Journal of Volcanology and Geothermal Research 193 (2010) 82–92

Contents lists available at ScienceDirect

Journal of Volcanology and Geothermal Research

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

# Eruption time series statistically examined: Probabilities of future eruptions at Villarrica and Llaima Volcanoes, Southern Volcanic Zone, Chile

Yvonne Dzierma [a,*], Heidi Wehrmann [b]

[a] SFB 574, University of Kiel, Otto-Hahn-Platz 1, 24118 Kiel, Germany
[b] SFB 574, Wischhofstr. 1-3, 24148 Kiel, Germany

ARTICLE INFO

*Article history:*
Received 12 October 2009
Accepted 20 March 2010
Available online 1 April 2010

*Keywords:*
eruption time series
distribution functions
probabilistic eruption forecasting
volcanic hazards
Villarrica Volcano
Llaima Volcano

ABSTRACT

Probabilistic forecasting of volcanic eruptions is a central issue of applied volcanology with regard to mitigating consequences of volcanic hazards. Recent years have seen great advances in the techniques of statistical analysis of volcanic eruption time series, which constitutes an essential component of a multi-discipline volcanic hazard assessment. Here, two of the currently most active volcanoes of South America, Villarrica and Llaima, are subjected to an established statistical procedure, with the aim to provide predictions for the likelihood of future eruptions within a given time interval.

In the eruptive history of both Villarrica and Llaima Volcanoes, time independence of eruptions provides consistency with Poissonian behaviour. A moving-average test, helping to assess whether the distribution of repose times between eruptions changes in response to the time interval considered, validates stationarity for at least the younger eruption record. For the earlier time period, stationarity is not entirely confirmed, which may artificially result from incompleteness of the eruption record, but can also reveal fluctuations in the eruptive regime. To take both possibilities into account, several different distribution functions are fit to the eruption time series, and the fits are evaluated for their quality and compared. The exponential, Weibull and log–logistic distributions are shown to fit the repose times sufficiently well. The probability of future eruptions within defined time periods is therefore estimated from all three distribution functions, as well as from a mixture of exponential distribution (MOED) for the different eruption regimes and from a Bayesian approach. Both the MOED and Bayesian estimates intrinsically predict lower eruption probabilities than the exponential distribution function, while the Weibull distributions have increasing hazard rates, hence giving the highest eruption probability forecasts.

This study provides one of the first approaches to subject historical time series of small eruptions (including those of Volcanic Explosivity Index = 2) of very active volcanoes to this type of statistical analysis. Since both Villarrica and Llaima are situated in a region of high population density, the eruption probabilities determined in this study present a valuable contribution to regional hazard assessment.



## 1. Introduction

The Andean Cordillera comprises four segments with active volcanic arcs, formed by subduction of the Nacza Plate beneath the South American Plate. Those are from North to South the Northern Volcanic Zone (NVZ, 5–2°S), the Central Volcanic Zone (CVZ, 14–27°S), the Southern Volcanic Zone (SVZ 33–46°S) and the Austral Volcanic Zone (AVZ, 49–55°S) (Lopez et al., 1995; Stern, 2004). The longest and most active of these, the SVZ, includes Villarrica and Llaima Volcano (Fig. 1), two of the most active volcanoes of South America, and has recently stirred much attention by the unexpected explosive eruption of Chaitén volcano, previously thought to be dormant or extinct. The high population density, industrial and agricultural use, and touristic attractiveness of the regions adjacent to the active volcanoes call for volcanic hazard assessment, aiming at eruption prediction and mitigation for these hazard-exposed zones.

Volcanic hazard assessment and mitigation can best be achieved by the continuous monitoring of various processes such as seismicity, amount and chemical characteristics of degassing, deformation etc., which in many cases allow for timely warning of an impending eruption (e.g. Scarpa and Tilling, 1996). However, on a longer time-scale, statistical analyses have also proved useful. Past eruption patterns have been statistically evaluated and were found to yield a prediction of expected future eruptions based on eruption frequency (e.g. Mendoza-Rosas and De la Cruz-Reyna, 2008). While this approach does not acknowledge interacting tectonic, geophysical, and geochemical processes as immediate eruption triggers, it uses this particular interaction to introduce a random behaviour in the eruption time series, which can then be investigated using the methods of standard life distribution/failure analyses (for an introduction to this topic, see, e.g., Marshall and Olkin, 2007; Cox and Oakes, 1984).

After the introduction of stochastic, i.e., non-deterministic, time series models for repose times by Wickman (1966), this approach has been successfully applied to various volcanoes around the world, notably the

* Corresponding author. Tel.: +49 431 880 3591; fax: +49 431 880 4456.
*E-mail address:* ydzierma@geophysik.uni-kiel.de (Y. Dzierma).

0377-0273/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.jvolgeores.2010.03.009

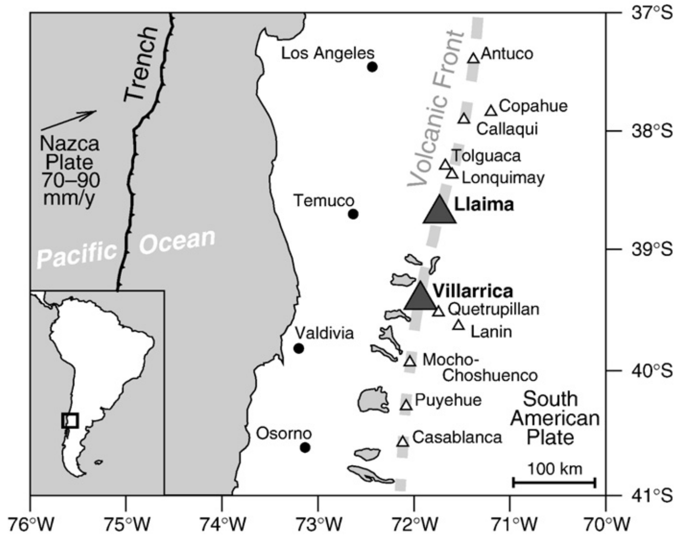


**Fig. 1.** Location map of Villarrica and Llaima Volcanoes.

Hawaiian volcanoes (Klein, 1982; Ho, 1990; Bebbington, 2008), New Zealand volcanoes (Bebbington and Lai, 1996a; Turner et al., 2008), Soufrière Hills (Connor et al., 2003), Italian volcanoes (Ho, 1990; Marzocchi et al., 2004; Bebbington, 2007; 2008), Mexico (Mendoza-Rosas and De la Cruz-Reyna, 2008, 2009) and others. To date, the only application to Chilean volcanoes appears to have been performed by Muñoz (1983), testing the volcanoes Villarrica, Llaima and Tupungatito for Poisson behaviour.

In the course of such recent advances in volcano statistics, we here implement these most elemental statistical techniques for the two selected volcanoes of the Chilean SVZ. Our aim is not to pursue a methodological advancement of the statistical theory, but to apply an established procedure to help estimate the likelihood of future eruptions within a given time interval, and to discuss possible shortcomings of the methodological approach. With this work we present a first step towards the wider aims of an integrative hazard assessment, which would have to encompass geophysical, volcanological and geochemical findings.

## 2. Volcanoes considered in this study

This study will focus on Villarrica and Llaima, two of the most active volcanoes of South America, whose eruptive record comprises about fifty eruptions over the past 400 years. Individual testing of other recently active volcanoes of the SVZ may be implementable, though possibly associated with more perceptible limitations because of insufficiencies in the documentation of the eruption record, or too low eruption frequencies. Chaitén, for example, cannot be statistically analysed since only one eruption (7420 BC) is known before the onset of new activity on May 2, 2008.

### 2.1. Villarrica Volcano

Villarrica Volcano is situated at the northern tip of the southern part of the large N–S trending Liquiñe–Ofqui Fault Zone (LOFZ), where it intersects with the W–E trending Mocha–Villarrica Fault Zone at 39°25′14″S / 71°56′23″W. It emerged in the Mid- to Late-Pleistocene by ejecting lava flows as well as violent fallouts, which led to an early caldera collapse at 95 ka BP. The following 80 ka were characterised by recurrent effusive and explosive activity, culminating in another caldera formation through release of several voluminous mafic pyroclastic flows at 14 ka BP. Continuous explosive activity rebuilt the current stratocone, interrupted by the 3.7 ka BP eruption of the Licán ignimbrite, which produced a smaller summit caldera. The historic activity has been mainly effusive with some strombolian explosions. At present, Villarrica Volcano contains a small lava lake inside the summit crater, subject to permanent degassing (Lara and Clavero, 2004). Despite this volcanologically rather weak activity pattern of present time, Villarrica's morphologic features and location in close proximity to the townships of Pucón and Villarrica and their touristic popularity, even smaller eruptions may cause a considerable hazard to people and property. Permanent monitoring is performed by the Observatorio Volcanológico de los Andes del Sur of the Servicio Nacional de Geología y Minería.

### 2.2. Llaima Volcano

Llaima Volcano, located slightly westward of the LOFZ at 38°41′30″S / 71°43′43″W, is a compound basaltic to andesitic stratovolcano that has grown since the Late-Pleistocene, initially dominated by effusive activity. An explosive stage started with a caldera-forming eruption at 13 ka BP and lasted till 7 ka BP, characterised by several large Plinian eruptions. The historical activity consists mostly of effusive behaviour, which is interrupted by numerous smaller explosions and accompanied by quiescent degassing. Recent explosive activity in 2008 and 2009, which forced evacuation of local population and caused damage to property, lead the Chilean government to support efforts of monitoring and surveillance (Dept. de Geología Aplicada of the Servicio Nacional de Geología y Minería, personal communication; Naranjo and Moreno, 2005; Stern, 2004; and references therein).

