



Contents lists available at ScienceDirect

Physics of the Earth and Planetary Interiors

journal homepage: www.elsevier.com/locate/pepi

Why is Probabilistic Seismic Hazard Analysis (PSHA) still used?



Francesco Mulargia^a, Philip B. Stark^b, Robert J. Geller^{c,*}

^a Dipartimento di Fisica e Astronomia, Settore di Geofisica, Università di Bologna, viale Bertoni Pichat 8, 40127 Bologna, Italy

^b Department of Statistics, Code 3860, University of California, Berkeley, CA 94720-3860, USA

^c Department of Earth and Planetary Science, School of Science, University of Tokyo, Hongo 7-3-1, Bunkyo-ku, Tokyo 113-0033, Japan

ARTICLE INFO

Article history:

Received 30 September 2016

Received in revised form 8 December 2016

Accepted 10 December 2016

Available online 16 December 2016

Keywords:

91.30.Px Earthquakes

91.30.Dk Seismicity

91.55.Jk Fractures and faults

91.30.Za Paleoseismology

ABSTRACT

Even though it has never been validated by objective testing, Probabilistic Seismic Hazard Analysis (PSHA) has been widely used for almost 50 years by governments and industry in applications with lives and property hanging in the balance, such as deciding safety criteria for nuclear power plants, making official national hazard maps, developing building code requirements, and determining earthquake insurance rates. PSHA rests on assumptions now known to conflict with earthquake physics; many damaging earthquakes, including the 1988 Spitak, Armenia, event and the 2011 Tohoku, Japan, event, have occurred in regions relatively rated low-risk by PSHA hazard maps. No extant method, including PSHA, produces reliable estimates of seismic hazard. Earthquake hazard mitigation should be recognized to be inherently political, involving a tradeoff between uncertain costs and uncertain risks. Earthquake scientists, engineers, and risk managers can make important contributions to the hard problem of allocating limited resources wisely, but government officials and stakeholders must take responsibility for the risks of accidents due to natural events that exceed the adopted safety criteria.

© 2016 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

“The current approach to the estimation of probabilities of extreme floods and droughts is based on analysis of extremes in historic streamflow or precipitation records. The main weakness of the analysis is that it takes no account of the actual climatic, hydrological and other geophysical mechanisms that produced the observed extremes. Rather, it is based on arbitrary postulates of preconceived probabilistic mechanisms so that the results do not reflect what is likely to happen in nature, but what would happen if these postulates were correct. The crucial aspect, i.e. whether they actually are correct, is not addressed at all. The usual excuse for this is that there are not enough data, the physical processes are not understood well enough and the planners, engineers, managers, etc. cannot wait with their decisions until these problems are resolved; hence, to comply with the practitioners’ requests for estimates of 100-year, 1000-year, and (lately) even 1,000,000-year floods and droughts, these crude methods have to be used.”

“Without an analysis of the physical causes of recorded floods, and of the whole geophysical, biophysical and anthropogenic

context which circumscribes the potential for flood formation, results of flood frequency analysis as [now practiced], rather than providing information useful for coping with the flood hazard, themselves represent an additional hazard that can contribute to damages caused by floods. This danger is very real since decisions made on the basis of wrong numbers presented as good estimates of flood probabilities will generally be worse than decisions made with an awareness of an impossibility to make a good estimate and with the aid of merely qualitative information on the general flooding potential.”

[Klemeš, 1989]

1. Introduction

“Earthquake prediction,” which is generally defined as the issuance of a science-based alarm of an imminent damaging earthquake with enough accuracy and reliability to justify measures such as evacuations, has been the goal of unsuccessful empirical research for the past 130 years (Geller et al., 1997; Geller, 1997; Kagan, 1997; Main, 1999; Kagan, 2014). Thus earthquake prediction is not presently a realistic option for mitigating earthquake hazards.

Much effort is now devoted to making earthquake hazard maps, and it is important to evaluate their performance. The national hazard map of the former Soviet Union seriously understated the

* Corresponding author.

E-mail addresses: francesco.mulargia@unibo.it (F. Mulargia), stark@stat.berkeley.edu (P.B. Stark), bob@eps.s.u-tokyo.ac.jp (R.J. Geller).

risk in several areas, most notably for the 1988 Spitak, Armenia, event (Muir-Wood, 1993). Similarly, the Japan national hazard map seriously understated the risk in the area struck by the 2011 Tohoku earthquake (Geller, 2011). Hazard maps do not agree well with subsequent seismicity for several other cases as well (Stein et al., 2012; but also see a critical comment by Frankel, 2013, and the reply by Stein et al., 2013). The failure of the hazard maps appears to be due to the use of models such as the elastic rebound model or characteristic earthquake model, which are contradicted by observations (Kagan et al., 2012; Geller et al., 2015).

Seismology's low predictive power notwithstanding, human society nevertheless has been, and will continue to be, faced with seismic safety issues that arise when deciding whether to build, and how to design, structures including nuclear power plants, dams, expressways, railways, high rise buildings, schools, hospitals, etc. The considerations involved in designing an earthquake-resistant structure are technical, financial, legal, political, and moral. The stakeholders desire that structures be designed and constructed to be sufficiently earthquake-resistant, while also being as inexpensive as possible. The stakeholders will obviously also seek to minimize their financial and criminal legal exposure in the event of an earthquake-induced failure.

Someone (or some group) must decide whether to build a particular structure. If the decision is positive then someone (or some group) must specify the ground motion that the structure should be designed to withstand. No one would argue that such decisions should be made without any input from earthquake scientists—seismologists, geologists, and earthquake engineers. On the other hand, it is essential that such consultants properly explain the uncertainties of the information they are providing, so that society as a whole can make appropriate judgements regarding the various tradeoffs (Stein and Geller, 2012). Such explanations must make clear the existence not only of “known unknowns” (formal statistical uncertainties within the framework of a particular model) but also (in the phrase coined by Donald Rumsfeld, a former U.S. Secretary of Defense) of “unknown unknowns” (uncertainties due to the limitations of the models being used), the effects of which may be much larger than those of “known unknowns” but are extremely difficult to quantify. This is actually a familiar issue in applications of statistics to evaluating results of physics experiments to determine whether or not a hypothesized physical effect (e.g., “the fifth force”) is real (e.g., Anderson, 1992).

2. Probabilistic Seismic Hazard Analysis (PSHA)

The method now called Probabilistic Seismic Hazard Analysis (PSHA) (although occasionally “assessment” is used in place of “analysis”) is based largely on work by Cornell (1968). Programs for making calculations based on the methods of Cornell (1968) were published by McGuire (1976) and Bender and Perkins (1987). A cursory literature search suggests that the term “PSHA” came into use in the mid-to late 1970's, but this search is far from complete. A somewhat light-hearted tutorial on PSHA was presented by Hanks and Cornell (1994). By now PSHA is in extremely widespread use as a standard part of the “due diligence” process for designing and building critical structure (e.g., Solomos et al., 2008; Hanks et al., 2009).

PSHA has been extensively debated over the years. A sample of contributions since 2000 includes Castanos and Lomnitz (2002), Stark and Freedman (2003), Stein et al. (2003), Wang et al. (2003), Frankel (2004), Musson (2004), Reiter (2004), Budnitz et al. (2005), Holzer (2005), Klügel (2005a,b,c,d,e,f), Krinitzsky (2005), Lomnitz (2005), Musson et al. (2005), Wang (2005a,b), Klügel (2007, 2008), and Musson (2012). Wang (2012) noted some fine points involving Cornell's (1968) derivation.

Much recent interest has been focused on testing PSHA estimates against actual seismicity (Kagan and Jackson, 2000; Schorlemmer et al., 2007). The Collaboratory for the Study of Earthquake Predictability is attempting to set standards while extending forecasting to shorter time intervals (Jordan, 2006; Lombardi and Marzocchi, 2010; Mak et al., 2014). However, such testing sets are in many cases too small to allow any real power (e.g., Iervolino, 2013).

2.1. Cornell's work

We begin our discussion of PSHA by reviewing Cornell's (1968) explanation of the motivations and assumptions of his work, quoting excerpts from his introduction (but omitting cited references).

- (1) “Owing to the uncertainty in the number, sizes, and locations of future earthquakes it is appropriate that engineers express seismic risk, as design winds or floods are, in terms of return periods.”
- (2) “[The] engineer should have available all the pertinent data and professional judgement of those trained in seismology and geology in a form most suitable for making [safety decisions] wisely. This information is far more usefully and completely transmitted through a plot of, say, Modified Mercalli intensity versus average return period than through such ill-defined single numbers as the ‘probable maximum’ or the ‘maximum credible’ intensity.”
- (3) “The locations and activities of potential sources of tectonic earthquakes may be many and different in kind; they may not even be well known. In some regions, for example, it is not possible to correlate past activity with known geological structure. In such circumstances the seismologist understandably has been led to express his professional opinion in terms of one or two single numbers, seldom quantitatively defined. It is undoubtedly difficult, in this situation, for the seismologist to avoid engineering influences; the seismologist's estimates will probably be more conservative for more consequential projects. But these decisions are more appropriately those of the design engineer who has at hand more complete information (such as construction costs, and system performance characteristics) upon which to determine the optimal balance of cost, performance, and risk.”
- (4) “In this paper a method is developed to produce for the engineer the desired relationships between such ground-motion parameters as Modified Mercalli Intensity, peak-ground velocity, peak-ground acceleration, etc., and their average return period for his site. The minimum data needed are only the seismologist's best estimates of the average activity levels of the various potential sources of earthquakes (e.g., a particular fault's average annual number of earthquakes in excess of some minimum magnitude of interest, say 4). “...the technique provides the method for integrating the individual influences of potential earthquake sources, near and far, more active or less, into the probability distribution of maximum annual intensity (or peak-ground acceleration, etc.). The average return period follows directly.”

2.2. Are the assumptions made by PSHA scientifically valid?

Cornell's (1968) method allows all of the information provided by geologists and seismologists to be aggregated and used as the input to a “black box” procedure for providing curves that give the relation between ground motion parameters and “return periods” at a particular site. Engineers can then use these curves as the basis for choosing the design parameters of structures without

having to get involved in the details of the Earth-science related issues.

