Conditional probability of distributed surface rupturing during normal faulting earthquakes

Maria Francesca Ferrario[1], Franz Livio[1]

[1] Università dell'Insubria, Dipartimento di Scienza e Alta Tecnologia, Como, 22100, Italy

Correspondence to: M. Francesca Ferrario (francesca.ferrario@uninsubria.it)

Abstract. Coseismic surface faulting is a significant source of hazard for critical plants and distributive infrastructures; it may occur either on the principal fault, or as distributed rupture on nearby faults. Hazard assessment for distributed faulting is based on empirical relations which, in the case of normal faults, were derived almost 15 years ago on a dataset of US earthquakes. We collect additional case histories worldwide, for a total of 21 earthquakes, and we calculate the conditional probability of distributed faulting as a function of distance from the principal fault. We found no clear dependency on the magnitude nor the time of occurrence of the earthquakes, but our data consistently show a higher probability of rupture when compared to the scaling relations currently adopted in engineering practice. We derive updated empirical regressions and show that results are strongly conditioned by the averaging of earthquakes effectively generating distributed faulting at a given distance and those which did not generate faulting; thus, we introduce a more conservative scenario, which can be included in a logic tree approach to consider the full spectrum of potential ruptures. Our results can be applied in the framework of probabilistic assessment of fault displacement hazard.

1 Introduction

Surface faulting is a significant source of hazard following moderate to strong earthquakes (i.e., M > ca. 6). The quantification of fault displacement hazard is critical for the engineering design of infrastructures and land use planning close to active faults. The avoidance criterion is usually applied for mitigating fault displacement hazard when fault strands location is certain; however, there are situations where crossing an active fault cannot be avoided (e.g., distributive infrastructures, pipelines). Moreover, faulting can occur either on the principal fault or on secondary fault strands. Currently, a probabilistic approach (i.e., PFDHA – Probabilistic Fault Displacement Hazard Analysis) is the suggested method to calculate the expected displacement due to surface faulting for siting nuclear power plants and critical facilities (e.g., ANSI/ANS-2.30, 2015). This approach was firstly proposed by Youngs et al. (2003) for the high-level nuclear waste repository in the Yucca Mountain, Nevada. The method is directly derived from probabilistic seismic hazard analysis (PSHA), firstly developed by Cornell (1968), which determines the annual rate of earthquakes in which a ground motion parameter exceeds a specific value, at a given location.

The earthquake approach for assessing PFDHA (Youngs et al., 2003) expresses the rate of displacement exceeding a given value as a function of i) the annual rate of occurrence of earthquakes with a given magnitude, ii) the probability of surface rupture along the principal fault and iii) the probability that off-fault rupture at a given distance from the principal fault exceeds a given displacement value.

According to Youngs et al. (2003), the attenuation function for fault displacement, i.e., the term iii) of the general PFDHA function, can be split into two terms (Eq. 1):

$$P_{kn}(D > d|m,r) = P_{kn}(Slip|m,r) \times P_{kn}(D > d|m,r,Slip) \tag{1}$$

Where k is the position of the site of interest, n is the seismogenic source, D is displacement at the site, d is a given displacement threshold, m is magnitude, r is distance from the principal fault to the site.

The first term is the conditional probability that some amount of displacement occurs at site k, i.e., it represents the actual occurrence of distributed faulting (D > 0). The second term is the conditional probability of exceeding a given level of displacement (d). In this paper, we focus exclusively on the first term of Eq. 1; this choice was driven by the fact that surface faulting can be an exclusion criterion for some plants (e.g., nuclear power plants).

While several efforts were usually devoted in describing and measuring surface faulting along the main rupture after strong earthquakes (e.g., Wells and Coppersmith, 1993, Pezzopane and Dawson, 1996; Field et al., 2015), less information is available on distributed faulting. Following the pivotal work of Youngs et al. (2003) on normal faults, the PFDHA approach was applied to strike-slip (Petersen et al., 2011) and reverse faults (Moss and Ross, 2011; Boncio et al., 2018; Nurminen et al., 2020). Regional datasets were built as well (e.g., Takao et al., 2013; Inoue et al., 2019 for Japan strike-slip and reverse faults). Slightly different methods and procedures were adopted for developing empirical relations of the probability of faulting. The decrease of the probability of faulting with distance from the principal fault was unequivocally identified by all the studies; for other driving parameters, a general consensus has still not been achieved: for instance, a magnitude dependency was included by Youngs et al. (2003), but not by Petersen et al. (2011). The hanging wall and footwall are considered separately by most of the authors dealing with dip-slip ruptures (but not by Takao et al., 2013).

The role of the local structural setting, fault architecture at depth and near-surface geology (e.g., cut lithologies, overburden load) has been highlighted after the analysis of recent earthquakes, which caused a complex pattern of rupture at surface (e.g., Bray et al., 1994; Milliner et al., 2015; Teran et al., 2015).

Concerning normal faults, the reference paper for PDFHA is still the work by Youngs et al. (2003), which was published almost 20 years ago and analyzed a dataset of US earthquakes. Since then, additional datasets were acquired (see Baize et al., 2019 for a comprehensive review on the efforts in building a unified database of fault displacement), and a general underestimation of the Youngs' relations in the far-field was recently pointed out (Ferrario and Livio, 2018). In this work, we present data and improved regression equations for the conditional probability of off-fault rupture. We supplemented the earthquakes already analyzed by Youngs et al. (2003) with additional case histories, for a total of 21 earthquakes. This figure more than doubled the dataset considered in the original work.

We provide an updated version of the conditional probability of distributed faulting as a function of distance from the principal fault. Our results broadly agree with those of Youngs et al. (2003), but systematically show a probability of occurrence higher than expected. We thus introduced a conservative scenario, to fully account for the range of observed

65 probabilities, that can be handled in a logic tree approach, as commonly done for PSHA studies.

2 Materials and methods

2.1 Dataset

We analyzed 21 events occurred between 1887 and 2016, ranging in magnitude (Mw) from 6.0 to 7.5. Nine of them are from the Western US and were already analyzed in terms of PFDHA by Youngs et al. (2003), which in turn used data compiled by

70 Pezzopane and Dawson (1996). Data for the remaining 12 earthquakes were collected from scientific literature or available databases (i.e., SURE Project, Baize et al., 2019). Table 1 lists the main details and data source for each earthquake.

The considered events occurred in different countries and all the 5 continents are represented with at least one case history: 9 earthquakes are from the US, 6 from Italy, 2 from Greece and 1 from Kenya, Turkey, New Zealand and China (Fig. 1). The geographical distribution of the events reflects either the regional seismotectonic setting, or the availability and accessibility

75 of data: the US, Italy and Greece are frequently hit by normal-faulting events; other regions, like China, are less represented in our database despite a strong earthquake activity. For the first time, an event from the African continent is included. We stress that the addition of events from different tectonic provinces and climatic conditions will enable a better characterization of distributed faulting, resulting in an overall more reliable scaling relation. In this sense, we follow the recommendation by Baize et al. (2019) and uploaded a shapefile with the rupture sections for the 13 events not already

80 available in the SURE database, in an online repository (see Data Availability Sect.).

The earthquakes' magnitude and year are plotted in the inset of Fig. 1: the mean magnitude of the events analyzed by Youngs et al. (2003) was 6.86, whereas the mean magnitude of the additional events is 6.58. The additional case histories are generally more recent in time with respect to the ones considered by Youngs, with the exception of the 1915 Fucino (Italy) and 1928 Laikipia-Subukia (Kenya) events. A particular effort was devoted to the collection of data on M < 6.5 earthquakes;

85 they were not adequately represented in the Youngs' database, but they pose a significant threat to society, being much more frequent than stronger earthquakes (Baize et al., 2019).

2.2 Methods

Our methodological workflow is illustrated in Fig. 2a. The input data are shapefiles mapping the traces of surface ruptures; these can be already available in a ready-to-use form (8 events included in Baize et al., 2019), or were created from

90 georeferencing and digitization of maps published in the scientific literature (13 events). We handled the maps with the best possible accuracy; as a general rule, the digitization of surface ruptures was performed at a scale ranging from 1:50.000 to 1:10.000, depending on the accuracy of the original maps.

