



POISSON PROCESS MODELLING OF THE TEMPORAL BEHAVIOUR OF VOLCANIC ERUPTIONS IN THE EAST AFRICAN RIFT SYSTEM

JAMES KINYANJUI¹, Dr. GEORGE MUHUA²

¹Email: jnkinyanjui@gmail.com

²School of Mathematics, University of Nairobi

Email: muhuageorge@gmail.com

DOI: [10.33329/bomsr.9.2.17](https://doi.org/10.33329/bomsr.9.2.17)



ABSTRACT

Volcanoes possess enormous destructive power and volcanic eruptions constitute a major natural hazard with potential for great social costs for eruptions of considerable magnitude. Understanding the temporal behaviour of volcanic eruptions is thus key to hazard assessments and prevention of future loss of life and damage to property. This paper aims to describe the temporal behaviour of volcanic eruptions in the East African Rift System in terms of a suitable Poisson model and consequently offer its predictive possibilities. The data used for the analysis was from the Smithsonian Institution's Global Volcanism Program, which is freely available on line. Three Poisson models with correspondingly unique intensity functions were chosen for analysis: the homogeneous model, the log-linear non-homogeneous model and the Weibull non-homogeneous model. The theory underpinning Poisson processes is briefly presented, the method of maximum likelihood is used to estimate the parameters of each model and the Akaike Information Criterion used to select the optimal model. The log-linear non-homogeneous Poisson model is found to best fit the empirical data with 95% confidence and based on it forecasts are issued which show an increase in eruptive activity.

Keywords: Homogeneous Poisson Process; Log-Linear Non-Homogeneous Poisson Process; Weibull Non-Homogeneous Poisson Process; Maximum Likelihood Estimation

1 Introduction

East Africa sits astride an area of seismic and geological importance with a significant number of active and dormant volcanoes. Of the 148 volcanoes found in Africa, 120 are found in the East African Rift System. Ethiopia contains 59 volcanoes followed in second place by Kenya with 22 volcanoes (Global distribution of volcanism: Regional and country profiles, 2015). Eruptions of considerable size are rare events yet the when they do occur they have the potential for devastating consequences. This was demonstrated in spectacular fashion in June 2011 when Nabro in Eritrea animated and erupted violently despite having had no historic eruptions and being thought to be extinct by the scientific community. The resulting ash cloud was dispersed northwesterly, disrupting air travel throughout the Horn of Africa and the Middle East. The eruption caused thousands to be evacuated and led to some fatalities (Global distribution of volcanism: Regional and country profiles, 2015). A volcanic eruption in a more densely populated area would have wreaked more havoc. Yet this is not a unique situation. Indeed, most volcanoes pose the greatest hazard over considerably long time scales in the order of decades and centuries; at longer time scales they have the potential for global impact and catastrophe. Even volcanoes thought to be dormant or extinct can suddenly erupt with little warning (as was the case with Nabro). The problem, then, becomes building probabilistic forecasts that account for this long-scale uncertainty using potential eruption scenarios and relevant data. An important consideration is that the historical record is short, biased and incomplete. The instrumented record is even more problematic, being shorter and, for most volcanoes, spanning only the last few decades of uninterrupted surveillance — a infinitesimal fraction of their long lifetime.

Many authors have contributed to the field of statistical volcanology with respect to modelling repose times. Poisson models seem to be the most popular among many authors despite the merits of other models because of the ease of their use and their ability to capture effectively the count and temporal aspect of volcanic activity. Wickman (1966) in his pioneering works introduced the idea of simple Poissonian behaviour of volcanoes. The Poisson process was defined as a model for describing random temporal-spatial events. He noted that certain volcanoes showed eruptive rates that were independent of time and were thus memoryless: past events had no bearing on future events. Such volcanoes were described as Simple Poissonian Volcanoes. This lack of memory naturally implied the use of an exponential distribution to model inter-event times. The exponential distribution is completely defined by its single parameter λ , which represents the rate of occurrence, which is constant in the case of stationarity. Other studies that used the homogeneous Poisson process include those of Klein (1982), Mulargia et al (1985), De la Cruz-Reyna (1991) and Dzierma and Wehrmann (2010). Reymont (1969) is unique in that he proposed a log-linear intensity function to model non-stationary eruptive series. Other studies came to the conclusion that non-homogeneous models were more appropriate in describing anomalous eruptive behaviour. These include the studies by Salvi et al (2006) and Smethurst et al (2009). Authors like Ho (1991), Bebbington and Lai (1996a, 1996b), Watts et al (2007) and Dzierma and Wehrmann (2010) considered Weibull Poisson models in their statistical analyses. This paper builds on all these efforts, providing an African perspective, of which the authors found no modelling enterprise. Three specific models used were selected for analysis based on their prevalent use and parsimony: the homogeneous model based on the need to incorporate a stationary model that could capture any underlying constant trend in the data especially in light of the potential that the data considered was incomplete and the log-linear and Weibull non-homogeneous models based on the the assumption the data set was complete and inherently non-stationary coupled with the need to capture any trend observed in the data. The Weibull model is particularly useful in

determining if volcanic eruptions can be termed Poissonian in behaviour because it incorporates a parameter that indicates waxing and waning of activity.

