ANALYSIS OF ERUPTIVE AND SEISMIC SEQUENCES TO IMPROVE THE SHORT- AND LONG-TERM ERUPTION FORECASTING

Ph.D. Thesis of:

Alexander Garcia Aristizabal

Tutor:

Dr. Warner Marzocchi

Ph.D. Coordinator:

Prof. Michele Dragoni

General abstract

Forecasting the time, location, nature, and scale of volcanic eruptions is one of the most urgent aspects of modern applied volcanology. The reliability of probabilistic forecasting procedures is strongly related to the reliability of the input information provided, implying objective criteria for interpreting the historical and monitoring data. For this reason both, detailed analysis of past data and more basic research into the processes of volcanism, are fundamental tasks of a continuous information-gain process; in this way the precursor events of eruptions can be better interpreted in terms of their physical meanings with correlated uncertainties. This should lead to better predictions of the nature of eruptive events.

In this work we have studied different problems associated with the long- and short-term eruption forecasting assessment. First, we discuss different approaches for the analysis of the eruptive history of a volcano, most of them generally applied for long-term eruption forecasting purposes; furthermore, we present a model based on the characteristics of a Brownian passage-time process to describe recurrent eruptive activity, and apply it for long-term, time-dependent, eruption forecasting (Chapter 1).

Conversely, in an effort to define further monitoring parameters as input data for short-term eruption forecasting in probabilistic models (as for example, the Bayesian Event Tree for eruption forecasting -BET_EF-), we analyze some characteristics of typical seismic activity recorded in active volcanoes; in particular, we use some methodologies that may be applied to analyze long-period (LP) events (Chapter 2) and volcano-tectonic (VT) seismic swarms (Chapter 3); our analysis in general are oriented toward the tracking of phenomena that can provide information about magmatic processes. Finally, we discuss some possible ways to integrate the results presented in Chapters 1 (for long-term EF), 2 and 3 (for short-term EF) in the BET_EF model (Chapter 4).

TO FILIPPO

Acknowledgments

Contents

List of Figures

List of Tables

Introduction

Forecasting the time, location, nature, and scale of volcanic eruptions is one of the most urgent aspects of modern applied volcanology. However, volcanoes are complex physical systems in which a completely deterministic description of the processes occurring before or during an eruption is practically impossible. Studies on active volcanoes involve different approaches that, in function of the time scale of interest, may provide useful information for eruption forecasting. For example, research that deals with documentation of the past eruptive activity by historic records, geologic mapping, stratigraphic studies, etc., may be the base for long-term eruption-forecasting and volcanic-hazard studies, as e.g., *Wickman* (1976); *Klein* (1982); *Mulargia et al.* (1985); *De la Cruz-Reyna* (1991); *Ho* (1991); *Burt et al.* (1994); *Bebbington and Lai* (1996b); *Marzocchi and Zaccarelli* (2006). This is a typical problem in which we do not have direct access to the physical processes, but we can have a record of the response of the system; in particular, if some characteristic properties of the response of the system (e.g. repose times) can be associated with a random variable, and if it is possible to express a probability function for the random variable, then it is possible to define a probabilistic model for the response of the considered system.

Conversely, if we are interested in short-term eruption forecasting, the use of monitoring information is fundamental since it may allow to identify changes in geophysical and geochemical parameters before, during, and after volcanic eruptions. Within this area we can find the bulk of geophysical research on volcanoes, including applications from different areas as volcano seismology, geodesy, geochemistry, etc. A general description of the most used geophysical and geochemical monitoring techniques may be found, for example, in texts as *Gasparini et al.* (1992); *Scarpa and Tilling* (1996); *Marzocchi and Zollo* (2008).

Given the extreme complexities and nonlinearities of volcanic processes, the combination of both –past activity records and monitoring data– provides the most comprehensive and logical basis to achieve the maximum possible level of knowledge about specific volcanic systems, but also for the forecasting in a probabilistic framework of their future activity. Some physical processes (e.g. planetary mechanics) are so regular that their histories can be successfully extrapolated into the future, and predictions can thus be done just based on a deterministic

basis. Conversely, physical phenomena such as volcanic eruptions are much more variable in nature, and in most of the cases extrapolating their past and current processes into the future can only be done (if possible at all) on a probabilistic basis; however, deterministic rules can be introduced in probabilistic models to improve predictions, then in this context probabilistic and deterministic approaches are not incompatible since the former, more general, can include information from deterministic models.

Beyond the natural scientific interest that volcanoes arise, there is a social fact that highlights the importance of the research on volcanoes. Millions of people around the world live close to an active volcano, and authorities in many places are becoming aware of the problem; for this reason there is an urgent need of reliable quantitative models for long- and short-term eruption forecasting and for volcanic hazard assessment.

Preamble of the thesis: The Bayesian Event Tree Model for Eruption Forecasting

The Bayesian Event Tree model (BET) of *Marzocchi et al.* (2004, 2008) is a quantitative tool to calculate and visualize probabilities related to eruption forecasting (EF) and volcanic hazard (VH) assessment. It is based on a fully probabilistic Bayesian scheme and introduces a Fuzzy approach to manage monitoring measurements, which provides the advantage of being applicable at different time scales. For example, during a quiet period of the volcano, long-term EF is estimated by accounting mainly for the past activity of the volcano; conversely, during unrest, the method allows mid- to short-term EF to be estimated by considering different patterns of pre-eruptive phenomena (e.g., *Newhall and Hoblitt*, 2002; *Marzocchi et al.*, 2004, 2008). Short- and long-term concepts refer to the expected characteristic time in which the volcanic processes show significant variations (and must not be confused with the forecasting time window): for instance, during unrest the time variations occur in time scales much shorter than the changes expected during a quiet phase of the volcano (for details see, e.g., *Marzocchi et al.*, 2008). In summary, BET applied to EF (BET_EF) can handle both volcanological and monitoring information, and may be dynamically used for long- and short-term eruption forecasting; furthermore, it takes properly and explicitly into account the epistemic (data- or knowledge-limited) and aleatory (stochastic) uncertainties: this guarantees reliable outputs, given reliable input information. For a general overview of the mathematical background, see Appendix A, and/or the electronic supplementary material in *Marzocchi et al.* (2008).

Aims and Outline of the thesis

The reliability of probabilistic forecasting procedures is strongly related to the reliability of the input information provided, implying objective criteria for interpreting the historical and monitoring data. For this reason both, detailed analysis of past data and more basic research

into the processes of volcanism, are fundamental tasks of a continuous information-gain process so that the precursor events of eruptions can be better interpreted in terms of their physical meanings with correlated uncertainties. This should lead to better predictions of the nature, time, location, and magnitude of eruptive events.

In this work we have studied different problems associated with the long- and short-term eruption forecasting assessment. The topics discussed in this Ph.D. thesis are the result of a close collaboration with many scientists, in particular: Dr. Warner Marzocchi (tutor, INGV-Roma), Dr. Jacopo Selva (INGV-Bologna), Dr. Laura Sandri (INGV-Bologna), Dr. Anna Maria Lombardi (INGV-Roma), and Dr. Eisuke Fujita (NIED-Japan); each one of them are coauthors of at least one of the research topics produced (paper-drafts already submitted or in preparation).

- **Long-term Eruption Forecasting**

  In **Chapter 1** we discuss different approaches for the analysis of the eruptive history of a volcano, most of them generally applied for long-term eruption forecasting purposes; in particular we compare and test different models as time-predictable, size-predictable, Poisson, Lognormal, Gamma, Weibull, and Inverse Gaussian. We present a model to describe recurrent eruptive activity, and apply it for long-term, time-dependent, eruption forecasting. This physically-motivated probabilistic model is based on the characteristics of the Brownian passage-time distribution; the physical process defining this model can be described by the steady rise of a state variable from a ground state to a failure threshold; adding Brownian perturbations to the steady loading produces a stochastic load-state process (a Brownian relaxation oscillator) in which an eruption relaxes the load state to begin a new eruptive cycle. The Brownian relaxation oscillator and Brownian passage-time distribution connect together physical notions of unobservable loading and failure processes of a point process with observable response statistics. The Brownian passage-time model is parameterized by the mean rate of event occurrence, μ, and the aperiodicity about the mean, α. We apply this model to analyze the eruptive history of Miyakejima volcano, Japan, and conclude that it provides a satisfactory description of the data, with better performance respect to other models often used in literature. Furthermore, since it is a physically-motivated model, it provides also an insight into the macro-mechanical processes driving the system. The work presented in Chapter 1 has been submitted (December 2009) to the *Bulletin of Volcanology*.

- **Volcano Seismology and Short-term Eruption Forecasting**

  In an effort to define further monitoring parameters as input data for short-term eruption forecasting in probabilistic models as BET, we analyze some characteristics of typical seismic activity recorded in active volcanoes; in particular, we use some methodologies that may be applied to analyze long-period (LP) events (Chapter 2) and volcano-

tectonic (VT) seismic swarms (Chapter 3); our analysis in general are oriented toward the tracking of phenomena that can provide information about magmatic processes.

The rationale behind the choice of volcano seismology data for our analysis is because nearly every recorded volcanic eruption has been preceded and/or accompanied by seismic activity beneath or near the volcano (e.g., *Aki*, 1992; *McNutt*, 1996). Seismology may be considered as one of the most important tools for monitoring volcanoes since it can provide information (i.e. strong seismic parameters) that, considered together with other monitoring parameters, can help volcanologists both to improve their understanding of some physical processes occurring inside volcanoes, and to produce better data to feed probabilistic eruption forecasting models. In this context, volcano seismology has a fundamental role since (1) it is probably the geophysical tool more widely used for the monitoring of active volcanoes, and (2) one of its central objectives is to understand the nature and dynamics of seismic sources associated with the injection and transport of magma and related hydrothermal fluids.

Active volcanoes are the source of a great variety of seismic signals, this is because volcanic seismic sources involve different kinds of physical processes. Two basic families of processes can usually be established to classify volcanic seismicity (e.g., *Chouet*, 1996): the first family consists of volumetric sources in which the fluids play an active role in the generation of elastic waves (that in a rough generalization we call Long-Period (LP) family), and the second consists of shear or tensile sources involving brittle rock failure (Volcano-Tectonic (VT) family).

In **Chapter 2** we consider LP seismicity; the source mechanisms associated with LP seismic events and volcanic tremor are intimately associated with fluids contained in a solid rock matrix. Both the frequency content and the characteristic long-lasting coda of these seismic signals are of fundamental importance since they represent evidences of the characteristic properties of the source, as have been highlighted by many authors using diverse models (e.g., *Aki et al.*, 1977; *Chouet*, 1981, 1982, 1985, 1988; *Ferrazzini and Aki*, 1987; *Fujita et al.*, 1995; *Neuberg et al.*, 2000; *Morrissey and Chouet*, 2001; *Fujita and Ida*, 2003). In particular, in Chapter 2 we present an alternative method based on probabilistic inverse theory to estimate the parameters of a simple conceptual model (presented by *Kumazawa et al.* (1990) about twenty years ago) as a method for physically-based spectral analysis. The probabilistic formulation leads to the definition of a (posterior) probability distribution in the model space $\sigma(\mathbf{m})$ which results of combining prior information (on data and model parameters) with new information obtained by measurements (the time series). This kind of formulation may be helpful for different kinds of geophysical problems, specially when it exists a nonlinear relationship between the observed data and the parameters of a given model. Using this formulation we have performed some numerical tests using synthetic time series, and we did apply the method to analyze the waveforms of Long-Period events from two

volcanoes: Cotopaxi (Ecuador) and Miyakejima (Japan). The procedure described may be an important tool for objectively monitor the characteristic complex frequencies of volcanic LP events; in this way, changes in the volcanic activity may be highlighted and families of events of particular interest could be identified for further analysis. Beyond this possible application for monitoring, our approach can also be used to analyze individual events as an alternative to the existent methodologies. The work presented in Chapter 2 has been submitted (January 2010) to the *Geophysical Journal International*.

On the other hand, in **Chapter 3** we analyze some properties of VT seismic swarms using a non-stationary epidemic-type, aftershock sequences (ETAS) modeling; we follow a procedure based on the model proposed by *Lombardi et al.* (2006) to analyze complex seismic swarms. We analyze seismic swarms in both tectonic and volcanic environments, in order to extract information useful to characterize (VT) swarms directly associated with dike migration. The final goal of this analysis is to get insights for a quantitative tool useful to interpret VT swarms in almost real-time during monitoring procedures, yielding physical constrains of the driving processes. To explore possible characteristic properties of VT swarms during dike intrusions, we analyze and compare different earthquake swarms selected from three general groups: (1) swarms from purely tectonic environments, (2) swarms occurred in volcanic areas but that did not preceded nor accompany eruptive activity, and (3) swarms in volcanic areas before/during an eruptive process. Then we characterize VT swarms of the third group in terms of the parameters of a non-stationary ETAS model in which both the background seismicity rate (λ_0) and p-value are allowed to change through time. The work presented in Chapter 3 at the moment is in preparation to be submitted to a international JCR journal.

