

FULL PAPER

Open Access



Earthquake interevent time distribution in Kachchh, Northwestern India

Sumanta Pasari* and Onkar Dikshit

Abstract

Statistical properties of earthquake interevent times have long been the topic of interest to seismologists and earthquake professionals, mainly for hazard-related concerns. In this paper, we present a comprehensive study on the temporal statistics of earthquake interoccurrence times of the seismically active Kachchh peninsula (western India) from thirteen probability distributions. Those distributions are exponential, gamma, lognormal, Weibull, Levy, Maxwell, Pareto, Rayleigh, inverse Gaussian (Brownian passage time), inverse Weibull (Frechet), exponentiated exponential, exponentiated Rayleigh (Burr type X), and exponentiated Weibull distributions. Statistical inferences of the scale and shape parameters of these distributions are discussed from the maximum likelihood estimations and the Fisher information matrices. The latter are used as a surrogate tool to appraise the parametric uncertainty in the estimation process. The results were found on the basis of two goodness-of-fit tests: the maximum likelihood criterion with its modification to Akaike information criterion (AIC) and the Kolmogorov-Smirnov (K-S) minimum distance criterion. These results reveal that (i) the exponential model provides the best fit, (ii) the gamma, lognormal, Weibull, inverse Gaussian, exponentiated exponential, exponentiated Rayleigh, and exponentiated Weibull models provide an intermediate fit, and (iii) the rest, namely Levy, Maxwell, Pareto, Rayleigh, and inverse Weibull, fit poorly to the earthquake catalog of Kachchh and its adjacent regions. This study also analyzes the present-day seismicity in terms of the estimated recurrence interval and conditional probability curves (hazard curves). The estimated cumulative probability and the conditional probability of a magnitude 5.0 or higher event reach 0.8–0.9 by 2027–2036 and 2034–2043, respectively. These values have significant implications in a variety of practical applications including earthquake insurance, seismic zonation, location identification of lifeline structures, and revision of building codes.

Keywords: Kachchh peninsula; Earthquake recurrence; Probability models; Conditional probability

Background

The Kachchh (Kutch) province of Gujarat, northwestern India, is a unique stable continental region (SCR) of the world that has experienced two large intraplate earthquakes within a span of 182 years (M_w 7.8 in 1819 and M_w 7.7 in 2001). In addition to these two disastrous events, it has a long history of infrequent but moderate earthquake occurrences, indicating slow but continuous stress release in these regions (Chandra 1977; Gupta et al. 2001; Rajendran and Rajendran 2001; Rastogi 2001, 2004; Jade et al. 2002). The Kachchh peninsula, which mostly falls in Zone V (highest seismicity and potential for magnitude 8.0 earthquakes) on the seismic zoning

map of India (BIS: Bureau of Indian Standards 2002), is an assemblage of several active faults. Those faults are the Allah Bund Fault (ABF), Island Belt Fault (IBF), Kachchh Mainland Fault (KMF), Bhuj Fault (BF), Katrol Fault (KF), Wagad Fault (WF), Nagar Parkar Fault (NPF), and the Kathiawar Fault (KTF) (Biswas 1987, 2005; Rajendran and Rajendran 2001; Rastogi 2001; Morino et al. 2008). The genesis of earthquakes from these faults, their source characterizations, stress-field analyses, and most importantly the vulnerability assessment of such seismically active zones bear significant importance for the safety of human lives and critical infrastructure such as power plants, refineries, schools, hospitals, shopping malls, and other lifeline structures (e.g., archeological monuments) situated in the nearby

* Correspondence: spasari@iitk.ac.in; sumanta.pasari@gmail.com
Department of Civil Engineering, Indian Institute of Technology, Kanpur
208016, India

major cities and industrial hubs in this heartland of northwestern India.

Earthquake studies based on physical and geological models are useful; however, the current paper focuses on empirical earthquake recurrence modeling. This modeling has become an integral part within many seismological societies and private organizations for nationwide hazard assessment and catastrophe insurance programs (Lee et al. 2011; Working Group on California Earthquake Probabilities 2013). Methods such as paleoseismic investigations of mapped faults or geodetic monitoring to determine strain accumulation patterns are helpful to forecast earthquakes. However, the primary issue with these approaches is that many disastrous earthquakes do not reappear on previously identified faults (Lee et al. 2011). To address this limitation, statistical seismology of earthquake occurrence and forecasting has become essential for seismic hazard assessment of large geographical areas (Jordan 2006; Shebalin et al. 2014).

In this paper, we analyze earthquake interevent times of magnitude 5.0 or higher events in the seismic-prone Kachchh region from thirteen probability distributions. Those distributions are exponential, gamma, lognormal, Weibull, Levy, Maxwell, Pareto, Rayleigh, inverse Gaussian (Brownian passage time), inverse Weibull (Frechet), exponentiated exponential, exponentiated Rayleigh (Burr type

X), and exponentiated Weibull models. We seek the best probability model(s) for earthquake forecasting, and subsequently generate a number of conditional probability curves (hazard curves) to assess the present-day seismicity in the Kachchh region.

Geology and seismotectonic framework of Kachchh

The Kachchh peninsula largely consists of Quaternary/Cenozoic sediments, Deccan volcanic rocks, and Jurassic sandstones (Biswas 1987; Gupta et al. 2001). Geomorphologically, the entire region can be broadly categorized into the following two major zones: the Rann area that often gets submerged by seawater, and the highland zone that comprises uplifts, elevated landforms, and residual depressions (Biswas 1987, 2005; Yadav et al. 2008). The Rann area is essentially an uninhabitable desert of mud flats and salt pans, whereas the highland zone is semi arid containing fluviomarine sediments and Banni plain grasslands (Biswas 1987).

Major structural features of the Kachchh region include several EW trending faults, folds, and a rift basin, which is bounded by the following two extensional faults: the south-dipping Nagar Parkar Fault (NPF) in the north and the north-dipping Kathiawar Fault (KTF) in the south (Fig. 1). Other major faults in this region are the north-dipping Allah Bund Fault (ABF), south-dipping Kachchh

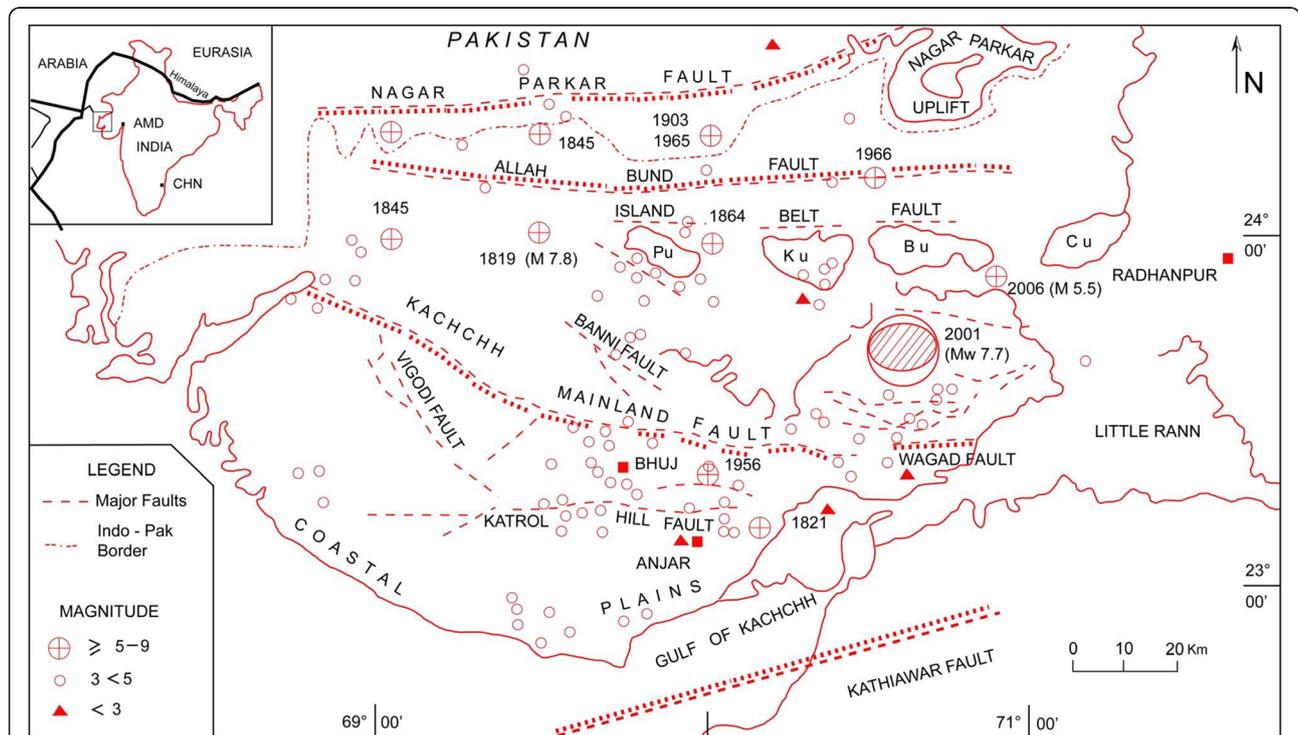


Fig. 1 Seismotectonic map of Kachchh region. This figure shows a detailed seismotectonic map of the Kachchh region comprising several faults, lineaments, uplifts, and other structural features; The Kachchh Mainland Uplift is in the south of Kachchh Mainland Fault, whereas the Wagad Uplift is in the north of Wagad Fault; PU, Pachham Uplift; KU, Khadir Uplift; BU, Bela Uplift; CU, Chobar Uplift. (Inset) Indian plate boundaries are highlighted to realize the geographical location of Kachchh peninsula. AMD, Ahmedabad; CHN, Chennai (modified after Rastogi 2001)

Mainland Fault (KMF), Katrol Fault (KF), Bhuj Fault (BF), and the south-dipping North Wagad Fault (NWF). These fault zones are the source of many moderate to large intraplate earthquakes in this region. For instance, the most disastrous M_w 7.7 (intensity X+ on the modified Mercalli intensity (MMI) scale) Republic Day Bhuj earthquake along the NWF (08:46 IST, January 26, 2001; epicentral location: 23.412° N, 70.232° E; focal depth: 23 km) caused 13,819 human deaths, US \$10 billion economic loss, and damaged over one million houses (Rastogi 2001). Another large event, the 1819 Rann of Kachchh earthquake (M_w 7.8), occurred along the ABF and created fault scarp approximately 4 to 6 m high in this region (Rajendran and Rajendran 2001; Morino et al. 2008). The Anjar 1956 (M_w 6.0) event occurred along the KF (Chung and Gao 1995). In addition to these major faults, a number of uplifts (Fig. 1), namely Kachchh Mainland Uplift, Kathiawar Uplift, Pachham, Khadir, Bela, Wagad, and Chobar uplift, and some minor NE/NW trending faults/lineaments characterize the tectonic setting of Kachchh region (Biswas 1987; Rastogi 2001).

