



HAL
open science

Conditional probability of surface rupture: A numerical approach for principal faulting

Lisa Mammarella, Francesco Visini, Paolo Boncio, Stephane Baize, Oona Scotti, Celine Beauval, Bruno Pace, Steve Thompson

► **To cite this version:**

Lisa Mammarella, Francesco Visini, Paolo Boncio, Stephane Baize, Oona Scotti, et al.. Conditional probability of surface rupture: A numerical approach for principal faulting. *Earthquake Spectra*, In press, 10.1177/87552930241293570 . irsn-04889059

HAL Id: irsn-04889059

<https://asn.hal.science/irsn-04889059v1>

Submitted on 15 Jan 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Conditional probability of surface rupture: A numerical approach for principal faulting

Earthquake Spectra

1–26

© The Author(s) 2024

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/87552930241293570

journals.sagepub.com/home/eqs

Lisa Mammarella¹ , Francesco Visini², Paolo Boncio¹,
Stephane Baize³, Oona Scotti³, Celine Beauval⁴,
Bruno Pace¹, and Steve Thompson, M.EERI⁵

Abstract

This study presents a numerical approach for probabilistic fault displacement hazard analysis (PFDHA), aimed at addressing an alternative solution with commonly used empirical methodologies. Our model utilizes probability distributions to compute the conditional probability of surface rupture (CPSR). Leveraging earthquake catalogs, we derived the hypocentral depth distribution (HDD) across eight globally distributed seismotectonic regions categorized by faulting kinematics (normal, reverse, strike-slip). We calculated the hypocentral depth ratio (HDR) distribution, to model rupture position from the hypocenter. Employing magnitude scaling relations we determined rupture widths (W) spanning magnitudes 5–8. User-input parameters, including fault style, average dip angle, and seismogenic depth, with associated uncertainties, derive the CPSR estimation of surface rupture occurrences. Our findings highlight seismogenic depth as the most influential parameter and reveal correspondences between empirical curves derived for specific regions, emphasizing the importance of site-specific rupture probability assessments over global datasets and underscores the significance of considering seismotectonic context when evaluating fault displacement hazard. The numerical code for CPSR calculation has been developed and is openly accessible on GitHub.

¹Dipartimento di Ingegneria e Geologia, Università degli Studi “G. d’Annunzio” Chieti – Pescara, Chieti, Italy

²Sezione di Chieti, Istituto Nazionale di Geofisica e Vulcanologia, Chieti, Italy

³Institut de Radioprotection et de Sûreté Nucléaire, Fontenay-aux-Roses, France

⁴CNRS, Grenoble-Alpes University, University of Savoie Mont Blanc, IRD, Gustave Eiffel University, ISTerre, Grenoble, France

⁵Lettis Consultants International, Inc., Concord, CA, USA

Corresponding author:

Lisa Mammarella, Dipartimento di Ingegneria e Geologia, Università degli Studi “G. d’Annunzio” Chieti – Pescara, Via dei Vestini, 31, IT 66100 Chieti, Italy.

Email: lisa.mammarella@unich.it

Keywords

PFDHA, probability of surface rupture, magnitude scaling relations, region-specific analysis, numerical model

Date received: 8 May 2024; accepted: 26 September 2024

Introduction

Fault displacement hazard analysis assesses the likelihood of ground rupture and the expected amount of displacement across a fault or fault zone given the occurrence of a ground rupturing earthquake. This hazard can be relevant for important facilities such as lifelines (pipelines, transit systems) and sensitive structures such as landfills, dams, and nuclear power plants located on or near an active fault. For situations where fault displacement hazard cannot be mitigated by avoidance, probabilistic fault displacement hazard analysis (PFDHA) is suggested (e.g. Moss and Ross, 2011; Petersen et al., 2011; Takao et al., 2013; Valentini et al., 2021; Youngs et al., 2003). As described by Youngs et al. (2003), the earthquake approach to PFDHA follows the commonly used formulation for probabilistic seismic hazard analysis (Cornell, 1968) wherein the magnitude-frequency distribution of earthquakes is determined based on a characterization of active fault sources, but fault displacement prediction equations (or fault displacement models) replace ground-motion prediction equations. Typically, PFDHA addresses three conditional probabilities: the conditional probability of surface rupture (CPSR), or the probability that an earthquake of magnitude m reaches the ground surface $P(\text{surface rupture} | m)$; the conditional probability that the surface rupture will intersect the facility of interest located a distance r from the fault $P(\text{rupture at site} | m, r, \text{surface rupture})$; and the conditional probability of displacement exceedance $P(D > d | m, r, \text{rupture at site})$.

This work focuses on the CPSR on principal fault for shallow crustal environments. The CPSR has a direct impact on the expected annual frequency of surface-fault ruptures at a site, and uncertainty in this probability can be a major contributor to total uncertainty in a PFDHA depending on the site and tectonic environment. Current models are based on either purely empirical approaches using the ergodic assumption (Moss et al., 2013; Moss and Ross, 2011; Pizza et al., 2023; Takao et al., 2013; Wells and Coppersmith, 1993; Youngs et al., 2003) or more numerical approaches that consider the specific width of the fault source (Otsuka, 1964; Pacific Gas and Electric Company (PG&E), 2010; Youngs et al., 2003). The purely empirical models are most common, and these rely on regional or global datasets of earthquakes with known magnitudes and knowledge of whether each earthquake did or did not produce surface-fault rupture. These empirical models commonly use logistic regression on data subsets to derive the CPSR as a function of earthquake magnitude. Wells and Coppersmith (1993) developed the first widely used model using worldwide data of all styles of faulting. Youngs et al. (2003) published additional empirical models for normal faulting using the database of Pezzopane and Dawson (1996) for the extensional Cordillera, the Great Basin, and the Basin and Range provinces of the western United States. Moss and Ross (2011) developed empirical models for reverse faulting building on the Lettis et al. (1997) worldwide database, and Moss et al. (2013) further expanded on the empirical modeling by documenting that the CPSR for reverse faulting appears to correlate with proxy values for VS30. Takao et al. (2013) developed empirical regressions based on a database of earthquakes in Japan that combined reverse and strike-slip styles of faulting.

More recently, Pizza et al. (2023) compiled a global database of earthquakes from 1992 to 2018 that did and did not produce surface rupture. Following similar approaches to earlier studies, their results differed, often substantially, from those of earlier studies both for specific styles of faulting and for all styles of faulting combined. Their study, combined with previous model results, highlights several features of the CPSR: first, that a particular tectonic region or specific style of faulting may have an important influence on CPSR; second, that the specific dataset used strongly affects the results; third, the reliability of the empirical method depends on the quality of the dataset and the ability to establish conclusively the occurrence or absence of surface ruptures. The Pizza et al. (2023) analysis determined that 60% of the updated database lacked surface rupture information. Furthermore, the set of earthquakes magnitude $m \geq 5$ that had confident documentation for the presence or absence of surface rupture was distributed subequally by magnitude, rather than exponentially distributed as should be expected for an unbiased dataset (Gutenberg and Richter, 1944). These features of empirical studies—apparent tectonic dependence, but also a dependence on specific database used, the challenge of obtaining a complete, unbiased database, and instances where known seismogenic thickness differs from the average thickness of the dataset, as exemplified by the 2019 Le Teil earthquake (e.g. Marconato et al., 2022; Ritz et al., 2020)—provide the motivation to explore numerical methods to calculate the CPSR.

Here, we propose a numerical modeling approach for the CPSR for principal faults that can be used in a PFDHA as an epistemic alternative to fully empirical approaches. The model calculates the CPSR based on location-specific estimates of down-dip fault geometry (i.e. fault dip and seismogenic depth). Given the fault's geometry, the model estimates the CPSR as a function of magnitude from three probability distributions: the position of the earthquake hypocenter along the fault's width, the total rupture width, and the up-dip and down-dip extents of fault rupture from the hypocenter. The framework for exploring down-dip earthquake rupture geometries is adopted from the Next Generation Attenuation for Central and Eastern North America (NGA-East) project (Youngs et al., 2021). The model can capture uncertainty in the fault geometry parameters from site-specific information, but constraints on hypocenter depth distributions, rupture widths as a function of magnitude, and the ratio of up-dip to down-dip rupture propagation from the hypocenter are determined from empirical data. To apply the method across various tectonic regimes, we examined seismotectonically homogeneous areas, such as extensional zones of Italy and the Great Basin province of the United States, contractional zones of Italy, Taiwan, and Japan, and strike-slip zones of New Zealand, Japan, and California, utilizing earthquake catalogs.

Model parameters

The model parameters considered to evaluate CPSR of a rupture are shown in cross-section view of Figure 1. The fault plane has a down-dip width W_F that is a function of the fault dip δ , and the local seismogenic depth Z_S . As we are concerned with the probability of surface rupture, the top of the fault plane is always at the ground surface ($Z = 0$). The rupture plane, indicated by the red line, has a width W and a vertical projection W_z .

Surface rupture is defined as the state where $W_{Z_{TOP}}$ reach the surface, constituting what we refer to as the *surface rupture condition*. The CPSR ($P(\text{surface rupture}|m)$) is function of the width of the rupture (W), which depends only on m , the dip angle of the fault (δ), the

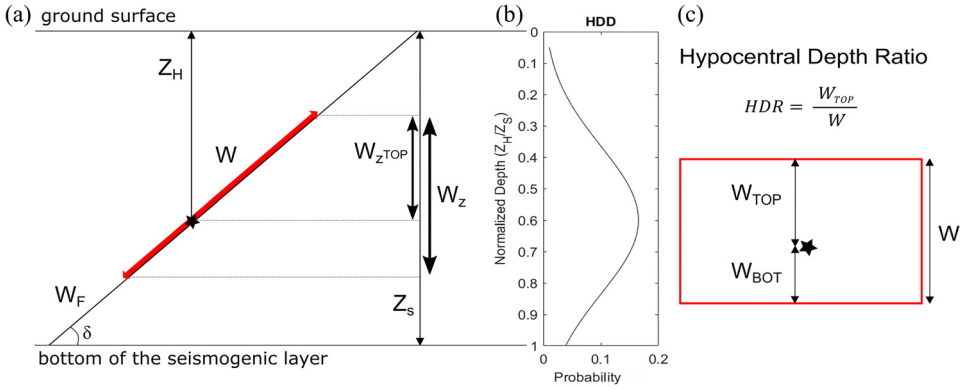


Figure 1. Schematic illustration of the model and parameters used in the calculation: (a) 2D representation of the fault and rupture geometry for an extensional domain, and (b) example of HDD computed for an extensional domain, and (c) definition of hypocentral depth ratio (HDR), following Chiou and Youngs (2008). Fault parameters: W = rupture width; W_F = fault width; δ = fault dip angle; Z_S = maximum seismogenic depth; Z_H = hypocentral depth; $W_{Z_{TOP}}$ = distance from the hypocenter to the top of the rupture; W_{BOT} = distance from the hypocenter to the bottom of the rupture; W_Z = vertical projection of the W ; $W_{Z_{TOP}}$ = vertical projection of the W_{TOP} ; star = hypocenter.

seismogenic depth (Z_S), the relative position of the rupture with respect to the hypocenter (HDR), and the hypocentral depth distribution (HDD):

$$P(\text{surface rupture}|m) = f(W, \delta, Z_S, HDR, HDD)$$

HDR and HDD are described more completely in the next subsections (see Table 1 for a list of abbreviations used in this article).