- **Integration in the Bayesian Event Tree (BET_EF)**

In **Chapter 4** we discuss some possible ways to integrate the results presented in Chapters 1 (for long-term EF), 2, and 3 (for short-term EF) in the BET_EF model. As discussed before, BET is a probabilistic model that merges all kinds of volcanological information to obtain probability of any relevant volcanic event, and handles both volcanological and monitoring information which allows to use it dynamically for long- and short-term eruption forecasting. For this reason BET is a valid platform to integrate the results of this work; for instance, the Brownian model described on Chapter 1 is an example of how a time-dependent, long-term probabilistic model based on the eruptive history of the volcano can be derived, which could be integrated within the long-term analysis of BET. On the other hand, the results obtained from volcano seismology analysis in Chapters 2 and 3, may be integrated as monitoring parameters that, together with parameters from other disciplines, may provide information to improve the short-term eruption forecasting assessment during a volcanic crisis. The strategy adopted here to introduce the information in the monitoring part of BET is based on the Fuzzy set

theory; in this way, two monitoring parameters are defined, which are the result of the respective analysis performed on LP and VT data, and the use of both theoretical models for interpretation, and Fuzzy logic to translate this information into *degrees* of anomaly for BET.

Chapter 1

A Brownian Model for Recurrent Volcanic Eruptions: an Application to Miyakejima Volcano (Japan)

"We have knowledge of the past, but we can't control it. We can control the future, but we have no knowledge of it".

Claude Shannon (1960)

Abstract

The definition of probabilistic models as mathematical structures to describe the response of a volcanic system is a plausible approach to characterize the temporal behavior of volcanic eruptions, and constitutes a tool for long-term eruption forecasting. This kind of approach is motivated by the fact that volcanoes are complex systems in which a completely deterministic description of the processes preceding eruptions is practically impossible. To describe recurrent eruptive activity we apply a physically-motivated probabilistic model based on the characteristics of the Brownian passage-time distribution; the physical process defining this model can be described by the steady rise of a state variable from a ground state to a failure threshold; adding Brownian perturbations to the steady loading produces a stochastic load-state process (a Brownian relaxation oscillator) in which an eruption relaxes the load state to begin a new eruptive cycle. The Brownian relaxation oscillator and Brownian passage-time distribution connect together physical notions of unobservable loading and failure processes of a point process with observable response statistics. The Brownian passage-time model is parameterized by the mean rate of event occurrence, μ, and the aperiodicity about the mean, α. We apply this model to analyze the eruptive history of Miyakejima volcano, Japan, and conclude that it provides a satisfactory description of the data, with better performance respect to

other models often used in literature (e.g. Weibull, Gamma, Lognormal, etc.). Furthermore, since it is a physically-motivated model, it provides also an insight into the macro-mechanical processes driving the system.

1.1   Introduction

Volcanoes can be viewed as complex physical systems in which a completely deterministic description of the processes occurring before or during an eruption is practically impossible. This fact motivates the definition and development of probabilistic models as mathematical structures to describe physical phenomena: this is a typical problem in which we do not have direct access to the physical processes, but we can have a record of the response of the system. In particular, if some characteristic properties of the response of the system (e.g. event times) can be associated with a random variable, and if it is possible to express a probability function for the random variable, then it is possible to define a probabilistic model for the response of the considered system.

A time series of eruptions from a single volcano can be treated as a stochastic point process with individual eruptions as (random) independent events in time. Statistical analysis of both repose time and erupted volume catalogs have been performed for a large number of volcanoes (e.g., *Wickman*, 1976; *Klein*, 1982; *Mulargia et al.*, 1985, 1987; *De la Cruz-Reyna*, 1991; *Burt et al.*, 1994; *Marzocchi and Zaccarelli*, 2006), mainly for those with frequent eruptive activity and where detailed catalogs exist. The main objective of this kind of analysis is to develop probabilistic models to understand the past eruptive activity of the volcano and to forecast its future behavior. Despite some general models for eruption occurrences exist, we argue that a more skillful forecast can be achieved trying also to capture the peculiarities that characterize the behavior of each single volcano. For this reason it is often necessary to perform tests for different trial models in order to identify which one explains better our data.

When treating eruptions as events in time, several simplifying assumptions must be made (*Klein*, 1982; *Ho*, 1991): although the onset date of an eruption is generally well constrained, the duration is harder to determine. In our analysis, we ignore eruption duration since we take the onset date as the most physically meaningful, and measure repose times from one onset date (of an eruptive cycle) to the next. In this way, our modeling intends to describe the waiting times of the long-term physical processes governing *renewed volcanic activity* (understood as episodes of new magmatic intrusions entering in the volcanic system and triggering new eruptive cycles). Once the volcanic system has been perturbed and a new eruptive cycle has started, the short-term behavior of the eruptive activity may follow different patterns during the gradual decline of activity; in this context, sporadic eruptive episodes in a short time window (respect to the repose time) after the onset of a new eruptive cycle cannot be described using the former long-term model. Thus, the definition of *repose time* from this point of view is not exactly equivalent to the classic concept of *non-eruptive period*; this assumption seems justified because most eruption durations are much shorter that typical effective repose

intervals (e.g., *Klein*, 1982).

Up to now distinct conceptual models have been proposed to describe the eruptive behavior of different volcanoes around the world. The most frequent solutions describe the eruptive activity in terms of (1) a homogeneous Poisson processes in time domain (e.g., *Klein*, 1982; *De la Cruz-Reyna*, 1991; *Marzocchi and Zaccarelli*, 2006), (2) Time-Predictable processes (e.g., *Burt et al.*, 1994; *Sandri et al.*, 2005) or (3) Size-Predictable processes (e.g., *Burt et al.*, 1994; *Marzocchi and Zaccarelli*, 2006). In our analysis, none of these existing models successfully explains the eruptive activity of Miyakejima volcano, which seems to show a more regular behavior. A Non-homogeneous Poisson process modeled using a Weibull process (e.g., *Ho*, 1991; *Bebbington and Lai*, 1996a,b) provides a quite better explanation of the observed data, however the Weibull process possesses some undesirable features that are difficult to explain in volcanological applications (e.g., *Bebbington and Lai*, 1996a).

In this chapter we present an analysis of the time series of repose times (as defined before), τ, and eruption volumes of Miyakejima volcano. In the first part we perform a completeness analysis of the catalog to define the period of uniformity in the data. Then we test different physically-based, trial models to evaluate which one provides the best description of the observations, and discuss their features for practical applications in eruption forecasting.

1.2 Miyakejima volcano and data set

1.2.1 Miyakejima volcano

Miyakejima island, located about 200 km south of Tokyo (Fig. 1.1), is one of the most active basaltic volcanoes in Japan. Its recurrent eruptive behavior has been suspected but up to now a detailed quantitative analysis based on its past activity has not been performed. In most historical eruptions, basaltic magma and scoria erupted mainly from flank fissures (*Tsukui and Suzuki*, 1998) and most eruptions lasted a short time (a day to a month). The latest eruptive cycle started in June 2000 and a caldera was formed at the submit; Since then, the volcano has been showing high activity levels for more than 9 years by now. On the basis of surface phenomena observed, many authors have divided this eruptive period in at least four stages (e.g., *Nakada et al.*, 2005; *Ueda et al.*, 2005): (1) magmatic intrusion (1 day), (2) summit subsidence (10 days), (3) Explosion (40 days), and (4) gas emissions accompanied by small seismic swarms, deformation and explosions (>9 years). The total volume of tephra erupted was about $0.009 km^3 (DRE)$, which is much smaller than the volume of the resulting caldera ($0.6 km^3$) (*Nakada et al.*, 2005).

Here we analyze a data set containing the repose periods and volumes of lava and tephra emitted by Miyakejima volcano based on the data published by *Tsukui and Suzuki* (1998) and from the Global Volcanism Program catalog (*Simkin and Siebert*, 2002-). The data set was updated introducing information of the last eruption (June 2000) from *Nakada et al.*
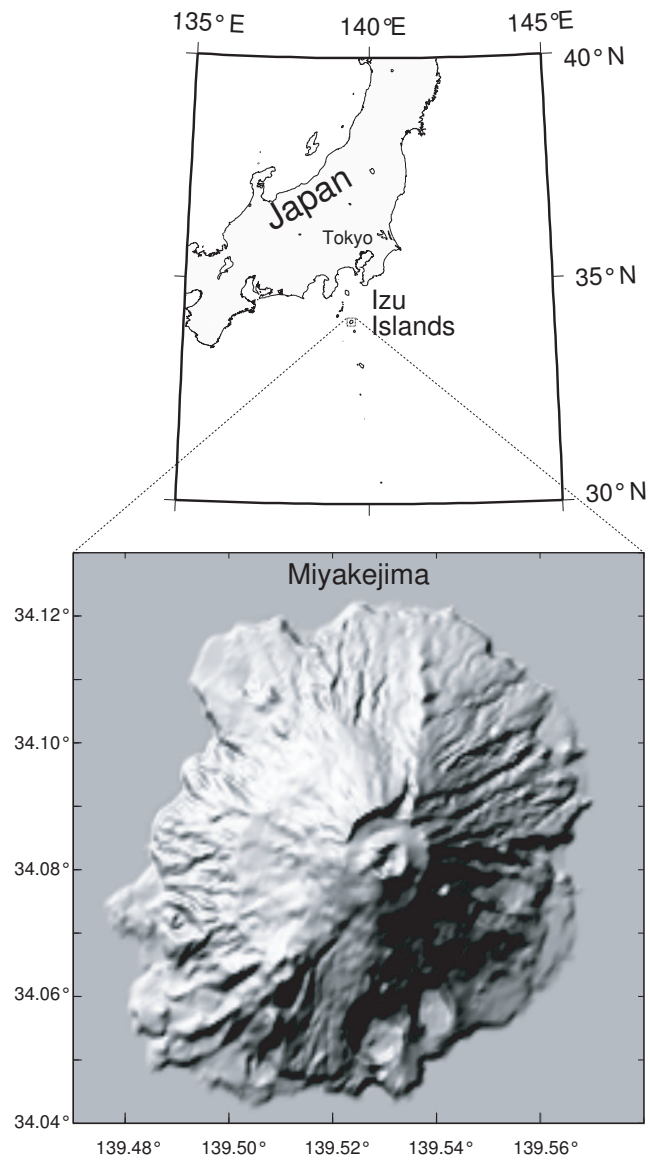
Figure 1.1: Map of the Miyakejima volcano and location in the Izu island group, Japan.

(2005). Table 1.1 is a summary of the eruptive history of Miyakejima merging all sources of information (*Tsukui and Suzuki*, 1998; *Nakada et al.*, 2005; *Simkin and Siebert*, 2002-).

Table 1.1: Summary of the eruptive history of Miyakejima (this table updates the catalog provided by *Tsukui and Suzuki* (1998). The volume of the 2000 eruption is from *Nakada et al.* (2005); some dates missing in *Tsukui and Suzuki* (1998) are from the Global Volcanism Program (*Simkin and Siebert*, 2002-). The VEI values are calculated based on the DRE volumes using the criteria defined by *Newhall and Self* (1982). Eruption numbers accompanied by (*) mark are those considered in the present study (after the completeness analysis).

No.	Kind	Date	Volume$^+$	VEI
29*	2000 Scoria	2000 Jun 27 (AD)	0.009	3
28*	1983 Scoria + Lava	1983 Oct 03 (AD)	0.007	2
27*	1962 Scoria + Lava	1962 Aug 24 (AD)	0.006	2
26*	1940 Scoria + Lava	1940 Jul 12 (AD)	0.015	3
25*	1874 Scoria + Lava	1874 Jul 03 (AD)	0.010	3
24*	1835 Lava	1835 Nov 11 (AD)	<0.001	<2
23*	1811 Scoria	1811 Jan 27 (AD)	<<0.010	2
	1769 (?) Lava		0.001	
22*	1763 (?) Lava	1763 Aug 17 (AD)	0.001	3
	Shinmio Explosion Breccia (SMB)		0.031	
	1763 Scoria		0.033	
21*	1712 Lava	1712 Feb 04 (AD)	0.001	2
20*	1643 Scoria	1643 Mar 31 (AD)	0.009	3
	1643 Lava		0.003	
19*	Kamakata Lava (KKL)	1595 Nov 22 (AD)	<0.001	<2
18*	Benkenezaki Lava (BKL)	1535 Mar (AD)	0.003	2
17*	Enokizawa Lava (EZL)	1469 Dec 24 (AD)	0.002	2
16	Son-ei Bokujo Ash (SBA)	1154 AD (?)	0.040	3
	1154 Scoria		<<0.010	
15	Nanto Lava (NTL)	1085 AD (?)	0.012	2
	Kamane Scoria (KMS)	10-11C AD (?)	0.010	
14	Miike Explosion Breccia (MKB)	838-886 AD	0.040	3
	Oyama Scoria		0.030	
	Oyama Lava		0.012	
13	Kazahaya Scoria (KHS)	832 AD	0.007	2
12	Mitoribata Scoria (MBS)	1290 yBP	0.007	2
11	Daihannya-yama Scoria (DHS)	500 AD	0.010	3
	Anegakata Lava		0.001	

Table 1.1: continues in next page

Table 1.1: (continued)

No.	Kind	Date	Volume[+]	VEI
10	Sabigahama Explosion Breccia (SHB)	320 AD	(?)<0.010	2
	Togahama-south Lava (TSL)		<0.001	
9	Togataira Ash (TGA)	260 AD	0.040	3
	Togataira Scoria (TGS)		0.005	
	Usuki-west Scoria (UWS)		n.d	
	Igayazawa Scoria		0.010	
8	Tairayama Lava (TYL)	2050 yBP	0.001	3
	Tairayama Scoria (TYS)		0.020	
7	Izu Scoria (IZS)	600 BC	0.050	3
6	Hatchodaira Accretionary Lapilli (HCA)	2500-3000 yBP	0.200	4
	Furumio Explosion Breccia (FMB)			
	Hatchodaira Scoria (HCS)		0.170	
	Nagane Scoria (NGS)	1450 BC	n.d.	
5	Tsubota Scoria (TBS)	3000 yBP(?)	0.010	2
4	Mizutamari Explosion Breccia (MZB)	3500 yBP(?)	0.062	3
3	Igaya-east Scoria (IES)	3660 yBP	<<0.010	2
2	Igaya Accretionary Lapilli (IGA)	4000 yBP	0.090	3
	Izushita Lava (ISL)		(?)0.001	
1	Ofunato Explosion Breccia (OFB)	7000-8000 yBP	0.150	4

[+] DRE Volumes, in km^3

1.2.2 Data set of eruptions and completeness of the catalog

In order to extract unbiased information from a catalog it is necessary to check for its completeness. This issue is well known in seismology where the completeness of catalogs is often checked by analyzing the Gutenberg-Richter law and/or the time evolution of the rate of occurrence of events. In volcanology however, the incompleteness of a catalog of eruptions may be more difficult to evaluate because there are just some weak indications that a general power law can hold (*Simkin and Siebert*, 2002-), and also because we know that volcanoes may have different eruptive regimes in their history and then the eruptive rate may change with time (e.g., *Ho*, 1991; *Marzocchi and Zaccarelli*, 2006; *Coles and Sparks*, 2006). Due to this dichotomy, when we talk about the *completeness* of a given eruptive sequence, we understand it as a period of "uniformity" in the data set.