Rationale and previous works

In this paper, we consider earthquake occurrence phenomena to be a statistical process. We assume that the earthquake interevent times are random variables associated with some probability distributions. Various properties of these underlying distributions provide important insights into earthquake recurrence intervals, elapsed time (time elapsed since last earthquake), residual time (time remaining to an earthquake), and related concepts, which may finally be integrated in a systematic manner to arrive at long-term earthquake forecasting in a specified zone of interest (SSHAC: Senior Seismic Hazard Analysis Committee 1997). Extensive studies of earthquake hazards and their statistical analyses are being undertaken globally to quantify anomalous behavior of earthquake risks in a seismically active region. These studies also facilitate the examination of hypothetical earthquake scenarios that aid in decisions for how much money a government should allocate for disaster utility or an insurance company should collect from an individual as a seismic-risk coverage premium (SSHAC: Senior Seismic Hazard Analysis Committee 1997; Lee et al. 2011; Working Group on California Earthquake Probabilities 2013).

Probabilistic earthquake modeling has crossed several milestones over the years. During the development stage of seismic renewal process theory, the exponential distribution (due to the Poissonian assumption of the number of earthquake events) used to be the favored distribution in representing a sequence of earthquake inter-arrival times (Cornell 1968; Hagiwara 1974; Baker 2008). Later, however, many researchers (e.g., Reid 1910; Anagnos

and Kiremidjian 1988; Matthews et al. 2002; Baker 2008) have pointed out that the time-independent Poissonian model has some disagreement with the physics of earthquake-generating mechanisms. Therefore, a number of time-dependent renewal models have subsequently evolved (Utsu 1984; Parvez and Ram 1997; Matthews et al. 2002; Tripathi 2006; Yadav et al. 2008, 2010; Working Group on California Earthquake Probabilities 2013; Pasari and Dikshit 2014a, b). A pioneering work by the Japanese researcher Utsu (1984) was notable in its contributions. Utsu (1984) used four renewal probability distributions (Weibull, gamma, lognormal, and exponential) to estimate earthquake recurrence intervals in Japan and its surrounding areas. Later, several researchers (e.g., Parvez and Ram 1997; Tripathi 2006; Yadav et al. 2008, 2010; Yazdani and Kowsari 2011; Chen et al. 2013; Pasari and Dikshit 2014a; Pasari 2015) applied similar methodologies for their respective geographic regions of study. The suitability of other probability models such as the Gaussian distribution (Papazachos et al. 1987), the negative binomial distribution (Dionysiou and Papadopoulos 1992), the Levy distribution (Sotolongo-Costa et al. 2000; Pasari and Dikshit 2014b), the Pareto group of distributions (Kagan and Schoenberg 2001; Ferraes 2003), the generalized gamma distribution (Bak et al. 2002), the Brownian passage time distribution (Matthews et al. 2002; Pasari and Dikshit 2014b), the Rayleigh distribution (Ferraes 2003; Yazdani and Kowsari 2011), the inverse Weibull distribution (Pasari and Dikshit 2014a), and the exponentiated exponential distribution (Pasari and Dikshit 2014c; Pasari 2015) has also been explored for identification of the most suitable probability model for a given earthquake catalog.

In this paper, we re-examine earthquake interevent studies conducted by Tripathi (2006) and Yadav et al. (2008) based on a versatile set of thirteen probability models to provide a fresh perspective on the most suitable probability distribution in this unique intraplate seismogenic zone of northwestern India. The results are discussed in conjunction with the probabilistic assessment of earthquake hazards within this region.

Earthquake data

We used a complete and homogeneous earthquake catalog (Yadav et al. 2008 in *Pure and Applied Geophysics* 165, 1813–1833) comprising 15 intra-continent earthquake events ($M \geq 5.0$) that occurred in the Kachchh region (23–25° N, 68–71° E) during 1819 to 2006. Details of these events are listed in Table 1, and their epicentral distributions are shown in Fig. 2. This catalog is the most recent updated catalog, as no $M \geq 5.0$ earthquakes have occurred in the Kachchh region since 2006.

Table 1 List of $M \geq 5.0$ earthquakes in Kachchh and its adjoining regions (Yadav et al. 2008)

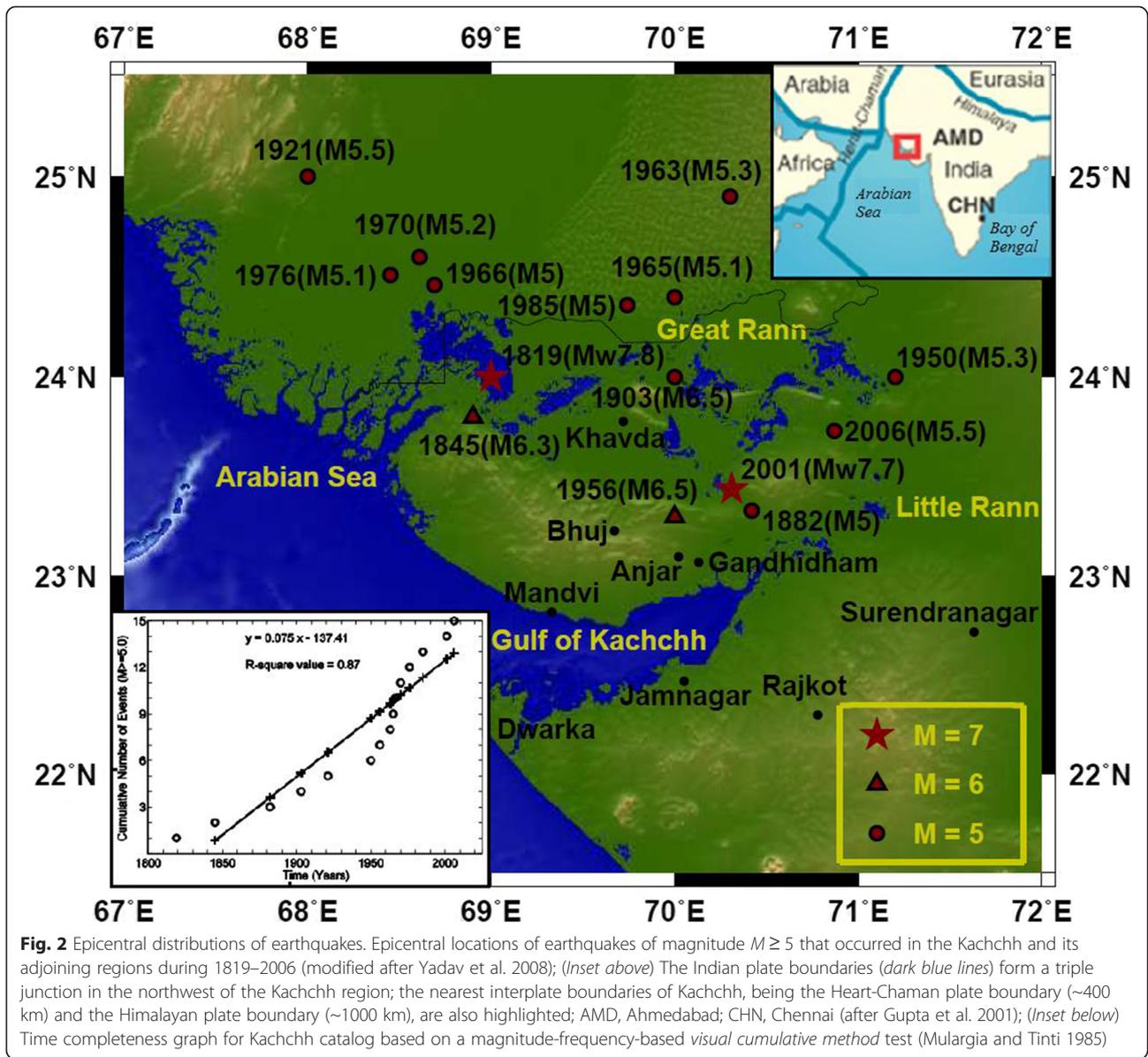
S.N.	Date			Location		Magnitude	Location
	Year	Month	Day	Latitude (N)	Longitude (E)		
1	1819	06	16	24.00	69.00	7.8	Kachchh
2	1845	06	19	23.80	68.90	6.3	Lakhpat
3	1882	12	15	23.33	70.42	5.0	Bhachau
4	1903	01	14	24.00	70.00	5.6	Kachchh
5	1921	10	26	25.00	68.00	5.5	Indus-Kachchh
6	1950	06	14	24.00	71.20	5.3	Kachchh
7	1956	07	21	23.30	70.00	6.5	Kachchh
8	1963	07	13	24.90	70.30	5.3	Thar, Pakistan
9	1965	03	26	24.40	70.00	5.1	Kachchh
10	1966	05	27	24.46	68.69	5.0	Thar, Pakistan
11	1970	02	13	24.60	68.61	5.2	Kachchh
12	1976	06	04	24.51	68.45	5.1	Allah Bund
13	1985	04	07	24.36	69.74	5.0	Kachchh
14	2001	01	26	23.44	70.31	7.7	Kachchh
15	2006	03	07	23.73	70.87	5.5	Gedi, Rapar

It should be noted that the present earthquake catalog consists of modern (instrumental) as well as historical (non-instrumental) events of Kachchh region from various sources. Those sources include the Indian Meteorological Department (IMD), United States Geological Survey (USGS), Harvard Centroid Moment Tensor (CMT) catalog, and some published literature (Yadav et al. 2008). The difficulty in converting the scale of earthquake data at different stages with different magnitudes or from different sources has been a consistent challenge in homogenizing this catalog (Biswas 1987; Rundle et al. 2003; Yadav et al. 2008). However, to examine the time completeness of the Kachchh catalog, we performed a magnitude-frequency-based *visual cumulative method* test (Mulargia and Tinti 1985). Under this test, a graph was constructed with time (in years) and the cumulative number of earthquake events. The equation of the best fit line (in a least-squares sense) was determined. A catalog is considered to be complete with respect to time if the trend of the data stabilizes to approximately a straight line (Mulargia and Tinti 1985). This cumulative straight-line approach is based on the fact that earthquake rates and moment releases are ultimately steady over sufficiently long time periods (Mulargia and Tinti 1985). However, one should keep in mind that this assumption may not be correct. The possibility of extended aftershock durations in low-strain-rate intra-continental regions (e.g., Stein and Liu 2009) and the possibility of substantial variations in seismic activity (e.g., Page and Felzer 2015) may raise questions about this assumption.

