Journal of Volcanology and Geothermal Research 176 (2008) 277–290

Contents lists available at ScienceDirect

Journal of Volcanology and Geothermal Research

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

A statistical method linking geological and historical eruption time series for volcanic hazard estimations: Applications to active polygenetic volcanoes

Ana Teresa Mendoza-Rosas [a], Servando De la Cruz-Reyna [b,*]

[a] Universidad Nacional Autónoma de México, Posgrado en Ciencias de la Tierra, Instituto de Geofísica, Ciudad Universitaria, México 04510 D.F., Mexico
[b] Universidad Nacional Autónoma de México, Instituto de Geofísica, Ciudad Universitaria, México 04510 D.F., Mexico

ARTICLE INFO

Article history:
Received 17 September 2007
Accepted 7 April 2008
Available online 30 June 2008

Keywords:
volcanic hazard
extreme-value theory
Weibull distribution
generalized Pareto distribution
non-homogeneous Poisson process

ABSTRACT

The probabilistic analysis of volcanic eruption time series is an essential step for the assessment of volcanic hazard and risk. Such series describe complex processes involving different types of eruptions over different time scales. A statistical method linking geological and historical eruption time series is proposed for calculating the probabilities of future eruptions. The first step of the analysis is to characterize the eruptions by their magnitudes. As is the case in most natural phenomena, lower magnitude events are more frequent, and the behavior of the eruption series may be biased by such events. On the other hand, eruptive series are commonly studied using conventional statistics and treated as homogeneous Poisson processes. However, time-dependent series, or sequences including rare or extreme events, represented by very few data of large eruptions require special methods of analysis, such as the extreme-value theory applied to non-homogeneous Poisson processes. Here we propose a general methodology for analyzing such processes attempting to obtain better estimates of the volcanic hazard. This is done in three steps: Firstly, the historical eruptive series is complemented with the available geological eruption data. The linking of these series is done assuming an inverse relationship between the eruption magnitudes and the occurrence rate of each magnitude class. Secondly, we perform a Weibull analysis of the distribution of repose time between successive eruptions. Thirdly, the linked eruption series are analyzed as a non-homogeneous Poisson process with a generalized Pareto distribution as intensity function. As an application, the method is tested on the eruption series of five active polygenetic Mexican volcanoes: Colima, Citlaltépetl, Nevado de Toluca, Popocatépetl and El Chichón, to obtain hazard estimates.

1. Introduction

Volcanic activity usually results from the interaction of many independent physical and geological processes acting over different time scales. The occurrence of volcanic eruptions may depend on the unknown nature of magma feeding from deeper sources, as well as the conditions of a previously resident magma, the nature of the magma mixing processes, the regional stresses, the local crustal composition and structure, the fluid distribution and composition under the volcano, the degree of fracturing, and even on some meteorological agents. These and other factors interact in complex ways introducing a random behavior on the time series of volcanic eruption occurrences.

On the other hand, volcanic eruptions may represent a serious threat on the people dwelling near a volcano, particularly when their perception of risk is negatively influenced by a large repose time, or by the lack of clear evidences of major past activity. Volcanic risk was first formally defined in UNDRO (1979) as a measure of the expected number of lives lost, persons injured, damage to property and disruption of economic activity as a result of a particular volcanic event. It was defined as the product of volcanic hazard, vulnerability and elements at risk (Fournier d'Albe, 1979). The volcanic hazard is consistently defined as the probability that a specific type of volcanic eruption occurs in a given area, within a given interval of time (De la Cruz-Reyna and Tilling, 2008) The volcanic risk is thus the probability of losing a certain percent of the value of a given exposed region over a given time interval caused by the possible occurrence of a particular volcanic eruption. Therefore, knowing the hazard allows designing adequate measures to reduce the risk through specific actions of vulnerability reduction.

Under the assumption that the past history of a volcano should reflect at least some relevant features of its expected future behavior, a careful analysis of the time series of past eruptions, that accounts for the scarcity of precise past eruption data, is essential to assess the hazard. The behavior of volcanic eruption time series of individual volcanoes shows a wide spectrum of possibilities. Some volcanoes show stationary patterns of activity, while others show time-dependent eruption rates. Nevertheless, combining the eruptions of large groups of volcanoes generates a definite homogeneous Poissonian behavior, as is the case of the overall global eruptive activity (De la Cruz-Reyna, 1991).

Early studies of volcanic time series were done by Wickman (1965, 1976) and Reyment (1969) used stochastic principles for the study of eruption patterns on specific volcanoes. However, the models presented

* Corresponding author.
E-mail address: sdelacrr@geofisica.unam.mx (S. De la Cruz-Reyna).

0377-0273/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.jvolgeores.2008.04.005

by Wickman did not distinguish among eruption of different types, and as he stated in his 1976 paper, such models were not tested against observed records. Other studies, analyzed specific volcanic eruption series, as was the case of the Hawaiian volcanoes (Klein, 1982) or Colima (De la Cruz-Reyna, 1993; Solow, 2001). Bebbington and Lai (1996a,b) examined whether the Weibull renewal model was adequate to describe the patterns of two New Zealand volcanoes.

Subsequent studies became increasingly sophisticated including for instance transition probabilities of Markov chains (Carta et al., 1981; Aspinall et al., 2006; Bebbington, 2007), change-point detection techniques (Mulargia et al., 1987; Burt et al., 1994), Rank-order statistics (Pyle, 1998), Bayesian analysis of volcanic activity (Ho, 1990; Solow, 2001; Newhall and Hoblitt, 2002; Ho et al., 2006; Marzocchi et al., 2008), non-homogeneous models (Ho, 1991a; Bebbington and Lai, 1996b), a mixture of Weibull distributions (Turner et al., 2007), and geostatistical hazard-estimation methods (Jaquet et al., 2000; Jaquet and Carniel, 2006).

Different parameters have been used as random variables to characterize the eruptive time series. Among them, the most frequently used are: the duration of eruptions, the interval between eruptions, the effusion rate; the volume or mass released, and the intensity of eruptions.

The probabilities of occurrence of future eruptions, and thus the volcanic hazard, may be estimated analyzing the sequence of past eruptions in a volcano, characterizing the eruptions by a measure of size that reflects their destructive potential, and assuming that the impact and effects of an eruption are proportional to both, the total mass or energy release (magnitude) and the rate of mass or energy release (intensity). The Volcanic Explosivity Index VEI is the quantity that characterizes eruptions based on those parameters (Newhall and Self, 1982). Frequently, an eruption has been defined ambiguously as a sudden, violent discharge of volcanic material, as well as a gentle, protracted pouring of lava or fumes. For our purpose we shall consider here only significant explosive eruptions, which usually are short-duration events when compared with the time between eruptions (also referred as repose time, even if minor or gentle effusive activity occurs). The volcanic eruption sequences of polygenetic volcanoes are thus considered here as point processes developing in the time axis, and the distribution of eruptions and the repose times between them are analyzed in different VEI categories or classes.

On the other hand, merging historical (usually describing more frequent smaller eruptions) and geological (usually describing larger, infrequent eruptions) eruptive data has been pointed as an important factor for a proper estimation of the likelihood of more damaging events (Marzocchi et al. 2004).

In this paper we propose a statistical methodology for estimating the volcanic hazard of future explosive eruptions using VEI — characterized sequences linking historical and geological records to obtain robust volcanic eruption time series. We first test the independence between successive eruptions to detect possible memory effects, and the stationarity, or time dependence of the explosive eruption sequences to find a possible non-homogeneity of the process. We then use a Weibull analysis to study the distribution of repose times between successive eruptions, and a non-homogeneous generalized Pareto–Poisson process (NHGPPP, as defined below) to obtain volcanic hazard estimations. We apply this method to Colima, Citlaltépetl, El Chichón, Nevado de Toluca and Popocatépetl volcanoes in México. Finally, the hazard estimates obtained with this and other methods are discussed and compared.

2. Methodology

The first step is testing the eruptive time series for independence between successive events and for the time dependence or stationarity of the process. The independence test is simply made by means of a serial correlation scatterplot (Cox and Lewis, 1966). The latter test is performed examining the repose period series for each VEI category and using a moving average test that reveals the possible existence of significantly different eruption rates, not attributable to the local rate changes expected in a stationary random process (Klein, 1982; De la Cruz-Reyna, 1996). These tests should be performed on a portion of the time series that satisfies a criterion of completeness, i.e. a portion in which no significant eruption data are missing, which in most cases is the historical eruption data set of intermediate-to-high VEI magnitudes.

A second step is the Weibull analysis of the repose periods between eruptions, which allows a quantitative description of both, stationary and non-stationary time series through the distribution shape parameter. The time-independence tests applied on the portions of the series assumed to be complete do not guarantee that the whole of the series has been stationary over its whole length. Therefore, the third step involving the link between the historical, usually complete, and the geological, probably incomplete eruptive series requires of a method that makes the estimation of hazard less sensitive to such condition. We propose here as the best estimate of the volcanic hazard

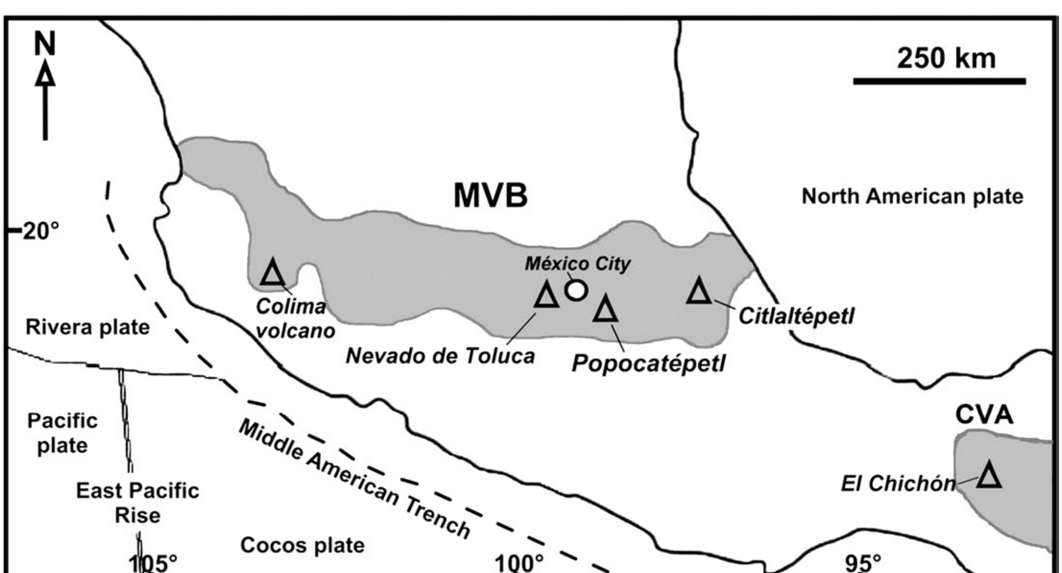

Fig. 1. Location of Colima, Nevado de Toluca, Popocatépetl, Citlaltépetl and El Chichón volcanoes.