To analyze the completeness of the catalog we plot the cumulative number of eruptions in time, as seen in Figure 1.2a, and identify changes in the statistics of the repose times. Changes in the statistics of the sequence are identified applying a *change point* strategy (CHPT). The change point hunting methods aim to find one or more statistically significant change points in a sequence of data. Here we use a method based on the two-sample Kolmogorov-Smirnov
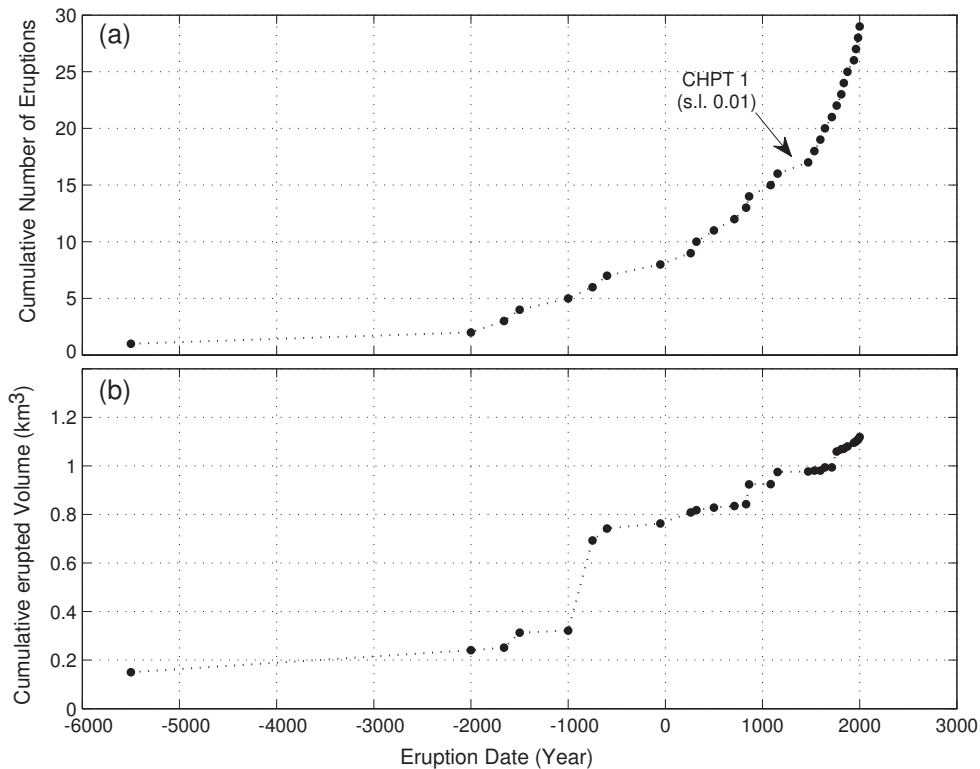
Figure 1.2: (a) Cumulative number of eruptions with time, and Change Point analysis, Miyakejima volcano. (b) The cumulative volume of erupted material is also shown for reference.

statistics (a non-parametric test for equal distributions), which has been proposed and tested by *Mulargia and Tinti* (1985) and *Mulargia et al.* (1987).

The cumulative plot in Fig. 1.2a shows a curve with a slope changing with time. Changes in the slope may be due to different factors as changes in the eruptive regime and/or under-reporting of eruptions in past time (incompleteness of the catalog). Applying the CHPT strategy, we identify one change point (with 0.01 s.l. threshold) between the eruptions 16 and 17 of the catalog (1154 and 1469 AD respectively, see table 1.1). It means that we can consider the catalog to be "uniform" in the period from 1469 up to now (13 eruptive episodes), and then our analysis is oriented to describe the eruptive regime of the volcano within this period. It is important to remark that volcanoes may change eruptive regime through time: our analysis and the derived eruption forecasting assessment is based on the eruptive regime shown by the volcano in the last (around) 540 years, and is valid only under the assumption that in the next future it will behave in the same way.

1.3 Test Models

Most of the models considered are based on generic distributions that characterize renewal processes that in general may have some physical meaning and applicability. A renewal process has the property that the inter-occurrence times are independent and identically-distributed, positive, random variables having a common distribution $F(\tau)$. We have also considered not renewal processes as the size-predictable (SPM) and the time-predictable (TPM) models as possible models to describe the activity of the volcano.

Our analysis consists of different test which can be summarized as: (1) test of a SPM and a TPM; (2) test of a homogeneous Poisson process in the time domain, (3) test of other possible renewal models describing different processes (e.g. recurrent, non-homogeneous Poisson, etc.); in this case we have considered the Lognormal, Gamma, Weibull, and Inverse Gaussian as possible candidate distributions.

1.3.1 Time predictable (TPM) and Size predictable (SPM) models

TPM and SPM are widely used in both seismological and volcanological literature. Both of them imply a functional relationship between size (of eruptions) and repose times. In the case of TPM, the time to the next eruption depends on the time required for magma entering the storage system to reach the eruptive level (*Burt et al.*, 1994). It can be described using a general definition of the form $\tau_i \propto [V_i]^\beta$ (*Sandri et al.*, 2005). A reliable application of a TPM requires that the size (e.g. erupted volume, explosivity index, etc.) of the eruptions has to be significantly correlated to the logarithm of the time to the next eruption (*Marzocchi and Zaccarelli*, 2006). The applicability of this model relies on two main assumptions: (1) eruptions occur when a threshold of the magma volume in the storage system is reached, and (2) the magma input in the storage system is a well defined function of the reservoir to be filled to reach that threshold; for example, the specific case of $\beta = 1$ means that the input rate is constant (*Marzocchi and Zaccarelli*, 2006).

On the other hand, in a SPM the duration of the repose time of the volcano (i.e. the time since the last eruption) is the parameter useful to forecast the size of the next eruption. As for the TPM case, the most general functional relationship between volumes and repose times for a SPM is of the form $V_i \propto \tau_i^\beta$ (e.g., *Marzocchi and Zaccarelli*, 2006). In this case, the model relies on two main assumptions: (1) the output of each eruption is determined only by the magma accumulated since the last eruption, and (2) as in the previous case, the magma enters in the plumbing system at a rate described by a well defined function of the magma volume in the reservoir.

1.3.2 Poisson Process in the time domain: random model of eruption occurrences

The Poisson process is an important model often used to describe the patterns of eruption occurrences in volcanoes (e.g., *Klein*, 1982; *Mulargia et al.*, 1985). It is mainly applicable to major eruptive activity involving a significant release of mass and energy (*De la Cruz-Reyna*, 1991). In a Poisson process the repose times follow an Exponential distribution. It characterizes a *random volcano*, which is one that is ready to erupt at any time. An alternative possibility is that the volcano is in some sense periodic, and that a certain repose time is favored; if eruptions were periodic, the distribution of repose times would be peaked instead of containing an exponentially decreasing number of larger times, as predicted by the Exponential model. In order to explore the degree of departure from a homogeneous Poisson process, we calculate the coefficient of variation, η, given by

$$\eta = \frac{\sigma}{\mu_\tau} \tag{1.1}$$

where μ_τ and σ are, respectively, the average and standard deviation of the repose times τ (*Marzocchi and Zaccarelli*, 2006). The coefficient η may help us to quantify if and how much the statistical distribution of τ differs from a Poisson process: for a Poisson process (and then an Exponential distribution of τ), $\eta = 1$; more clustered distributions have $\eta > 1$ and for more regular recurrent times $\eta < 1$ (e.g., *Cox and Lewis*, 1966; *Marzocchi and Zaccarelli*, 2006).

1.3.3 Renewal models described by Weibull, Gamma and Lognormal distributions

Different probabilistic models are often used in studies in the area of life-testing or reliability theory, in general when the random variable τ represents the lifetime or time to failure of a system. To analyze the intrinsic characteristics of this kind of probabilistic models, it is often more informative to consider the hazard function (also known as hazard rate, or intensity function) of the model than to look at the shape of the PDF or CDF directly; for this reason we make extensive use of the hazard function to compare the different renewal models considered. The hazard function describes the instantaneous failure rate, or the conditional density of failure at a given time, considering the information that no event occurred until that time. A more detailed description of the hazard function concept can be found in section 1.5.

The Exponential distribution (and then its implicit homogeneous Poisson process) has a constant hazard function, highlighting its characteristic no-memory property. However, when processes like wearing, improvement, learning, growth, etc. are implicit in the physical system, then it is necessary to consider models where the hazard function must be a decreasing or increasing function of time. The Weibull, Gamma, Lognormal and Inverse Gaussian (the last one is described in the next section) are models with those characteristics that are widely used in literature.

The Weibull is one of the models most used in volcanological applications (e.g., *Ho*, 1991, 1996; *Bebbington and Lai*, 1996a,b). The Weibull process (WEI(ν,θ)) is one of the possible generalizations of the Exponential case. If the volcanism is waning or developing, the model is generalized to allow the rate of volcanic events (which is constant in the homogeneous case) to be a decreasing or increasing function of time (*Ho*, 1996). This can be defined as a non-homogeneous Poisson process (*Bain*, 1978). The Gamma distribution (GAM(ν,θ)) provides an alternative generalization of the Exponential distribution but with different characteristics respect to the Weibull; in fact, if we consider the hazard function for the Gamma model, the event rate may increase some initially, but after some time the system would reach a stable condition and from then on would be as likely to fail in one time interval as in another (*Bain*, 1978). This is considerably different in the Weibull model where, for $\theta > 1$, the hazard function tends to infinity as the time tends to infinity. Finally, we also consider the possibility of a Lognormal model, which has been considered for periodicity tests by some authors (e.g., *Bebbington and Lai*, 1996a). In this case, the hazard function has a similar behavior as the Gamma model but with the difference that the asymptotic event-rate goes to zero as the time goes to infinity.

1.3.4 Brownian Passage-Time Model

A particularly interesting renewal model is the Brownian passage-time Model. It was originally introduced by *Matthews et al.* (2002) and *Ellsworth et al.* (1999) to provide a physically-motivated renewal model for earthquake recurrence. It is based on the properties of the Brownian relaxation oscillator (BRO). A Brownian passage-time model considers an event (earthquakes in Matthews' model or renewed eruptive activity in our case) as a realization of a point process in which new eruptive activity will occur when a state variable (or a set of them) reaches a threshold (X_f) and at which time the state variable returns to a base ground level (X_0). Adding Brownian perturbations to steady loading of the state variable X produces a stochastic load-state process. An eruption relaxes the load state to the characteristic ground level and begins a new cycle. The load-state process is a BRO, while intervals between events have a distribution known as *Brownian passage-time distribution*. Note that this is the name used in physics literature; in statistics literature it is often known as *Inverse Gaussian* or *Wald* distribution (*Matthews et al.*, 2002).

