.P.O.
1927b	1927/11/04	Lompoc*	Lucas, M. C. (2019). This study.
1932	1932/12/20	Nevada*	Coast and Geodetic Survey (1984). United States Earthquakes, 1928–1935, Department of the Interior, U.S. Geological Survey, OFR 84-928.
1933	1933/03/10	Long Beach	Hough and Graves (2020)
1934	1934/12/31	Colorado River*	Abstracts of earthquake reports for the Pacific Coast and the eastern mountain region 1 July 1934 to 30 September 1934. Department of Commerce, U.S. Coast and Geodetic Survey Field Station 510 Custom House San Francisco, California.
1937	1937/03/25	Buck Ridge*	Coast and Geodetic Survey, United States Earthquakes, 1928–1935, Department of the Interior, U.S. Geological Survey, OFR 84-928.
1940	1940/05/18	Imperial Valley*	Abstracts of earthquake reports for the Pacific Coast and the eastern mountain region 1 April 1940 to 30 June 1940. MSA-26, Department of Commerce, U.S. Coast and Geodetic Survey, Seismological Field Survey 214, Old Mint Building San Francisco 3, California.
1942	1942/10/21	Fish Creek*	Abstracts of earthquake reports for the Pacific Coast and the eastern mountain region 1 October 1943 to 31 December 1947. MSA-40, Department of Commerce, U.S. Coast and Geodetic Survey, Seismological Field Survey 214, Old Mint Building San Francisco 3, California.
1946	1946/03/15	Walker Pass*	Abstracts of earthquake reports for the Pacific Coast and the eastern mountain region 1 January 1946 to 31 March 1946. MSA-49, Department of Commerce, U.S. Coast and Geodetic Survey, Seismological Field Survey 214, Old Mint Building San Francisco 3, California.
1947	1947/04/10	Manix*	Abstracts of earthquake reports for the Pacific Coast and the eastern mountain region 1 April 1947 to 30 June 1947. MSA-54, Department of Commerce, U.S. Coast and Geodetic Survey, Seismological Field Survey 214, Old Mint Building San Francisco 3, California.
1948a	1948/12/04	Desert Hot Springs*	Abstracts of earthquake reports for the Pacific Coast and the eastern mountain region 1 October 1948 to 31 December 1948. MSA-60, Department of Commerce, U.S. Coast and Geodetic Survey, Seismological Field Survey 214, Old Mint Building San Francisco 3, California.
1948b	1948/12/29	Northeast California*	Abstracts of earthquake reports for the Pacific Coast and the eastern mountain region 1 October 1948 to 31 December 1948. MSA-60, Department of Commerce, U.S. Coast and Geodetic Survey, Seismological Field Survey 214, Old Mint Building San Francisco 3, California.
1952a	1952/07/21	Kern County	Salditch *et al.* (2018)
1952b	1952/11/22	Bryson*	Abstracts of earthquake reports for the Pacific Coast and the eastern mountain region, 1 July 1951 to 30 September 1951. MSA-71, 214 Old Mint Building, San Francisco 3, California.

CHIMP, California Historical Intensity Mapping Project; DYFI, "Did You Feel It?".
*Events reinterpreted in this study.
(Continued next page.)