We used a numerical approach for computing the CPSR described in Appendix 1. To summarize, we convolve the distributions of input parameters and obtain a sample of $W_{Z_{TOP}}$ with associated probabilities. Then, we sum the sampled probabilities satisfying the surface rupture condition to define a magnitude-dependent probability of surface rupture.

Estimating model parameters: W , HDD, HDR, and adjusted HDR

Rupture width— W

Empirical scaling relations are used to compute W from moment magnitude (M_W). As our analysis is focused on shallow crustal (rather than subduction) environments, we explored the magnitude range $5 \leq M_W \leq 8$. Two published sources of empirical relations are presented below as examples. We considered scaling relations for W and not rupture length or rupture area because we are interested in whether the surface rupture condition occurs at the surface. Our model for W comes from two different magnitude scaling relations, namely Leonard (2014) and Thingbaijam et al. (2017). These models are called L14 and T17, respectively, in the following sections.

L14 developed coefficients for a linear equation to estimate M_W from $\log_{10} W$ based on simple scaling law assumptions and fitting empirical data. Separate relations were developed for actively deforming (interplate) and stable continental regions (intraplate), and for

Table 1. Variables (Var.) and Abbreviations (Abb.) used in this article

Var.	Explanation	Abb.	Explanation
M_w	Moment magnitude	STZ	Seismotectonic zone
W	Rupture width (down-dip)	HDR	Hypocentral depth ratio
W_Z	Vertical projection of the rupture width	HDD	Hypocentral depth distribution
W_{TOP}	Distance from hypocenter to rupture top	D90	90% seismicity cut-off depth
W_{ZTOP}	Vertical projection of W_{TOP}	L14	Magnitude scaling relation provided by Leonard (2014)
W_F	Fault width (seismogenic down-dip)	T17	Magnitude scaling relation provided by Thingbaijam et al. (2017)
δ	Fault dip angle		
Z_H	Hypocentral depth		
Z_S	Maximum seismogenic depth		

strike-slip and dip-slip styles of faulting. In the proposed model, both interplate and intra-plate regressions are included, and users can select between these regressions by specifying their preferred type. However, the comparison within the article focuses exclusively on interplate regions. The equations of L14 are self-consistent and may be rearranged to solve for W as a function of M_w . An important constraint in the L14 model is a fixed value for the exponent in the power law relationship between W and rupture length ($\beta = 2/3$) that is applied to all styles of faulting. This constraint results in a fixed slope in the $\log_{10} W - M_w$ relation of 0.4 in the L14 models.

T17 developed empirical earthquake source-scaling laws based on the Finite-Fault Source Model Database (SRCMOD) (Mai and Thingbaijam, 2014). Linear equation coefficients to estimate $\log_{10} W$ from M_w and the standard deviation of the prediction were developed for four faulting environments: shallow crustal reverse, normal, strike-slip, and subduction zone interface. Unlike L14, T17 define the slope of the $\log_{10} W - M_w$ relationship based on orthogonal regression that fits the empirical data.

Model variability in W within each scaling relation is based on a calculated or estimated standard deviation in the $\log_{10} W - M_w$ relationship with an upper limit based on W_F . T17 provide calculated standard deviations. For the L14 relations, we adopted a value of 0.15 for the standard deviation in $\log_{10} W$ based on the measured values in Wells and Coppersmith (1994). Figure 2 shows an example of how uncertainties in W are implemented for a normal-faulting earthquake with $M_w = 6.5$ on a fault with $W_F = 17$ km and L14 scaling. The distribution of W follows a log-normal distribution, and the model considers the full range of probabilities within ± 1 standard deviation. For the sake of clarity, rupture widths are categorized based on their saturation relative to the fault width. Specifically, rupture widths lower than W_F are considered small ruptures, while values greater than W_F are regarded as large ruptures. For large ruptures, the model recognizes the saturation of fault width and imposes the condition of surface rupture as being met.

Hypocentral depth distribution

The HDD and Z_S are deduced from earthquake catalog analysis. We present HDDs for a variety of tectonically homogeneous areas (or seismotectonic zones, STZs) with the expectation that many (if not most) site-specific studies will not have sufficient data or resources to calculate a local HDD, and that HDDs from different areas may be used for a site-specific analysis. Earthquake catalogs from eight different STZs within six interplate regions were compiled and analyzed to give a modeling framework suitable for different

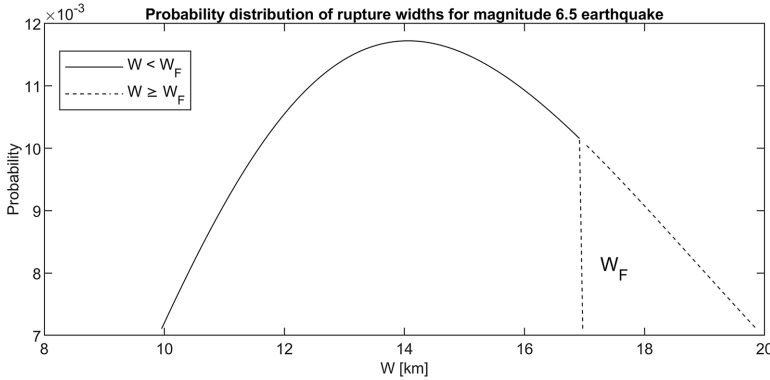


Figure 2. The plot illustrates the uncertainties of rupture widths (1 standard deviation from the mean) according to Leonard (2014) magnitude scaling relation, normal faulting regime, and a $W_F = 17$ km. Rupture widths lower than W_F are considered small ruptures, while values exceeding the W_F are regarded as large ruptures. In this specific example, $\delta = 50^\circ$ and $Z_S = 13$ km.

seismotectonic environments. Where possible, data were limited to events with low formal hypocentral errors. We set earthquake selection criteria requiring the estimated horizontal error to be less than 3 km and the vertical error to be less than 5 km. In cases where this information was unavailable, catalog data were judged to be acceptable based on general statements about data quality stated by the provider.

As the HDD is a probability distribution normalized by Z_S , our first interest is to estimate Z_S . For each STZ, we selected earthquakes with magnitudes equal to or greater than 2.5 and hypocentral depths shallower than the Moho depth. To achieve this, regional-scale numerical models of Moho depth were sought for each seismic zone. The Moho depth models were obtained from the literature and are detailed in the section “Data used to estimate HDD and HDR.” This strategy allowed us to exclude subduction earthquakes or deep crustal events that do not belong to the upper brittle crust.

The Z_S for each STZ can be evaluated as the depth above which a defined percentage of observed seismicity occurs, known as the cut-off depth (e.g. Hauksson and Meier, 2019; Maggini and Caputo, 2021; Omuralieva et al., 2012; Sibson, 1982; Zheng et al., 2021). In this study, we estimated Z_S using the cut-off depth at which 90% of the cataloged earthquakes occur (D90). The D90 of a STZ was computed based on sampling 20×20 km areas at 5 km interval spacing, following the grid strategy of Latorre et al. (2023). A minimum of 25 earthquakes was required for a valid D90 calculation. In practice, for each square (20×20 km area): (1) only crustal events are selected by ignoring earthquakes deeper than the average Moho depth; (2) D90 is calculated; and (3) the hypocentral depths are normalized to D90. The HDD for the entire STZ is determined by stacking the normalized depth histograms from all the analyzed squares.

Because the HDD of interest is that of moderate to large earthquakes that can produce displacements of engineering significance, we attempted to calculate HDDs for different minimum magnitudes. There are insufficient numbers of events in the target magnitude range of interest (generally $m \geq 5$). We found that calculating HDD based on magnitudes greater than 3.5 provided a reasonable balance between having enough samples for statistical reliability and having statistics on the earthquake size range of greatest interest.

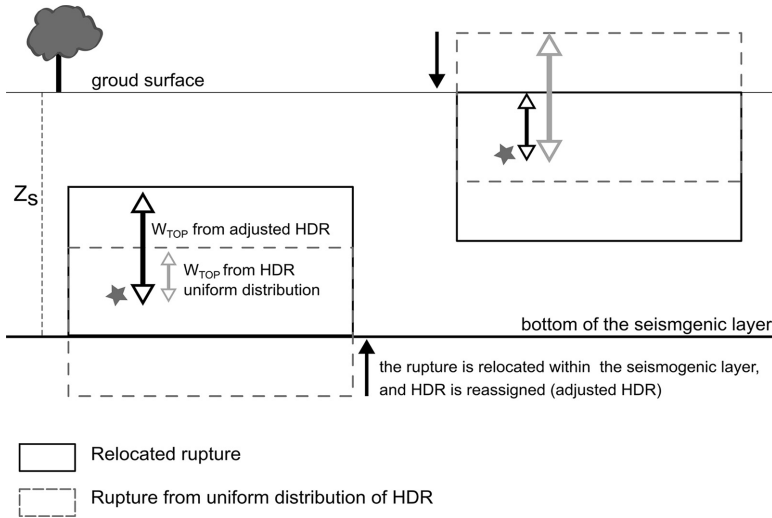


Figure 3. Illustrative sketch showing how the ruptures are reallocated. The ground surface and bottom of the seismogenic layer represent the boundary layers. The rectangles depict the rupture before (dashed line) and after (solid line) reallocation. The hypocenter is marked with a star. The solid black arrow indicates the direction of the rupture’s repositioning.

Hypocentral depth ratio and adjusted HDR

The HDR provides the rupture’s position relative to the hypocenter. Initially defined by Chiou and Youngs (2008), HDR represents the ratio between the distance from the hypocenter to the top of the rupture (W_{TOP}) and the full rupture width (Figure 1c). Special cases for the HDR include the value $HDR = 0$, which implies nucleation at the top of the rupture, allowing propagation only downward (i.e. all down-dip), and $HDR = 1$, which indicates propagation solely upward (i.e. all up-dip). Initial probability (prior) distributions of HDR follow a uniform probability distribution in the range $0 < HDR < 1$. The final distribution involves reallocating ruptures because, in our model, ruptures from all possible combinations of variables could extend beyond the upper and lower boundary layers (i.e. $Z = 0$ and $Z = Z_S$). To address this issue, we reallocated the ruptures that violated the boundary layers by reassigning the HDR value. We will refer to the distribution obtained from this reassignment as “adjusted” HDR. Figure 3 provides an illustrative example of how HDRs are reassigned. Consider a scenario where the rupture extends below the boundary layer (dashed rectangle in Figure 3), with an initial HDR value. To ensure correct repositioning of this rupture inside the boundary layer, we define an adjusted HDR considering the Z_H value in the new rupture location (gray star in Figure 3). This approach is similarly applied when ruptures exceed the upper boundary layer. Finally, only the uniform probability values that meet the rupture criterion are summed.