In the conceptual Model of *Matthews et al.* (2002), the loading of the system has two components: (1) a constant-rate loading component, λt, and (2) a random component, $\epsilon(t) = \sigma W(t)$, that is defined as a Brownian motion (where W is a standard Brownian motion and σ is a nonnegative scale parameter). Standard Brownian motion is simply integrated stationary increments where the distribution of the increments is Gaussian (which might be motivated by central-limit arguments if we consider perturbations as the sum of many small, independent contributions), with zero mean and constant variance. The Brownian perturbation process for the state variable $X(t)$ is defined as (Fig. 1.3):

$$X(t) = \lambda t + \sigma W(t) \tag{1.2}$$

An event will occur when $X(t) \geq X_f$; event times are seen as "first passage" or "hitting" times of Brownian motion with drift (*Matthews et al.*, 2002). The BRO are a family of stochastic renewal processes defined by four parameters: the drift or mean loading (λ), the perturbation rate (σ^2), the ground state (X_0), and the failure state (X_f). On the other hand, the recurrence properties of the BRO (repose times) are described by a Brownian passage-time distribution which is characterized by two parameters: (1) the *mean time* or period between events, (μ), and (2) the *aperiodicity* of the mean time, α, which is equivalent to the familiar coefficient of variation (defined in equation 1.1). The probability density for the Brownian passage-time model is given by:

$$f(t; \mu; \alpha) = \left(\frac{\mu}{2\pi\alpha^2 t^3}\right)^{\frac{1}{2}} \mathbf{e}^{\left\{-\frac{(t-\mu)^2}{2\alpha^2\mu t}\right\}} \tag{1.3}$$

The state variable $X(t)$ is a formal parameter of a point process model and represents a constant-rate mean path that embodies a macroscopic view of a uniform loading of the volcanic system. It may summarize the macro-mechanics of the volcanic system controlled by one or more physical variables. An explicit definition of the driving physical parameters may be unrealistic and impossible to demonstrate from our analysis. Independently of its physical nature, the state variable should be a parameter that accumulates with time during repose episodes, up to a critical value beyond which the system becomes perturbed enough and a new eruptive process may be triggered. Then, the eruptive process *relaxes* the system and the state variable returns to a ground level and a new cycle starts. The perturbation factor $\epsilon(t)$ represents the total sum of all other factors which may play a role in the recurrent eruptive process considered and/or that may randomly disturb the state variable producing the aperiodicity of the mean time between eruptions (e.g. effects from tectonic environment, changes in the magma rate supply, compositional changes, etc.). Fig. 1.3(a to d) exhibits four simulated BRO paths (Eq. 1.2) with $\lambda = 1$, $[X_f - X_0] = 1$, and $\sigma = 0.1, \ 0.3, \ 0.5, \ 1.0$. In particular, let's consider the typical length of a time interval between "failures" (jumps in $X(t)$) and random variations in the lengths of these intervals. The load state $X(t)$ must traverse the distance $[X_f - X_0]$ and does so at average rate λ, so recurrence intervals will have average length $\mu = [X_f - X_0]/\lambda$ (e.g., *Matthews et al.*, 2002). As can be seen in Fig. 1.3, relatively small values of σ produce quite regular paths that closely resemble the deterministic relaxation oscillator and generate nearly μ-periodic recurrence; on the other hand, as σ grows, the periodicity of the system may be *hidden* due to the effect of random perturbations.

1.4   Data Analysis and Results

In order to find which of the trial models described in section 1.3 better explains our data, we estimate the model parameters for each candidate model (using a Maximum Likelihood -MLE- approach) and use the Akaike Information criteria -AIC- (*Akaike*, 1974) to provide a measure of the goodness of fit of each model. The AIC is a tool for model selection based
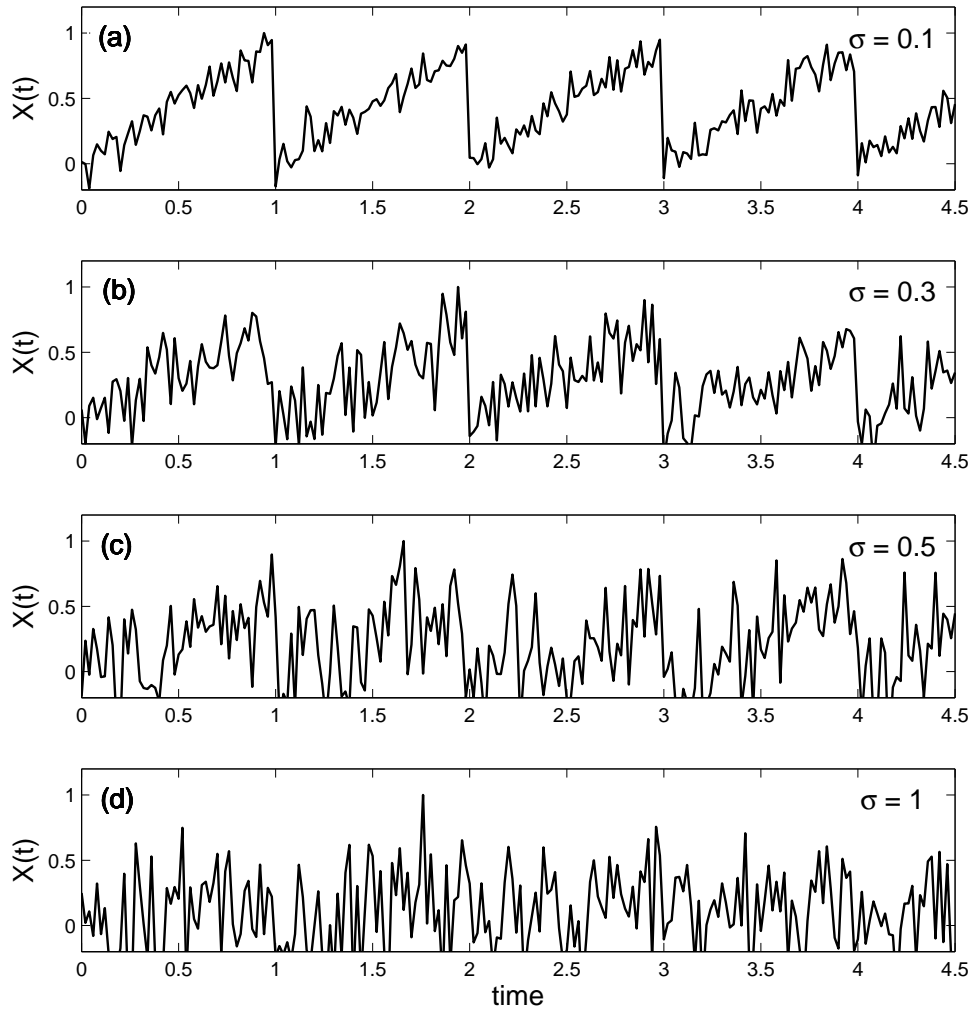
Figure 1.3: Load-state paths for a Brownian relaxation oscillator with $\lambda = 1$, and $[X_f - X_0 = 1]$: (a) $\sigma = 0.1$; (b) $\sigma = 0.3$; (c) $\sigma = 0.5$; (d) $\sigma = 1.0$.
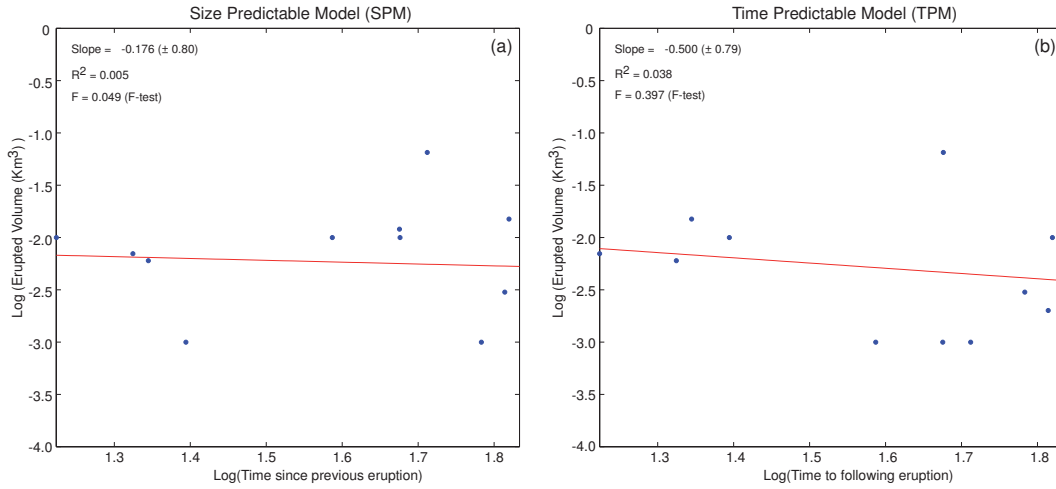
Figure 1.4: Plot of times since previous eruption (a) and times to the following one (b) against the erupted volume (in km^3) of Miyakejima volcano, to test a Size Predictable Model (a) and a Time Predictable Model (b), respectively.

on the concept of entropy, and offers a relative measure of the information lost when a given model is used to describe some data (a trade off between accuracy and complexity of the model). On the other hand, SPM and TPM models are tested by regression analysis of repose times and sizes (erupted volumes and VEI).

1.4.1   Finding the best model fitting the observed data

Fig. 1.4 shows the plots of erupted volumes against the time since the last eruption (Fig. 1.4a) for a SPM, and against the time to next eruption (Fig. 1.4b) for a TPM. In those figures we cannot see any clear relationship between sizes (i.e. volumes) and the times (from last eruption or to the next one). This is confirmed by the R-square statistic (reported in each panel), which in both cases is very low ($R^2 < 0.06$). Furthermore, if we consider the slope of the best fitted line and the associated errors (see Fig. 1.4), in both cases the slope of the fitted line does not significantly differs from zero; using a F-test (Ho : slope = 0), the hypothesis Ho cannot be rejected (at a significance level of 0.05). Then there is no strong evidence suggesting either a SPM or TPM as a successful model for Miyakejima, so at least from these data we cannot extract any information, even if we group sizes using the VEI.

Next we test the hypothesis of a homogeneous Poisson process, in which the repose times τ follow an Exponential distribution. Fig. 1.5 shows the empirical cumulative distribution of observed data and the best (maximum likelihood) Exponential distribution fitting the data. A one-sample Kolmogorov-Smirnov test rejects this hypothesis (at a significance level of 0.05). Furthermore, if we calculate the coefficient of variation η (equation 1.1) we get $\eta = 0.51$, which confirms the non-random distribution and the possibility of a recurrent behavior (i.e.
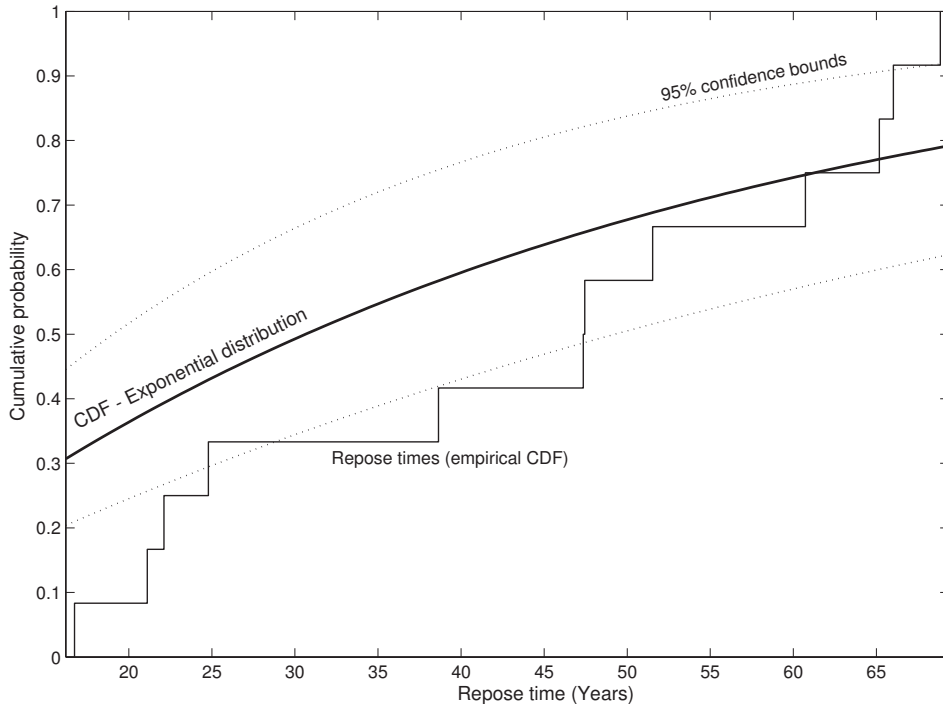
Figure 1.5: Cumulative distribution function (CDF) of the best fitting Exponential distribution, and the empirical CDF of the observed repose times.

$\eta < 1$).

Now we test a set of trial probabilistic models (which were briefly described in section 1.3) to find which provides a better explanation of the repose times observed in the eruptive history of Miyakejima volcano (during the last 540 years). In this part we evaluate models that can describe more regular (recurrent) repose times. The characteristic parameters of each model (Lognormal, Gamma, Weibull and Brownian passage-time -BPT-) were estimated using a maximum likelihood approach. Table 1.2 summarizes the functional form of the PDF, estimated (MLE) parameters (and uncertainties), and the AIC (*Akaike*, 1974) for all the probabilistic models considered. Using a Kolmogorov-Smirnov test, we cannot reject the Lognormal, Gamma, Weibull and BPT hypothesis (at s.l. of 0.05), which means that, from a statistical point of view, all these probability models could explain the observed data. Fig. 1.6 shows the Cumulative Distribution Function (CDF) of the candidate distributions and the empirical CDF of the observed data (τ); as reference, the CDF of the Exponential distribution is also included. As we can see in the plot, many of the candidate distributions may successfully explain the data, however, based on the maximum likelihood value and the AIC, it is possible to select our preferred model (lower AIC) which can be considered as the one that provides the best explanation of the observed data (table 1.2).
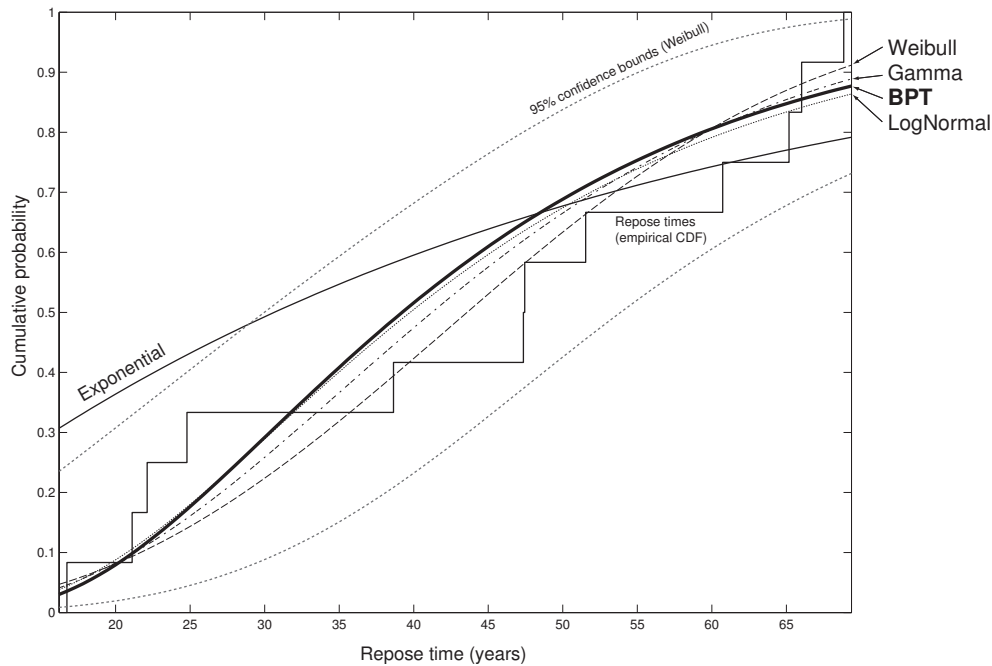
Figure 1.6: Plot of the Cumulative Distribution Function (CDF) of Brownian passage-Time, Lognormal, Gamma and Weibull models, and the empirical CDF of the observed repose times. The CDF of the Exponential model is also included for reference.