For the Kachchh catalog, it is observed that the graph between time and cumulative number of events has a

linear relationship with an R-square value greater than 0.85 (Inset of Fig. 2). Thus, the studied homogeneous catalog may be considered to be complete with respect to time.

The intensities, casualties, and fault characteristics of some of these earthquakes have been studied in detail (Chung and Gao 1995; Rajendran and Rajendran 2001; Rastogi 2001, 2004; Negishi et al. 2002; Mandal et al. 2005; Mandal and Horton 2007; Morino et al. 2008; Kayal et al. 2012). The 16 June 1819 great Rann of Kachchh earthquake (M_w 7.8; intensity XI on MMI scale) that occurred near the northwestern international border (~100 km northwest of Bhuj town) rumbled the entire region, causing 1500 deaths in Kachchh and 500 in Ahmedabad (~250 km from epicenter). This low-angle reverse faulting earthquake formed a fault scarp of 4–6 m in height and 90 km in length trending EW along the ABF (Rajendran and Rajendran 2001; Rastogi 2001). The isoseismals of this earthquake were nearly elliptical with a principal axis oriented in the ENE direction (Rajendran and Rajendran 2001; Yadav et al. 2008). The next damaging earthquake (M_w 6.3; intensity VII on MMI scale) occurred in the Lakhpat region on 19 June 1845. Another well-studied (Chung and Gao 1995) destructive earthquake is the 1956 Anjar earthquake (M_w 6.0; intensity IX on MMI scale) that occurred along the KF near Anjar and destroyed much of the infrastructure and buildings, claiming 115 deaths (Chung and Gao 1995; Gupta et al. 2001; Rastogi 2001). The most recent catastrophic earthquake (M_w 7.7, intensity X+ on the MMI scale) in the western part of the



peninsular Indian shield occurred on 26 January 2001, followed by more than 2000 aftershocks, including a few with magnitudes above 5.0 (Mandal et al. 2005; Yadav et al. 2008; Kayal et al. 2012). This event caused extensive damage in the neighboring areas, leaving nearly 14,000 people dead, 167,000 injured, and one million people homeless (Miyashita et al. 2001; Rastogi 2001). Apart from structural damages, intense liquefaction and fluidization occurred in an area of 300 km × 200 km covering Rann, Banni plain, and several other saline-marshy lowland regions (Rastogi 2001; Kayal et al. 2012). The fault plane solutions of the main shock of the 2001 event suggested a reverse faulting (with some strike-slip component) mechanism, which is analogous to the 1956 Anjar earthquake or the 1819

Rann of Kachchh earthquake (Chung and Gao 1995; Rastogi 2001; Kayal et al. 2002; Negishi et al. 2002).

In view of the above threats from earthquake damage, several national and international initiatives have been performed for earthquake hazard analysis and catastrophic insurance programs in the seismically active Kachchh region (Kayal et al. 2012; Choudhury et al. 2014). The present study contributes to this body of work by estimating earthquake recurrence intervals and associated hazards in a probabilistic environment.

Methods and results

Our methodology in this paper broadly consists of the following three steps: model description, parameter estimation, and model validation. We briefly discuss each

method and present the results therein for a better visualization of the concepts.

Model description

Thirteen different probability models are considered in the present analysis. The probability density functions of these distributions are shown in Table 2. The respective model parameters, their domains, and specific roles are also listed (Johnson et al. 1995; Murthy et al. 2004). With the known density function $f(t)$ of a positive random variable, T , it is straightforward to obtain the cumulative distribution function $F(t)$, survival function $S(t)$, hazard function $h(t)$, and reverse

hazard function $r(t)$ as $F(t) = \int_0^t f(u) du$, $S(t) = 1 - F(t)$,

$$h(t) = \frac{f(t)}{S(t)}, \text{ and } r(t) = \frac{f(t)}{F(t)}.$$

From Table 2, we observe that domains of all distributions except the Pareto distribution are the entire positive real line. These distributions also offer various representations depending on the shape parameter. In particular, the shapes of the hazard rate function and reversed hazard rate function have significant importance in understanding whether the residual time (time remaining to a future event) is increasing or decreasing for an increasing elapsed time (Sornette and Knopoff 1997; Matthews et al. 2002). On the other hand, the heavy-tailedness property of a model (e.g., lognormal, Levy, Pareto) allows diversification in seismic risk analysis and associated applications (SSHAC: Senior Seismic Hazard Analysis Committee 1997).

In order to calculate the conditional probability of an earthquake for a known elapsed time τ , we introduce a random variable V , corresponding to a waiting time ν . The conditional probability of an earthquake in the time interval $(\tau, \tau + \nu)$, knowing that no earthquake occurred during previous τ years, is then calculated as

$$P(V \leq \tau + \nu | V \geq \tau) = \frac{F(\tau + \nu) - F(\tau)}{1 - F(\tau)} \quad (\nu > 0) \quad (1)$$

Parameter estimation

The maximum likelihood estimation (MLE) method was adopted for parameter estimation not only because of its flexibility and wide applicability, but also for its ability to provide the uncertainty (asymptotic) measure in the estimation. The MLE method yields consistent estimators that are often desirable in any statistical analysis (Johnson et al. 1995). In brief, the MLE method estimates parameter values by maximizing the likelihood (log) function on the basis of observed sample values. Thus, it is more realistic for practical applications.

In recent years, significant research has focused on quantifying the uncertainties and intricacies of the estimation process. Nevertheless, precise uncertainty information is rarely available because exact distributions of the estimated model parameters are mostly unknown (Johnson et al. 1995). Therefore, a method based on the Fisher information matrices (FIM) is frequently utilized as a surrogate tool to appraise the parametric uncertainty in terms of the variances and confidence bounds of the estimated parameters (Hogg et al. 2005).

Let $I_{p \times p}(\theta)$ be the information matrix; $\theta = (\theta_1, \theta_2, \dots, \theta_p)$, for some integer p , denotes the vector of parameters. Then, $I_{p \times p}(\theta)$ can be calculated (Hogg et al. 2005) as

$$\begin{aligned} I_{p \times p}(\theta) &= (I_{ij}(\theta))_{i,j=1,2,\dots,p} \\ &= E \left(- \frac{\partial^2 \ln f(T; \theta)}{\partial \theta_i \partial \theta_j} \right)_{i,j=1,2,\dots,p} \\ &= \frac{1}{n} E \left(- \frac{\partial^2 L(T; \theta)}{\partial \theta_i \partial \theta_j} \right)_{i,j=1,2,\dots,p} \end{aligned} \quad (2)$$

Here E is the expectation operator; $L(T; \theta)$ is the log-likelihood function of n sample data points $\{t_1, t_2, t_3, \dots, t_n\}$. The FIM $I_{p \times p}(\theta) = (I_{ij}(\theta))_{i,j=1,2,\dots,p}$ can also be expressed in terms of the density function $f(t; \theta)$, the hazard function $h(t; \theta)$ (Efron and Johnstone 1990), or the reversed hazard function $r(t; \theta)$ (Gupta et al. 2004) as shown below.

$$\begin{aligned} (I_{ij}(\theta))_{i,j=1,2,\dots,p} &= E \left[\left(\frac{\partial \ln f(T; \theta)}{\partial \theta_i} \right) \left(\frac{\partial \ln f(T; \theta)}{\partial \theta_j} \right) \right]_{i,j=1,2,\dots,p} \\ &= E \left[\left(\frac{\partial \ln h(T; \theta)}{\partial \theta_i} \right) \left(\frac{\partial \ln h(T; \theta)}{\partial \theta_j} \right) \right]_{i,j=1,2,\dots,p} \\ &= E \left[\left(\frac{\partial \ln r(T; \theta)}{\partial \theta_i} \right) \left(\frac{\partial \ln r(T; \theta)}{\partial \theta_j} \right) \right]_{i,j=1,2,\dots,p} \end{aligned} \quad (3)$$

The FIM defined above is a symmetric and positive semi-definite matrix. It is important to note that although the random variables $(\partial/\partial\theta)\ln f(T; \theta)$, $(\partial/\partial\theta)\ln h(T; \theta)$ or $(\partial/\partial\theta)\ln r(T; \theta)$ produce different first moments, their second moments are identically equal to the elements of $I_{p \times p}(\theta)$. The information matrix is often combined with the Cramer-Rao lower bound theorem (Hogg et al. 2005) to asymptotically estimate the variance-covariance matrix $\sum \hat{\theta}$ of the estimated parameters $(\hat{\theta})$ as $\sum \hat{\theta} \geq [nI(\hat{\theta})]^{-1}$; $\hat{\theta}$ is the maximum likelihood estimate of θ . In addition, the $(1 - \delta)\%$ two-sided asymptotic confidence bounds on the parameters are easily obtained from the following inequality:

Table 2 Probability distributions, density functions, and associated Fisher Information Matrices (FIM)