The adjusted HDR distributions were compared with empirical observations of HDRs obtained from earthquake source databases to test the pertinence of our modeling approach (see section “How consistent are observed and adjusted HDRs?”).

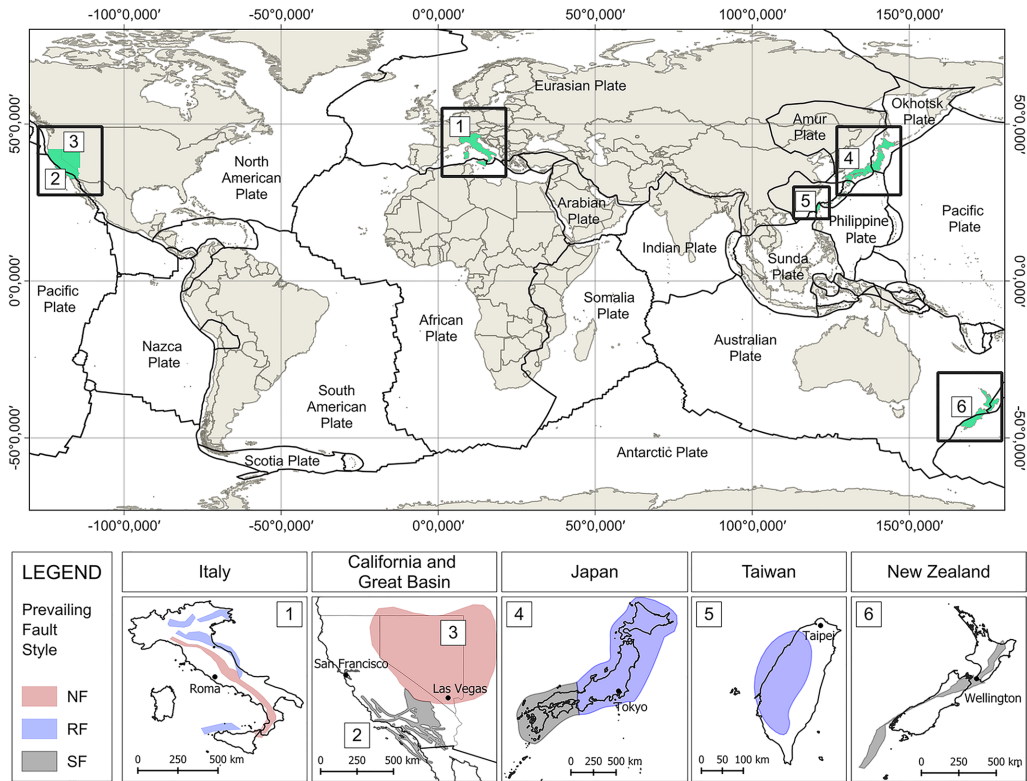


Figure 4. Map of continents and major tectonic plates (Bird, 2003). Black squares include the regions examined in this study; the prevailing fault style is indicated by different colors: NF = normal faulting; RF = reverse faulting; SF = strike-slip faulting. (1) Italy; (2) California; (3) Great basin; (4) Japan; (5) Taiwan; and (6) New Zealand.

Data used to estimate HDD and HDR

Seismotectonic zones

We distinguished eight STZs according to prevailing tectonic regimes within six global regions (Figure 4) as follows: (1) Italy, partitioned into two STZs, one predominantly normal and one predominantly reverse, with zoning according to Visini et al. (2021, 2022) (Figure 5a); (2) California, in which a predominantly strike-slip STZ was defined (Supplemental Figure S1—Supplementary material); (3) the Great Basin, which is a STZ with predominantly normal faulting (Supplemental Figure S2); (4) Japan, which is partitioned into a southern STZ of predominantly strike-slip faulting and a northern STZ of predominantly reverse faulting (Supplemental Figure S3); (5) Taiwan, whose western portion is characterized as predominantly reverse faulting (Supplemental Figure S4); and (6) New Zealand, which has a STZ predominantly characterized by strike-slip faulting as zoned in Seebeck et al. (2022, 2023) and Thingbaijam et al. (2024) (Supplemental Figure S5).

In cases where existing seismotectonic zoning based on predominant faulting styles was unavailable, we used alternative methods to define the zones. For California's strike-slip faults, we utilized a 10-km buffer around the traces of all active strike-slip faults according to SCEC Community Fault Model (CFM) (Plesch et al., 2007). The Great Basin area was

treated as a single normal faulting area, based on fault (US Geological Survey, 2023), stress (Heidbach et al., 2016), and geodetic (Hammond et al., 2024) data. For Taiwan, we analyzed the central-western part of the island (Coastal Plain, Western Foothills, Pingtung Plain, and Hsuenhsan Range), characterized by dominant contraction and reverse faulting (Chen et al., 2017; Franklin and Huang, 2022; Lin et al., 2010). We excluded from the analysis the Central Range, characterized by mixed focal mechanisms (Heidbach et al., 2022), and the contractional zone along the east coast (Coastal Range), due to the difficulty in separating crustal from subduction events. For Japan, we relied on the stress map provided by Uchide et al. (2022) identifying homogeneous seismotectonic areas for reverse and strike-slip faulting. Northern Japan is characterized by prevailing reverse faulting, while the southern area exhibits predominantly strike-slip faulting.

Earthquake catalogs

To analyze the HDD within each STZ, we selected databases of relocated seismicity based on their completeness and accuracy. For our analysis, we selected the following: the CLASS catalog (Latorre et al., 2023) for Italy; the Hauksson et al. (2012) catalog (up to 2023) for California; the Mueller (2019) catalog for the Great Basin; the JUICE catalog (Yano et al., 2017) for Japan; the Wu et al. (2008) catalog for Taiwan; and the Eberhart-Philips and Reyners (2022) catalog for New Zealand.

As mentioned in the “Hypocentral depth distribution” section, we filtered the earthquake catalogs to exclude events below the shallow continental crust by using numerical models of Moho depths where such data were available. The numerical models utilized for the lateral variation of Moho depth are as follows: Di Stefano et al. (2011) for Italy; Li et al. (2022) for California; Farag et al. (2022) for Japan. Fixed values were used to cut the earthquake depths where specific numerical data were unavailable: for the Great Basin, we employed a value of 30 km in accordance with Shehata and Mizunaga (2022); for Taiwan, a value of 40 km was utilized based on Goyal and Hung (2021); for New Zealand, a value of 40 km was adopted based on Ellis et al. (2024).

Earthquake source databases for observed HDR

The SRCMOD database is a collection of finite-fault rupture models that provide detailed representations of earthquake rupture surfaces. These models are developed using various data sources, including geological, geodetic, and seismological data, to better understand the geometry and kinematics of faults involved in earthquakes. To calculate the HDR, information such as hypocentral depth, rupture width, fault dip angle, and slip distribution is required, and all of this data is included in the SRCMOD database.

The NGA-W2 (Next Generation Attenuation for Western US) database (Ancheta et al., 2014) is an expansive collection of recorded ground motions utilized for developing and validating ground-motion prediction equations used in seismic hazard analysis. The database includes seismic parameters useful for calculating HDR, including hypocentral depth, rupture width, and fault dip angle.

Data processing

This section presents the data processing for the computation of HDD and observed HDR.

Hypocentral depth distribution

We conducted HDD calculations for the eight STZs illustrated in Figure 4. In this section, we use the STZs in Italy to illustrate our processing methods. Results pertaining to the remaining STZs are provided in the Supplementary material. Figure 5a illustrates the extensional (red polygon) and compressional areas (blue polygons) according to the seismic zoning defined by Visini et al. (2022). Gray circles represent earthquakes with quality factor (qf) AB in the CLASS catalog (CLASS-AB). These data adhere to formal vertical error thresholds of 82% ≤ 4 km, and 50% ≤ 2 km, which is suitable for the objectives of this study and correspond to excellent and good quality according to Latorre et al. (2023). Additional filtering of the CLASS-AB catalog was performed with the criteria that vertical and horizontal errors exceeding 5 and 3 km, respectively, were removed.

The algorithm to compute the HDD is configured to select a minimum of 25 earthquakes per 20×20 km square. Typically, each square contains hundreds of earthquakes, ensuring robust statistical analysis. Figure 5b illustrates an example of earthquake depth distribution within the highlighted black square in Figure 5a. The red line corresponds to the calculated D90 for the same square, while the black line represents the average Moho depth. To represent the HDD in the model, we assumed that a normal distribution would be most appropriate for each STZ. A Gaussian distribution with an upper truncation at $HDD = 1$ was applied to the histograms. Figure 5c displays the histogram depicting the normalized frequency depth distribution for the entire Italian extensional area, obtained by staking the histograms of each square.

Figure 6 illustrates the normal distributions calculated for the HDD across all eight STZs examined in the study. Our analysis reveals a tendency for rupture to nucleate in the deeper sections of the seismogenic crust regardless of fault style. Specifically, we observe a preference for deep nucleation in regions such as Japan and New Zealand, with peak values of 0.79 and 0.72, respectively. Conversely, in areas like Taiwan, California, Great Basin, and Italy, nucleation predominantly occurs in the low-to-mid sections of the seismogenic crust, with peaks around 0.60–0.67. Understanding whether these differences arise from variations in seismic network configuration and localization procedures or are due to specific seismotectonic setting needs additional investigations that lie beyond the scope of this study. In addition, we provide three aggregate HDDs in Supplemental Table S1 categorized by fault style. This may offer a general selection for users who want to consider global average conditions instead of or in addition to zone-specific distributions.

Hypocentral depth ratio

The following sections offer a comprehensive analysis of the two global databases used to compute the observed HDR, namely SRCMOD and NGA-W2 databases. The aim is to compare the adjusted HDR, obtained by reallocating rupture events extending beyond the upper and lower boundary layers of the model, with the observed HDR. The comparison is discussed in detail in section “How consistent are observed and adjusted HDRs?”