Cornell's method and its later extensions fill a need for a seemingly objective “due diligence” procedure, which neatly separates the roles of Earth scientists and engineers. Proponents of PSHA seem to regard its validity as firmly established, and to dismiss criticisms as unwarranted. For example, [Musson \(2012\)](#) made the following statement:

“It might seem odd to be writing about confirmation of the validity of probabilistic seismic hazard assessment (PSHA) in 2011, given that the method has been successfully applied in countless studies worldwide over the last 40 years. However, the fact that papers still occasionally find their way into print attacking the method as mathematically invalid seems to indicate that there is still some requirement, if small, to demonstrate the soundness of the method.”

As quoted above, [Musson \(2012\)](#) alludes to many successful applications of PSHA in “countless studies worldwide over the last 40 years.” Since this statement was neither sourced nor clarified it is not clear which PSHA studies were being alluded to nor what qualified them as “successes.” One can only guess, but presumably this statement is based on classifying all PSHA studies that were accepted by the clients or regulators as “successes.” However, the real success or failure of PSHA studies can only be evaluated by comparison to subsequent seismicity, and, as noted by [Muir-Wood \(1993\)](#), [Geller \(2011\)](#), and [Stein et al. \(2012\)](#), many large damaging earthquakes have occurred in areas categorized as low risk by PSHA-based hazard maps.

As noted by [Musson \(2012\)](#), most debate about PSHA has been focused on mathematical issues, which basically involve questions of internal consistency. In contrast, this paper is about physics and statistics, not mathematics. We consider the question of whether or not PSHA is scientifically valid, in view of what is now known about the earthquake source process. To our knowledge this question has not heretofore been systematically investigated. We discuss below what is now known about the physics and statistics of earthquake occurrence. We then show in Section 4 that the assumptions made by PSHA are not consistent with this knowledge.

3. Earthquake phenomenology—what is known?

As noted by [Ben-Menahem's \(1995\)](#) review of the history of seismological research, earthquake kinematics has been emphasized and progress toward understanding the physics of earthquake occurrence has been slow; this is still true today. This lack of progress of progress should probably be attributed to the difficulty of the problem rather than to a lack of appreciation of its importance. Studies by [Kanamori and Heaton \(2000\)](#) and [Mulargia et al. \(2004\)](#) may serve as starting points for further work.

Significant effort has been made towards applying the statistical mechanics of critical phenomena to earthquake occurrence (cf. [Rundle et al., 1995](#); [Turcotte, 1997](#); [Bak et al., 2002](#); [Mulargia and Geller, 2003](#); [Corral, 2003](#); [Turcotte et al., 2009](#)). The simple models (sandpiles, blocks and sliders, etc.) considered by such studies can provide some insights into earthquake-like phenomena. However, it is not clear whether these simple models are sufficiently realistic to assist in understanding actual earthquakes.

Seismic and geodetic data allow post hoc inference of the details of the mechanism of a given earthquake, and results from rock mechanics experiments have contributed to the physical interpretation of such results ([Scholz, 2002](#)). However, these experimental results have not advanced efforts to predict future large earthquakes. For example, after the 1999 Izmit, Turkey,

earthquake, [Parsons et al. \(2000\)](#) argued, on the basis of the Coulomb Failure Stress criterion, that “The Big One” in Istanbul was imminent, but no such event has occurred. Among the reason for the failure of such predictions is that it is practically impossible to measure point by point in the crust (a) the past stress history, (b) the material properties, and (c) the flow map of crustal fluids, which play a much more important triggering role than Coulomb stress itself ([Mulargia and Bizzarri, 2014, 2015](#)). More generally, see Section 10.1 of [Geller et al. \(2015\)](#) for a brief discussion of some other limitations on the applicability of laboratory rock mechanics experiments to actual earthquakes in the crust.

Two basic empirical statistical laws governing earthquake occurrence are well known: the Gutenberg-Richter (G-R) Law, which gives the relation between magnitude and frequency of occurrence, and the Omori Law, which gives the decay of after-shock activity with time. We now briefly review and summarize these laws. Note that these empirical laws specify frequency distributions, not probability distributions. See [Freedman \(1997\)](#) and [Stark \(2013, 2017\)](#) for a discussion of this important distinction.

3.1. Gutenberg-Richter law and modern extensions

As stated by [Gutenberg and Richter \(1956\)](#), the G-R Law was

$$\log_{10} N = a - bM, \quad (1)$$

where N is the number of earthquakes with magnitude great than or equal to M , and a and b are constants. Most studies find $b \approx 1$. If Eq. (1) held scale-invariantly as M becomes arbitrarily large, then the total energy released by earthquakes would diverge, so some deviation from scale invariance of Eq. (1) is required for large magnitudes. This could be either a hard cutoff (a maximum magnitude), or a soft cutoff (a “corner magnitude”) above which the frequency-magnitude curve falls off, say, exponentially, from the G-R straight line. (Because of atomic size there must also be deviation from scale-invariance as M becomes extremely small, but we will not discuss this further.)

Until the mid-1970s global earthquake catalogs used body wave magnitudes, m_b , and surface wave magnitudes, M_s ([Geller and Kanamori, 1977](#)). However, as seismology progressed it became generally recognized that earthquake magnitudes measured from 20 s surface wave amplitudes saturated at about $M_s = 8.5$ ([Kanamori and Anderson, 1975](#); [Geller, 1976](#)), and that the scalar moment, M_0 given by

$$M_0 = \mu S \bar{D}, \quad (2)$$

where μ is the rigidity of the material bounding the fault, S is the fault area, and \bar{D} is average slip, should instead be used to quantify the size of earthquakes. In order to preserve continuity with the magnitude scale, [Kanamori \(1977\)](#) defined the moment-magnitude, M_w , as follows:

$$\log_{10} M_0 = 1.5 M_w + 9.1, \quad (3)$$

where the scalar moment M_0 is quantified in units of N m. If M_0 is quantified in units of dyn cm then the constant in Eq. (3) becomes 16.1 rather than 9.1. Using Eq. (3) ensures that the magnitudes of extremely large earthquakes are properly quantified, while at the same time ensuring that smaller earthquakes have M_w values roughly equal to their M_s values.

3.2. Kagan's and related work on moment distribution

[Kagan \(2002\)](#) studied the moment distribution of earthquakes using data from the Harvard catalog (now Columbia University GCMT catalog), which is a basically homogenous catalog in which all moment values are determined using methods that have not

significantly changed since its inception. He analyzed global shallow earthquake data from the period 1977–2000.

In this subsection only we follow the notation of Kagan (2002). He denotes the scalar moment by M rather than by M_0 and the moment magnitude by m rather than by M_w . Kagan (2002) begins by rewriting the classical G-R law (Eq. (1)) to account for the fact that earthquake catalogs have a magnitude threshold m_t below which the catalog is incomplete. He wrote

$$\log_{10} N(m) = a_t - b(m - m_t) \quad \text{for } m_t \leq m, \quad (4)$$

where $N(m)$ is the number of earthquakes with magnitude greater than or equal to m and a_t is the (base 10) logarithm of the number of earthquakes with magnitude greater than or equal to m_t .

Using Eq. (3) we can define the slope parameter for the moment distribution, β , in relation to the slope parameter for the magnitude distribution, b , as follows:

$$\beta = \frac{2}{3}b. \quad (5)$$

If we differentiate Eq. (4) and use Eq. (5) we obtain the following functional form for the derivative of the frequency distribution, $\phi(M)$:

$$\phi(M) = \beta M_t^\beta M^{-1-\beta}, \quad (6)$$

where M_t^β is the moment for the threshold magnitude m_t . While this may look like a probability distribution (a Pareto distribution), actually treating it as a probability distribution is a huge leap, with no sound physical basis (cf. Freedman, 1997; Stark, 2013, 2017).

Fig. 1 (modified from Kagan, 2002) shows the data from the Harvard catalog (now the GCMT catalog) for all shallow (depth ≤ 70 km) earthquakes with magnitude greater than about 5.8 from 1977–2000. The top panel shows all the data, and the lower panel shows a close-up of the lower right part of the upper panel. The classic G-R curve for $b = 1$ (i.e., for $\beta = 2/3$) fits the data well up to about magnitude 7.7 or so, and then the data fall increasingly below the G-R curve, as required by physics if the energy released by earthquakes is to remain finite (see Kanamori, 1977). Kagan fits three distributions to the data in Fig. 1: the tapered G-R, the gamma distribution, and the truncated Pareto. A visual comparison shows that all of the last three distributions do a reasonable job of fitting the data, but which, if any, is “best” has not, to our knowledge, been established by objective testing. The equations for these three distributions are given by Kagan (2002) and will not be repeated here.

The tapered G-R distribution and gamma distribution both have in common that they overlay G-R at lower magnitudes but fall off exponentially (i.e., below the G-R curve) at magnitudes above the “corner magnitude,” which is a parameter in each of these distributions. The implication of this work for seismic hazard assessment is that rather than thinking in terms of a “hard” maximum magnitude we should be thinking in terms of a “soft” corner magnitude, with no “hard” upper limit on the magnitude. It is not clear whether this is sufficiently appreciated by workers on seismic hazard estimates.

Kagan and Jackson (2013) consider regional variations in corner magnitudes for different subduction zones. Although such differences may exist, the shortness of the available record makes it difficult to estimate them. This also applies to attempts to find regional differences in corner magnitude for intraplate regions, as the data are even sparser. Another important question is whether b -values vary spatially or temporally. As this topic is outside the main area of this paper a detailed discussion is omitted. However, as noted by Kemer and Hiemer (2015), claims of regional or temporal variations in b -values have a high likelihood of being artifacts.