For the two volcanoes studied here, information on past eruptions is taken from the Smithsonian Global Volcanism Program website (www.volcano.si.edu), compiled from Naranjo and Moreno (2005), and Lara and Clavero (2004). Following the convention by Klein (1982), repose times are taken to be the interval of time between the *onset* of successive eruptions, which neglects the duration of the individual eruptions.

## 3. Eruption record

Statistical evaluation of eruption frequencies has been shown to be more reliable for large eruptions than for smaller ones (e.g. De la Cruz-Reyna, 1991). In this paper, we filter the eruption records of the target volcanoes by the Volcanic Explosivity Indices (VEIs), as assigned in literature data for the individual eruptions. The VEI has been defined by Simkin and Siebert (1994), refined after Newhall and Self (1982), as a measure of eruption magnitude, considering predominantly the erupted tephra volumes, supported by several other eruption parameters such as eruption duration, column height, and qualitative descriptive terms (Newhall and Self, 1982).

As a large number of explosive eruptions occurred at both Llaima and Villarrica in the past centuries, we will focus here on the historical data only (Tables 1a and 1b).

In the historical record, most eruptions from Villarrica did not exceed VEI = 2. Eruptions smaller than VEI = 2, of which the deposits are more likely removed and therefore unidentifiable, yield a higher risk of being missed in the chronological eruption documentation. In particular in the earlier periods of the eruption record, it is likely that the record of smaller eruptions may be incomplete. For this reason, only eruptions with VEI ≥ 2 are considered in this study.

For consistency and comparability of the results, the limit of VEI ≥ 2 will also be applied to Llaima Volcano, although it would also be possible to do the statistical assessment using a VEI ≥ 3 limit for Llaima.

## 4. Poisson processes

In the simplest and ideal case, the occurrence of a sequence of events such as volcanic eruptions can often be modelled as a Poisson process, a stochastic process in which events

(1) occur seldom (the probability of two or more events occurring contemporaneously is virtually zero),

**Table 1a**
Historical eruptions of Villarrica Volcano, with VEI.

| Year | VEI | Year | VEI | Year | VEI | Year | VEI |
|---|---|---|---|---|---|---|---|
| 1558 | 2 | 1815 | 1 | 1908 | 2 | 1963 | 3 |
| 1562 | 2 | 1822 | 2 | 1909 | 2 | 1964 | 2 |
| 1594 | 2 | 1832 | 2 | 1915 | 1 | 1971 | 2 |
| 1647 | 1 | 1837 | 2 | 1920 | 2 | 1977 | 1 |
| 1657 | 1 | 1853 | 2 | 1921 | 2 | 1980 | 2 |
| 1688 | 1 | 1859 | 2 | 1922 | 2 | 1983 | 1 |
| 1716 | 1 | 1864 | 2 | 1927 | 2 | 1984 | 2 |
| 1730 | 2 | 1869 | 2 | 1929 | 1 | 1991 | 2 |
| 1737 | 2 | 1874 | 2 | 1933 | 2 | 1992 | 1 |
| 1742 | 2 | 1875 | 2 | 1935 | 1 | 1994 | 1 |
| 1745 | 1 | 1877 | 2 | 1938 | 1 | 1995 | 1 |
| 1751 | 1 | 1879 | 2 | 1938 | 2 | 1996 | 1 |
| 1759 | 1 | 1880 | 2 | 1947 | 1 | 1996 | 1 |
| 1775 | 2 | 1883 | 2 | 1948 | 2 | 1998 | 1 |
| 1777 | 1 | 1893 | 2 | 1948 | 3 | 2003 | 1 |
| 1780 | 1 | 1897 | 2 | 1956 | 1 | 2004 | 1 |
| 1787 | 2 | 1904 | 2 | 1958 | 1 | 2008 | 1 |
| 1790 | 1 | 1906 | 2 | 1960 | 1 | | |
| 1806 | 2 | 1907 | 2 | 1961 | 1 | | |

**Table 1b**
Historical eruptions of Llaima Volcano, with VEI.

| Year | VEI | Year | VEI | Year | VEI | Year | VEI |
|---|---|---|---|---|---|---|---|
| 1640 | 4 | 1889 | 2 | 1932 | 3 | 1990 | 2 |
| 1751 | 2 | 1892 | 2 | 1937 | 2 | 1992 | 1 |
| 1759 | 2 | 1893 | 2 | 1938 | 1 | 1994 | 2 |
| 1822 | 2 | 1895 | 2 | 1941 | 2 | 1995 | 2 |
| 1852 | 2 | 1903 | 2 | 1942 | 2 | 1997 | 1 |
| 1862 | 3 | 1907 | 2 | 1944 | 2 | 1998 | 2 |
| 1864 | 3 | 1912 | 2 | 1945 | 3 | 1998 | 2 |
| 1866 | 2 | 1914 | 2 | 1946 | 2 | 2002 | 1 |
| 1869 | 2 | 1917 | 2 | 1949 | 2 | 2003 | 2 |
| 1872 | 2 | 1922 | 2 | 1955 | 3 | 2007 | 2 |
| 1875 | 2 | 1927 | 2 | 1964 | 2 | 2008 | 3 |
| 1877 | 2 | 1929 | 2 | 1971 | 2 | | |
| 1883 | 2 | 1930 | 2 | 1979 | 2 | | |
| 1887 | 2 | 1932 | 2 | 1984 | 2 | | |

(2) independently of one another (the probability of an event in any time interval is independent of what happened up to the start time of the interval) and
(3) with constant probability (there is no time trend, "memoryless property").

(The resulting formulae are discussed in Section 6.1; for a mathematical definition and derivation of these properties, see, e.g., Cox and Lewis, 1966).

### *4.1. Test for independence*

In a first step, it is tested whether successive eruptions are time independent from each other. This is done by calculating the correlation of successive repose times from serial correlation scatter plots (Cox and Lewis, 1966; Mendoza-Rosas and De la Cruz-Reyna, 2008). The correlation coefficient $R$ is calculated as

$$R = \sqrt{\frac{\sum\left(y_{est,i} - \bar{y}\right)^2}{\sum\left(y_i - \bar{y}\right)^2}} \quad (1)$$

which, in the special case of an assumed linear relationship, can be written as

$$R = \frac{\sum\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right)}{\sqrt{\left(\sum\left(x_i - \bar{x}\right)^2\right)\left(\sum\left(y_i - \bar{y}\right)^2\right)}} \quad (2)$$

where the sum is taken over the $n-1$ repose time values on the $x$-axis ($x_i$) and the repose time values on the $y$-axis ($y_i$), and their corresponding means $\bar{x}$ and $\bar{y}$. In the more general, nonlinear case, the estimated $y_{est,i}$ are calculated as any nonlinear function of the $x_i$. The square of the correlation coefficient (the coefficient of variation $R^2$) represents the fraction of the variation explained by the model to the total variation.

For a known correlation coefficient, we test whether the observed correlation could have arisen by pure chance (the null hypothesis "no correlation") by comparing the value $R\sqrt{\frac{n-2}{1-R^2}}$ with the tabulated values for the Student's $t$-distribution for 5% confidence level (e.g. Spiegel and Stephens, 2008).

Fig. 2B shows the serial correlation scatter plot for Villarrica repose times, with correlation coefficient $R=0.16$. This value does not yield enough statistical evidence to discard the null hypothesis that the values are uncorrelated (at the 5% level of significance). Therefore, this is consistent with the hypothesis of no correlation, and data points are close to the axes, both pointing to independence of successive repose times. However, this may be an artefact generated by the longest repose time of 136 years in the early eruption record.

Indeed, before 1730, eruptions occur rarely, with large repose times in between, whereas after 1730, they become fairly regular. While this observation might reflect a real eruption frequency increase in 1730, such that Villarrica entered regime of higher activity, a more likely cause may be that the earlier historical record is incomplete. In the documentation, assigned VEI values of older eruptions are more at risk of being too low, as the erupted volumes may be underestimated because of removal of the respective deposit by erosion. Additionally, entire small eruptions may be missed because the deposits could be too disturbed. But even if only eruptions after 1730 are taken into account (inset in Fig. 2A), the repose times give a very low correlation coefficient of 0.19, still consistent with the assumption of independence of successive eruptions.

In the case of Llaima, Fig. 2B shows that again the rate of eruptions seems to increase markedly at some time in the past, in this case around 1850 AD. As for Villarrica, this may have been caused by a real change in eruption activity, but might also reflect incompleteness of the historical record. Therefore, the inset in Fig. 2B repeats the test, but restricted to the time after 1850.

For both time intervals, the correlation of successive Llaima repose times is weak, the correlation coefficients are 0.26 and 0.25, respectively. As for Villarrica, this is consistent with the assumption of independence. In addition to the calculated correlation coefficient, this independence also becomes evident from visual inspection of the figure, which displays large scatter between data pairs.

Despite the mentioned limitations of the data, we do not find compelling evidence to discredit the assumption of independence of successive eruptions and assume that memory effects are lacking.