2 Data

The sampling frame for the sample used for the paper is a global database of documented eruptions known to have occurred over the last 12,500 years compiled by the Smithsonian Institution's Global Volcanism Program (<http://www.ieor.columbia.edu/>). This catalogue contains all documented geological and historical-observation eruptions known to have occurred and consists of a set of just under 10,000 eruptions from approximately 10,500 B.C.E. till present day. This paper limits itself in scope to eruptions from 1 January 1919 to 1 January 2019, encompassing 13 volcanoes that have erupted in this period out of a total of the 120 volcanoes found in the East African Rift (see Appendix). The volcanoes are spread out across eight countries located in the East African Rift System: Ethiopia, Eritrea, Djibouti, Kenya, Uganda, Tanzania, Rwanda, and the Democratic Republic of Congo. The database contains a variety of information related to eruptions; however, the variable of interest was date of onset¹, which was either given in exact or approximated form. The completeness of the data is beyond the scope of this paper but it is assumed that for purposes of analysis a fairly accurate snapshot is proffered.

3 Methodology

3.1 Homogeneous Poisson Process

A simple Poisson process is a mathematical model that describes a temporal-spatial series of events that occur randomly and independent of each other. The broad characteristics of a homogeneous Poisson process are: events occur singly with probability of near zero that two events occurring simultaneously; the rate of occurrence of events is constant; the probability of future events is independent of the past; and lack of time trend, i.e., stationarity. One useful characteristic to investigate is the distribution of inter-event times, which for a homogeneous Poisson process have an exponential distribution which is completely defined by a single parameter commonly denoted by λ , which represents the rate of occurrence of events or arrivals (Cox & Lewis, 1966).

Definition 1. A collection of random variables $\{N(t): t \in [0; \infty)\}$ indexed by time t is called a continuous- time stochastic process. Further, such a stochastic process is a (homogeneous) Poisson process if:

- (a) starting from $N(0) = 0$ the process $N(t)$ takes non-negative integers $0, 1, 2, \dots$ for all $t \geq 0$;
- (b) the increment $N(t + s) - N(t)$ is surely non-negative for any $s > 0$;
- (c) the increments $N(t_1), N(t_2) - N(t_1), N(t_n) - N(t_{n-1})$ are independent for any $0 < t_1 < t_2 \dots \dots \dots t_{n-1} < t_n$;
- (d) the increment $N(t + s) - N(t)$ has a distribution which is dependent on the values $s > 0$ but independent of $t > 0$; and
- (e) the increment $N(t + s) - N(t)$ has a Poisson distribution with mean λ_t , i.e., for any $s, t \geq 0$.

¹ The onset dates for the 69 events considered were transformed into a series of inter-event times and cumulative event times for purposes of modelling

$$\Pr(N(t+s) - N(s) = n) = \frac{(\lambda t)^n \exp(-\lambda t)}{n!} \quad (1)$$

The definition given by Ross (2010) lays the foundation for looking at the Poisson process as an integrated collection of three random variables: a counting process, a sequence of arrival or onset times and a sequence of inter-arrival or inter-event times.

A stochastic process satisfying (a) and (b) is called a counting process in which $N(t)$ represents the total number of 'events' (from here onwards 'events' will refer to volcanic eruptions). Properties (c) and (d) are respectively called the independent and stationary increments.

Events counted by a Poisson process $\{N(t), t \geq 0\}$ are called Poisson events. Now, let T_n denote the time when the n -th Poisson event occurs. T_n is called the arrival, event, onset or occurrence time (in this case the onset of an eruption) and we can then define the inter-arrival, inter-event or (in a volcanological context) repose times W_n as

$$W_n = T_n - T_{n-1}; \quad n = 1, 2, \dots \quad (2)$$

where $T_0 = 0$ by convention and for convenience.

3.2 Non-Homogeneous Poisson Process

The restrictions in the properties of the homogeneous Poisson process make it inadequate for many other real world systems and natural phenomena which are prone to wild unpredictability but nonetheless possess Poisson-like characteristics with parameters dependent on time. Such systems can be described by non-homogeneous (or non-stationary) Poisson processes. While Poisson models have been found to be adequate in certain volcanic systems, universal application for all volcanic systems is found to be untenable as certain data sets show considerable deviation. The introduction of the non-homogeneous Poisson process as a generalization of the homogeneous Poisson model allows for some of this randomness to be captured. A non-homogeneous Poisson process satisfies the same assumptions as a homogeneous Poisson process but with λ dependent on time, i.e., $\lambda(t)$. Utilizing a non-homogeneous time-dependent Poisson process in the context of volcanic activity implies that a number of underlying processes conflate together and the balance of these processes is a function of time (Sanchez, 2014). For a NHPP the time-dependent intensity function $\lambda(t)$ takes different forms.

Definition 2. The counting process $\{N(t), t \geq 0\}$ is said to be a NHPP with intensity function $\lambda(t), t \geq 0$ if it satisfies the following conditions:

- (a) $N(0) = 0$ almost certainly;
- (b) $N(t)$ has independent increments;
- (c) $\Pr(N(t+h) - N(t) = 0) = 1 - \lambda(t)h + o(h)$;
- (d) $\Pr(N(t+h) - N(t) = 1) = \lambda(t)h + o(h)$;
- (e) $\Pr(N(t+h) - N(t) \geq 2) = o(h)$.