We then divided the ruptures in principal and distributed faults; "principal" faulting shows longer continuity and higher displacement with respect to distributed faults and corresponds to the surface expression of the rupture along the seismogenic source at depth. Consequently, we classified as "distributed faults" (DF) all the ground breaks along structures either connected or not to the main fault, which occurred in response to principal faulting (e.g., Youngs et al., 2003; ANSI/ANS-2.30, 2015). We discarded only ruptures explicitly referred to gravitational phenomena, according to the reports. Different approaches have been adopted in the literature for defining the ruptures to be processed in further analyses: for instance, Petersen et al. (2011) discarded the "triggered" ruptures (i.e., not structurally connected to the principal fault), while Nurminen et al. (2020) ranked ruptures due to reverse earthquakes into different categories (ranked from 1 to 3). We did not attempt to further categorize the distributed ruptures in our database due to the lack of reliable information for some of the case histories; moreover, from an engineering perspective the occurrence of DF itself is much more relevant than the nature of the triggering process (e.g., Youngs et al., 2003).

The conditional probability of distributed faulting was computed following the earthquake approach, as defined by Youngs et al. (2003). This was the first work to introduce PFDHA and can be categorized as a gridding approach (Nurminen et al., 2020). We derived a raster of the Euclidean distance from the principal fault with a 500-m grid resolution and counted the number of pixels for each distance class, up to 20 km in the hanging wall and 15 km in the footwall; these limits allow to encompass all the observed DF, which reached the bins 15-15,5 km in the hanging wall and 12-12,5 km in the footwall, respectively. The conditional probability of faulting (P) is defined as the number of pixels containing distributed faulting divided by the total number of pixels in each distance class (Fig. 2a) and was calculated separately for the hanging wall and footwall blocks.

Firstly, for each event, we computed the probability of DF as a function of distance from the principal rupture - P(x). We also investigated the possible influence of magnitude range and year of occurrence (i.e., historical or modern events), assessing whether to include or not these parameters in the analyses (see Sect. 2.3). Then, we calculated the mean value for each distance class and fitted empirical regression to the dataset. We tested different functional forms, following those ones proposed in the literature (i.e., power form – Petersen et al. (2011); exponential form – Youngs et al. (2003)); in the following we provide the fitting coefficients for Eq. 2, a functional form equivalent to the one used by Youngs et al. (2003), which was the one better performing in terms of fitting.

$$P(x) = \frac{e^{(a+b*(\ln(x+c))}}{1+e^{(a+b*(\ln(x+c))}} \qquad (2)$$

Where x is distance from the principal fault in km, *a*, *b* and *c* are fitting coefficients.

2.3 Assumptions and limitations

Overall, PFDHA analysis requires a significant amount of subjective choices; a common methodology, or a thorough comparison of different methods, is still lacking. The gridding method adopted in the current research is not devoid of

limitations, which we address in the following. Nevertheless, our primary goal was to assess the performance of the only scaling relation available for normal faults, and to update such regression with new case histories. For this reason, we basically replicated the work by Youngs et al. (2003).

A first issue regards the scale, resolution and completeness of the original maps and data. For all the selected events, detailed rupture maps are available; nevertheless, a large variability in the quality of data is present, since we investigated events occurred between 1887 and 2016. An increasing quality of the reported data for more recent events could be expected: modern technologies and integrated approaches encompassing extensive fieldwork and remote sensing (e.g., InSAR, optical correlation techniques etc.) can capture ground deformation of few centimeters, which in the past could have easily gone undetected (see Livio et al., 2017 for a more detailed discussion). A working hypothesis is that older events are characterized by a higher epistemic uncertainty due to less reliable technologies or natural censoring of smaller displacements if measurements are taken long after the earthquake occurrence (e.g., Stirling et al., 2002). The variability of modern events should indeed be aleatory. We explored this issue by dividing the dataset in two subsets, namely events occurred in the XXI century (4 events) and those occurred before that date (17 events).

A second issue is related to the definition of the principal fault, which in turn affects the computation of distances. The delineation of the principal fault can be straightforward in some rupture sections but can be more complex in other sections (Petersen et al., 2011), where the surface rupture is discontinuous or structural complexities (multiple parallel strands, gaps between ruptures) are present. Figure 2 shows examples from the Gulf of Corinth (Greece) and Edgecumbe (New Zealand) events: in the first case, the principal fault shows a relatively simple trace at the two ends, while in the central part a complex pattern of ruptures is present. The Edgecumbe event ruptured the Edgecumbe fault and 10 secondary segments. In the assessment of the principal fault, we referred to the fault plane modeled by Beanland et al. (1990), which defines the Edgecumbe fault as the main fault rupture; the Onepu fault (see Fig. 2c) lies along-strike of the Edgecumbe fault, but shows a much smaller displacement (maximum values of 2.5 m and 26 cm, respectively) and thus is considered as a secondary rupture.

The issues defined above are not trivial, because the gridding method implicitly assumes that the mapping data for the ruptures are complete, and results depend on the grid size. The grid size of 500 m is quite coarse, to compensate for the possible underestimation due to incomplete mapping (Youngs et al. 2003). Other approaches have been explored in the literature, such as to consider different grid sizes (e.g., Petersen et al., 2011). More recently, a "slicing" approach has been introduced in the analysis of reverse earthquakes (Boncio et al., 2018; Nurminen et al., 2020); this method makes no assumptions on the completeness of the database and does not depend on the grid size.

Finally, in case of multiple earthquakes in few days/months (e.g., 1981 Gulf of Corinth: Jackson et al. 1982; 2016 Central Italy: Brozzetti et al., 2019), it may not be possible to attribute each surface rupture to its causative event. Repeated rupture of the same fault strand is certainly a theme to be investigated, but we believe that this kind of uncertainty does not heavily affect our results.

5

3 Results

160 Firstly, we computed the conditional probability of distributed faulting for each single event; then, we explored the role of magnitude and year of occurrence as factors affecting this value.

3.1 The role of magnitude

Firstly, we grouped the case histories into magnitude classes (i.e., M < 6.5; 6.5 ≤ M < 7.0; M ≥ 7.0). Figure 3 shows the conditional probability of faulting as a function of the distance from the principal fault. Positive values correspond to the
165 hanging wall, whereas negative values refer to the footwall. Each symbol represents P(x) at a specific distance for a single earthquake. Points on the x-axis indicate that no distributed faulting occurred in the given distance class; black asterisks show the mean values. The probability of rupture drops off quickly while moving farther from the main fault, with a steeper decrease in the footwall than in the hanging wall. The ratio between the probability of faulting in the 0-500 m and 1000-1500 m classes is 7:1 for the hanging wall and 14:1 for the footwall, pointing to a fundamental difference between the near and far
170 fields. Some peaks in the far-field are visible as well (e.g., in Fig. 3a at distance of 7-8 km in the hanging wall).

Figure 3b shows the stacked values of conditional probability, with earthquakes ordered according to magnitude; again, it is clear the decrease of P(x) with distance, but no clear trend is apparent for increasing magnitude values. Thus, in the following analyses we do not explicitly include a magnitude term in the scaling relations; moreover, it is important to highlight the relatively small sample set and that the magnitude determination for the older events (which go back to the end
175 of the XIX century) may bear a significant degree of uncertainty.

The mean values shown in Fig. 3 correspond quite well with the scaling relations by Youngs et al. (2003), but we underline that mean values derive from balancing a dichotomous variable, because distributed rupture can either occur or not occur. On one hand, earthquakes actually producing distributed faulting show probabilities much higher than the mean value; on the other hand, several earthquakes do not produce faulting at all at a certain distance from the principal fault. Figure 4 better
180 clarifies this point: we calculated the percentage of earthquakes generating distributed faulting for each distance class ("events with no DF" corresponds to the points on the x-axis in Fig. 3). Most earthquakes produce distributed faulting in the first kilometers from the principal fault, while only 30% generate surface faulting at 7 km in the hanging wall, and this value decreases to 3 km in the footwall. Another aspect worth mentioning is that real data are constrained down to probabilities as low as ca. 10^{-2} (see "No data" field in Fig. 3a); this lower threshold is constrained by the number of pixels for each distance
185 bin, which in turn depends on the grid size of the analysis and the length of the principal fault. Probabilities lower than this threshold derive from the averaging of the earthquakes producing DF with those not producing DF.