Table 1
Historical eruptions of Colima volcano. (Adapted from De la Cruz- Reyna, 1993; Global Volcanism Program, http://www.volcano.si.edu; and Observatorio del Volcán de Colima; http://www.ucol.mx)

Year	VEI	Year	VEI
1560	2	1881	2
1576	3	1885	1
1585	4	1886	3
1590	3	1889	3
1606	4	1890	4
1611	3	1893	2
1612	2	1895	1
1622	4	1903	3
1690	3	1904	1
1749	2	1908	3
1770	3	1909	2
1795	2	1913	4
1804	2	1960	1
1818	4	1975	1
1869	3	1987	1
1872	3	1991	1
1873	1	1994	2
1874	1	1999	2
1877	1	2003	2
1879	1	2005	3
1880	1		

from the integrated time series a procedure based on the use of a non-homogenous Poisson process with a generalized Pareto distribution (GPD) as intensity function, referred herein as a non-homogeneous generalized Pareto–Poisson process (NHGPPP) (Coles, 2001). Such a process describes a series of independent, non-overlapping physical events occurring in a space A with an intensity density $\lambda(x_i)$, where x_i are the A-domain variables in which process develops. In our case, x_i are the coordinates t (time) and eruption size (VEI: magnitude–intensity) of a two-dimensional space, where the domain is limited by the historical and geological available eruption data. The process is homogeneous if λ is constant. If λ depends on either variable, the process is non-homogeneous.

Hazard and risk estimates based on catalogues assumed to be complete may include a very few or not include at all "rare" large events, with very low or unknown occurrence rates. Hazard estimates from such databases may be underrated. On the other hand, dealing with extreme values means using a set containing very few, probably incomplete data. These few data, most likely extracted from the geological record, are represented by the right tail of the repose-time distribution. Although they may have a little influence on the distribution itself, they certainly should have a significant influence on the hazard estimation. The GPD is a robust tool which allows modeling extreme values, such as the "rare", very high-magnitude eruptions, allowing for a better fit of the whole distribution. Additionally, it is less sensitive to the possible time dependence of the large-magnitude eruption sequence, since it only considers the number of exceedances over a threshold of a series that may be stationary or not.

2.1. Historical and geological time series of volcanic magnitudes

The occurrence behavior of explosive eruptions is similar to that of earthquakes, (and many other natural phenomena) in the sense that the frequency of the events decreases as their size or magnitude increases. In the case of earthquakes, the distribution of magnitudes in a region can be described by the frequency–magnitude distribution of earthquakes based on the Ishimoto and Iida (1939) and Gutenberg-Richter (1944) law

$$\log\ N = A - BM \tag{1}$$

where N is the cumulative number of earthquakes with magnitude $\geq M$, and A and B are constants that describe the power law decay of occurrences with increasing magnitude over a given time interval.

The analysis of historical data classified by VEI (Newhall and Self, 1982) shows that the VEI magnitude M_{vei} of N events may be represented as a random variable with a distribution function $N=a10^{-bM_{\text{vei}}}$. When the eruption data are analyzed in terms of the number of eruptions occurring over time intervals, this relation may be more clearly expressed in terms of the eruption occurrence rate of each class magnitude $\lambda_{M_{\text{vei}}}$ as

$$\log\lambda_{M_{\text{vie}}} = a - bM_{\text{vei}}. \tag{2}$$

Notice that Eq. (2) relates the eruption size M_{vei} with the eruption rate (number of eruptions per unit time in the magnitude class M_{vei}) unlike Eq. (1), which relates the cumulative number of earthquakes exceeding a certain magnitude. Using the cumulative number of eruptions can be used equally well, since in both cases the linearity of the relations is maintained, and only the values of the coefficients are different (see for instance Palumbo, 1997). In the present work, we prefer to use the non-cumulative occurrence rates of the VEI categories since they directly provide a more intuitive perception of the probabilities of occurrence of eruptions in each magnitude class.

Table 2
Dates and magnitudes of major Holocenic eruptions of Colima volcano inferred from eruption deposits. (Cortés et al., 2005)

Year BP	VEI
2300	>4
3600	>4
7040	>4

Table 3
The historical and geological eruptive activity of Citlaltépetl volcano. (De la Cruz-Reyna and Carrasco-Nuñez,2002)

	Year	VEI
Historical	1867	2
	1846	2
	1687	2
	1569–89	2
	1545	2
	1533–39	2
	Year BP	VEI
Geological	4100	≥4
	8500–9000	≥4
	13000	≥4

Table 4
Volcanic Explosivity Indexes of known eruptions of Popocatépetl volcano reported since the 16th century. (De la Cruz-Reyna and Tilling, 2008)

Year	VEI
1512	2
1519	3
1539–1540	2
1548	2
1571	2
1592	2
1642	2
1663	2
1664	3
1665	2
1697	2
1720	1
1804	1
1919–1920	2
1921	2
1925–1927	2
1994–1997	2
2000	3
2001–present	1–2

Table 5
The geological activity record of Popocatépetl volcano (Siebe and Macías, 2004)

Year B.P.	Eruption type
1200	Plinian
1700	Plinian
2150	Plinian
5000	Plinian
7100	Plinian
9100	Plinian
10,700	Plinian
14,000	Plinian
23,000	Plinian (sector collapse and massive debris flows)

De la Cruz-Reyna (1991) estimated the values $a=3.494$ and $b=0.789$ for the global volcanic activity based on the historical eruption data of Newhall and Self (1982) in the VEI range 3–6. Later, Simkin and Siebert (1994, 2000), integrated eruption data for various time intervals: 20, 200, 1000 and 2000 yr. Based on their Fig. 10 in volcanoes of the World (1994), the best fit for the eruption data in the VEI range 2–6 yields was graphically determined to be $a=5.8$ and $b=0.785$. Although a strongly depends on the length of the sampled period, the slope b seems to be a constant Gusev et al. (2003) obtain $b=0.75$ with the same graphical method on the Simkin and Siebert (1994) plot.

We can conclude that since the VEI is a composite estimate of mass magnitude and/or mass rate intensity, depending on which data are available, and considering that the VEI of many of the explosive eruptions listed are based on intensity related parameters (such as eruptive column height), the VEI is an appropriate parameter to characterize the eruption size for hazard calculation purposes.

The above analysis has been used on individual volcanoes to estimate the eruption rates for different VEI magnitudes (De la Cruz-Reyna, 1991, 1993; De la Cruz-Reyna and Carrasco-Núñez, 2002; De la Cruz-Reyna and Tilling, 2008) using both, the historical and the geological eruption records to obtain self-consistent series.

Here, we use Eq. (2) to link the historical and geological eruptive series.

We construct different models of the distribution of large events using the available geological information, and selecting the model which best fits the eruption rates obtained from Eq. (2). From the best estimates of the large-eruption rates, we may infer the number of eruptions that have exceeded a threshold. To calculate the probability of occurrence of more large-magnitude eruptions exceeding a threshold, we use the NHGPPP.

2.2. Repose period distribution

To analyze the characteristics of the repose periods of successive volcanic eruptions produced by a specific volcano in a given magnitude class, and particularly, the stationarity of the process, we use the Weibull distribution on the complete portion of the catalogue. This distribution has been widely applied in statistical quality control, reliability analysis of system components, earthquake hazard assessment, and many other applications (see for instance Johnson, 1966; Ferráes, 2003). It has also been used to model volcanic eruption sequences (Ho, 1991b, 1995; Bebbington and Lai, 1996b).

Table 6
The geological eruptive activity of Nevado de Toluca volcano

Eruption (name)	Years BP	VEI	References
1	3300		Macías et al. (1997b)
2 (UTP)	10,500	5	Arce et al. (2003)
3 (MTP)	12,100	4	Cervantes (2001)/Arce et al. (2005)
4	13,000	4	Arce et al. (2006)/D'Antonio et al. (2008)
5 (LTP)	21,700	4	Capra et al. (2006)
6	28,000	4	García-Palomo et al. (2002)/D'Antonio (2008)
7 (OPF)	36,000–39,000		García-Palomo et al. (2002)
8	37,000		García-Palomo et al. (2002)
9 (Pink)	42,000		Arce et al. (2003)

Table 7
Volcanic Explosivity Indexes of known eruptions of El Chichón volcano (Macías et al., 2007)

Years BP	VEI
25	5
550	4
900	3
1250	4
1500	3
1600	
1900	
2000	2–3
2500	2–3
3100	
3700	4
7500/7700	

The 2-parameter cumulative Weibull distribution function is

$$F(t) = 1 - e^{-\left(\frac{t}{\alpha}\right)^k}, \tag{3}$$

where α is a scale parameter, and k is a shape parameter.

The shape parameter is of particular interest because it characterizes the failure rate trends, i.e., reflects the stationary or non-stationary character of the time series (Yang and Xie, 2003). There are different methods to estimate the Weibull parameters (Johnson and Kotz, 1953; Ho, 1991b, 1995). In the present paper, we obtain the distribution parameters using a fairly simple graphical method (Bebbington and Lai, 1996a). The probability of having a repose period of duration greater than t has been thus obtained form the survival function $1-F(t)$.

2.3. Estimation of volcanic hazard using extreme-event statistics

In this section, the recent history, the geological record and the extreme-value techniques are used to obtain estimates of the probability of intermediate- to high-magnitude eruption events.

2.3.1. Extreme-value theory

Two methods are normally used to sample the original data of extreme events: the annual (or any other adequate time interval) maximum (AM), and the peaks or exceedances over a threshold (EOT) series. AM series are composed by the largest event occurring in a given sample time interval, so the series length equals the number of recording intervals. According to some theorems originally due to Fisher and Tippett (1928) and Gnedenko (1943), a series of sample maxima like an AM series can be described by a generalized extreme-value distribution, which includes the Gumbel, Fréchet and Weibull distributions (Gumbel, 1958). In the EOT analysis, the samples are not collected at fixed intervals and it has several important advantages over the AM approach, as it adapts better to heavy-tailed distributions and makes a more efficient use of information since it permits to include more cases through the choice of the threshold (Beguería, 2005). The EOT analysis includes all the values of the variable that exceed an a-priori determined threshold, u, defined by the transformed variable

$$Y = X - u, \tag{4}$$

for every case where $X>u$. The EOT considers all the excesses, i.e., the events above a certain level u. This level is fixed according to the model needs and provides a physically based definition of what must be considered an extreme event. Lang et al. (1999) reviewed different systematic methods for the choice of the threshold value, since it usually has a strong subjective component. Pickands (1975) demonstrated that if X is an identically independent distributed variable, a threshold value u can be found that makes the process converge to a generalized Pareto distribution (GPD). The GPD is described by a shape parameter β, a scale parameter θ, and a location parameter u (threshold), and has the following cumulative distribution function:

$$G_{\beta,\theta}(y) = 1 - \left(1 - \frac{\beta y}{\theta}\right)^{1/\beta} \quad \text{for } \beta \neq 0, \qquad G_{\beta,0}(y) = 1 - e^{-y/\theta} \quad \text{for } \beta = 0 \tag{5}$$

where $y=x-u$ is a realization of an excess. The generalized Pareto distribution has a relation with the generalized extreme-value (GEV) distribution Coles (2001). The Weibull distribution, which is a particular case of GEV to estimate extremes, needs a discrete time series generated sequentially at equidistant time intervals unlike the GPD. The GPD distribution contains the exponential distribution as a special case, when $\beta=0$ (second expression in Eq. (5)). For $\beta<0$ the distribution is long-tailed, and for $\beta>0$ it becomes upper-bounded with endpoint at $-\theta/\beta$. This condition should be used with caution unless there is physical evidence of upper bounding.