This definition from Ross (2010) introduces another way of defining a Poisson process. Parts (c) and (d) of the definition may look awkward at first sight but are, in fact, insightful and intuitive. They state that having two or more events in a small time interval is extremely unlikely while the probability of a single event is approximately proportional to the length of that small interval. The

notation $o(h)$ refers to some function g for which $\lim_{h \rightarrow 0} \frac{g(h)}{h} = 0$. The intensity function $\lambda(t)$ is a function of time and is often called the instantaneous arrival rate.

The distribution of the number of events in an interval is as follows:

$$Pr(N(t + s) - N(t) = n) = \frac{[\Lambda(t+s) - \Lambda(t)]^n e^{-[\Lambda(t+s) - \Lambda(t)]}}{n!} \tag{3}$$

This logically follows from the HPP since both processes have independent increments. This means that the distribution of $N(t + s) - N(t)$ is, in fact, Poisson with parameter $\Lambda(t + s) - \Lambda(t)$.

The relationship between the average number of events which occur in the interval $(0, t]$ and the intensity function can be expressed as

$$m(t) = E[N] = \int_0^t \lambda(u) du = \Lambda(t) - \Lambda(0) \tag{4}$$

This expectation function $m(t)$ completely defines the NHPP and is a monotonic non-decreasing right-continuous function such that

$$0 \leq \int_0^t \lambda(u) du < \infty$$

for all bounded subsets \mathbb{R} of the state space S of the process.

This concept can be extended to the number of events between times t and $t + s$ to yield

$$E[N(t + s) - N(t)] = \int_0^{t+s} \lambda(u) du = \Lambda(t + s) - \Lambda(t) \tag{5}$$

The above results were adopted from *C, inlar (2013)*.

For a NHPP $\{N(t), t \geq 0\}$ with intensity function $\lambda(t)$, the integrated function between two successive event times (which is, in fact, the mean function) T_n and T_{n+1} follows an exponential distribution with unit mean, i.e.,

$$m(T_n, T_{n+1}) = \int_{T_n}^{T_{n+1}} \lambda(t) dt \tag{6}$$

In addition, as a result of the independent increments property in non-overlapping intervals, $m(T_0, T_1), m(T_1, T_2), \dots, m(T_{n-1}, T_n)$ are iid exponential random variables (Smethurst, 2009). This serves as a useful link between the HPP and NHPP, which can be exploited, for example, if NHPP event times are to be converted to HPP event times.

The final set of results are adopted from *Cox & Lewis (1966)*.

For a NHPP $\{N(t), t \geq 0\}$ with mean function $m(t)$ with an intensity function $\lambda(t)$ which is absolutely continuous, the arrival times t_1, t_2, \dots, t_n for n observed events in the interval $(0, T]$ are distributed as order statistics from a sample with probability density function

$$f(t) = \frac{\lambda(t)}{\Lambda(T) - \Lambda(0)} \tag{7}$$

We use the equality $Wn > t \Leftrightarrow N(t) < n$ and observe that $Pr(T > t) = Pr(N(t + s) - N(t) = 0)$, i.e., time of the next arrival from start of observation is greater than t only if there is no event in the interval $(t, t + s]$. Using Eq. 3 it can be seen that for any $t, s \geq 0$

$$Pr(N(t + s) - N(t) = 0) = e^{-[\Lambda(t+s) - \Lambda(t)]} \tag{8}$$

and the cdf of arrival time becomes

$$F_T(t) = 1 - e^{-[\Lambda(t+s)-\Lambda(t)]} \tag{9}$$

To obtain the density function of conditional arrival time we get the derivative of the cdf wrt to s.

$$\frac{d}{ds} F_T(t) = \lambda(t + s)e^{-[\Lambda(t+s)-\Lambda(t)]} \tag{10}$$

For n observed events in the interval (0,T] at times t_1, t_2, \dots, t_n the joint density becomes

$$\begin{aligned} &\lambda(t_1)e^{-[\Lambda(t_1)-\Lambda(0)]}\lambda(t_2)e^{-[\Lambda(t_2)-\Lambda(t_1)]} \dots \dots \dots \lambda(t_n)e^{-[\Lambda(t_n)-\Lambda(t_{n-1})]}e^{-[\Lambda(T)-\Lambda(t_n)]} \\ &= e^{-[\Lambda(T)-\Lambda(0)]} \prod_{i=1}^n \lambda(t_i) \\ &= e^{-\int_0^T \lambda(u)du} \prod_{i=1}^n \lambda(t_i) \end{aligned} \tag{11}$$

where the term $e^{-[\Lambda(T)-\Lambda(t_n)]}$ denotes the probability that no event occurs in the interval $(t_n, T]$.

Eq. 11 can also be expressed as

$$[\Lambda(T) - \Lambda(0)]^n e^{-[\Lambda(T)-\Lambda(0)]} \prod_{i=1}^n \frac{\lambda(t_i)}{[\Lambda(T) - \Lambda(0)]^n} \tag{12}$$

if we consider unordered event times.