## 5. Test for stationarity

Non-stationarity of the dataset (namely, the repose time series), meaning that the probability distribution or parameters may depend on the time interval considered, can be detected by a moving average test (Klein, 1982; De la Cruz-Reyna, 1996; Mendoza-Rosas and De la Cruz-Reyna, 2008). To implement this test, the average of each five successive repose times is calculated and the results are plotted as a function of time. This test is performed to detect non-homogeneity of the assumed Poisson process, in the respect that the eruption rate-parameter may change over time. While smoothing over short-term fluctuations, this provides insight into a possible trend in repose times. A condition for this test is the completeness of the eruption record. On the other hand, if stationarity is assumed, the test can serve to detect possible incompleteness of the dataset.

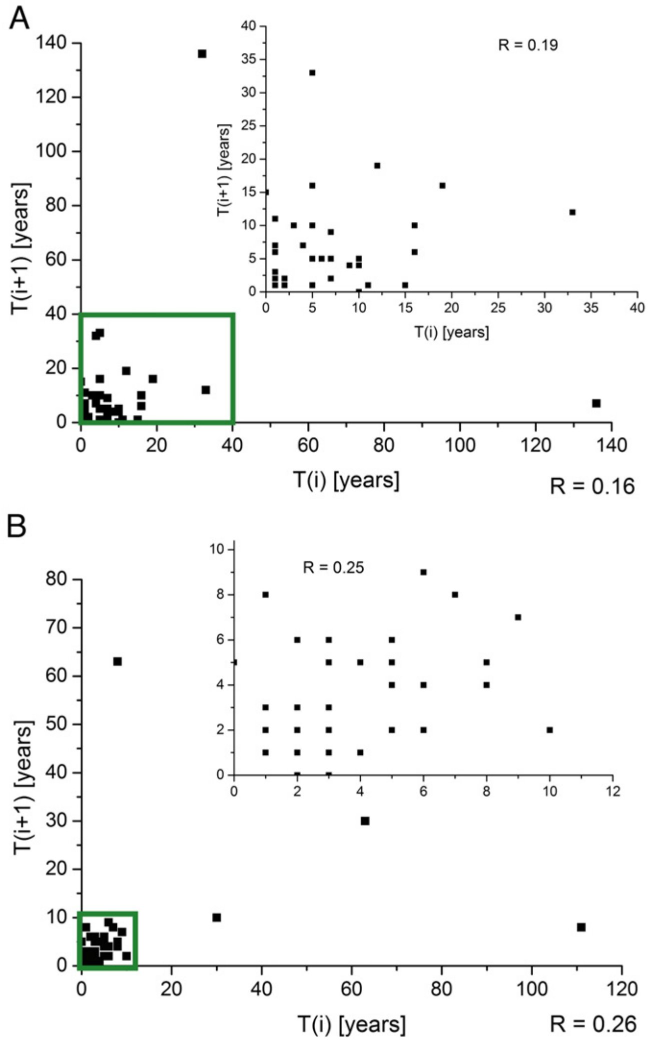


**Fig. 2.** A,B. Serial correlation scatter plot of repose times for (A) Villarrica, from 1558 (inset from 1730), (B) Llaima, from 1640 (inset from 1850).

For the total historical eruption record of Villarrica, the average 5-pt repose time amounts to 9.0 years. The first two 5-pt repose time averages, however, fall well outside the 95 % confidence limit of this value (Fig. 3A). As already inferred from Fig. 2A, this shows that stationarity is not confirmed for the early eruption record, indicating that either a change in the eruption frequency took place, or that the early eruption record is not complete.

If we restrict the analysis to eruptions after 1730, the inset in Fig. 3A shows that the variations in 5-pt average repose times are consistent with random variations around the mean (6.4 years). Although the diagram hints at a decrease in average repose times in the 19th century followed by some minor variations, statistically, these changes are not significant and may be pure fluctuations. Hence, the data are consistent with the assumption of stationarity.

For Llaima, the 5-pt repose time averages also appear to decrease steadily until about 1870 (Fig. 3B); after that, the averages are much more constant. The first average falls outside the 95% confidence limit of the mean (5.6 years), which agrees with our previous apprehension that the eruption record before 1850 may not be complete — or that the eruption regime may have changed to stronger activity around this time. Including only eruptions after 1850 in the analysis (inset in Fig. 3B), the data are again consistent with the assumption of stationarity.

## 6. Survival time fit

Define the repose time distribution function that "the probability that the observed repose time $T$ is smaller than or equal to $t$" by

$$F(t) = P\{T \leq t\} \tag{3}$$

with values in the interval [0,1]. This cumulative distribution function $F(t)$ is non-decreasing in $t$, and $T$ is the random variable for the repose time between the eruptions. The corresponding survival function $S(t)$ is then

$$S(t) = P\{T > t\} = 1 - F(t) \tag{4}$$

The observed number of repose times greater than time $t$ is then obtained by scaling the survival function with the total number of observed reposes.

In this application, we consider only basic distributions that are "sufficiently good-natured" so that the density function $f$ exists, where

$$f(t) = \frac{d}{dt}F(t) \tag{5}$$

is a measure of the instantaneous failure probability at time $t$. Several functions are commonly used to approximate the observed repose time distribution.

### *6.1. Exponential distribution*

The simplest case of a stationary Poisson process gives rise to an exponential distribution function

$$F_{\mathrm{exp}}(t) = 1 - \exp(-\lambda t) \tag{6}$$

$$S_{\mathrm{exp}}(t) = \exp(-\lambda t) \tag{7}$$

$$f_{\mathrm{exp}}(t) = \lambda \exp(-\lambda t) \tag{8}$$

The distinguishing characteristic of the exponential function is that the statistical hazard rate

$$r(t) = \lim_{\Delta \searrow 0} \frac{P\{t < T \leq t + \Delta | T > t\}}{\Delta} = \frac{f(t)}{S(t)} = \lambda \tag{9}$$

is constant, i.e., the probability of an eruption occurring in the next small increment of time does not depend on the time that has already elapsed since the last eruption occurred.

This means that if we define the residual life distribution $S_x$ by

$$S_x(t) = \frac{S(x+t)}{S(x)} \tag{10}$$

i.e., the survival function as a function of time $t$ after a given waiting time (age) $x$, given survival up to that time, the exponential distribution is characterized by the property that $S_{x,exp}(t) = S_{exp}(t)$. This means, an exponential distribution is not affected by aging; i.e., the eruptive regime at the time considered does not suffer from exhaustion of the driving forces consumed by previous eruptions.

### *6.2. Weibull distribution*

It may be argued on physical grounds that the hazard rate should be allowed to systematically increase/decrease with time to include regimes of increasing volcanic activity or waning/extinguishing activity. This can be accomplished by the Weibull distribution, commonly used in