3.2 The role of the dataset age

We then grouped the case histories according to the year of occurrence, i.e., 17 earthquakes occurred before year 2000 (mean magnitude: 6.75) and 4 more recent earthquakes (mean magnitude: 6.47). We tested the hypothesis that older earthquakes

190   may show lower probability of distributed faulting due to the incompleteness of dataset or less reliable measures. However, we found no systematic bias between the two subsets and any clear pattern arises when stacking the events according to the year of occurrence (Fig. 5).

3.3 Deriving scaling relations: regular and conservative scenarios

In the following we derive the equations of the conditional probability of distributed faulting as a function of distance from
195   the principal fault. We calculate the mean probability values for each class distance and fit the data with Eq. 2. When trying to fit all the data, we found a tradeoff between the near and far-field: including in the regression the points in the vicinity of the principal fault gives birth to higher residuals in the far-field. This drawback was already pointed out by Petersen et al. (2011), who excluded the first two off-fault rupture probability measurements. We tested several functional forms and we selected Eq. 2 because it was the best performing in the different scenarios. The obtained regression coefficients and good of
200   fit parameters are provided in Table 2, while the curves are plotted in Fig. 6. The overall pattern matches the results of Youngs et al. (2003), but we note that observed data are consistently higher than predicted by their equations, in particular in the hanging wall.

As pointed out earlier, computing the mean value of all the earthquakes for each distance bin results in the average of two end members: events without distributed faulting at a given distance; and events that generated distributed faulting at that
205   distance. The conditional probability of faulting is zero for the former category, while it is variable, within a certain range, for the latter and it comes out that, when occurring, the probability is much higher than the mean value. For this reason, we introduced a conservative scenario, computing the conditional probability only for those data that recorded distributed faulting at a given distance. The points are shown in Fig. 6, together with the curves obtained in the same way as for the regular scenario. The relevance of this conservative scenario lies in the fact that for some projects it may be necessary to
210   consider the worst-case scenario or return period longer than the norm (Wells and Kulkarni, 2014; Cline et al., 2018).

Between 1 and 5 km from the principal fault, the conservative scenario predicts a probability of faulting that is ca. 3 times the regular scenario, while at about 10 km from the principal fault the conservative scenario is one order of magnitude higher than the regular one. The residual plots (observed minus predicted values, Fig. 6c) show that for the regular scenario observed values better match with the predicted values, with a higher discrepancy at 1-3 km in the hanging wall; beyond ca.
215   10 km in the hanging wall, residuals are always negative, i.e., the equation overpredicts the values. The conservative scenario shows a higher dispersion of the values, resulting in overall higher residuals and lower goodness of fit parameters (Table 2). In the database used to derive the equations, distributed faulting was observed up to 15 and 12 km from the principal fault in the hanging wall and footwall, respectively. Due to the limited numbers of observation points beyond ca. 10 km, extrapolation beyond such limits should be gingerly considered.

4 Discussion

We compute the conditional probability of distributed faulting for dataset of 21 normal faulting earthquakes and provide updated empirical relations assessing the decrease of faulting with distance. Our data show a very steep decrease in the probability of faulting at distances higher than 1 km from the principal fault, pointing to a fundamental difference between near and far-fields. The pattern of fault rupture in the near field has been recently explored by several authors, by means of observation of actual faulting, numerical or analytical models (e.g., Fletcher and Spelz, 2009; Teran et al., 2015; Gold et al., 2015; Loukidis et al., 2009; Treiman, 2010). Several factors have been pointed out as conditioning the surface expression of faulting, including type of fault movement, fault dip, amount of displacement, geometrical complexity of fault traces, rock type, thickness and nature of the materials above bedrock (Bray et al., 1994; Avar and Hudyma 2019). The integration of modern field and remote technologies allows to capture the finer details of surface rupture; a geologically sound interpretation of such data is pivotal to the understanding of surface rupture, which in turn is essential for engineering design and eventually mitigation measures. The regressions presented in the current work derive from a global dataset; we stress that, if site-specific data are available, they should be properly considered. For instance, a different behavior between Japan and US ruptures has been identified (Inoue et al., 2018; Petersen and Chen, 2018; Suzuki and Annaka, 2018); if such variations are systematic, they can be to some extent predictable and thus help in assessing the hazard.

From the methodological point of view, the biggest caveat of our analysis is arguably the 500-m grid size, which is quite coarse and possibly smooths differences between events. We adopted this size because we wanted to verify the performance of the regressions by Youngs et al. (2003). The assumption that any distributed faulting in the 500 x 500 m cell is the same as the occurrence of rupture at the point of interest leads to an overestimation of the rate of occurrence; the large pixel size was chosen by Youngs et al. (2003) as a compromise between this effect and the possible incomplete mapping.

Beside the grid size, the definition of principal faulting itself can be tricky. Several earthquakes may occur in a short time interval, making it difficult to identify the causative event for each ground rupture. Discontinuous, sub-parallel strands with similar displacement, or a complex pattern of faulting are a challenge for defining principal vs distributed faulting. The updated empirical regressions broadly confirm the results obtained by Youngs et al. (2003), highlighting the soundness of their approach and its wide applicability in different regions worldwide. The addition of new case histories allows just an incremental improvement in the equations. The conditional probability of faulting does not show dependency on magnitude or time of occurrence, thus allowing to combine all the events in a single dataset for obtaining empirical regressions. The lack of dependency on magnitude can be explained by the fact that we analyze a global dataset, whereas Youngs et al. (2003) focused their work on earthquakes occurred in the Western US. Moreover, it has been shown that the choice of the grid size has a much larger effect than the moment magnitude of the events (Suzuki and Annaka, 2018; Takao et al., 2018).

Concerning the year of occurrence, in a previous analysis of Italian Apennines events we claim that older events and maps may be affected by higher epistemic uncertainties (Ferrario and Livio, 2018). In the current study, we tested again this hypothesis and found no systematic difference between earthquakes occurred several decades apart; thus, we argue that the

issue of data completeness is not simply a matter of the time of occurrence, but is influenced by other factors like the territorial and climatic setting (e.g., arid vs humid climate) and the potential of preservation of surface ruptures. Another possible explanation for the lack of dependency on the year of occurrence is the limited number of modern case histories (only 4 events post-2000). The acquisition of future datasets is critical for assessing eventual systematic biases and will allow to test different grid sizes.

The most striking observation when analyzing the updated regressions is the systematic higher probabilities of faulting, in particular in the hanging wall, with respect to the original formulation by Youngs et al., (2003). The output of the analysis is strongly influenced by the bins where no distributed faulting was observed (i.e., points on the x-axis in Fig. 3 and 5). In order to account for the full spectrum of potential occurrences, we introduce a conservative scenario, where conditional probabilities of faulting are computed without the bins with no distributed faulting. This scenario results in higher expected probabilities, and the residual plot (Fig. 6c) shows that observed values are lower than expected between 3 and 7 km in the hanging wall, whereas are higher than expected at distances of 7-12 km. A possible explanation to this behavior may be the structural architecture of normal faults, which commonly show a horst and graben setting: the sector 3-7 km may correspond to the block which undergo lowering and/or tilting, and the sector 7-12 km should correspond to the location of the antithetic structure. This hypothesis can be tested on well-documented case histories, where detailed information is available on the structural architecture of the shallow subsoil. As a first approximation, this pattern is quite evident for the 1980 Irpinia and 2016 Norcia earthquakes, and more subdued for the 1954 Fairview Peak, 1959 Hebgen Lake, 1970 Geidz and 1987 Edgecumbe events. The position at surface of the antithetic structure is driven e.g., by the change in dip of the principal fault at depth (e.g., Caskey et al., 1996); this points to the possibility of introducing deterministic constraints in the estimation of the expected distributed faulting. Consistently, we here point out that the use of elastic dislocation models of deformation (e.g., Okada, 1985) and, in turn, of induced Coulomb stress transfer on receiving pre-existing faults (e.g., Toda et al., 2011) may be useful for better predicting the probability of DF (e.g., Gurpinar et al., 2017, Livio et al., 2017), especially where the current models show higher residuals, e.g., at ca. 7-12 km in the hanging wall.