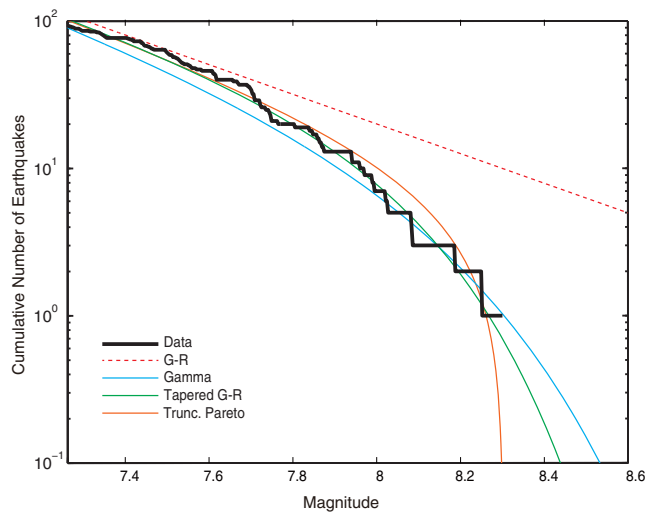
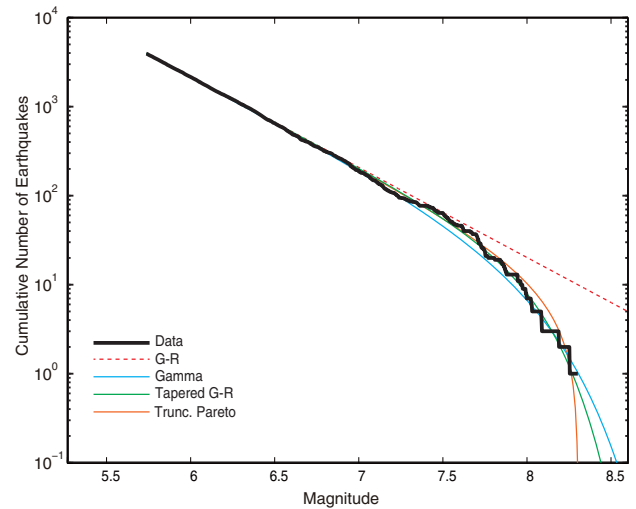


Fig. 1. (a) Global magnitude-moment relation after Kagan (2002). (b) Closeup view of lower right corner of upper panel.

3.3. Omori law and earthquake clustering

As mentioned above, the two fundamental empirical laws of seismology are the Gutenberg-Richter Law and Omori’s Law (Omori, 1895; Utsu et al., 1995). Recent work has generalized Omori’s Law in important ways, as we now discuss. These new results have significant implications for earthquake hazard estimation.

Our discussion is based on the Italian seismic catalog of Gasperini et al. (2013), which is the result of a careful equalization of bulletin records (without any forced compliance to the Gutenberg-Richter or Omori laws), correcting for the individual responses of each sub-net, for the type of instruments, operational periods, methods of analysis used, etc. (Lolli et al., 2014).

The catalog starts on January 1, 1981, is complete for magnitudes above 2.5, and consists of 20,547 events, including 19 M5.5+ destructive earthquakes. It has a non-constrained b value of 0.99 and fits the G-R law well (see Fig. 2). Fig. 3 shows the instantaneous rate of event occurrence versus time for different magnitudes, illustrating through its vertical “pinnacles” one main feature of earthquake occurrence, namely, clustering in time. As shown by Fig. 4, for $M > 2.5$ events, each cluster has a rapid increase in the rate of event occurrence at the beginning of a clus-

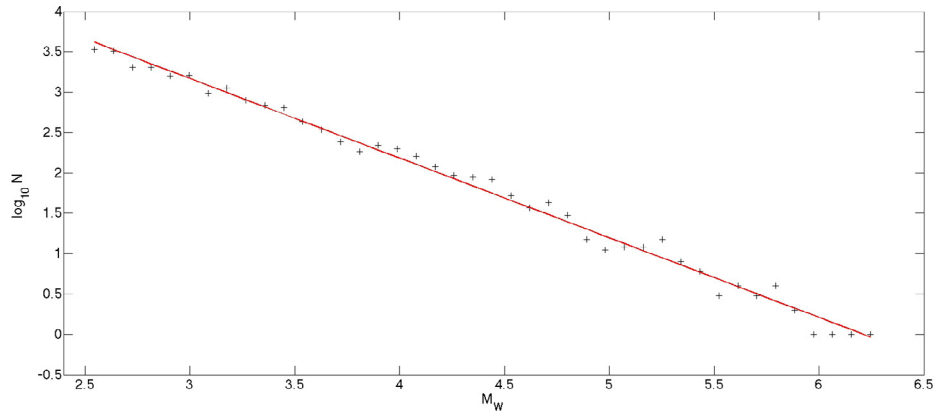


Fig. 2. The event density versus magnitude for the New Italian Seismic Catalog (Gasparini et al., 2013) with the related Gutenberg-Richter curve.

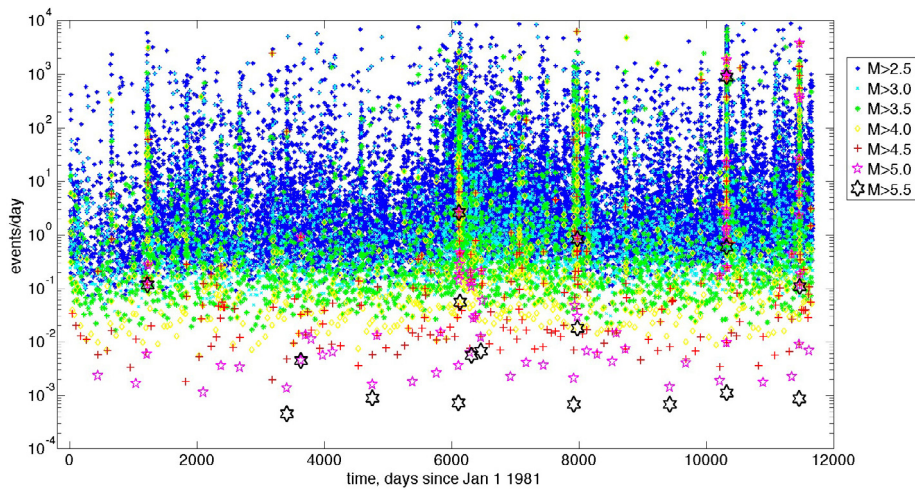


Fig. 3. The rate of event occurrence at various magnitudes versus time for the New Italian Seismic Catalog (Gasparini et al., 2013).

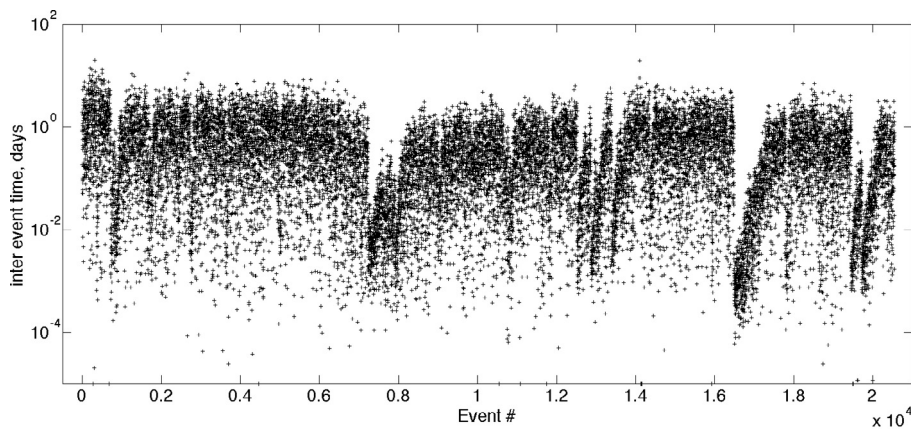


Fig. 4. The inter-event times versus sequential event number for the $M \geq 2.5$ earthquakes in the New Italian Seismic Catalog (Gasparini et al., 2013).

ter (and thus a sharp decrease in inter-event time), and a subsequent smooth decrease in the rate of occurrence (and thus a smooth increase in the inter-event time). Fig. 5 shows a close-up of the rightmost portion of Fig. 4.

The behavior seen in Figs. 4 and 5 is well represented by the Omori power law

$$\lambda(t) = \frac{k}{(c + t)^d}, \quad (7)$$

with $d \sim 1$, and where $\lambda(t)$ is the frequency of event occurrence at time t after the initial event, and k and c are constants (see Utsu et al., 1995).

Clustering in time (Kagan and Jackson, 1991) is accompanied by clustering in space, as is apparent in Fig. 6. Note that a clear simultaneous time–space picture is readily obtainable due to the geographic shape of Italy, which is essentially a North–South segment with earthquakes distributed along its structural backbone, the Apennines.

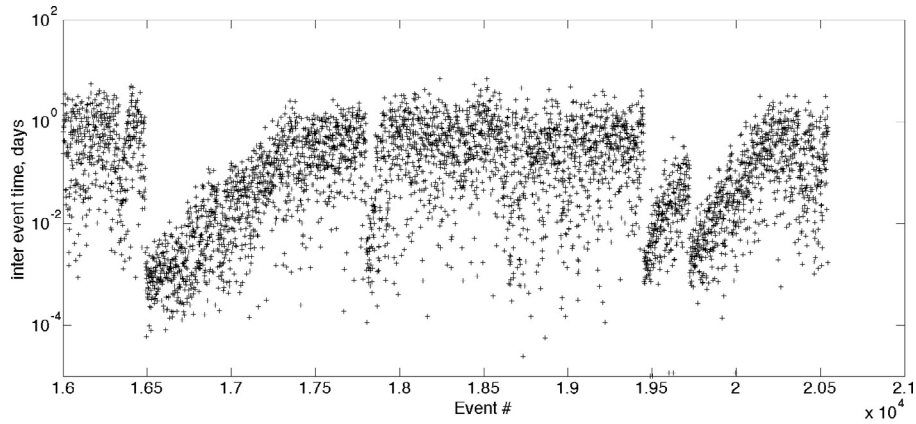


Fig. 5. The same as Fig. 4 zoomed in on the last part to show the details of event clustering.

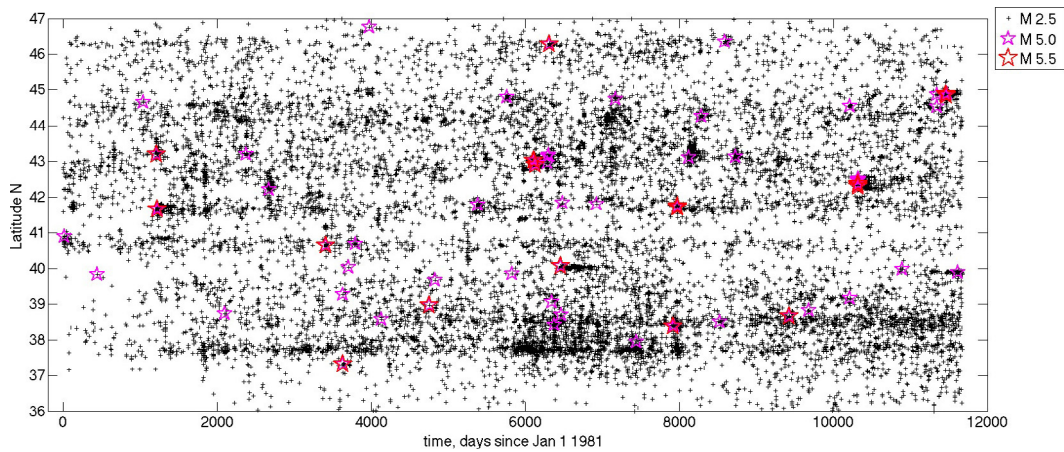


Fig. 6. Space–time diagram of the $M \geq 5.5$ and the $M \geq 5.0$ events shown together with the $M \geq 2.5$ events in the New Italian Seismic Catalog (Gasperini et al., 2013). Events are plotted in latitude vs. time due to the fact that Italy has a (tilted) N-S alignment.