Distribution	Density function	Domain	Role	Parameters	FIM $I(\theta)$
Exponential	$\frac{1}{a} e^{-\frac{t}{a}}$	$t > 0$	a -scale	$a > 0$	$\frac{1}{a^2}$
Gamma ^a	$\frac{1}{\Gamma(\beta)} \frac{t^{\beta-1}}{a^\beta} e^{-\frac{t}{a}}$	$t > 0$	a -scale β -shape	$a > 0$ $\beta > 0$	$\begin{bmatrix} \frac{\beta}{a^2} & \frac{1}{a} \\ \frac{1}{a} & \psi'(\beta) \end{bmatrix}$
Lognormal	$\frac{1}{t\beta\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\ln t - a}{\beta}\right)^2\right]$	$t > 0$	a -log-scale β -shape	$-\infty < a < \infty$ $\beta > 0$	$\begin{bmatrix} \frac{2}{\beta^2} & 0 \\ 0 & \frac{1}{\beta^2} \end{bmatrix}$
Weibull ^a	$\frac{\beta}{a^\beta} t^{\beta-1} e^{-\left(\frac{t}{a}\right)^\beta}$	$t > 0$	a -scale β -shape	$a > 0$ $\beta > 0$	$\begin{bmatrix} \frac{\beta^2}{a^2} & -\frac{1}{a}(1 + \psi(1)) \\ -\frac{1}{a}(1 + \psi(1)) & \frac{1}{\beta^2}(\psi'(1) + \psi^2(2)) \end{bmatrix}$
Levy	$\sqrt{\frac{a}{2\pi}} \frac{e^{-\frac{a}{t}}}{t^{3/2}}$	$t > 0$	a -scale	$a > 0$	$\frac{1}{2a^2}$
Maxwell	$\sqrt{\frac{2}{\pi}} \frac{t^2}{a^3} \exp\left[-\frac{1}{2}\left(\frac{t}{a}\right)^2\right]$	$t > 0$	a -scale	$a > 0$	$\frac{6}{a^2}$
Pareto	$\frac{\beta}{t^{\beta+1}}$	$t > a$	a -scale β -shape	$t > a > 0$ $\beta > 0$	$\begin{bmatrix} \frac{\beta}{a^2(\beta+2)} & -\frac{1}{a(\beta+1)} \\ -\frac{1}{a(\beta+1)} & \frac{1}{\beta^2} \end{bmatrix}$
Rayleigh	$\frac{t}{a^2} \exp\left(-\frac{t^2}{2a^2}\right)$	$t > 0$	a -scale	$a > 0$	$\frac{4}{a^2}$
Inverse Gaussian (Brownian Passage Time)	$\sqrt{\frac{\beta}{2\pi t^3}} \exp\left[-\frac{\beta(t-a)^2}{2a^2 t}\right]$	$t > 0$	β/a -shape	$a > 0$ $\beta > 0$	$\begin{bmatrix} \frac{1}{2\beta^2} & 0 \\ 0 & \frac{\beta}{a^3} \end{bmatrix}$

Table 2 Probability distributions, density functions, and associated Fisher Information Matrices (FIM) (Continued)

Inverse Weibull (Frechet) ^a	$\beta a^\beta t^{-\beta-1} e^{-\left(\frac{t}{a}\right)^{-\beta}}$	$t > 0$	a -scale β -shape	$a > 0$ $\beta > 0$	$\begin{bmatrix} \frac{\beta^2}{a^2} & \frac{1}{a}(1 + \psi(1)) \\ \frac{1}{a}(1 + \psi(1)) & \frac{1}{\beta^2}(\psi'(1) + \psi^2(2)) \end{bmatrix}$
Exponentiated exponential ^a	$a\beta(1 - e^{-at})^{\beta-1} e^{-at}$	$t > 0$	$1/a$ -scale β -shape	$a > 0$ $\beta > 0$	$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ $a_{11} = \frac{1}{a^2} \left[1 + \frac{\beta(\beta-1)}{\beta-2} (\psi'(1) - \psi'(\beta-1)) + (\psi(\beta-1) - \psi(1))^2 \right]$ $- \frac{\beta}{a^2} [\psi'(1) - \psi(\beta) + (\psi(\beta) - \psi(1))^2]; \beta \neq 2$ $a_{12} = a_{21} = \frac{1}{a} \left[\frac{\beta}{\beta-1} (\psi(\beta) - \psi(1)) - (\psi(\beta+1) - \psi(1)) \right]; \beta \neq 1$ $a_{22} = \frac{1}{\beta^2}$
Exponentiated Rayleigh (Burr Type X) ^b	$\frac{2\beta t}{a^2} e^{-\left(\frac{t}{a}\right)^2} \left(1 - e^{-\left(\frac{t}{a}\right)^2}\right)^{\beta-1}$	$t > 0$	a -scale β -shape	$a > 0$ $\beta > 0$	-
Exponentiated Weibull ^b	$\frac{\beta\gamma}{a} \left(\frac{t}{a}\right)^{\beta-1} e^{-\left(\frac{t}{a}\right)^\beta} \left(1 - e^{-\left(\frac{t}{a}\right)^\beta}\right)^{\gamma-1}$	$t > 0$	a -scale β -shape γ -shape	$a > 0$ $\beta > 0$ $\gamma > 0$	-

^a $\psi(x)$ and $\psi'(x)$ denote the digamma function and its first derivative

^bThe FIMs of exponentiated Rayleigh and exponentiated Weibull distributions are not calculated, as the FIM of exponentiated Rayleigh distribution contains highly non-linear implicit terms (Kundu and Raqab 2005), and the FIM of exponentiated Weibull distribution is not completely known (Pal et al. 2006)

$$\hat{\theta} - z_{\delta/2} \cdot \sqrt{[I_{ii}(\hat{\theta})]_{i=1,2,\dots,p}} < \theta < \hat{\theta} + z_{\delta/2} \cdot \sqrt{[I_{ii}(\hat{\theta})]_{i=1,2,\dots,p}} \tag{4}$$

$[I_{ii}(\hat{\theta})]_{i=1,2,\dots,p}$ is the vector of diagonal entries of the FIM and $z_{\delta/2}$ is the critical value corresponding to a significance level of $\delta/2$ on the standard normal distribution (Hogg et al. 2005).

The FIMs of 11 distributions are provided in Table 2. We excluded the exponentiated Rayleigh and exponentiated Weibull distributions because the FIM of the exponentiated Rayleigh distribution contains highly non-linear implicit formulae (Kundu and Raqab 2005), and the FIM of exponentiated Weibull distribution is not completely known (Pal et al. 2006).

Table 3 presents the maximum likelihood estimated parameter values along with their asymptotic standard deviations and confidence bounds. For the Pareto distribution, we calculated the exact variances (Quandt 1966) of the estimated parameters using the following formula:

$$\begin{aligned} \sigma^2(\hat{\alpha}) &= \frac{\alpha^2 n \beta}{(n\beta - 2)(n\beta - 1)^2} && \left(n > \frac{2}{\beta}\right) \\ \sigma^2(\hat{\beta}) &= \frac{\beta^2 n^2}{(n-2)^2(n-3)} && (n > 3) \end{aligned} \tag{5}$$

Table 3 also reveals many insights into the distributional properties. For example, estimated shape parameters for the gamma, lognormal, Weibull, and exponentiated exponential are greater than 1.0, indicating that the associated hazard curves are monotonically increasing. A similar observation is noted for the exponentiated Weibull distribution, with $\beta\gamma > 1$ (Pal et al. 2006). It should be emphasized that the asymptotic standard deviations of estimated model parameters, which, in this case, vary from 0.007673 (exponentiated exponential) to 4.725639 (gamma), do not necessarily correspond to a good or a bad fit of the underlying distribution. Rather, the impact of the asymptotic standard deviations could be examined to understand the accuracy of the model fitness (Hogg et al. 2005).

Model validation

Two popular goodness-of-fit tests, namely the maximum likelihood criterion and its modification, called the Akaike information criterion (AIC), and Kolmogorov-Smirnov (K-S) minimum distance criterion, are used for model selection and validation. A brief description of these methods is presented below.

The maximum likelihood criterion is entirely based on the MLE method. It uses the log-likelihood values to prioritize the competing models. The higher the likelihood value, the better is the model (Johnson et al. 1995). The

maximum likelihood criterion, however, assumes that the number of parameters in each model is the same. To relax this presumption and to account for model complexity due to a greater number of parameters, the AIC was employed. This is an extension of the maximum likelihood criterion. The AIC values are calculated as $2k - 2L$, where k is the number of parameters in the model and L is the log-likelihood value. The model with the minimum AIC value is marked as the most economical model (Johnson et al. 1995).

The K-S minimum distance criterion prioritizes the competing models based on their ‘‘closeness’’ to the empirical distribution function $F_n(t)$, defined as $F_n(t) = \frac{\text{Number of } t_i \leq t}{n}$. $F_n(t)$ defined in this manner becomes a step function. Now, if we assume that there are two competitive models F and G , then, the corresponding K-S distances are calculated as

$$\begin{aligned} D_1 &= \sup_{-\infty < t < \infty} |F_n(t) - F(t)| \\ D_2 &= \sup_{-\infty < t < \infty} |F_n(t) - G(t)| \end{aligned} \tag{6}$$

where \sup_t denotes the supremum of the set of distances. If $D_1 < D_2$, model F is chosen; otherwise we choose model G . Monte Carlo simulations are often utilized to generate thousands of data points to obtain the maximum distance between the empirical model and the fitted model. However, it is recommended to prioritize the competitive models based on their overall fit as revealed in various K-S plots (Johnson et al. 1995; Murthy et al. 2004).

The non-linear equation solver package *fsolve()* function of the MATLAB 7.10 software (MATLAB 2010) was used for the MLE estimations. The initial solution vectors for these estimations are usually determined from the graphical parametric approximations (Murthy et al. 2004). In particular, for the initial approximations of the scale and shape parameters of the three-parameter exponentiated Weibull distribution, we used MLE estimates of corresponding Weibull parameters.