This, then, yields the conditional density function of T_i as shown below.

$$f_T(t_i | N(t) = n) = \frac{\lambda(t)}{\Lambda(T) - \Lambda(0)}; i = 1, 2, \dots, n \tag{13}$$

3.3 Parameter Estimation

Parameters of each model were obtained using MLE. MLE involves optimizing the likelihood function with the goal of estimating parameters which make it more probable to observe the given data. The advantage of MLE it takes into account the real distribution of the data and is robust in case of deviation from normality, a key assumption of OLS estimation (Myung, 2003).

Let us consider a random sample from an unknown population. MLE attempts to make inference about the population that generated that sample. Assume we have a set of iid random variables (t_1, t_2, \dots, t_n) , each indexed by a unique parameter vector $\underline{\theta} = \{\theta_1, \theta_2, \dots, \theta_k\}$ whose values can lie anywhere in the parameter space Θ . To obtain the ML estimates of θ we need to get the likelihood function, which is the joint distribution of the observed sample.

$$\mathcal{L}(\underline{\theta}; t) = \prod_{i=1}^N f_i(t_i; \underline{\theta}) \tag{14}$$

The goal of MLE is to find the specific values that maximize the likelihood function over the parameter space, i.e,

$$\hat{\underline{\theta}} = \underset{\theta \in \Theta}{argmax} \mathcal{L}(\underline{\theta}; t) \tag{15}$$

This maximum point is called the maximum likelihood estimate.

This entails the selection of parameter values that make the observed data most probable. It is customary and convenient to deal with the log-likelihood, the natural logarithm of the likelihood function.

$$\ell(\underline{\theta}; t) = \mathcal{L}(\underline{\theta}; t) \tag{16}$$

If $\ell(\underline{\theta}; t)$ is a differentiable function then the maxima are the solutions to the likelihood equations obtained by getting the derivative with respect to θ and setting the results to zero, i.e.,

$$\frac{\partial \ell}{\partial \theta_1} = 0, \frac{\partial \ell}{\partial \theta_2} = 0 \dots \dots \dots \frac{\partial \ell}{\partial \theta_k} = 0 \tag{17}$$

The joint distribution of arrival times conditional on number of arrivals was derived (see Eq. 11) and was found to be as follows:

$$f_T(t_i | N(t) = n) = e^{-\int_0^T \lambda(u)du} \prod_{i=1}^n \lambda(t_i) \tag{18}$$

For HPP with intensity function $\lambda(t) = \lambda$ the ML estimator becomes

$$\hat{\lambda} = \frac{n}{T} \tag{19}$$

For a log-linear NHPP with intensity function $\lambda(t) = e^{\gamma_0 + \gamma_1 t}$ the ML estimators become

$$\hat{\gamma}_0 = \ln\left(\frac{n\hat{\gamma}_1}{e^{\hat{\gamma}_1 T} - 1}\right); \sum_{i=1}^n t_i + \frac{n}{\hat{\gamma}_1} = \frac{nTe^{\hat{\gamma}_1 T}}{e^{\hat{\gamma}_1 T} - 1} \tag{20}$$

which have no closed-form solution.

For a Weibull NHPP with intensity function $\lambda(t) = \beta/\eta)(t/\eta)^{\beta-1}$ the ML estimators become

$$\hat{\beta} = \frac{T}{\sum_{i=1}^n \ln(\frac{T}{t_i})}; \hat{\eta} = T/n^{1/\hat{\beta}} \tag{21}$$

3.4 Akaike Information Criterion

The AIC is an estimator of resampling prediction error and therefore a measure of the relative quality of a statistical model for a given set of data. A statistical models can never perfectly represent the process that generated a sample; inevitably some information gets lost in the process. AIC estimates the relative amount of information lost by a given model and gives a score expressing this loss: the less information a model loses, the better the quality of that model. In determining the amount of information lost AIC performs a balancing act between model fit (as determined by maximized likelihood) and parsimony (as determined by k , the dimension of the parameter vector), penalizing both overfitting and underfitting. If a number of models are considered, then, in the most simplistic sense, the model with the lowest score is the one selected (Aho et al, 2014).

Let us a consider a model with k parameters and let L be the value of the maximized likelihood of the model. The AIC score of the model is given below.

$$AIC = -2\ln L(\hat{\theta}) + 2k \tag{22}$$

3.5 Goodness of Fit

The Kolmogorov-Smirnov test (K-S test) is a formal statistical test used to augment the customary plots used to check goodness of fit. The K-S test is a non-parametric and agnostic test used to detect differences between distributions. It examines a single maximum difference between distributions. If a statistical difference exists, the test does not provide insight into the cause of the difference nor does it indicate the nature of the common distribution if there is no statistical difference between the two distributions. The differences could be as a result of difference in: location; variation; skewness; kurtosis; and modality, or presence of outliers, among other things (Daniel, 1990).

4 Results

The first task was to produce a plot to have a sense of how the eruptions were arranged in time. While more mathematically rigorous methods exist for delineating eruptive regimes (for example Mulargia et al (1987)), a visual representation can also reveal much about the nature of eruptive activity.