Another major phenomenological feature accompanying clustering is scale invariance. A fundamental result of Bak et al. (2002) and Corral (2003) is that an identical picture exists (Fig. 7) for the rescaled recurrence times of all the events independent of magnitude, as is shown for the distribution density function $D(\tau)$ of the inter-event times τ , renormalized to the occurrence rate over a time interval in which it is approximately constant

$$R = \lambda \neq \lambda(t). \tag{8}$$

(In discussing scale invariance here we neglect what happens above the corner magnitude.) This same scale-invariant picture was found to occur in many regions, irrespective of the spatial extent, magnitude threshold, and the seismic catalog from which the data were taken (Corral, 2003, 2004). Our analysis of the Italian catalog differs from Corral (2003, 2004, 2005a) in that that we deal with the whole catalog, without any selection of stationary periods, i.e., without using Eq. (8), to avoid any arbitrary choice. This leads to a little blurring of the picture (compare Fig. 7 to similar figures in Corral, 2003,

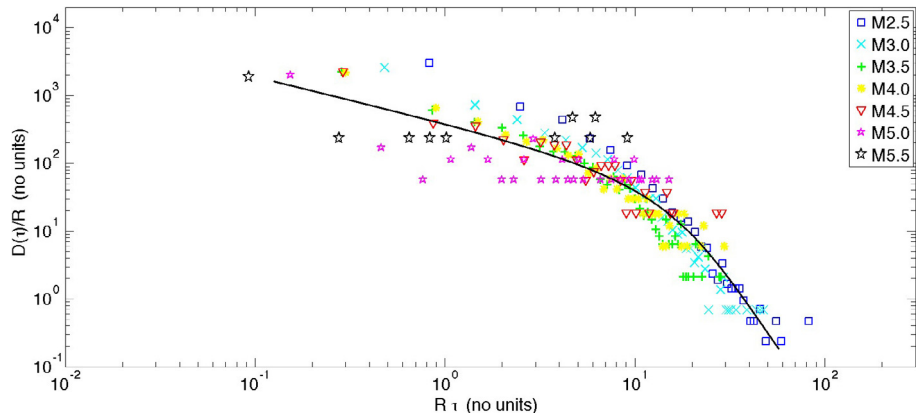


Fig. 7. The rescaled density of the rescaled interevent-times (see text) for all the event magnitude classes of the New Italian Seismic Catalog (Gasperini et al., 2013). The best-fitting gamma density is also shown.

2005a), but nevertheless retains the same general features. The mathematical environment in which this can be framed is the renormalization group (Corral, 2003, 2005a).

In quantitative terms, the scaling law

$$D(\tau) = Rf(R\tau) \quad (9)$$

holds in general, due to the self-similarity in space and time of earthquake occurrence. The functional form of f is close to a generalized gamma distribution

$$f(\theta) = \frac{|\delta|}{a^\gamma \Gamma(\gamma/\delta)} \frac{1}{\theta^{1-\gamma}} e^{-(\theta/a)^\delta}, \quad (10)$$

where $\theta = R\tau$ is a dimensionless recurrence time, $\gamma \sim 0.7$, $\delta \sim 1$, $a \sim 1.4$, and Γ is the gamma function. Thus f decays as a power law for $\theta < 1$ and exponentially for $\theta > 1$.

The crucial physical implication is not just the specific functional form of Eq. (10)—which we call the Bak-Corral equation—but the universality that it implies through Eq. (9), which appears to apply irrespective of event size and spatial domain. This constitutes a basic property of earthquakes, which appear always to occur in clusters with an identical scale-invariant behavior. Rescaled by rate of occurrence R , this scale-invariant pattern occurs identically both in “obvious” clusters as well as in periods where no obvious clustering is apparent (i.e., in periods which Corral calls “stationary”). This can be viewed as a generalization of Omori’s law, but its meaning is more fundamental: while Omori’s law just governs the decay of the occurrence rate following a large earthquake, i.e. “the aftershocks of a mainshock,” the Bak-Corral equation governs the occurrence of all earthquakes, independent of their size and of the region considered.

The behavior discussed above is phenomenological and does not imply any specific physical model. In practical terms, it implies: (1) positive clustering at short time scales, with events tending to occur closer to each other in time and space, but (2) negative clustering at long time scales, with sparse events tending to be even sparser. This leads to the counter-intuitive consequence that the waiting time until the next earthquake becomes longer the longer one has been waiting for it. This is at odds with both elastic rebound, which would imply shorter waiting times as the elapsed time becomes longer, and with Poisson models, for which the waiting time does not depend on the elapsed time (Davis et al., 1989; Corral, 2003; Mulargia, 2013).

Earthquakes do exhibit “memory” in several respects and there exist a number of correlations in seismicity (Corral, 2005a, 2006). The effect of memory is obvious in Omori’s law and in the time evolution of any cluster, as illustrated by Figs. 4 and 6. Note that this also occurs identically outside major clusters, in the time intervals called “stationary” by Corral (2003, 2007), due to the fact that clustering is always present.

The existence of memory may appear surprising when considering large areas, since one might think that the superposition of the seismicity of many different areas could be viewed as the sum of many independent processes, so that the seismicity of a broad region should converge to a Poisson process (Cox and Lewis, 1966). However, as discussed above, this does not occur in practice for Italy (Fig. 7) or for the whole world (Corral, 2004).

We can summarize earthquake memory effects as follows. (1) Pre-earthquake conditions are all alike due to the lack of correlation of the magnitude with previous seismicity, confirmed also by paleoearthquake data for single faults (Weldon et al., 2004; Cisternas, 2005). (2) The larger an earthquake, the shorter the time until the next one, due to the negative correlation between magnitudes and the occurrence time of the next earthquake. (3) The shorter the time from the last earthquake, the shorter the time

until the next one. (4) The longer the time since the last earthquake, the longer the expected time until the next one (Davis et al., 1989; Corral, 2005b).

No physical or statistical model developed to date matches this phenomenology. In particular, the ETAS model (Ogata, 1988, 1999), which combines the Gutenberg-Richter law, the Omori law, and an aftershock production rule, is not scale-invariant.

4. PSHA: assumptions vs. reality

As discussed in Section 2, in the almost 50 years since Cornell (1968) was published, PSHA has become a standard tool for hazard analysis, despite receiving some criticism. Current procedures for implementing PSHA are discussed in detail by Hanks et al. (2009) and Kammerer and Ake (2012). PSHA is based on the classic Cornell (1968) approach. The seismic hazard at a given site is estimated (Stirling, 2014) based on (a) the location of the site with respect to known or assumed earthquake sources, (b) the assumed recurrence behavior of these earthquakes sources, and (c) the computed ground motion for the earthquakes at the given site.

In this paper we make the case that PSHA is fundamentally flawed and should therefore be abandoned. The results generated by PSHA should therefore not be used by society as the basis for decisions on seismic hazard mitigation. Our main concern is not with the numerical inputs used by various PSHA practitioners, but with the methodology itself.

As discussed below, our case rests on three main arguments. (1) PSHA makes assumptions that contradict what is known about seismicity. (2) PSHA fundamentally misuses the concept of “probability.” (3) In practice PSHA does not work; as noted above (and see Stein et al., 2012) many recent destructive earthquakes occurred in regions that PSHA identified as low risk.

We hope that researchers who disagree will respond by attempting to present physics-based data and analyses that refute the above three arguments. If this cannot be done, then, like other once widely accepted theories or methods that did not agree with observations or experiments, PSHA should be scrapped. We emphasize that the arena for the debate on PSHA should be that of physics, rather than mathematics, taking fully into account what is now known about the phenomenology of earthquake occurrence, as discussed above in Section 3.

4.1. PSHA cookbook

In more detail, the “PSHA cookbook” (e.g., NEHRP, 1997; USGS-WGCEP, 1999; Kammerer and Ake, 2012) follows the following basic steps:

- (1) identify seismically active regions, structures, and faults using earthquake catalogs, geologic evidence, and geodetic strains (derived from GPS data);
- (2) for each fault or region, estimate the average event rate, using both instrumental and historical seismicity data, and geodetic strains;
- (3) for each source, using a physical model, infer the average future occurrence rate $\lambda = N/\Delta t$, (where N is the number of events in the time interval Δt), commonly expressing it in terms of the recurrence time: $\tau = 1/\lambda$;
- (4) attenuate the spectrum of ground motion from each source to the given site through an empirical function derived from the ground motion prediction equation (GMPE);
- (5) forecast the ground motion, parameterized as an engineering variable such as the peak ground acceleration (PGA), in terms of the probability of exceedance p or its reciprocal,

the expected return period, $T = 1/p$, of a given level of shaking; for example, a 10 percent probability of exceedance in 50 years is equivalent to a return period of 475 years;

- (6) sum the exceedance probabilities of the different sources to account for the fact that a given site is potentially subject to shaking due to a variety of earthquakes sources;
- (7) use the forecast values as design parameters for the building or other structure being designed.

Additional steps can also be introduced. Beyond specific choices about extrapolating the smaller magnitude event rates to the “maximum credible earthquake” (Wang, 2008), the most common additions are as follows (NEHRP, 1997; Kammerer and Ake, 2012):

- (8) assume that the earthquakes will repeat identically not only on the same faults, but with the same size and mechanism and produce a spectrum identical to past occurrences, and use this to forecast the expected ground motion at each frequency (the response spectrum) in terms of exceedance probabilities;
- (9) estimate the local amplification at each site from the local shear wave response, which is primarily determined using the shear wave velocity profile versus depth in the uppermost 30 m (V_{s30}).

We now consider whether the above cookbook can be expected to provide reliable and accurate hazard analyses, given the available data and present knowledge of the physics of earthquake occurrence.