We claim that the regular and conservative scenarios can be seen as alternative branches in a logic tree; the relative weights of the two branches can be tuned according to the degree of conservativity of the project and to professional judgement. We argue that the conservative scenario can be particularly important in the analysis of normal faults, for two main reasons: i) normal faults have a higher probability of generating surface faulting along the principal fault when compared to strike-slip and reverse faults (Moss and Ross, 2011; Suzuki and Annaka, 2018; Takao et al., 2018); ii) the conditional probability of distributed faulting for normal events is much higher than for strike-slip and reverse earthquakes. The latter point is better illustrated in Fig. 7, where we compare our regressions for normal faulting with the original formulation of Youngs et al. (2003) and the one proposed by Takao et al. (2018) for strike-slip and reverse Japanese events, developed using the same grid size of 500 m.

5 Conclusions

We develop an updated regression for the conditional probability of distributed faulting as a function of distance. We propose two alternative scenarios, to take into account the wide range of ruptures: the regular scenario computes the average value of all the analyzed events, while the second scenario provides a more conservative estimate. We believe that a periodic update of the database allows to derive more robust relations: for this reason, measures of distributed faulting and data implementation in a common framework (e.g., Baize et al., 2019) should be a standard practice following every ground-rupturing event. In the effort of common data mining, the shapefiles of ruptures for the new case histories and the table with conditional probability of faulting derived in the present research are made available in an open repository (see Data Availability Sect.).

Here we focus on the conditional probability of faulting and do not consider the amount of faulting (i.e., we treat faulting as a binomial variable YES/NO). Future research should consider also the amount of displacement, to fully implement our result in a PFHDA perspective. PFDHA is a young science relying on empirical data, thus still holds a significant degree of epistemic uncertainty, which can be reduced by enlarging the dataset of case histories.

The assessment of uncertainties in PFDHA analyses is not yet fully developed and methodological choices have a substantial effect on the total hazard (Moss and Ross, 2010; Treiman, 2010; Wells and Kulkarni, 2014; Cline et al., 2018); logic trees can be used to consider the full distribution of rupture characteristics. Finally, a critical comparison of different methods and procedures should be pursued and the comparability with the original works (e.g., Youngs et al., 2003) should be guaranteed.

References

Ambraseys, N. N.: Earthquake hazard in the Kenya Rift: the Subukia earthquake 1928. Geophys. J. Int., 105, 253-269, (1991)

10

Ambraseys, N. N., and Tchalenko, J. S.: Seismotectonic aspects of the Gediz, Turkey, Earthquake of March 1970. Geophys. J. R.astr. SOC., 30, 229-252, 1972.

Amoroso, S., Bernardini, F., Blumetti, A. M., Civico, R., Doglioni, C., Galadini, F., Galli, P., Graziani, L., Guerrieri, L., Messina, P., Michetti, A. M., Potenza, F., Pucci, S., Roberts, G., Serva, L., Smedile, A., Smeraglia, L., Tertulliani, A., Tironi, G., Villani, F., and Vittori, E.: Quaternary geology and Paleoseismology in the Fucino and L'Aquila basins. 6th INQUA, 19-24 April 2015, Pescina (AQ) - Fucino basin. Periodico semestrale del Servizio Geologico d'Italia - ISPRA e della Società Geologica Italiana Geol.F.Trips, 8, 88 pp., DOI 10.3301/GFT.2016.02, 2016.

ANSI/ANS-2.30: Criteria for assessing tectonic surface fault rupture and deformation at nuclear facilities, Published by the American Nuclear Society, 2015.

Avar, B. B., and Hudyma, N. W.: Earthquake Surface Rupture: A Brief Survey on Interdisciplinary Research and Practice from Geology to Geotechnical Engineering. Rock Mechanics and Rock Engineering https://doi.org/10.1007/s00603-019-02006-0, 2019

Baize, S., Nurminen, F., Sarmiento, A., Dawson, T., Takao, M., Scotti, O., et al.: A worldwide and unified database of surface ruptures (SURE) for fault displacement hazard analyses. Seismol Res. Lett. 91, 499–520. doi:10.1785/0220190144 2019, 2019.

Beanland, S., Berryman, K.R., and Blick G. H.: Geological investigations of the 1987 Edgecumbe earthquake, New Zealand. N.Z.J. Geol. Geophys. 32(1), 73-91, 1989.

Beanland, S., Blick, G. H., and Marby, D. J.: Normal Faulting in a Back Arc Basin: geological and geodetic characteristics of the 1987 Edgecumbe Earthquake, New Zealand. Journ of Geoph. Res. (95), B4, 4693-4707, 1990.

Blumetti, A. M., Esposito, E., Ferreli, L., Michetti, A. M., Porfido, S., Serva, L., and Vittori, E.: New data and reinterpretation of the November 23, 1980, M 6.9 Irpinia-Lucania earthquake (Southern Apennines) coseismic surface effects, Large scale vertical movements and related gravitational processes. Studi Geologici Camerti Special Issue, 2002, 19–27, 2002.

Boncio, P., Liberi, F., Caldarella, M., and Nurminen, F.C.: Width of surface rupture zone for thrust earthquakes: implications for earthquake fault zoning. Nat. Hazards Earth Syst. Sci. 18, 241–256. doi:10.5194/nhess-18-241-2018, 2018.

Bray, J. D., Seed, R. B., and Cluff, L. S.: Earthquake fault rupture propagation through soil. Journal of Geotechnical Engineering, 120, 3, 1994.

Brozzetti, F., Boncio, P., Cirillo, D., Ferrarini, F., de Nardis, R., Testa, A., et al.: High-resolution field mapping and analysis of the August–October 2016 coseismic surface faulting (central Italy earthquakes): Slip distribution, parameterization, and comparison with global earthquakes. Tectonics, 38. https://doi.org/10.1029/2018TC005305, 2019.

Caskey, S. J., Wesnousky, S. G., Zhang, P., and Slemmons, D. B.: Surface faulting of the 1954 Fairview Peak (MS 7.2) and Dixie Valley (MS 6.8) earthquakes, central Nevada. Bulletin of the Seismological Society of America, 86(3), 761-787, 1996.

Cello, G., Deiana, G., Mangano, P., Mazzoli, S., Tondi, E., Ferreli, L., et al.: Evidence for surface faulting during the September 26, 1997, Colfiorito (Central Italy) earthquakes. Journal of Earthquake Engineering, 2(2), 303–324. https://doi.org/10.1080/13632469809350324, 1998.

Civico R, Pucci S, Villani F, Pizzimenti L, De Martini PM, Nappi R, the Open EMERGEO Working Group: Surface ruptures following the 30 October 2016 M w 6.5 Norcia earthquake, central Italy. J Maps 14(2):151–160. https://doi.org/10.1080/17445647.2018.1441756, 2018.

Cline, K. M., Cline, M. L., Blanco, J., Quittmeyer, R., Kimball, J., Rizzo, P. C., and Bavec, M.: Probabilistic Fault Displacement Hazard Analysis for Regulatory Decision-Making: A Case Study Using State-of-the-Practice Fault Characterization Methods. In: IAEA-TECDOC-CD-1833 ISBN 978–92–0–158917–0 ISSN 1684–2073, Vienna, 2018.

Cornell, C. A.: Engineering seismic risk analysis, Bull. Seismol. Soc. Am., 22, 2, 1583-1606, 1968.

Ferrario, M. F., and Livio, F.: Characterizing the distributed faulting during the 30 October 2016, Central Italy earthquake: A reference for fault displacement hazard assessment. Tectonics, 37. https://doi.org/10.1029/2017TC004935, 2018.

Field, E. H., Biasi, G. P., Bird, P., Dawson, T. E., Felzer, K. R., Jackson, D. D., ... and Milner, K. R.: Long-term time-dependent probabilities for the third Uniform California Earthquake Rupture Forecast (UCERF3). BSSA, 105(2A), 511-543, 2015.

Fletcher, J. M., and Spelz, R. M.: Patterns of Quaternary deformation and rupture propagation associated with an active low-angle normal fault, Laguna Salada, Mexico: evidence of a rolling hinge? Geosphere 2009;5;385-407 doi: 10.1130/GES00206.1, 2009.

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

Gürpinar, A., Serva, L., Livio, F. and Rizzo, P.C.: Earthquake-induced crustal deformation and consequences for fault displacement hazard analysis of nuclear power plants. Nuclear Engineering and Design 311, 69–85. https://doi.org/10.1016/j.nucengdes.2016.11.007, 2017.