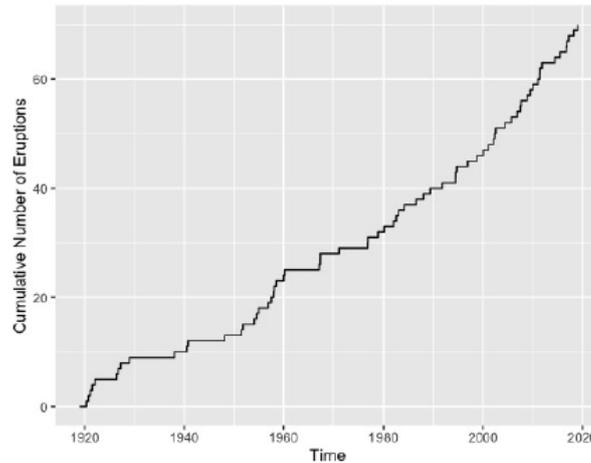


Figure 1: Step Plot of Cumulative Number of Eruptions against Time

The data shows two noticeable regimes: a regime featuring a number of long repose and another (starting from around 1980) dominated by shorter repose. The slight curvature was an indication that the eruption rate was non-constant. To check if this was the case the intensity was plotted as a function of time using a non-parametric method with points smoothed out using a Gaussian kernel function.

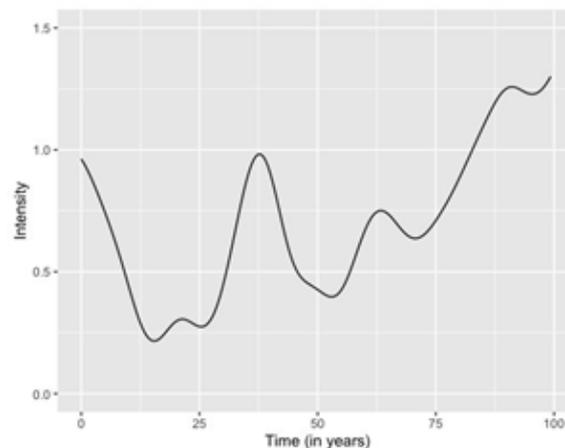


Figure 2: Line plot of Empirical Intensity

4.1 Model Fitting

MLE was used to obtain parameter estimates for the three models (R codes used in fitting the data adopted from Gelissen (2016a)). The results are shown in the table below.

Table 1: Results of Model Fitting

	Parameter Estimates	ℓ	AIC
HPP	$\hat{\lambda} = 0.69$	-94.6034	191.2068
Log-Linear NHPP	$\hat{\gamma}_0 = -0.911036781$ $\hat{\gamma}_1 = 0.009976769$	-91.8111	187.6222
Weibull NHPP	$\hat{\beta} = 1.167440$ $\hat{\eta} = 2.660038$	-93.8176	191.6352

Based on the AIC the best model is the log-linear NHPP. The fitted models and empirical data were plotted on the same axes to see if the choice of the log-linear NHPP was justified.

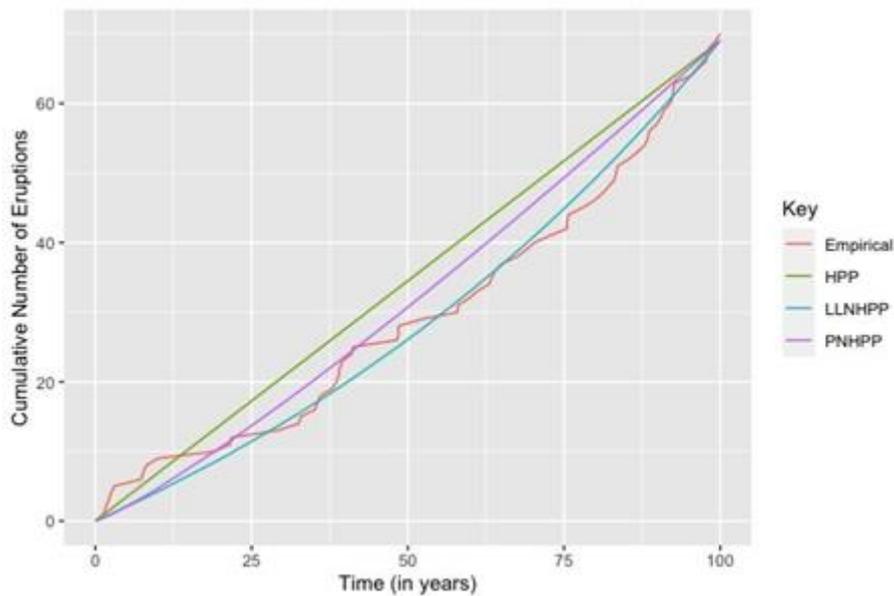


Figure 3: Plot of Cumulative Number of Eruptions Against Observed and Fitted Event Times

The plot shows that the log-linear NHPP was the best model of the three considered: it gave the best fit to the data.