4.2. Insufficiency of seismicity data

PSHA estimates the seismicity rates of a given region based on data on past earthquakes. Ideally, this requires a complete catalog of all significant earthquakes for an extensive time period. Unfortunately, however, catalogs based on instrumental recordings are available only since about 1900, and are only reliable after about 1960. Historical data for about 2000 years is available for some countries (e.g., Italy, China, Japan), but even though the dates and approximate epicenters are probably reliable, it is difficult to obtain reliable magnitudes and focal mechanisms, and the catalogs are incomplete in unknown ways. Even the Italian macroseismic catalog, which for historical and geographic reasons is of high quality—the area of Italy is small and it has been a site of civilization for over two millennia—rarely reports more than one significant occurrence in any given seismic region (Boschi et al., 2000). In principle, palaeoseismology can extend the time span of seismic catalogues to thousands of years, allowing the inclusion of several large earthquakes on the same tectonic structure (Sieh et al., 1989). However, exposed faults are rare, so such direct estimates are generally unavailable (Mulargia and Geller, 2003).

An indirect estimate can be achieved by extrapolating the G-R law from lower magnitudes, making use of the main phenomenological feature of earthquakes: scale invariance. Thus the behavior of large earthquakes, for which data are insufficient, can be inferred from small seismicity, for which data are available. However, in addition to the general risks and inaccuracies of any extrapolation, this procedure is also subject to great uncertainty depending on whether a maximum event size (Wang, 2008) or a corner magnitude (Kagan, 2002) is assumed.

The lack of a sufficiently long time series may lead to errors, since often “the largest historic magnitude” is taken as an upper limit, which is inappropriate (Mulargia, 2013). The 2011 Tohoku earthquake is a good example of a magnitude 9.0 earthquake happening where a maximum magnitude 8.0 had been wrongly

specified by official Japanese government forecasts (Geller, 2011; Stein et al., 2012; Kagan and Jackson, 2013).

If, hypothetically, seismic catalogs 10⁵ years long were available, seismicity might appear to be nearly time independent on that time scale, because the time scale of geological processes, on the order of $\sim 10^6$ years, is much longer than that of inter-earthquake intervals in seismically active regions, $10^2 \sim 10^4$ years. Unfortunately, only very short seismic catalogs are available, so it is unlikely that the available seismicity data are time-independent. Also Liu et al. (2011) found considerable variability in the spatial pattern of seismicity in North China over the past 2000 years.

In conclusion, even if there were no methodological problems afflicting PSHA, which, as we discuss below, is not the case, the insufficiency of the available seismicity data means that it is highly questionable that PSHA can provide reliable and accurate hazard analyses.

4.3. PSHA's dubious basic assumption

The word denoted by the first letter of the acronym PSHA is “probabilistic.” PSHA proponents appear to take for granted that earthquakes truly occur randomly in time and space, so it is reasonable and natural to estimate the spatio-temporal probability distribution of events, and to use that estimate as the foundation for risk assessment (see, e.g., Musson, 2012). Values of such probabilities are key inputs to PSHA, which then produces numbers such as “the annual probability of a peak ground acceleration of 0.2 g” at a particular site.

PSHA assumes that “the probability of an earthquake at a particular site” or the probability of “how often, both in terms of inter-event time and magnitude-frequency, earthquakes occur” exists. There are data on occurrence dates, locations, and magnitudes (subject to various errors and catalog incompleteness uncertainties) of past earthquakes. Let us call these “frequency” data. The information PSHA requires is the “probability” of future earthquakes. PSHA simply conflates the observed empirical frequencies with probabilities.

But frequencies are not probabilities, nor are they necessarily estimates of probabilities. If seismicity were the result of a random process, its empirical frequencies could be used to estimate the underlying probabilities. However, unless we have exogenous knowledge that the empirically observed frequency of past seismicity is due to something random, defining a quantity called the “probability of future seismicity” is just an arbitrary assumption about the mechanism that generates seismicity.

There are several conceivable potential justifications for modeling seismicity as random. For instance, there might be an underlying theoretical basis in physics—as there is in quantum mechanics—but that does not appear to be the case for earthquake physics. Or modeling seismicity as random might lead to a compact and statistically adequate description of the data, but in fact seismicity is empirically inconsistent with random seismic models that have been proposed (Luen, 2010; Stark, 2013). Or perhaps modeling seismicity as random might lead to useful predictions or hazard analyses, but, as we have discussed above, they do not. For further discussion of the fundamental issues involved here see also Stark (2017).

The above notwithstanding, Musson (2012), a prominent PSHA proponent, argues as follows.

“In a die-based experiment, the statement ‘the probability of rolling a two on this die is 0.166667’ (i.e., this die is fair) could be tested by making a suitable number of rolls and observing directly the number of twos that are rolled as a fraction of the total number. In seismology, a similar solution could be applied

to answering the question ‘What is the probability that tomorrow there will be an earthquake larger than 6 Mw somewhere in the world?’ It would be sufficient to collate the data for the past 1000 days and observe on how many days an earthquake above 6 Mw was recorded. To answer the question ‘What is the annual probability of 0.2 g PGA at my site?’ is more intractable, as the data are insufficient, and there may actually be no past observations of the target condition.

“The test of a statement about earthquake ground motion would be to make observations over a long enough number of years (say, 100,000) and count the number of times an acceleration over 0.2 g is recorded. Or even better (but even more impractical), observe the same year 100,000 times in independent parallel universes. This would give a very classical frequentist estimate of the probability, based on simple counting.

“While one cannot in actuality venture into multiple parallel universes, one can very easily do the next best thing, which is to simulate what would happen if one could. One can think of a PSHA study as consisting of two parts: a model (represented by a PSHA input file) and a process (represented by a seismic hazard program). The model is essentially a conceptualization of the seismic process expressed in numerical form, describing (1) where earthquakes occur, (2) how often they occur, both in terms of inter-event time and magnitude-frequency, and (3) what effects they have. With these three elements, one describes everything that determines the statistical properties of the seismic effects that will occur at a given site in the future. This is, therefore, all that one needs to simulate that future. One does not know precisely what will happen in the future, one only knows the aggregate properties of the seismicity and the ground motion propagation. Therefore simulations need to be stochastic. For each simulation, the probability density functions for earthquake occurrence in the model are randomly sampled to produce one possible outcome compatible with the model.”

It appears to us that Musson's argument is circular: it commits the well known fallacy of assuming that the real world is governed by probability in the same way that games of chance are (e.g., Taleb, 2007), and it rests on conflating probability with frequency and on conflating estimates of parameters with the true values of the parameters. Musson's argument implicitly assumes that earthquakes are generated by a stationary random process that is known except for a few parameters. Then he asserts that by observing seismicity for a long enough time period we could estimate the values of those parameters well. Once we know the parameters well, we can simulate additional (synthetic) seismicity, which we can then use to find other probabilities numerically. But where is the evidence that games of chance are a good model for earthquakes—i.e., that earthquakes are generated by a stationary random process? None is presented by Musson.

The models used in the common formulations of PSHA are the characteristic earthquake model and the Poisson process model. We now discuss the validity of these models.

4.4. The characteristic earthquake model

The characteristic earthquake model views large earthquakes in a particular region as the result of a cycle in which elastic strain steadily accumulates and is then released by the “characteristic earthquake” for that region when some limit is reached. Each repetition of the cycle is posited to result in a repetition of the “same” characteristic earthquake. This model is based on classic ideas proposed by Gilbert (1884) and Reid (1911), and was given widespread currency by Schwartz and Coppersmith (1984). In a recent paper (Geller et al., 2015), the present authors reviewed

the characteristic earthquake model and discussed its failure to agree with observed data. Needless to say, it is a basic principle of physics that any model or theory, however intuitively appealing it may be, and regardless of whether or not it has a majority of supporters in some community, must be discarded if it does not agree with observed data, and we argued that this is clearly the case for the characteristic earthquake model.

Kagan et al. (2012) pointed out that even though it has been refuted by a series of studies, the characteristic earthquake model, together with related concepts such as seismic gaps, and the seismic cycle, continues to be widely used. How can this situation persist? The following comment by Parsons et al. (2012) provides a clue:

“The controversy between the characteristic and Gutenberg-Richter [sic] models was partly developed through discussions by Nishenko and Sykes (1993), Jackson and Kagan (1993), Kagan and Jackson (1995) and Rong et al. (2003). This debate did not have a clear conclusion and is ongoing. However, the recent great M = 9.1, 2004 Sumatra and M = 9.0, 2011 Tohoku earthquakes do not seem to fit the definitions of characteristic earthquakes that were expected in their respective regions. The approximate interval between great earthquakes in the Nankai zone is about 150–600 yr, meaning that empirical resolution of this issue could require many hundreds of years.”

We can see three factors at work here. First, it is up to editors and referees whether or not to allow authors to continue to invoke refuted theories and models simply by asserting that the “debate did not have a clear conclusion and is ongoing.” In order to be allowed to make such an assertion the authors should have been required either to cite a specific recent study that supports the model in question and refutes the criticisms thereof, or to make such a case themselves. We see no evidence of either in Parsons et al. (2012).

A second key point is the framing of the issue. Parsons et al. (2012) have framed the issue as “are great earthquakes in the Nankai zone characteristic or not?” It seems to us, however, that this framing is an unacceptable departure from the fundamental principles of physics, which seeks universal laws. In the absence of a specific and well documented reason why the physics of great earthquakes in the Nankai zone should be different from that of great earthquakes at other subduction zones throughout the world, Occam's Razor, the principle of choosing the simplest possible hypothesis, tells us we should test the hypothesis “all great subduction earthquake are characteristic,” rather than testing each individual subduction zone as a separate case. In fact, work by the UCLA group (e.g., Rong et al., 2003) has done just that, consistently obtaining the result that the characteristic earthquake model (which can also be called the seismic gap model) is not successful beyond chance. While it might take hundreds of years to obtain a meaningful result at one particular site, the UCLA group has repeatedly shown that ten years is sufficient to obtain a meaningful result if data from all of the world's subduction zones are used.

Finally, a third point is the hypothetical question of what would happen if, as suggested by Parsons et al. (2012), we were to wait hundreds of years or longer and obtain data on four or five great earthquakes in the Nankai zone. Even if these great earthquakes appeared to be periodic (or somewhat periodic), as was pointed out by Savage (1993), that could either be because the characteristic earthquake model was correct, at least for the Nankai zone, or it could be accidental. Thus a test of the type proposed by Parsons et al. (2012) would not be conclusive.