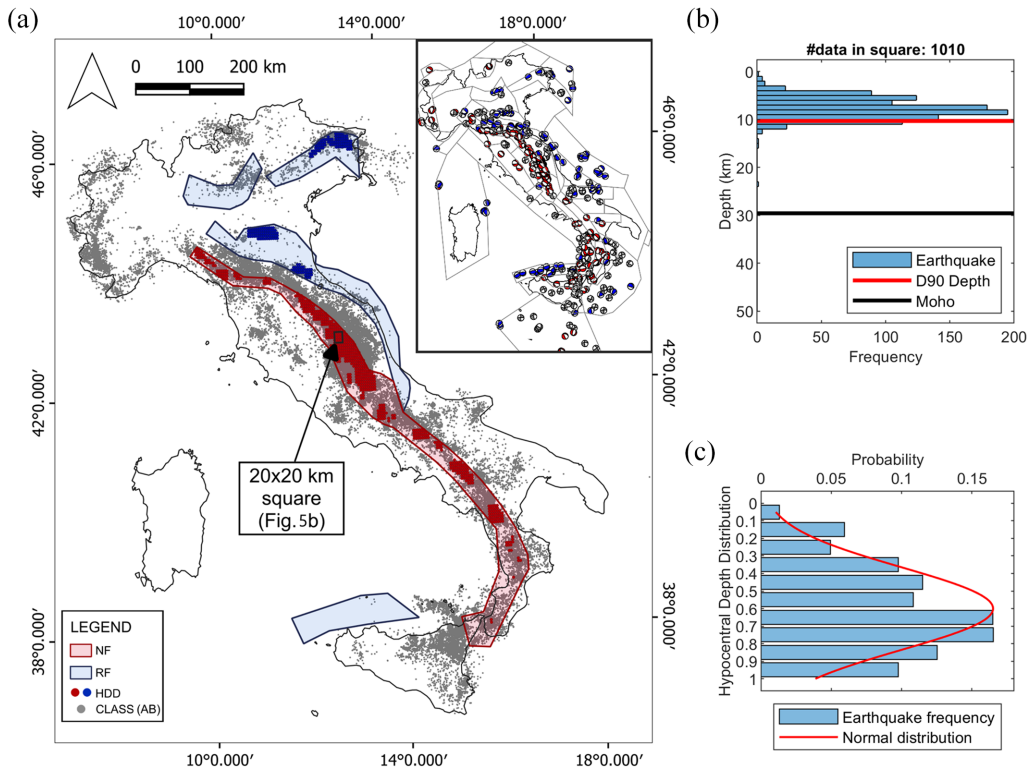


Figure 5. (a) Epicentral distribution of the CLASS catalog in Italy (gray circles) filtered by qf AB. Zones with dominant normal faulting (NF) and reverse faulting (RF) regimes are indicated by red and blue polygons, respectively. Red and blue circles correspond to grid nodes that meet the calculation conditions for HDD (20×20 km squares with number of events ≥ 25). Inset at top right: gray solid lines are the boundaries of seismotectonic zones in Visini et al. (2022); focal mechanisms are from RCMT catalog (Pondrelli, 2002): red = normal, blue = reverse, gray = strike-slip. (b) Example of histogram showing the hypocentral depths that occur in the black square; the red line represents D90, while the black line indicates the average depth of the Moho within the square. The histogram and the corresponding D90 were computed for all the red and blue circles in the zones. (c) Results of the HDD normalized to D90 for the normal faulting regime. The red line depicts a normal distribution fitted to the observed data (blue bars).

SRCMOD data processing. We selected events from the SRCMOD database since 1906 with $m \geq 5$. To aggregate data according to fault style, we applied a classification scheme using the rake angle, defining four categories: (1) normal faults for rake ranging from -120° to -60° ; (2) reverse faults for rake between 60° and 120° ; (3) strike-slip faults for rake between -30° and 30° , $>150^\circ$, or $<-150^\circ$; and (4) oblique faults for rake in the ranges -60° to -30° , -120° to -150° , 120° to 150° , and 30° to 60° . Within the category of oblique faulting, we manually assigned each event to strike-slip, reverse, or normal styles of faulting based on a supplementary literature review. This process involved reallocating events with oblique rake angles to one of the end-member faulting styles, as well as resolving conflicting events with inaccurate rake angles. For instance, in the SRCMOD database, there are seven reference tags corresponding to the earthquake in Turkey on 17 August 1999. Among these references, the study by Sekiguchi and Iwata (2002) stands out because it interprets a rake

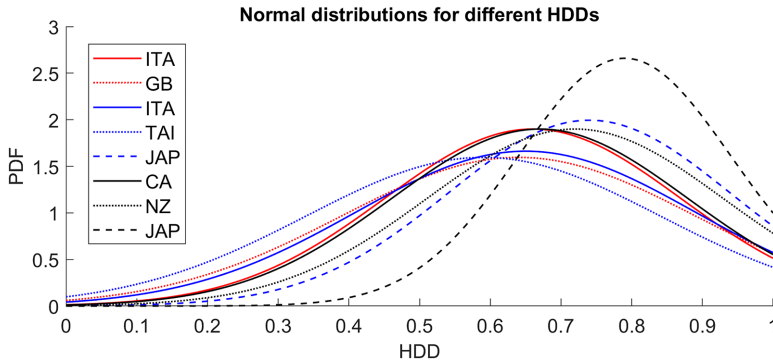


Figure 6. Comparison of the hypocentral depth distributions (HDDs), represented as probability density function (PDF), calculated for eight homogeneous seismotectonic zones with different fault regimes (red = normal faulting; blue = reverse faulting; black = strike-slip faulting). ITA = Italy, GB = Great Basin, TAI = Taiwan, JAP = Japan, CA = California, NZ = New Zealand.

angle of 75° . Consequently, it falls within our classification of oblique faults. In this specific case, the fault style has been updated to strike-slip because the event is characterized by strike-slip kinematics according to most publications (e.g. Bouchon et al., 2002; Cakir et al., 2003; Delouis et al., 2002; Hayes, 2017; Tibi et al., 2001). We classified both continental and oceanic faults during the analysis process. However, for the HDR computation, only continental faults were considered, with subduction, outer-rise, and oceanic transforms excluded.

For each event, the SRCMOD database contains one (single) or more (multiple) event solutions. We performed a sensitivity analysis of key geometric parameters—such as rupture width, dip angle, depth, and rake angle—and found significant variability among these parameters. Based on this analysis, we selected the most suitable candidate from the multiple solutions considering various techniques and data used by different research teams in the study of the same seismic event. Our preference was for rupture models with joint methodologies, which integrate multiple data sources and approaches. When this option was not available, or when there were multiple models based on joint methodologies, we selected the model constrained by the highest number of stations.

A common concern about rupture dimensions derived from finite-fault modeling approaches is the potential overestimation of the rupture area, which can occur due to low or even zero slip at the adjusted fault edges. To mitigate this issue, we adopted the trimming approach proposed by Somerville et al. (1999) to remove rows from the slip model if their average displacement was less than 0.3 times the overall average slip. Within the database, there are instances of multi-segment events. In such cases, we considered only the information contained in the segment including the hypocenter. Among the events selected for HDR computation, some already reached the ground surface before trimming was applied. However, this did not pose any calculation issues, as the HDR was consistently determined by trimming both the total width (W) and the upper width (W_{top}), ensuring the stability of the HDR value.

NGA-W2 data processing. The NGA-W2 catalog contains pre-processed data, obviating the need for a separate analysis of fault parameters. The database includes only one

interpretation of style of faulting and fault rupture parameters for each earthquake. Only data declared to be derived from source-inversion models were chosen. Slip trimming was unnecessary, as custom investigations were conducted for each case by NGA-W2 compilers with expert consultation (Ancheta et al., 2014; Bozorgnia et al., 2014; Chiou et al., 2008; Chiou and Youngs, 2008).

The comparison of observed HDR between the SRCMOD and NGA-W2 databases is illustrated in panel a of Figure 7 (histograms 1 to 6). From the SRCMOD database, we analyzed 24 normal, 29 reverse, and 51 strike-slip events, covering a magnitude range from Mw 5.0 to 7.9. The NGA-W2 database includes 13 normal, 28 reverse, and 36 strike-slip events, with magnitudes ranging from Mw 5.2 to 7.9. There is an overlap of 8% for normal, 16% for reverse, and 16% for strike-slip earthquakes, indicating the percentage of events that are common to both databases. For the SRCMOD database, the mean and standard deviation for HDR values for normal, reverse, and strike-slip faults are 0.59 ± 0.13 , 0.59 ± 0.21 , and 0.57 ± 0.22 , respectively. Events in the NGA-W2 database have HDR mean and standard deviation values for normal, reverse, and strike-slip earthquakes of 0.85 ± 0.16 , 0.59 ± 0.28 , and 0.67 ± 0.23 , respectively. The greatest difference between datasets is in normal earthquakes, where there is a difference in mean values of 0.26 compared with a difference of 0.10 for strike-slip earthquakes and no difference for reverse earthquakes. Surprisingly, the normal earthquakes also had the lowest standard deviations of the three faulting styles within each database. No correlation was found between magnitude and HDR for either database (Supplemental Figure S6).

Kolmogorov–Smirnov (KS) tests were performed on each dataset to evaluate the similarity of the HDR distribution and determine whether the data could be grouped (e.g. dip-slip as representative of normal and reverse faults). For both databases, we compared pairs of HDR distributions as follows: (1) normal and reverse, (2) reverse and strike-slip, (3) normal and strike-slip, and (4) dip-slip (normal + reverse) and strike-slip. For the SRCMOD database, the KS test indicates that none of the compared pairs are significantly different. In contrast, for NGA-W2, the KS test shows a mix of significant and non-significant outcomes (e.g. normal faults and dip-slip HDR). Therefore, we opted to keep them separated based on faulting style.

How consistent are observed and adjusted HDRs?

Here, we refer back to the processing method introduced in the section “Hypocentral depth ratio and adjusted HDR” and check if the adjusted HDRs are consistent with observations contained in the SRCMOD and NGA-W2 databases. Panel b of Figure 7 illustrates the outcomes for the adjusted HDR categorized into three distinct normalized hypocenter depth classes (HDD, Figure 1b): Z_H/Z_S between 0.1 and 0.3, Z_H/Z_S between 0.4 and 0.6, and Z_H/Z_S between 0.7 and 0.9, referring, respectively, to shallow, intermediate, and deep nucleation depths.

We were unable to determine the normalized hypocentral depth for each event in the SRCMOD and NGA-W2 databases due to the lack of Z_S information. For comparison purposes, we computed a pseudo-HDD (Figure 7, panel a, histograms 7–12), as a proxy of the normalized HDD calculated in our study. Pseudo-HDD was defined as the ratio of hypocentral depth to the total distance between the bottom of the rupture and the surface, enabling standardized comparisons of seismic depths across different events.

To compare the observed HDR (Figure 7, panel a, histograms 1–6) with the adjusted HDR (Figure 7, panel b) for each fault style, we first examine the three adjusted HDRs—grouped by depth classes—with respect to the observed HDR distribution. Next, we identify the most representative depth class (Z_H/Z_s in Figure 7b). Finally, we evaluate the pseudo-HDD (Figure 7, panel a, histograms 7–12).

For instance, consider the comparison for normal faulting (first column of Figure 7) between the observed HDR from the SRCMOD database (Figure 7a-1) and the adjusted HDR (Figure 7b-1,4,7). Among the three depth classes—shallow (Figure 7b-1), intermediate (Figure 7b-4), and deep (Figure 7b-7)—the one most compatible with the observed HDR in Figure 7a-1 lies between the intermediate and deep classes. This is because the observed HDR shows a peak for central values (i.e. 0.5), with the remaining observations tending toward the bottom of the rupture, that is, HDR values close to 1. Therefore, we envision an adjusted distribution of HDR that falls between Figure 7b-4 and Figure 7b-7, corresponding to a medium-deep depth class. Then we compare it with the Pseudo-HDD. This allows us to verify the consistency with the hypocentral depth derived as a proxy from the same dataset used to calculate the observed HDR. In this case, the adjusted result is found to be compatible with the observed one. A similar discussion can be made when comparing the observed HDR from NGA-W2 (Figure 7a-4). In this case, the correspondence is found for deep class, which is consistent with the pseudo-HDD shown in Figure 7a-10. Analogous comparisons can be made for reverse and strike-slip faults for both databases. Overall, the consistency observed between the representative class of Pseudo-HDD and the depth class associated with the adjusted HDR support our modeling approach.