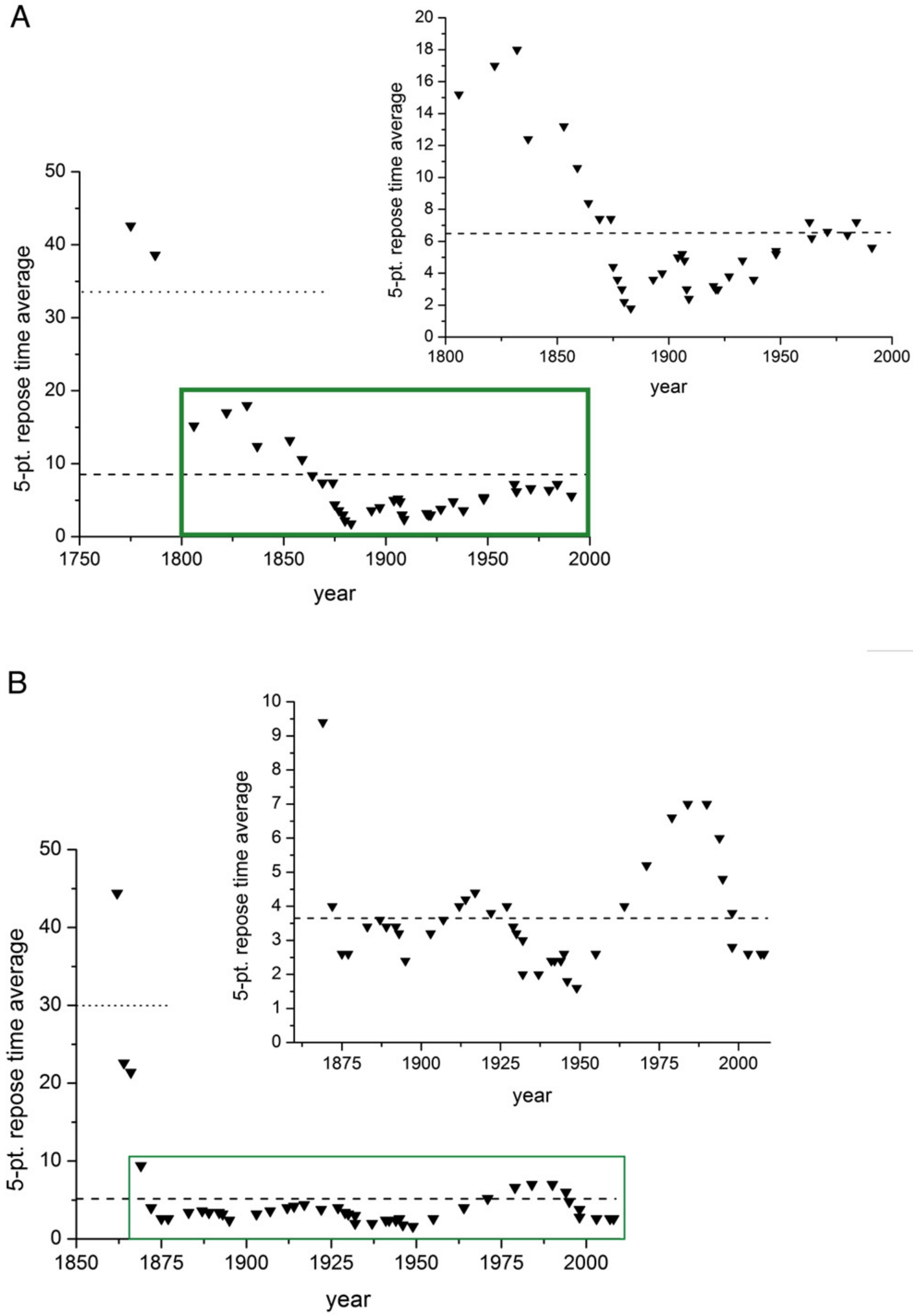


**Fig. 3.** A,B. 5-point moving averages of Villarrica and Llaima repose times, respectively (plotted at the end of the last repose). Dashed line: mean, dotted line: 95% confidence interval of the mean. A) Villarrica repose times (inset from 1730); B) Llaima repose times (inset from 1850).

failure analysis and successfully applied to various volcanoes (Ho, 1991; Bebbington and Lai, 1996a,b; Watt et al., 2007):

$$S_{WB}(t) = \exp\left\{-(\lambda t)^{\alpha}\right\} \tag{11}$$

where $\alpha$ is a power parameter, usually referred to as the "shape parameter". For $\alpha = 1$, the Weibull distribution includes the exponential distribution as a special case, but it also accommodates the possibilities of increasing or decreasing hazard rates if $\alpha > 1$ or $\alpha < 1$, respectively. As the Weibull distribution represents a model of simple failure, it best illustrates scenarios that consider this failure after a given time as a result of only one dominant process in the system. Volcanologically, this could be projected to repose times of magma maturation between the eruptions in a continuous or temporally regular pattern, and eruption onset when gas pressure through magma differentiation builds up beyond a critical threshold. In such scenario, the value of $\alpha$ could be used to express for example changes in the replenishment rate of a shallow chamber from a deeper source, resulting in episodes of systematic waxing or waning activity.

### 6.3. Log–logistic distribution

Finally, it has been argued that competing processes may act in a way that particular parameters increase the probability of an eruption, while other, contemporaneously counteracting parameters cause a decrease in eruption probability. For example, vent closure as commonly occurs at Llaima Volcano between the eruptions combined with only limited permanent gas release, hence volatile accumulation in the melt through magma storage in shallow reservoirs, will lead to a positive feedback process of pressure increase, thus contributing to explosion probability enhancement. In contrast, processes such as open system degassing, as takes place at the lava lake of Villarrica, will facilitate a continuous quiescent pressure release, therefore rather leading to relaxation of the system and lowering the likelihood of an eruption. Such influencing factors can be phenomenologically addressed by the introduction of, e.g., shape parameters. Interacting processes such as those described can give rise to a log-logistic distribution (Pareto III distribution):

$$S_{\log}(t) = \frac{1}{1 + (t/b)^{\alpha}} \tag{12}$$

which includes a scale parameter $b$ and a shape parameter $\alpha$. A log–logistic distribution can sometimes achieve a better fit particularly to very long or short repose intervals (Connor et al., 2003).

### 6.4. Quality of the fits

Since our aim in this study is to show the applicability of the standard statistical modelling schemes to any volcano of interest, we will indiscriminately try to fit all three possible distribution function models to the observed repose time distribution functions, using the method of least-squares fitting (equivalent to maximum-likelihood if the errors are assumed to be normally distributed). A standard software package can be used for the fits (e.g., R, Matlab or Origin; here, OriginPro 8.0 and Origin 7.0 were used). The fit curves, parameters and statistical measures ($\chi^2$ at the respective degrees of freedom (DoF) and coefficients of determination $R^2$, which indicate the proportion of variance in the dataset that is accounted for by the statistical model) are provided in Fig. 4A,B for both volcanoes.

#### 6.4.1. Akaike information criterion

The Akaike Information Criterion (AIC) is a tool for model selection, which penalises both the misfit and a large number of parameters (Akaike, 1973; Bebbington, 2007; Turner et al., 2008). In the case of least squares fitting to the same dataset, the AIC is calculated as

$$AIC = n \cdot \ln \frac{\sum (y(x_i) - y_i)^2}{n} + 2k \tag{13}$$

where $n$ is the number of data points fitted, $k$ the number of free parameters, and $y(x_i)$ the points of the model $y(x)$ used to fit the data points $(x_i, y_i)$. In the case of a small number of data points $n$, a correction needs to be applied, giving the adjusted

$$AIC_c = AIC + \frac{2k(k+1)}{n-k-1} \tag{14}$$

(Sigiura, 1978; Burnham and Anderson, 1998). Results are given in Table 2. On the basis of the $AIC_c$, the best fit for Llaima is achieved by the Weibull function, while for Villarrica the exponential function is preferred.

The absolute value of the $AIC_c$ is irrelevant, the criterion provides only a relative comparison between different models, in which the lowest value points to the best fit. While this allows us to select the best of the models, the $AIC_c$ is not a goodness-of-fit test. Therefore, the absolute quality of the fit needs to be assessed with a standard hypothesis test such as the Kolmogorov–Smirnov-test.

#### 6.4.2. Kolmogorov–Smirnov-test

The goodness of the three fits is evaluated using the Kolmogorov–Smirnov–(K–S-) test (Table 3) (e.g., Gibbons, 1976). In contrast to the $AIC_c$, where the misfit is evaluated by the root-mean-square deviation,

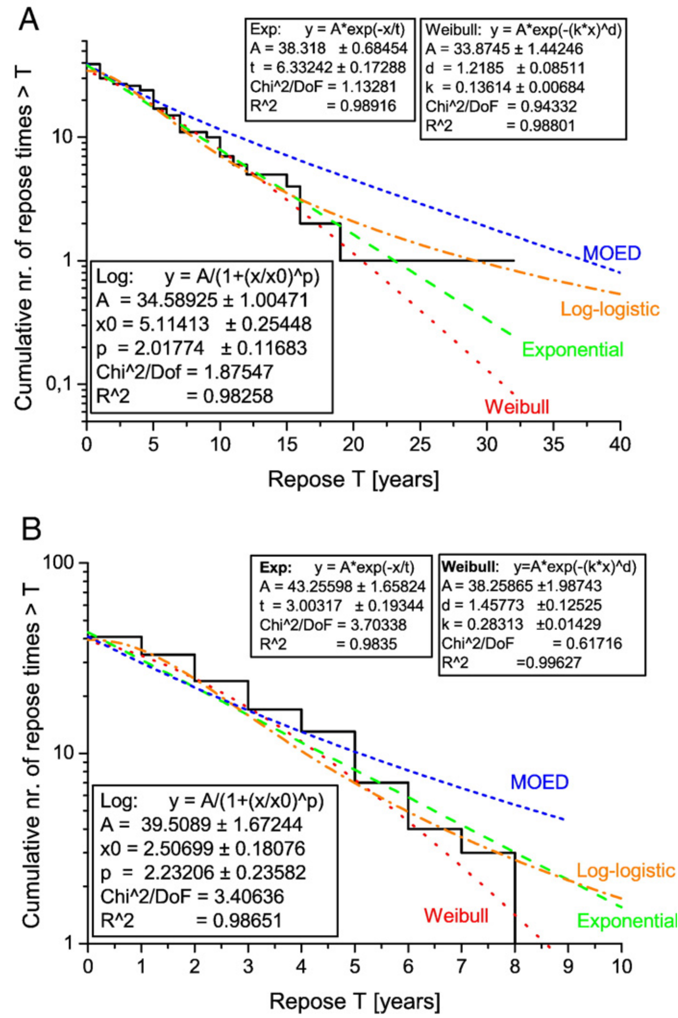


**Fig. 4.** A,B. Cumulative repose time fits for (A) Villarrica (from 1730), (B) Llaima (from 1850). Exponential, Weibull and log-logistic distributions were fit to the data with the Origin software package using least-squares fitting. Fit parameters and results (reduced $\chi^2$, i.e., $\chi^2$ divided by the number of degrees of freedom (DoF), $R^2$) are given in the boxes. The MOED is also plotted for comparison, see Section 7).

the K–S-test is a nonparametric test which assesses the maximum deviation of the model from the data and evaluates the probability that such a deviation might have arisen by chance. The advantage of the K–S-test over, e.g., $\chi^2$, is that it is an exact test, i.e., valid independently of the number of data points, and nonparametric, that is, independent of the underlying probability distribution.

The $\chi^2$-test, in contrast, would be easier to reconcile with the $AIC_c$ results, because the degree of misfit is evaluated in the same way. Another disadvantage of using the K–S-test as a measure of goodness of fit is that it does not account for the reduction of the number of degrees of freedom when parameters are estimated from the data. In this case, the K–S-values are conservative and possibly overly restrictive. The K–S-test provides the maximum of the absolute deviations, which is a measure of the goodness of fit, but a very restrictive condition in fitting. (For a more ample discussion of $\chi^2$ vs.