4.2 Goodness of Fit

The log-linear intensity function was used to simulate a set of NHPP event times (R code used in the simulation adopted from Gelissen (2016b)). The eruption data was compared with the simulated data to check how well the model performed. This was done visually and through the K-S test.

The plot showed a fairly good fit. The K-S test was performed to confirm the visual conclusion. The null hypothesis was that the two event times were similar in distribution. The K-S test statistic was $d = 0.0875$ compared against a critical value of $d = 0.236$ (p value = 0.9315). The null hypothesis was therefore not rejected and the conclusion drawn was that the two event times have similar distributions at 5% level of significance.

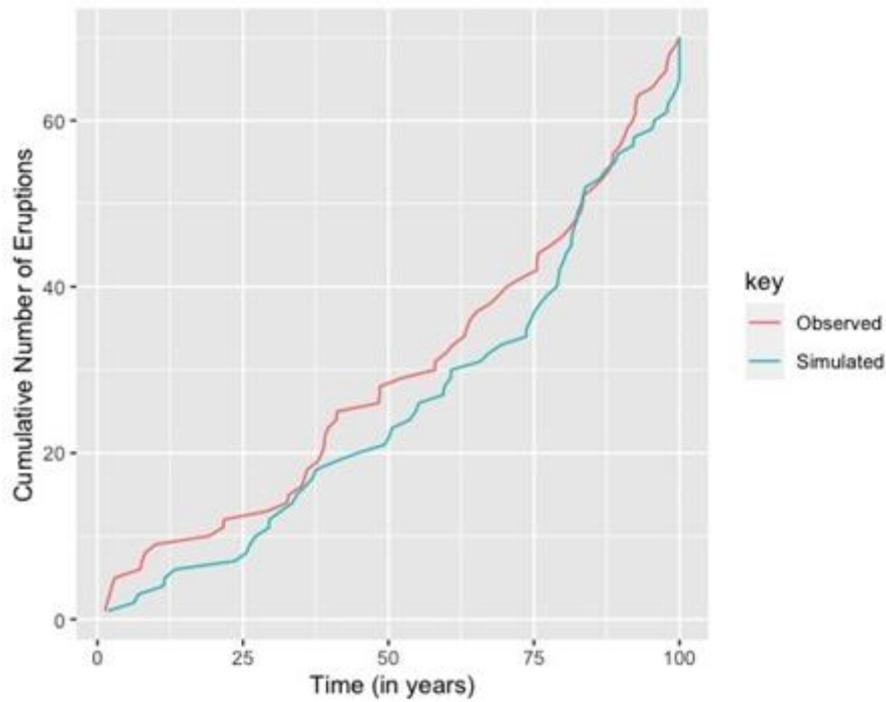


Figure 4: Plot of Cumulative Number of Eruptions against Observed and Simulated Event Times

4.3 Forecasting

The log-linear model was then used to forecast the cumulative number of eruptions in the next hundred years, together with confidence intervals for the estimates. The plot shown below illustrates this.

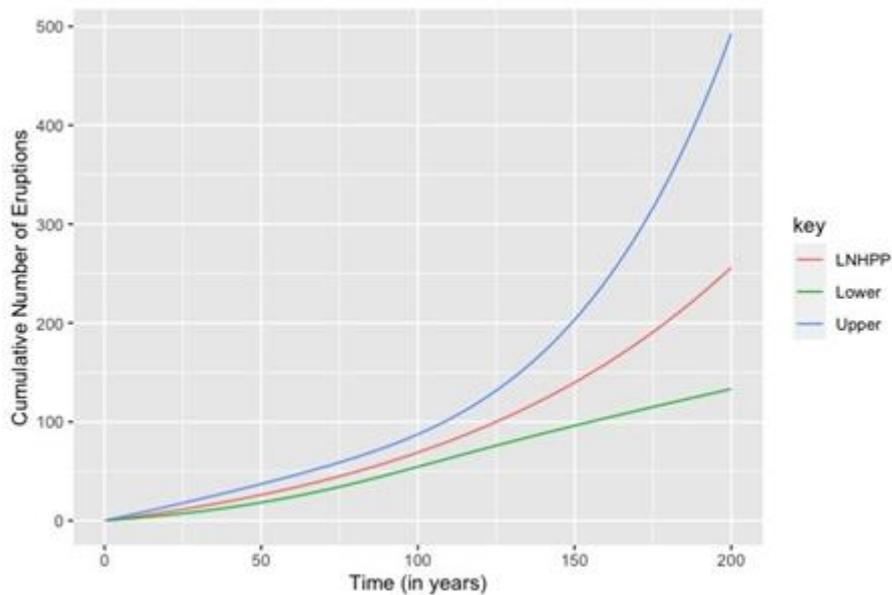


Figure 5: Plot of Cumulative Number of Eruptions Against Times

We might want, for example, to predict the number of eruptions between January 2019 and December 2043, i.e., $E[N(125) - N(100)]$. The model predicts that there will be about 31 (whole-number approximate of 30.96372) eruptions. This result was obtained by using Eq. 5. The confidence interval of this estimate is (18.49463, 51.83949). The standard errors were calculated by R using the delta method formula, i.e.,

$$Var [\hat{\Lambda}(t)] = \left(\frac{\partial \hat{\Lambda}(t)}{\partial \underline{\theta}} \Big|_{\underline{\theta}=\hat{\underline{\theta}}} \right)^2 Var [\hat{\lambda}(t)]$$

