A New Bayesian Time-Predictable Modeling of Eruption Occurrences and Forecasting: Application to Mt Etna volcano and Kilauea volcano

L Passarelli[a,b,*], B. Sansò[b], L. Sandri[a], W. Marzocchi[c]

[a]*Istituto Nazionale di Geofisica e Vulcanologia, sezione di Bologna via M.A. Franceschini, 31 40128 Bologna Italy*
[b]*University of California, Santa Cruz 1156 High Street Santa Cruz CA, USA*
[c]*Istituto Nazionale di Geofisica e Vulcanologia, sezione di Roma, via di Vigna Murata, 605 00143 Roma Italy*

Abstract

In this paper we propose a model to forecast eruptions in a real forward perspective. Specifically, the model provides a forecast of the next eruption after the end of the last one, using only the data available up to that time. We focus our attention on volcanoes with open conduit regime and high eruption frequency. We assume a generalization of the classical time predictable model to describe the eruptive behavior of open conduit volcanoes and we use a Bayesian hierarchical model to make probabilistic forecasts. We apply the model to Kilauea volcano eruptive data and Mount Etna volcano flank eruption data.

The aims of the proposed model are: 1) to test whether or not the Kilauea and Mount Etna volcanoes follow a time predictable behavior; 2) to discuss the volcanological implications of the time predictable model parameters inferred; 3) to compare the forecast capabilities of this model with other models present in literature. The results obtained using the MCMC sampling algorithm show that both volcanoes follow a time predictable behavior. The numerical values inferred for the parameters of the time predictable model suggest that the amount of the erupted volume could change the dynamics of the magma chamber refilling process during the repose period. The probability gain of this model compared with other models already present in literature is appreciably greater than zero. This means that our model provides better forecast than previous models and it could be used in a probabilistic volcanic hazard assessment scheme.

Keywords: Effusive volcanism, Bayesian hierarchical modeling, Mount Etna, Kilauea, Probabilistic forecasting, Volcanic hazards.

---

*Corresponding author:
Istituto Nazionale di Geofisica e Vulcanologia, sezione di Bologna via M.A. Franceschini, 31 40128 Bologna Italy, phone:(+39)0514151520, fax:(+39)0514151498, e-mail:passarelli@bo.ingv.it

Email addresses: passarelli@bo.ingv.it (L Passarelli), bruno@ams.ucsc.edu (B. Sansò), sandri@bo.ingv.it (L. Sandri), warner.marzocchi@ingv.it (W. Marzocchi)

1. Introduction

One of the main goals in modern volcanology is to provide reliable forecast of volcanic eruptions with the aim of mitigating the associated risk. The extreme complexity and non linearity of a volcanic system make deterministic prediction of the evolution of volcanic processes rather impossible (e.g. Marzocchi 1996; Sparks 2003). Volcanic systems are intrinsically stochastic. In general, eruption forecasting involves two different time scales: i) a *short-term* forecasting, mostly based on monitoring measures observed during an episode of unrest (e.g., Newhall and Hoblitt 2002, Marzocchi *et al.* 2008 among others); ii) a *long-term* forecasting, usually made during a quiet period of the volcano, and mostly related to a statistical description of the past eruptive catalogs (e.g. Klein, 1982, Bebbington and Lai, 1996a among others). Here, we focus our attention only on this second issue.

In a *long-term* eruption forecast perspective we believe that an incisive and useful forecast should be made before the onset of a volcanic eruption, using the data available at that time, with the aim of mitigating the associated volcanic risk. In other words, models implemented with forecast purposes have to allow for the possibility of providing "forward" forecasts and should avoid the idea of a merely "retrospective" fit of the data available. Models for forecasting eruptions should cover a twofold scope: fit the eruption data and incorporate a robust forecast procedure. While the first requirement is mandatory, the latter one is not commonly used in statistical modeling of volcanic eruptions. By carrying out and testing a forecast procedure on data available at the present, one could make enhancement in the forecast matter and reveal the model limitations.

Different methods have been presented in the past years aiming at the identification of possible recurrence or correlation in the volcanic time and/or volume data for long-term eruption forecast. Klein (1982), Mulargia *et al* (1985) and Bebbington and Lai (1996a and 1996b) studied the time series of volcanic events looking at the mean rate of occurrence. Sandri *et al.* (2005) applied a generalized form of time predictable model to Mount Etna eruptions using regression analysis. Marzocchi and Zaccarelli (2006) found different behavior for volcanoes with "open" conduit regime compared to those with "closed" conduit regime. Open conduit volcanoes (Mt Etna, Kilauea volcano there tested) seem to follow a so-called *Time Predictable Model*. While closed conduit volcanoes seem to follow a homogeneous Poisson process. De La Cruz-Reyna (1991) proposed a load-and-discharge model for eruptions in which the time predictable model could be seen as a particular case. Bebbington (2008) presented a stochastic version of the general load-and-discharge model also including a way to take into account the history of the volcano discharging behavior. In this paper the author studied the time predictability as a particular case of his model with application to Mount Etna, Mauna Loa and Kilauea data series. A different hierarchical approach has been presented by Bebbington (2007) using Hidden Markov Model to study eruption occurrences with application to Mount Etna flank eruptions. This model is able to find any possible underlying volcano activity resulting in changes of the volcanic regime. Salvi *et al* (2006) carried out analysis for Mt Etna flank eruption using an Non Homogeneous Poisson process with a power law intensity, while Smethurst *et al* (2009) applied a Non Homogeneous Poisson process with a piecewise linear intensity to Mt Etna flank eruptions.

In a recent paper Passarelli *et al* (2010) proposed a Bayesian Hierarchical

Model for interevent time-volumes distribution using the time predictable process with application to Kilauea volcano. The model presents a new Bayesian methodology for an open conduit volcano that accounts for uncertainties in observed data. Besides, the authors present and test the forecast ability of the model retrospectively on the data through a forward sequential procedure. While the model seems to produce better forecasts that other models in the literature, it produces fits to eruption volumes and interevent times that are too large, reducing the forecast performances. This is due to the use of normal distributions for the log-transformed data. This is a restrictive distributional assumption that creates very long tails. Here we propose a more general modeling strategy that allows for more flexible distributions for the interevent times and volumes data.

Using the same framework of Passarelli et al (2010), we will model the interevent times and volumes data through distributions with exponential decay (Klein, 1982, Mulargia, 1985, Marzocchi, 1996, Bebbington, 1996a , 1996b and 2007, Salvi *et al*, 2006, Smethurst *et al*, 2009). This provide a general treatment of the volume and interevent time series, hopefully improving the forecast capability of the model. As eruptive behavior we use the Generalized Time Predictable Model (Sandri *et al*, 2005 and Marzocchi and Zaccarelli, 2006). This model assumes: 1) eruptions occur when the volume of magma in the storage system reaches a threshold value, 2) magma recharging rate of the shallow magma reservoir could be variable and 3) the size of eruptions is a random variable, following some kind of statistical distribution. Under these assumptions, the time to the next eruption is determined by the time required for the magma entering the storage system to reach the eruptive threshold. The more general form for a time-predictable model is a power law between the erupted volume and the interevent time:

$$r_i = cv_i^b \tag{1}$$

where, if the parameter b is equal to unity we are in a classical time predictable system (see De La Cruz Reyna 1991, Burt *et al.* 1994). If b is equal to 0 the system is not time predictable. If $b > 1$ or $0 < b < 1$ we have a non-linear relationship implying a longer or shorter interevent time after a large volume eruption compared to a classical time predictable system. The goal of the present work is to infer the parameters of Equation (1).

In the remainder of this paper, we focus our attention on some specific issues: 1) to discuss the physical meaning and implications of parameters inferred; 2) to verify if the model describes the data satisfactorily; 3) to compare the forecasting capability of the present model with other models previously published in literature using the sequential forward procedure discussed in Passarelli *et al* (2010). In the first part of this paper, we will introduce the generality of the model by considering three stages: 1) a model for the observed data; 2) a model for the process and 3) a model for the parameters (Wikle, 2002). Then we will discuss how: 1) to simulate the variables and parameters of the model; 2) to check the model fit; 3) to use the model to assess probabilistic forecast in comparison with other statistical published models. The last part of the paper contains the application of the model to Kilauea volcano and Mount Etna eruptive data.

2. A Bayesian Hierarchical Model for Time-Predictability

In the following sections we present a detailed description of our proposed model. We denote it as Bayesian Hierarchical Time Predictable Model II (BH_TPM II), while the model proposed in Passarelli et al. (2010) is denoted as BH_TPM. In Section 2.1 we discuss the measurement error model. In Section 2.2 we consider a model for the underlying process, which is based on the exponential distribution. In Section 2.3 we discuss the distributions that are placed on the parameters that control the previous two stages of the model. In Section 2.4 we introduce the simulation procedure and in Section 2.5 we consider model assessment and forecasting of volcanic eruptions.

2.1. Data model

The dataset for this model has n pairs of observations: volumes and interevent times denoted as d_{v_i} and d_{r_i} respectively. We assume independence between the measurement errors of interevent times and volumes. This is justified by the fact that these two quantities are measured using separate procedures. Dependence between times and volumes will be handled at the process stage, following the power law in Equation (1). In addition, we assume that, conditional on the process parameters, the interevent times or volumes are independent within their group. This is a natural assumption within a hierarchical model framework. It is equivalent to assuming that the volumes (times) are exchangeable between them. Exchangeability implies that all permutations of the array of volumes (times) will have the same joint distribution. Exchangeability is weaker than independence, and it is implied by it.

Our measurement error model assumes a multiplicative error for the observations. This follows from BH_TPM where is was assumed that

$$\log(d_{r_i}) = \log r_i + \log \epsilon_{r_i} \tag{2}$$

with $\log \epsilon_{r_i} \sim N(0, \sigma^2_{D_{r_i}})$ where $\sigma^2_{D_{r_i}} = (\frac{\Delta d_{r_i}}{d_{r_i}})^2$ (for more details see Passarelli *et al*, 2010). The analogous assumption $\log(d_{v_i}) = \log v_i + \log \epsilon_{v_i}$ and $\log \epsilon_{v_i} \sim N(0, \sigma^2_{D_{v_i}})$ where $\sigma^2_{D_{v_i}} = (\frac{\Delta d_{v_i}}{d_{v_i}})^2$, was considered for the volumes. Exponentiating on both sides of Equation (2) we have

$$d_{r_i} = \epsilon_{r_i} r_i \tag{3}$$

which is the data stage model we propose in BH_TPM II.

The error in Equation (3) follows a probability distribution with positive support. We choose an inverse gamma distribution. This is a flexible distribution defined by two parameters which will provide computational advantages. We fix the two defining parameters by assuming $\mathrm{E}(\epsilon_{r_i}) = 1$ and calculating $\mathrm{var}(\epsilon_{r_i})$ using a delta method approximation. Specifically, from the assumption that $\log \epsilon_{r_i} \sim N(0, \sigma^2_{D_{r_i}}))$, we have that $\mathrm{E}(\log \epsilon_{r_i}) = 0$ and $\mathrm{var}\left(\log \epsilon_{r_i}\right) = \sigma^2_{D_{r_i}} = (\frac{\Delta d_{r_i}}{d_{r_i}})^2$. Thus

$$\mathrm{var}\left(\epsilon_{r_i}\right) = \sigma^2_{D_{r_i}} \left[g'\left(\mathrm{E}\left(\frac{\Delta d_{r_i}}{d_{r_i}} \right)\right)\right]^2 = \left(\frac{\Delta d_{r_i}}{d_{r_i}} \right)^2$$

where $g(x) = \exp(x)$ and g' is the first derivative.