**Table 2**
$AIC_c$ results for fits to Villarrica and Llaima repose time functions.

| | Villarrica | Llaima |
|---|---|---|
| Exponential | 9.37 | 21.62 |
| Weibull | 27.14 | 15.47 |
| Log–logistic | 30.82 | 25.89 |

**Table 3**
K–S differences for fits to Villarrica and Llaima repose time functions.

| | Villarrica | Llaima |
|---|---|---|
| Exponential | 0.169 | 0.263 |
| Weibull | 0.205 | 0.204 |
| Log-logistic | 0.162 | 0.199 |

K–S-testing, see, e.g., Gibbons, 1976). On the other hand, if the distributions fit to the data are consistent with both $\chi^2$ and the K–S-test, this provides some independent insight into the quality of the fit.

For Villarrica Volcano, both the exponential and log-logistic distributions give K–S-differences smaller than that of the 0.05 level of significance (0.17 and 0.16, respectively, compared with a cut-off value of 0.21), so the models are consistent with the observations. The Weibull distribution, however, performs worse than the exponential distribution (K–S value of 0.21). This is a somewhat unreasonable contrast, because the Weibull distribution includes the exponential distribution as a special case with a constant shape parameter, therefore it should perform at least as well as the exponential distribution. The large difference is obtained at the origin, due to the fact that the scale factor $A$ differs markedly from the total number of observed repose times. If $A$ is fixed in the fitting process, the resulting Weibull distribution is very similar to the exponential distribution. Indeed, the exponential distribution offers a good fit particularly to small repose times, whereas it tends to zero quickly for longer repose times. The Weibull distribution, if fit with variable $A$, remedies this problem and provides a good fit to longer repose times with a long tail of the distribution function, but does not reproduce very short repose time statistics well, particularly not for a repose time value of 0. If we leave out the value at $T=0$ in the K–S-test for Villarrica and limit the fit to finite repose times, a K–S-difference of only 0.15, well below the 0.21 of the 0.05 limit, is obtained. It is justifiable to ignore $T=0$, because the source data provide many eruption dates only as calendar years, rendering it impossible to properly distinguish between 0 and 1. For instance, and the only instance in this time series, there were two VEI≥2 Villarrica eruptions in the calendar year 1948, which are mathematically treated as repose time $T=0$, while in fact they had a repose time of about ½ year and would have been treated differently had they both occurred two months later, so that the second eruption would have fallen into 1949 — a repose time of one year (i.e., $T=1$) would then have been used for the analysis. Considering eruption dates on a monthly to weekly time resolution would contribute to a better precision for the short repose times.

In the case of Llaima, both the Weibull and log-logistic distributions give a good fit to the data within the 0.05 level of significance (both 0.20), whereas the exponential distribution fails the test at this level (0.26). While it may seem paradox that the exponential distribution fails the K–S-test, despite a better $AIC_c$ than the log–logistic distribution, the lower and hence better $AIC_c$ is due solely to the lower number of parameters and does not reflect less misfit. The combination of the K–S-results with the $AIC_c$ indicates that a purely stationary Poisson process is not a good description of the Llaima eruption sequence, and the time series appear to be more likely affected by an increasing hazard rate.

## 7. Different eruption regimes — MOED

Recently, Mendoza-Rosas and De la Cruz-Reyna (2008; 2009) introduced the mixture of exponentials distribution (MOED) into volcano statistics analysis as a way to take into account possible non-stationarity of the eruption process. In their approach, the historical eruption record is divided into a number of different regimes: time intervals over which the eruption rate can reasonably be assumed to be constant. Statistical methods exist that allow for a mathematically sound and objective definition of regime boundaries in time series, which can involve application of hidden Markov models (Mulargia and Tinti, 1986; Mulargia et al., 1987; Bebbington, 2007). However, within the scope of this work, we decided for the intuitive, visual definition of regimes as used for MOED estimation and MOED-based Bayesian analyses (De la Cruz-Reyna, 1996; Mendoza-Rosas and De la Cruz-Reyna, 2008, 2009), which is based on visual identification of piecewise linear regimes in a plot of cumulative number of eruptions vs. time (Fig. 5A,B). Although it may seem at first sight that the choice of the start- and end-points of the intervals is arbitrary to some degree, the main features are stable: while the regime start and end times may be moved by one or two decades at most, this does not involve a significant change in the eruption rate for this regime, nor in the results of the method.

The MOED is defined as the weighted sum over the individual regime distribution functions, giving a piecewise exponential behaviour:

$$F_{\mathrm{MOED}}(t) = \sum_{i=1}^{m} w_i(1-\exp(-\lambda_i t)), \tag{15}$$

$$S_{\mathrm{MOED}}(t) = \sum_{i=1}^{m} w_i \exp(-\lambda_i t) \tag{16}$$

$$\sum_{i=1}^{m} w_i = 1 \tag{17}$$

where $m$ is the total number of identified regimes, $\lambda_i$ is the eruption rate (number of eruptions in regime $i$ divided by duration of the regime), and $w_i$ are weighting factors given by

$$w_i = \frac{D_t - D_i}{\sum_{i=1}^{m}(D_t - D_i)} \tag{18}$$

with $D_i$ being the duration of each regime and $D_t$ the total duration of the sampled interval of time.

Following the procedure outlined by Mendoza-Rosas and De la Cruz-Reyna (2008), we use the chosen eruption regimes to determine the parameters of a mixture of exponentials distribution (Table 4a and b). Note that this method does not *fit* a MOED to give best agreement with the data, but *reads* the MOED parameters from the plot. In consequence, the agreement of the MOED with the data will depart from the optimum on one hand, but on the other hand, this method is much easier and more stable than fitting, and certainly more intuitive than using a readily implemented fitting routine. It is also directly insightful because the changes between phases of higher and lower volcanic activity become apparent in the plots.

For Villarrica, three episodes of differing eruption rates can be visually derived from Fig. 5A. The highest activity is bracketed by two phases of somewhat lower eruption rates, which method-wise diminishes the scenario that regimes of lower activity are only artefacts of insufficient eruption documentation aggravating when reaching further back in the past within this time span.

For Llaima, Fig. 5B shows that the data until 1862 can be grouped together in one regime, which has considerably lower eruption frequency than the following episode. We already supposed that the eruptive record up to the year 1850 might not be complete, and have up to now chosen this year as lower boundary of consideration. This new analysis, however, suggests that we should only consider Llaima eruptions after 1862. As mentioned before, it is of course possible that this flatter slope images a real regime of lower activity, implying that it should be included in the analysis. However, since this is the oldest part of the historic record, and the rate is saliently different from all later regimes, it is more plausible that this effect is due to incompleteness of the record and this time interval should be skipped. This also agrees with the outcome of the stationarity check performed above.

As for the three other distributions, the agreement of the MOED with the data is examined with the K–S-test and produces differences of 0.136 and 0.272 for Villarrica and Llaima, respectively. This gives an excellent agreement with observations for Villarrica, while failing for Llaima. We must, however, take into account that this is based on data