The estimated parameters of the GPD can characterize the mean value of the excesses. There are many methods (maximum likelihood, (ML), Coles, 2001; Reiss and Thomas, 2001; moments (MM), probability weighted moments (PWM), generalized probability weighted moments (GPWM), Hosking and Wallis, 1987; and others) to estimate the parameters. The fitting method should be chosen carefully because it may produce upper bound estimations which can be inconsistent with the observed data. For instance, Dupuis (1996) found inconsistencies in the GPD upper bound estimated parameters obtained with the MM and PWM fitting methods. This inconsistency occurs when one or more sampled observations exceed the estimated upper bound. The problem of inconsistency of a fitting method with the observed data requires appropriate attention (Simiu, 1995; Ashkar and Nwentsa Tatsambon, 2007).

2.3.2. The non-homogeneous generalized Pareto–Poisson process

The original development of this characterization is due to Pickands (1971); however, Smith (1989) was the first to convert the model into a tool for inference. The Generalized Pareto–Poisson Process consists of two components (Davison and Smith, 1990; Coles, 2001; Reiss and Thomas, 2001): (i) the occurrences of exceedances of some high threshold u (i.e., $X_i>u$, for some value of i) may be described as a Poisson process (with rate parameter λ_e) and (ii), the excesses over threshold u (i.e., X_i-u, for some i) have a GP distribution (with scale and shape parameters, θ and β). In the case of volcanic eruptions, the magnitude of the eruptions and the time of their occurrence, are viewed as points in a two-dimensional space, which formally is the realization of a point process (Cox and Isham, 1980). The intensity measure of this two-dimensional Poisson process on $B=[t_1,t_2]x(u,\infty)$ with $[t_1,t_2]\subset[0,1]$ is given by

$$\Lambda(B) = (t_2 - t_1)\left[1 - \frac{\beta(x-u)}{\theta}\right]^{1/\beta}, \tag{6}$$

where β, and θ are the parameters of the GPD (Eq. (5)) (Brabson and Palutikof, 2000; Lin, 2003). An important property of the GPD is the threshold stability (Hosking and Wallis, 1987). If $Y=X-u$ is a variable distributed like a GPD with a shape parameter β, it continues distributed like GPD with an identical shape parameter β for any higher truncation value $u+q$. Another related property of the GPD (Davison and Smith, 1990) refers to the mean excess: if $Y=X-u$ is a GP-distributed variable, then the mean excess over threshold u is

$$E(x-u|x>u) = \frac{\theta - \beta u}{1+\beta} \tag{7}$$

for $\beta>-1$, $u>0$ and $\theta-u\beta>0$. This implies that the conditional mean exceedance over a threshold, u, is a linear function of u. Furthermore, $E(x-u|x>u)$ is the mean of excesses of the threshold u, for which the sample mean of the excesses of u provides an empirical estimate. The sample mean excess is the sum of excesses over the threshold u divided by the number of data points which exceed u. The sample mean excess is an empirical estimate of conditional mean exceedances and β and θ of GPD can be determined by slope and intercept of sample mean excess plot. Hence, $E(x-u|x>u)$ is linear in u with slope $\frac{-\beta}{1+\beta}$ and intercept $\frac{\theta}{1+\beta}$. Davison and Smith 1990; Díaz, 2003; Lin, 2003; Beguería, 2005). The real data series at different threshold values can be tested by the mean excess plot, i.e. the plot of the average excess over a threshold against

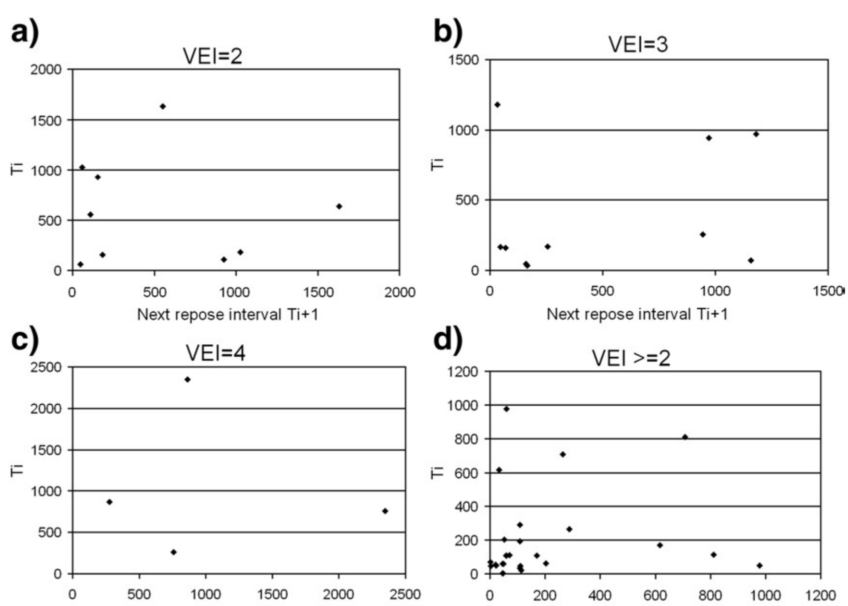

Fig. 2. Serial correlation diagram of successive repose intervals T_i and T_{i+1} (from Table 1, measured in months) between eruptions of Colima volcano with a) VEI = 2, b) VEI = 3, c) VEI = 4 and d) VEI ≥ 2. The highly scattered pattern indicates independent adjacent repose intervals.

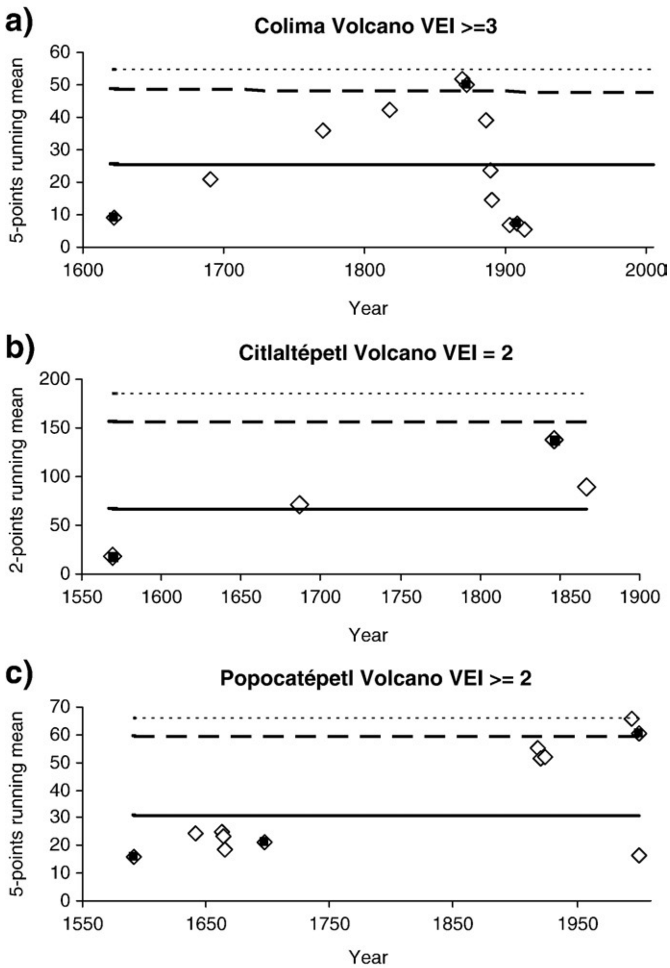

Fig. 3. Moving averages of consecutive reposes of a) Colima, b) Citlaltépetl, and c) Popocatépetl volcanoes. The diamonds represent averages of five (two for Citlaltépetl volcano) consecutive repose plotted at the date of eruption ending the fifth (second for Citlaltépetl volcano) repose interval making every fifth (second for Citlaltépetl volcano) point (filled square) independent (Klein, 1982; De la Cruz-Reyna, 1996). The solid horizontal line is the mean of all the repose periods; the thin dotted line represents the 95% upper confidence level, and the thick dashed line represents the 90% upper confidence level.

the value of the threshold, for a given set of threshold values. The sample mean excess plot is given by:

$$\overline{x}_u = \frac{\sum_{i:x_i>u} (x_i - u)}{N_u} \quad (8)$$

Where N_u is the number of excess x_i over a threshold u (McNeil and Saladin,1997; Martínez, 2003; Lin, 2003).

The mean excess plot is also a diagnostic plot that should be drawn before fitting any model, providing guidance about what threshold to use. The key feature is that if Y is GPD then the mean excess over a threshold u, for any $u>0$, is a linear function (Eq. (7)) of u. Therefore, if the variable follows a GPD over a threshold value u, the mean excess plot should appear approximately linear at those points.

Table 8
Two possible models of the VEI magnitude distributions for the major Holocenic eruptions of Colima volcano

Years BP	Colima 1	Colima 2
2300	5	5
3600	5	5
7040	6	5

Table 9
Two possible models of the VEI magnitude distributions for the major geologic eruptions of Citlaltépetl volcano

Years BP	Citlaltépetl 1	Citlaltépetl 2
4100	4	4
8500	4	4
13,000	4	5

The generalized Pareto–Poisson process may be seen as a limiting form of the joint point process of exceedance times and excess values over the threshold.

3. Applications to active volcanoes

3.1. Case studies and data sets

We present here five case studies for the estimation of volcanic hazard with the proposed method. Colima, Nevado de Toluca, Citlaltépetl, Popocatépetl and El Chichón volcanoes (Fig. 1), are among the most active in México and they represent a significant threat to a large population dwelling in their neighborhoods. Except for the last, all of them are located in the Mexican Volcanic Belt, a Miocene–Quaternary province (Ferrari et al., 1999) that crosses the central part of México. El Chichón is located in the NW end of the Chiapas Volcanic Arc (CVA), which is associated with the subduction of the Cocos plate under the North American plate, but complicated by the geometry of the plate boundary fault system (Damon and Montesinos, 1978; Mora et al., 2007).

Colima volcano (19.512° N, 103.617° W) is the active volcano in México with the highest eruption rate, with a historical record (Table 1) of 41 eruptions in the past 500 years. For the present study, the records (De la Cruz-Reyna, 1993; Bretón et al., 2002) have been completed with recent data published in www.ucol.mx (Observatorio del Volcán de Colima) and www.volcano.si.edu (Global Volcanism Program, Smithsonian Institution).

The ancestral Colima volcano was formed in the late Pleistocene on the southern flank of Nevado de Colima, an older volcano located to the North of the current Colima volcano. About 10,700 B.P., (Cortés et al., 2005), this andesitic volcano rose to a presumed height of 4100 m. During a Bezimyanny–St Helens type eruption, the ancestral Colima volcano collapsed southwards, forming a 5-km-wide horseshoe-shaped caldera, and a massive volcanic debris avalanche deposit. This avalanche blanketed an area of about 1500 km^2, reaching up to 70 km from the former summit. The deposit has a volume estimated in 10 km^3. Soon after this avalanche, the currently active cone of Colima began to grow within the caldera. It is assumed that its mean magma production rate is about 0.3 km^3/1,000 yr, (Luhr and Carmichael, 1990). Table 2 lists the major holocenic eruptions of Colima volcano.