The subject of this paper is the appropriateness of using PSHA in applications where lives and the safety of major infrastructure

hang in the balance. Such applications should be based only on well validated methods with a sound physical basis. The spectrum of recent views on the characteristic earthquake model lies between complete refutation (Kagan et al., 2012; Geller et al., 2015), and “ongoing debate” (Parsons et al., 2012); no recent work in peer-reviewed journals makes a strong data-based case in support of the characteristic earthquake model. Thus use of the characteristic earthquake model in real-world hazard analysis applications should be eschewed.

4.5. Exceedance probability, return time, and Poisson process model

In the Cornell formulation PSHA estimates are generally based on the return time, τ , and the exceedance probability, p , both sometimes also used in hydrology (Gupta, 1989). Note that the use of an exceedance probability implicitly assumes that earthquakes occur at random: some (stationary) stochastic mechanism generates seismicity, so the parameter p exists in the world and can be estimated from observed occurrence data.

The simplest stochastic model proposed for earthquakes is a Poisson process, for which waiting times between events are independent, identically distributed exponential (memoryless) random variables. The memoryless property of the exponential distribution contradicts the phenomenology—earthquake clustering—discussed above in Section 3.3. For Poisson processes, the theoretical “recurrence time” is simply the reciprocal of the theoretical rate of events per unit time. However, inter-arrival times for a Poisson process are highly variable: the “recurrence time” would be useless for prediction, even if it were known perfectly.

Because actual seismicity has more clustering than a Poisson process can account for, it is common to estimate the rate in Poisson models of seismicity after first “declustering” the catalogs to try to make them more compatible with the Poisson assumption. Declustering methods attempt to remove foreshocks and aftershocks, leaving only mainshocks. However, the definition of mainshock is effectively circular and algorithm-dependent: “mainshocks” are those events that the declustering algorithm does not delete. Also, in many cases aftershocks can be destructive, so removing them from a hazard assessment seems inappropriate.

Declustering algorithms may be divided into two main classes (Luen and Stark, 2012): mainshock window methods and linked-window methods. Mainshock window methods remove all earthquakes in a space–time window around every “mainshock,” but the definition of “mainshock” is subjective and varies from author to author. A prototype of this approach is the algorithm of Gardner and Knopoff (1974). Linked-window methods (e.g., Reasenber, 1985) allow a posteriori identification of mainshocks, aftershocks, independent events and, for some specific algorithms, also foreshocks. However, the choice of parameters is subjective, as is the choice among the various options. In conclusion, declustering discards the principal phenomenological feature of earthquake occurrence—clustering—and replaces it by a subjective entity—mainshocks. As such, it appears to be illogical and thus unacceptable. Furthermore, current declustering approaches are faulty even from a strictly statistical viewpoint, since the residual catalogs they produce do not fit the Poissonian assumption (Luen and Stark, 2012).

4.6. Logic trees: Does the sum of ignorance lead to knowledge?

Some applications of PSHA seek to combine all possible options into a “logic tree,” more or less subjectively assigning a probability to each branch, which will result in a wide variety of scenarios; these are presumed to provide a thorough basis for decision makers. This approach acknowledges our ignorance, but treats it by trying to reach an expert consensus. Such techniques have their origin

in the Delphi Method (Dalkey and Helmer, 1963), which was used to assimilate expert opinions on possible outcomes of a nuclear war. In practice, a large—but not too large—group of experts (typically 7–12 in number) is convened and brainstorming is conducted to search for a consensus. What the numbers obtained from a logic tree mean is less than clear. Failures of engineered systems with a small number of well-understood components whose failure modes are known might sometimes be fruitfully modeled as a tree with known nodes and branches and estimable probabilities at each node, at least as a thought experiment. But there is no evidence that this is meaningful or useful for phenomena as complex and poorly understood as earthquakes.

The use of “expert opinion” in PSHA studies implicitly acknowledges the unscientific nature of such endeavors. For example, expert opinion is not needed to make calculations involving Newton’s law of gravitation; barring computational blunders, a high-school student and a Nobel prize-winning physicist will calculate the same forces from the same input data.

As described in Section 2.1, the essence of Cornell’s (1968) approach is to separate the tasks of the seismologist and engineer. The job of the former is to provide “best estimates of the average activity levels of the various potential sources of earthquakes,” and the latter does the rest. This division of labor has the effect of shielding the engineer from responsibility if the “best estimates” fail to match subsequent earthquake activity. The engineer is also shielded from the need to study and understand the current uncertainties of the seismological state of the art.

5. Discussion and conclusions

Let us now return to the basic issue discussed in the introduction. Society will continue to consider building large-scale structures, including major industrial facilities such as nuclear power plants, in earthquake-prone regions. At the end of the day, someone or some group must decide either to approve or reject the proposed construction. Society’s needs being what they are, it seems inevitable that at least some, and perhaps most, of these applications will be approved. In the event of approval, someone or some group must decide on quantitative criteria for the strength and duration of ground motion (and tsunami, etc.) that must be designed for.

The tradeoffs in this process are obvious. Safety is desired, but so is lower cost. There is great implicit, and often direct, pressure on geoscience consultants, architects, and engineers to reduce costs by lowering the design criteria. On the other hand, in the event of an earthquake-caused failure of the structure or of critical equipment (e.g., a nuclear reactor) inside the structure, all parties involved in this process have strong incentives to minimize their financial and criminal liability. PSHA, as based on Cornell (1968), is an ideal solution to this conundrum. Geoscience consultants feed information into the PSHA machine, the crank is turned, and a seemingly objective number pops out. Basing decisions on this number effectively absolves engineers, architects, project developers, and owners from responsibility and liability, since the PSHA methodology is deeply entrenched as a standard element of “due diligence.” However, the fly in the ointment is that, as discussed above, the numbers that pop out of the PSHA machine have no theoretical or empirical justification, and thus do not quantify risk reliably or accurately.

The situation in our field is thus the same as that in hydrology, as discussed by Klemeš (1989) in the quotation at the beginning of this paper. Society would be far better off if all participants in the earthquake design process honestly acknowledged the grave shortcomings of PSHA and scrapped it rather than continuing to provide numbers for use in life-and-death earthquake safety decisions

based on an as yet unvalidated theory which, as shown above, disagrees with the physics and statistics of earthquake occurrence.

What should we be doing instead? The answer does not lie in making minor (or even major) adjustments to PSHA as now practiced and relaunching it, as the problems go far deeper. What has to happen instead is that rather than leaving the choice of earthquake design criteria up to only geoscience and engineering consultants, all participants—including project developers and owners, risk managers, engineers, and architects—have to “take ownership” of the inherent uncertainties of earthquake hazards, rather than simply tasking everything to geoscience and engineering consultants. In recent years administrative procedures in the European Union have adopted what is called the “precautionary principle” (Foster et al., 2000). The implementation of this principle in individual cases is still evolving, but it may provide some guidance here.

In the roughly 50 years since the work of Cornell (1968) PSHA has grown into a minor industry, and we recognize that there may be strong resistance to our call for its abandonment. However, the facts must be squarely faced: PSHA is based on flawed physical models, and its hazard forecasts do not agree with subsequent seismicity. That being the case, a new way forward must be sought, starting again from square one.

5.1. Some thoughts on the way forward

Having pointed out the grave flaws of PSHA-based due diligence, it behooves the authors to make some suggestions about where we might go from here. One major conclusion of this paper is that responsibility for seismic safety measures must be shared among all parties, rather than just being tasked to geoscience and engineering consultants. The obvious place to start is by looking at risks, rather than just hazards. In many countries (e.g., Greece, Turkey, Armenia, Italy, etc.), a majority of the building stock has little or no earthquake resistance. Even in Japan and California, many older buildings can be readily identified as lacking sufficient strength. A reasonable first step might be to focus on the vulnerability of individual buildings and systems. This might include studies of the buildings' and systems' capacity to resist various earthquakes, including beyond current-code earthquakes.

A second step is an open acknowledgment that “absolute safety” is impossible to achieve. Just as in automotive and air traffic, there will always be accidents with casualties; the only realistic goal is to reduce their occurrence and effects as much as possible, within the limits of society's ability and willingness to pay (Meehan, 1984; Geller et al., 2013). If risk is acknowledged, then open discussion of how to make structures “fail soft” will become possible.

A third step concerns how to most economically make progress and how to minimize litigation. This requires a clear sharing of responsibilities and costs among all of the stakeholders. The measures required (new laws and government regulations) are outside the authors' professional expertise, but we suggest that a first step should be for geoscience and engineering professionals to communicate the full extent of the uncertainties, including an unqualified repudiation of PSHA, to lawmakers and cognizant government officials.

Fourth, correct information on seismic risk must be supplied to the public in understandable form, so that each individual can take appropriate measures to prepare for future earthquakes.

Finally, everyone involved in seismic safety issues should acknowledge the shortcomings of PSHA, and its use as an unquestioningly relied on black box must cease.

Acknowledgements

We thank Slava Gusiakov for calling the paper by Klemeš to our attention. We thank Peter Yanev and two anonymous referees for

constructive comments. We thank David Yuen and Hou Yanping for their assistance and encouragement throughout the publication process. We thank Kei Hasegawa for assistance in drafting Fig. 1. This work was partially supported by a Grant (No. 16K05531) from the Japan Society for the Promotion of Science.