Based on parameters listed in table 1.2, our preferred model is the Brownian passage-time. This result is very important since the BPT model may be directly linked to a physical system which may provide significant insights for the interpretation of the observed eruptive behavior. For instance, the mean repose time μ of the BPT (44.2 ± 6.5 years), or its reciprocal, the mean rate of occurrence, is the natural scale parameter of first-order interest, as it measures the typical frequency at which eruptions occur. Changing the mean re-scales time but does not otherwise alter the shape of the probability distribution. The aperiodicity ($\alpha = 0.51 \pm 0.01$) is chosen as a second parameter because (1) it is a natural shape parameter of the BPT family, and (2) it is a dimensionless measure of irregularity in the event sequence (*Matthews et al.*, 2002); in other words, it is a measure of the aperiodicity of the mean. As α tends to 0, the sequence tends to be perfectly periodic, while as α grows, the sequence tends to a "Poisson-like" process. Somehow, the BPT model may be regarded as a delayed Poisson process (*Ellsworth et al.*, 1999), for which the failure rate is zero for a finite time following an event and then steps up to an approximately constant failure rate at all succeeding times; then, as α grows, this kind of *delay* time tends to be smaller (for a given μ).

Due to its intrinsic characteristics the BPT model can be considered as a powerful tool to characterize volcanic systems since, -given that there exists a homogeneous (uniform) catalog of past eruptive episodes-, it could be used to characterize a wide range of processes that

Table 1.2: Candidate distributions, PDF, estimated (MLE) model parameters and uncertainties, and Akaike Information Criteria (AIC)

Model	Probability Density	Parameters (MLE)	AIC
Brownian Passage-time (μ, α)	$\left(\frac{\mu}{2\pi\alpha^2 t^3}\right)^{\frac{1}{2}} e^{\left\{-\frac{(t-\mu)^2}{2\alpha^2\mu t}\right\}}$	$\mu = 44.2 \ (\pm 6.5)$ $\alpha = 0.51 \ (\pm 0.01)$	**86.8842**
Weibull (ν, θ)	$\theta\nu^\theta t^{\theta-1} e^{\left\{-\nu^\theta t^\theta\right\}}$	$\nu^{-1} = 49.9 \ (\pm 5.6)$ $\theta = 2.7 \ (\pm 0.7)$	107.3074
Gamma (ν, θ)	$\frac{\nu^\theta t^{\theta-1}}{\Gamma(\theta)} e^{-\nu t}$	$\nu^{-1} = 9.1 \ (\pm 3.8)$ $\theta = 4.9 \ (\pm 1.9)$	108.2612
Lognormal (μ, σ)	$\frac{1}{\sigma t \sqrt{2\pi}} e^{\left\{-\frac{(ln(t)-\mu)^2}{2\sigma^2}\right\}}$	$\mu = 3.68 \ (\pm 0.15)$ $\sigma = 0.51 \ (\pm 0.11)$	109.1462
Exponential (ν)	$\nu e^{-\nu t}$	$\nu^{-1} = 44.2 \ (\pm 12.8)$	116.9338

span from "Poisson-like" (i.e. completely random) up to "perfectly periodic" processes. On the other hand, the direct link with a conceptual physical system may provide insights into the physical processes associated with the eruptive behavior of the volcano. Furthermore, if we analyze the hazard function of the BPT model (see section 1.5), some interesting features may be discussed for a time-dependent, long-term eruption forecasting.

1.4.2   Distribution analysis of erupted volumes

Based on the results described in previous sections we reject the hypothesis of a Poissonian process as the model explaining the eruptive behavior of Miyakejima volcano. The probabilistic model providing the best explanation of the data is a Brownian passage-time, which allows us to infer some kind of periodicity in the eruptive process. Furthermore, from the eruption size data (volumes and VEI), it was not possible to find any evidence of a TPM or SPM. The question that arises is then, within this framework, how should the erupted volumes be distributed? We perform an analysis of the erupted volumes (within the period of completeness of the catalog) and the results are summarized in Fig. 1.7. A Lognormal distribution provides a good explanation of the erupted volumes data (hypothesis not rejected using a one sample Kolmogorov-Smirnov test at a s.l. of 0.05); it means that there exists a *preferred* or more common erupted volume (the mean erupted volume is $0.012 \pm 0.004 \ km^3$). In other words, we can consider that the logarithm of the erupted volumes are normally distributed. This result can support the hypothesis of a recurrent source model, as suggested by the Brownian passage-time distribution for the repose times, and may also explain the poor resolution of the TPM, since if there is a *preferred* size and a *preferred* repose time, then the data in a time-size space should tend to group in a cluster. From this point of view our BPT model would not be incompatible with a TPM model, as is discussed in section 1.6.

1.5   Implications of a Brownian model for Eruption Forecasting Assessment

Repose times for recurrent eruptive activity that follows a Brownian passage-time distribution may be used to define a model for time-dependent, long-term eruption forecasting. This distribution has some noteworthy properties as (1) the probability of having renewed eruptive activity at time $t = 0$ is 0 (i.e. just after the last eruptive period); (2) as $t \to \infty$ the hazard function is finite. In other words, it increases steadily from zero at $t = 0$ to a finite maximum near the mean recurrence time. The first property should be analyzed carefully since it may lead to misunderstanding if used improperly. As described in the introductory part, in our analysis we ignore eruption duration since we take the onset date as the most physically meaningful; then we measure repose times from one onset date to the next. Following this approach, our modeling intends to describe the waiting times of the long-term physical processes governing the onset of new volcanic activity defined as episodes of renewed magmatic intrusions entering in the volcanic system and triggering new eruptive cycles. When
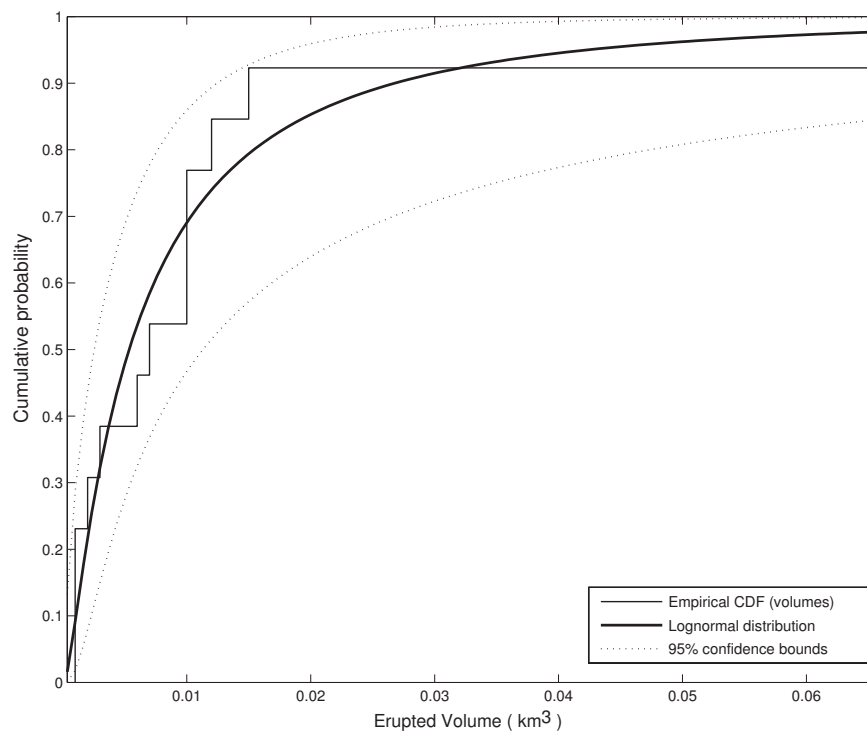
Figure 1.7: CDF (empirical and Lognormal distribution) of the erupted volumes in Miyake-jima volcano from 1469.

an eruption starts, the short-term behavior of eruptive activity might follow different patterns during the gradual decline of activity; then, sporadic eruptive episodes in a short time window after the onset of a new eruptive cycle cannot be described using the former long-term model.

Let's consider the random series of events $t_1 < t_2 < \ldots < t_i \ldots$, and then the repose times $\tau_i = t_{i+1} - t_i$, $(i = 1, 2, \ldots)$. If the sequence of random variables $\{\tau\}$ is independent and distributed according to a function $F(\tau)$, then the original series of events $\{t_i\}$ is called a renewal process. For a history-dependent point process, the conditional intensity function $\lambda(t|H_t)$ of the form

$$\lambda(t|H_t) = h(x) = \frac{f(x)}{1 - F(x)} = \frac{f(x)}{S(x)} \tag{1.4}$$

for $x = (t - t_L)$, defines the hazard function. Here, H_t is a history of occurrence times $\{t_i; \; t_i < t\}$ before time t, including the information that no event occurred neither in the intervals $(t, \; t_{i+1})$ nor in the interval $(t_L, \; t)$, and where t_L is the last occurrence before the considered time t (e.g., *Bain*, 1978; *Ogata*, 1999). Then $h(x)$ is the ratio of the probability density function $f(x)$ to the survival function $S(x)$ and it may be defined as the event rate at time t conditional on survival until time t (or later).

The hazard function describes instantaneous failure rate, or the conditional density of failure at time x given that no event occurred until time x. An increasing hazard function at time x indicates that an event is more likely to occur in a given increment of time $(x, x + \Delta x)$ than it would be in the same increment of time in an earlier age. It is also useful in the specification of a point process since it may be directly linked with the probabilistic forecast of an event occurrence.

Fig. 1.8 shows the hazard function of the BPT model for Miyakejima volcano (see also Fig. 1.6 for the corresponding cumulative distribution functions). Hazard functions of the other candidate models are also included for comparison. For the BPT model, the failure rate is zero (0) immediately after an event, then it grows to a peak and then asymptotically tends to a finite value at long times compared to the mean time. Fig. 1.8 can help to understand the different behavior of the different candidate models and to compare them with the BPT model. The main characteristic of the Exponential model is the constant hazard function, implying a random occurrence of volcanic eruptions in time. All the other models are more or less similar up to the mean recurrence time, at which point their behavior diverges. For example, the hazard function of the Weibull model starts at zero and increases to infinity, while for the Lognormal model, the asymptotic failure rate goes to zero. The Gamma model also has a finite asymptotic failure rate, but the function grows more smoothly.