The log-likelihood, AIC, and K-S distance values for each competing probability model were calculated and are presented in Table 4. It is observed that the exponential model has the lowest AIC value, indicating this model was the most economic model to fit the present data. The AIC values of gamma, lognormal, Weibull, inverse Gaussian, exponentiated exponential, exponentiated Rayleigh, and exponentiated Weibull distributions are larger than those of the exponential model, and these AIC values themselves are close to each other. Therefore, these seven competitive models may be categorized into a common group of models that provide an intermediate fit to the Kachchh catalog. The rest (Levy, Maxwell, Rayleigh, and inverse Weibull) have larger AIC values, indicating a poor fit to the present catalog. On the other hand, the K-S distances corresponding to the

Table 3 Estimated parameter values along with their asymptotic standard deviations and confidence bounds

Model	Parameter values		Asymptotic standard deviation		Confidence interval (95 %)	
					Lower	Upper
Exponential	α	13.346380	σ_α	3.566970	6.355119	20.337641
Gamma	α	11.312436	σ_α	4.725639	2.050184	20.574688
	β	1.179797	σ_β	0.398276	0.399176	1.960418
Lognormal	α	2.167443	σ_α	0.271801	1.634713	2.700173
	β	1.013243	σ_β	0.191485	0.637932	1.388554
Weibull	α	14.173818	σ_α	3.347283	7.613143	20.734493
	β	1.191600	σ_β	0.248309	0.704914	1.678286
Levy	α	5.089593	σ_α	1.923685	1.319170	8.860016
Maxwell	α	9.953759	σ_α	1.086044	7.825113	12.082405
Pareto ^a	α	1.169863	σ_α	0.232367	0.714424	1.169863
	β	0.497375	σ_β	0.174958	0.154457	0.840293
Rayleigh	α	12.190815	σ_α	1.629066	8.997846	15.383784
Inverse Gaussian	α	13.346380	σ_α	4.543217	4.441675	22.251085
	β	8.226885	σ_β	3.109476	2.132312	14.321458
Inverse Weibull	α	5.169580	σ_α	1.499738	2.230094	8.109066
	β	0.970010	σ_β	0.202133	0.573829	1.366191
Exponentiated exponential	α	0.089602	σ_α	0.007673	0.074563	0.104641
	β	1.336289	σ_β	0.368175	0.614666	2.057912
Exponentiated Rayleigh ^b	α	22.635345	–			
	β	0.463758				
Exponentiated Weibull ^b	α	15.964143	–			
	β	1.319327				
	γ	0.847971				

^aFor Pareto distribution, we calculated (Quandt 1966) exact standard deviations ($\sigma_\alpha, \sigma_\beta$) of the estimated parameters; also, the upper confidence bound is capped at 1.169863, as $\alpha < t$

^bThe parametric uncertainties of exponentiated Rayleigh and exponentiated Weibull distributions are not calculated, as the FIMs are not completely known (Pal et al. 2006)

exponential and gamma models are the smallest, implying that these two models provide the best fit. In contrast, the K-S distances corresponding to Levy, Maxwell, Pareto, Rayleigh, inverse Gaussian, and exponentiated Rayleigh are the largest, implying that these models have a poor fit. The remaining models (lognormal, Weibull, inverse Weibull, exponentiated exponential, and exponentiated Weibull) have intermediate K-S values, indicating their intermediate fit to the present earthquake catalog. To support and extend this discussion on model fitness and model comparisons, Fig. 3 provides a number of K-S graphs that examine the overall fit of the competitive models.

From Fig. 3, it is observed that some pairs of distributions are very close to each other, making them almost indistinguishable. Examples are the exponential and gamma, the gamma and exponentiated exponential, the Weibull and exponentiated Weibull, and the exponentiated exponential and exponentiated Weibull distributions. However, some other pairs have differences in overall fit. These include the

lognormal and inverse Gaussian, the exponential and exponentiated Rayleigh, the exponential and exponentiated exponential, the exponential and exponentiated Weibull, and the Rayleigh and exponentiated Rayleigh distributions. The K-S plots of Levy, Maxwell, Pareto, Frechet, and Rayleigh distributions clearly indicate a poor fit to the present data. It is also observed that, unlike the Weibull and exponentiated Weibull “parent-child” pair, the exponential and exponentiated exponential pair and the Rayleigh and exponentiated Rayleigh pair do not preserve a close fit to the Kachchh catalog. In addition, the abscissa values where the maximum K-S distances are achieved are not identical for all studied distributions.

In summary, it is concluded from Fig. 3 and Table 4 that the exponential model or the Poissonian random distribution provides the best fit to the present data. The gamma, lognormal, Weibull, inverse Gaussian, exponentiated exponential, exponentiated Rayleigh, and exponentiated Weibull models provide an intermediate fit. The remaining models, namely Levy, Maxwell, Pareto, Rayleigh, and inverse

Table 4 Model selection and validation from two goodness-of-fit tests: the maximum log-likelihood criterion ($\ln L$) and the Kolmogorov-Smirnov (K-S) minimum distance criterion

Distribution	Maximum likelihood		K-S min. distance K-S distance ^a
	$\ln L$	AIC	
Exponential	-50.2774	102.5549	0.1228 (15.8136)
Gamma	-50.02s05	104.0411	0.1227 (06.9809)
Lognormal	-50.3935	104.7871	0.1495 (15.8136)
Weibull	-49.9635	103.9271	0.1505 (06.9809)
Levy	-53.9911	109.9821	0.1709 (20.0959)
Maxwell	-59.9866	121.9731	0.2819 (06.9809)
Pareto	-54.1220	112.2439	0.3488 (06.9809)
Rayleigh	-53.6749	109.3498	0.3055 (05.1123)
Inverse Gaussian	-50.6296	105.2592	0.1790 (15.8136)
Inverse Weibull	-51.8923	107.7907	0.1496 (05.1123)
Exponentiated exponential	-49.9800	103.9601	0.1406 (06.9809)
Exponentiated Rayleigh	-49.9891	103.9783	0.1714 (06.9809)
Exponentiated Weibull	-49.9611	105.9223	0.1549 (06.9809)

$\ln L$ – log-likelihood value, AIC – value of Akaike information criterion

^aNumber in parenthesis is the abscissa value, where K-S distance is achieved

Weibull models, provide a poor fit to the present earthquake catalog. One possible reason for such a poor fit of Levy, Maxwell, Pareto, Rayleigh, and inverse Weibull models could be the non-consideration of the smaller magnitude events in the present study ($M < 5.0$). In general, the heavy-tailed models such as Levy and Pareto, or the extreme value distributions such as Frechet, offer a good fit to a data in which the frequency of smaller events are higher than the frequency of larger events (Johnson et al. 1995). Thus, for the present catalog that consists of only larger events ($M \geq 5.0$), the heavy-tailed models do not provide a suitable fit.

Earthquake hazard assessment

After analyzing the relative model fitness of different probability distributions, the earthquake hazards of the Kachchh region will be assessed in terms of the estimated recurrence interval and the conditional probability values. A number of conditional probability curves are also generated for different elapsed times ($\tau = 0, 5, \dots, 60$ years) to appraise the long-term earthquake forecasting in Kachchh and its adjoining regions. The best fit and the intermediate fit probability models were used in this calculation.

The mean recurrence interval for a magnitude 5.0 or higher event in the Kachchh region was calculated to be 13.35 ± 10.91 years (from exponential distribution), whereas the estimated cumulative probability values were found to be 0.8–0.9 by 2027–2036. The conditional probability values (using Eq. 1) and the associated conditional probability curves (hazard curves) were also obtained for different

elapsed times. The conditional probability values for an elapsed time of 9 years (i.e., March 2015) are tabulated in Table 5, while the conditional probability curves for the Kachchh region are presented in Fig. 4.

Table 5 shows that the uncertainties of parametric estimations are favorably accounted by providing a range of conditional probability values; 95 % confidence intervals of the estimated parameters are used (refer Table 3). For the exponentiated Rayleigh and the exponentiated Weibull distributions, the “absolute” conditional probabilities are presented because of the non-availability of their parametric uncertainties (Pal et al. 2006). It is observed that the conditional probability values of a magnitude 5.0 or higher event reaches 0.8–0.9 by 2034–2043. Moreover, it is observed that the upper confidence values of estimated parameters usually lead to lower conditional probabilities in comparison with the probabilities obtained from the lower confidence values of the estimated parameters. For gamma and lognormal distributions, the differences of the upper and lower conditional probability values are higher than for the exponential, Weibull, inverse Gaussian, and exponentiated exponential distributions. For smaller waiting times, the conditional probabilities of lognormal and inverse Gaussian distributions are larger compared with other competitive distributions. However, for larger waiting times, these probability values gradually become smaller compared with the other distributions. In fact, probability values from all distributions are observed to converge to the highest probability value for large waiting times (about 60 years).

In addition to the tabulated conditional probability values, a few conditional probability curves are plotted in Fig. 4 to examine the long-term seismicity of the study region. These hazard curves have many direct and indirect applications in city planning, designing seismic insurance products, location identification of lifeline structures, seismic zonation, and revision of building codes (SSHAC: Senior Seismic Hazard Analysis Committee 1997; Yadav et al. 2008).

It is concluded that the results from the present study are largely consistent with the prior research by Tripathi (2006) and Yadav et al. (2008). Tripathi (2006) conducted a probabilistic hazard assessment for the Kachchh region from three probability models (gamma, lognormal, and Weibull distributions). The forecasting was based on an earthquake catalog with ten $M \geq 5.0$ events. The results revealed high probability values of earthquake occurrence after 28–42 years for an $M \geq 5.0$ event and after 47–55 years for an $M \geq 6.0$ event, with reference to the last event in 2001 (Tripathi 2006). In a similar effort, Yadav et al. (2008) applied the same set of three probability models to an updated earthquake catalog of the Kachchh region to examine probabilistic earthquake hazards in terms of the estimated recurrence interval and the conditional probability values. The MLE was used for parameter estimations. They