Recall that a random variable X that follows an inverse gamma distribution with parameters α_{r_i} and β_{r_i} has expected value $E(X) = \frac{\beta_{r_i}}{\alpha_{r_i}-1}$ and variance $\mathrm{var}(X) = \frac{\beta_{r_i}^2}{(\alpha_{r_i}-1)^2(\alpha_{r_i}-2)}$. We then have that

$$\frac{\beta_{r_i}}{\alpha_{r_i}-1} = 1 \quad \text{and} \quad \frac{\beta_{r_i}^2}{(\alpha_{r_i}-1)^2(\alpha_{r_i}-2)} = \left(\frac{\Delta d_{r_i}}{d_{r_i}}\right)^2 .$$

Solving for α_{r_i} and β_{r_i} gives $\alpha_{r_i} = (\frac{d_{r_i}}{\Delta d_{r_i}})^2 + 2$ and $\beta_{r_i} = (\frac{d_{r_i}}{\Delta d_{r_i}})^2 + 1$ where $\frac{\Delta d_{r_i}}{d_{r_i}}$ is the relative error. Analogous calculations can be done for the volumes. The joint distributions for the measurement errors $\epsilon_r = (\epsilon_{r_1}, \ldots, \epsilon_{r_n})$ and $\epsilon_v = (\epsilon_{v_1}, \ldots, \epsilon_{v_n})$ result in

$$[\epsilon_r|\alpha_{r_i}, \beta_{r_i}] = \prod_{i=1}^{n} \Gamma^{-1}(\alpha_{r_i}, \beta_{r_i}) \quad \text{and} \quad [\epsilon_v|\alpha_{v_i}\beta_{v_i}] = \prod_{i=1}^{n} \Gamma^{-1}(\alpha_{v_i}, \beta_{v_i}) \tag{4}$$

where $\alpha_{v_i} = \left(\frac{d_{v_i}}{\Delta d_{v_i}}\right)^2 + 2$ and $\beta_{v_i} = \left(\frac{d_{v_i}}{\Delta d_{v_i}}\right)^2 + 1$. Here we use $[X]$ to denote the distribution of a random variable X and Γ^{-1} to denote an inverse gamma.

The distribution of the observed variables d_{r_i} and d_{v_i} can be obtained from the error distributions specified by the expression in (4). Noting that $\left|\frac{d\epsilon_{r_i}}{d(d_{r_i})}\right| = \frac{1}{r_i}$ and using the change of variables formula for probability density functions, we have that

$$[d_r|\alpha_{r_i}, \beta_{r_i}, r_i] = \prod_{i=1}^{n} \Gamma^{-1}(\alpha_{r_i}, \beta_{r_i} r_i) \quad \text{and} \quad [d_v|\alpha_{v_i}, \beta_{v_i}, v_i] = \prod_{i=1}^{n} \Gamma^{-1}(\alpha_{v_i}, \beta_{v_i} v_i). \tag{5}$$

The expression in (5) will be used to obtain the likelihood function for our data.

2.2. Process model

The starting point for the model pertaining the unobserved quantities r_i is the assumption that volcanic eruptions correspond to a homogeneous Poisson process. A Poisson process in times has the property that the number of events that occur during a given time interval follow a Poisson distribution with mean proportional to the length of the interval. Additionally the time between consecutive events is distributed as an exponential random variable (Klein, 1982, Mulargia, 1985, Marzocchi, 1996, Bebbington and Lai, 1996a, 1996b). Thus we assume that $r_i \sim \text{Exp}(\lambda)$ implying that the joint distribution of $r = (r_1, \ldots, r_n)$ is given by $[r|\lambda] = \prod_{i=1}^{n} \text{Exp}(\lambda)$. Given the distributional assumption for the interevent times we can obtain the distribution of the volumes v_i using Equation (1). Recalling that $r_i = cv_i^b$ and $\left|\frac{dr_i}{dv_i}\right| = cbv_i^{b-1}$, the change of variable formula for probability density functions yields $[v_i] = cb\lambda v_i^{b-1}\mathrm{e}^{-\lambda c v_i^b}$ Written in distributional form we have: $v_i \sim \text{Wb}\left(b, \left(\frac{1}{\lambda c}\right)^{\frac{1}{b}}\right)$ where $Wb(\cdot,\cdot)$ denotes a Weibull distribution. The joint distribution for the volumes $v = (v_1, \ldots, v_n)$ is given as

$$[v|\lambda, b, c] = \prod_{i=1}^{n} \text{Wb}\left(b, \left(\frac{1}{\lambda c}\right)^{\frac{1}{b}}\right). \tag{6}$$

This completes the specification of the second stage of our model.

2.3. Parameters model

To complete our model we need to specify distributions for the parameters b, c and λ. Our choices are based on prior information obtained from previous modeling efforts. In a Bayesian setting, like the one proposed in this work, we have the ability to include structural information, like the one used to build the second stage model, as well as prior information. The final product consists of the posterior distribution of all model parameters. This contains a blend of the information provided by all the stages of the model: data, process and prior knowledge.

We choose for λ a gamma distribution with known parameters, from now on hyperparameters. This is denoted as have: $\lambda \sim \text{Ga}(\alpha_\lambda, \beta_\lambda)$ where

α_λ and β_λ are calculated by fitting the interevent times data with a gamma distribution, via maximum likelihood estimation. For the time predictable equation parameters, i.e. b and c, we use normal distributions with moments calculated using the posterior distributions taken from BH_TPM (Passarelli et al., 2010). Thus $[b] = \mathrm{N}(\mu_b, \sigma_b^2)$ and $[c] = \mathrm{N}(\mu_c, \sigma_c^2)$.

By choosing the values of the hyperparameters we are introducing a certain degree of subjectivity in our modeling. We believe that this is a desirable feature of the Bayesian approach, as it allows to incorporate knowledge from similarly behaved open conduit volcanoes. We remark the subjective approach allowed in Bayesian Statistics could be a suitable tool in modeling geophysical phenomena where available data are scarce. This provides the possibility of incorporating knowledge obtained from other sources in a probabilistic way, through the prior distributions. This allows for the introduction of physical and/or statistical constraints, when available, on the parameters governing the examined phenomenon. In principle this methodology could be helpful to improve the understanding of a particular system. We want to point out, though, that subjective statistical modeling choices need careful justification, possibly relying on physical or phenomenological constraints.

2.4. Posterior and full conditional distributions

The three stage model specification developed in the previous sections produces a posterior distribution for the model parameters r, v, b, c and λ that, using Bayes theorem, can be written as

$$[r, v, b, c, \lambda | d_r, d_v \Delta d_r \Delta d_v] \propto \tag{7}$$

$$[d_r|\alpha_{d_r}, \beta_{d_r}, r][d_v|\alpha_{d_v}, \beta_{d_v}, v][v|c, \lambda, b][r|\lambda][\lambda][b][c] \; .$$

To make inference about the posterior distribution specified by Equation (7) we draw samples from it using Markov chain Monte Carlo (MCMC) methods (Gelman et al. 2000, Gilks et al, 1996). This requires the full conditional distributions for each parameter in the model. In the equations below we specify each of them using the notation $[X|\ldots]$ to indicate the distribution of variable X conditional on all other variables.

$$[r_i|\ldots] \propto r_i^{\alpha_{r_i}} \exp\left\{-r_i\left(\lambda + \frac{\beta_{r_i}}{d_{r_i}}\right)\right\} = \mathrm{Ga}\left(\alpha_r + 1\,,\, \lambda + \frac{\beta_{r_i}}{d_{r_i}}\right)$$

$$[v_i|\ldots] \propto v_i^{\alpha_{v_i}+b-1} \exp\left\{\left(\lambda c v_i^b + \frac{\beta_{v_i} v_i}{d_{v_i}}\right)\right\}$$

$$[\lambda|\ldots] \propto \lambda^{2n+\alpha_\lambda-1} \exp\left\{-\lambda\left(\beta_\lambda + c\sum_{i=1}^{n} v_i^b + \sum_{i=1}^{n} r_i\right)\right\} = \mathrm{Ga}\left(\alpha_\lambda + 2n\,,\, \beta_\lambda + c\sum_{i=1}^{n} v_i^b + \sum_{i=1}^{n} r_i\right)$$

$$[c|\ldots] \propto c^n \exp\left\{-c\lambda\sum_{i=1}^{n} v_i^b + \frac{\mu_c c}{2\sigma_c^2} - \frac{c^2}{2\sigma_c^2}\right\}$$

$$[b|\ldots] \propto \prod_{i=1}^{n}\left(bv^{b-1}\right) \exp\left\{-\lambda c\sum_{i=1}^{n} v_i^b + \frac{\mu_b b}{2\sigma_b^2} - \frac{b^2}{\sigma_b^2}\right\}$$

The full conditional distributions of $r_i, i = 1, \ldots, n$ and λ can be sampled directly in Gibbs steps, as they correspond to gamma distributions. The full conditionals of the other parameter do not have standard forms. So we use Metropolis steps to obtain samples from them. Once samples from the MCMC are obtained we discard the first part of the chain as a burn-in phase (see for example Gilks et al., 1996); then we do a “thinning” of the chain by subsampling the simulated values at a fixed lag k. This strategy ensures that, setting k to some high enough value, successive draws of the parameters are approximately independent (Gelman, 1996). To define the lag we use the auto-correlation function as shown later in the text.

2.5. Model Checking and Forecasting procedure

We have presented, so far, the hierarchical structure of the model and the fitting procedure for the model parameters, based on MCMC sampling. We now address the issues of (1) testing the goodness of the proposed model and (2) forecasting future interevent times.

Bayesian model checking is based on the idea that predictions obtained from the model should be compatible with actual data. So our strategy consists of simulating data from the predictive posterior distribution and comparing them to actual observations. The predictive posterior distribution quantifies the uncertainty in future observations given the observed data. By denoting $\tilde{r}$ future values of interevent times we have that the posterior predictive is

$$[\tilde{r} \mid Data] = \int_{\mathbb{R}^+} [\tilde{r} \mid \lambda][\lambda \mid Data] d\lambda \tag{8}$$

where $\mathbb{R}^+$ denotes the parameter space. To obtain samples from the distribution in Equation (8) we start from the MCMC samples of λ. Suppose we have N of them and denote them as λ^j. Conditional on λ^j, for $j = 1, \ldots, N$ we simulate $\tilde{r}^j$ from $[\tilde{r} \mid \lambda^j]$, which are products of exponentials. We obtain N synthetic catalogs with n pairs of interevent time and volume data.

These are compared to the observed data using descriptive statistics. As descriptive statistics we choose the mean number of events or rate of occurrence, maximum, minimum, median and standard deviation for both real and synthetic data.

To test the ability of the model to forecast future volumes and interevent times we use a sequential approach. We proceed by fitting the model to the first data pair, then we add the data of the second event to the model fitting. We continue adding data sequentially until the last event. This provides an assessment of the number of data needed for the model to effectively "learn" the model parameters. Therefore, we are able to decide the minimum amount of data needed to define the learning phase for the model. For the remaining part of data (i.e. voting phase), we use the sequentially sampled parameters to generate the distribution for the next event (interevent time). We can thus compare the forecasted interevent times with the observed data and with forecasts from other published methods (see *forward* procedure discussed in Passarelli *et al*, 2010).

A close look at Equation (8) reveals a practical forecasting problem. We observe that the posterior predictive distribution of the interevent times depends on the distribution of the interevent times given the parameter λ. While this is statistically correct, it is not a realistic forecasting procedure. In fact, in a generalized time predictable system the time to the next eruption is strongly dependent on the volume of the previous eruption. More explicitly, in our current framework, after the end of the n-th eruption we have samples of λ that are simulated using only the information up to $(d_{r_{(n-1)}}, d_{v_{(n-1)}})$. We would like to incorporate the information on d_{v_n}. We do this by resampling the posterior realizations of λ using the Sampling

Importance Resampling algorithm (hereafter SIR), (Rubin,1988, Smith and Gelfand, 1992) together with Bayes theorem.