Citlaltépetl or Pico de Orizaba volcano (19.03°N, 92.27°W) is an ice-capped, andesitic, currently dormant, active stratovolcano. With an elevation of 5675 m a.s.l., it is one of the highest active volcanoes in North America. The Citlaltépetl record of historical activity, its high relief,

Table 10
Four possible models of the VEI magnitude distributions for the major geologic eruptions of Popocatépetl volcano

Years BP	Popocatépetl 1	Popocatépetl 2	Popocatépetl 3	Popocatépetl 4
1100	4	4	4	4
1700	4	4	4	4
2150	4	4	4	4
5000	4	4	4	4
7100	4	4	4	4
9100	4	4	4	4
10,700	4	4	5	5
14,000	4	5	5	6
23,000	5	5	5	5

Table 11
Three possible models of the VEI magnitude distributions for the historical and geologic eruptions of El Chichón volcano

Years BP	Reported VEI	Chichón 1 VEI	Chichón 2 VEI	Chichón 3 VEI
25	5	5	5	5
550	4	4	4	4
900	3	3	3	3
1250	4	4	4	4
1500	3	3	3	3
1600		3	3	3
1900		3	3	4
2000	3	3	3	3
2500	3	3	3	3
3100		4	3	5
3700	4	4	4	4

and the widespread distribution of its products are clear indications that it is a potentially high-risk volcano. In addition to this, over 750,000 people live within a radius of 40 km. As in many other cases, population growth around the volcano resulted from fertility of soils and a relatively low eruption rate. The evolution of Citlaltépetl can be divided into three distinct constructional stages (Carrasco-Núñez, 2000): (1) Growth of an ancestral large stratovolcano (Torrecillas cone), resulting from a predominantly effusive phase; (2) Construction of a superimposed cone (Espolón de Oro), and extrusion of various peripheral silicic domes; (3) Building of the present cone (Citlaltépetl) and emplacement of several overlapping dacitic domes. The historical and geological activity of Citlaltépetl is summarized in Table 3 (De la Cruz-Reyna and Carrasco-Núñez, 2002).

Popocatépetl volcano (19.02°N, 98.62°W) is located within a densely populated region about 70 km southeast of downtown Mexico City and 40 km west of the city of Puebla, with over 20 million people vulnerable to direct hazards associated with a major explosive eruption. Because of the large exposed population, this volcano is arguably the most risky in the country. This 5454 m a.s.l. volcano geologic past clearly indicates that it is capable of producing major catastrophic eruptions: three Plinian events have occurred within the past 5000 yr B.P., well within the period of human settlement in central Mexico (Siebe et al., 1996; Siebe and Macías, 2004). The modern cone of Popocatépetl consists of interlayered andesitic to dacitic lava flows and pyroclastic deposits (Robin, 1984), all erupted in the last 23,000 years following a cataclysmic eruption that included edifice collapse and debris avalanche emplacement (Siebe and Macías, 2004). Fortunately, the current eruptive episode – beginning in December 1994 after being dormant for nearly seven decades – has consisted to date of relatively minor activity which has characterized Popocatépetl's activity since the 14th century (De la Cruz-Reyna et al., 1995). The historical activity (De la Cruz-Reyna and Tilling, 2008) is shown in Table 4, and the geological record is summarized in Table 5 (Siebe and Macías, 2004).

Nevado de Toluca volcano (19.108°N, 99.758°W) is the fourth highest peak in Mexico with an altitude of 4680 m a.s.l (Fig. 1). It is located 21 and 80 km southwest of the cities of Toluca and Mexico, respectively, in the central sector of the Trans-Mexican volcanic belt. The volcano, which has no historical eruptions, has a complex geological record. Studies on the stratigraphy (Macías et al., 1997; Arce, 1999; Cervantes, 2001), the flank collapses (Capra and Macías, 2000), the structural geology (García-Palomo et al., 2000), and the overall geology (García-Palomo et al., 2002) have permitted to reconstruct an eruptive history, of which five major geological events listed in Table 6 have been widely recognized. The volcano has been quiet in the last 3.3 ka, but, as shown by their stratigraphic record, Plinian eruptions may occur after long periods of inactivity. Should this happen, it would represent a significant hazard at least for the entire Lerma basin, where more than 1 million people live (Capra et al., 2006).

El Chichón volcano (17.36°N, 92.23°W) has an elevation of 1100 m a.s.l. The volcano has in its summit a 1 km wide crater with gently dipping outward slopes. Its average depth is 140 m, and contains a central lake. The 1982 eruption of El Chichón volcano in southeastern Mexico had a strong social and environmental impact (Luhr 1991; Rampino and Self 1984; Schönwiese 1988). Recent studies on the stratigraphy of the volcano and new radiocarbon ages show that at least 11 eruptions prior to the 1892 eruption have occurred at El Chichón in the past 8000 years. Explosive events, most of them producing block-and-ash flow and surge deposits, occurred around 550, 900, 1250, 1500, 600, 1900, 2000, 2500, 3100, 3700 and 7700 yr BP. (Espíndola et al., 2000, Macías et al., 2007). Table 7 summarizes the historical and geological records.

3.2. Analysis of historical eruption series

3.2.1. Independence of repose interval durations between historical eruptions

In each of the cases, the independence of the repose intervals between eruptions of VEI categories high enough to assure completeness of the catalogs has been tested, by means of serial correlation scatterplots, in which the duration of each interval T_i+1 between two successive eruptions for each VEI magnitude category is plotted against the previous repose interval durations T_i. Diagrams showing large dispersion, a relatively high concentration of points near the axes, and very low correlation coefficients may be characterized as random process of independent events (Cox and Lewis, 1966).

For Colima volcano, the repose times in months between successive historical eruptions calculated from Table 1 are shown in Fig. 2. The correlation coefficients between set T_i and set T_i+1 for each VEI category 2, 3, 4, and ≥2 are: −0.096, 0.256, −0.156, and 0.152, respectively, indicating a very low serial correlation.

Similar scatterplots were calculated for historical data (complete portion) of the other volcanoes. The scatter diagrams also show large dispersions and higher concentrations of points near the axes. The correlation coefficient between the sets T_i and T_i+1 for Citlaltépetl volcano for VEI 2 is 0.00165 and for Popocatépetl volcano for VEI 2 and VEI≥2 categories is −0.174 and −0.097, respectively.

In all cases no significant correlation was found and thus independence may be assumed.

3.2.2. Test for stationarity of historical eruption sequences

To verify a possible non-stationary character of the eruption histories, and test for the existence of regimes with significantly different eruption rates, we performed a moving average test (Klein, 1982) on each of the historical series. Fig. 3a plots the moving averages of five consecutive reposes of Colima for the VEI≥3 eruptions. One point is plotted for each possible sequence of five successive reposes at the date of the last eruption of each sequence. The overall mean repose of 25.24 years is represented by the solid horizontal line. The probability

Table 12
Estimated magnitudes of the geologic eruptions of Colima, Popocatépetl, Citlaltépetl and El Chichón volcanoes based on the power law (2)

Colima 1 Year BP	VEI	Popocatépetl 2 Year BP	VEI	Citlaltépetl 1 Year BP	VEI	El Chichón 2 Year BP	VEI
2300	5	1100	4	4100	4	25	5
3600	5	1700	4	8500–9000	4	550	4
7040	6	2150	4	13,000	4	900	3
		5000	4			1250	4
		7100	4			1500	3
		9100	4			1600	4
		10,700	4			1900	3
		14,000	5			2000	3
		23,000	5			2500	3
						1900	3
						2000	3
						2500	3

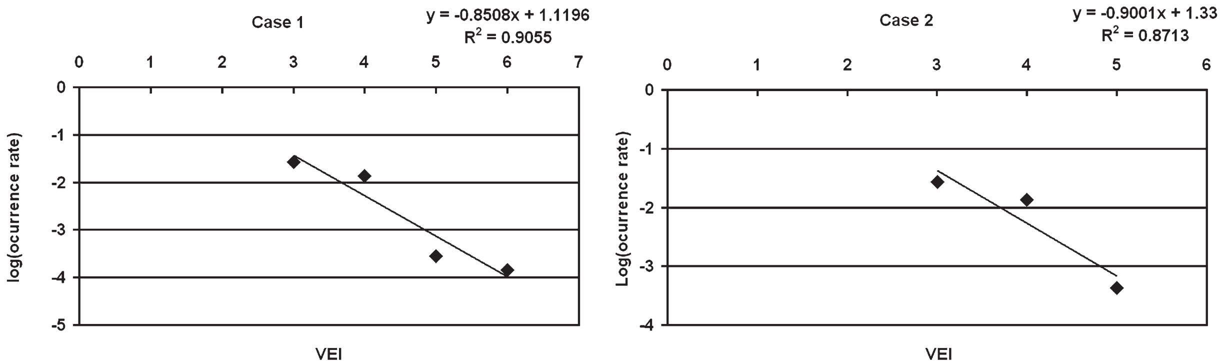

Fig. 4. Best linear fits of Eq. (2) for all the geological activity models of Colima volcano listed in Table 8. The lower magnitude eruption rates correspond to the historical record (Table 1). The highest regression coefficient corresponds to the case 1 (model "Colima 1").

that a given set of successive repose intervals departs from a certain amount from the mean by chance is represented by the thin dotted line (95% confidence level) and the thick (90% confidence level) dashed line. If a certain average departs from the average more than the amount indicated by the broken lines, one may say that the eruption sequence is non-stationary for the corresponding confidence level.

Error limits were computed from the combined chi-square and binomial distributions (Klein, 1982). Fig. 3a shows a weak non-stationary behavior for Colima volcano, because one of the points exceeds the 90% level, but not the 95% confidence level. Therefore, we cannot reject the hypothesis that Colima may have a non-stationary behavior, with at least two alternating eruption regimes (De la Cruz-Reyna, 1996).

A similar analysis of the sequence of available historical volcanic eruptions for Citlaltépetl volcano shows no evidence that the eruptive process for the VEI=2 magnitude eruptions may be non-stationary (Fig. 3b). On the other hand the analysis of Popocatépetl volcano for the VEI≥2 eruptions shows a weakly non-stationary eruptive series similar to the case of Colima volcano (Fig. 3c).

3.3. Analysis eruption series

The available information, from historical records (Tables 1, 3 and 4), and from the deposits of major eruptions at Colima (Table 2), Citlaltépetl (Table 3), Popocatépetl (Table 5), Nevado de Toluca (Table 6) and El Chichón (Table 7) volcanoes is not sufficient to assign precise VEI values to all the eruptions, although some constrictions may be set on their relative sizes. Tables 8–11 show sets of likely values (models) of the VEI of those eruptions. For Nevado de Toluca, we use a single model based on only five eruptions occurred between 28,000 and 10,500 yr B.P., using published VEI data, or estimating them from erupted volume values reported in the references cited in Table 6. In the other cases, the VEI's of eruptions in which no volume or intensity data were available were estimated probing the best fit to the VEI values of other eruptions based on the power law described by Eq. (2).