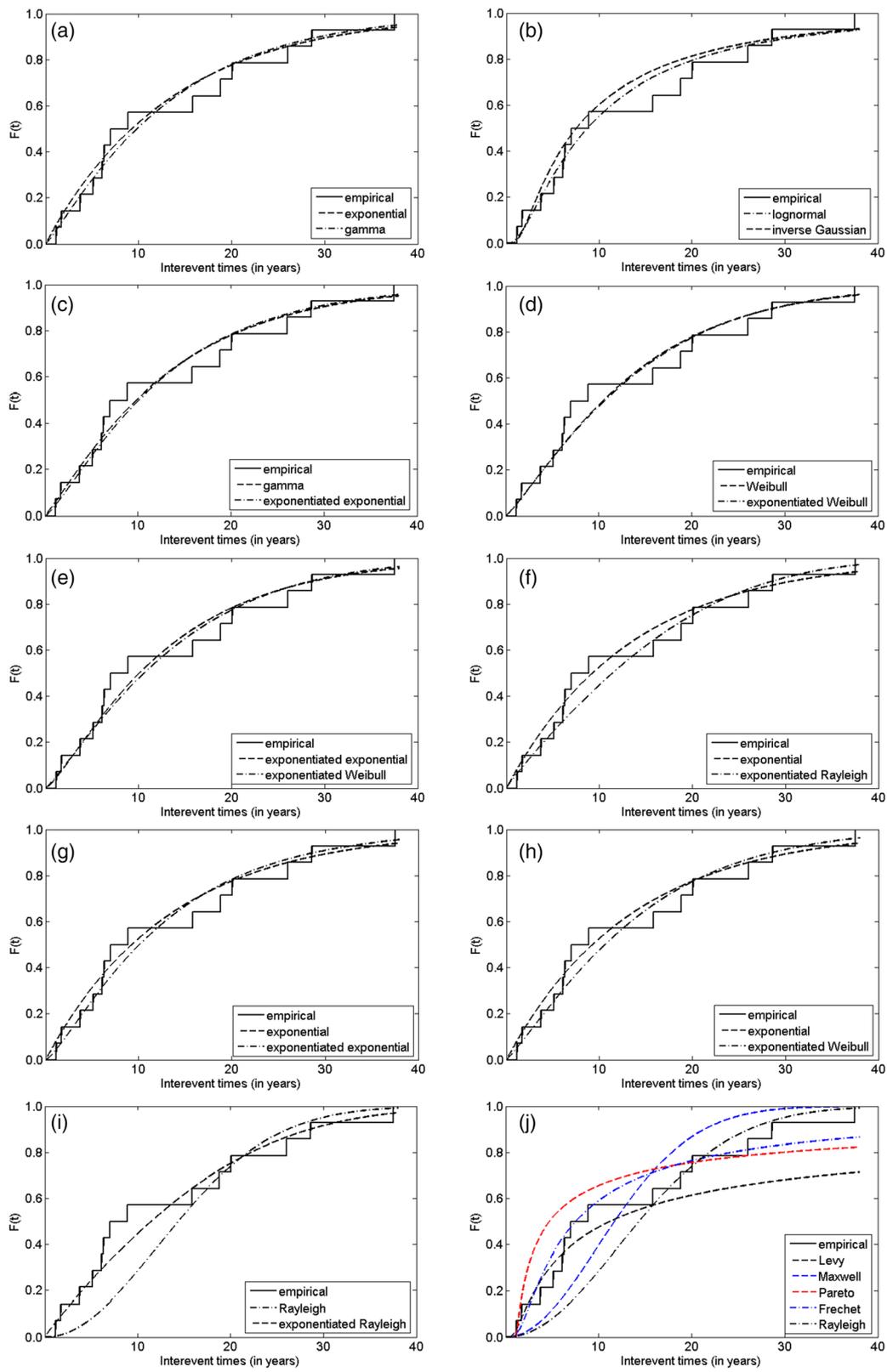


Fig. 3 (See legend on next page.)

(See figure on previous page.)

Fig. 3 K-S graphs for model comparison. This figure shows a number of K-S plots for the following studied models: **(a)** exponential and gamma, **(b)** lognormal and inverse Gaussian, **(c)** gamma and exponentiated exponential, and **(d)** Weibull and exponentiated Weibull, **(e)** exponentiated exponential and exponentiated Weibull, **(f)** exponential and exponentiated Rayleigh distributions, **(g)** exponential and exponentiated exponential, **(h)** exponential and exponentiated Weibull, **(i)** Rayleigh and exponentiated Rayleigh, and **(j)** all poor fitting distributions

calculated estimated recurrence intervals to be 13.34 ± 10.91 years and the conditional probability values to be 0.8–0.9 in 2027–2034. The Weibull model provided the best relative fit ($\ln L = -49.954$), whereas the gamma had an intermediate fit ($\ln L = -49.957$) and the lognormal had relatively poor fit ($\ln L = -50.384$) to the Kachchh catalog. Yadav et al. (2008) mentioned that the difference in the likelihood functions of gamma and Weibull was negligible, implying their similar nature of model fitness.

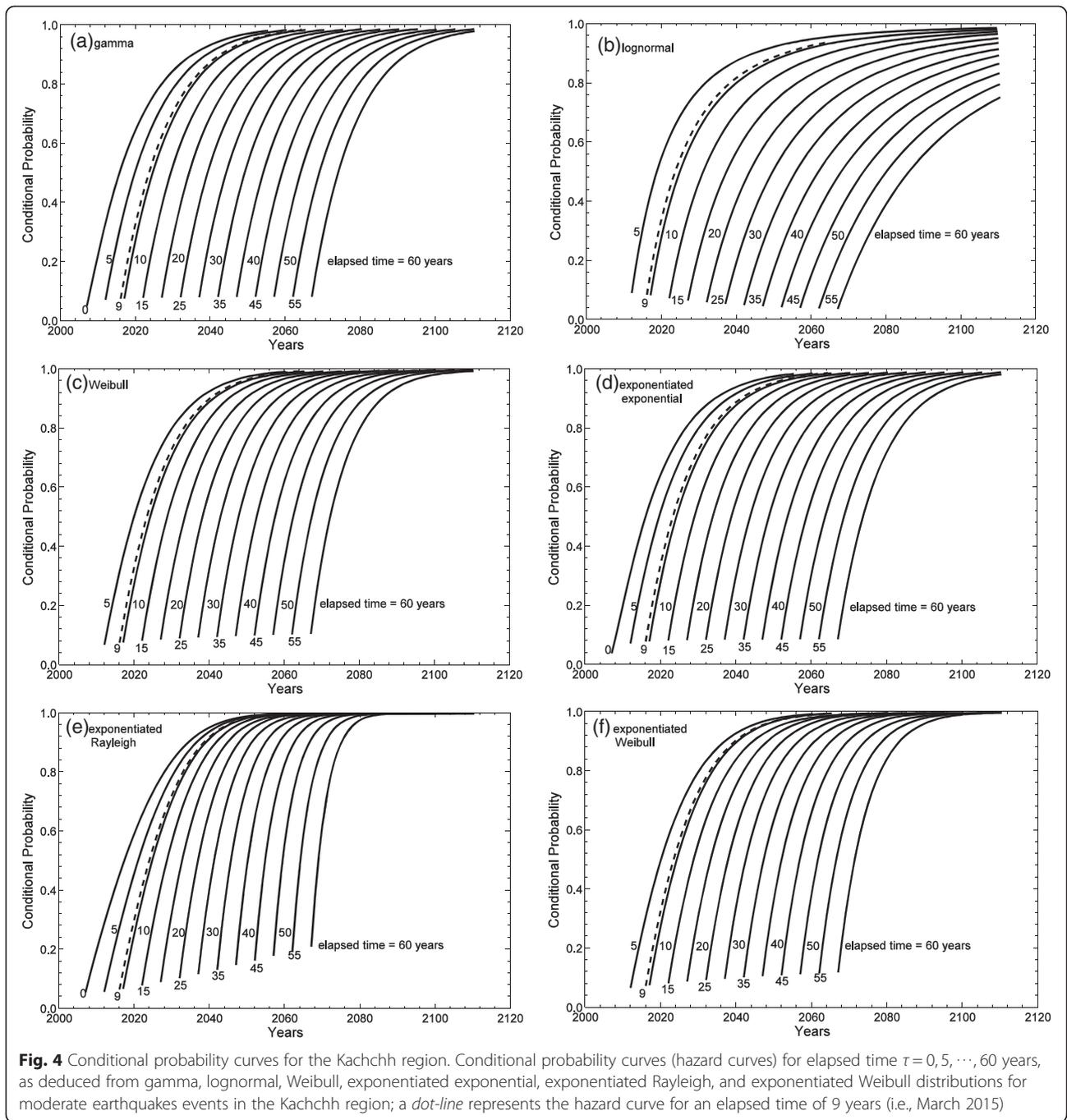
Discussions

Kachchh and its adjoining regions suffer from a number of high intensity yet infrequent intraplate earthquakes (Gupta et al. 2001; Rastogi 2001). The recent 2001 Bhuj event (Mw 7.7), which caused a huge loss of 14,000 lives, reignited questions on our understanding of earthquake genesis (Gupta et al. 2001; Choudhury et al. 2014). Scientists have been trying to examine earthquake processes and associated hazards in Kachchh from many different

Table 5 Estimated conditional probability values for an elapsed time of 9 years (i.e., March 2015)

<i>v</i>	Year	Exponential	Gamma	Lognormal	Weibull
1	2016	0.07 (0.05–0.15)	0.08 (0.02–0.42)	0.08 (0.05–0.16)	0.07 (0.05–0.08)
4	2019	0.26 (0.18–0.47)	0.27 (0.07–0.88)	0.29 (0.18–0.49)	0.27 (0.19–0.28)
7	2022	0.41 (0.29–0.67)	0.43 (0.13–0.98)	0.44 (0.29–0.68)	0.43 (0.33–0.43)
10	2025	0.53 (0.39–0.79)	0.55 (0.18–0.99)	0.55 (0.37–0.79)	0.57 (0.46–0.54)
13	2028	0.62 (0.47–0.87)	0.65 (0.24–0.99)	0.63 (0.44–0.86)	0.67(0.58–0.63)
16	2031	0.70 (0.54–0.92)	0.73 (0.30–0.99)	0.69 (0.50–0.90)	0.75 (0.67–0.69)
19	2034	0.76 (0.61–0.95)	0.79 (0.36–0.99)	0.74 (0.56–0.93)	0.81 (0.75–0.76)
22	2037	0.81 (0.66–0.97)	0.83 (0.41–0.99)	0.78 (0.60–0.95)	0.86 (0.79–0.82)
25	2040	0.85 (0.71–0.98)	0.87 (0.46–0.99)	0.82 (0.64–0.96)	0.90 (0.83–0.87)
28	2043	0.88 (0.75–0.99)	0.90 (0.51–0.99)	0.84 (0.67–0.97)	0.92 (0.85–0.91)
31	2046	0.90 (0.78–0.99)	0.92 (0.56–0.99)	0.86 (0.70–0.98)	0.94 (0.88–0.94)
34	2049	0.92 (0.81–0.99)	0.94 (0.60–0.99)	0.88 (0.72–0.98)	0.96 (0.90–0.96)
37	2052	0.94 (0.84–0.99)	0.95 (0.64–0.99)	0.90 (0.75–0.99)	0.97 (0.91–0.97)
40	2055	0.95 (0.86–0.99)	0.96 (0.67–0.99)	0.91 (0.77–0.99)	0.98 (0.93–0.98)
<i>v</i>	Year	Inverse Gaussian	Exponentiated exponential	Exponentiated Rayleigh ^a	Exponentiated Weibull ^a
1	2016	0.09 (0.06–0.14)	0.08 (0.07–0.08)	0.06	0.07
4	2019	0.29 (0.22–0.43)	0.28 (0.28–0.29)	0.25	0.27
7	2022	0.43 (0.34–0.60)	0.44 (0.44–0.46)	0.40	0.43
10	2025	0.54 (0.44–0.72)	0.57 (0.56–0.59)	0.54	0.56
13	2028	0.62 (0.52–0.79)	0.67 (0.65–0.69)	0.65	0.67
16	2031	0.68 (0.58–0.85)	0.74 (0.72–0.77)	0.75	0.75
19	2034	0.73 (0.63–0.88)	0.80 (0.78–0.83)	0.82	0.81
22	2037	0.77 (0.67–0.91)	0.85 (0.83–0.88)	0.87	0.86
25	2040	0.81 (0.71–0.93)	0.88 (0.86–0.91)	0.92	0.90
28	2043	0.83 (0.74–0.95)	0.91 (0.89–0.93)	0.94	0.93
31	2046	0.86 (0.77–0.96)	0.93 (0.91–0.95)	0.96	0.95
34	2049	0.87 (0.79–0.97)	0.95 (0.93–0.96)	0.98	0.96
37	2052	0.89 (0.81–0.98)	0.96 (0.94–0.97)	0.99	0.97
40	2055	0.90 (0.83–0.98)	0.97 (0.96–0.98)	0.99	0.98