Inoue, N., Kitada, N., Tonagi, M., and Irikura, K.: Study on the evaluation method for fault displacement: probabilistic approach based on Japanese earthquake rupture data. In: IAEA-TECDOC-CD-1833 ISBN 978–92–0–158917–0 ISSN 1684–2073, Vienna, 2018.

Inoue, N., Kitada, N., Shibuya, N., Omata, M., Takahama, T., Tonagi, M., and Irikura, K.: Probabilistic Evaluation of Off-Fault Displacements of the 2016 Kumamoto Earthquake. Pure Appl. Geophys. https://doi.org/10.1007/s00024-019-02345-7, 2019.

Jackson, J. A., Gagnepain, J., Houseman, G., King, G.C.P., Papadimitriou, P., Soufleris, C., and Virieux, J.: Seismicity, normal faulting, and the geomorphological development of the Gulf of Corinth (Greece): the Corinth earthquakes of February and March 1981. EPSL, 57, 377-397, 1982.

Livio, F., Michetti, A. M., Vittori, E., Gregory, L., Wedmore, L., Piccardi, L., et al.: Surface faulting during the August 24, 2016, central Italy earthquake (Mw 6.0): Preliminary results. Annals of Geophysics, 59, 1–8. https://doi.org/10.4401/ag-7197, 2016.

Livio, F., Serva, L., and Gürpinar, A.: Locating distributed faulting: Contributions from InSAR imaging to Probabilistic Fault Displacement Hazard Analysis (PFDHA). Quaternary International, 451, 223–233, 2017.

Loukidis, D., Bouckovalas, G. D., and Papadimitriou, A. G.: Analysis of fault rupture propagation through uniform soil cover. Soil Dynamics and Earthquake Engineering 29. 1389–1404, 2009.

Milliner, C. W. D., Dolan, J. F., Hollingsworth, J., Leprince, S., Ayoub, F., and Sammis, C. G.: Quantifying near-field and off-fault deformation patterns of the 1992 Mw 7.3 Landers earthquake. Geochemistry, Geophysics, Geosystems, 16, 1577–1598. https://doi.org/10.1002/2014GC005693, 2015.

Moss, R. E. S., and Ross, Z. E.: Probabilistic fault displacement hazard analysis for reverse faults. Bull. Seismol. Soc. Am. 101 (4), 1542–1553. doi:10.1785/0120100248, 2011.

Nurminen, F., Boncio, P., Visini, F., Pace, B., Valentini, A., Baize, S., and Scotti O.: Probability of Occurrence and Displacement Regression of Distributed Surface Rupturing for Reverse Earthquakes. Front. Earth Sci. 8:581605. doi:10.3389/feart.2020.581605, 2020.

Okada, Y.: Surface deformation due to shear and tensile faults in a half-space. Bulletin of the Seismological Society of America 75 (4), 1135-1154, 1985.

Petersen, M. D., Dawson, T. E., Chen, R., Cao, T.,Wills, C. J., Schwartz, D. P., et al.: Fault displacement hazard for strike-slip faults. Bull. Seismol. Soc. Am. 101 (2), 805–825. doi:10.1785/0120100035, 2011.

Petersen, M. D., and Chen, R: Empirical fault displacement hazard methods applied in the United States. In: IAEA-TECDOC-CD-1833 ISBN 978–92–0–158917–0 ISSN 1684–2073, Vienna, 2018.

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

Stirling, M., Rhoades, D., and Berryman, K.: Comparison of Earthquake Scaling Relations Derived from Data of the Instrumental and Preinstrumental Era. Bulletin of the Seismological Society of America, 92, 2, 812–830, 2002.

Suzuki, Y., and Annaka, T.: Probabilistic hazard analysis for secondary fault In: IAEA-TECDOC-CD-1833 ISBN 978–92–0–158917–0 ISSN 1684–2073, Vienna, 2018.

Takao, M., Annaka, T., and Kurita, T.: Application of probabilistic fault displacement hazard analysis in Japan. Journal of Japan Association for Earthquake Engineering, 13(1), 17–36, 2013.

Takao, M., Annaka, T., and Kurita, T.: Establishment of evaluation formulae for Probabilistic Fault Displacement Hazard analysis (PFDHA) in Japan. In: IAEA-TECDOC-CD-1833 ISBN 978–92–0–158917–0 ISSN 1684–2073, Vienna, 2018.

Teran, O. J., Fletcher, J. M., Oskin, M. E., Rockwell, T. K., Hudnut, K. W., Spelz, R. M., et al.: Geologic and structural controls on rupture zone fabric: A field-based study of the 2010 Mw 7.2 El Mayor–Cucapah earthquake surface rupture. Geosphere, 11(3), 899–920. https://doi.org/10.1130/GES01078.1, 2015.

Toda, S., Stein, R. S., Sevilgen, V., and Lin, J.: Coulomb 3.3 Graphic-rich deformation and stress-change software for earthquake, tectonic, and volcano research and teaching—user guide. US Geological Survey open-file report, 1060(2011), 63, 2011.

Treiman, J. A.: Fault Rupture and Surface Deformation: Defining the Hazard. Environmental & Engineering Geoscience, 16, 1, 19–30, 2010.

Wells, D.L. and Coppersmith, K.J.: Likelihood of surface rupture as a function of magnitude (abs.). Seismological Research Letters 64 (1), 54, 1993.

Wells, D. L., and Kulkarni, V. S.: Probabilistic Fault Displacement Hazard Analysis – sensitivity analyses and recommended practices for developing design fault displacements. Proceedings of the 10th National Conference in Earthquake Engineering, Earthquake Engineering Research Institute, Anchorage, AK, 2014.

Xu, X., Tan, X., Yu, G., Wu, G., Fang, W., Chen, J., Song, H. and Shen, J.: Normal- and oblique-slip of the 2008 Yutian earthquake: Evidence for eastward block motion, northern Tibetan Plateau. Tectonophysics 584, 152–165, 2013.