TABLE A1 (continued)
References for Earthquake Data Sources

CHIMP Label	Date (yyyy/mm/dd)	Location	Reference
1954a	1954/03/19	San Jacinto	Abstracts of earthquake reports for the Pacific Coast and the eastern mountain region 1 January 1954 to 31 March 1954. MSA-81, Department of Commerce, U.S. Coast and Geodetic Survey Seismological Field Station 14 Old Mint Building San Francisco 3, California. MSA-83.
1954b	1954/07/06	Nevada*	Abstracts of earthquake reports for the Pacific Coast and the eastern mountain region 1 July 1954 to 30 September 1954. MSA-83, Department of Commerce, U.S. Coast and Geodetic Survey Seismological Field Station 14 Old Mint Building San Francisco 3, California. MSA-83.
1966	1966/09/12	Northern California-Truckee*	von Hake and Cloud (1984)
1968	1968/04/08	Borrego*	Abstracts of earthquake reports for the United States 1 April 1968 to 30 June 1968. MSA-137, Department of Commerce, Seismological Field Survey, San Francisco, California.
1971	1971/02/09	Sylmar	DYFI, DYFI intensity summary geocoded UTM aggregated (10 km spacing).txt, available at https://earthquake.usgs.gov/earthquakes/eventpage/ci3347678/executive (last accessed January 2019).
1979	1979/10/15	Imperial Valley	DYFI, DYFI intensity summary geocoded UTM aggregated (1 km spacing).txt, available at https://earthquake.usgs.gov/earthquakes/eventpage/ci3352060/executive (last accessed January 2019).
1980a	1980/05/25	Mammoth*	Stover, C. W., and C. A. von Hake. United States Earthquakes, 1980. OFR 84-980. Department of the Interior, U.S. Geological Survey.
1980b	1980/11/08	Eureka*	Stover, C. W., and C. A. von Hake. United States Earthquakes, 1980. OFR 84-980. Department of the Interior, U.S. Geological Survey.
1983	1983/05/02	Coalinga	DYFI, DYFI intensity summary geocoded UTM aggregated (1 km spacing).txt, available at https://earthquake.usgs.gov/earthquakes/eventpage/nc1091100/executive (last accessed January 2019).
1984a	1984/04/24	Morgan Hill*	Stover (1988)
1984b	1984/11/23	Round Valley*	Stover (1988)
1986a	1986/07/08	North Palm Springs	DYFI, DYFI intensity summary (city or zip code aggregated).txt, available at https://earthquake.usgs.gov/earthquakes/eventpage/ci700917/executive (last accessed January 2019).
1986b	1986/07/21	Chalfant Valley*	Stover, C. W., and Brewer, L. R. (1986). U.S. earthquake reports.
1987a	1987/10/01	Whittier Narrows*	Jim Dewey, U.S. earthquakes, 1987, unpublished, personal comm. (2019)
1987b	1987/11/24	Superstition Hills*	Jim Dewey, U.S. earthquakes, 1987, unpublished, personal comm. (2019)
1989	1989/10/17	Loma Prieta	DYFI, DYFI intensity summary geocoded (1 km spacing).txt, available at https://earthquake.usgs.gov/earthquakes/eventpage/nc216859/executive (last accessed January 2019).
1991	1991/06/28	Sierra Madre	DYFI, DYFI intensity summary geocoded (1 km spacing).txt, available at https://earthquake.usgs.gov/earthquakes/eventpage/ci2021449/executive (last accessed March 2019).
1992a	1992/04/22	Joshua Tree*	Salditch and Gallahue (2019). This study.
1992b	1992/04/25	Rio Dell	DYFI, DYFI intensity summary geocoded (1 km spacing).txt, available at https://earthquake.usgs.gov/earthquakes/eventpage/nc269151/executive (last accessed March 2019).
1992c	1992/06/28	Landers	DYFI, DYFI intensity summary geocoded (1 km spacing).txt, available at https://earthquake.usgs.gov/earthquakes/eventpage/ci3031111/executive (last accessed January 2019).
1992d	1992/06/28	Big Bear	DYFI, DYFI intensity summary aggregated by city or zip code.txt, available at https://earthquake.usgs.gov/earthquakes/eventpage/ci3031425/executive (last accessed January 2019).
1993	1993/05/17	Big Pine*	Salditch and Gallahue (2019). This study.

CHIMP, California Historical Intensity Mapping Project; DYFI, "Did You Feel It?".
*Events reinterpreted in this study.
(Continued next page.)