Fig. 4, illustrates the loglinear relationship between the VEI magnitudes and occurrence rates from Eq. (2) for the Colima volcano eruptive history models. The regression coefficients indicate that the best fit is for the case "Colima 1" of Table 8. Applying the same procedure to Popocatépetl, Citlaltépetl, and El Chichón, we conclude that the best estimations of magnitudes for the geologic records are, "Citlaltépetl 1", "Chichón 2" and "Popocatépetl 2" (Table 12).

3.3.1. Analysis of repose-time series

In this part of the study, we attempt to find the survival Weibull functions (Eq. (3)) that best fit the repose-time distributions of the volcanoes referred in Section 3.1. To do this we only use the VEI

Table 13
The Weibull distribution parameters for the indicated volcanoes

VEI	Shape parameter	Scale parameter
Colima		
2	0.52	3.03
3	0.85	3.89
4	0.41	2.48
>2	0.37	1.02
Citlaltépetl		
2	0.7	4.9
Popocatépetl		
2	1.01	2.75
3	1.47	19.26
Nevado de Toluca		
>3	1	1.44
El Chichón		
>2	1.19	3.97

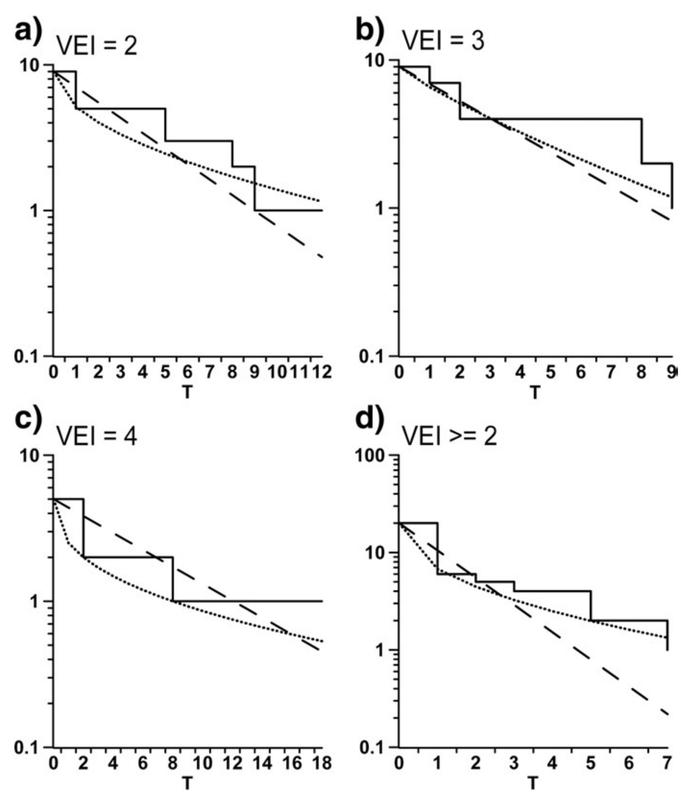

Fig. 5. Distribution of observed repose intervals for a) VEI = 2, b) VEI = 3, c) VEI = 4, and d) VEI≥2 with duration greater than *T* decades (steps) for eruptions at Colima volcano in the period 1560 to present. The survival Weibull distribution (dotted line) better fits the data than the exponential distribution (dashed line).

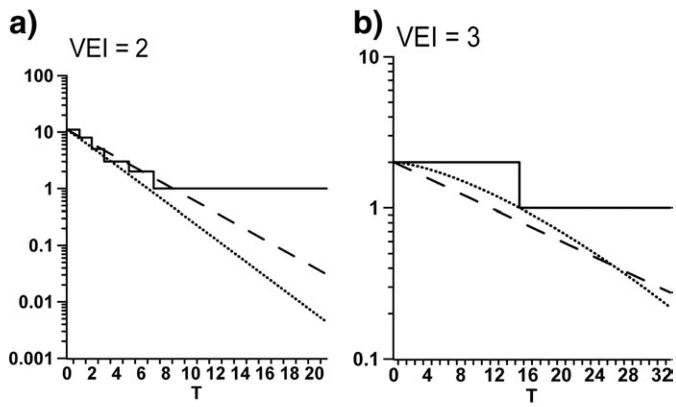

Fig. 6. Distribution of observed repose intervals with duration greater than *T* decades (steps) for a) VEI 2 and b) VEI 3 eruptions at Popocatépetl volcano in the period 1510 to present. The survival Weibull distribution (dotted line) shows a slightly better fitting than the exponential distribution (dashed line).

categories that made possible to consider the eruptive series as complete. Such eruptive series are the historical data listed in Tables 1, 3 and 4, and the geological series of Nevado de Toluca and El Chichón. Although In the last two cases it is difficult to sustain completeness, we are including the portions of Table 6 in which date and VEI data have been published for Nevado de Toluca, and El Chichón 2 model of Table 11 as an example of the Weibull representation of available data. The resulting Weibull distribution parameters are summarized in Table 13.

The comparison between the exponential and Weibull distributions is shown in Figs. 5–9. In various cases, the Weibull survival function provided better fits to the repose-time data than the exponential function, because the shape parameter accounts for the non-stationary character of some of the series. Stationary repose-time series may be equally well described by both distributions.

3.4. Assessment of volcanic hazard from geological and historical eruption series.

The volcanic hazards for Colima, Nevado de Toluca, Popocatépetl, Citlaltépetl and El Chichón, volcanoes were estimated using the *non-homogeneous generalized Pareto–Poisson* process described in a previous section. We use the number of excesses (Eq. (4)) inferred from

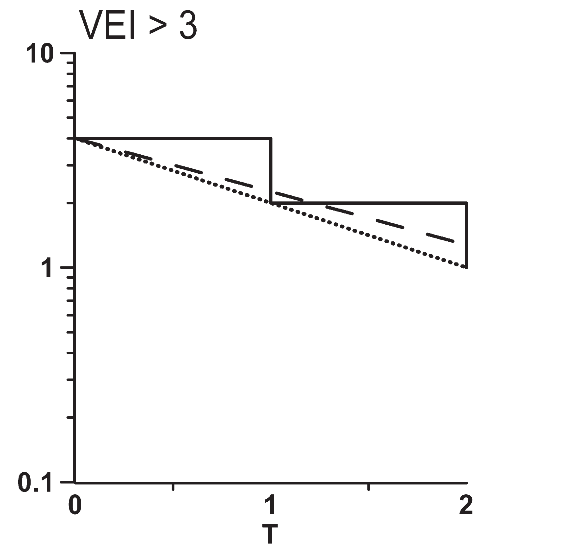

Fig. 8. Distribution of observed repose intervals (steps) with duration greater than *T* (in units of 4,000 years) for eruptions at Nevado de Toluca volcano in the last 28,000 years. The survival Weibull distribution (dotted line) fits the data better than the exponential distribution (dashed line).

the eruptive rates (Eq. (2)) of the geological and historical data (Tables 1, 3, 4, 6 and 12) to calculate the probabilities of occurrence of eruptions in the different magnitude classes.

First we use a graphical method to estimate the parameters from the linear regression (using Eq. (7)) of the plot of the mean of the excesses (obtained with Eq. (8)) vs their thresholds (Davison and Smith, 1990).

The linear fittings of the means of the excesses and the means of the exceedances vs their thresholds, obtained as described in Section 2.3, are illustrated in Fig. 10. The good fittings of the mean exceedances and the fair fittings of the mean excesses indicates that the method is satisfactory. The problem of the fair fitting of the mean excesses may be addressed considering not one, but two regression lines, one for the lower threshold values, and other for threshold 4 and above, and calculating the NHGPPP parameters for each line. However, in this case we have used the single mean excess lines for the

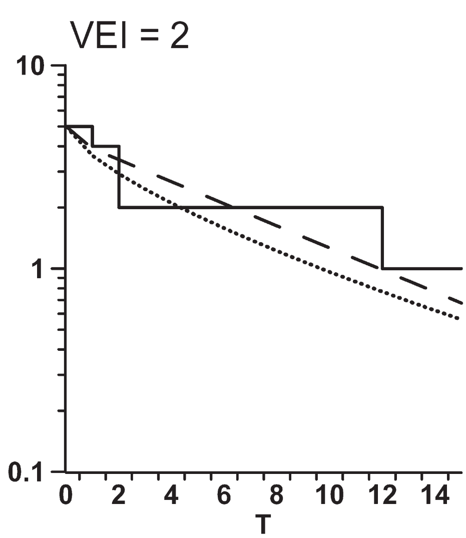

Fig. 7. Distribution of observed repose intervals with duration greater than *T* decades (steps) for VEI = 2 eruptions at Citlaltépetl volcano in the period 1530 to present. The survival Weibull distribution (dotted line) and the exponential distribution (dashed line), show similar degrees of fitting.

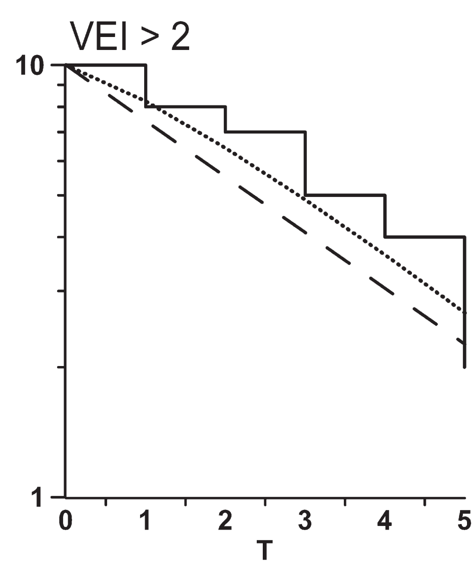

Fig. 9. Distribution of observed repose intervals (steps) with duration greater than *T* (in units of 100 years) for eruptions at El Chichón volcano in the last 3700 years. The survival Weibull distribution (dotted line) fits the data better than the exponential distribution (dashed line).

probability calculations. The Pareto generalized parameters for each volcano are summarized in Table 14.

We used Eq. (6) to calculate the intensity of the NHGPPP and to obtain the probabilities of eruption occurrences. Since the approach of exceedances do implicitly assume that the scale measuring the phenomena is open, and considering that the VEI scale ends in 8, we subtracted the probabilities of eruptions exceeding that magnitude from the probabilities of exceeding VEI's lower than 8. Table 15 and Fig. 11 show the probabilities of at least one eruption exceeding a given VEI occurring in the stipulated time intervals for each activity model. We also compare the results obtained with the NHGPPP with volcanic hazard estimates obtained from direct application of the Binomial and simple Poisson distributions for the same eruption series.

Inspection of these results shows that the probabilities of occurrence of eruptions in the lower magnitude classes, calculated with the method proposed here differ very little from the standard Binomial–Poisson methods. However, the probabilities of occurrence of eruptions exceeding moderate magnitudes are significantly increased. This difference arises from the added information that the GPD (Eq. (5)) introduces in the NHGPPP when the estimated eruption rates of large-magnitude eruptions are introduced.