^aAbsolute conditional probability values are presented, as the parametric uncertainties are unknown



approaches. These include paleoseismic and active fault mapping (e.g., Rajendran and Rajendran 2001; Morino et al. 2008), seismology (e.g., Mandal et al. 2005; Choudhury et al. 2014), GPS geodesy (e.g., Jade et al. 2001; Miyashita et al. 2001; Reddy and Sunil 2008), and geotechnical investigations (e.g., Vipin et al. 2013). This study, in contrast, focused on stochastic earthquake recurrence modeling from thirteen different probability distributions. A statistical strategy was developed to evaluate earthquake hazards by specifying the

estimated recurrence interval and conditional probability values from these probability distributions.

An alternative approach to assessing earthquake hazards is to simulate historical earthquake events and quantify the relative chances of occurrence in each subdivision of the study region. The results could be combined with geodetic observations to determine strain accumulation, or with paleoseismic investigations to reconstruct the chronology of the past events (Lee et al. 2011). These combined methods are useful for examining earthquake

risks in various parts of the study region, and thus have drawn significant attention from earthquake insurance agencies. Nevertheless, one limitation of these studies is that the devastating earthquakes do not always occur on the mapped faults. Therefore, the city planners and the seismic insurance product designers often require an empirical earthquake hazard model for a large geographical region of interest.

In recent years, there have been enormous efforts with alarm-based earthquake forecasting techniques. Examples are the *pattern informatics* (PI) approach (e.g., Rundle et al. 2003) that uses a pattern recognition technique to capture seismicity dynamics of an area, the *relative intensity* (RI) approach (e.g., Zechar and Jordan 2008) that utilizes smoothed historical seismicity based on extrapolated rate of occurrence of small events, and a *moment ratio* (MR) based method (e.g., Talbi et al. 2013) that uses the ratio of first- and second-order moments of earthquake interevent times as a precursory alarm to forecast large earthquakes. The comparison of these competitive forecasting models encompasses a number of likelihood testing methods, such as the N-test for data consistency in expected number space, the L-test for data consistency in likelihood-space, and the R-test for relative performance checking of seismicity models (Schorlemmer et al. 2007). The working group on Regional Earthquake Likelihood Models (RELM), supported by the Southern California Earthquake Centre (SCEC) and United States Geological Survey (USGS), or the group of Collaboratory for the Study of Earthquake Predictability (CSEP), facilitates such testing methods as a part of their earthquake research and forecasting programs (Jordan 2006; Schorlemmer et al. 2007; Shebalin et al. 2014). These simulation-based model-testing strategies usually require a controlled environment with a complete list of small to moderate historical events and a detailed seismotectonic map of the study region. In addition, the test region comprising smaller grids should experience sufficient earthquake events (during the test period) to evaluate the alarm-based earthquake predictions (Schorlemmer et al. 2007). In the present study, however, the models could not be compared in the suite of N-test, L-test, and R-test because of the lack of seismotectonic understanding and sufficient sophistication of seismic events in the Kachchh region (e.g., microseismicity, deformation rates, and fault maps). For this reason, two statistical goodness-of-fit tests were employed to examine the performance of each studied distribution such as: the maximum likelihood criterion with its modification to AIC and the K-S minimum distance criterion.

While the methodology described here focuses on a statistical analysis of estimating earthquake interevent times and conditional probabilities for earthquakes of magnitude 5.0 and higher in the Kachchh region, a physical correlation

of the obtained results would be valuable and may be considered as our future work. Nevertheless, we believe that the research and results provided in this paper demonstrate resurgence in statistical seismology and should be utilized for large global seismic databases.

Conclusions

The present investigation led to the following conclusions:

1. The thirteen probability models that were studied may be categorized into three groups on the basis of their performance against the current earthquake catalog of Kachchh region. The best fit came from the exponential distribution. An intermediate fit came from the gamma, lognormal, Weibull, inverse Gaussian, exponentiated exponential, exponentiated Rayleigh, and exponentiated Weibull models. The remainder of the models, namely Levy, Maxwell, Pareto, Rayleigh, and inverse Weibull, fit poorly to the present data.
2. The hazard curves for different elapsed times reveal high seismicity in the geographic region of interest. The expected mean recurrence interval of a magnitude 5.0 or higher earthquake was calculated to be 13.35 ± 10.91 years, whereas the estimated cumulative probability and conditional probability values reached 0.8–0.9 by 2027–2036 and 2034–2043, respectively.
3. The estimated seismic recurrence intervals and conditional probability values are largely consistent with the previous studies. Nevertheless, identification of the most suitable probability model and related coverage (e.g., uncertainty estimation, hazard assessment) provide additional support to seismologists and earthquake professionals to quantitatively compare various models to improve earthquake hazard analyses and associated applications in the seismically active Kachchh region.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

SP carried out all numerical computations and drafted the initial version of the manuscript. OD has greatly edited the manuscript and helped in fine-tuning of the results and discussions. Both the authors have thoroughly read and approved the final manuscript to submit in the EPS journal.

Authors' information

SP is a Post-Doctoral Fellow in the Department of Civil Engineering, Indian Institute of Technology (IIT) Kanpur, India. He has earned his PhD in Civil Engineering and Masters in Mathematics from IIT Kanpur. He has been interested in modeling natural hazards from statistical seismology and geodetic (GPS) viewpoint. OD is a Professor in the Department of Civil Engineering, IIT Kanpur. His major research areas include remote-sensing applications, GIS, GPS, and natural hazard management problems. He has published more than 30 papers in reputed international and national journals. He has supervised more than 70 Masters students and 4 Doctoral students.

Acknowledgements

We sincerely thank the two anonymous reviewers for their valuable comments and suggestions that have improved the quality of this paper. We are thankful to the editor-in-chief Prof. Yasuo Ogawa and the associate editor Prof. Azusa Nishizawa for their quick editorial handling and continuous encouragements toward improvement of the manuscript. We are grateful to Prof. Debasis Kundu, Prof. Teruyuki Kato, Prof. Ronald Burgmann, and Prof. John Rundle for their important suggestions at various stages of this study. We also thank Dr. RBS Yadav from Kurukshetra University for introducing the concept of stochastic earthquake recurrence modeling. We acknowledge the help of Dr. Santiswarup Sahoo in preparation of some of the figures. The first author (SP) acknowledges IIT Kanpur for the Post-Doctoral Fellowship. The generic mapping tool (GMT) system (Wessel and Smith 1995) and the MATLAB software (MATLAB 2010) were used for plotting and numerical computation purposes.