$Var [\hat{\Lambda}(t)]$ is, in fact, the inverse of the *Hessian* (a matrix of second-order partial derivatives of the log-likelihood) and the square root of the diagonal gives the standard errors of the parameter estimates. By the partitioning of a Poisson process a forecast on the number of eruptions of a particular volcano and of a particular VEI² can be issued. The number of observed eruptions for Nyamuragira, for example, are 33 and so the model predicts it will have approximately 15 eruptions [(33/69)*31] from January 2019 to December 2043. The number of eruptions of VEI≥3 is 5 and so the model predicts approximately 6 eruptions [(13/69)*31] in the next 25 years.

Probabilities can also be computed for a particular number of eruptions over an interval of choice. For instance, the probability of two or more eruptions from January 2019 to December 2021, i.e., $Pr(N(103) - N(100) \geq 2) \approx 0.8439$. This result was calculated using Eqs. 5 and 3. The plot below illustrates predictions for three different cumulative eruptions.

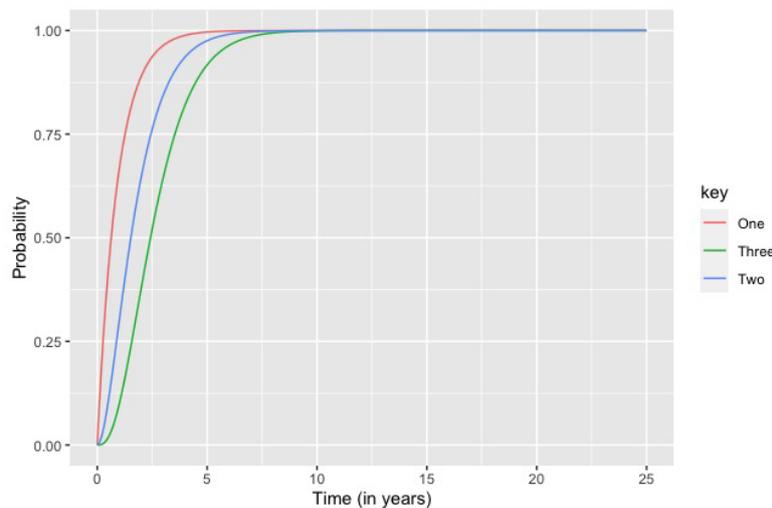


Figure 6: Plot of Predicted Probabilities for a 25-Year Period for One or More, Two or More and Three or More Eruptions

5 Conclusion

This paper sought to find a Poisson model most appropriate in describing a curtailed catalogue of eruptions. Of the three models considered a NHPP with a log-linear intensity function was found to best explain the data. The most parsimonious models are usually preferred and in this case the most parsimonious model considered was the single-parameter HPP. However, it was not justified with respect to the model results. In fact, a cursory examination of the cumulative reposes showed the data was likely non-stationary, making the HPP the least obvious candidate for selection as a model. While the quality of the eruption data was not called into question because it was beyond the scope

² The Volcanic Explosivity Index is a relative measure of the explosiveness of volcanic eruptions developed by Newhall and Self in 1982. The volume of products, eruption cloud height, and qualitative observations (using terms ranging from 'gentle' to 'mega-colossal') are used to determine the explosivity value. The scale is open-ended with the largest volcanic eruptions in history given magnitude 8. A value of 0 is given for non-explosive eruptions, defined as less than 10,000 cubic metres of tephra ejected and 8 representing a mega-colossal explosive eruption that can eject a trillion cubic metres of tephra and have a cloud column height of over 20 km. The scale is logarithmic (similar to the Richter scale used in the study of earthquakes), with each interval on the scale representing a tenfold increase in observed ejecta, with the exception of between VEI-0, VEI-1 and VEI-2.

of the paper, the issue of incomplete eruption records, which usually favour rejection of stationary models (Dzierma and Wehrmann (2010), for example, discuss the issue of volcanic data and stationarity), is a long-standing concern in statistical volcanology. Indeed, the data considered showed a number of long repose in the early record. The Weibull model performed the poorest. Of note, however, is the fact that the shape parameter $\beta = 1.167440$. Ho (1991), for example, identified the shape parameter of his Weibull process model as the indicator of waxing or waning of eruptive activity. Because $\beta > 1$, it can be concluded that there was increase in volcanic eruptions with time. This coincides with the increasing trend that the log-linear model predicts. The forecasting possibilities of the log-linear model were demonstrated with the model predicting of an increase in eruptive activity with time though the monotonic-increasing nature of the model (in general contravention of the real-life observable behaviour of volcanoes) means that only short-term forecasting of, say, one or two decades, will give plausible results.

Acknowledgements: Special thanks to Leonard Salasya for checking my output and Dr James Tuju for proofreading.