References

- Anderson, P.W., 1992. The Reverend Thomas Bayes, needles in haystacks, and the fifth force. *Phys. Today* 45 (1), 9–11.
- Bak, P., Christensen, K., Danon, L., Scanlon, T., 2002. Unified scaling law for earthquakes. *Phys. Rev. Lett.* 88, 178501, 1–4.
- Bender, B., Perkins, D.M., 1987. Seisrisk III: A Computer Program for Seismic Hazard Estimation. U.S. Geological Survey Bulletin 1772. <<https://pubs.er.usgs.gov/publication/b1772>>.
- Ben-Menahem, A., 1995. A concise history of mainstream seismology: origins, legacy, and perspectives. *Bull. Seismol. Soc. Am.* 85, 1202–1225.
- Boschi, E., Guidoboni, E., Ferrari, G., Mariotti, D., Valensise, G., Gasperini, P., 2000. Catalogue of strong Italian earthquakes from 461 B.C. to 1997. Introductory texts and CD-ROM, Version 3 of the Catalogo dei Forti Terremoti in Italia. *Ann. Geofis.* 43, 609–868.
- Budnitz, R.J., Cornell, C.A., Morris, P.A., 2005. Comment on J.U. Klügel's “Problems in the application of the SSHAC probability method for assessing earthquake hazards at Swiss nuclear power plants.” *Eng. Geol.* 82, 76–78.
- Castanos, H., Lomnitz, C., 2002. PSHA: is it science? *Eng. Geol.* 66, 315–317.
- Cisternas, M. et al., 2005. Predecessors of the giant 1960 Chile earthquake. *Nature* 437, 404–407.
- Cornell, C.A., 1968. Engineering seismic risk analysis. *Bull. Seismol. Soc. Am.* 58, 1583–1606.
- Corral, A., 2003. Local distributions and rate fluctuations in a unified scaling law for earthquakes. *Phys. Rev. E* 68, 035102(R).
- Corral, A., 2004. Long-term clustering, scaling, and universality in the temporal occurrence of earthquakes. *Phys. Rev. Lett.* 92, 108501.
- Corral, A., 2005a. Renormalization-group transformations and correlations of seismicity. *Phys. Rev. Lett.* 95, 028501.
- Corral, A., 2005b. Time-decreasing hazard and increasing time until the next earthquake. *Phys. Rev. E* 71, 017101.
- Corral, A., 2006. Dependence of earthquake recurrence times and independence of magnitudes on seismicity history. *Tectonophysics* 424, 177–193.
- Corral, A., 2007. Statistical features of earthquake temporal occurrence. *Lect. Notes Phys.* 705, 191–221.
- Cox, D.R., Lewis, P.A.W., 1966. *The Statistical Analysis of Series of Events*. Methuen, London.
- Dalkey, N., Helmer, O., 1963. An experimental application of the Delphi method to the use of experts. *Manage. Sci.* 9, 458–467.
- Davis, P.M., Jackson, D.D., Kagan, Y.Y., 1989. The longer it has been since the last earthquake, the longer the expected time till the next? *Bull. Seismol. Soc. Am.* 79, 1439–1456.
- Foster, K.R., Vecchia, P., Repacholi, M.H., 2000. Science and the precautionary principle. *Science* 288, 979–981.
- Frankel, A., 2004. How can seismic hazard around the New Madrid seismic zone be similar to that in California? *Seismol. Res. Lett.* 75, 575–586.
- Frankel, A., 2013. Comment on “Why earthquake hazard maps often fail and what to do about it.” *Tectonophysics* 592, 200–206.
- Freedman, D.A., 1997. Some issues in the foundation of statistics. In: van Fraassen, B.C. (Ed.), *Topics in the Foundation of Statistics*. Springer, Netherlands, pp. 19–39.
- Gardner, J.K., Knopoff, L., 1974. Is the sequence of earthquakes in Southern California, with aftershocks removed, Poissonian? *Bull. Seismol. Soc. Am.* 64, 1363–1367.
- Gasperini, P., Lolli, B., Vannucci, G., 2013. Empirical calibration of local magnitude data sets versus moment-magnitude in Italy. *Bull. Seismol. Soc. Am.* 103, 2227–2246.
- Geller, R.J., 1976. Scaling relations for earthquake source parameters and magnitudes. *Bull. Seismol. Soc. Am.* 66, 1501–1523.
- Geller, R.J., 1997. Earthquake prediction: a critical review. *Geophys. J. Int.* 131, 425–450.
- Geller, R.J., 2011. Shake-up time for Japanese seismology. *Nature* 472, 407–409.
- Geller, R.J., Kanamori, H., 1977. Magnitudes of great shallow earthquakes from 1904 to 1952. *Bull. Seismol. Soc. Am.* 67, 587–598.
- Geller, R.J., Jackson, D.D., Kagan, Y.Y., Mulargia, F., 1997. Earthquakes cannot be predicted. *Science* 275, 1616–1617.
- Geller, R.J., Epstein, W., Nöggerath, J., 2013. Fukushima – two years later. *Seismol. Res. Lett.* 84, 1–3.
- Geller, R.J., Mulargia, F., Stark, P.B., 2015. Why we need a new paradigm of earthquake occurrence. In: Morra, G. et al. (Eds.), *Subduction Dynamics: From Mantle Flow to Mega Disasters*, Geophysical Monograph 211. American Geophysical Union, Washington, DC, USA, pp. 183–191.
- Gilbert, G.K., 1884. A theory of the earthquakes of the Great Basin, with a practical application. *Am. J. Sci. Ser. 3* (27), 49–54.
- Gupta, R.S., 1989. *Hydrology and Hydraulic Systems*. Prentice Hall, Upper Saddle River, NJ.
- Gutenberg, B., Richter, C.F., 1956. Magnitude and energy of earthquakes. *Ann. Geofis.* 9, 1–15 (republished in *Annals of Geophysics*, 53, 7–12, 2010).