Let $\theta_{n-1} = b, c, \lambda$ be the samples obtained from our model using the first $n-1$ data. For the n-the interevent time we have

$$[\tilde{r_n} \mid d_{v_n}] = \int_{R^+} [\tilde{r_n} \mid d_{v_n}, v_{n-1}, \theta_{n-1}][\theta_{n-1} \mid d_{v_n}, v_{n-1}] d\theta_{n-1} \tag{9}$$

Obtaining samples from the predictive distribution in Equation (9) requires samples of $[\theta_{n-1} \mid d_{v_n}, v_{n-1}]$, which are not available. Our MCMC algorithm produces samples of $[\theta_{n-1} \mid d_{v_{n-1}}, v_{n-1}]$ instead. Using Bayes theorem we have that

$$[\theta_{n-1} \mid d_{v_n}, v_{n-1}] \propto [d_{v_n} \mid v_{n-1}, \theta_{n-1}][\theta_{n-1} \mid v_{n-1}] \ . \tag{10}$$

In Equation (10) we recognize $[d_{v_n} \mid v_{n-1}, \theta_{n-1}]$ as the inverse gamma distribution used for volume data in Equation (5). $[\theta_{n-1} \mid v_{n-1}]$ is the posterior distribution for parameters λ, b and c up to the first $n-1$ events. The SIR algorithm consists of resampling the output from the MCMC, say θ^j_{n-1}, with replacement, using the normalized weights defined as

$$w^*(\theta^i_{n-1}) = \frac{w(\theta^i_{n-1})}{\sum_{j=1}^m w(\theta^j_{n-1})}$$

where $w(\theta^i_{n-1}) = [d_{v_n} \mid v^i_{n-1}, \theta^i_{n-1}]$. The weights w correspond to the inverse gamma distribution in Equation (5) for the observed volume of the n-th event conditional on the sampled volumes of the previous event and the remaining parameter, as simulated by the MCMC. The output from the SIR algorithm can be used within Equation (9) to obtain the desired samples of the n interevent time. An brief description of the SIR algorithm is given in Appendix A.

Finally we use the notion of probability gain or information content, as proposed by Kagan and Knopoff, 1987, to make explicit comparisons of different forecasting methods. We calculate the information gain for the present model with respect to other statistical models in the literature. Let A and B be two statistical models, the probability gain is defined as the difference between their log-likelihoods, i.e.:

$$PG = \sum_{i=m}^{n} (l_A(\delta d_{r_i}) - l_B(\delta d_{r_i})). \tag{11}$$

Here l_A and l_B are the natural logarithm of the likelihoods for Model A and B respectively and $m, \ldots, n$ denote the voting phase. These are calculated in a temporal window δd_{r_i} of one month around the observed interevent time in the voting phase. If PG is greater than zero, Model A has better forecasting performance than Model B, if PG is zero the two models are equivalent. Together with the total probability gain given by equation (11), we can calculate the "punctual" probability gain, i.e. the probability for each event $l_A(\delta d_{r_i}) - l_B(\delta d_{r_i})$ with $i = m, \ldots, n$ (Passarelli et al, 2010).

3. Application to Kilauea volcano and Mount Etna

We apply the BH_TPM II to Kilauea volcano and Mt Etna eruption data. Marzocchi and Zaccarelli, 2006 have found that Kilauea volcano and Mt. Etna follow a time predictable eruptive behavior. They also stated that these volcanoes are in open conduit regime because of their high eruptive frequency and, consequently, short duration of interevent times. Bebbington, 2007 have showed evidences of the time-predictable character of Mt.

Etna flank eruptions using a catalog starting in 1610 AD. The same results on time-predictability are attained by Sandri et al., 2005 only focusing on the Mt Etna flank eruptions in the period 1971-2002. Passarelli *et al*, 2010 have found time-predictability of Kilauea volcano for eruptive catalog starting in 1923 AD.

These findings led us to use Kilauea and Mt Etna as test cases for our proposed model. Our goals in this paper is to test: 1) whether or not they follow a time predictable behavior; 2) the reliability of the assumptions used in the model; 3) improvements in using the information given by the volume measurement errors; 4) the ability to fit the observed data, and 5) the forecast capability of the model compared with models previously published in literature for Kilauea and Mt Etna.

3.1. Kilauea volcano

Kilauea volcano is the youngest volcano on the Big Island of Hawaii. The subaerial part of Kilauea is a dome-like ridge rising to a summit elevation of about 1200 m, is about 80 km long, 20 km wide and covers an area of about 1500 km^2. Kilauea had a nearly continuous summit eruptive activity during the 19th century and the early part of the 20th century. During the following years, Kilauea's eruptive activity had shown little change. After 1924, summit activity had become episodic and after a major quiescence period during 1934-1952, the rift activity raised increasing the volcanic hazard (Holcomb, 1987). It is widely accepted that Kilauea has its own magma plumbing system extending from the surface to about 60 km deep in the Earth, with a summit shallow magma reservoir at about 3 km depth. The shallow magma reservoir is an aseismic zone beneath the South zone of the

Kilauea caldera and it is surrounded on two sides by active rift conduits (Klein et al, 1987).

The eruption history of Kilauea volcano directly documented dates back to 18*th* century, however before the 1923 the eruption record is spotty and in most of the events the erupted volume is unknown. Therefore, we limit our analysis to the 42 events after 1923 AD (please refer to Passarelli et al., 2010 for more details on the Kilauea catalog completeness). The data are listed in Table (1) where we report the onset date of each eruption together with the volume erupted (lava + tephra) and the relative interevent time. The volume of the 1924/05/10 event is taken from http://www.volcano.si.edu/ and is only the tephra volume. Since the eruption that began in 1983 is still ongoing with a volume erupted greater than 3 km^3, we have 41 pairs of data of interevent time (i.e. the time between the onset of i-*th* and the onset of (i+1)-*th* eruptions) and volume erupted (in the i-*th* eruption).

In the next two subsections we will present the results of the model for the Kilauea dataset.

3.1.1. Results for variables and parameters

We begin with a discussion of the choice of hyperparameter values. For interevent times we choose an error (Δd_{r_i}) of 1 day for all data in the catalog. For the volumes we assume relative errors ($\Delta d_{v_i}/d_{v_i}$) of 0.25 for data before the 1960 AD (i.e. $i = 1, \ldots, 13$) and of 0.15 for data after the 1960 AD (i.e. $i = 14, \ldots, 41$) (see discussion in Passarelli et al., 2010). Other hyperparameters for the distributions of b and c, are chosen by matching the first two moments of the output of the BH_TPM, i.e. $\mu_b = 0.2$, $\sigma_b = 0.1$, $\mu_c = 200$ days/10^6m^3 and $\sigma_c = 50\,\text{days}/10^6\text{m}^3$ (see Passarelli *et al*, 2010

Figure 4).

We run an MCMC simulation for 201,000 iterations with a burn-in of 1,000 iterations and a thinning of one every 20 iterations. We checked the output for convergence and approximate independence of the final sample. In Figure 1 we show the MCMC realizations of r_i and v_i (blue stars), obtained using the whole catalog, and compare with the observed data (red pluses). The plots indicate that the model is able to accurately reproduce the data and that measurement errors have a realistic impact in the estimation uncertainty of the true interevent times and volumes.

Figure 2 shows the posterior distributions of b, c and λ using all data. As the distribution of b (top left panel) is concentrated within the [0,1] interval, with mean 0.45 and standard deviation 0.05, we infer that the Kilauea volcano has a time predictable behavior. This is compatible with the findings in Passarelli et al. (2010). For the distribution of c (top right panel), which is function of the average magma recharge process, we find that the distribution is mostly contained within the interval [100,240] days/10^6 m^3, with mean 164 days/$10^6\mathrm{m}^3$ and a standard deviation 24days/$10^6\mathrm{m}^3$. In the bottom left panel we have the posterior distribution for λ, the time of occurrence of the number of events over the length of the catalog. Most of this distribution is contained in the interval $[1.5, 3]$ $\times 10^{-3}\,\mathrm{days}^{-1}$ and has mean is 2.0×10^{-3} $\times 10^{-3}\,\mathrm{days}^{-1}$ and standard deviation 0.3×10^{-3} $\times 10^{-3}\,\mathrm{days}^{-1}$. This results are compatible with the time of occurrence calculated directly from the data with Maximum Likelihood Estimation (MLE) techniques, which yields $\lambda_{MLE} = 1.9 \times 10^{-3}$ $\times 10^{-3}\,\mathrm{days}^{-1}$ with 95 % confidence interval $[1.4, 2.5] \times 10^{-3}$ $\times 10^{-3}\,\mathrm{days}^{-1}$. Figure 3 corresponds to the sequential version of Figure 2. The plots are obtained using the approach discussed in

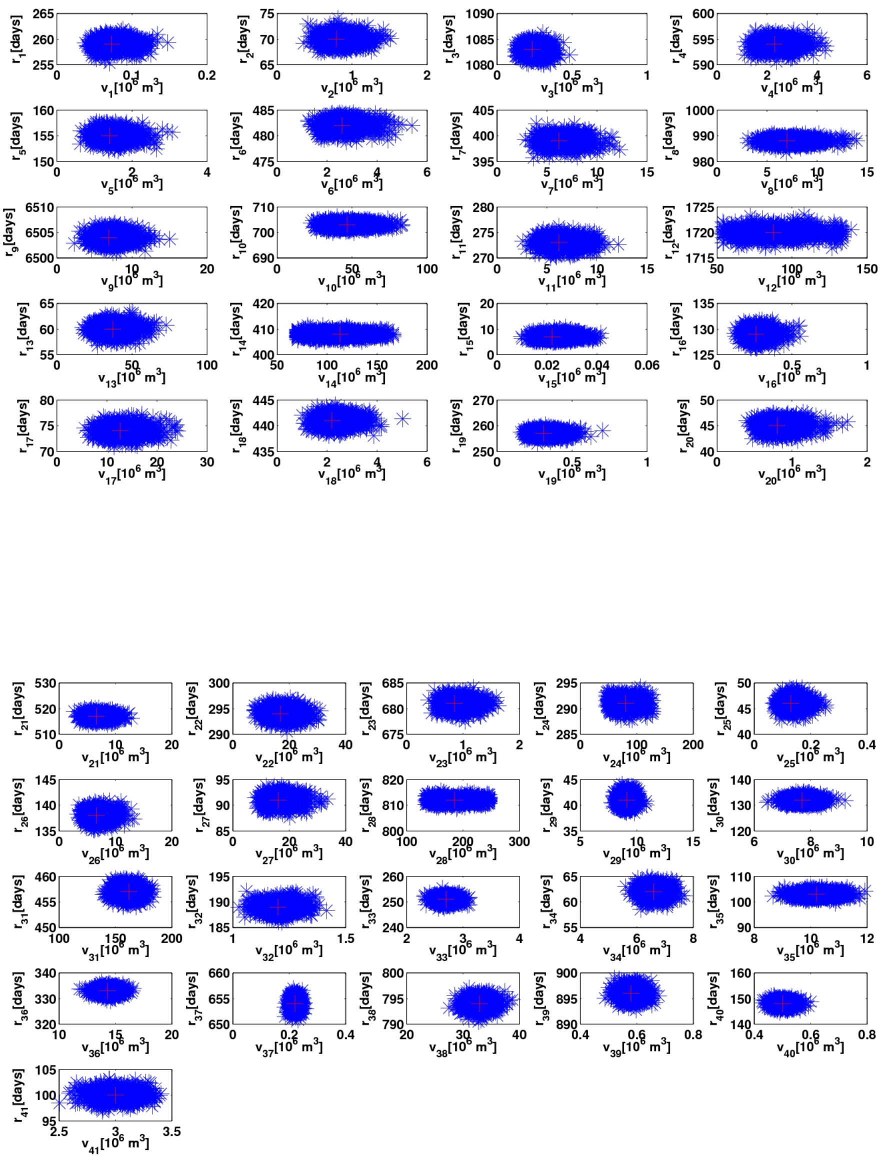

Figure 1: *Blue stars show the posterior distributions of pairs of simulated variables (interevent times r_i and volumes v_i). The top panel corresponds to $i = 1, \ldots, 20$ and the the bottom panel to $i = 21, \ldots, 41$.*

section 2.5.