4. Discussion and conclusions

The relatively simple methodology proposed in this paper allows the use of historical and geological eruption time series to obtain more precise estimates of the volcanic hazard. The method considers the limitations inherent to each of those series: short sample time, probable absence of large events and incomplete reporting of very small magnitudes for the historical series; incomplete reporting of small and intermediate magnitudes and uncertainties in the age and magnitude of major eruptions for the geological series. It also considers the

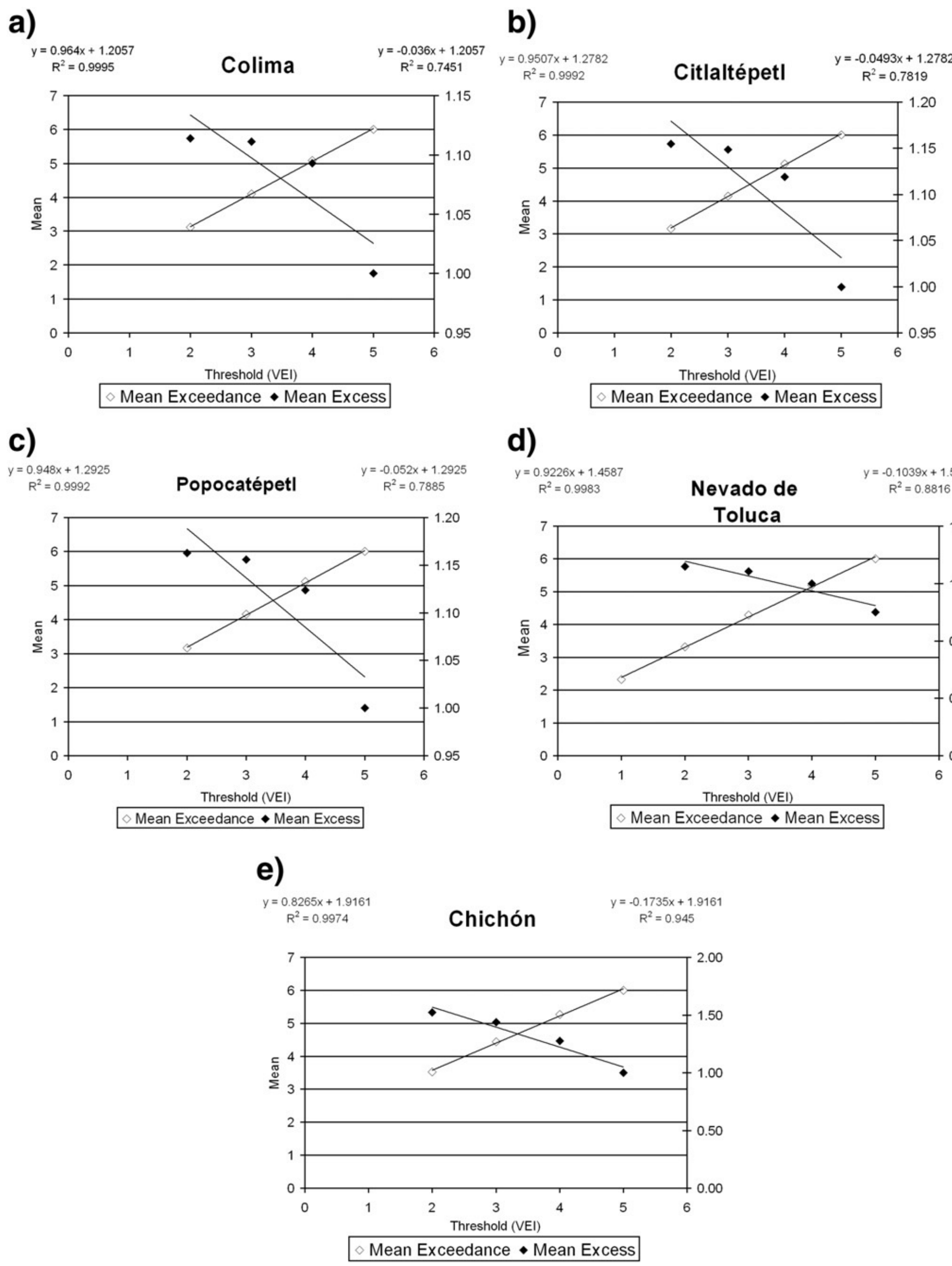

Fig. 10. Plots of exceedance and excess mean vs. threshold for a) Colima, b) Citlaltépetl, c) Popocatépetl, d) Nevado de Toluca and e) El Chichón volcanoes.

possibility of time-dependent eruption rates (non-stationary eruption series). These problems are addressed assuming a characteristic behavior of most natural phenomena: the inverse relationship between occurrence rate and magnitude. (De la Cruz-Reyna, 1991, 1996; De la Cruz- Reyna and Carrasco-Núñez, 2002). The use of relation (Eq. (2)), describing this behavior, permits linking the historical and geological time series, greatly expanding the data base, both in duration and eruption magnitude range.

If an eruptive series is non-stationary, the hazard estimate depends on the time it is made. For instance, the hazard evaluation of a volcano which shows a succession of high and low occurrence rate regimes may be inaccurate if only the current regime is taken into account. Thus determining the degree of time dependence of the occurrence rate is an essential step. The method proposed here provides a criterion to determine such a condition, and how to deal with it. The degree and nature of the non-stationarity is revealed by the Weibull analysis on the repose-time distribution between successive eruptions. The adjustable shape parameter of this distribution can describe variable eruptive rates, allowing a better description of the time-dependent distribution of reposes.

To deal with the difficulties derived from the lack of catalogue completeness of very low and very high eruption magnitudes in the historical series, and of low and intermediate magnitudes in the geological series, the hazard estimates for the full series is done linking both time scales using the power law distribution (Eq. (2)). Although the validity of such relationship for groups of volcanoes has proven effective (De la Cruz-Reyna, 1991), its use on individual volcanoes may be questioned (Marzocchi and Zaccarelli, 2006), particularly in the case of strongly time-dependent eruptive series, its applications to individual volcanoes showing a stationary or a quasi-stationary behavior using mean eruption rates calculated over periods long enough to include the representative rate variations, render satisfactory results as shown in the examples presented here. The method also allows testing different models of geologic past behavior, when the uncertainties of the date and magnitude of older eruptions are high, and search for the most likely combination of date-magnitude that is consistent with the more recent and reliable data, even for weakly time-dependent (quasi-stationary) series. In this respect, no problem raises form the historical and the geological records having overlapping VEI categories. For example, in the case of El Chichón we used the VEI 4 data in both subsets, with good fittings in both the mean exceedances and the mean excesses vs their thresholds.

Once a representative eruption rate has been determined, the volcanic hazard estimations based on extreme values are obtained using a NHGPPP. This method renders eruption probabilities of exceedance for each VEI magnitude category, i.e. hazard estimates that takes into account all the above mentioned factors, since it gives the appropriate weights to the more reliable (yet incomplete) historical data, and to the scarce large-magnitude geological data.

In general terms, the application of a NHGPPP in the final stage of the method emphasizes the effect of large magnitudes in the hazard estimation.

The application of this method to the eruption sequences of the Popocatépetl, Citlaltépetl and Colima volcanoes was compared with published results of other hazard estimates (De la Cruz-Reyna, 1993; De la Cruz-Reyna and Carrasco-Núñez, 2002; De la Cruz-Reyna and Tilling, 2008). Although the results were similar, a characteristic difference when comparing with the Poisson and Binomial distribution based hazard estimates was that the exceedance probabilities calculated with the NHGPPP were increasingly larger for VEI magnitudes greater than 3.

These results should thus be taken into account in the assessment of volcanic risk and in the design of prevention and response measures, particularly for major eruptions to which larger areas may be 100% vulnerable.

Table 14
The Pareto distribution parameters to calculate the NHGPPP for the indicated volcanoes

	Colima 1	Citlaltépetl 1	Popocatépetl 2	Nevado de Toluca	El Chichón 2
Shape parameter	0.037	0.052	0.055	0.116	0.143
Scale parameter	1.251	1.344	1.363	1.746	2.191

Table 15
Volcanic eruption hazards of Colima, Citlaltépetl, Nevado de Toluca, Popocatépetl and El Chichón volcanoes as probabilities of occurrence of at least one eruption exceeding a VEI magnitude over different time periods, for different models of the past activity. The probabilities were calculated using the NHGPP

Years	Colima 1	Popocatépetl 2	Citlaltépetl 1	Nevado de Toluca	El Chichón 2
VEI>2					
20	0.63290	0.10311	0.03851	0.01487	0.05835
50	0.90840	0.23792	0.09348	0.03675	0.13913
100	0.96840	0.41839	0.17810	0.07211	0.25758
500	0.87936	0.91781	0.62160	0.31095	0.74175
VEI>3					
20	0.35806	0.04982	0.01810	0.00816	0.03611
50	0.66361	0.11980	0.04461	0.02028	0.08759
100	0.86989	0.22482	0.08718	0.04012	0.16666
500	0.87935	0.70881	0.36435	0.18444	0.57378
VEI>4					
20	0.17236	0.02276	0.00812	0.00424	0.02118
50	0.37367	0.05587	0.02018	0.01056	0.05195
100	0.59816	0.10843	0.03992	0.02100	0.10071
500	0.87180	0.43014	0.18338	0.10030	0.39558
VEI>5					
20	0.07391	0.00974	0.00344	0.00204	0.01153
50	0.17328	0.02415	0.00857	0.00509	0.02848
100	0.31210	0.04763	0.01705	0.01015	0.05589
500	0.75173	0.21337	0.08197	0.04955	0.24023
VEI>6					
20	0.02802	0.00375	0.00131	0.00087	0.00554
50	0.06806	0.00934	0.00328	0.00217	0.01375
100	0.12975	0.01857	0.00654	0.00433	0.02718
500	0.44890	0.08817	0.03213	0.02138	0.12381

Acknowledgements

We thank Olivier Jaquet and an anonymous referee for their comments and suggestions that greatly contributed to improve this paper. This work was carried out with the support of CONACYT grant 195487.

References

Arce, J.L., 1999. Reinterpretación de la erupción Pliniana que dio origen a la Pómez Toluca Superior, Volcán Nevado de Toluca [M.S. Thesis]: México, D.F., Universidad Nacional Autónoma de México, 101 pp.
Arce, J.L., Macías, J.L., Vázquez-Selem, L., 2003. The 10.5 ka Plinian eruption of Nevado de Toluca volcano, Mexico: stratigraphy and hazard implications. Geol. Soc. Amer. Bull 115 (2), 230–248.
Arce, J.L., Cervantes, K.E., Macías, J.L., Mora, J.C., 2005. The 12.1 ka Middle Toluca Pumice: a dacitic Plinian–subplinian eruption of Nevado de Toluca volcano in Central Mexico. J. Volcanol. Geotherm. Res. 147, 125–143.
Arce, J.L., Macias, J.L., Gardner, J.E., Layer, P.W., 2006. A 2.5 ka History of dacitic magmatism at Nevado de Toluca, Mexico: petrological, 40Ar/39Ar dating, and experimental constraints on petrogenesis. J. Petrol. 47 (3), 457–479.
Ashkar, F., Nwentsa Tatsambon, C., 2007. Revisiting some estimation methods for the generalized Pareto distribution. J. Hydrol. 346, 136–143.
Aspinall, W.P., Carniel, R., Jaquet, O., Woo, G., Hincks, T., 2006. Using hidden multi-state Markov models with multi-parameter volcanic data to provide empirical evidence for alert level decision-support. J. Volcanol. Geotherm. Res. 153, 112–124.
Bebbington, M.S., 2007. Identifying volcanic regimes using hidden Markov models. Geophys. J. Int. 171, 921–942.
Bebbi`ngton, M.S., Lai, C.D., 1996a. Statistical analysis of New Zealand volcanic occurrence data. J. Volcanol. Geotherm. Res. 74, 101–110.
Bebbington, M.S., Lai, C.D., 1996b. On nonhomogeneous models for volcanic eruptions. Math. Geol. 28, 585–600.