- Hanks, T.C., Cornell, C.A., 1994. Probabilistic seismic hazard analysis: a beginner's guide. In: Fifth Symposium on Current Issues Related to Nuclear Power Plant Structures, Equipment, and Piping, Dec. 14–16, 1994, Lake Buena Vista, Florida, USA.
- Hanks, T.C., Abrahamson, N.A., Boore, D.M., Coppersmith, K.J., Knepprath, N.E., 2009. Implementation of the SSHAC Guidelines for Level 3 and 4 PSHAs – Experience Gained from Actual Applications. U.S. Geological Survey Open-File Report 2009-1093. <<https://pubs.er.usgs.gov/publication/ofr20091093>>.
- Holzer, T.L., 2005. Comment on "Comparison between probabilistic seismic hazard analysis and flood frequency analysis" by Z. Wang. *Eos Trans. Am. Geophys. Union* 86, 303.
- Iervolino, I., 2013. Probabilities and fallacies: why hazard maps cannot be validated by individual earthquakes. *Earthq. Spectra* 29, 1125–1136.
- Jackson, D.D., Kagan, Y.Y., 1993. Reply to Nishenko and Sykes. *J. Geophys. Res.* 98, 9917–9920.
- Jordan, T., 2006. Earthquake predictability, brick by brick. *Seismol. Res. Lett.* 77, 3–6.
- Kagan, Y.Y., 1997. Are earthquakes predictable? *Geophys. J. Int.* 131, 505–525.
- Kagan, Y.Y., 2002. Seismic moment distribution revisited: I. Statistical results. *Geophys. J. Int.* 148, 520–541.
- Kagan, Y.Y., 2014. Earthquakes: Models, Statistics, Testable Forecasts. American Geophysical Union, Washington, DC.
- Kagan, Y.Y., Jackson, D.D., 1991. Long-term earthquake clustering. *Geophys. J. Int.* 104, 117–133.
- Kagan, Y.Y., Jackson, D.D., 1995. New seismic gap hypothesis: five years after. *J. Geophys. Res.* 100, 3943–3960.
- Kagan, Y.Y., Jackson, D.D., 2000. Probabilistic forecasting of earthquakes. *Geophys. J. Int.* 143, 438–453.
- Kagan, Y.Y., Jackson, D.D., 2013. Tohoku earthquake: a surprise? *Bull. Seismol. Soc. Am.* 103, 1181–1194.
- Kagan, Y.Y., Jackson, D.D., Geller, R.J., 2012. Characteristic earthquake model, 1884–2011, R.I.P. *Seismol. Res. Lett.* 83, 951–953.
- Kamer, Y., Hiemer, S., 2015. Data-driven spatial b-value estimation with applications to California seismicity: to b or not to b. *J. Geophys. Res.* 120, 5191–5214.
- Kammerer, A.M., Ake, J.P., 2012. Practical Implementation Guidelines for SSHAC Level 3 and 4 Hazard Studies, Office of Nuclear Regulatory Research. U.S. Nuclear Regulatory Commission, NUREG-2117, Rev. 1.
- Kanamori, H., 1977. The energy release in great earthquakes. *J. Geophys. Res.* 82, 2981–2987.
- Kanamori, H., Anderson, D.L., 1975. Theoretical basis of some empirical relations in seismology. *Bull. Seismol. Soc. Am.* 65, 1073–1095.
- Kanamori, H., Heaton, T.H., 2000. Microscopic and macroscopic physics of earthquakes. In: Rundle, J.B., Turcotte, D.L., Klein, W. (Eds.), *Geocomplexity and the Physics of Earthquakes*, Geophysical Monograph, vol. 120. American Geophysical Union, Washington, D.C., pp. 147–164.
- Klemeš, V., 1989. The improbable probabilities of extreme floods and droughts. In: Starosolszky, O., Medler, O.M. (Eds.), *Hydrology of Disasters*. James and James, London, pp. 43–51.
- Klügel, J.U., 2005a. Problems in the application of the SSHAC probability method for assessing earthquake hazards at Swiss nuclear power plants. *Eng. Geol.* 78, 285–307.
- Klügel, J.U., 2005b. Reply to Musson et al. comment. *Eng. Geol.* 82, 56–65.
- Klügel, J.U., 2005c. Reply to Krinitsky comment. *Eng. Geol.* 82, 69–70.
- Klügel, J.U., 2005d. Reply to Lomnitz comment. *Eng. Geol.* 82, 74–75.
- Klügel, J.U., 2005e. Reply to Budnitz et al. comment. *Eng. Geol.* 82, 79–85.
- Klügel, J.U., 2005f. Reply to Wang comment. *Eng. Geol.* 82, 89–90.
- Klügel, J.U., 2007. Error inflation in probabilistic seismic hazard analysis. *Eng. Geol.* 90, 186–192.
- Klügel, J.U., 2008. Seismic hazard analysis: Quo vadis? *Earth-Sci. Rev.* 88, 1–32.
- Krinitsky, E.L., 2005. Comment on J.U. Klügel's "Problems in the application of the SSHAC probability method for assessing earthquake hazards at Swiss nuclear power plants". *Eng. Geol.* 82, 66–68.
- Liu, M., Stein, S., Wang, H., 2011. 2000 years of migrating earthquakes in North China: how earthquakes in mid-continent differ from those at plate boundaries. *Lithosphere* 3(3) (2), 128. <http://dx.doi.org/10.1130/L129.1>.
- Lolli, B., Gasperini, P., Vannucci, G., 2014. Empirical conversion between teleseismic magnitudes (mb and Ms) and moment magnitude (Mw) at the Global, Euro-Mediterranean and Italian scale. *Geophys. J. Int.* 199, 805–828.
- Lombardi, A.M., Marzocchi, W., 2010. Exploring the performances and usability of the CSEP suite of tests. *Bull. Seismol. Soc. Am.* 100, 2293–2300.
- Lomnitz, C., 2005. Comment on J.U. Klügel's "Problems in the application of the SSHAC probability method for assessing earthquake hazards at Swiss nuclear power plants". *Eng. Geol.* 82, 71–73.
- Luen, B., 2010. Earthquake Prediction: Simple Methods for Complex Phenomena. Ph.D. Dissertation. University of California, Berkeley.
- Luen, B., Stark, P.B., 2012. Poisson tests of declustered catalogues. *Geophys. J. Int.* 189, 691–700.
- Main, I. (moderator), 1999. Is the reliable prediction of individual earthquakes a realistic scientific goal? *Nature Debates*. <<http://www.nature.com/nature/debates/earthquake/>>.
- Mak, S., Clements, R.A., Schorlemmer, D., 2014. The statistical power of testing probabilistic seismic-hazard assessments. *Seismol. Res. Lett.* 85, 781–783.
- McGuire, R.K., 1976. Fortran Computer Program for Seismic Risk Analysis. U.S. Geological Survey Open-File Report 76-67. <<https://pubs.er.usgs.gov/publication/ofr7667>>.
- Meehan, R.L., 1984. *The Atom and the Fault*. MIT Press, Boston.
- Muir-Wood, R., 1993. From global seismotectonics to global seismic hazard. *Ann. Geophys.* 36, 153–168.
- Mulargia, F., 2013. Why the next large earthquake is likely to be a big surprise. *Bull. Seismol. Soc. Am.* 103, 2946–2952.
- Mulargia, F., Bizzarri, A., 2014. Anthropogenic triggering of large earthquakes. *Sci. Rep.* 4, 6100. <http://dx.doi.org/10.1038/srep06100>.
- Mulargia, F., Bizzarri, A., 2015. Fluid pressure waves trigger earthquakes. *Geophys. J. Int.* 200, 1279–1283.
- Mulargia, F., Geller, R.J. (Eds.), 2003. *Earthquake Science and Seismic Risk Reduction*. Kluwer, Dordrecht. 347 pp.
- Mulargia, F., Castellaro, S., Ciccotti, M., 2004. Earthquakes as three stage processes. *Geophys. J. Int.* 158, 98–108.
- Musson, R.M.W., 2004. Comment on "Communicating with uncertainty, a critical issue with probabilistic seismic hazard analysis" by Z. Wang. *Eos Trans. Am. Geophys. Union* 85, 235–236.
- Musson, R.M.W., 2012. PSHA validated by quasi observational means. *Seismol. Res. Lett.* 83, 130–134.
- Musson, R.M.W., Toro, G.R., Coppersmith, K.J., Bommer, J.J., Deichmann, N., Bungum, H., Cotton, F., Scherbaum, F., Slejko, D., Abrahamson, N.A., 2005. Evaluating hazard results for Switzerland and how not to do it: a discussion of "Problems in the application of the SSHAC probability method for assessing earthquake hazards at Swiss nuclear power plants" by J.U. Klügel. *Eng. Geol.* 82, 43–55.
- NEHRP – National Earthquake Hazards Reduction Program, 1997. NEHRP Recommended Provisions for Seismic Regulations for New Buildings and Other Structures, Part 1: Provisions. Building Seismic Safety Council, Washington, D.C.
- Nishenko, S.P., Sykes, L.R., 1993. Comment on "Seismic gap hypothesis: ten years after" by Y.Y. Kagan and D.D. Jackson. *J. Geophys. Res.* 98, 9909–9916.
- Ogata, Y., 1988. Statistical models for earthquake occurrences and residual analysis for point processes. *J. Am. Stat. Assoc.* 83, 9–27.
- Ogata, Y., 1999. Seismicity analysis through point-process modeling: a review. *Pure Appl. Geophys.* 155, 471–507.
- Omori, F., 1895. On the aftershocks of earthquakes. *J. College Sci. Imperial Univ. Tokyo* 7, 111–200.
- Parsons, T., Toda, S., Stein, R.S., Barka, A., Dieterich, J.H., 2000. Heightened odds of large earthquakes near Istanbul: an interaction-based probability calculation. *Science* 288, 661–665.
- Parsons, T., Console, R., Falcone, G., Murru, M., Yamashina, K., 2012. Comparison of characteristic and Gutenberg-Richter models for time-dependent $M \geq 7.9$ earthquake probability in the Nankai-Tokai subduction zone, Japan. *Geophys. J. Int.* 190, 1673–1688.
- Reasenber, P.A., 1985. Second-order moment of Central California seismicity, 1969–1982. *J. Geophys. Res.* 90, 5479–5495.
- Reid, H.F., 1911. The elastic rebound theory of earthquakes. *Univ. California Publ. Bull. Dept. Geol. Sci.* 6, 413–444.
- Reiter, L., 2004. When are ground motion estimates too high? *Seismol. Res. Lett.* 74, 282.
- Rong, Y., Jackson, D.D., Kagan, Y.Y., 2003. Seismic gaps and earthquakes. *J. Geophys. Res.* 108, 2471. <http://dx.doi.org/10.1029/2002JB002334>.
- Rundle, J.B., Klein, W., Gross, S., Turcotte, D.L., 1995. Boltzmann fluctuations in numerical simulations of nonequilibrium threshold systems. *Phys. Rev. Lett.* 75, 1658–1661.
- Savage, J.C., 1993. The Parkfield prediction fallacy. *Bull. Seismol. Soc. Am.* 83, 1–6.
- Scholz, C.H., 2002. *The Mechanics of Earthquakes and Faulting*. Cambridge University Press, New York.
- Schorlemmer, D., Gerstenberger, M., Wiemer, S., Jackson, D.D., Rhoades, D.A., 2007. Earthquake likelihood model testing. *Seismol. Res. Lett.* 78, 17–29.
- Schwartz, D.P., Coppersmith, K.J., 1984. Fault behavior and characteristic earthquakes—examples from the Wasatch and San Andreas fault zones. *J. Geophys. Res.* 89, 5681–5698.
- Sieh, K., Stuiver, M., Brillinger, D., 1989. A more precise chronology of earthquakes produced by the San Andreas Fault in Southern California. *J. Geophys. Res.* 94, 603–623.
- Solomos, G., Pinto, A., Dimova, S., 2008. A Review of the Seismic Hazard Zonation in National Building Codes in the Context of EUROCODE 8. European Commission Joint Research Centre, Report EUR 23563 EN.
- Stark, P.B., 2013. Ontology of Earthquake Probability: Metaphor, presentation at SAMSI Symposium on Dynamics of Seismicity, Earthquake Clustering and Patterns, Research Triangle Park, North Carolina, USA, October 9–11, 2013. <<https://www.samsi.info/programs-and-activities/other-workshops-and-post-doc-seminars/2013-dynamics-of-seismicity-earthquake-clustering-and-patterns-in-fault-networks-october-9-11-2013/>>.
- Stark, P.B., 2017. Pay no attention to the model behind the curtain. In: Saltelli, A., Guimarães Pereira, A. (Eds.), *Significant Digits: Responsible use of Quantitative Information* (in press).
- Stark, P.B., Freedman, D., 2003. What is the chance of an earthquake? In: Mulargia, F., Geller, R.J. (Eds.), *Earthquake Science and Seismic Risk Reduction*, NATO Science Series IV: Earth and Environmental Sciences, vol. 32. Kluwer, Dordrecht, The Netherlands, pp. 201–213.
- Stein, S., Geller, R.J., 2012. Communicating uncertainties in natural hazard forecasts. *Eos Trans. Am. Geophys. Union* 93, 361–362.
- Stein, S., Tomasello, J., Newman, A., 2003. Should Memphis build for California's earthquakes? *Eos Trans. Am. Geophys. Union* 84, 177–185.
- Stein, S., Geller, R.J., Liu, M., 2012. Why earthquake hazard maps often fail and what to do about it. *Tectonophysics* 562–563, 1–25.

- Stein, S., Geller, R.J., Liu, M., 2013. Reply to comment by Arthur Frankel on “Why earthquake hazard maps often fail and what to do about it.” *Tectonophysics* 592, 207–209.
- Stirling, M.W., 2014. The continued utility of probabilistic seismic-hazard assessment. In: Wyss, M. (Ed.), *Earthquake Hazard, Risk and Disasters*. Academic Press, pp. 359–376.
- Taleb, N., 2007. *The Black Swan*. Random House.
- Turcotte, D.L., 1997. *Fractals and Chaos in Geology and Geophysics*. Cambridge University Press, New York.
- Turcotte, D.L., Scherbakov, R., Rundle, J.B., 2009. Complexity and earthquakes. In: Kanamori, H. (Ed.), *Earthquake Seismology, Treatise on Geophysics*, vol. 4. Elsevier, Amsterdam, pp. 675–700.
- USGS-WGCEP, 1999. *Earthquake Probabilities in the San Francisco Bay Region: 2000 to 2030. A Summary of Findings By Working Group on California Earthquake Probabilities Open-File Report 99-517*.
- Utsu, T., Ogata, Y., Matsu'ura, R.S., 1995. The centenary of the Omori formula for a decay law of aftershock activity. *J. Phys. Earth* 43, 1–33.
- Wang, Z., 2005a. Comparison between probabilistic seismic hazard analysis and flood frequency analysis. *Eos Trans. Am. Geophys. Union* 86, 45, 51–5.
- Wang, Z., 2005b. Comment on JU. Klügel's “Problems in the application of the SSHAC probability method for assessing earthquake hazards at Swiss nuclear power plants.” *Eng. Geol.* 78, 86–88.
- Wang, Z., 2008. A technical note on seismic microzonation in the central United States. *J. Earth Syst. Sci.* 117, 749–756.
- Wang, Z., 2012. Comment on “PSHA Validated by Quasi Observational Means” by R. M.W. Musson. *Seismol. Res. Lett.* 83, 714–716.
- Wang, Z., Woolery, E., Shi, B., Kiefer, J., 2003. Communicating with uncertainty, a critical issue with probabilistic seismic hazard analysis. *Eos Trans. Am. Geophys. Union* 84, 501–508.
- Weldon, R., Scharer, K., Fumal, T., Biasi, G., 2004. Wrightwood and the earthquake cycle: what a long recurrence record tells us about how faults work. *GSA Today* 14, 4–10.