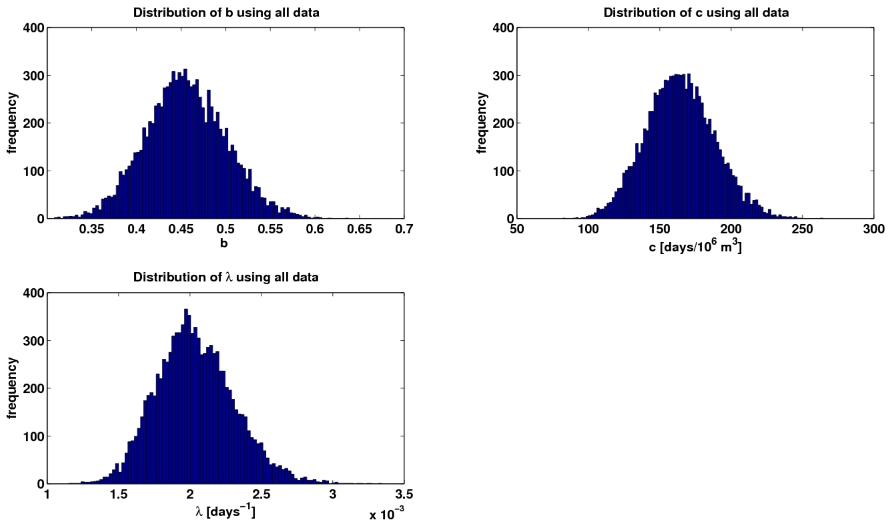

Figure 2: *Posterior distributions for BH_TPMII parameters obtained using all data in the catalog: top left panel refers to b, top right to c and bottom left to λ.*

The results obtained imply a power law relationship between interevent times and volumes. As discussed in Passarelli *et al*, 2010, this non linear association underlines the role played by the magma discharging process in the eruption frequency. Such relationship implies the possibility of having a non constant input rate in the magma storage system. Therefore, a large erupted volume may trigger the increasing of the magma upwelling process inside a shallow reservoir. We expect a shorter quiescence period after an eruption characterized by a large volume compared with a process where the magma recharging rate is constant (i.e. classical time predictable model). A simple explanation is the existence of an additional gradient

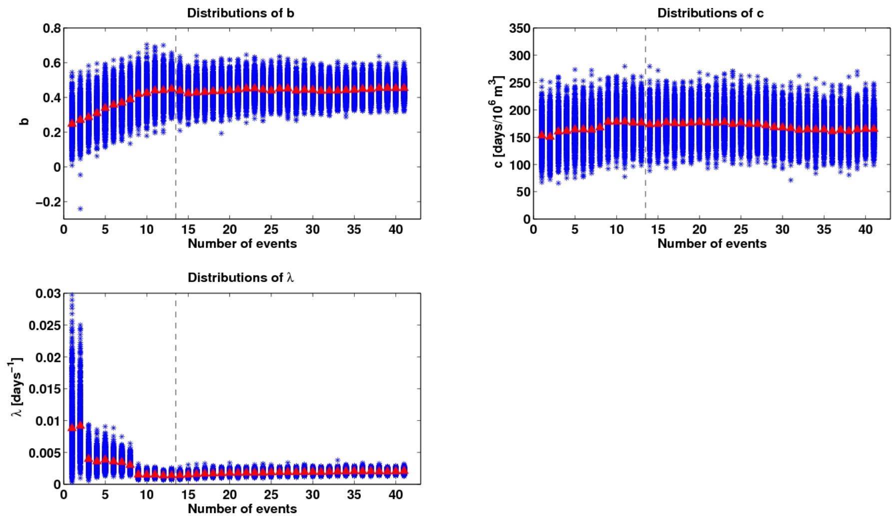

Figure 3: *Posterior distributions of: b parameter in top left panel, c parameter in top right panel and λ in the bottom left panel, all calculated using the sequential procedure discussed in the text. Black dashed line represents the learning phase. Red triangles are the mean of each distribution.*

of pressure due to the drainage process of the shallow magma system by a large erupted volume. This pressure gradient may increase the magma upwelling process from the deep crust into the shallow storage system. Non constant magma input rate for the shallow magma reservoir for Kilauea volcano has been found by Aki and Ferrazzini (2001) and Takada, (1999). This non-stationarity should be take into account in modeling the magma chamber dynamics at Kilauea volcano.

3.1.2. Model checking and Forecasts

We use the ability of our approach to quantify uncertainties in future predictions given the observed data to check the validity of our model. We simulate 10,000 synthetic catalogs using the procedure described in Section 2.5. We then calculate for both, synthetic catalogs and observed data, the rate of occurrence, the maximum, the minimum, the median and the standard deviation. Figure 4 shows the comparisons between the histograms of the synthetic data and the corresponding observed values. Predictions are in good agreement with observed values for the rate of occurrence, the minimum and the median. The are some discrepancies for the maximum and, consequently, for the standard deviation. In these cases the observed value falls in the tails of the predictive distributions. This is due to the fact that the maximum corresponds to the 18 years of quiescence of the Kilauea volcano (i.e. 1934-1952 AD). This is a extraordinary long period of rest for the Kilauea and it could be considered as an extreme value. The second longest interevent time is about 5 years of quiescence (i.e. 1955-1959 AD). Such value falls right at the center of the distribution with p-value=0.7. In summary, the model is capable of reproducing the data, with the exception of future extreme events that correspond to the tails of the predictive distribution.

We use the sequential approach of Section 2.5 to evaluate the model's forecast performance and compare it with published results for the Kilauea volcano's interevent times. Here we compare our results with those from the homogeneous Poisson process (Klein et al., 1982), the Log-Normal model (Bebbington and Lai, 1996b), the Generalized Time Predictable Model (GTPM) (Sandri *et al.*, 2005) and the BH_TPM (Passarelli et al., 2010).

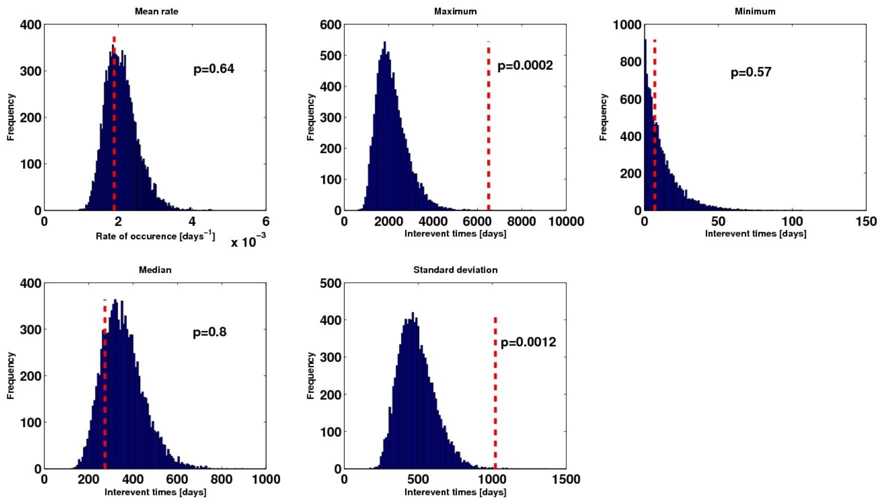

Figure 4: *Histograms of samples from the posterior predictive distributions of several summaries of the interevent times for the Kilauea (Red bars). Red dashed lines denote the corresponding observed values. p-values correspond to the proportion of samples above the observed values.*

The homogeneous Poisson implies a totally random and memoryless eruptive behavior. In the Log-Normal model interevent times are described using a log-normal distribution. The mode of a log-normal distribution could reveal a certain degree of ciclicity in the eruptive behavior for Kilauea volcano. The GTPM consists of a linear regression among pairs of interevent times and volumes. The BH_TPM is a hierarchical model where the interevent times and volumes are described via log-normal distributions and uses the logarithm of the generalized time predictable model equation as eruptive behavior.

To gauge the role of the information provided by the volumes in the sequential estimation of the interevent times we compare the MCMC samples of λ with those obtained after the SIR procedure. The results are shown in Figure 5. From the figure it is clear that the information provided by the volumes shrinks and shifts the distribution of λ . We use the resampled λ values to calculate the probability gains with respect to the other four models considered. The results are plotted in Figure 6 where we show the "punctual" probability gain and we report the total probability gain as calculated using equation (11). As indicated by positive total probability gains in all cases, our model shows an improvement in forecasting capability when compared to any of the other four models. The largest gain is observed for the Poisson model (panel a) where the model provides better forecasts for 20 out of 27 eruptions. The largest global gain is obtained testing against the GTPM (panel d). This latter results is likely due to the inclusion of information on measurement error. The smallest overall gain is achieved with respect to BH_TPM (panel b). This is not surprising as BH_TPM is the closets model to BH_TPMII among the ones considered.

Overall we observe that BH_TPMII has better forecasting performance than any of the four competing models in more than 50% of the events. Thus BH_TPMII seems to be more reliable for probabilistic hazard assessments that the other models considered.

Finally we investigate possible linear associations between the pointwise probability gains and the interevent times or volumes in each of the four considered cases. We only find a significant correlation (p-value≤ 0.01) for the case of the homogeneous Poisson process. In this case there is a clear inverse relationship. This implies that the longer the interevent time the worse our forecast is. This is justified by the fact that for long quiescence periods

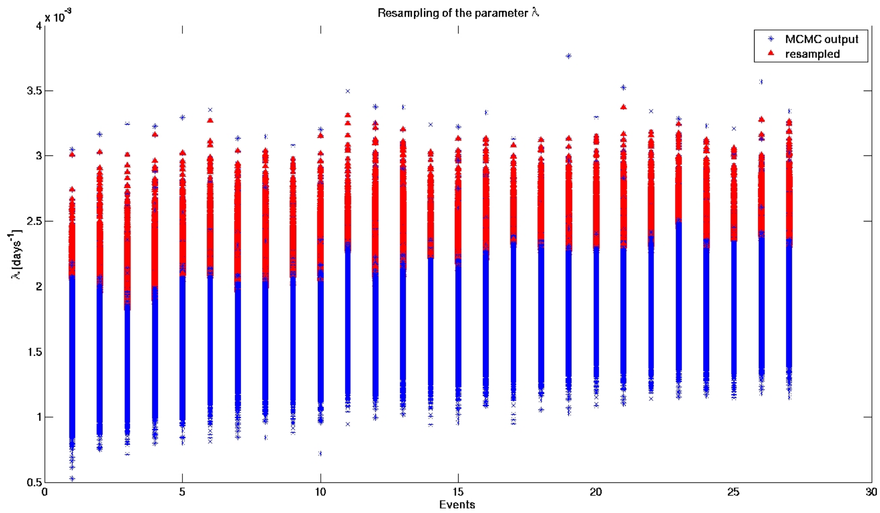

Figure 5: *Sequentially updated posterior samples of λ's in the voting phase (events from 14 to 41). Blue stars corresponds to MCMC output. Red triangles correspond to resampling after observing the corresponding volumes.*

the Kilauea volcano could become memoryless with transition from open to closed conduit regime (see Marzocchi and Zaccarelli, 2006). In addition, considering the events as a point in time (see Bebbington, 2008) together with the fact that we do not consider intrusions not followed by eruptions (Takada, 1999, Dvorak and Dzurisin, 1993) could be distorting. Finally another possible explanation could be related to possible modification of the shallow magma reservoir geometry after an eruption (Gudmundsson, 1986).