Youngs, R. R., Arabasz, W. J., Anderson, R. E., Ramelli, A. R., Ake, J. P., Slemmons, D. B., et al.: A methodology for probabilistic fault displacement hazard analysis (PFDHA). Earthq. Spectra. 19 (1), 191–219. doi:10.1193/1.1542891, 2003.
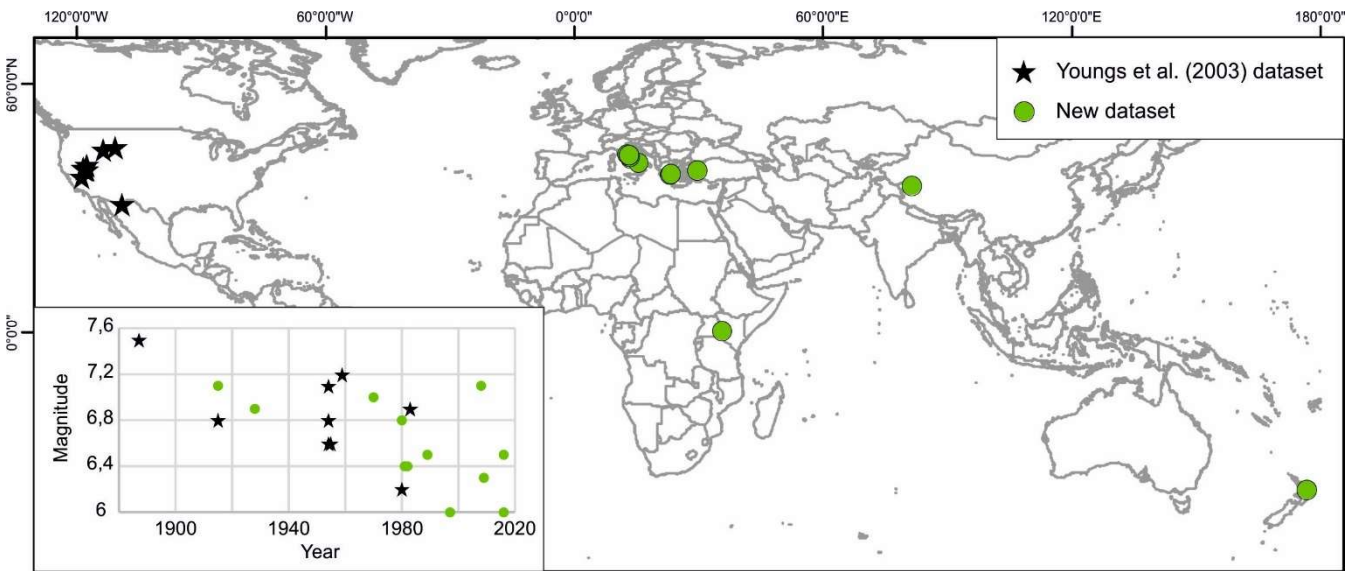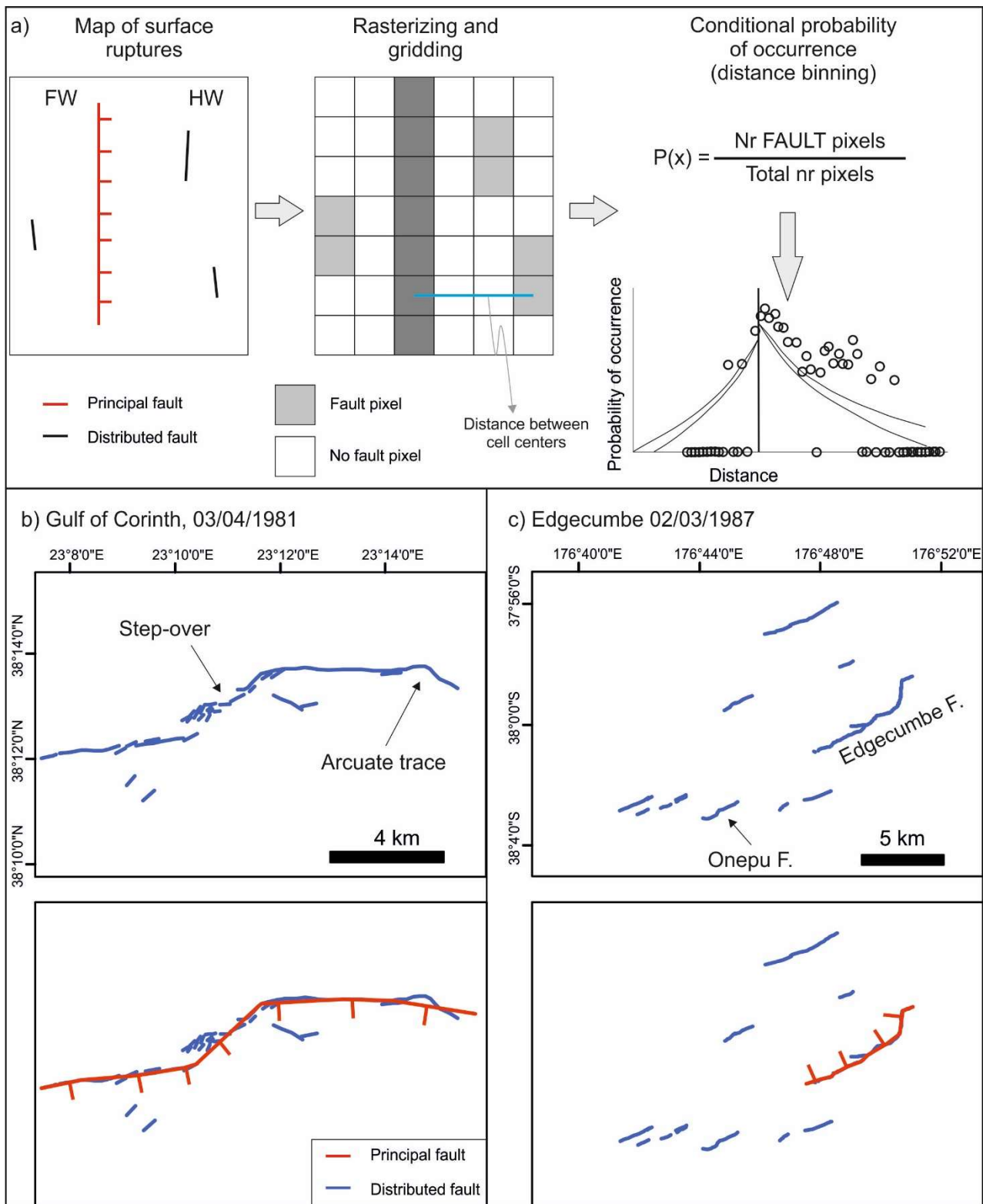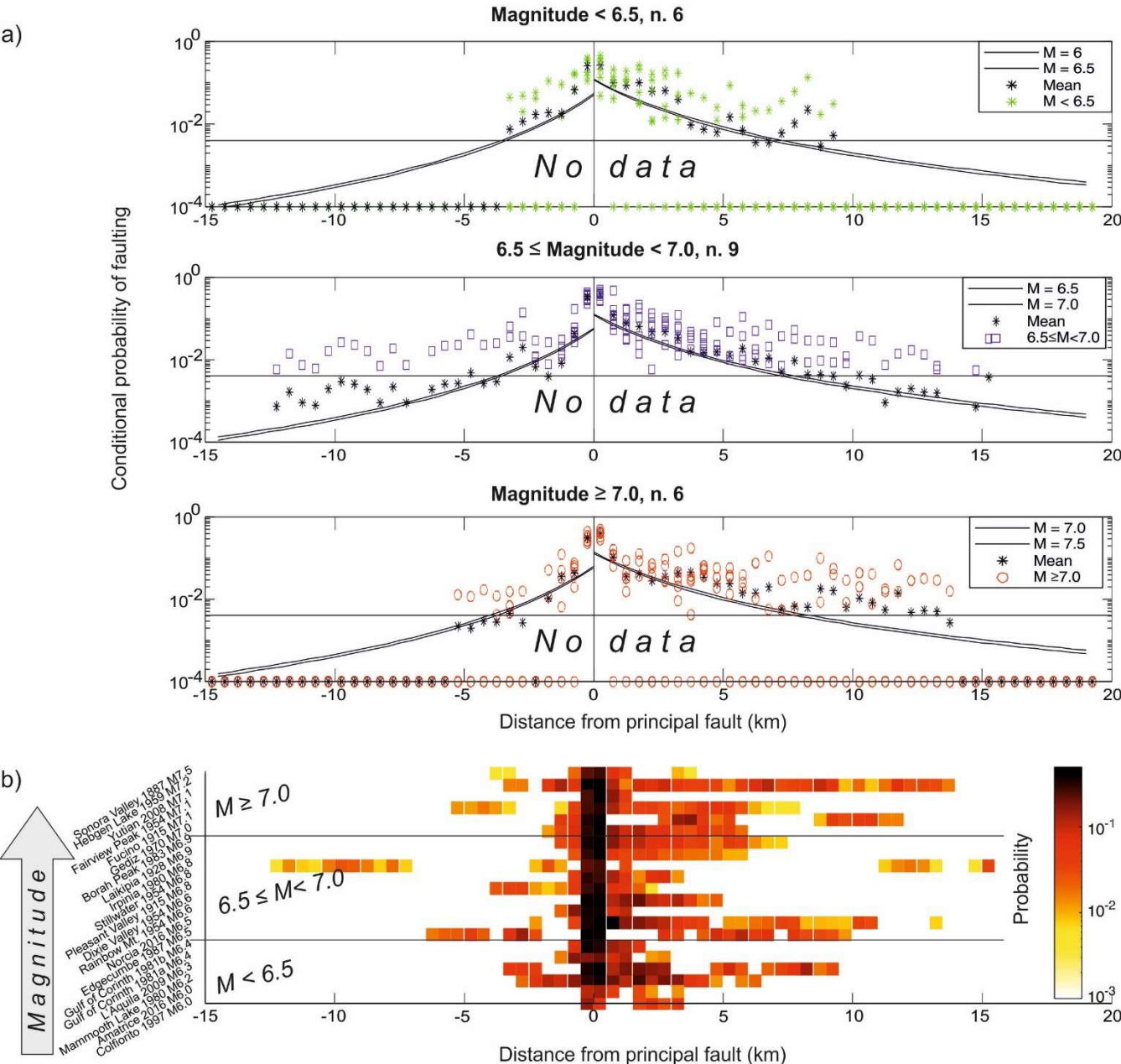
Figure 1: Location of the analyzed events. The plot in the inset shows the distribution of the events according to magnitude and time of occurrence

14

a) Map of surface ruptures | Rasterizing and gridding | Conditional probability of occurrence (distance binning)

FW   HW

$$P(x) = \frac{Nr\ FAULT\ pixels}{Total\ nr\ pixels}$$

— Principal fault
— Distributed fault

Fault pixel
No fault pixel

Distance between cell centers

Probability of occurrence
Distance

b) Gulf of Corinth, 03/04/1981

23°8'0"E   23°10'0"E   23°12'0"E   23°14'0"E

38°14'0"N
38°12'0"N
38°10'0"N

Step-over

Arcuate trace

4 km

c) Edgecumbe 02/03/1987

176°40'0"E   176°44'0"E   176°48'0"E   176°52'0"E

37°56'0"S
38°00'0"S
38°40'S

Edgecumbe F.