We can calculate the conditional probability $Pr_{(x, \; x+\Delta t)}$ that an eruption will happen in a time interval $(x, x + \Delta t]$, given an interval of $x = (t - t_L)$ years since the occurrence of the previous eruption. Let **T** be the time to the next eruption, then $Pr_{(x, \; x+\Delta t)} = P(x < \mathbf{T} \leq$
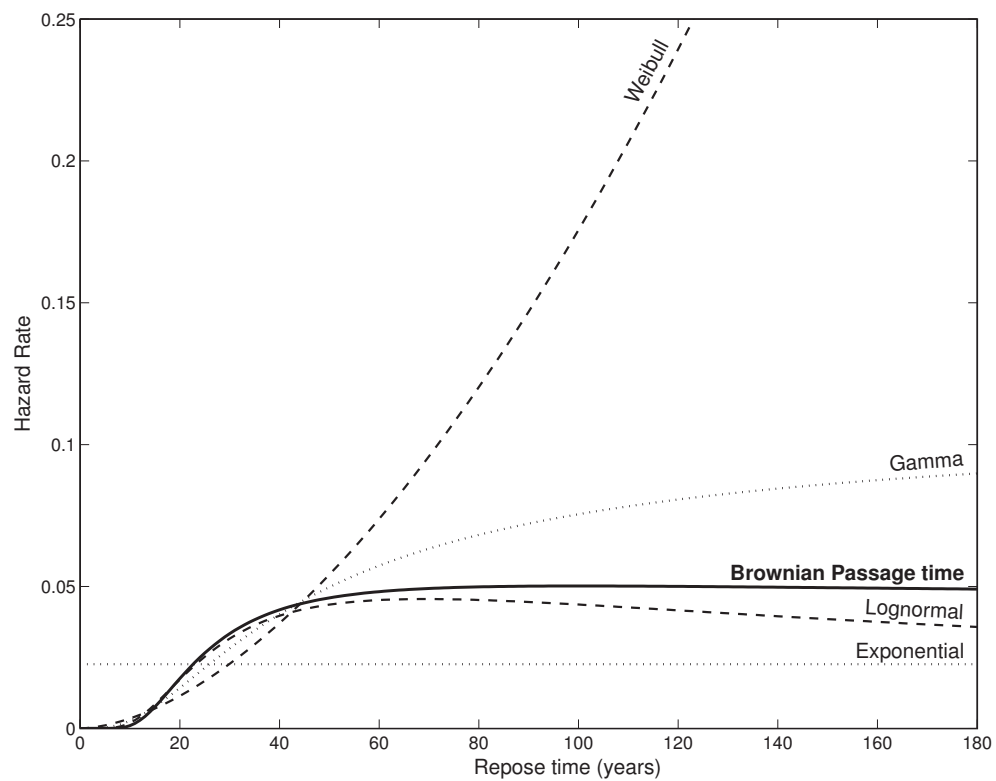
Figure 1.8: Hazard function of the Brownian passage-time (BPT) model and all the other candidate distributions (Exponential, Gamma, Lognormal and Weibull).
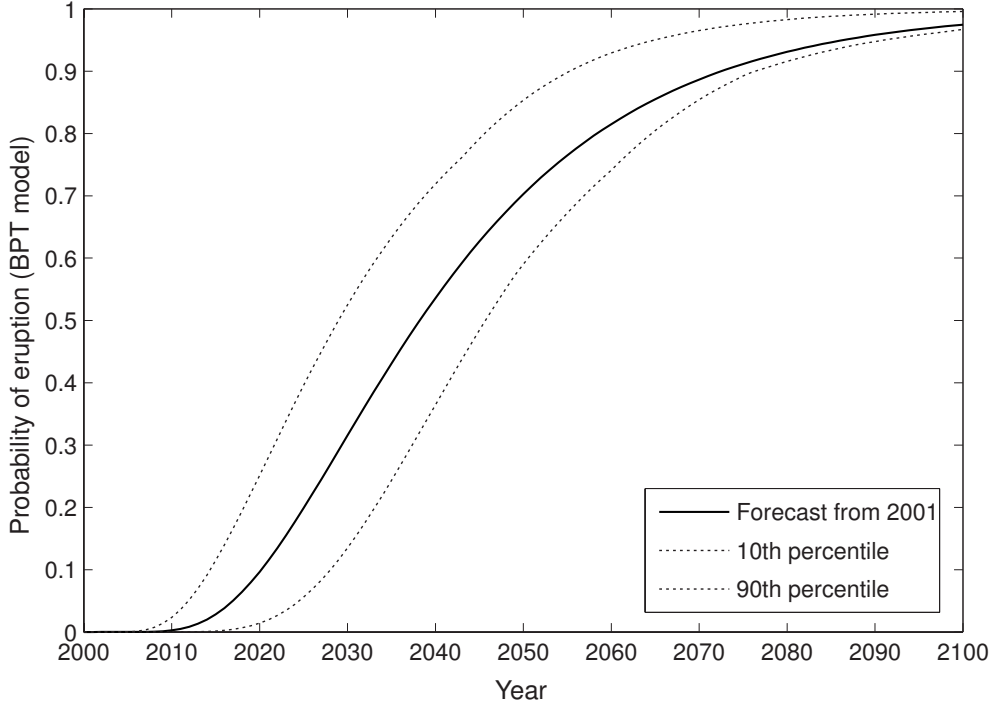
Figure 1.9: Eruption Forecasting for Miyakejima volcano, using a Brownian passage-Time model. Probability of eruption from 2001 (just after the last eruption) evaluated for Δt in the interval [1, 100]

$(x + \Delta t) \mid \mathbf{T} > x)$ (for x being the time since last eruption, as defined before).

If $F(\tau)$ denotes the (cumulative) distribution function of the repose times τ, then $F(x) = Pr(\mathbf{T} \leq x)$, and $F(x + \Delta t) = Pr(\mathbf{T} \leq (x + \Delta t))$ for $x \geq 0$, while the survival time function $S(x)$ is $S(x) = 1 - F(x) = Pr(\mathbf{T} > x)$, for $x \geq 0$. Then, the probability that an eruption occurs in the next Δt interval is (e.g., *Bowers et al.*, 1997)

$$
\begin{aligned}
Pr_{(x,\ x+\Delta t)} &= \frac{\int_x^{x+\Delta t} f(s)ds}{1-F(x)} \\
&\approx \frac{F(x+\Delta t)-F(x)}{1-F(x)}
\end{aligned}
\tag{1.5}
$$

The approximation in equation 1.5 is valid for small Δt; it can be interpreted as the conditional probability that an eruption will occur in the time interval $(x, x+\Delta t]$, given an interval of x years since the occurrence of the last event. We can use this equation to calculate probabilities of eruption and forecast future events. For example, Fig. 1.9 is the evolution of $Pr_{(x,\ x+\Delta t)}$ as seen from the time immediately after the last eruption in 2000 for different values of Δt.

1.6 Discussion

Renewal processes characterized by five different probabilistic models, plus a TPM and a SPM, have been applied to analyze the repose times between eruptive episodes of Miyakejima volcano during the last 540 years (when the catalog has been considered complete). From our analysis we conclude that the probabilistic model that better explains the observed data is a Brownian passage-time. This model is based upon a simple physical model resembling the characteristic cycles of volcanic activity, the Brownian relaxation oscillator, and is parameterized by the mean rate of event occurrence, μ, and the aperiodicity about the mean, α.

The Brownian passage-time family differs from other usual candidate distributions for long-term eruption forecasting in that it may reflect physical properties of the macro-mechanics of volcanic processes. The Brownian relaxation oscillator and Brownian passage-time distribution connect together physical notions of unobservable loading and failure processes of a point process with observable response statistics (i.e. event recurrence in time).

Up to now, the definition of a general model to describe eruptive activity has been a difficult task due to different factors as the intrinsic complexity of eruptive processes and the difficulty of getting complete catalogs with sufficient observations. Even if we do not pretend to define it as a general model, the BPT may be considered as a first-order approximation to describe different kinds of volcanic systems, which can span from random volcanoes (Poisson-like processes), up to perfectly periodic systems. The non-homogeneous Poisson process model of *Ho* (1991) characterized by a Weibull distribution has been a first attempt of generalization to describe with a single model different kind of processes. However, the Weibull is a rather empirical model that possesses some intrinsic undesirable features difficult to support from a physical point of view in volcanological applications. For example, hazard rate functions of Weibull variates (e.g. see Fig. 1.8) either start at zero and increase to infinity or start at a finite value and decrease to zero. This asymptotic behavior may be unrealistic in many physical systems and specifically in a volcanological application may lead to unnecessary assumptions.

Conversely, the BPT model possesses many interesting features which make it a plausible model to describe the activity of different volcanoes. If we consider its hazard function, the failure rate is zero immediately after an event. Then it grows to a peak and then declines to a finite asymptotic rate at times long compared to the mean rate. These are unique properties among the set of candidate models considered. These properties provide a more realistic asymptotic behavior of the failure rate. The BPT model may be regarded as a delayed Poisson process (*Ellsworth et al.*, 1999), for which the failure rate is zero for a finite time following an event and then steps up to an approximately constant failure rate at all succeeding times.

To measure how much the BPT model approaches whether a Poisson-like or a periodic process, we can consider the α parameter. As we did mention before, the more periodic the process, the more α approaches zero. The value $\alpha = 0.51 \pm 0.01$ found in this work for the aperiodicity in Miyakejima volcano indicates probably for the first time a clear recurrent

behavior in a volcanic system. To compare eruptive activity of different volcanoes with the results obtained in Miyakejima, we did analyze some catalogs from published works in other volcanic areas: for instance, we considered the data from (1) Mt Ruapehu and (2) Mt. Ngauruhoe -New Zealand- (tables 2 and 3 in *Bebbington and Lai* (1996b)), (3) Kilauea and (4) Mauna Loa -Hawaii- (tables 1 and 2, respectively, in *Klein* (1982)), and (5) Mt. Etna (*Marzocchi and Zaccarelli*, 2006).

Fig. 1.10 is a plot of the estimated parameters α and μ assuming a BPT model for the volcanoes cited before. The μ (y axis) is just an scale parameter measuring the mean recurrence time. However, the (dimensionless) α parameter (x axis) may provide a good framework to compare different volcanic systems; for instance, if we consider the results in Fig. 1.10 it is evident that all considered volcanoes but Miyakejima have $\alpha > 1$. This is very consistent with the different results provided by the authors; for example, if we consider Mauna Loa ($\alpha = 1.28 \pm 0.4$) and Kilauea ($\alpha = 3.02 \pm 1.49$) volcanoes, α parameter indicates that those volcanoes have more Poisson-like behavior, which is in agreement with the results of *Klein* (1982) who concluded that Hawaiian eruptions are largely random phenomena displaying no periodicity. For Ruapehu ($\alpha = 1.26 \pm 0.36$) and Ngauruhoe ($\alpha = 1.4 \pm 0.33$) volcanoes, α parameter indicates also a Poisson-like behavior, which is also in agreement with the results of *Bebbington and Lai* (1996b); in those cases, the authors examined both the homogeneous Poisson and Weibull as possible models to describe the eruptive patterns of both volcanoes, concluding that both of them are more Poisson-like processes even if they are satisfactorily modeled by a Weibull renewal process.

Another important consideration should be done with respect to the TPM/SPM models. As discussed in section 1.4.1, it is not possible to define a clear relationship between repose times and eruption sizes from Miyakejima volcano; additionally, in section 1.4.2 we found that there is a *preferred* erupted volume. Given the recurrent behavior of Miyakejima volcano (inferred from the α value of the BPT model), we argue that it is coherent that for *preferred* repose times it is possible to have *preferred* erupted volumes. It means that it is possible that a TPM or SPM model could coexist within our BPT model for recurrent volcanic activity. In fact, The BRO may be extended to models that are no renewal processes; in particular stochastic versions of TPM and SPM may be derived from randomization boundary conditions in the Brownian oscillator (*Matthews et al.*, 2002).

1.7   Concluding remarks

The intrinsic complexity of volcanic systems motivates the definition of probability models as mathematical structures to describe the response of the considered systems. The analysis of the past eruptive activity of Miyakejima volcano suggests that the probability model that provides the best description of the observed repose times is a Brownian passage-time model. The BPT family of distributions describes the response of a conceptual physical system defined
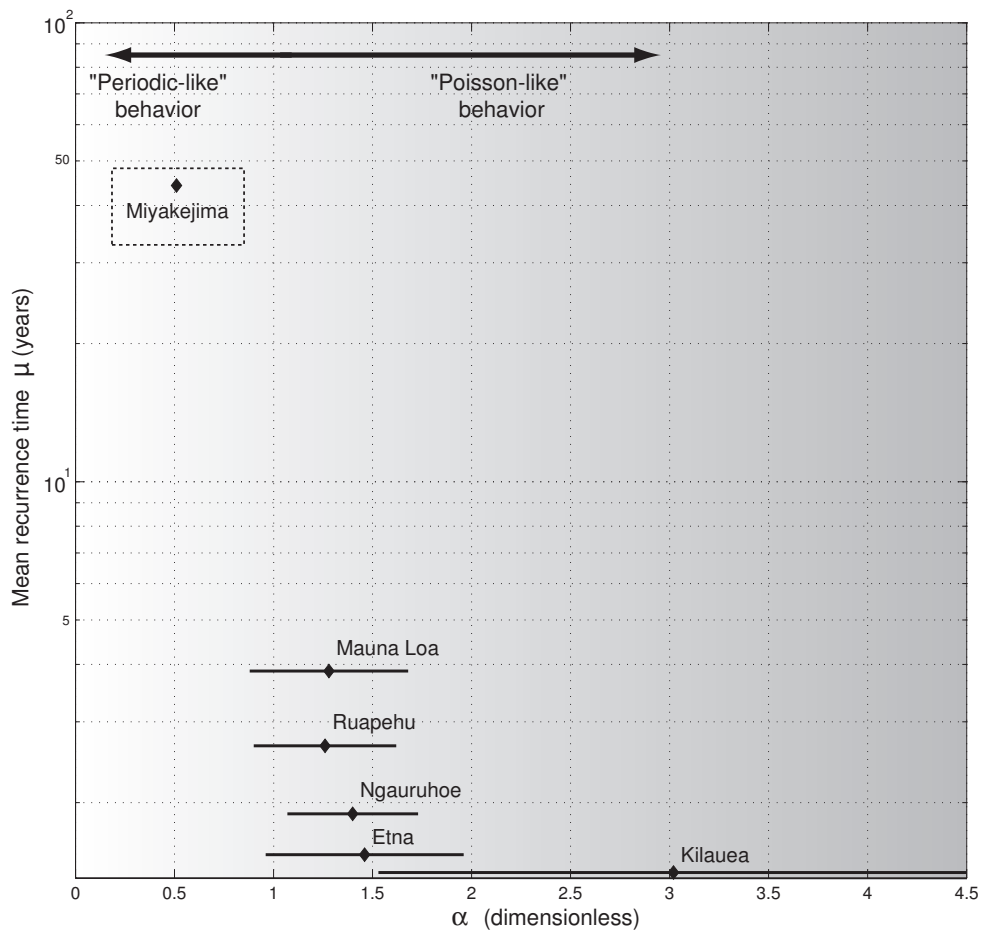
Figure 1.10: Estimated parameters (α and μ) of the Brownian passage-time model for different volcanoes: Miyakejima (Japan), Etna (Italy), Ruapehu and Ngauruhoe (New Zealand), Mauna Loa and Kilauea (Hawaii). For the source of the data see the text.

as a Brownian relaxation oscillator (BRO). BRO and BPT together connect physical notions of unobservable loading and failure processes of a point process with observable event-time statistics. BPT model is characterized by two parameters: the mean repose time (μ) and the aperiodicity of the mean (α). While μ is just an scale parameter that provides information about the typical frequency at which events occur, α is a dimensionless parameter that measures the aperiodicity of the mean response of the system, and for this reason this parameter may be useful to compare different volcanoes spanning from periodic-like to Poisson-like systems.