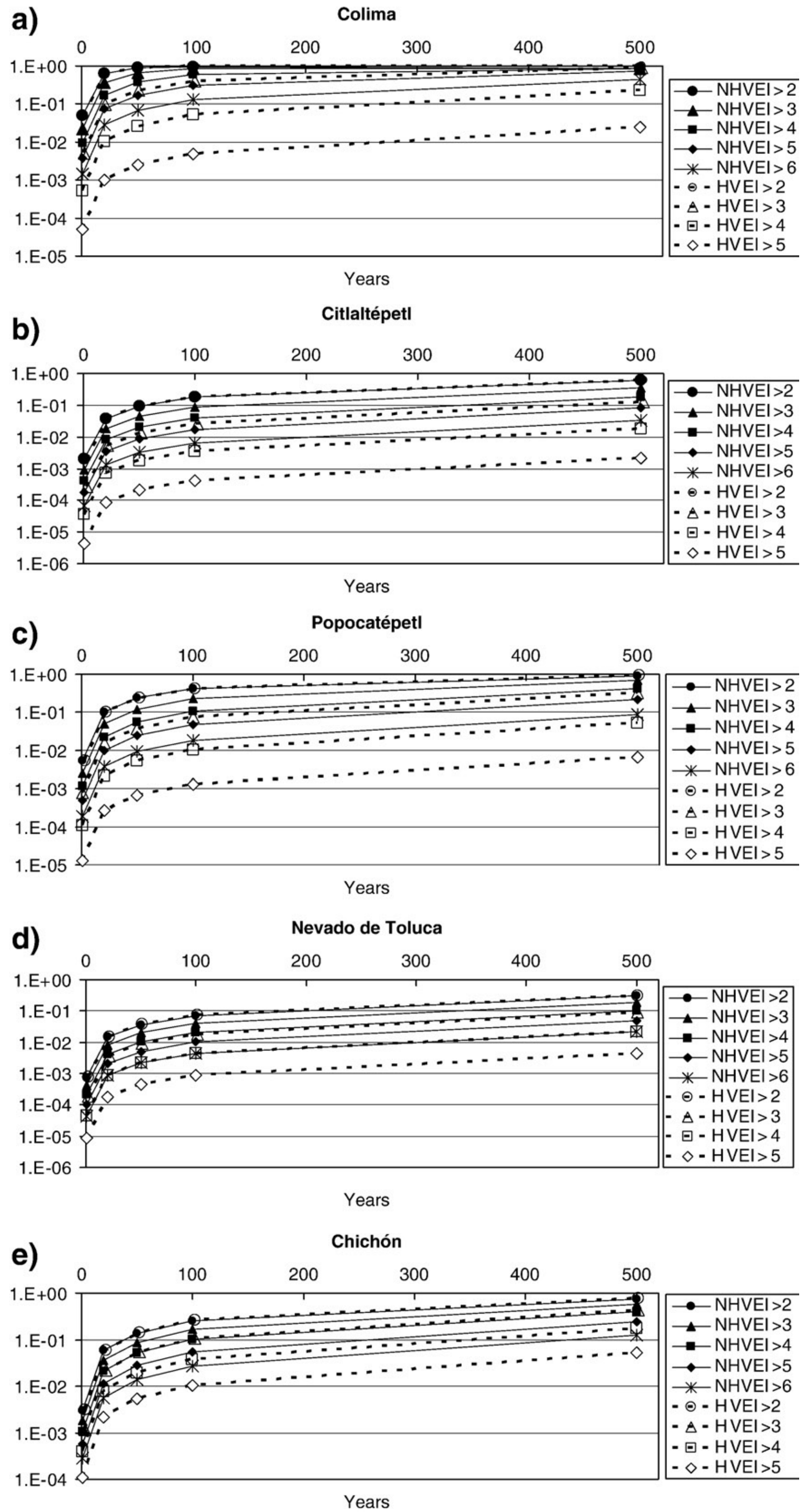

Fig. 11. Probabilities calculated with NHGPPP (filled symbols marked NH) and Homogeneous Poisson distribution (open symbols marked H) of at least one eruption, with a VEI magnitude greater than a given VEI threshold for a) "Colima 1", b) "Citlaltépetl 1" c) "Popocatépetl 2", d) Nevado de Toluca and e) "El Chichón 2" eruptive series and models.

Beguería, S., 2005. Uncertainties in partial duration series modelling of extremes related of the choice of the threshold value. J. Hydrol. 303, 215–230.

Brabson, B.B., Palutikof, J.P., 2000. Test of the generalized Pareto distribution for predicting extreme wind speeds. J. Appl. Meteorol. 39, 1627–1640.

Bretón, M., Ramírez, J., Navarro, C., 2002. Summary of the historical eruptive activity of Volcán de Colima, Mexico 1519–2000. J. Volcanol. Geotherm. Res. 117, 21–46.

Burt, M.L., Wadge, G., Scott, W.A., 1994. Simple stochastic modelling of the eruption history of a basaltic volcano: Nyamuragira, Zaire. Bull. Volcanol. 56, 87–156.

Capra, L., Macías, J.L., 2000. Pleistocene cohesive debris flows at Nevado de Toluca volcano, central Mexico. J. Volcanol. Geotherm. Res. 102, 149–167.

Capra, L., Carreras, L., Arce, J.L., Macías, J.L., 2006. The lower Toluca pumice: a 21,700 yr B.P. Plinian eruption of the Nevado de Toluca volcano, Mexico. Geological Society of America Special Volume "Neogene–Quaternary Continental Margin Volcanism: A Perspective from Mexico", vol. 42, pp. 155–174.

Carrasco-Núñez, G., 2000. Structure and proximal stratigraphy of Citlaltépetl volcano (Pico de Orizaba), México. Cenozoic tectonic and volcanism in Mexico. Geol. Soc. Amer. Bull. 334, 247–262.

Carta, S., Figari, R., Sartoris, G., Sassi, E., Scandone, R., 1981. A statistical model for Vesuvius and its volcanological implications. Bull. Volcanol. 44, 129–151.

Cervantes, K., 2001. La Pómez Blanca Intermedia: Depósito producido por una erupción Pliniana-subPliniana del Volcán Nevado de Toluca hace 12,100 años. [M.S. Thesis]: México, D.F., Universidad Nacional Autónoma de México, 86 pp.

Coles, S., 2001. An Introduction to Statistical Modeling of Extreme Values. Springer-Verlang, London. 224 pp.

Cortés, A., Garduño-Monroy, V.H., Navarro-Ochoa, C., Komorowski, J.C., Saucedo, R., Macías, J.L., Gavilanes, J.C., 2005. Carta Geológica del Complejo Volcánico de Colima con Geología del Complejo Volcánico de Colima. Instituto de Geología UNAM. México D.F. Cartas Geológicas y mineras 10.

Cox, D.R., Lewis, P.A.W., 1966. The Statistical Analysis of Series of Events. Methuen and Co., London. 285 pp.

Cox, D.R., Isham, V., 1980. Point Process. Chapman and Hall, London. 188 pp.

Damon, P., Montesinos, E., 1978. Late Cenozoic volcanism and metallogenesis over an active Benioff Zone in Chiapas, Mexico. Ariz. Geol. Soc. Dig. 11, 155–168.

D'Antonio, M. 2008. Reconstrucción de los Eventos Eruptivos de hace 28 y 13 KA asociados al emplazamiento de flujos de bloques y ceniza en el volcán Nevado de Toluca (México). [PhD. Thesis]: México, D.F., Universidad Nacional Autónoma de México.

D'Antonio, M., Capra, L., Sarocchi, D., Bellotti, F., 2008. Reconstrucción del evento eruptivo asociado al emplazamiento del flujo piroclástico El Refugio hace 13 ka, volcán Nevado de Toluca (México). Rev. Mex. Cienc. Geol. 25 (1), 115–134.

Davison, A.C., Smith, R.L., 1990. Models for exceedances over high thresholds. J. R. Stat. Soc. 52 (B), 393–442.

De la Cruz-Reyna, S., 1991. Poisson-distributed patterns of explosive activity. Bull. Volcanol. 54, 57–67.

De la Cruz-Reyna, S., 1993. Random patterns of occurrence of explosive eruptions at Colima Volcano, México. J. Volcanol. Geotherm. Res. 55, 51–68.

De la Cruz-Reyna, S., 1996. Long-term probabilistic analysis of future explosive eruptions. In: Scarpa, R., Tilling, R.I. (Eds.), Monitoring and Mitigation of Volcanic Hazards. Springer, Berlín. 599–629 pp.

De la Cruz Reyna, S., Carrasco Núñez, G., 2002. Probabilistic hazard analysis of Citlaltépetl (Pico de Orizaba) Volcano, Eastern Mexican volcanic belt. J. Volcanol. Geotherm. Res. 113, 307–318.

De la Cruz Reyna, S., Tilling, R., 2008. Scientific and public responses to the ongoing volcanic crisis at Popocatépetl Volcano, México: importance of an effective hazards warning system. J. Volcanol. Geotherm. Res. 170, 121–134.

De la Cruz-Reyna, S., Quezada, J.L., Peña, C., Zepeda, O., Sánchez, T., 1995. Historia de la actividad del Popocatépetl (1354–1995). En: Volcán Popocatépetl, Estudios Realizados Durante la Crisis de 1994–1995. CENAPRED-UNAM, México, D.F. 3–22.

Díaz, A., 2003. Teoría de Valores Extremos para Sucesiones de Variables Aleatorias Dependientes. [Actuary. Thesis]: México, D.F., Facultad de Ciencias. Universidad Nacional Autónoma de México.

Dupuis, D.J., 1996. Estimating the probability of obtaining nonfeasible parameter estimates of the generalized Pareto distribution. J. Stat. Comput. Simul. 54, 197–209.

Espíndola, J.M., Macías, J.L., Tilling, R.I., Sheridan, M.F., 2000. Volcanic history of El Chichón Volcano (Chiapas, Mexico) during the Holocene, and its impact on human activity. Bull. Volcanol. 62, 90–104.

Ferráes, G.S., 2003. Probabilistic prediction of the next large earthquake in the Michoacán Fault — segment of the Mexican subduction zone. Geofís. Int. 42, 69–81.