References

- [1]. Aho, K., Derryberry, D., & Peterson, T. (2014). Model selection for ecologists: the worldviews of AIC and BIC. *Ecology*, 95(3), 631–636.
- [2]. Bebbington, M. S., & Lai, C. D. (1996a). On non-homogeneous models for volcanic eruptions. *Math. Geol.*, 28(5), 585–600.
- [3]. Bebbington, M. S., & Lai, C. D. (1996b). Statistical analysis of New Zealand volcanic occurrence data. *J. Volcanol. Geotherm. Res.*, 74(1-2), 101–110.
- [4]. Brown, S. K., Sparks, R. S. J., Mee, K., Vye-Brown, C., Ilyinskaya, E., Jenkins, S., Loughlin, S. C., et al. (2015). *Global distribution of volcanism: Regional and country profiles. Report IV of the GVM/IAVCEI contribution to the UN-ISDR global assessment report on disaster risk reduction 2015*. Global Volcano Model and International Association of Volcanology and Chemistry of the Earth's Interior.
- [5]. C, inlar, E (2013). *Introduction to stochastic processes*. Dover Publications.
- [6]. Cox, D. R., & Lewis, P. A. W. (1966). *The statistical analysis of series of events*. John Wiley & Sons.
- [7]. Daniel, W. W. (1990). *Applied nonparametric statistics* (2nd Ed.). PWS-Kent.
- [8]. De la Cruz-Reyna, S. (1991). Poisson-distributed patterns of explosive eruptive activity. *Bulletin of Volcanol.*, 54(1), 57–67.
- [9]. Dzierma, Y., & Wehrmann, H. (2010). Statistical eruption forecast for the Chilean southern volcanic zone: typical frequencies of volcanic eruptions as baseline for possibly enhanced activity following the large 2010 Concepcion earthquake. *Nat. Hazards. Earth Sys. Sci.*, 10, 2093–2108.
- [10]. Gelissen, S. (2016a). *R code for fitting a nonhomogeneous temporal Poisson process model* [R Source code]. <https://www.blogs2.datall-analyse.nl>
- [11]. Gelissen, S. (2016b). *R code for fitting a nonhomogeneous temporal Poisson process model using the spatstat package* [R Source code]. <https://www.blogs2.datall-analyse.nl>
- [12]. Ho, C. H. (1991). Non-homogeneous Poisson model for volcanic eruptions. *Math. Geol.*, 23(2), 167–173.
- [13]. Klein, F. W. (1982). Patterns of historical eruptions at Hawaiian volcanos. *J. Volcanol. Geotherm. Res.*, 12(1-2), 1–35.
- [14]. Mulargia, F., Tinti, S., & Boschi, E. (1985). A statistical analysis of flank eruptions on Etna volcano. *J. Volc. Geotherm. Res.*, 23, 263–272.

- [15]. 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.
- [16]. Myung, I. J. (2003). Tutorial on Maximum Likelihood Estimation. *Journal of Mathematical Psychology*, 47(1), 90–100.
- [17]. Newhall, C., & Self, S. (1982). The Volcanic Explosivity Index (VEI): An estimate of explosive magnitude for historical volcanism. *Journal of Geophysical Research*, 87(C2), 1231–1238.
- [18]. Reyment, R. A. (1969). Statistical analysis of some volcanologic data regarded as series of point events. *Pure Appl. Geophys.*, 74(1), 57–77.
- [19]. Ross, S. M. (2010). *Introduction to probability models* (10th ed.). Academic Press.
- [20]. Salvi, F., Scandone, R., & Palma, C. (2006). Statistical analysis of the historical activity of Mount Etna, aimed at the evaluation of volcanic hazard. *J. Volcanol. Geotherm. Res.*, 154, 159–168.
- [21]. Sanchez, L. A. (2014). *Statistical analysis and computer modelling of volcanic eruptions* (Doctoral dissertation). <https://ir.lib.uwo.ca/etd/1912>
- [22]. Smethurst, L., James, M. R., Pinkerton, H., & Tawn, J. A. (2009). A statistical analysis of eruptive activity on Mount Etna, Sicily. *Geophys. J. Int.*, 179, 655–666.
- [23]. Watt, S. F. L., Mather, T. A., & Pyle, D. M. (2007). Vulcanian explosion cycles: Patterns and predictability, *Geology*, 35(9), 839–842.
- [24]. Wickman, F. E. (1966). Repose-period patterns of volcanoes. *Ark. Kem. Mineral. Geol.*, 4(4), 291- 367.

Appendix: List of Volcanoes

Name	Location	Number of Eruptions*	Date of Last Eruption*
Alu-Dalafilla	Ethiopia	1	03-11-2008
Ardoukoba	Djibouti	1	07-11-1978
The Barrier	Kenya	1	31-12-1921
Dabbahu	Ethiopia	1	26-09-2005
Dalol	Ethiopia	2	04-01-2011
Erta Ale	Ethiopia	3	02-07-1967
Ol Doinyo Lengai	Tanzania	15	09-04-2017
Manda Hararo	Ethiopia	2	28-06-2009
Manda-Inakir	Djibouti/Ethiopia	1	31-12-1928
Nabro	Eritrea	1	13-06-2011
Nyamuragira	DRC	33	18-04-2018
Nyiragongo	DRC	7	17-05-2002
Visoke	DRC/Rwanda	1	01-08-1957

*for time period considered