Onepu F.

5 km

— Principal fault
— Distributed fault

15

435

Figure 3: a) **Conditional probability of faulting as a function of distance from the principal fault. Earthquakes are grouped according to their magnitude (M < 6.5; 6.5 ≤ M < 7.0; M ≥ 7.0); colored symbols represent the values for each earthquake and distance class, black asterisks are the mean values. Black lines are the scaling relations by Youngs et al. (2003; their Eq. 7). The**
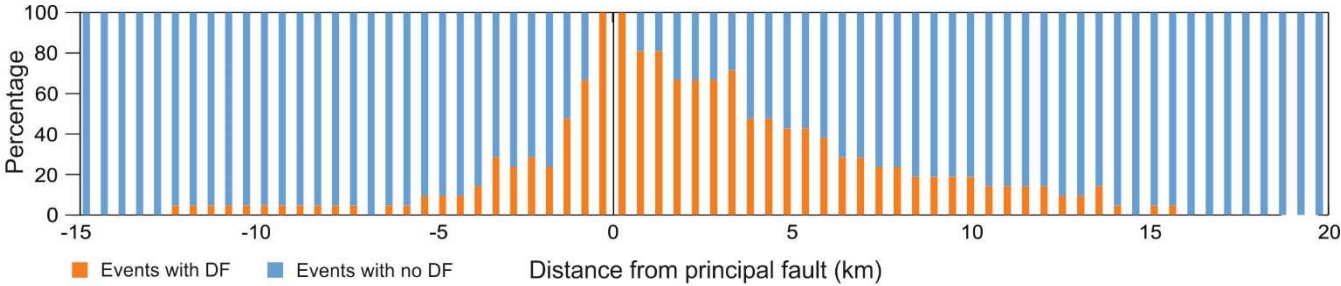
440

Figure 4: percentage of earthquakes showing distributed faulting vs no faulting for each distance class.
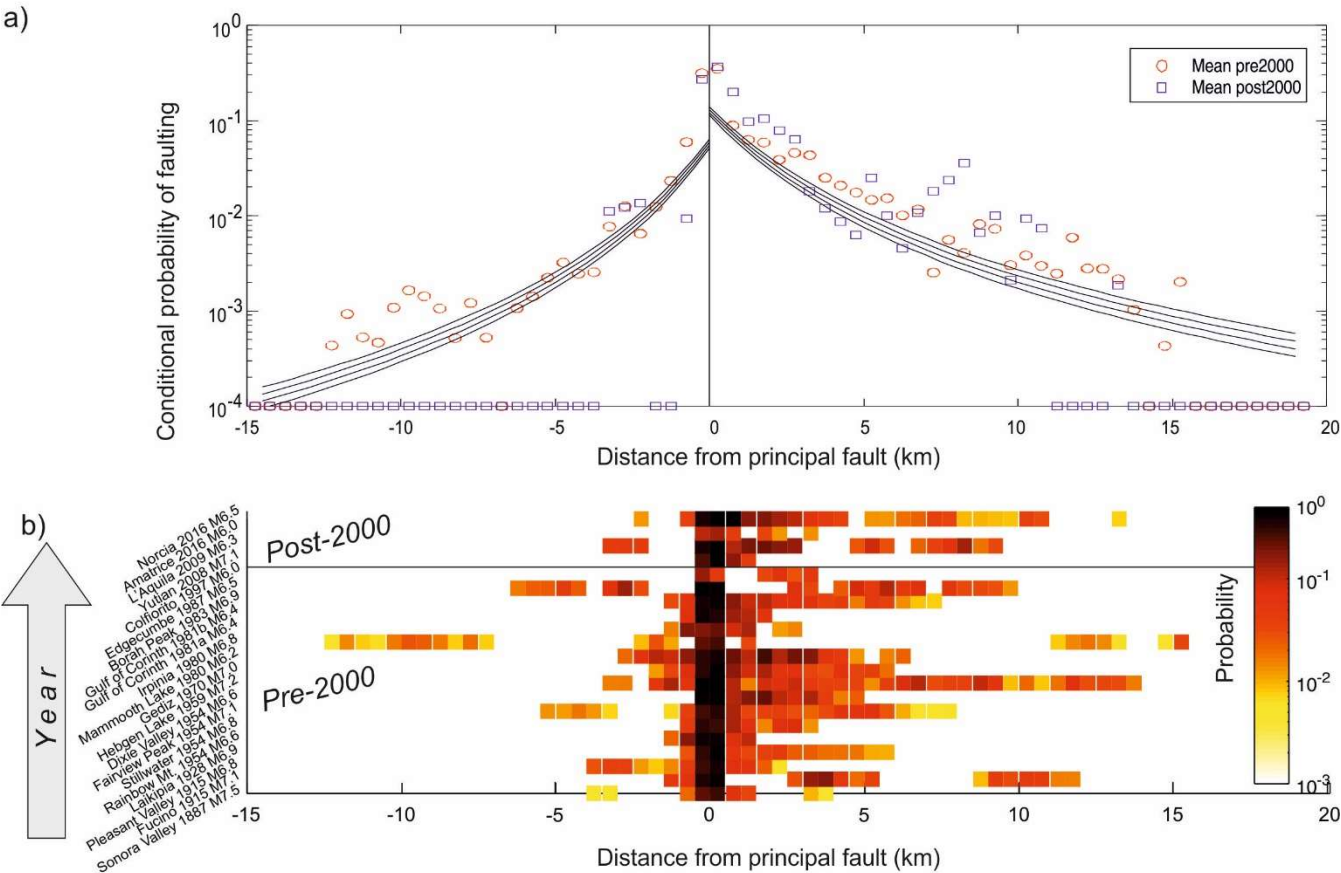