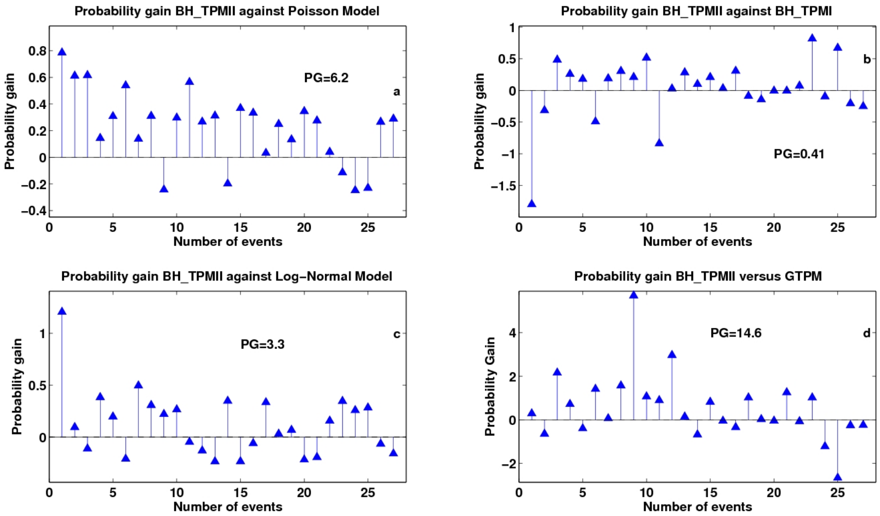

Figure 6: *"Punctual probability gain" of the BH_TPMII for each event after the learning phase against: in panel **a** Poisson Model (Klein, 1982), in panel **b** BH_TPM (Passarelli et al, 2010), in panel **c** Log-Normal Model (Bebbington* and *Lai, 1996b) and in panel **c** Generalized Time Predictable Model (Sandri et al., 2005). Values greater than zero indicate when BH_TPM model performs better forecast than the reference models. Positive values indicate that BH_TPMII has better forecasting ability than the alternative model. Global probability gains are reported as "PG" in each of the four cases.*

3.2. Mount Etna volcano

Mount Etna volcano is a basaltic stratovolcano located in the North-Eastern part of the Sicily Island. It is one of the best known and monitored volcano in the world and records of its activity date back to several centuries B.C. The sub-aerial part of Mount Etna is 3,300 m high covering an area of approximately 1,200 km^2. Two styles of activity occur at Mt Etna: a quasi-continuous paroxysmal summit activity, often accompanied with explosions, lava fountains and minor lava emission; a less frequent flank eruptive activity, typically with higher effusion rate originate from fissures that open downward from the summit craters. The flank activity is sometimes accompanied by explosions and lava spattering; recently, two flank eruptions have been highly explosive and destructive, the 2001 and 2002-2003 events (Behncke and Neri, 2003, Andronico *et al*, 2005, Allard *et al*, 2006).

At present there are petrological, geochemical and geophysical evidences for a 20-30 km deep reservoir controlling the volcanic activity (Tanguy *et al*, 1997), but it is still debated whether or not Mt Etna has one o more shallower plumbing systems. Results from seismic tomography do not reveal any low velocity zone in the uppermost part of the volcanic edifice, while a high-velocity body at depth of < 10 km b.s.l. is interpreted as a main solidified intrusive body (Chiarabba *et al*, 2000, Patanè *et al*, 2003). However, a near-vertical shallower plumbing system has been recently inferred at about 4.5 km b.s.l. using deformation data (Bonforte *et al*, 2008 for a review). It is widely accepted that a central magma conduit feeds the near-continuous summit activity, while lateral eruption are triggered by lateral draining of magma from its central conduit. Only few events appear to be independent from the central conduit being fed by peripheral dikes (see Acocella and Neri, 2003 among others).

The recorded eruptive activity for Mt Etna dates back to 1500 B.C. (Tanguy *et al*, 2007). Unfortunately, the eruptive catalog can be considered complete only since 1600 AD for flank eruptions (Mulargia *et al*, 1985). Instead summit activity, was recorded carefully only after the World War II (Andronico and Lodato, 2005) and only after 1970 all summit eruptions were systematically registered (Wadge, 1975, Mulargia *et al*, 1987). Thus the Mt

Etna catalog is considered complete since 1970 AD for summit eruptions. There are several catalogs for Mt Etna eruptions available in the literature, the most recent ones being those compiled by Behncke *et al* (2005), Branca and Del Carlo (2005) and Tanguy *et al* (2007); the Andronico and Lodato (2005) catalog is detailed only for events in the $20th$ century. In this study we use only the flank eruptions since 1600 AD using the Behncke *et al* (2005) catalog as it appears the most complete, at least for volume data. We also integrate and double-check the volume data for the $20th$ century events with the Andronico and Lodato (2005) catalog. The Behcke *et al* (2005) catalog lists events up to 2004/09/07 eruption, so we update it for 2006 AD and 2008 AD eruptions using information available in Burton *et al* (2005) and Behncke *et al* (2008). A raw estimation for the volume of the 2008/05/13 eruption was kindly provided by Marco Neri (Marco Neri personal communication, 2010).

The choice of using only lateral eruptions needs qualification. Although it could be arguable and could explain only one aspect of the eruption activity at Mt Etna volcano, we are pushed in this direction by the quality of data available. Besides, from a statistical point of view, it is better not to use an incomplete dataset with the awareness of the risk of losing one piece of information, than using incomplete data and find false correlations (Bebbington, 2007). Flank eruptions, however, constitute one of the most important threat for a volcanic hazard assessment at Mt Etna (see Behncke *et al*, 2005 and Salvi *et al*, 2006 among others). Thus, in our opinion, the choice of using only flank eruptions seems the best available in a volcanic hazard assessment perspective. In Table 2 the data of flank eruptions at Mt Etna are reported; we indicate the onset date, interevent times (d_{r_i}) and volumes (d_{v_i}). There are 63 eruptive events and consequently 62 pairs of interevent time and volume data.

The next two subsections are organized as follows: we will show first the results obtained for the model parameters both using all data and the sequential procedure discussed in Section 2.5, the ability of the model to fit the data (model checking) and the forecasts obtained. We will compare them with previously published models, when the comparison is possible.

3.2.1. Results for variables and parameters

In order to apply the model to the Mt Etna flank eruptions, first we need to specify the measurements errors ($\Delta d_{r_i}, \Delta d_{v_i}$) and the hyperparameters (μ_b, σ_b^2, μ_c and σ_c^2) for the priors distribution for b and c. In the Behncke et al. (2005) catalog there is no mention about the interevent time errors whereas relative errors are given for volume data. Therefore, we assign an error of 1 day for Δd_{r_i} for interevent times. According to Behncke et al. (2005) we assign relative errors as follows: $\Delta d_{v_i}/d_{v_i} = 0.25$ for $i = 1, \ldots, 43$, $\Delta d_{v_i}/d_{v_i} = 0.05$ for $i = 44, \ldots, 60$ and $\Delta d_{v_i}/d_{v_i} = 0.25$ for $i = 61, 62$. The latter errors are relative to the 2006 and 2008 AD events not in Behncke *et al* (2005) catalog; where volumes are the first raw estimate not reparametrized yet (Marco Neri personal communication, 2010). For the hyperparameters we choose the same parameters as the Kilauea case.

The obtained simulations are presented in Figures 7 and 8. As in the Kilauea case, the model reliably reproduces the assumed measurement errors. In Figure 8 we present the results for the model parameters b, c and λ using all data. As the distribution of b (top left panel in Figure 8) is within [0,1] with mean and standard deviation $\overline{b} = 0.30$ and $\overline{\sigma_b} = 0.04$ respectively, we conclude that Mt Etna flank eruptions follow a generalized time predictable eruptive behavior. For the distribution of c (top right panel) we find a value within [200,460] days/10^6 m^3 with mean $\overline{c}$ = 330days/10^6m^3 and error (1 standard deviation) $\overline{\sigma_c}$ = 40days/10^6 m^3. In the bottom left panel we have the posterior distribution for the time of occurrence λ. This is concentrated in the interval [3.5,8] x10^{-4} days^{-1}. The mean value and standard deviation are $\overline{\lambda} = 5.4 \times 10^{-4}$ days^{-1} and $\overline{\sigma_\lambda} = 0.6 \times 10^{-4}$ days^{-1} respectively. This result is totally compatible with the occurrence time calculated directly by the data with MLE technique, i.e. $\lambda_{MLE} = 4.2 \times 10^{-4}$ days^{-1} with 95 % confidence interval $[3.2, 5.4] \times 10^{-4}$ days^{-1}. Figure 9 presents the sequential estimation of parameters b, c and λ.

From the values corresponding to the posterior distributions of b and c we are lead to speculate about the role played by the magma chamber feeding system in the eruption frequency as we have speculated in Section 3.1.1. Mt Etna volcano seems to act as a non-stationary volcano (Mulargia *et al*, 1987), and the non-stationarity could also imply some sort of cyclicity in the eruption frequency (Behncke and Neri, 2003, Allard *et al*, 2006). This possible non-stationarity should be taken into account in modeling the magma chamber dynamics at Mt Etna volcano.

3.2.2. Model checking and Forecasts

The results of the model check are presented in Figure 10. It is immediate to realize the agreement of the synthetic simulations (blue bars) with values calculated from the data (red bar) for the rate of occurrence, minimum and median. For the rate of occurrence where the p-value=0.94, we can speculate that the model predicts interevent times slightly longer that the observed one. Although the model works well for minimum, median and rate, it is less satisfactorily for the maximum and, as a consequence, for the standard deviation. For these cases the observed value falls in the tails of the predictive distribution. This can be imputed by the fact that the maximum observed interevent time is relative to a long quiescence period from 1702 to 1755 AD and can be considered an extreme value. By considering the second longest interevent time in catalog, i.e. quiescence period from 1614 AD to 1634 AD, it is compatible with the synthetic maximum distribution with p-value=0.7.

Summing up, as for Kilauea data, BH_TPM II model is able to capture the main data features except for the extreme value that fall within the tail of the predictive distribution.

Using the sequential approach discussed in Section 2.5 now we test the forecast ability of the present model. But, before we embark in this comparison, we present the results of the SIR procedure used to resample the λ^j's with the information provided by the erupted volumes. Figure 11 shows the comparison of the MCMC output with the resampled draws. It is clear that the information provided by the volume data in the SIR procedure shrinks and shifts the λ^j distributions

There are several statistical model in literature for the eruptive data series of Mt Etna. The models are: BH_TPM proposed by Passarelli *et al* (2010); A Non Homogeneous Poisson process with a power law intensity proposed by Salvi *et al* (2006); A Non homogeneous Poisson process with piecewise linear intensity by Smethurst *et al* (2009); the GTPM proposed by Sandri *et al* (2005), and the Hidden Markov Models of Bebbington (2007). The latter model allows the detection of change in volcanic activity using

Hidden Markov Models. The activity level of Mt Etna volcano is tested through the onset count data, the interevent time data and the quiescence time data (interonset in the Bebbington 2007 terminology) together with time and size-predictable model. Unfortunately, we were not able to apply the sequential procedure to the Bebbington (2007) model due to its intrinsic complexity, so we do not perform the probability gain test against it.

We have already discussed the BH_TPM and GTPM in the previous sections. Salvi et al (2006) proposed a model based on a non homogeneous Poisson process (NHPP). The intensity of the process has a power law time dependence, whose parameters are estimated using MLE. The intensity can increase or decrease with time, depending on the value of the exponent. This provides the ability to fit any trend in eruptive activity. In Smethurst *et al* (2009), a different (NHPP) was proposed, using a piecewise linear intensity, fitted with numerical MLE. The intensity of the process is constant for eruption before 1970 AD, and then increases linearly with time. The model has a change point that is not easy to handle under our sequential procedure, as the proposed method to estimate it requires the use of all the data. Adding one data point at a time may produce a different estimated change point (see Gasperini *et al*, 1990). In addition, the estimation of the parameters of the process in the Smethurst *et al* (2009) model is subject to numerical stability issues that may complicate a sequential approach.

To tackle the change point problem and compute "forward" probabilities of eruptions, we use two different approaches. The first one is to fix the change point (i.e. 1964 AD) at the values estimated in Smethurst *et al* (2009) and simulate sequentially the other two model parameters.

The second approach consists of empirically estimating the trend for the process intensity, calculated under the sequential procedure. As we show in Figure 12, after the learning phase, we examine and evaluate the trend for the intensity λ_{MLE} (blue stars in the graph), calculated by adding one data at a time, assuming a homogeneous Poisson process. We find that the intensity shows a slow increase with important fluctuations up to the change point found by Smethurst et al. (2009) (black dashed line). Then, after the change point, the intensity rises more markedly. We estimate its trend using linear regressions. In Figure 12, we denote positive significant slopes using green lines. Other cases correspond to red lines. It is clear from the graph that there are no significant trends up to four events after the change point found by Smethurst *et al* (2009). This delay in the detection of the chance point is due to the sequential nature of the forward procedure. Hence to evalute probabilities sequentially, we consider an Homogeneous Poisson process up to four events after the change point of Smethurst *et al* (2009) and then a linearly increasing intensity.