Note that shallow values estimated for the adjusted HDR (i.e. Figure 7b-1,2,3) are either absent or notably scarce in the observed data. However, in the calculation HDR values located in the shallow part of the seismogenic crust are counterbalanced by the HDD, which indeed exhibits very low probabilities for shallow values.

Results

In this section, we discuss the results for normal faults, using a general case study as an example (i.e. aggregates HDD for normal fault, Supplemental Table S1). The analyses for reverse (Supplemental Figure S7) and strike-slip faults (Supplemental Figure S8) are provided in the Supplementary materials. We compared CPSR curves for different fault angles and seismogenic thicknesses (Figure 8). The CPSR is calculated by aggregating cases that meet the *surface rupture condition* (i.e. when Wz_{TOP} reach the surface—Figure 1a).

The curves shown in Figure 8 were computed using the L14 and T17 magnitude scaling relationships across a magnitude range from 5 to 8, reflecting most of the data range provided by the authors. The CPSR curves exhibit a characteristic pattern: an initial gradual increase in surface rupture probability, which is less than 20% and rises slowly with M_w , followed by a steeper increase above a magnitude threshold. This threshold depends on the seismogenic thickness and fault dip, marking the transition from small to large ruptures.

For a seismogenic thickness (Z_s) of 15 km, using L14 relationships, the magnitude threshold ranges approximately between M_w 6.2 and 6.5 for fault dip angles ranging from 40° to 60° (Figure 8a). When using T17 relationships, this threshold shifts to approximately M_w 5.8 and 6.2. For a reduced seismogenic thickness of 10 km and an average fault dip of 50° , the magnitude threshold decreases to around $M_w \simeq 5.9$ for L14 and $M_w \simeq 5.4$ for T17 relationships (Figure 8b).

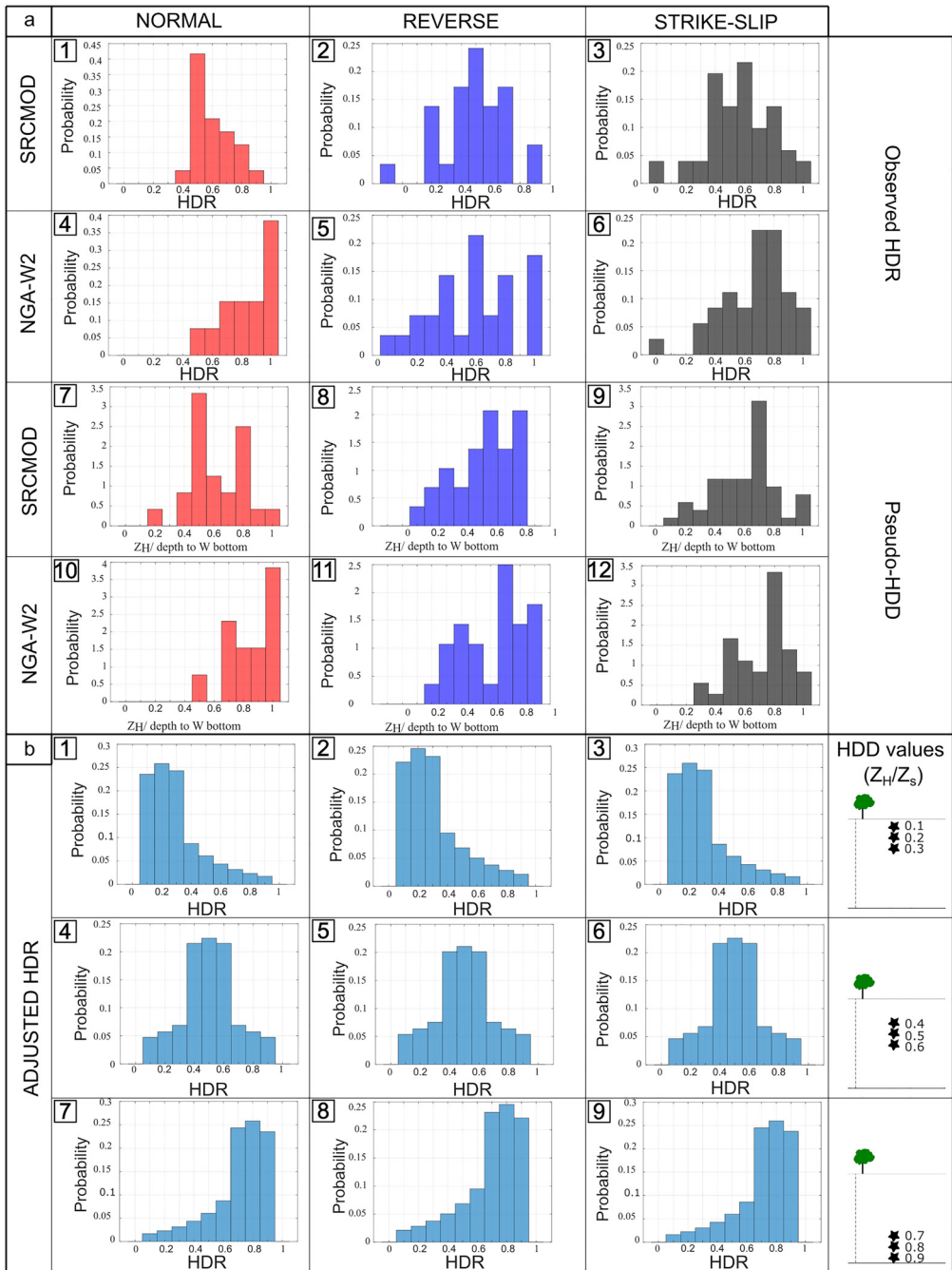


Figure 7. Comparison between observed and adjusted HDR based on fault style (normal, reverse, strike-slip) and depth. In panel a, histograms 1 to 6 depict results from finite-fault rupture models in SRCMOD and NGA-W2 databases, while histograms 7 to 12 represent Pseudo-HDD computed from the same databases, assuming the seismogenic depth extends to the bottom of the rupture. Panel b displays the adjusted HDR divided into shallow, intermediate, and deep depth classes, with corresponding illustrations of the seismogenic crust and hypocenter positions (black stars). Hypocentral depth values are normalized to the seismogenic depth.

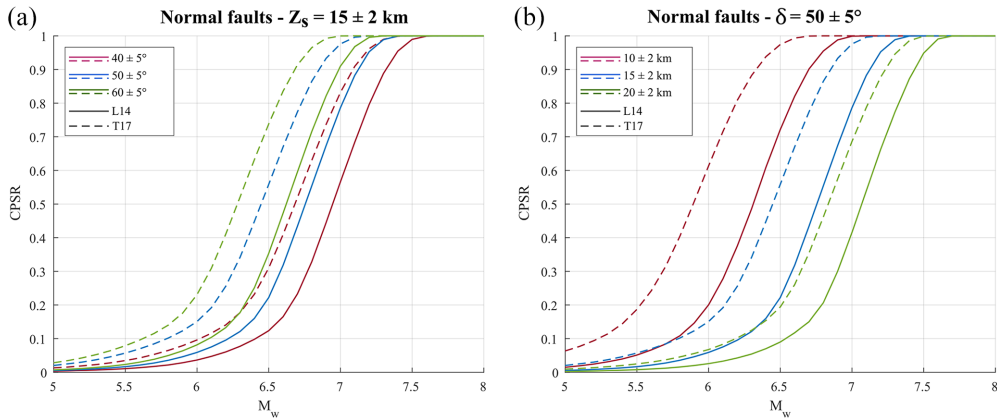


Figure 8. Comparative analysis of calculated CPSR curves for normal faults for (a) seismogenic thickness $Z_s = 15 \pm 2$ km and (b) dip angle $\delta = 50 \pm 5^\circ$. Solid lines (L14) refer to Leonard (2014) magnitude scaling relations and dashed lines (T17) refer to Thingbaijam et al. (2017) magnitude scaling relations. The distinct colors denote variations in (a) average fault dip angle and (b) average seismogenic depth. The dip angles were selected based on the fault angle analysis from the SRCMOD database, with an average dip angle δ of 50° .

As expected, the probability of surface rupture increases with fault dip angle (Figure 8a) and decreases with increasing seismogenic thickness (Figure 8b). The results differ significantly depending on the empirical relationships used. The CPSR estimated with L14 relationships lead to lower probability values than those relying on T17. This difference is more pronounced for shallower seismogenic depth.

Discussion

Our model is numerically driven and primarily based on geometric relationships, without incorporating the physics of rock behavior or stress variations. We recognize this as a limitation, as these shallow fault properties are key contributors to surface rupture (Dalguer et al., 2008; Kagawa et al., 2004) and are not explicitly considered in our approach. This research aims at improving the methodology and developing a numerical code for estimating the probability of having surface rupture along the trace of a fault rupture.

So far, purely empirical methods have been developed which tend to mix up information about the seismotectonic setting and/or fault geometry (e.g. Moss and Ross, 2011; Petersen et al., 2011; Pizza et al., 2023; Youngs et al., 2003). A previous case study of numerical approach can be ascribed to PG&E (2010) who developed an integrated approach for the creation of a probability curve for surface rupture on the Shoreline fault. They combined empirical data (e.g. Wells and Coppersmith, 1994) and numerical models imposing that for the maximum expected magnitude, the probability of surface rupture should be 97.5%. Thus, the model developed is exclusive to Shoreline case study. Moreover, while the approach developed by PG&E (2010) is to use a predefined rupture function to smooth the transition between the empirical data, our methodology deviates by considering ruptures of a range of magnitudes.

From the results obtained in this study, we observed how fault parameters affect the CPSR (Figure 8) and that the most influential parameter deals with the seismogenic depth

(Z_S). As a result, the investigations conducted for PFDHA should be region-specific developing probability curves according to the seismotectonic area of interest. In fact, there is a similarity between the curves resulting from the site-specific PG&E (2010) model and ours, and it lies in the slope of the probability curve, stating that the surface rupture probability curve is indeed a function of the seismogenic thickness.

The slope of the CPSR is also influenced by the HDD. While we have not specifically addressed this variation, as results across different seismotectonic areas are consistently confined to the deeper sections of the seismogenic crust (Figure 6), it might be valuable to enhance the investigations conducted for HDD. In this study, we applied a normal distribution to all examined STZs to model the calculated frequencies (Figure 5c). Future work can be developed to better account the most appropriate distribution for HDD calculation in different seismotectonic regions.