Figure 5: a) Conditional probability of faulting for the subset of pre-2000 and post-2000 events (n. 17 and 4, respectively); colored symbols represent the mean values. Black lines are the scaling relations by Youngs et al. (2003; their Eq. 7). The "No data" field illustrates the region where no empirical observations are present, and probabilities are extrapolated by fitting. b) comparison between the 21 earthquakes, ranked according to the year of occurrence; each line represents a single event.
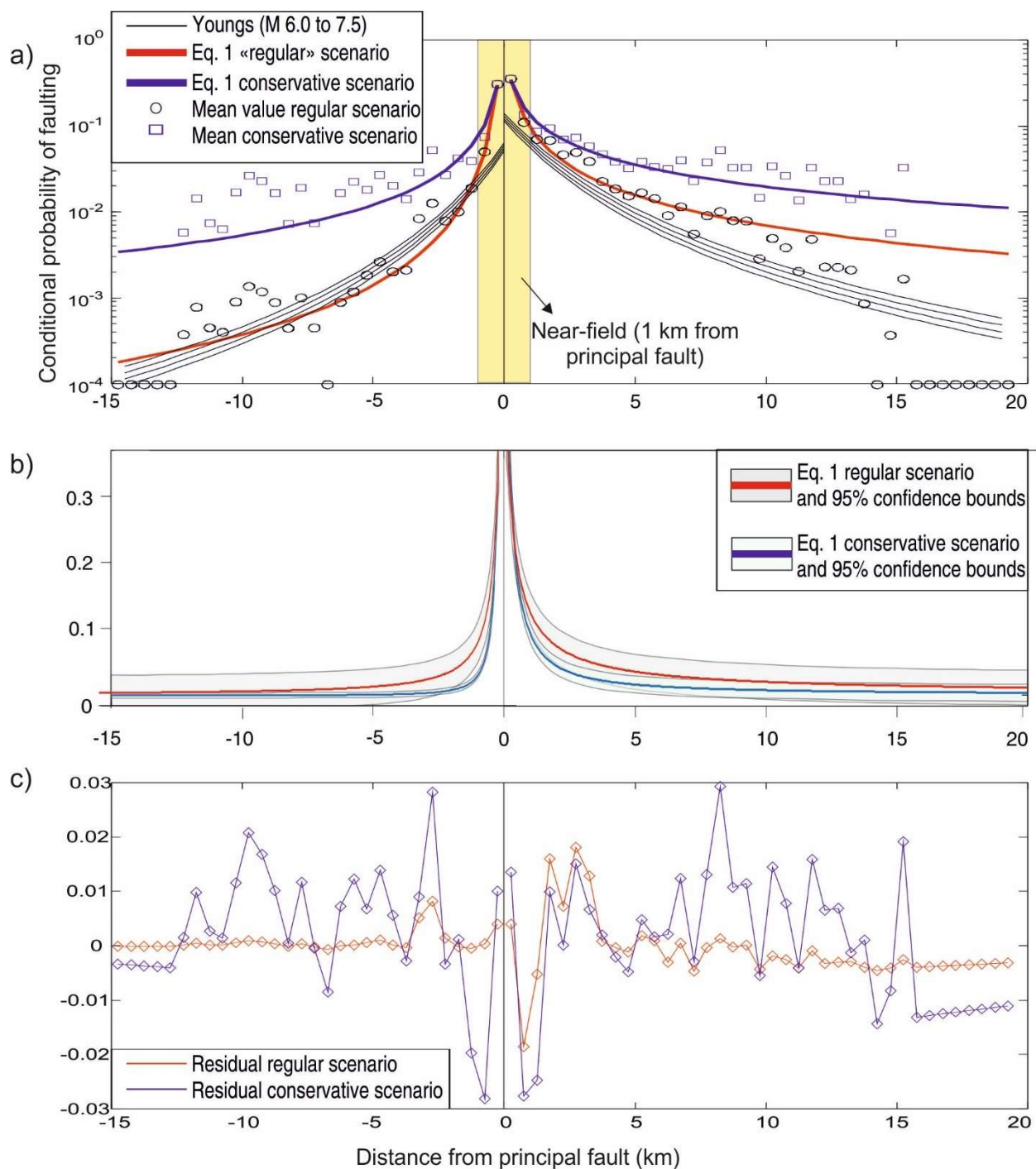
445

450

Figure 6: a) Conditional probability of faulting as a function of distance from the principal fault. Black lines are the scaling relations by Youngs et al. (2003; their Eq. 7). Colored lines are the equations proposed in the present research ("regular" and "conservative" scenarios). B) Proposed equations and 95% confidence bounds (y-axis: linear scale). C) Residual plot (observed – predicted).
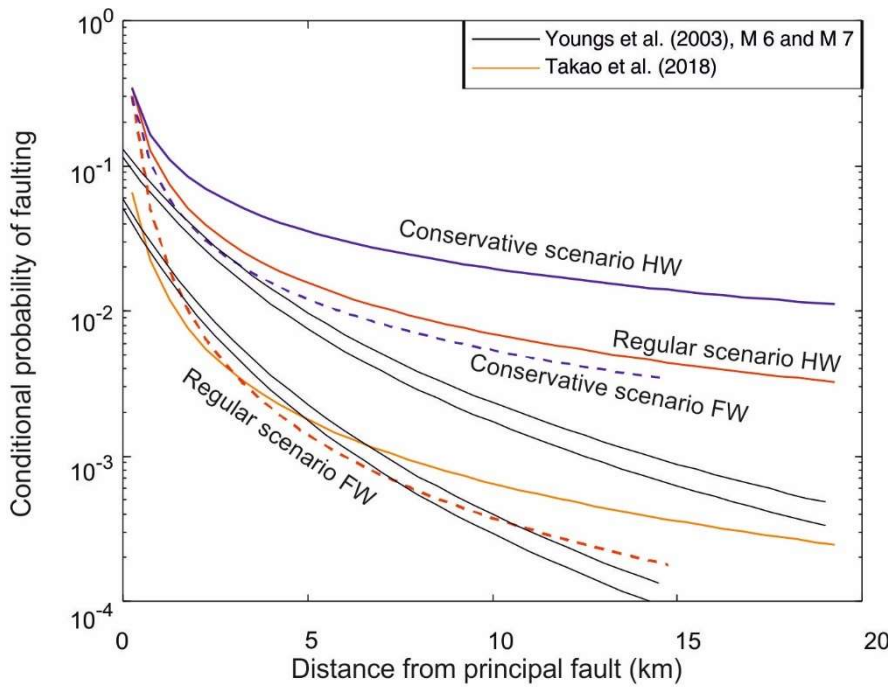
Figure 7: Comparison between the conditional probability of faulting for normal events obtained in this study and the ones by Youngs et al. (2003; the curves for M 6 and 7 are shown as a reference) and for strike-slip and reverse events in Japan (Takao et al., 2018); full lines: hanging wall, dashed lines: footwall.

Table 1: Details on the 21 earthquakes analyzed in the present study. Date expressed as year, month, day. Reference list the source of the map of surface ruptures; * in the reference column highlights the events already included in the work by Youngs et al. (2003).

ID	Date (yyyy-mm-dd)	Location	Country	M	Reference
01	1887-05-03	Sonora Valley	US	7.5	Baize et al. (2019) *
02	1915-01-13	Fucino	Italy	7.1	Amoroso et al. (2016)
03	1915-10-03	Pleasant Valley	US	6.8	Baize et al. (2019) *
04	1928-01-06	Laikipia-Subukia	Kenya	6.9	Ambrasyes (1991)
05	1954-07-06	Rainbow Mountain	US	6.6	Pezzopane and Dawson (1996) *
06	1954-08-23	Stillwater	US	6.8	Pezzopane and Dawson (1996) *
07	1954-12-16	Fairview Peak	US	7.1	Baize et al. (2019) *
08	1954-12-16	Dixie Valley	US	6.6	Baize et al. (2019) *
09	1959-08-18	Hebgen Lake	US	7.2	Baize et al. (2019) *
10	1970-03-28	Gediz	Turkey	7.0	Ambraseys and Tchalenko (1972)
11	1980-05-25	Mammoth Lake	US	6.2	Baize et al. (2019) *
12	1980-11-23	Irpinia	Italy	6.8	Blumetti et al. (2002)

19

13	1981-02-25	Gulf of Corinth	Greece	6.4	Jackson et al. (1982)
14	1981-03-04	Gulf of Corinth	Greece	6.4	Jackson et al. (1982)
15	1983-10-28	Borah Peak	US	6.9	Baize et al. (2019) *
16	1987-03-02	Edgecumbe	New Zealand	6.5	Beanland et al. (1989)
17	1997-09-26	Colfiorito	Italy	6.0	Cello et al., (1998)
18	2008-03-20	Yutian	China	7.1	Xu et al. (2013)
19	2009-04-06	L'Aquila	Italy	6.3	Baize et al. (2019)
20	2016-08-24	Amatrice	Italy	6.0	Livio et al. (2016)
21	2016-10-30	Norcia	Italy	6.5	Civico et al. (2018)

465 **Table 2: Data fitting and goodness of fit parameters for the traditional and conservative scenarios (see Eq. 2); regressed parameters, with 95% confidence bounds, are indicated. HW: hanging wall, FW: footwall, SSE: sum of squares due to error, RMSE: root mean squared error.**

Model	a	b	c	R2	Adj. R2	SSE	RMSE
HW, regular scenario	-2.254 (-2.318, -2.189)	-1.175 (-1.228, -1.122)	1e-05	0.9890	0.9887	0.001464	0.006291
FW, regular scenario	-3.459 (-3.638, -3.28)	-1.903 (-2.255, -1.55)	1.008e-05 (-0.08461, 0.08463)	0.9989	0.9988	9.969e-05	0.001922
HW, conservative scenario	-1.888 (-2.132, -1.643)	-0.8802 (-1.05, -0.7099)	1.009e-05 (-0.1406, 0.1406)	0.9559	0.9534	0.005919	0.01282
FW, conservative scenario	-2.505 (-2.934, -2.076)	-1.181 (-1.62, -0.7428)	1.006e-05 (-0.2128, 0.2128)	0.9548	0.9515	0.00399	0.01216