Ferrari, L., López-Martínez, M., Aguirre-Díaz, G., Carrasco- Núñez, G., 1999. Space–time patterns of Cenozoic arc volcanism in Central Mexico: from Sierra Madre Occidental to the Mexican volcanic belt. Geology 27, 303–306.

Fisher, R.A., Tippett, L.H.C., 1928. Limiting forms of the frequency distributions of the largest or smallest member of a sample. Proc. Camb. Philos. Soc. 24, 180–190.

Fournier d'Albe, E.M., 1979. Objectives of volcanic monitoring and prediction. J. Geol. Soc. Lond. 136, 321–326.

García-Palomo, A., Macías, J.L., Garduño, V.H., 2000. Miocene to Holocene structural evolution of the Nevado de Toluca volcano region, central Mexico. Tectonophysics 318, 281–302.

García-Palomo, A., Macías, J.L., Arce, J.L., Capra, L., Garduño, V.H., Espíndola, J.M., 2002. Geology of Nevado de Toluca Volcano and surrounding areas, central Mexico. Geological Society of America Map and Chart Series MCH089.

Gnedenko, B.V., 1943. Sur la Distribution Limite du Terme d'une Serie Aleatoire. Ann. Math. 44, 423–453.

Gumbel, E.J., 1958. Statistics of Extremes. Columbia University Press, New York. 375 pp.

Gusev, A.A., Ponomareva, V.V., Braitseva, O.A., Melekestsev, I.V., Sulerzhitsky, L.D., 2003. Great explosive eruptions on Kamchatka during the last 10,000 years: self-similar irregularity of the output of volcanic products. J. Geophys. Res. 108 (B2), 2126. doi:10.1029/2001JB000312.

Gutenberg, R., Richter, C.F., 1944. Frequency of earthquakes in California. Bull. Seismol. Soc. Am. 34, 185–188.

Ho, C.-H., 1990. Bayesian analysis of volcanic eruptions. J. Volcanol. Geotherm. Res. 43, 91–98.

Ho, C.-H., 1991a. Time trend analysis of basaltic volcanism for the Yucca Mountain site. J. Volcanol. Geotherm. Res. 46, 61–72.

Ho, C.-H., 1991b. Nonhomogeneous Poisson model for volcanic eruptions. Math. Geol. 23, 167–173.

Ho, C.H., 1995. Sensitivity in volcanic hazard assessment for the Yucca Mountain high-level nuclear waste repository site: the model and the data. Math. Geol. 27, 239–258.

Ho, C.H., Smith, E.I., Keenan, D.L., 2006. Hazard area and probability of volcanic disruption of the proposed high-level radioactive waste repository at Yucca Mountain, Nevada, USA. Bull. Volcanol. 69, 117–123.

Hosking, J.R.M., Wallis, J.R., 1987. Parameter and quantile estimation for the generalized Pareto distribution. Technometrics 29 (3), 333–349.

Ishimoto, M., Iida, K., 1939. Observations of earthquakes registered with the micro-seismograph constructed recently. Bull. Earthq. Res. Inst. Univ. Tokyo 17, 443–478.

Jaquet, O., Carniel, R., 2006. Estimation of volcanic hazard using geostatistical models. In: Mader, H.M., Coles, S.G., Connor, C.B., Connor, L.J. (Eds.), Statistics in Volcanology. IAVCEI Publications n.1. Geological Society, London, pp. 89–103.

Jaquet, O., Löw, S., Martinelli, B., Dietrich, V., Gilby, D., 2000. Estimation of volcanic hazards based on Cox stochastic processes. Phys. Chem. Earth 25, 571–579.

Johnson, N.L., 1966. Cumulative sum control charts and the Weibull distribution. Technometrics 8, 481–491.

Johnson, N.L., Kotz, S., 1953. Continuous Univariate Distributions — 1. Houghton Mifflin Co., Boston. 300pp.

Klein, F.W., 1982. Patterns of historical eruptions at Hawaiian volcanoes. J. Volcanol. Geotherm. Res. 12, 1–35.

Lang, M., Ouarda, T.B.M.J., Bobée, B., 1999. Towards operational guidelines for over-threshold modeling. J. Hydrol. 225, 103–117.

Lin, X.G., 2003. Statistical modelling of severe wind gust. International Congress on Modelling and Simulation, Townsville, 14–17 July, vol. 2, pp. 620–625.

Luhr, J.F., 1991. Volcanic shade causes cooling. Nature 354, 104–105.

Luhr, J.F., Carmichael, I.S.E., 1990. Geology of Volcán de Colima. Boletín Instituto de Geología UNAM, México D.F. 102 pp.

Macías, J.L., García-Palomo, A., Arce, J.L., Siebe, C., Espíndola, J.M., Komorowski, J.C., Scott, K.M., 1997. Late Pleistocene–Holocene cataclysmic eruptions at Nevado de Toluca and Jocotitlán volcanoes, central Mexico. In: Kowallis, B.J. (Ed.), Proterozoic to Recent Stratigraphy, Tectonics, and Volcanology, Utah, Nevada, Southern Idaho and Central Mexico. Brigham Young University Geology Studies, pp. 493–528.

Macías, J.L., Arce, J.L., Scolamacchia, T., Rouwet, D., García, A., Espíndola, J.M., Taran, Y., 2007. El Chichón volcano. 25th Anniversary. A Commemorative Conference. Field Trip Guide. 25 pp.

Martínez, P.A., 2003. Un Modelo Estadístico para Valores Extremos en Series de datos Actuariales. [Actuary. Thesis]: México, D.F., Facultad de Ciencias. Universidad Nacional Autónoma de México.

Marzocchi, W., Zaccarelli, L., 2006. A quantitative model for the time-size distribution of eruptions. J. Geophys. Res. 111, B04204. doi:10.1029/2005JB003709.

Marzocchi, W., Sandri, L., Selva, J., 2008. BET_EF: a probabilistic tool for long- and short-term eruption forecasting. Bull. Volcanol. 70, 623–632.

Marzocchi, W., Sandri, L., Gasparini, P., Newhall, C., Boschi, E., 2004. Quantifying probabilities of volcanic events: the example of volcanic hazards at Mount Vesuvius. J. Geophys. Res. 109, B11201. doi:10.1029/2004JB003155.

McNeil, A.J., Saladin, T., 1997. The peaks over thresholds method for estimating high quantiles of loss distributions. Colloquium 28th International ASTIN Bulletin. Web reference: http://www.ma.hw.ac.uk/~mcneil/ftp/cairns.pdf.

Mora, J.C., Jaimes-Viera, M.C., Garduño-Monroy, V.H., Layer, P.W., Pompa- Mera, V., Godinez, M.L., 2007. Geology and geochemistry characteristics of the Chiapanecan volcanic arc (Central Area), Chiapas Mexico. J. Volcanol. Geotherm. Res. 162, 43–72. doi:10.1016/j.jvolgeores.2006.12.0009.

Mulargia, F., Gasperini, P., Tinti, S., 1987. Identifying different regimes in eruptive activity: an application to Etna volcano. J. Volcanol. Geotherm. Res. 34, 89–106.

Newhall, C.G., Self, S., 1982. The Volcanic Explosivity Index (VEI): an estimate of explosive magnitude for historical volcanism. J. Geophys. Res. 87C2, 1231–1238.

Newhall, C.G., Hoblitt, R.P., 2002. Constructing event trees for volcanic crises. Bull. Volcanol. 64, 3–20.

Palumbo, A., 1997. Long-term forecasting of large volcanic eruptions. J. Volcanol. Geotherm. Res. 78, 179–183.

Pickands, J., 1971. The two dimensional Poisson process and extremal processes. J. Appl. Probab. 8, 745–756.

Pickands, J., 1975. Statistical inference using extreme order statistics. Ann. Stat. 3 (1), 119–131.

Pyle, D.M., 1998. Forecasting sizes and repose times of future extreme volcanic events. Geology 26, 367–370.

Rampino, M.R., Self, S., 1984. Sulphur-rich volcanic eruptions and stratospheric aerosols. Nature 310, 677–679.

Reiss, R.D., Thomas, M., 2001. Statistical Analysis of Extreme Values, 2nd ed. Birkhauser, Basel. 316pp.

Reyment, R.A., 1969. Statistical analysis of some volcanologic data. Regarded as series of point events. PAGEOPH 74 (3), 57–77.

Robin, C., 1984. Le Volcan Popo (Mexique): structure, evolution petrologique et risques. Bull. Volcanol. 47, 1–23.

Schönwiese, C.D., 1988. Volcanic activity parameters and volcanism–climate relationships within the recent centuries. Atmosfera 1, 141–156.

Siebe, C., Macías, J.L., 2004. Volcanic Hazards in the Mexico City Metropolitan Area from Eruptions at Popocatépetl, Nevado de Toluca, and Jocotitlán Stratovolcanoes and Monogenetic Scoria Cones in The Sierra Chichinautzin Volcanic Field. Geological Society of America. Fieldtrip Books, 1.
Siebe, C., Abrams, M., Macías, J.L., Obenholzner, J., 1996. Repeated volcanic disasters in preHispanic time at Popocatépetl, Central Mexico: past key to the future? Geology 24, 399–402.
Simiu, E., 1995. Estimation of extreme wind speeds. Indian Society for Wind Engineering. Sponsored by University of Roorkee, New Delhi. International Association for Wind Engineering. State of the Art in Wind Engineering. Proceedings. International Conference on Wind Engineering, 9th. Davenport Sixtieth Birth Anniversary Volume. January 9–13, New Delhi, India, vol. 1, pp. 109–123.
Simkin, T., Siebert, L., 1994. Volcanoes of the World, 2nd ed. Geoscience Press, Tucson, Ariz.
Simkin, T., Siebert, L., 2000. Earth's volcanoes and eruptions: an overview. In: Sigurdsson, H. (Ed.), Encyclopedia of Volcanoes. Academic Press, San Diego, pp. 249–261.
Smith, R.L., 1989. Extreme value analysis of environmental time series: an application to trend detection in ground-level ozone (with discussion). Stat. Sci. 4, 367–393.
Solow, A.R., 2001. An empirical Bayes analysis of volcanic eruptions. Math. Geol. 33 (1), 95–102.
Turner, M.B., Cronin, S.J., Bebbington, M.S., Platz, T., 2007. Developing probabilistic eruption forecasts for dormant volcanoes: a case study from Mt Taranaki, New Zealand. Bull. Volcanol. 70, 507–515. doi:10.1007/s00445-007-0151-4.
UNDRO, 1979. Natural disasters and vulnerability analysis. Office of the United Nations Disaster Relief Coordinator (UNDRO). Report of Expert Group Meeting (9–12 July 1979). UNDRO, Geneva.
Wickman, F.E., 1965. Repose period patterns of volcanoes, 5. General discussion and a tentative stochastic model. Ark. Mineral. Geol. 4, 351–367.
Wickman, F.E., 1976. Markov models of repose-period patterns of volcanoes. In: Merriam, DF (Ed.), Random Processes in Geology. Sprínger-Verlag, Berlín, pp. 135–161.
Yang, Z.L., Xie, M., 2003. Efficient estimation of the Weibull shape parameter based on a modified profile likelihood. J. Stat. Comput. Simul. 73, 115–123.