In Figure 9, we compare the CPSR curves from this study with the empirical regression for Japan (Figure 9a), New Zealand (Figure 9b), and Great Basin (Figure 9b), which are the only published empirical regressions. We also compare our and empirical regional curves with the most recent global empirical regression of PEA23 (Pizza et al., 2023). The input parameters for the CPSR curves are the seismogenic depth, corresponding to the average D90, and the HDD derived from the considered STZ (Supplemental Table S1). The average dip angle is the mean of the fault dip distribution extracted from the SRCMOD database. For simplicity, we used for the comparison only the magnitude scaling relations proposed by L14 and relative uncertainties.

In particular, Figure 9a shows the reverse case study comparing the empirical regression by TEA13 (Takao et al., 2013) with the CPSR computed in this study for Japan. Specifically, we used a seismogenic depth of 10 km (mean value from the computed D90) with an uncertainty of 2 km, and an average dip angle of $40^\circ \pm 5^\circ$. The regressions by TEA13 are also valid for strike-slip faults. Therefore, we conducted a similar analysis for strike-slip faults considering an average seismogenic depth of 11 km for the strike-slip zone of Japan with an uncertainty of 2 km, and an average dip angle of $80^\circ \pm 5^\circ$. Comparing the CPSR curves with the TEA13 regression reveals a good match, particularly for the reverse faulting case.

The CPSR curves were computed using parameters specific to the strike-slip STZ of Japan, derived from data analysis in this study. Since the empirical regressions by TEA13 are based on earthquakes occurring exclusively within Japan, the observed agreement between the numerical and empirical curves can be attributed to their regional characteristics. For the global empirical regressions by PEA23, a good agreement is observed only for strike-slip faults.

For the strike-slip regime, we conducted a further comparison with a study conducted for New Zealand. The mean D90 obtained for this region is deepest with respect to all the STZs (over 20 km); however, existing literature revealed consistency with our results (Ellis et al., 2024; Seebeck et al., 2023). Thus, to compute the CPSR (Figure 9b), we used a 27-km seismogenic depth according to the average D90, with 2 km of uncertainty and an average dip angle of $80^\circ \pm 5^\circ$. So far, there is no empirical regression developed for this area; thus, for comparison, we used the area of probability of surface rupture obtained by Nicol et al. (2016)—NEA16. This area encompasses all possible relationships for earthquakes in New Zealand and was derived using New Zealand historical earthquakes. We also conducted a sensitivity test (Z_S test in Figure 9b), considering a seismogenic depth of 21 km based on the average D90 computed in this study for all type of fault in New

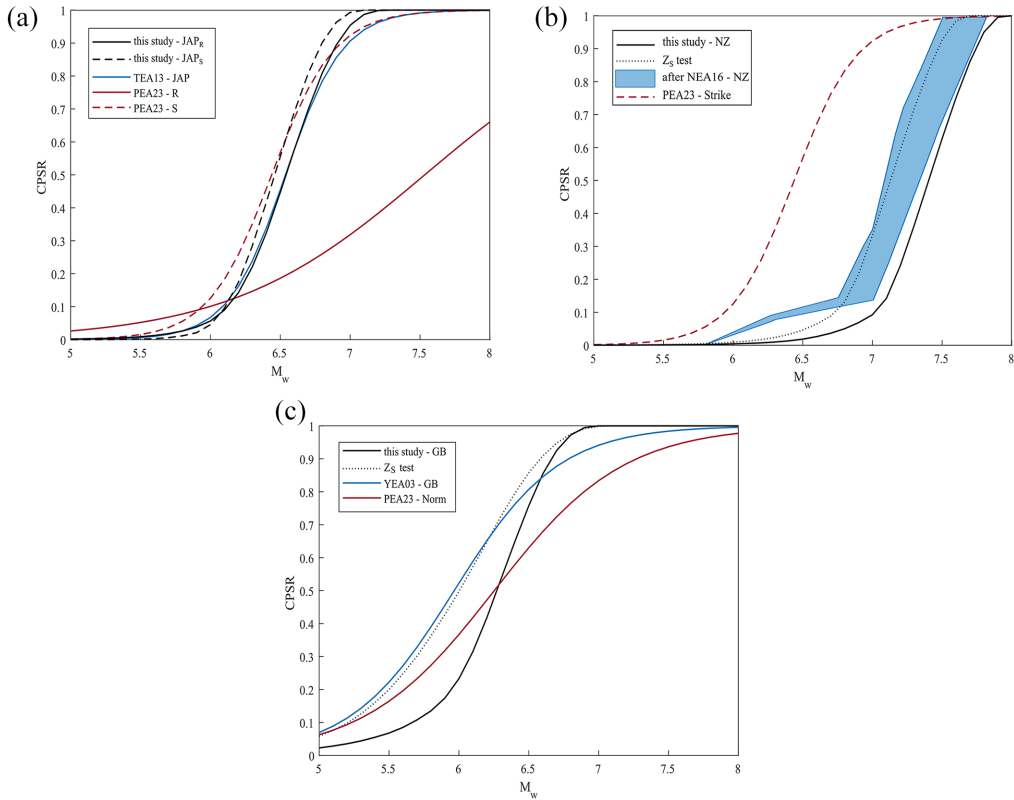


Figure 9. Comparison of CPSR curves computed in this study with empirical regressions available in the literature; the black lines represent the CPSR curves for (a) Japan (JAP), (b) New Zealand (NZ), and (c) Great Basin (GB). Light blue lines refer to the empirical regression by Takao et al. (2013)—TEA13 and Youngs et al. (2003)—YEA03; light blue area refers to the surface rupture probability area of New Zealand historical earthquakes (post 1845) after Nicol et al. (2016)—NEA16. “Z_S test” (dotted black line) refers to the CPSR curves calculated in this study, using alternative Z_S values. Red lines refer to the empirical regressions by Pizza et al. (2023)—PEA23.

Zealand. The two CPSR curves compared seem to delimit the area modeled by NEA16 for magnitudes greater than 6.7. According to NEA16, New Zealand historical surface ruptures have magnitudes greater than 7 and only one event characterized by a lower magnitude has generated surface rupture, which is associated with the Taupō volcanic zone, thus linked to very shallow seismogenic thicknesses. This suggests that the higher probability values predicted by NEA16 (Figure 9b) for magnitudes lower than 6.7 might be related to this rupture event characterized by reduced seismogenic depth. For this case study, the global curve of PEA23 is out of range and this is possibly related to the deepest seismogenic thickness that characterizes New Zealand.

The empirical regression developed by Youngs et al. (2003; YEA03) could be interpreted as regional empirical regression for normal faults since regressions include the extensional cordillera of Western United States and subdomains therein, such as the Northern Basin and Range and Great Basin (Supplemental Figure S9a). Figure 9c shows the comparison between the empirical regression for the GB by YEA03 and the CPSR

from this study. The continuous black line was computed considering a seismogenic depth of 11 km, according to the Great Basin average D90, and an uncertainty of 2 km. The average dip angle was set at $60^\circ \pm 5^\circ$. For this regional example, some discrepancies are observed between the empirical regression and the CPSR. This could be attributed to the large variability of the seismogenic depth in the Great Basin or shallow geological complexities. Examining the earthquake catalog and hypocentral depths of earthquakes in more detail, we identified a concentration of shallow earthquakes located along the Sierra Nevada volcanic zone. Therefore, we performed a sensitivity test changing Z_S and its uncertainty. The best fit was found by calculating a CPSR for the Great Basin including an uncertainty of 5 km and a seismogenic depth of 9 km (Z_S test in Figure 9c) and keeping the other parameters the same as the previous analysis. Our sensitivity test, which adjusted seismogenic depth and its uncertainty, highlights the importance of accurately defining region-specific parameters.

This comparison aims to illustrate how controlling parameters (e.g. Z_S) influence the CPSR and how integrating global data (e.g. PEA23 relations) with varying parameters leads to more refined empirical curves. CPSR curves tend to exhibit steeper slopes compared to most empirical regressions (Supplemental Figure S9). This discrepancy may stem from the mixing of diverse data (e.g. seismogenic depth, fault dips) in global empirical regressions. The analysis shows that while global models offer broad insights, they may not fully capture the local variations required for precise site-specific assessments.

So far, we have discussed the CPSR computed using the magnitude scaling relation derived by L14. However, as shown in Figure 8, our curves change with the scaling relation used. Those differences are probably related to the different methods used by the authors to derive the rupture width relation. Also, we have noticed that the discrepancies between the W value computed using the L14 and T17 relation increases with the magnitude. Further investigation into the specific characteristics of each calculation method is warranted to understand the source of these discrepancies and ensure the reliability of our results. In the model proposed herein, we have solely considered the width of the rupture, as our primary interest lies in understanding how frequently the rupture intersects the surface. It would be advantageous to enhance the model by incorporating the parameter of rupture length as well. Utilizing the scaled relationship of area to define the size of the rupture would be particularly insightful, as area relationships tend to provide better constraints.

Users can explore uncertainties by constructing a logic tree and generating multiple curves through variations in input parameters. This approach facilitates a comprehensive understanding of uncertainty, allowing users to visualize the breadth of possibilities associated with the calculated curve.

Conclusion

We have developed a methodology and numerical code to estimate the CPSR on a principal (seismogenic) fault for use in PFDHA. The CPSR is a function of earthquake magnitude, and the shape of the probability distribution is dependent on a model fault's seismogenic width and style of faulting. Our numerical approach, which was developed previously for ground-motion modeling (Chiou and Youngs, 2008; Youngs et al., 2003), utilizes a probabilistic framework for estimating hypocenter depths, rupture widths, and the position of earthquake ruptures on a modeled fault in the down-dip (width) dimension. The numerical method differs from purely empirical approaches that aggregate data

from different tectonic settings and face questions about data completeness. The most significant result of our analysis is the impact of seismogenic depth on the computed CPSR. This highlights the utility of the numerical approach in settings where the local fault dip and regional seismogenic thickness may be constrained by data, and especially where local conditions may deviate from global averages. The numerical approach may be used in a hazard evaluation as an epistemic alternative to, or replacement of, the empirical approaches. We have made our numerical code openly available as a MATLAB script on GitHub, enabling end-users to access and utilize it for their own analyses.

Acknowledgments

The authors wish to express their gratitude to Dr Jack Baker, Editor-in-Chief of *Earthquake Spectra*, and the two anonymous reviewers for their suggestions and comments, which have improved the quality of the manuscript. We thank the Fault Displacement Hazard Initiative (<https://www.risksciences.ucla.edu/nhr3/fdhi/home>) working group for fruitful discussions. The authors acknowledge the researchers who have shared their data, which were essential for conducting the analyses in this study.


Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was funded by the agreement Ref IRSN\LW 2023-0135\ between IRSN and University of Chieti—Pescara, “Fault displacement hazard on principal fault: probability of occurrence, slip distribution and role of surface geology,” which is funding the PhD research of L.M.