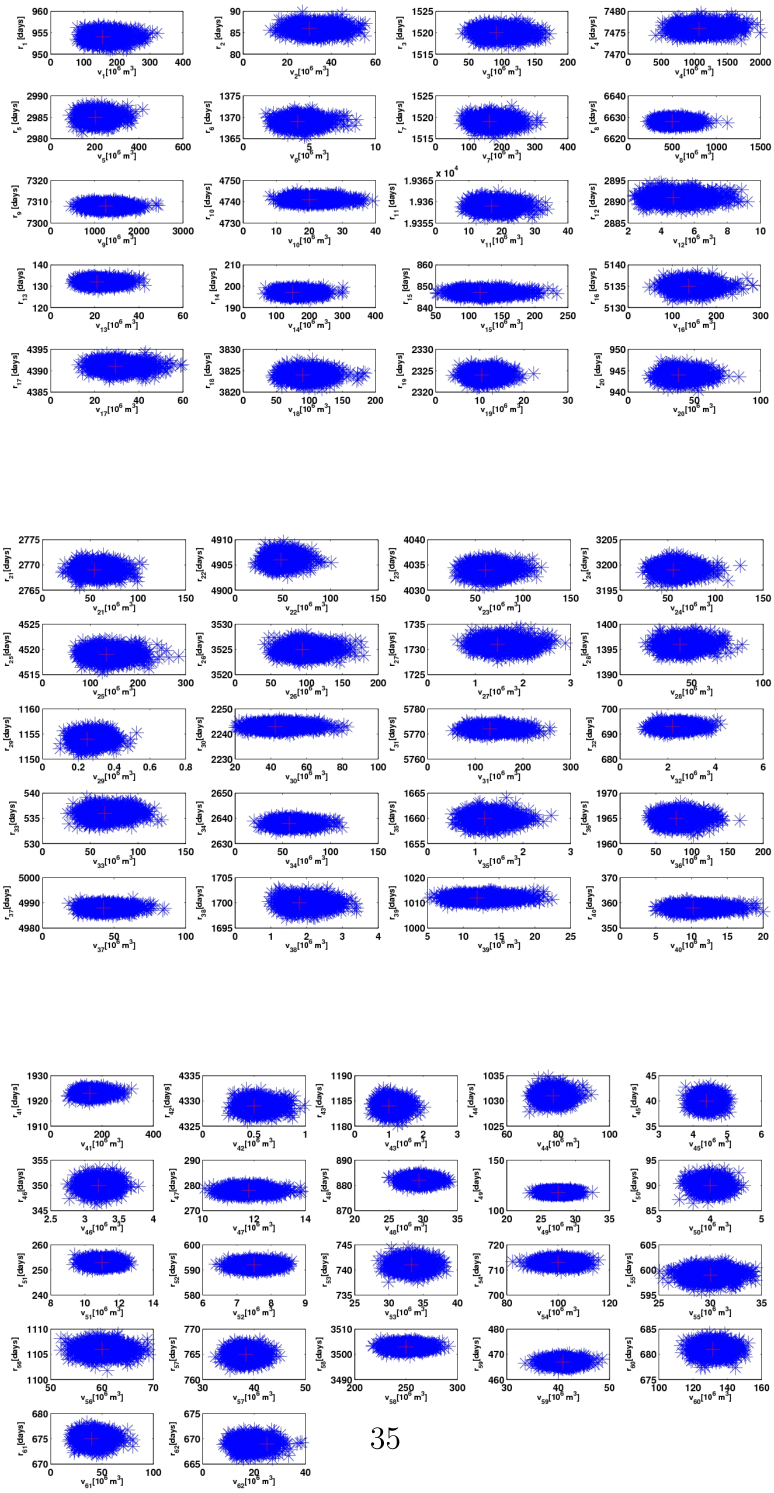

Figure 7: *Same as Figure 1. From top to bottom the first panel corresponds to* r_i *and* v_i $i = 1, \ldots 20$*, the second panel corresponds to* $i = 21 \ldots 40$ *and the third panel corresponds to* $i = 40, \ldots, 62$.

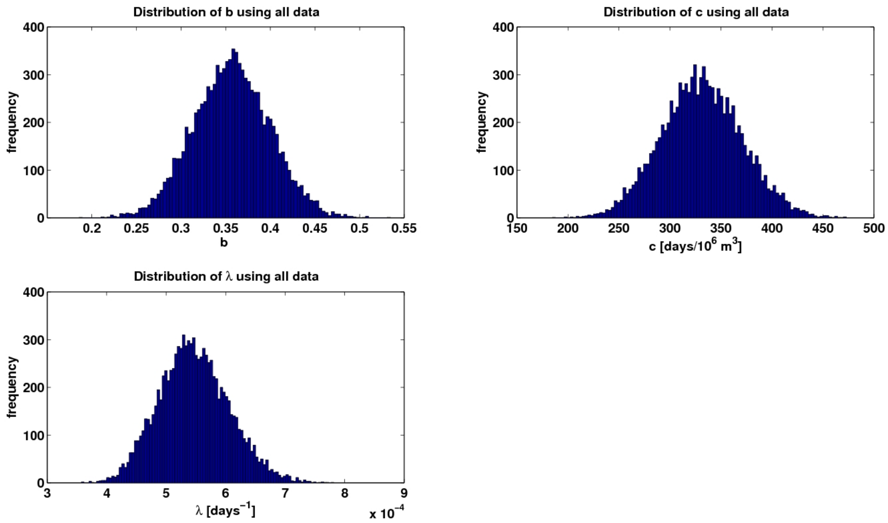

Figure 8: *Posterior distributions for BH_TPMII parameters obtained using all data in the catalog: top left panel refers to b, top right to c and bottom left to λ.*

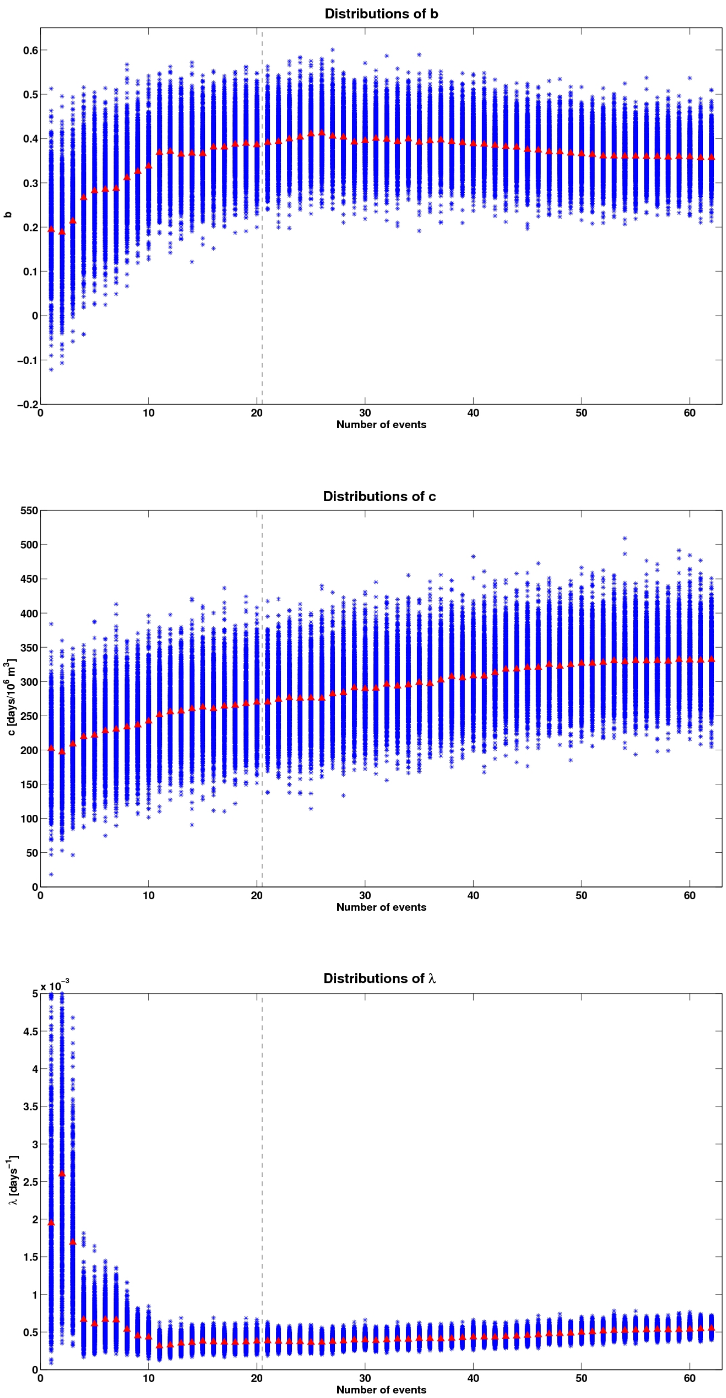

Figure 9: *Posterior distributions of: b, top panel; c, middle panel, and λ, bottom panel. All distributions are calculated using the sequential procedure discussed in the text. Black dashed lines represent the learning phase. Red triangles correspond to the means*

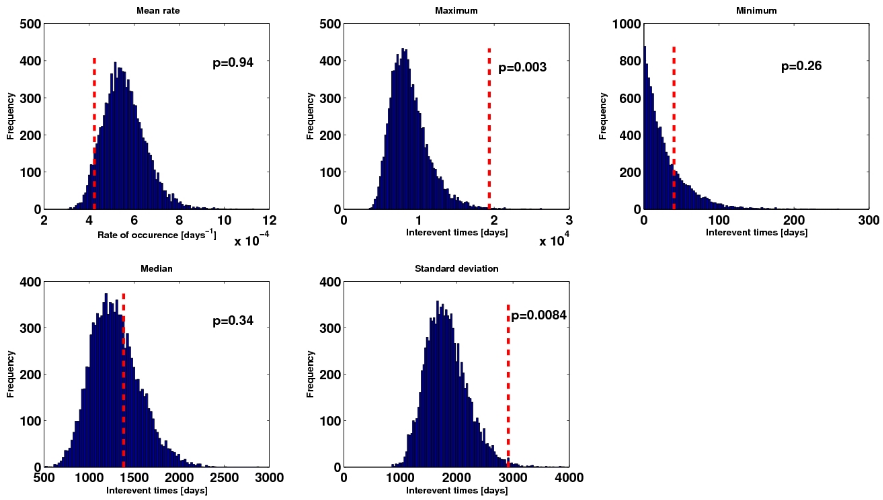

Figure 10: *As Figure 4, histograms of samples from the posterior predictive distributions of several summaries of the interevent times for the Mt Etna (Blue bars). Red dashed lines denote the corresponding observed values. p-values correspond to the proportion of samples above the observed values.*

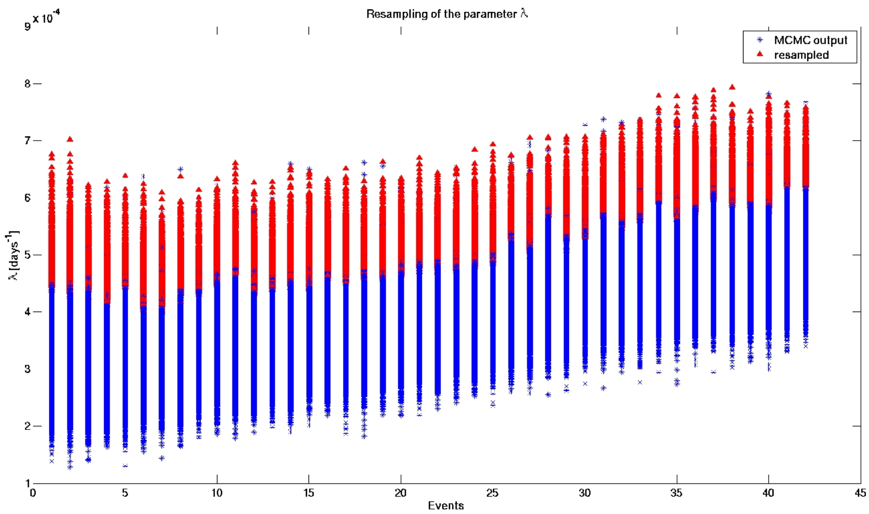

Figure 11: *As Figure 5, the SIR procedure is applied to samples of λ obtained after the learning phase as required for the sequential approach used (i.e. events from 20 to 62). Blue stars correspond to the MCMC output and red ones to resampled draws.*