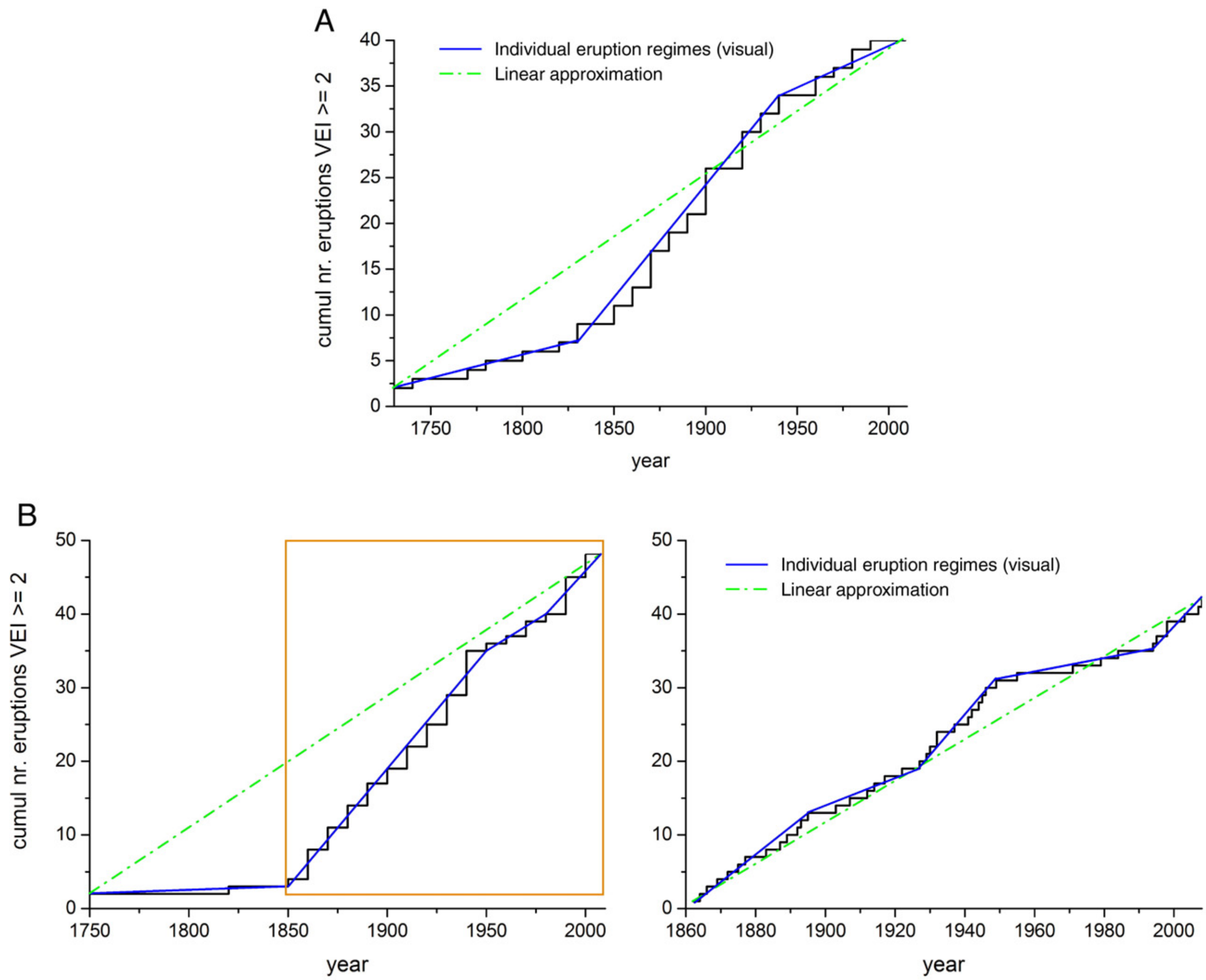


**Fig. 5.** A,B. Cumulative number of eruptions of (A) Villarrica (from 1730), and (B) Llaima (inset: eruptions from 1862). Blue/dashed lines: visually identified regimes of constant eruption rate; (green) dashed-dotted lines: constant eruption rate approximation to the total time interval.

read off from the plot without any kind of optimisation. Given the fact that the K–S-differences are comparable in size to those of the fits, the MOED model performs moderately well. Furthermore, there is no reason why the MOED should be a more adequate representation of the data than for example a single, but optimised, exponential distribution. To determine the MOED, exponential regimes are identified piecewise from the data, then the weighted mean of the individual functions is taken to compile the total function. This mixture of the exponentials is then applied to the entire data set for any given period, thus immanently accompanied by the possibility of deviations from the data.

**Table 4**
Parameters for MOED.

| Regime | Start year | End year | Nr. erupt. | Rate $\lambda_i$ | Weight $w_i$ |
|---|---|---|---|---|---|
| *Villarrica* | | | | | |
| 1 | 1730 | 1859 | 10 | 0.0769 | 0.2688 |
| 2 | 1859 | 1948 | 24 | 0.2667 | 0.3405 |
| 3 | 1948 | 2009 | 6 | 0.0968 | 0.3907 |
| *Llaima* | | | | | |
| 1 | 1862 | 1955 | 31 | 0.3333 | 0.1852 |
| 2 | 1955 | 1994 | 6 | 0.1539 | 0.3670 |
| 3 | 1994 | 2009 | 7 | 0.4667 | 0.4478 |

The only way to achieve a better representation of the data seizing their piecewise definition is to also adjust the distribution function to a piecewise exponential definition, in the way that

$$F_{\text{PED}}(t) = \sum_{i=1}^{m(t)} \{1 - \exp[-\lambda_i(t - t_i)]\} \quad (19)$$

where $m(t)$ is the regime which includes the time $t$, and $t_i$ is the time when the $i$th regime starts. This piecewise-exponential distribution is represented by the blue lines in Fig. 5A,B, which is continuous but only piecewise differentiable for $m>1$. However, this function – while giving an excellent approximation to the data in Fig. 5A,B – is not suitable for further analysis, for several reasons.

Firstly, the function depends on the start year, i.e., the expected distribution of repose times starting in year $t_1$ is generally not identical to the distribution of repose times starting in year $t_2 \neq t_1$, as long as repose times long enough to traverse from one regime to the next are considered. To find the distribution of repose times, we would therefore have to integrate over all possible starting times. While this is possible, it is cumbersome and doubtful whether the improvement in the results can justify the additional computation effort.

Secondly and more importantly, the ultimate aim of carrying out this analysis is a prediction of the present and future volcanic hazards. To infer the most realistic forecast from the time series, the present-day eruption rate is the most relevant section of the eruption record to be subjected to the statistical analysis. Therefore, defining a distribution function for an eruption rate in a limited time period that had terminated in the past is reduced to an academic exercise with no application for hazard research. This distribution would express only this particular time period and would not be valid for any later episodes, thus producing only results irrelevant for forecasting present-day eruption probabilities. The only ways to estimate the present-day eruption rate are therefore (1) to concentrate only on the most recent regime, which is considered to be ongoing, and use the observed eruption rate to extrapolate to the future, or (2) to use the weighted mean of past eruption rates, because there is no reason to believe that the last observed regime is still continuing, it is possible that we just entered a new regime with unknown eruption rate. This last assumption, using the weighted mean of all past eruption rates, is the underlying idea behind the use of the MOED, although the exact functional form of the weights may be subject to debate (see above). Therefore, while not providing as good a fit to the observed eruption data, previous investigators (Mendoza-Rosas and De la Cruz-Reyna, 2008, 2009) have considered the MOED useful for the estimation of future eruptions. As we are applying the most common techniques here, also the MOED will be used for eruption forecasting later in this work.

## 8. MOED-based Bayesian analysis of present eruption rate

When interpreted from a Bayesian point of view, what the MOED does is to provide an estimate of the eruption rate used for the forecast, based on the length of time period each eruption rate has been observed to occur in the past. This is to some degree intuitive, and it may be argued that a non-informative prior should be used instead. However, we believe that an educated guess of prior information, subsequently improved based on the data (the Bayesian approach), provides a better estimate than using no information at all. Nonetheless, also the latter approach will be followed in Section 9.

The weighting scheme used in the MOED estimation is to some degree debatable, as has been discussed. In the context of a Bayesian analysis, the underlying idea is that the eruptions follow an exponential distribution, but the value of the eruption rate at the present time is unknown and needs to be estimated. The eruption rates observed in the past in different eruption regimes can give a tentative estimate of the possible values of λ, therefore we weigh the regimes according to their length by

$$w_i^* = \frac{D_i}{D_t} \tag{20}$$

to determine the prior expectation of the eruption rate

$$\lambda^*_{\text{prior}} = \sum_i w_i^* \lambda_i \tag{21}$$

as the average over the observed eruption rates $\lambda_i$ up to now.

Taking advantage of Bayesian theory, the posterior expectation for the eruption rate can be calculated based on Bayes' theorem that

$$P(\lambda_i|y) = \frac{P(y|\lambda_i)P(\lambda_i)}{\sum_i^n P(y|\lambda_i)P(\lambda_i)} \tag{22}$$

With $P(\lambda_i|y)$ the probability of each $\lambda_i$ reviewed in the light of an experimental outcome $y$ = "no eruption at the present time", $P(\lambda_i) = w_i$ the prior (estimated "intrinsic") probability of $\lambda_i$ and $P(y|\lambda_i)$ the likelihood of observing $y$ = "no eruption" given $\lambda_i$, calculated as the probability of event "zero eruptions" in a Poisson distribution with parameter $\lambda_i$:

$$P(0,\lambda_i) = \frac{(\lambda_i \Delta t)^0 \exp(-\lambda_i \Delta t)}{0!} = \exp(-\lambda_i \Delta t) = \exp(-n_i) \tag{23}$$

where $n_i$ is the number of eruption in the $i$th regime (see De la Cruz-Reyna, 1996 for a full explanation and, e.g., Gelman et al., 2009, for details on the Bayesian approach).

The $P(\lambda_i|y)$ can then be used as weights in the calculation of the posterior expected value of λ in the light of the known data:

$$\lambda_{\text{posterior}} = \sum_i P(\lambda_i|y)\lambda_i \tag{24}$$

The calculations and results are given in Table 5.