ORCID iD

Lisa Mammarella  <https://orcid.org/0009-0009-8414-1495>

Data and resources

All data utilized in this study are sourced from published papers and web links referenced in the manuscript. The current version of the CPSR code developed in this work is available on GitHub repository (<https://github.com/MammarellaLisa/CPSR>—last accessed September 2024).

Supplemental material

Supplemental material for this article is available online.

References

- Ancheta TD, Darragh RB, Stewart JP, Seyhan E, Silva WJ, Chiou BS-J, Wooddell KE, Graves RW, Kottke AR, Boore DM, Kishida T and Donahue JL (2014) NGA-West2 database. *Earthquake Spectra* 30(3): 989–1005.
- Bird P (2003) An updated digital model of plate boundaries. *Geochemistry, Geophysics, Geosystems* 4(3): 1027.

- Bouchon M, Toksoz MN, Karabulut H, Bouin M-P, Dietrich M, Aktar M and Edie M (2002) Space and time evolution of rupture and faulting during the 1999 Izmit (Turkey) earthquake. *Bulletin of the Seismological Society of America* 92(1): 256–266.
- Bozorgnia Y, Abrahamson NA, Al Atik L, Ancheta TD, Atkinson GM, Baker JW, Baltay A, Boore DM, Campbell KW, Chiou BS-J, Darragh R, Day S, Donahue J, Graves RW, Gregor N, Hanks T, Idriss IM, Kamai R, Kishida T, Kottke A, Mahin SA, Rezaeian S, Rowshandel B, Seyhan S, Shahi S, Shantz T, Silva W, Spudich P, Stewart JP, Watson-Lamprey J, Wooddell K and Youngs R (2014) NGA-West2 research project. *Earthquake Spectra* 30(3): 973–987.
- Cakir Z, De Chabaliere JB, Armijo R, Meyer B, Barka A and Peltzer G (2003) Coseismic and early post-seismic slip associated with the 1999 Izmit earthquake (Turkey), from SAR interferometry and tectonic field observations. *Geophysical Journal International* 155(1): 93–110.
- Chen SK, Wu YM, Hsu YJ and Chan Y-C (2017) Current crustal deformation of the Taiwan Orogen reassessed by cGPS strain-rate estimation and focal mechanism stress inversion. *Geophysical Journal International* 210(1): 228–239.
- Chiou BS-J and Youngs RR (2008) An NGA model for the average horizontal component of peak ground motion and response spectra. *Earthquake Spectra* 24(1): 173–215.
- Chiou BS-J, Darragh R, Gregor N and Silva W (2008) NGA project strong-motion database. *Earthquake Spectra* 24(1): 23–44.
- Cornell CA (1968) Engineering seismic risk analysis. *Bulletin of the Seismological Society of America* 58: 1583–1606.
- Dalguer LA, Miyake H, Day SM and Irikura K (2008) Surface rupturing and buried dynamic-rupture models calibrated with statistical observations of past earthquakes. *Bulletin of the Seismological Society of America* 98(3): 1147–1161.
- Delouis B, Giardini D, Lundgren P and Salichon J (2002) Joint inversion of InSAR, GPS, teleseismic, and strong-motion data for the spatial and temporal distribution of earthquake slip: Application to the 1999 Izmit mainshock. *Bulletin of the Seismological Society of America* 92(1): 278–299.
- Di Stefano R, Bianchi I, Ciaccio MG, Carrara G and Kissling E (2011) Three-dimensional Moho topography in Italy: New constraints from receiver functions and controlled source seismology. *Geochemistry, Geophysics, Geosystems* 12(9): Q09006.
- Eberhart-Phillips D and Reyners M (2022) New Zealand earthquake catalogue 2001-2011 relocated with 3-D velocity model nzwide2.3 (update 2023). *Zenodo*. Epub ahead of print 1 June. DOI: 10.5281/zenodo.6604627.
- Ellis S, Bannister S, Dissen R, Eberhart-Phillips D, Boulton C, Reyners M, Funnell R, Mortimer N, Upton P, Rollins C and Seebeck H (2024) New Zealand fault-rupture depth model v.1.0: A provisional estimate of the maximum depth of seismic rupture on New Zealand's active faults. *Bulletin of the Seismological Society of America* 114(1): 78–94.
- Farag T, Sobh M and Mizunaga H (2022) 3D constrained gravity inversion to model Moho geometry and stagnant slabs of the Northwestern Pacific plate at the Japan Islands. *Tectonophysics* 829: 229297.
- Franklin KR and Huang MH (2022) Revealing crustal deformation and strain rate in Taiwan using InSAR and GNSS. *Geophysical Research Letters* 49(21): e2022GL101306.
- Goyal A and Hung SH (2021) Lateral variations of Moho depth and average crustal properties across the Taiwan Orogen from *H-V* stacking of P and S receiver functions. *Geochemistry, Geophysics, Geosystems* 22(3): e2020GC009527.
- Gutenberg B and Richter CF (1944) Frequency of earthquakes in California. *Bulletin of the Seismological Society of America* 34(4): 185–188.
- Hammond WC, Kreemer C and Blewitt G (2024) Robust imaging of fault slip rates in the Walker Lane and Western Great Basin from GPS data using a multi-block model approach. *Journal of Geophysical Research: Solid Earth* 129(3): e2023JB028044.
- Hauksson E and Meier MA (2019) Applying depth distribution of seismicity to determine thermo-mechanical properties of the seismogenic crust in Southern California: Comparing lithotectonic blocks. *Pure and Applied Geophysics* 176(3): 1061–1081.

- Hauksson E, Yang W and Shearer PM (2012) Waveform relocated earthquake catalog for Southern California (1981 to June 2011). *Bulletin of the Seismological Society of America* 102(5): 2239–2244.
- Hayes GP (2017) The finite, kinematic rupture properties of great-sized earthquakes since 1990. *Earth and Planetary Science Letters* 468: 94–100.
- Heidbach O, Rajabi M, Reiter K, Ziegler M and WSM Team (2016) World Stress Map Database Release 2016. V. 1.1. GFZ Data Services. Available at: <https://doi.org/10.5880/WSM.2016.001> (accessed September 2023)
- Heidbach O, Liang W.-T, Morawietz S, von Specht S and Ma K-F (2022) Stress Map of Taiwan 2022, GFZ Data Services, doi.org/10.5880/WSM.Taiwan2022.
- Kagawa T, Irikura K and Somerville PG (2004) Differences in ground motion and fault rupture process between the surface and buried rupture earthquakes. *Earth, Planets and Space* 56: 3–14.
- Latorre D, Di Stefano R, Castello B, Michele M and Chiaraluce L (2023) An updated view of the Italian seismicity from probabilistic location in 3D velocity models: The 1981–2018 Italian catalog of absolute earthquake locations (CLASS). *Tectonophysics* 846: 229664.
- Leonard M (2014) Self-consistent earthquake fault-scaling relations: Update and extension to stable continental strike-slip faults. *Bulletin of the Seismological Society of America* 104(6): 2953–2965.
- Lettis WR, Wells DL and Baldwin JN (1997) Empirical observations regarding reverse earthquakes, blind thrust faults, and quaternary deformation: Are blind thrust faults truly blind? *Bulletin of the Seismological Society of America* 87(5): 1171–1198.
- Li T, Yao J, Wu S, Xu M and Tong P (2022) Moho complexity in Southern California revealed by local PmP and teleseismic Ps waves. *Journal of Geophysical Research: Solid Earth* 127(2): e2021JB023033.
- Lin KC, Hu JC, Ching KE, Angelier J, Rau R-J, Yu S-B, Tsai C-H, Shin T-C and Huang M-H (2010) GPS crustal deformation, strain rate, and seismic activity after the 1999 Chi-Chi earthquake in Taiwan. *Journal of Geophysical Research: Solid Earth* 115(7): B07404.
- Maggini M and Caputo R (2021) Seismological data versus rheological modelling: Comparisons across the Aegean Region for improving the seismic hazard assessment. *Journal of Structural Geology* 145: 104312.
- Mai PM and Thingbaijam KKS (2014) SRCMOD: An online database of finite-fault rupture models. *Seismological Research Letters* 85(6): 1348–1357.
- Marconato L, Leloup PH, Lasserre C, Jolivet R, Caritg S and Grandin R (2022) Insights on fault reactivation during the 2019 November 11, Mw 4.9 Le Teil earthquake in southeastern France, from a joint 3-D geological model and InSAR time-series analysis. *Geophysical Journal International* 229(2): 758–775.
- Moss RES and Ross ZE (2011) Probabilistic fault displacement hazard analysis for reverse faults. *Bulletin of the Seismological Society of America* 101(4): 1542–1553.
- Moss RES, Stanton KV and Buelna MI (2013) The impact of material stiffness on the likelihood of fault rupture propagating to the ground surface. *Seismological Research Letters* 84(3): 485–488.
- Mueller CS (2019) Earthquake catalogs for the USGS national seismic hazard maps. *Seismological Research Letters* 90(1): 251–261.
- Nicol A, Van Dissen RJ, Stirling MW and Gerstenberger MC (2016) Completeness of the paleoseismic active-fault record in New Zealand. *Seismological Research Letters* 87(6): 1299–1310.
- Omuralieva AM, Hasegawa A, Matsuzawa T, Nakajima J and Okada T (2012) Lateral variation of the cutoff depth of shallow earthquakes beneath the Japan Islands and its implications for seismogenesis. *Tectonophysics* 518–521: 93–105.
- Otsuka M (1964) Earthquake magnitude and surface fault formation. *Journal of Physics of the Earth* 12: 19.
- Pacific Gas and Electric Company (PG&E) (2010) *Evaluation of secondary fault rupture hazard from the Shoreline Fault Zone*. PG&E Letter no. DCL-10, 26 February. Washington, DC: US Nuclear Regulatory Commission.