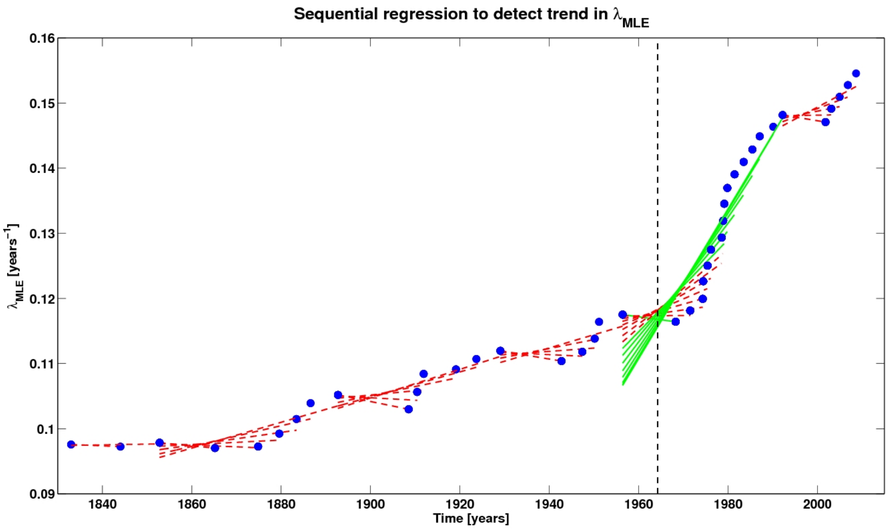

Figure 12: *Trend detection for the intensity of a homogeneous Poisson process using the sequential procedure. Blue stars correspond the intensity calculated sequentially via MLE by adding one data point at a time. Red lines represent non significant regressions (at 1% level), green lines represents significant regressions. The black dashed line is the change point estimated by Smethurst et al 2009. Sequential estimation allows the detection of the change point only four events after the change point found by Smethurst et al., 2009.*

Finally we present the results for the probability gain in Figure 13. As it is shown in the inset of each panel, PG's are always greater than zero, showing the present model performs better than the other ones. In particular, the forecasting test against the homogeneous Poisson process (Panel a) shows only 14 eruptions out of 42 with a negative "punctual" probability gain, corroborating the fact that Mt Etna flank eruptions are non stationary in time (Mulargia *et al* 1987, Bebbington, 2007, Salvi *et al* 2006 and Smethurst *et al*, 2009). In testing against BH_TPM (Panel b), only 17 eruptions have a negative probability gain indicating that modeling Mt Etna interevent times with log-normal distributions does not seem to be the best choice. The result in Panel c against the GTPM is the best one and remarks the limitation of a regression technique in modeling linear relationship between the logarithm of interevent times and of volumes, without using measurement errors. Salvi *et al* (2006) model, in Panel d, performs worse forecasts compared with BH_TPMII, confirming that a power law intensity is not appropriate for Mt Etna eruption occurrences (Smethurst *et al* 2009). In Panel e, the test against the Smethurst *et al* (2009) model, with fixed change point as they found, is the worse one, although the PG is still slightly positive. On one hand, this test shows that modeling the intensity with a linear increasing function for events in the last 40 years seems more appropriate. At the same time, it shows some limitations: a close look at Subplot e shows that event 38 has a very high gain in favor of the BH_TPMII. This event is the 2001 AD eruption, started after 10 years of quiescence. Therefore, the Smethurst *et al* (2009) model, with the ad hoc fitted piecewise linear intensity, could be misleading for real forecasting purposes as the observed eruption frequency decreases in the future. Finally we present, in Panel f, the probability gain against the modified Smethurst *et al* (2009) model following the specification discussed in the previous paragraph for the "forward" application. Here the probability gain is considerably higher than that in Panel e, although the linear intensity fits better the last part of the catalog.

As a summary, it seems that, the BH_TPMII shows better results in forecasting for more than 50% of the eruptive events manifesting a higher reliability. However, we have to remark that the Smethurst *et al* (2009) model is preferable if the Mt Etna flank eruptive frequency keeps increasing in the next years.

We investigate some possible linear relationship between the "punctual" probability gains and the interevent times or volumes using linear regression analysis. We do not find any correlation between volumes and probability gain. The only significant relationship (p-value≤ 0.01) is an inverse linear relationship between "punctual" probability gain calculated against the homogeneous Poisson process and interevent times. The inverse relationship implies that we systematically perform worse forecast for long interevent times. We can justify this results stating that for long quiescence periods the volcano becomes memoryless with transition from open and closed conduit regime (see Marzocchi and Zaccarelli, 2006 and Bebbington, 2007). Another explanation could be related to the complexity of the volcano eruption system not considered in this model. The time predictable model seems more appropriate when the eruptions are close in time. Conversely, when the quiescence period are extremely long, other compelling physical processes may control the volcanic activity. Finally, by neglecting the summit activity we lose one piece of information related to the amount of erupted

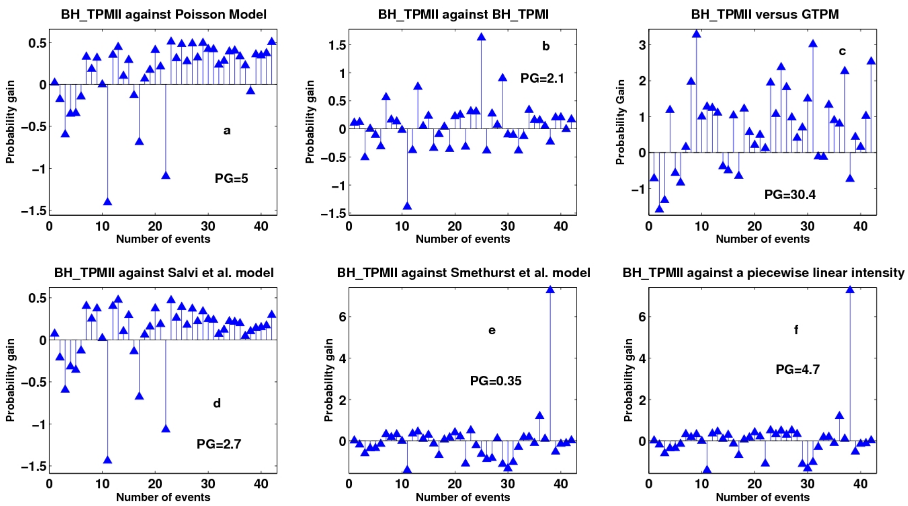

Figure 13: *"Punctual probability gain" of the BH_TPMII for each event after the learning phase with respect to: Poisson Model (Klein, 1982) (Panel **a**); BH_TPM (Passarelli et al, 2010) (Panel **b**); GTPM (Sandri et al, 2005) (Panel **c**); Salvi et al. 2006 (Panel **d**), ; Smethurst et al. 2009 (Panel **e**); Modified piecewise linear model of Smethurst et al, 2009 (Panel **f**). Values greater than zero indicate that BH_TPM model performs better than the reference models. The inset in each panel is the total Probability gain.*

volume from summit crater during the quiescence period. This may introduce a bias that could explain the inverse relationship.

4. Conclusion

In this work we propose a Bayesian Hierarchical model to fit a time predictable model for open conduit volcanoes (BH_TPMII). The use of Bayesian Hierarchical model provides a suitable tool to take into account the uncertainties related to the eruption process as well as those relative to the data, parameters, and variables. We have applied the model to the Kilauea eruptive catalog from 1923 to 1983 AD and to Mount Etna flank eruptions from 1607 to 2008 AD. The results show that both volcanoes have a time predictable eruptive behavior where interevent times depend on the previous volume erupted. The numerical values of the time predictable model parameters inferred, suggest that the amount of the erupted volume could change the dynamics of the magma chamber refilling process during the repose period.

The model shows a good fit with the observed data for both volcanoes and is also able to capture extreme values as a tail behavior of the distributions. The forecasts obtained by BH_TPM II are superior to those provided by other statistical models for both volcanoes. In particular we have improved the forecast performance compared to that of BH_TPM. It is important to notice that a model based on a NHPP, as the one developed in Smethurst et al (2009), could provide better forecast if the flank eruptive activity on Mt Etna keeps increasing in time in the same fashion as it did in the last 40 years; any change from this trend may cause wrong forecasts of the Smethurst et al's (2009) model. Finally, we remark again that the model proposed here may be used for real prospective long-term forecasts to Kilauea and Mount Etna volcano.

Appendices

A. Sampling Importance Resampling algorithm

The Sampling Importance Resampling (SIR) is a non iterative procedure proposed by Rubin (1988). The SIR algorithm generates an approximately independent and identically distributed (i.i.d.) sample of size m from the target probability density function $f(x)$. It starts by generating M ($m \leq M$) random numbers from a probability density function $h(x)$ as inputs to the algorithm. The output is a weighted sample of size m drawn from the M inputs, with weights being the importance weights $w(x)$. As expected, the output of the SIR algorithm is good if the inputs are good ($h(x)$ is close to $f(x)$) or M is large compared to m.

The SIR consists of two steps: a sampling step and an importance resampling step as given below:

1. (Sampling step) generate $X_1, \ldots, X_M$ i.i.d. from the density $h(x)$ with support including that of $f(x)$;
2. (Importance Resampling Step) draw m values $Y_1, \ldots, Y_m$ from $X_1, \ldots, X_M$ with probability given by the importance weights:

$$w^*(X_1, \ldots, X_M) = \frac{w(X_i)}{\sum_{j=1}^{M} w(X_j)} \quad \text{for} \quad i = 1, \ldots, M.$$

   where $w(X_j) = f(X_j)/h(X_j)$ for all j.

The resampling procedure can be done with or without replacement.

References

[48] Acocella, V., Neri, M., 2003. What makes flank eruptions? The 2001 Etna eruption and its possible triggering mechanisms. *Bull Volcanol*, 65:517–529, doi:10.1007/s00445-003-0280-3.

[48] Aki, K., Ferrazzini, V., 2001. Comparison of Mount Etna, Kilauea, and Piton de la Fournaise by a quantitative modeling of their eruption histories, *J. Geophys. Res.*, 106, B3, 4091–4102.

[48] Allard, P., Behncke, B., D'Amico, S., Neri, M., Gambino, S., 2006. Mount Etna 1993-2005: anatomy of an evolving eruptive cycle. *Earth Science Reviews* 78 85–114 doi:10.1016/j.earscirev.2006.04.002.

[48] Andronico, D., Lodato, L., 2005. Effusive activity at Mount Etna volcano (Italy) during the 20th century: a contribution to volcanic hazard assessment.*Nat. Hazards*, 36, 407–443.

[48] Andronico, D., Branca, S.,Calvari, S., Burton, M., Caltabiano, T., Corsaro, A.R., Del Carlo, P., Garf, G., Lodato, L., Miraglia,L., Murè, F., Neri, M., Pecora, E., Pompilio, M., Salerno, G., Spampinato, L., 2005. A multidisciplinary study of the 200203 Etna eruption: insights into a complex plumbing system. 2005. *Bull Volcanol* 67:314–330 doi:10.1007/s00445-004-0372-8.

[48] Bebbington M.S., 2007. Identifying volcanic regimes using Hidden Markov Models. *Geophys. J. Int.* 171, 921-941.

[48] Bebbington M, 2008. Incorporating the eruptive history in a stochastic model for volcanic eruptions. *J. Volcanol. Geotherm. Res.* 175, 325-333.

[48] Bebbington M.S., Lai C. D. 1996a. Statistical analysis of New Zealand volcanic occurrence data, *J. Volcanol. Geotherm. Res.*, 74, 101–110.

[48] Bebbington M.S., Lai C. D. 1996b. On nonhomogeneous models for volcanic eruptions, *Math. Geology*, 28,5, 585–600.

[48] Behncke B, Neri M 2003. Cycles and trends in the recent eruptive behaviour of Mount Etna (Italy). *Can J Earth Sci* 40:1405–1411 doi:10.1139/E03052.

[48] Behncke, B., Neri, M., Nagay, A., 2005. Lava flow hazard at Mount Etna

(Italy): new data from GIS-based study, in *Kinematics and dynamics of lava flow*, Manga, M., and Ventura, G., eds, Geological Society of America, Special Paper 396, 189-208,doi:10.1130/2005.2396(13).