## 9. Bayesian analysis based on non-informative Gamma prior

If no prior knowledge of the eruption rate is assumed, a non-informative Gamma prior may serve to describe the distribution of possible eruption rates (Solow, 2001; Varley et al., 2006; Gelman et al., 2009).

$$\pi(\lambda) = \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} \exp(-\beta\lambda) \tag{25}$$

where for the non-informative prior case $\alpha = \beta = 0$. The posterior distribution is then of the same form, with parameters $\alpha_{\text{post}} = \alpha + n_{\text{total}} = n_{\text{total}}$ and $\beta_{\text{post}} = \beta + t_{\text{total}} = t_{\text{total}}$, $n_{\text{total}}$ being the total number of eruptions observed in the total time $t_{\text{total}}$:

$$P(n \text{ eruptions in time } t_0) = \left(\frac{\beta_{\text{post}}}{\beta_{\text{post}} + t_0}\right)^{\alpha_{\text{post}}} \left(\frac{\beta_{\text{post}}}{t_0} + 1\right)^{-n} \tag{26}$$

The probability of at least one eruption occurring in the time $t_0$ then becomes

$$\begin{aligned} P(T \le t_0) &= 1 - P(n = 0 \,\text{eruptionsintime}\, t_0) \\ &= 1 - \left(\frac{\beta_{\text{post}}}{\beta_{\text{post}} + t_0}\right)^{\alpha_{\text{post}}} \end{aligned} \tag{27}$$

## 10. Forecasting eruption probabilities in comparison of the methods

The different distributions as well as the results from the Bayesian analysis are used here to determine the probability of witnessing at least one VEI≥2 eruption in a specific time $t$ in the future by evaluating the formula (Marshall and Olkin, 2007)

$$P(T \leqq s + t | T > s) = 1 - \frac{1 - F(s+t)}{1 - F(s)} \tag{28}$$

where $s$ is the time that elapsed since the last eruption, which reaches till the present (Tables 6a, 6b and Fig. 6A,B). The fraction term of the

**Table 5**
Calculation of posterior eruption rates. For Villarrica Volcano: $\lambda_{\text{posterior}} = 0.0961\ \text{yr}^{-1}$; for Llaima Volcano: $\lambda_{\text{posterior}} = 0.193\ \text{yr}^{-1}$.

| Erupt rate $\lambda_i$ (yr$^{-1}$) | Duration $\Delta t$ [yr] | Prior $P(\lambda_i)$ | Likelihood $P(y\|\lambda_i)$ | Posterior $P(\lambda_i\|y)$ |
|---|---|---|---|---|
| *Villarrica* | | | | |
| 0.0769 | 129 | 0.462 | 4.54e−5 | 0.0371 |
| 0.2667 | 89 | 0.319 | 3.78e−11 | 2e−8 |
| 0.0968 | 61 | 0.219 | 2.48e−3 | 0.963 |
| *Llaima* | | | | |
| 0.3333 | 93 | 0.633 | 3.44e−14 | 2.9e−11 |
| 0.1539 | 39 | 0.265 | 2.48e−3 | 0.876 |
| 0.4667 | 15 | 0.102 | 9.12e−4 | 0.124 |

**Table 6a**
Probability of at least one VEI ≥ 2 eruption of Villarrica Volcano within the next T years, in %.

| $T$ (years) | Exp. | Weibull | Log–log | MOED | Bayes (MOED) | Non-informative |
|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | 14.6 | 18.4 | 9.7 | 8.6 | 9.2 | 13.3 |
| 2 | 27.1 | 33.5 | 18.0 | 16.5 | 17.5 | 24.8 |
| 5 | 54.6 | 64.6 | 37.2 | 35.9 | 38.2 | 50.7 |
| 10 | 79.4 | 88.1 | 57.1 | 58.6 | 61.7 | 75.4 |
| 20 | 95.8 | 98.8 | 76.5 | 82.4 | 85.4 | 93.7 |
| 50 | 100 | 100 | 92.7 | 98.6 | 99.2 | 99.9 |
| 100 | 100 | 100 | 97.6 | 100 | 100 | 100 |

**Table 6b**
Probability of at least one VEI ≥ 2 eruption of Llaima Volcano within the next T years, in %.

| T (years) | Exp. | Weibull | Log–log | MOED | Bayes (MOED) | Non-informative |
|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | 28.3 | 24.2 | 29.6 | 25.7 | 17.6 | 25.6 |
| 2 | 48.6 | 46.7 | 54.7 | 43.7 | 32.0 | 44.6 |
| 5 | 81.1 | 86.6 | 85.9 | 72.8 | 61.9 | 76.8 |
| 10 | 96.4 | 99.4 | 96.0 | 89.7 | 85.5 | 94.4 |
| 20 | 99.9 | 100 | 99.0 | 98.0 | 97.9 | 99.6 |
| 50 | 100 | 100 | 99.9 | 100 | 100 | 100 |

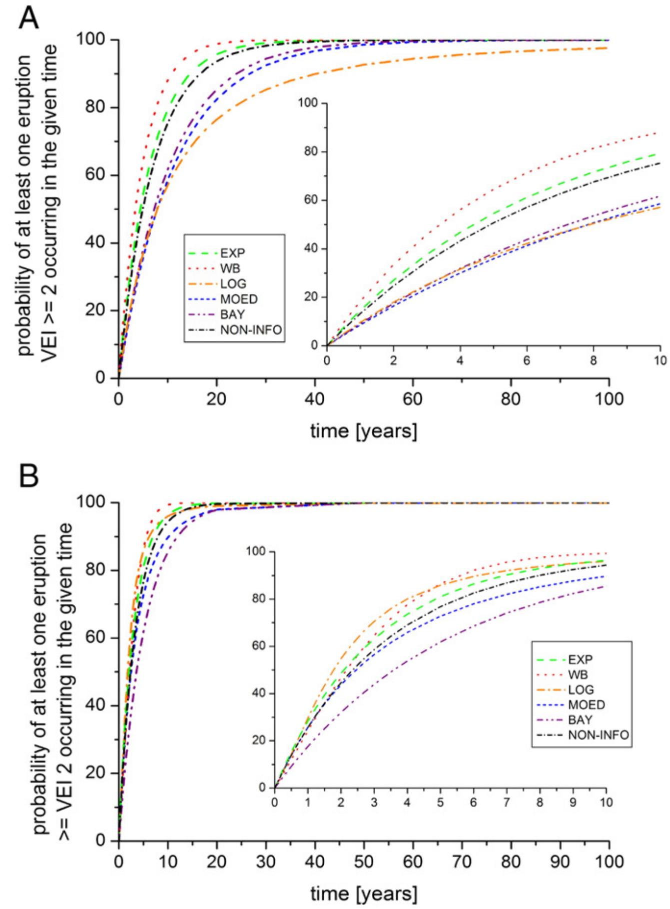


**Fig. 6.** A,B. Probability of at least one VEI ≥ 2 eruption of Villarrica (A) and Llaima (B), as a function of time, in %.

formula is the probability of no eruption occurring within time $s+t$ given our knowledge that none has occurred by time $s$ — this is the residual life distribution defined in Eq. (10). Vice versa, this can then be applied to estimating the probability that at least one eruption occurs in the given time.

In general, the Weibull distribution gives the highest probabilities of an eruption occurring within a given time in the future, and reaches values close to 100% fastest. Depending on the volcano, the exponential and log–logistic distributions give intermediate hazards, whereas the predicted time to the next eruption is longer from the MOED and Bayesian estimates.

The high hazard rate produced for both volcanoes by the Weibull distribution is primarily a result of the shape parameter $\alpha>1$, which fits particularly short repose times well, usually better than the exponential fit. The resulting Weibull distributions are both found to be increasing hazard rate ($\alpha>1$), which is by no means less plausible a priori than would be decreasing hazard rate shapes ($\alpha<1$). A somewhat stronger disagreement in the fit is observed for longer repose times, which would rather lead to a long-tailed behaviour with $\alpha<1$.

The log–logistic distribution predicts eruption likelihoods developing over larger time scales than the Weibull and exponential distribution in the near future for Villarrica, and the 100% level is approached only slowly in the long-term. Since the log-logistic distribution proved to be a good representation of the data, this points to competing processes rather contributing to relaxation of the system and a relatively higher probability of eruption triggering by simple failure. For Llaima, the log-logistic model tends to produce similar results as the Weibull and exponential distributions, in which the 100% probability is reached sooner.

The MOED gives overall lower hazards and reaches eruption probabilities close to 100% comparatively late, which is due to the fact that as a mixture of exponential distributions, the MOED itself has a decreasing hazard rate (Marshall and Olkin, 2007).

The MOED-based Bayesian estimate for the present eruption rate is in any case lower than the eruption rate obtained from an exponential distribution fit. Since both estimates for future eruption probabilities are based on the same functional form (the exponential distribution), but with different eruption rate parameter (lower for the Bayesian estimate), it is a natural consequence that the Bayesian prediction consistently gives much lower probabilities for an eruption to occur within a given time span. This takes into account the theoretical consideration that high eruption rate regimes are supposed to be shorter than low eruption rate regimes (Mendoza-Rosas and De la Cruz-Reyna, 2009) and we can therefore expect to find ourselves in a low eruption rate regime rather than in a high rate regime — particularly if the time since the last eruption has been long.

In comparison, the non-informative Bayesian prediction gives intermediate probabilities within the range of the other approaches. This is not surprising since this estimate is a smooth model generalising for the entire time span of observation and the total number of eruptions.

Such diverging predictions derived from different models should raise alertness from understanding calculated eruption hazard as unimpeachable. The fact that Llaima last erupted in 2008 makes it plausible to believe that the 1994–2008 eruption regime is still ongoing, which justifies the extrapolation of the past hazard rate to the present and future. There is, however, no confirmation on geologic grounds that a new regime is not just commencing. The unexpected Chaitén eruption that started in 2008 provides a good example that the mere absence of eruptions at a volcano for almost 10,000 years does not necessarily mean it is extinct. This is especially true for volcanoes such as the ones this study is focused on, where the large-scale driving force that acts on a much longer time scale, here the subduction of the Nazca Plate beneath the South American Plate, shows no signs of deceleration or attenuation.