- Petersen MD, Dawson TE, Chen R, Cao T, Wills CJ, Schwartz DP and Frankel AD (2011) Fault displacement hazard for strike-slip faults. *Bulletin of the Seismological Society of America* 101(2): 805–825.
- Pezzopane SK and Dawson TE (1996) *Fault Displacement Hazard: A Summary of Issues and Information, Seismotectonic Framework and Characterization of Faulting at Yucca Mountain*. Reston, VA: US Geological Survey.
- Pizza M, Ferrario MF, Thomas F, Tringali G and Livio F (2023) Likelihood of primary surface faulting: Updating of empirical regressions. *Bulletin of the Seismological Society of America* 113: 2106–2118.
- Plesch A, Shaw JH, Benson C, Bryant WA, Carena S, Cooke M, Dolan J, Fuis G, Gath E, Grant L, Hauksson E, Jordan T, Kamerling M, Legg M, Lindvall S, Magistrale H, Nicholson C, Niemi N, Oskin M, Perry S, Planansky G, Rockwell T, Shearer P, Sorlien C, Peter Süss M, Suppe J, Treiman J and Yeats R (2007) Community Fault Model (CFM) for southern California. *Bulletin of the Seismological Society of America* 97(6): 1793–1802.
- Pondrelli S (2002) *European-Mediterranean Regional Centroid-Moment Tensors Catalog (RCMT)* (up to February 2023). Istituto Nazionale di Geofisica e Vulcanologia (INGV). Available at: <https://doi.org/10.13127/rcmt/euromed> (accessed September 2023).
- Ritz JF, 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): 10.
- Seebeck H, Van Dissen R, Litchfield N, Barnes P, Nicol A, Langridge R, Barrell DJA, Villamor P, Ellis S, Rattenbury M, Bannister S, Gerstenberger M, Ghisetti F, Sutherland R, Fraser J, Nodder S, Stirling M, Humphrey J, Bland K, Howell A, Mountjoy J, Moon V, Stahl T, Spinardi F, Townsend D, Clark K, Hamling I, Cox S, De Lange W, Wopereis P, Johnston M, Morgenstern R, Coffey G, Eccles JD, Little T, Fry B, Griffin J, Townend J, Mortimer N, Alcaraz S, Massiot C, Rowland J, Muirhead J, Upton P, Hirschberg H and Lee J (2022) *New Zealand Community Fault Model—Version 1.0*. GNS Science report 2021/57, July 2022. Lower Hutt, New Zealand: GNS Science, 96 pp. (July 2022).
- Seebeck HR, Van Dissen N, Litchfield P, Barnes P, Nicol A, Langridge R, Barrell DJA, Villamor P, Ellis S, Rattenbury M, Bannister S, Gerstenberger M, Ghisetti F, Sutherland R, Fraser J, Nodder S, Stirling M, Humphrey J, Bland K, Howell A, Mountjoy J, Moon V, Stahl T, Spinardi F, Townsend D, Clark K, Hamling I, Cox S, De Lange W, Wopereis P, Johnston M, Morgenstern R, Coffey G, Eccles JD, Little T, Fry B, Griffin J, Townend J, Mortimer N, Alcaraz S, Massiot C, Rowland J, Muirhead J, Upton P, Hirschberg H and Lee J (2023) *New Zealand Community Fault Model—Version 1.0: An improved geological foundation for seismic hazard modelling*. *New Zealand Journal of Geology and Geophysics*. Epub ahead of print 19 March. DOI: 10.1080/00288306.2023.2181362.
- Sekiguchi H and Iwata T (2002) Rupture process of the 1999 Kocaeli, Turkey, earthquake estimated from strong-motion waveforms. *Bulletin of the Seismological Society of America* 92(1): 300–311.
- Shehata MA and Mizunaga H (2022) Moho depth and tectonic implications of the western United States: Insights from gravity data interpretation. *Geoscience Letters* 9(1): 23.
- Sibson RH (1982) Fault zone models, heat flow, and the depth distribution of earthquakes in the continental crust of the United States. *Bulletin of the Seismological Society of America* 72(1): 151–163.
- Somerville P, Irikura K, Graves R, Sawada S, Wald D, Abrahamson NA, Iwasaki Y, Kagawa T, Smith N and Kowada A (1999) Characterizing crustal earthquake slip models for the prediction of strong ground motion. *Seismological Research Letters* 70(1): 59–80.
- Takao M, Tsuchiyama J, Annaka T and Kurita T (2013) Application of probabilistic fault displacement hazard analysis in Japan. *Journal of Japan Association for Earthquake Engineering* 13: 17–36.
- Thingbaijam KKS, Mai PM and Goda K (2017) New empirical earthquake source-scaling laws. *Bulletin of the Seismological Society of America* 107(5): 2225–2246.

- Thingbaijam KKS, Rattenbury MS, Van Dissen RJ, Gerstenberger MC, Ristau J and Fitzenz DD (2024) Characterization of focal mechanisms for upper crustal distributed seismicity in Aotearoa New Zealand. *Seismological Research Letters* 95(1): 150–158.
- Tibi R, Bock G, Xia Y, Baumbach M, Grosser H, Milkereit C, Karakisa S, Zünbül S, Kind R and Zschau J (2001) Rupture processes of the 1999 August 17 Izmit and November 12 Düzce (Turkey) earthquakes. *Geophysical Journal International* 144: F1–F7.
- Uchide T, Shiina T and Imanishi K (2022) Stress map of Japan: Detailed nationwide crustal stress field inferred from focal mechanism solutions of numerous microearthquakes. *Journal of Geophysical Research: Solid Earth* 127(6): e2022JB024036.
- US Geological Survey (2023) Earthquake Hazards Program, 18 March. Available at: <https://www.usgs.gov/programs/earthquake-hazards/faults> (accessed September 2023)
- Valentini A, Fukushima Y, Contri P, Ono M, Sakai T, Thompson SC, Viallet E, Annaka T, Chen R, Moss RES, Petersen MD, Visini F and Youngs RR (2021) Probabilistic Fault Displacement Hazard Assessment (PFDHA) for nuclear installations according to IAEA safety standards. *Bulletin of the Seismological Society of America* 111(5): 2661–2672.
- Visini F, Meletti C, Rovida A, D’Amico V, Pace B and Pondrelli S (2022) An updated area-source seismogenic model (MA4) for seismic hazard of Italy. *Natural Hazards and Earth System Sciences* 22(8): 2807–2827.
- Visini F, Pace B, Meletti C, Marzocchi W, Akinci A, Azzaro R, Barani S, Barberi G, Barreca G, Basili R, Bird P, Bonini M, Burrato P, Busetti M, Cosimo Carafa MM, Cocina O, Console R, Corti G, D’Agostino N, D’Amico S, D’Amico V, Dal Cin M, Falcone G, Fracassi U, Gee R, Kastelic V, Lai CG, Langer H, Maesano FE, Marchesini A, Martelli L, Monaco C, Murru M, Peruzza L, Poli ME, Pondrelli S, Rebez A, Rotondi R, Rovida A, Sani F, Santulin M, Scafidi D, Selva J, Slejko D, Spallarossa D, Tamaro A, Tarabusi G, Taroni M, Tiberti MM, Tusa G, Tuvè T, Valensise G, Vannoli P, Varini E, Zanferrari A and Zuccolo E (2021) Earthquake rupture forecasts for the mps19 seismic hazard model of Italy. *Annals of Geophysics* 64(2): SE220.
- Wells D and Coppersmith KJ (1993) Likelihood of surface rupture as a function of magnitude.
- Wells D and Coppersmith KJ (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.
- Wu YM, Chang CH, Zhao L, Teng T-L and Nakamura M (2008) A comprehensive relocation of earthquakes in Taiwan from 1991 to 2005. *Bulletin of the Seismological Society of America* 98(3): 1471–1481.
- Yano TE, Takeda T, Matsubara M and Shiomi K (2017) Japan unified high-resolution relocated catalog for earthquakes (JUICE): Crustal seismicity beneath the Japanese Islands. *Tectonophysics* 702: 19–28.
- Youngs RR, Arabasz WJ, Anderson RE, Ramelli AR, Ake JP, Slemmons DB, McCalpin JP, Doser DI, Fridrich CJ, Swan FH, Rogers AM, Yount JC, Anderson LW, Smith KD, Bruhn RL, Knuepfer PLK, Smith RB, dePolo CM, O’Leary DW, Coppersmith KJ, Pezzopane SK, Schwartz DP, Whitney JW, Olig SS and Toro GR (2003) A Methodology for Probabilistic Fault Displacement Hazard Analysis (PFDHA). *Earthquake Spectra* 19(1): 191–219.
- Youngs RR, Goulet CA, Bozorgnia Y, Kuehn N, Al Atik L, Graves RW and Atkinson GM (2021) NGA-East ground-motion characterization model Part II: Implementation and hazard implications. *Earthquake Spectra* 37(1 suppl.): 1283–1330.
- Zheng Q, Liu X, Zhang H, Gu X, Fang M, Wang L and Adeeb S (2021) Reliability evaluation method for pipes buried in fault areas based on the probabilistic fault displacement hazard analysis. *Journal of Natural Gas Science and Engineering* 85: 103698.

Appendix I

We outline the numerical method used to estimate the conditional probability of surface rupture (CPSR) on principal fault. The model considers a fault plane with a down-dip width (W_F) determined by the fault’s dip angle (δ) and the seismogenic depth (Z_S), with the

top of the fault plane always at the ground surface (Figure 1a). The CPSR is computed by convolving the distributions of input parameters to generate samples of $W_{Z_{TOP}}$, each associated with a probability (e.g. Figure 1b). Surface rupture occurs when $W_{Z_{TOP}}$ reach the surface, and the corresponding probabilities are summed to derive a magnitude-dependent probability of surface rupture.

Step 1: Starting from the moment magnitude M_w , we calculate the possible values of W and its pdf (probability density function). M_w is correlated to a distribution of W values, denoted as $W(M)$, with an associated probability distribution function $p_W(W|M)$.

Thus, we can express W as a variable conditional on M :

$$W \sim p_W(W|M)$$

Where $p_W(W|M)$ is the pdf of W given the magnitude M .

Step 2: We multiply each value of W by a set of dip values.

We have a set of dip values (δ), which follows a distribution with mean μ_δ , standard deviation σ_δ , and an associated pdf $p_\delta(\delta)$.

The product of W with $\sin(\delta)$ yields a new variable W_z :

$$W_z = W \times \sin(\delta)$$

The distribution of W_z is given by the convolution of the distributions of W and δ :

$$W_z \sim p_{W_z}(W_z) = p_W(W) \cdot p_\delta(\delta)$$

Step 3: Now, we multiply each resulting value (W_z) from Step 2 by all possible HDR values r , which follow a probability distribution described by the pdf, denoted as $p_r(r)$:

$$W_{z_{TOP}} = W_z \times r$$

The resulting pdf for $W_{z_{TOP}}$ is the convolution of $p_{W_z}(W_z)$ and $p_r(r)$:

$$W_{z_{TOP}} \sim p_{W_{z_{TOP}}}(W_{z_{TOP}}) = p_{W_z}(W_z) \cdot p_r(r)$$

Step 4: We compare each value of $W_{z_{TOP}}$ with the Z_H . Z_H is the hypocentral depth and range between 0 and Z_s . Uncertainty on Z_s is explored by a distribution with mean μ_{Z_s} standard deviation σ_{Z_s} , and an associated pdf $p_{Z_s}(Z_s)$:

$$Z_s \sim p_{Z_s}(Z_s)$$

The resulting pdf for Z_H is estimated for various STZ and the assumed distribution (HDD) is described by a normal distribution with the convolution of $p_{Z_s}(Z_s)$:

$$Z_H \sim p_{Z_H}(Z_H)$$

Step 5: To verify the boundary conditions:

if $W_z \geq Z_s \mid W_{z_{TOP}} \geq Z_H \rightarrow g = 1$ (*surface rupture condition*)

$$P(\text{surface rupture} | m) = \sum_{g=1}^G (p(W|M) \cdot p(\delta) \cdot p(Z_s) \cdot p(r) \cdot p(Z_H))$$

G represents all possible cases that meet the *surface rupture condition*.