Received: 14 May 2015 Accepted: 22 July 2015

Published online: 15 August 2015

References

- Anagnos T, Kiremidjian AS (1988) A review of earthquake occurrence models for seismic hazard analysis. *Probab Eng Mech* 3(1):1–11
- Bak P, Christensen K, Danon L, Scanlon T (2002) Unified scaling law for earthquakes. *Phys Rev Lett* 88(17):178501–178504
- Baker JW (2008) An introduction to probabilistic seismic hazard analysis. [http://www.stanford.edu/~bakerjw/Publications/Baker_\(2008\)_Intro_to_PSHA_v1_3.pdf](http://www.stanford.edu/~bakerjw/Publications/Baker_(2008)_Intro_to_PSHA_v1_3.pdf). Accessed on 15 March 2015
- BIS: Bureau of Indian Standards (2002) Indian standard criteria for earthquake resistant design of structures, part 1—general provisions and buildings. IS 1893(Part 1):39
- Biswas SK (1987) Regional tectonic framework structure and evolution of the western marginal basins of India. *Tectonophysics* 135:307–327
- Biswas SK (2005) A review of structure and tectonics of Kutch basin, western India with special reference to earthquakes. *Current Sci* 88(10):1592–1600
- Chandra U (1977) Earthquakes of peninsular India—a seismotectonic study. *Bull Seismol Soc Am* 67:1387–1413
- Chen C, Wang JP, Wu YM, Chan CH (2013) A study of earthquake inter-occurrence distribution models in Taiwan. *Nat Hazards* 69(3):1335–1350
- Choudhury P, Chopra S, Roy KS, Rastogi BK (2014) A review of strong motion studies in Gujarat state of western India. *Nat Hazards* 71:1241–1257
- Chung WY, Gao H (1995) Source parameter of the Anjar earthquake of July 21, 1956, India, and its seismotectonic implications for the Kutch rift basin. *Tectonophysics* 242:281–292
- Cornell CA (1968) Engineering seismic risk analysis. *Bull Seismol Soc Am* 58:1583–1606
- Dionysiou DD, Papadopoulos GA (1992) Poissonian and negative binomial modeling of earthquake time series in the Aegean area. *Phys Earth Planetary Inter* 71:154–165
- Efron B, Johnstone I (1990) Fisher information in terms of the hazard function. *Ann Stat* 18:38–62
- Feraes SG (2003) The conditional probability of earthquake occurrence and the next large earthquake in Tokyo. *Jpn J Seismol* 7:145–153
- Gupta HK, Rao NP, Rastogi BK, Sarkar D (2001) The deadliest intraplate earthquake. *Science* 291:2101–2102
- Gupta RD, Gupta RC, Sankaran PG (2004) Some characterization results based on the (reversed) hazard rate function. *Commun Stat: Theory Methods* 33(12):3009–3031
- Hagiwara Y (1974) Probability of earthquake occurrence as obtained from a Weibull distribution analysis of crustal strain. *Tectonophysics* 23:313–318
- Hogg RV, Mckean JW, Craig AT (2005) Introduction to mathematical statistics. 6th edn, PRC Press, p 718
- Jade S, Mukul M, Parvez IA, Ananda MB, Kumar PD, Gaur VK (2002) Estimates of co-seismic displacement and post-seismic deformation using global positioning system geodesy for the Bhuj earthquakes of 26 January 2001. *Current Sci* 82:748–752
- Johnson NL, Kotz S, Balakrishnan N (1995) Continuous univariate distributions. Vol. 2, 2nd edn. Wiley-Interscience, p 752
- Jordan TH (2006) Earthquake predictability, brick by brick. *Seismol Res Lett* 77(1):3–6
- Kagan YY, Schoenberg F (2001) Estimation of the upper cutoff parameter for the tapered Pareto distribution. *J Appl Probab* 38:158–175
- Kayal JR, De R, Ram S, Sriram BP, Gaonkar SG (2002) Aftershocks of the 26 January, 2001 Bhuj earthquake in western India and its seismotectonic implications. *J Geol Soc India* 59:395–417
- Kayal JR, Das V, Ghosh U (2012) An appraisal of the 2001 Bhuj earthquake (Mw7.7, India) source zone: fractal dimension and b value mapping of the aftershock sequence. *Pure Appl Geophys* 169:2127–2138
- Kundu D, Raqab MZ (2005) Generalized Rayleigh distribution: different methods of estimation. *Comput Stat Data Anal* 49:187–200
- Lee YT, Turcotte DL, Holliday JR, Sachs MK, Rundle JB, Chen CC, Tiampo KF (2011) Results of the Regional Earthquake Likelihood Models (RELM) test of earthquake forecasts in California. *Proc Natl Acad Sci* 108(40):16533–16538
- Mandal P, Horton S (2007) Relocation of aftershocks, focal mechanisms and stress inversion: Implications toward the seismo-tectonics of the causative fault zone of Mw7.6 2001 Bhuj earthquake (India). *Tectonophysics* 429:61–78
- Mandal P, Chadha RK, Satyamurty C, Raju IP, Kumar N (2005) Estimation of site response in Kachchh, Gujarat, India, region using H/V spectral ratios of aftershocks of the 2001 Mw 7.7 Bhuj earthquake. *Pure Appl Geophys* 162:2479–2504
- MATLAB: Matrix Laboratory, version 7.10.0 (2010), The MathWorks Inc., Natick, Massachusetts, United States
- Matthews MV, Ellsworth WL, Reasenber PA (2002) A Brownian model for recurrent earthquakes. *Bull Seismol Soc Am* 92(6):2233–2250
- Miyashita K, Vijaykumar K, Kato T, Aoki Y, Reddy CD (2001) Postseismic crustal deformation deduced from GPS observations, in a Comprehensive Survey of the 26 January 2001 Earthquake (Mw 7.7) in the State of Gujarat, India. In: Sato T et al. (ed) pp 46–50, Ministry of Education, Culture, and Sports, Science and Technology, Tokyo. (Available online: http://www.sthirosaki-u.ac.jp/~tamao/Gujarat/print/Gujarat_4.pdf).
- Morino M, Malik JN, Mishra P, Bhuiyan C, Kaneko F (2008) Active fault traces along Bhuj fault and Katrol hill fault, and trenching survey at Wandhay, Kachchh, Gujarat, India. *J Earth Syst Sci* 117(3):181–188
- Mulgaria F, Tinti S (1985) Seismic sample area defined from incomplete catalogs: an application to the Italian territory. *Phys Earth Planetary Sci* 40(4):273–300
- Murthy DNP, Xie M, Jiang R (2004) Weibull models. John Wiley and Sons, New Jersey, p 383
- Negishi H, Mori J, Sato T, Singh R, Kumar S, Hirata N (2002) Size and orientation of the fault plane for the 2001 Gujarat, India earthquake Mw7.7 from aftershock observations: a high stress drop event. *Geophys Res Lett* 29(20):10-1–10-4
- Page M, Felzer K (2015) Southern San Andreas fault seismicity is consistent with the Gutenberg-Richter Magnitude-Frequency distribution. *Bull Seismol Soc Am* 105(4); doi: 10.1785/0120140340
- Pal M, Ali MM, Woo J (2006) Exponentiated Weibull distribution. *Commun Stat Theory Methods* 32:1317–1336
- Papazachos BC, Papadimitriou EE, Kiratzi AA, Papaioannou CA, Karakaisis GF (1987) Probabilities of occurrence of large earthquakes in the Aegean and the surrounding area during the period of 1986–2006. *Pure Appl Geophys* 125:592–612
- Parvez IA, Ram A (1997) Probabilistic Assessment of earthquake hazards in the north-east Indian Peninsula and Hindukush regions. *Pure Appl Geophys* 149:731–746
- Pasari S (2015) Understanding Himalayan tectonics from geodetic and stochastic modelling. Unpublished PhD Thesis, Indian Institute of Technology Kanpur, p 376
- Pasari S, Dikshit O (2014a) Impact of three-parameter Weibull models in probabilistic assessment of earthquake hazards. *Pure Appl Geophys* 171:1251–1281. doi:10.1007/s00024-013-0704-8
- Pasari S, Dikshit O (2014b) Distribution of earthquake interevent times in northeast India and adjoining regions. *Pure Appl Geophys*. doi:10.1007/s00024-014-0776-0
- Pasari S, Dikshit O (2014c) Three parameter generalized exponential distribution in earthquake recurrence interval estimation. *Nat Hazards* 73:639–656. doi:10.1007/s11069-014-1092-9
- Quandt RE (1966) Old and new methods of estimation and the Pareto distribution. *Metrika* 10:55–82
- Rajendran CP, Rajendran K (2001) Characteristics of deformation and past seismicity associated with the 1819 Kutch earthquake, Northwestern India. *Bull Seismol Soc Am* 91:407–426
- Rastogi BK (2001) Ground deformation study of Mw7.7 Bhuj earthquake of 2001. *Episodes* 24:160–165
- Rastogi BK (2004) Damage due to the Mw7.7 Kutch, India earthquake of 2001. *Tectonophysics* 390:85–103

- Reddy CD, Sunil PS (2008) Post seismic crustal deformation and strain rate in Bhuj region, western India, after the 2001 January 26 earthquake. *Geophys J Int* 172:593–606
- Reid HF (1910) The mechanics of the earthquake, the California earthquake of April 18, 1906, vol 2. Report of the State Investigation Commission, Carnegie Institution of Washington, Washington
- Rundle JB, Turcotte DL, Shchepakov R, Klein W, Sammis C (2003) Statistical physics approach to understanding the multiscale dynamics of earthquake fault systems. *Rev Geophys* 41(4):1019
- Schorlemmer D, Gerstenberger MC, Wiemer S, Jackson DD, Rhoades DA (2007) Earthquake likelihood model testing. *Seismol Res Lett* 78:17–29
- Shebalin PN, Narteau C, Zechar JD, Holschneider M (2014) Combining earthquake forecasts using differential probability gains. *Earth Planets Space* 66:37
- Sornette D, Knopoff L (1997) The paradox of the expected time until the next earthquake. *Bull Seismol Soc Am* 87:789–798
- Sotolongo-Costa O, Antoranz JC, Posadas A, Vidal F, Vazquez A (2000) Levy flights and earthquakes. *Geophys Res Lett* 27(13):1965–1968
- SSHAC: Senior Seismic Hazard Analysis Committee (1997) Recommendations for probabilistic seismic hazard analysis: guidance on uncertainty and use of experts. US Nuclear Regulatory Commission Report CR-6372, UCRL-ID-122160, Vol. 2, Washington, DC, p 888 (<http://pbadupws.nrc.gov/docs/ML0800/ML080090004.pdf>)
- Stein S, Liu M (2009) Long aftershock sequences within continents and implications for earthquake hazard assessment. *Nature* 462:87–89
- Talbi A, Nanjo K, Zhuang J, Satake K, Hamadache M (2013) Interevent times in a new alarm-based earthquake forecasting model. *Geophys J Int* 194(3):1823–1835
- Tripathi JN (2006) Probabilistic assessment of earthquake recurrence in the January 26, 2001 earthquake region of Gujarat, India. *J Seismol* 10:119–130
- Utsu T (1984) Estimation of parameters for recurrence models of earthquakes. *Bull Earthq Res Inst, Univ Tokyo* 59:53–66
- Vipin KS, Sitharam TG, Kolathayar S (2013) Assessment of seismic hazard and liquefaction potential of Gujarat based on probabilistic approaches. *Nat Hazards* 65:1179–1195
- Wessel P, Smith WHF (1995) New version of the generic mapping tools released. *EOS Trans Am Geophys Union* 76:p329
- Working Group on California Earthquake Probabilities (2013) The uniform California earthquake rupture forecast, Version 3 (UCERF 3): USGS open file report 2013–1165 and California geological survey special report 228 (<http://pubs.usgs.gov/of/2013/1165/>)
- Yadav RBS, Tripathi JN, Rastogi BK, Das MC, Chopra S (2008) Probabilistic assessment of earthquake hazard in Gujarat and adjoining region of India. *Pure Appl Geophys* 165:1813–1833
- Yadav RBS, Tripathi JN, Rastogi BK, Das MC, Chopra S (2010) Probabilistic assessment of earthquake recurrence in northeast India and adjoining regions. *Pure Appl Geophys* 167:1331–1342
- Yazdani A, Kowsari M (2011) Statistical prediction of the sequence of large earthquakes in Iran. *IJE Trans B: Appl* 24(4):325–336
- Zechar JD, Jordan TH (2008) Testing alarm-based earthquake predictions. *Geophys J Int* 172:715–724

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com