## 11. Applicability of the method

The eruption time series of the very active volcanoes Villarrica and Llaima have not been successfully subjected to the implemented statistical procedure in previous investigations. De la Cruz-Reyna (1991) found globally Poisson-distributed explosive eruptive activity only for VEI $\geq$ 4 and to some degree for VEI $\geq$ 3, but observed severe shortcomings for VEI $\leq$ 2. However, since this method has not yet been ubiquitously applied to many volcanoes in the world, also possible distortions for less explosive eruptions are not thoroughly verified. In this study, we are forced to use VEI $\geq$ 2 eruptions (Villarrica) because of the scarcity of eruptions with assigned larger VEI. In the course of scrutinising the approach, we follow the philosophy that an estimate with limitations will be far more useful than no estimate at all. In addition, this work may help to reveal possible systematic deviations from the expectations, and hence, to adjust and better establish the technique. Moreover, it is questionable whether it is appropriate to simply ignore smaller eruptions from hazard considerations. We acknowledge the fact that despite the focus on eruptions with VEI $\geq$ 2, also smaller eruptions and non-explosive eruptions releasing lava flows, debouching flank collapses, triggering lahars, with an eruptive regime characterised largely by intrusions, thus leading to permanent quiescent gas emissions, yield a potential of considerable damage to people, property, and ecosystems, or even devastation of the surroundings. These smaller explosive and effusive eruptions are also able to stepwise relax the system by pressure and stress release, likely to affect the explosiveness of and the repose time elapsing until the next eruption.

## 12. Conclusions

The implementation of standard failure-analysis statistical techniques to the eruption time series of the young volcanoes Villarrica and Llaima in the Chilean Southern Volcanic Zone with historical VEI $\geq$ 2 eruption records permits to forecast the probabilities of future eruptions within given time periods. From the Weibull, exponential and log-logistic distributions, the 100% likelihood of a future eruption is generally reached in a shorter time span than from MOED-based analyses. The non-informative gamma prior for the Bayesian analysis gives intermediate probabilities. For Llaima, all distributions forecast at least one eruption with a VEI $\geq$ 2 to occur within the next 20 years at a probability of >90%, which is reached at Villarrica within 50 years.

Despite the discussed limitation of the method, we offer this approach as a useful tool for stepping towards hazard evaluation, which should obviously never build up on statistical analyses alone. Volcanic settings have been observed many times to be of untamed nature, and statistical eruption forecasting as presented in this study can do no more than illustrating the likelihood that eruptions will occur within a time slot considered. For a successful disaster prevention, management and mitigation, short-term hazard assessment in the daily use should be based on informative monitoring techniques such as regional recording of seismic activity and quiescent gas release, deformation measurements by GPS networks, constraints on magma compositions and tectonic features. Joining these sources, and supporting them by statistical forecasting, will provide a multi-parameter approach, which can indicate much more reliably and at shorter notice if a change in the activity level of the volcanoes is to be expected.

## Acknowledgements

This manuscript benefited from many constructive comments of Shane Cronin and an anonymous reviewer. This is contribution number 183 to Sonderforschungsbereich 574 “Volatiles and Fluids in Subduction Zones” at Kiel University, funded by the German Research Foundation.

## References

Akaike, H., 1973. Information Theory and an Extension of the Maximum Likelihood Principle. In: Petrov, B.N., CSAKI, F. (Eds.), Second International Symposium on Information Theory. Akademiai Kiado, Budapest, pp. 267–281.
Bebbington, M.S., 2007. Identifying volcanic regimes using Hidden Markov Models. Geophys. J. Int. 171, 921–942. doi:10.1111/j.1365-246X.2007.03559.x.
Bebbington, M., 2008. Incorporating the eruptive history in a stochastic model for volcanic eruptions. J. Volcanol. Geoth. Res. 175, 325–333. doi:10.1016/j.volgeores.2008.03.013.
Bebbington, M.S., Lai, C.D., 1996a. Statistical analysis of New Zealand volcanic occurrence data. J. Volcanol. Geoth. Res. 74, 101–110.
Bebbington, M.S., Lai, C.D., 1996b. On nonhomogenous models for volcanic eruptions. Math. Geol. 28/5, 585–600.
Burnham, K.P., Anderson, D.R., 1998. Model Selection and Inference: A Practical Information-Theoretic Approach. Springer-Verlag, New York, New York, USA. 353 pp.
Connor, C.B., Sparks, R.S.J., Mason, R.M., Bonadonna, C., 2003. Exploring links between physical and probabilistic models of volcanic eruptions: the Soufrière Hills Volcano, Montserrat. Geophys. Res. Lett. 30/13, 1701. doi:10.1029/2003GL017384.
Cox, D.R., Lewis, P.A.W., 1966. The statistical Analysis of Series of Events. Methuen & Co., London. 285 pp.
Cox, D.R., Oakes, D., 1984. Analysis of survival data. Chapman and Hall, London. 201 pp.
De la Cruz-Reyna, S., 1991. Poisson-distributed patterns of explosive activity. Bull. Volcanol. 54, 57–67.
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, New York, Berlin.
Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 2009. Bayesian Data Analysis. Chapman & Hall/CRC Texts in Statistical Science, Florida.
Gibbons, J.P., 1976. Nonparametric method for quantitative analysis. Holt, Rinehart and Winston, New York, USA.
Ho, C.-H., 1990. Bayesian analysis of volcanic eruptions. J. Volcanol. Geoth. Res. 43, 91–98.
Ho, C.-H., 1991. Time trend analysis of basaltic volcanism for the Yukka Mountain site. J. Volcanol. Geoth. Res. 46, 61–72.
Klein, F.W., 1982. Patterns of historical eruptions at Hawaiian volcanoes. J. Volcanol. Geoth. Res. 12, 1–35.
Lara, L.E., Clavero, J., 2004. Villarrica Volcano (39.5°S), Southern Andes, Chile, vol. 61. Servicio Nacional de Geología y Minería.
Lopez, L., Cembrano, J., Moreno, H., 1995. Geochemistry and tectonics of the Chilean Southern Andes basaltic Quaternary volcanism (37°–46°S). Rev. Geol. Chile 22 (2), 219–234.
Marshall, A.W., Olkin, I., 2007. Life Distributions: Structure of Nonparametric, Semiparametric, and Parametric Families, Spinger Series in Statistics.
Marzocchi, W., Sandri, L., Gasparini, P., Newhall, C., Boschi, E., 2004. Quantifying probabilities of volcanic events: the example of volcanic hazard at Mount Vesuvius. J. Geophys. Res. 109, B11201. doi:10.1029/2004JB003155.
Mendoza-Rosas, A.T., De la Cruz-Reyna, S., 2008. A statistical method linking geological and historical eruption time series for volcanic hazard estimations: application to active polygenetic volcanoes. J. Volcanol. Geoth. Res. 176, 277–290. doi:10.1016/j.volgeores.2008.04.005.
Mendoza-Rosas, A.T., De la Cruz-Reyna, S., 2009. A mixture of exponentials distribution for a simple and precise assessment of the volcanic hazard. Nat. Hazards Earth Syst. Sci. 9, 425–431.
Mulargia, F., Tinti, S., 1986. Un approccio di soluzione al probleme del punto di svolta. Statistica 46, 47–57.
Mulargia, F., Gasperini, P., Tinti, S., 1987. Identifying different regimes in eruptive activity: an application to Etna Volcano. J. Volcanol. Geoth. Res. 34, 89–106.
Muñoz, M., 1983. Eruption patterns of the Chilean volcanoes Villarrica, Llaima, and Tupungatito. Pageoph 121 (5–6), 835–852.
Naranjo, J.A., Moreno, H., 2005. Geología del Volcán Llaima, Región Araucaría & Carta Geológica de Chile Basica, No. 88.
Newhall, C.G., Self, S., 1982. The Volcanic Explosivity Index (VEI): an estimate of explosive magnitude for historical volcanism. J. Geophys. Res. 87/C2, 1231–1238.
Scarpa, R., Tilling, R.I., 1996. Monitoring and mitigation of volcanic hazards. Springer, New York, Berlin.
Sigiura, N., 1978. Further analysis of the data by akaike's information criterion and the finite corrections. Comm. Statistics—Theory and Methods 7/1, 13–26. doi:10.1080/03610927808827599.
Simkin, T., Siebert, L., 1994. Volcanoes of the World, 2nd ed. Geosciences Press, Tucson, Arizona, USA.
Solow, A.R., 2001. An empirical Bayes analysis of volcanic eruptions. Math. Geol. 33/1, 95–102.
Spiegel, M.R., Stephens, L.J., 2008. Theory and Problems of Statistics, Schaum's Outline Series, Forth edition. McGraw-Hill, New York. doi:10.1036/0071485848.
Stern, C.R., 2004. Active Andean volcanism: its geologic and tectonic setting. Rev. Geol. Chile 31–2, 161–206.
Turner, M.B., Cronin, S.J., Bebbington, M.S., Platz, T., 2008. Developing probabilistic eruption forecast for dormant volcanoes: a case study from Mt Taranaki, New Zealand. Bull. Volcanol. 70, 507–515. doi:10.1007/s00445-007-0151-4.
Varley, N., Johnson, J., Ruiz, M., Reyes, G., Martin, K., 2006. Applying statistical analysis to understanding the dynamics of volcanic explosions. In: Mader, H.M., Coles, S.G., Connor, C.B., Connor, L.J. (Eds.), Statistics in Volcanology. Special Publication of IAVCEI, 1. Geological Society, London, pp. 57–76.
Watt, S.F.L., Mather, T.A., Pyle, D.M., 2007. Vulcanian explosion cycles: patterns and predictability. Geology 35/9, 839–842. doi:10.1130/G23562A.1.
Wickman, F.E., 1966. Repose-period patterns of volcanoes. Ark. Min. Geol. 4/7, 291–367.