For the Miyakejima volcano, the mean repose time is $\mu = 44.2 \pm 6.5$ years, while the dimensionless aperiodicity parameter is $\alpha = 0.51 \pm 0.01$. This value for α parameter is an evidence of recurrent eruptive activity of Miyakejima volcano; furthermore, this is probably the first documented case of periodic behavior in a volcanic system.

BPT model provides some insights for time-dependent, long-term eruption forecasting. For instance, if we consider the hazard function, some noteworthy properties can be defined: the probability of having renewed eruptive activity just after an eruptive cycle is very low, then it increases steadily from zero to a finite maximum near the mean recurrence time. Finally, for times greater than the mean recurrence time the hazard function tends to a finite constant value, indicating that for long repose times the system tends to behave as a Poisson process. In Chapter4 (section 4.2), we propose a methodology to consider any time-dependent or time-independent, long-term, eruption forecasting model based on the eruptive history of the volcano (as for example the Brownian passage-time model described here) into the Bayesian Event Tree model for eruption forecasting (BET_EF) of *Marzocchi et al.* (2008).

Application of probabilistic inverse theory to physically-based spectral analysis of Long-Period seismicity, and its possible use for volcano monitoring

"Whenever a theory appears to you as the only possible one, take this as a sign that you have neither understood the theory nor the problem which it was intended to solve."

Karl Popper
(Objective Knowledge: An Evolutionary
Approach, 1972)

Abstract

We present an alternative method based on probabilistic inverse theory to estimate the parameters of a simple conceptual model (presented by *Kumazawa et al.* (1990) about twenty years ago) as a method for physically-based spectral analysis. The probabilistic formulation leads to the definition of a (posterior) probability distribution in the model space $\sigma(\mathbf{m})$ which results of combining prior information (on data and model parameters) with new information obtained by measurements (the time series). This kind of formulation may be helpful for different kind of geophysical problems, specially when it exists a nonlinear relationship between the observed data and the parameters of a given model. Using this formulation we have performed some numerical tests using synthetic time series, and we did apply the method to analyze the waveforms of Long-Period (LP) events from two volcanoes: Cotopaxi (Ecuador)

and Miyakejima (Japan). The procedure described may be an important tool for objectively monitor the characteristic complex frequencies of volcanic LP events; in this way, changes in the volcanic activity may be highlighted and families of events of particular interest could be identified for further analysis. Beyond this possible application for monitoring, our approach can also be used to analyze individual events as an alternative to the existent methodologies.

2.1   Introduction

Active volcanoes are the source of a great variety of seismic signals, this is because volcanic seismic sources involve different kinds of physical processes. Moreover, nearly every recorded volcanic eruption has been preceded and/or accompanied by seismic activity beneath or near the volcano (e.g., *Aki*, 1992; *McNutt*, 1996), then seismology may be considered as one of the most important tools for monitoring volcanoes since it can provide information (i.e. strong seismic parameters) that considered together with other monitoring parameters, can help volcanologists to improve their understanding of the physical processes inside volcanoes and to produce better probabilistic eruption forecasting models (e.g., *Marzocchi et al.*, 2008).

The source mechanisms associated with Long-Period (LP, also known as Low-Frequency LF) seismic events and volcanic tremor are intimately associated with fluids contained in a solid rock matrix. Both the frequency content and the characteristic long-lasting coda of these seismic signals are of fundamental importance since they represent evidences of the characteristic properties of the source, as have been highlighted by many authors using diverse models (e.g., *Aki et al.*, 1977; *Chouet,* 1981, 1982, 1985, 1988; *Ferrazzini and Aki*, 1987; *Fujita et al.*, 1995; *Fujita and Ida*, 2003; *Neuberg et al.*, 2000; *Morrissey and Chouet*, 2001).

The observed waveform includes both source mechanism and path effects, however, we normally deal with data of observation stations located close to the source (few wavelengths) and then we assume that most of the waveform characteristics are mainly associated with source processes. In fact, in the recorded signals we can distinguish some properties that support such assumption. The waveforms normally show an emergent onset, which means a gradually increasing amplitude, until reach a maximum. Amplitude then diminishes with a decay rate that change from event to event, generating coda waves which can last from few seconds up to even some minutes. Their spectra are sharply peaked, showing generally a dominant peak and, sometimes, several subdominant peaks (harmonics). Spectral analysis of the signals also shows that there is little variation of the dominant frequency with station location, and during a period of steady activity (where normally many events take place in relatively short time intervals), the wide variation in amplitude of the events does not corresponds with any evident variation in the dominant frequency, suggesting that the amplitude of the signal is more related to the force generating the signal rather than the size of the source region. The relative consistency of the spectrum of this kind of events observed at various locations may indicate that many of the waveform properties observed in the records are due to a source effect rather than path and site effects (*Aki and Koyanagi*, 1981). Obser-

vations made at different volcanoes have shown that this type of events can not be explained by a simple instantaneous faulting but are rather interpreted as oscillations of a resonator in relation to both magmatic and hydrothermal activity (e.g., *Aki et al.*, 1977; *Aki and Koyanagi*, 1981; *Ferrazzini and Aki*, 1987; *Chouet*, 1988; *Neuberg et al.*, 2000; *Fujita and Ida*, 2003).

The description of the oscillation frequencies of LP signals is fundamental for the inference of the characteristic properties of the resonator system. The fast Fourier transform (*Cooley and Tukey*, 1965) is one of the most commonly used techniques for spectral analysis. In order to provide a physically-based method for spectral analysis, *Kumazawa et al.* (1990) proposed an alternative method with higher spectral resolution which is based on the characteristic properties of a linear dynamic system. This method provides both the decay characteristics and the oscillation period of individual signal components. *Yokoyama et al.* (1997) extended the Sompi method to the inhomogeneous equation of motion including an external force term, and *Nakano et al.* (1998) presented a method to quantify the source excitation function and characteristic frequencies of LP events based on the former inhomogeneous model.

Our objective in this chapter is to present an alternative method for the estimation of the characteristic parameters of the homogeneous model of *Kumazawa et al.* (1990). Our main interest is pointed toward the analysis of the tails of exponentially-decaying harmonic LP waveforms. Our approach is based on a probabilistic formulation of inverse problems in which prior models (randomly generated) are combined with the information provided by the physical theory (forward model) and the measurements (data) in order to define a probability density representing the *posterior* information. This posterior probability density describes all the information we have about the problem (model and data). It may be multimodal, or not have a mathematical expectation, have infinite variances or some other pathologies, but it constitutes the complete solution to the problem. To obtain samples of the posterior probability distribution (and then information about the model parameters) we use Markov-chain Monte Carlo (MCMC) techniques. This approach, even if it may require more computer time to obtain results, may provide some important benefits respect to the classical maximum likelihood approach, specially due to the nonlinear relationship between the observed data and the model parameters of the problem.

2.2 Physically-based spectral analysis and representation of LP events

It is widely accepted that the source of LP events consists of excitation and the subsequent resonance of a volcanic fluid system in the source region. To produce monotonic and harmonic waveforms, it is reasonable to take as possible resonators fluid-filled structures like dykes, cracks, sills, etc. The excitation itself is still not well understood and it may be due to different mechanisms as pressure or stress perturbations caused by migration of fluids, vesiculation and collapse of bubbles, small ruptures close to the fluid-filled structure, etc. If the excita-

tion is a time-localized function present only during a short time interval at the beginning of the event (e.g., *Nakano et al.*, 1998) then, in many cases the resonance of the fluid-filled structure should be the source process producing the dominant spectral peaks observed in the tail of the signals. To describe the main characteristics of the tails of LP events (i.e. frequencies and decay rate), we can approximate it as a function generated by a superposition of simple decaying sinusoids (e.g., *Nakano et al.*, 1998; *Kumagai and Chouet*, 1999, 2000, 2001).

Spectral analysis, as the discrete Fourier transform, is performed decomposing a given function into a linear combination of a complete set of orthogonal basis functions, where uniqueness of the decomposition is guaranteed. Harmonic oscillations are widely observed in physical phenomena, then it is reasonable that the decomposition of a time series into Fourier components has a great popularity among spectral analysis methods. Due to the inevitable presence of random noise in the observed signal, statistical modeling has been also successfully introduced in many spectral estimation theories (e.g., *Kay and Marple*, 1981). However, as in many other physical systems, the LP waveforms show a clear trend of amplitude to decay with time, then the decomposition of the time series into purely harmonic components may be translated into a loss of information. Since the ultimate purpose of spectral analysis of time series is in general to obtain information on the physical system which generated the signal, *Kumazawa et al.* (1990) introduced an spectral analysis method that is based not only on mathematics or information theory but also on physical concepts. The method (called 'Sompi') is based on the Autoregressive (AR) model, and consist of extracting a set of deterministic and coherent signals from a data set of finite length, which are regarded as the observed samples of the realization of a hypothetical linear dynamic system (*Kumazawa et al.*, 1990).

In the approach proposed by *Kumazawa et al.* (1990), a given time series $\{x(t)\}$ is decomposed into a linear combination of finite-number coherent sinusoids with amplitudes exponentially decaying (or growing) with time such as

$$x(t) = \sum_{i=1}^{n_w} A_i e^{\gamma_i t} cos(\omega_i t + \theta_i) + \epsilon(t) \tag{2.1}$$

where n_w is the (finite) number of sinusoids (wave elements), and ω_i, γ_i, A_i and θ_i are real constants, and $\epsilon(t)$ is incoherent noise. The method uses individual spectral lines in the complex frequency space to provide both the decay properties (characterized by γ) and the oscillation frequency (characterized by ω) of individual signal components.

In this kind of spectral analysis, the signal in the given time series is represented by the superposition of "wave elements" (*Kumazawa et al.*, 1990). Each wave element is specified by the four real parameters outlined before, where ω and γ correspond to the real and imaginary parts of the complex angular frequency, respectively (positive and negative γ indicate that the instantaneous amplitude of the relevant wave element grows or decays exponentially with time), and A and the phase angle θ represent the complex amplitude.

The ordinary (real) frequency f is defined by

$$f = \frac{\omega}{2\pi} \tag{2.2}$$

and the "growth rate" (or "gradient") parameter **g** is defined by (*Kumazawa et al.*, 1990)

$$g = \frac{\gamma}{2\pi} \tag{2.3}$$

then, the dissipation factor Q^{-1} (e.g., *O'connell and Budiansky*, 1978)) is given by

$$Q^{-1} \simeq \frac{2\gamma}{\omega} = \frac{2g}{f} \tag{2.4}$$

for small values of Q^{-1}.

In the last years this approach has been used to analyze different kinds of data; some examples are the applications to the Earth's free oscillations (*Hori et al.*, 1989) and to nuclear magnetic resonance spectroscopy (*Matsuura et al.*, 1990). Specifically, it starts to be widely used by many authors in volcano seismology to describe spectral characteristics of LP seismic signals from different volcanoes (e.g., *Nakano et al.*, 1998; *Kumagai et al.*, 2002; *Molina et al.*, 2004; *Kumagai et al.*, 2005; *De Angelis*, 2006) specially after the works of *Kumagai and Chouet* (1999, 2000, 2001) where the authors performed the link between a volcanic physical system (a fluid-filled crack) and the characteristic complex frequencies of the simulated signals (by forward modeling), opening the way to use the characteristic complex frequencies of LP events (determined with the Sompi method of *Kumazawa et al.* (1990)) to infer the acoustic properties of a crack containing magmatic or hydrothermal fluids at the source.

In the original work of *Kumazawa et al.* (1990), the parameters of the model were calculated based on a Maximum likelihood approach. In that paper there is a complete description of the analog physical system at the base of the model. In the next sections we will describe an alternative setting of the problem to determine the parameters of the model in a framework based on probabilistic inverse theory; we then perform inference on model parameters using Monte Carlo techniques. As pointed out by *Kumazawa et al.* (1990), we also highlight here that this kind of spectral analysis is different of the spectral decomposition into orthogonal basis functions. It should be understood as a *physical spectral analysis* where time series are modeled based on the characteristic properties of a linear dynamic system with no external forces acting on it.

2.3 Probabilistic Inverse Theory: Overview and Setting of the Problem

2.3.1 Probabilistic Inverse Theory

Given a physical system, the *forward* or *direct* problem consists, by definition, in using a physical theory to predict the outcome of possible experiments, which in classical physics have a unique solution. The *inverse problem* arises when we do not have a perfect model but we have a valid set of data. Following the approach proposed by *Tarantola* (2005), the inverse problem is seen as a problem of *conjunction of states of information* (theoretical, experimental and prior information, generally on model parameters) taking a probabilistic point of view. The axioms of probability theory apply to different situations. One is the traditional statistical analysis of random phenomena, another one is the description of (more or less) subjective states of information on a system. For instance, estimation of the uncertainties attached to any measurement usually involves both uses of probability theory: some uncertainties contributing to the total uncertainty are estimated using statistics, while some other uncertainties are estimated using informed scientific judgement (e.g. the quality of an instrument, effects not explicitly taken into account, etc.).