[48] Behncke, B., Calvari, S., Giammanco, S., Neri, M., Pinkerton, H., 2008. Pyroclastic density currents resulting from the interaction of basaltic magma with hydrothermally altered rock: an example from the 2006 summit eruptions of Mount Etna. *Bull Volcanol*, 70:1249–1268 doi:10.1007/s00445-008-0200-7.

[48] Bonforte, A., Bonaccorso, A., Guglielmino, F., Palano, M., Puglisi, G.,2008. Feeding system and magma storage beneath Mt. Etna as revealed by recent inflation/deflation cycles, *J Geophys Res*, 113 B05406, doi:10.1029/2007JB005334.

[48] Branca, S., Del Carlo, P., 2005. Types of eruptions of Etna volcano AD 1670-2003: implication for short term eruptive behavior. *Bull Volcanol* 67: 732 – 742, doi:10.1007/s00445-005-0412-z.

[48] Burt M.L., Wadge G., Scott W.A., 1994. Simple stochastic modelling of eruption history of basaltic volcano: Nyamuragira, Zaire, *Bull. Volcanol.*, 56:87–97.

[48] Burton, M., Neri, M., Andronico, D., Branca, S., Caltabiano, T., Calvari, S., Corsaro, R. A., Del Carlo, P., Lanzafame, G., Lodato, L., Miraglia, L., Salerno, G., Spampinato, L., 2005. Etna 2004-2005: an archetype for geodynamically-controlled effusive eruptions. *Geophys Res Lett*, 32, L09303, doi:10.1029/2005GL022527.

[48] Chiarabba, C., Amato, A., Boschi, E., Barberi, F., 2000. Recent seismicity and tomographic modeling of the Mount Etna plumbing system. *J Geophys Res*, 105, B5, 10923–10938.

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

[48] Dvorak, J.J., Dzurisin, D., 1993. Variations in magma-supply rate at Kilauea volcano, Hawaii, *J. Geophys. Res.*, 98, 22255-22268.

[48] Gasperini, P, Gresta, S, Mulargia, F, 1990. Statistical analysis of seis-

mic and eruptive activities at Mount Etna during 1978-1987, *J Volcanol Geotherm Res*, 40, 317 – 325.

[48] Gelman,A., 1996 Inference and monitoring convergence, in *Markov chain Monte Carlo in practice*, pp 131–143, eds Gilks W.R., Richardson S., Spiegelhalter D.J., Chapman & Hall, London, 2nd edn.

[48] Gelman A., Carlin J.B., Stern H.S., Rubin D.B.,2000. *Bayesian Data Analysis*,1st edn, Chapman & Hall/CRC,Boca Raton-Florida.

[48] Gilks W.R., Richardson S., Spiegelhalter D.J., 1996. Introducing Markov chain Monte Carlo, in *Markov chain Monte Carlo in practice*, pp 1–19, eds Gilks W.R., Richardson S., Spiegelhalter D.J., Chapman & Hall, London, 2nd edn.

[48] Gudmundsson A, 1986. Possible effect of aspect ratios of magma chambers on eruption frequency. *Geology* 14, 991-994.

[48] Holcomb, R.T.,Eruptive history and long term behavior of Kilauea volcano. In *Volcanism in Hawaii*, Decker R.W., Wrigth, T.L., Stauffer, P.H. (Eds), U.S. Geol. Surv. Prof. Pap. 1350 pp. 261–350.

[48] Kagan Y.Y., Knopoff L., 1987. Statistical Short-Term Earthquake Prediction, *Science*, 236(4808), 1563–1567

[48] Klein F.W.,1982 Patterns of historical eruptions at Hawaiian Volcanoes, *J. Volcanol. Geotherm. Res.*, 12 , 1–35

[48] Klein F.W., Koyanagi,R.Y., Nakata, J.S., Tanigawa, W.R., The seismicity of Kilauea's magma system, in *Volcanism in Hawaii*, pp. 1019–1185 U.S. Geol. Surv. Prof. Pap. 1350.

[48] Marzocchi W.,1996. Chaos and stocasticity in volcanic eruptions the case of Mount Etna and Vesuvius, *J. Volcanol. Geotherm. Res.*,70, 205-212.

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

[48] Marzocchi, W., Sandri, L., Gasparini, P., Newhall, C. G., Boschi, E., 2004. Quantifying probabilities of volcanic events: the example of volcanic hazard at Mt. Vesuvius. *J. Geophys. Res.*, 109, B11201, doi:10.1029/2004JB003155.

[48] Marzocchi W., Sandri L., Selva J.,2008. BET_EF: a probabilistic tool for long- and short-term eruption forecasting, *Bull. Volcanol.*, 70, 623–632, doi:10.1007/s00445-007-0157-y

[48] Mulargia F.,Tinti S., Boschi E.,1985. A statistical analysis of flank eruptions on Etna Volcano, *J Volcanol Geotherm Res* , 23, 263–272.

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

[48] Newhall C.G., Hoblitt,R.P., 2002. Constructing event trees for volcanic crises, *Bull. Volcanol.*, 64, 3–20, doi:10.1007/s004450100173.

[48] Patanè, D., De Gori, P., Chiarabba, C., Bonaccorso, A., 2003. Magma ascent and the pressurization of Mount Etna's volcanic system, *Science*, 299, 29 March 2003.

[48] Passarelli, L., Sandri, L., Bonazzi, A., Marzocchi, W., 2010. Bayesian Hierarchical Time Predictable Model for eruption occurrence: an application to Kilauea Volcano, published on-line Apr 8 2010 *Geophys J Int* doi:10.1111/j.1365-246X.2010.04582.x

[48] Rubin, D. B. 1988. Using the SIR algorithm to simulate posterior distributions, in *Bayesian Statistics-3*, pp. 395402, eds Bernardo, J.W., DeGroot, M. H., Lindley, D. V., Smith, A.F.M. Oxford: Oxford University Press.

[48] 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, doi:10.1016/j.jvolgeores.2006.01.002.

[48] Sandri L., Marzocchi W., Gasperini P., 2005. Some inside on the occurrence of recent volcanic eruptions of Mount Etna volcano (Sicily,Italy), *J. Geophys. Res.*, 163, 1203–1218.

[48] 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.

[48] Smith, A.F.M., Gelfand, A.E., 1992. Bayesian statistics without fear: a

sampling resampling prespective, *Am Statist*, 46, 84–88.

[48] Sparks, R.S.J., 2003. Forecasting volcanic eruptions,*Earth Planet. Sci. Lett.*, 210, 1–15.

[48] Takada, A., 1999. Variations in magma supply and magma partitioning: the role of the tectonic settings, *J. Volcanol. Geotherm. Res.*, 93, 93–110.

[48] Tanguy, J.C., Condomines, M., Kieffer, G., 1997. Evolution of the Mount Etna magma: constraints on the present feeding system and eruptive mechanism, J Volcano Geotherm Res, 221–250.

[48] Tanguy, J.C., Condomines, M.,Le Goff,M., Chillemi, V., La Delfa, S., Patanè, G., 2007. Mount Etna eruptions of the last 2,750 years: revised chronology and location through archeomagnetic and ^{226}Ra-^{230}Th dating. *Bull Volcanol* 70:55–83, doi:10.1007/s00445-007-0121-x.

[48] Wadge, G., Walker, W.P.L., Guest, J.E., 1975. The output of Etna volcano, *Nature*, 255, 385–387.

[48] Wikle C. K.,2003. Hierarchical Models in Environmental Science, *Internat. Statist. Rev.*, 71-2, 181–199.

Table 1: Catalog of eruptive events at Kilauea volcano

Eruption #	**Onset** yyyymmdd	**Interevent time** [days]	**Volume** lava e tephra $[10^6 m^3]$
1	1923 08 25	259	0.073
2	1924 05 10	70	0.79
3	1924 07 19	1083	0.234
4	1927 07 07	594	2.30
5	1929 02 20	155	1.40
6	1929 07 25	482	2.60
7	1930 11 19	399	6.20
8	1931 12 23	988	7.00
9	1934 09 06	6504	6.90
10	1952 06 27	703	46.70
11	1954 05 31	273	6.20
12	1955 02 28	1720	87.60
13	1959 11 14	60	37.20
14	1960 01 13	408	113.20
15	1961 02 24	7	0.022
16	1961 03 03	129	0.26
17	1961 07 10	74	12.60
18	1961 09 22	441	2.20
19	1962 12 07	257	0.31
20	1963 08 21	45	0.80
21	1963 10 05	517	6.60

Eruption #	**Onset**	**Interevent time**	**Volume**
	yyyymmdd	[days]	lava e tephra [$10^6 m^3$]
22	1965 03 05	294	16.80
23	1965 12 24	681	0.85
24	1967 12 05	291	80.30
25	1968 08 22	46	0.13
26	1968 10 07	138	6.60
27	1969 02 22	91	16.10
28	1969 05 24	812	185.00
29	1971 08 14	41	9.10
30	1971 09 24	132	7.70
31	1972 02 03	457	162.00
32	1973 05 05	189	1.20
33	1973 11 10	251	2.70
34	1974 07 19	62	6.60
35	1974 09 19	103	10.20
36	1974 12 31	333	14.30
37	1975 11 29	654	0.22
38	1977 09 13	794	32.90
39	1979 11 16	896	0.58
40	1982 04 30	148	0.50
41	1982 09 25	100	3.00
42	1983 01 03		ongoing

Table 2: Catalog of eruptive events at Mount Etna volcano

Eruption #	Onset	Interevent time	Volume
	yyyymmdd	[days]	lava e tephra [$10^6 m^3$]
1	1607 06 28	954	158.00
2	1610 02 06	86	30.00
3	1610 05 03	1520	91.71
4	1614 07 01	7476	1071.00
5	1634 12 19	2985	203.03
6	1643 02 20	1369	4.12
7	1646 11 20	1519	162.45
8	1651 01 17	6628	497.53
9	1669 03 11	7308	1247.50
10	1689 03 14	4741	20.00
11	1702 03 08	19359	16.94
12	1755 03 09	2891	4.73
13	1763 02 06	132	21.08
14	1763 06 18	197	149.96
15	1764 01 01	847	117.20
16	1766 04 27	5135	137.25
17	1780 05 18	4391	29.35
18	1792 05 26	3824	90.13
19	1802 11 15	2324	10.43
20	1809 03 27	944	38.19

Eruption #	**Onset** yyyymmdd	**Interevent time** [days]	**Volume** lava e tephra $[10^6 m^3]$
21	1811 10 27	2769	54.33
22	1819 05 27	4906	47.92
23	1832 10 31	4034	60.74
24	1843 11 17	3199	55.70
25	1852 08 20	4519	134.00
26	1865 01 03	3525	94.33
27	1874 08 29	1731	1.47
28	1879 05 26	1396	41.93
29	1883 03 22	1154	0.25
30	1886 05 19	2243	42.52
31	1892 07 09	5772	130.58
32	1908 04 29	693	2.20
33	1910 03 23	536	65.20
34	1911 09 10	2638	56.60
35	1918 11 30	1660	1.20
36	1923 06 17	1965	78.50
37	1928 11 02	4988	42.50
38	1942 06 30	1700	1.80
39	1947 02 24	1012	11.90
40	1949 12 02	358	10.20
41	1950 11 25	1923	152.00
42	1956 03 01	4329	0.50
43	1968 01 07	1184	1.00

Eruption #	**Onset**	**Interevent time**	**Volume**
	yyyymmdd	[days]	lava e tephra [$10^6 m^3$]
44	1971 04 05	1031	78.00
45	1974 01 30	40	4.40
46	1974 03 11	350	3.20
47	1975 02 24	278	11.80
48	1975 11 29	882	29.40
49	1978 04 29	118	27.50
50	1978 08 25	90	4.00
51	1978 11 23	253	11.00
52	1979 08 03	592	7.50
53	1981 03 17	741	33.30
54	1983 03 28	713	100.00
55	1985 03 10	599	30.03
56	1986 10 30	1106	60.00
57	1989 11 09	765	38.40
58	1991 12 14	3503	250.00
59	2001 07 17	467	40.90
60	2002 10 27	681	131.50
61	2004 09 07	675	40.00
62	2006 07 14	669	25.00
63	2008 05 13		35.00