TABLE A1 (continued)
References for Earthquake Data Sources

CHIMP Label	Date (yyyy/mm/dd)	Location	Reference
1994	1994/01/17	Northridge	DYFI, DYFI intensity summary geocoded (10 km spacing).txt, available at https://earthquake.usgs.gov/earthquakes/eventpage/ci3144585/executive (last accessed January 2019).
1999	1999/10/16	Hector Mine	DYFI, DYFI intensity summary geocoded (10 km spacing).txt, available at https://earthquake.usgs.gov/earthquakes/eventpage/ci9108652/executive (last accessed January 2019).
2000	2000/09/03	Northern California	DYFI, DYFI intensity summary geocoded (10 km spacing).txt, available at https://earthquake.usgs.gov/earthquakes/eventpage/nc21123384/executive (last accessed March 2019).
2001	2001/08/10	Northern California	DYFI, DYFI intensity summary geocoded (10 km spacing).txt, available at https://earthquake.usgs.gov/earthquakes/eventpage/nc21188442/executive (last accessed March 2019).
2003	2003/12/22	San Simeon	DYFI, DYFI intensity summary geocoded (10 km spacing).txt, available at https://earthquake.usgs.gov/earthquakes/eventpage/nc21323712/executive (last accessed January 2019).
2004	2004/09/28	Shandon	DYFI, DYFI intensity summary geocoded (10 km spacing).txt, available at https://earthquake.usgs.gov/earthquakes/eventpage/ci14095628/executive (last accessed March 2019).
2005a	2005/06/12	Anza	DYFI, DYFI intensity summary geocoded (10 km spacing).txt, available at https://earthquake.usgs.gov/earthquakes/eventpage/ci14151344/executive (last accessed March 2019).
2005b	2005/06/14	Mendocino	DYFI, DYFI intensity summary geocoded (10 km spacing).txt, available at https://earthquake.usgs.gov/earthquakes/eventpage/usp000dt25/executive (last accessed January 2019).
2007	2007/10/30	San Francisco Bay	DYFI, DYFI intensity summary geocoded (1 km spacing).txt, available at https://earthquake.usgs.gov/earthquakes/eventpage/nc40204628/executive (last accessed March 2019).
2008a	2008/02/21	Wells, Nevada	DYFI, DYFI intensity summary geocoded (1 km spacing).txt, available at https://earthquake.usgs.gov/earthquakes/eventpage/nn00234425/executive (last accessed January 2019).
2008b	2008/07/29	Chino Hills	DYFI, DYFI intensity summary geocoded (1 km spacing).txt, available at https://earthquake.usgs.gov/earthquakes/eventpage/ci14383980/executive (last accessed March 2019).
2009	2009/05/17	Lennox	DYFI, DYFI intensity summary geocoded (10 km spacing).txt, available at https://earthquake.usgs.gov/earthquakes/eventpage/ci10410337/executive (last accessed March 2019).
2010	2010/04/04	Baja	DYFI, DYFI intensity summary geocoded (1 km spacing).txt, available at https://earthquake.usgs.gov/earthquakes/eventpage/usp000jhr6/executive (last accessed January 2019).
2014	2014/08/24	Napa	DYFI, DYFI intensity summary geocoded (1 km spacing).txt, available at https://earthquake.usgs.gov/earthquakes/eventpage/nc72282711/executive (last accessed January 2019).
2019a	2019/07/04	Ridgecrest	DYFI, cdi_geo_1km.txt, available at https://earthquake.usgs.gov/earthquakes/eventpage/ci38443183/executive (last accessed August 2019).
2019b	2019/07/06	Ridgecrest	DYFI, cdi_geo_1km.txt, available at https://earthquake.usgs.gov/earthquakes/eventpage/ci38457511/executive (last accessed August 2019).

CHIMP, California Historical Intensity Mapping Project; DYFI, "Did You Feel It?".
*Events reinterpreted in this study.

Manuscript received 18 February 2020
Published online 1 July 2020