The most general way to describe an *state of information* is to define a probability density over the parameter space; it follows that the results of the measurements (data), the a priori information on model parameters, and the information on the physical correlations between observable and model parameters can all be described using probability densities. The general problem can be set as a problem of *combining* all of this information (*Tarantola*, 2005). Using this kind of approach, the solution of the problem and the analysis of uncertainty can be performed in a fully nonlinear way; in all usual cases, the results obtained with this method may reduce to those obtained using more conventional approaches.

Purely probabilistic formulations of inverse theory started to be proposed around 1970 (e.g., *Kimeldorf and Wahba*, 1970). *Jackson* (1979) explicitly introduced prior information in the context of linear inverse problems, an approach that was generalized by *Tarantola and Valette* (1982a,b) to nonlinear problems.

If we have a physical theory that can be used to solve the forward problem, i.e., that given an arbitrary model **m** it allows us to predict the theoretical data values $\mathbf{d}_{calc}$ that an ideal measurement should produce (if **m** was the actual system), the generally non linear function that associates to any model **m** the theoretical data values **d** is a (generally not unique) *parameterization* of the system. Independently of any particular parameterization, it is possible to introduce an abstract space of points (**M**), each of which represents a conceivable model of the system. Once a particular parameterization has been chosen, with each point **m** of the model space **M** a set of numerical values $\{m^1, ..., m^n\}$ is associated. This corresponds to the definition of a kind of *system of coordinates* over the model space (*Tarantola*, 2005; *Menke*, 1989).

On the other hand, to obtain information on model parameters, we perform some measurements of some observable parameters, so it is possible to define a *data space* (**D**); any conceivable (exact) result of the measurement then corresponds to a particular point **d** on the space **D**. The coordinates **d**= $\{d^1, ..., d^i\}$ (where i belongs to some discrete and finite index set) are then the *components* in the data space.

The physical theory can be used to solve the direct problem, i.e., that given an arbitrary model **m**, it allows us to predict the theoretical (error-free) data values $\mathbf{d}_{calc}$ that an ideal measurement should produce. The function that associates to any model **m** the theoretical data values $\mathbf{d}_{calc}$ can be denoted as:

$$d^i_{calc} = f^i(m^1, m^2, \ldots) , \quad \text{or}: \quad \mathbf{d}_{calc} = \mathbf{g}(\mathbf{m}) \tag{2.5}$$

for $i = 1, 2, \ldots$. The (usually nonlinear) operator $\mathbf{g}(\cdot)$ is called the *forward operator* and it expresses our mathematical model of the physical system. The predicted values in general cannot be identical to the observed ones because of measurement uncertainties and modeling imperfections. These two sources of error generally produce uncertainties of the same order of magnitude. The proper way of putting together measurements and physical predictions – each with its own uncertainties– is still a matter in progress (*Tarantola*, 2005). In Tarantola's approach, it is proposed to treat both sources of information symmetrically, stating that the more general way of describing any state of information is to define a probability density. In this context, the error-free equation 2.5 is replaced with a probabilistic correlation between model parameters **m** and observable parameters **d**.

The joint probability density describing the correlations that correspond to our physical theory together with the inherent uncertainties of the theory (due to imperfect parameterization or any fundamental lack of knowledge) will be denoted as $\Theta(\mathbf{d}, \mathbf{m})$. Nontrivial complications arise when the relation between **d** and **m** is not linear, which appear when trying to properly define the notion of conditional probability density. This topic is affronted in *Mosegaard and Tarantola* (2002). In general, in some situations it is possible, for every model **m**, do slightly better than to exactly predict an associated value **d**: one may, for every model **m**, exhibit a probability density for **d**, denoted as $\theta(\mathbf{d} \mid \mathbf{m})$. Then, the joint probability density can be written as the product of a conditional and a marginal. Taking for the marginal of the model parameters the homogeneous probability density μ_M (for details see *Tarantola* (2005)) then gives:

$$\Theta(\mathbf{d}, \mathbf{m}) = \theta(\mathbf{d} \mid \mathbf{m}) \, \mu_M(\mathbf{m}) \tag{2.6}$$

On the other hand, the observations are represented by the set of parameters **d** (*data*). However, the result of a measurement is not just a set of "observed values" but a state of information acquired on some observable parameter. If $\mathbf{d} = \{d^1, d^2, \ldots, d^n\}$ represents the set of

observable parameters, the result of the measurement can be represented by a probability density $\rho_D(\mathbf{d})$. Although the instrumental errors are an important source of data uncertainties, in geophysical measurements there are other sources of relevant uncertainty as, for example, the environmental noise. In the same way, the prior information on model parameters can be described by a probability density in the model space $\rho_M(\mathbf{m})$. As by definition the prior information on model parameters is independent of observations, the whole information we have in both model and observable parameters can then be described by the *joint probability density*:

$$\rho(\mathbf{d}, \mathbf{m}) = \rho_D(\mathbf{d}) \, \rho_M(\mathbf{m}) \tag{2.7}$$

The states of information defined by equation 2.6 (which represents the information on the physical correlations between $\mathbf{d}$ and $\mathbf{m}$ as obtained by a physical model) and equation 2.7 (which represents both information obtained on observable parameters (data) and prior information on model parameters), combine to produce the *posterior* state of information $\sigma(\mathbf{d}, \mathbf{m})$. Following *Tarantola* (2005), the probability density representing the posterior information is defined as

$$\sigma(\mathbf{d}, \mathbf{m}) = k \, \frac{\rho(\mathbf{d}, \mathbf{m}) \, \Theta(\mathbf{d}, \mathbf{m})}{\mu(\mathbf{d}, \mathbf{m})} \tag{2.8}$$

where $\mu(\mathbf{d}, \mathbf{m})$ and k represent the homogeneous state of information and a normalization constant, respectively.

With the posterior information $\sigma(\mathbf{d}, \mathbf{m})$ defined in the space $\mathbf{D} \times \mathbf{M}$, the *marginal probability density* $\sigma_M(\mathbf{m})$ provide the posterior information in the model space $\mathbf{M}$ (*Tarantola*, 2005), and can be considered as the most general solution (for model parameters) of a given problem:

$$\sigma_M(\mathbf{m}) \;=\; \int_{\mathbf{D}} d\mathbf{d} \, \sigma(\mathbf{d}, \mathbf{m}) \;=\; k \, \rho_M(\mathbf{m}) \int_{\mathbf{D}} d\mathbf{d} \frac{\rho_D(\mathbf{d}) \, \theta(\mathbf{d}|\mathbf{m})}{\mu_D(\mathbf{d})} \tag{2.9}$$

2.3.2   Setting the Problem and General Solution

The forward operator $\mathbf{g}(\cdot)$, which links the model parameters $\mathbf{m} = \{A_i, \ \gamma_i, \ \omega_i, \ \theta_i\}$, $i = (1, 2, \ldots, n_w)$ with the observed waveform $\mathbf{d} = \{x(t)\}$, is the linear combination of coherent sinusoids with amplitudes exponentially decaying or growing with time such as defined by *Kumazawa et al.* (1990):

$$\mathbf{d} = \mathbf{g}(\mathbf{m}) = \sum_{i=1}^{n_w} A_i e^{\gamma_i t} \cos(\omega_i t + \theta_i) \tag{2.10}$$

If we assume that modeling uncertainties are negligible compared to observational uncertainties, and that the data space is a linear space (which implies that the homogeneous probability density over the data space is constant, for details see *Tarantola* (2005), and *Menke* (1989)), then $\theta(\mathbf{d}|\mathbf{m}) = \delta(\mathbf{d} - \mathbf{g}(\mathbf{m}))$, where $\mathbf{d} = \mathbf{g}(\mathbf{m})$ is the exact solution of the forward problem. Then, equation 2.9 gives:

$$\sigma_M(\mathbf{m}) = k' \, \rho_M(\mathbf{m}) \, \rho_D(\, \mathbf{g}(\mathbf{m}) \,) \tag{2.11}$$

where k' is a constant, and $\rho_D(\, \mathbf{g}(\mathbf{m}) \,) = L(\mathbf{m})$ is the likelihood function giving a measure of how good a model $\mathbf{m}$ is explaining the data.

There are three important elements in equation 2.11: (a) the prior probability density in the model space, $\rho_M(\mathbf{m})$, which can be replaced by its homogeneous limit $\mu_M(\mathbf{m})$ if no prior information is available, (b) the probability density $\rho_D(\mathbf{d})$ which describe the results of our measurements and the attached uncertainties, and (c) the nonlinear function $\mathbf{g}(\mathbf{m})$ solving the direct problem.

Now we can replace the function defining our forward operator (equation 2.10) in equation 2.11 to define the general solution for the posterior distribution in the model space. To do it, we need to assign a probability function to $\rho_D(\mathbf{d})$ to describe the measurements and uncertainties of observed data. Due to the nature of seismic records, seismic noise is overlapped to the signals of interest; in our specific problem (as in all applications that use the seismic waveforms in time domain) the noise perturbing the signal of interest may be of different nature and difficult to describe; in general terms we arbitrarily can decompose it in two wide components: (1) a random component which may be associated, e.g., with measurement uncertainties, and (2) an "autocorrelated" component, which may be described as coherent oscillations whose source is not the same as the source of the target event analyzed (e.g. wave oscillations from other seismic events or from activity at the Earth's surface, etc.). In our analysis, the probability function $\rho_D(\mathbf{d})$ represents the random, non-autocorrelated component of the seismic noise; if there is an autocorrelated component (e.g. coherent waves overlapping the signal of interest), they probably will be considered as *spurious* wave elements in the solution; in section 2.5 we discuss a strategy to select an *optimum* number of wave elements to describe the observed data in order to avoid to over-fit the signal with this kind of spurious wave elements.

In this part we concentrate in the definition of a probability function to describe the random component of the noise; in many geophysical problems (and also in the case of the Kumazawa's solution in the Sompi method) the noise in measurements is generally described using a Gaussian distribution. In our approach we do not have any constrain and then we can introduce a general description of the data using a more general family of distribution functions.

Let's consider the probability density function defined by:

$$f_p(x) = \frac{p^{1-1/p}}{2\sigma_p \Gamma(1/p)} \, \mathbf{exp} \left\{ -\frac{1}{p} \frac{|x - x_0|^p}{(\sigma_p)^p} \right\}$$
(2.12)

which defines a *Generalized Gaussian* family of well known functions centered at x_0; it defines a generalized family of error distribution that includes, e.g., the symmetric Exponential ($p = 1$) and Gaussian ($p = 2$) distributions, and as limiting cases (as $p \to \infty$) it includes all continuous uniform distributions (box function) on bounded intervals of the real line (e.g., *Varanasi and Aazhang*, 1989; *Tarantola*, 2005).

Plugging equations 2.12 for $\rho_D(\mathbf{d})$ and 2.10 for the direct model, in equation 2.11, the general equation to define the posterior information $\sigma_M(\mathbf{m})$ in the model space for our problem is

$$\sigma_M(\mathbf{m}) = K \, \rho_M(\mathbf{m}) \, \mathbf{exp} \left\{ -\frac{1}{p\sigma_p^p} \sum_{j=1}^{s} |\mathbf{d}_{cal}^j - \mathbf{d}_{obs}^j|^p \right\}$$
(2.13)

where $\mathbf{m} = \{A_i, \gamma_i, \omega_i, \phi_i \; ; \;\; i = 1, 2, \dots, n_w\}$, are the $\{4 \times n_w\}$ model parameters for n_w wave elements, s the number of samples of the waveform, and p the order of the Generalized Gaussian, which can be chosen analyzing a noise sample (before or after the specific event) to select the distribution function from that family which best describes the noise characteristics; $\mathbf{d}_{cal}^j$ is the prediction of the $j-th$ sample due to a model $\mathbf{m}$ using the direct model defined in equation 2.10, and $\mathbf{d}_{obs}^j$ is the $j-th$ sample effectively observed.

Equation 2.13 represents the posterior joint probability density for the model parameters and constitutes the general solution to our problem. In section 2.4 we show an strategy based on a Markov-chain Monte Carlo method to sample the posterior distribution $\sigma_M(\mathbf{m})$ in order to get information about model parameters.

2.4   Exploring the Posterior Distribution using a Markov-chain Monte Carlo Approach

When the number of model parameters is small and the computation of $\sigma_M(\mathbf{m})$ for any arbitrary $\mathbf{m}$ does not consume too much computer time, it is feasible to define a grid over the model space, sample it systematically, and use those results directly to discuss the information obtained on model parameters. However, as the number of model parameters grows, the number of required points increases too rapidly since more dimensions are added, then the systematic exploration of the model space must be replaced with a random exploration. In this section we discuss a strategy to solve our problem by exploring the model space using a Markov-chain Monte Carlo (MCMC) approach based on the Metropolis algorithm (*Metropolis and Ulam*, 1949; *Metropolis et al.*, 1953; *Hastings*, 1970), and the implementation of a

Simulated Annealing strategy (*Kirkpatrik et al.*, 1983) to improve the process of finding the maximum of the posterior distribution.

2.4.1 Implementation of a Metropolis algorithm

A MCMC method attempts to simulate direct draws from some complex distributions of interest. It is random (Monte Carlo) and has no memory, since each step depends only on the previous step (Markov chain). The basic idea of the Metropolis algorithm is to perform a random walk that, sampling some initial (proposal) probability $\rho(\mathbf{m})$ and using a probabilistic rule, modifies the walk (some proposed moves are accepted, some are rejected) in such a way that the guided random walk samples the target distribution.

Our goal is to draw samples from a distribution of the form $\sigma_M(\mathbf{m}) = \kappa L(\mathbf{m})$ (e.g. Eq.. 2.13), where the normalizing constant κ may not be known (generally it is very difficult to compute). The Metropolis algorithm (*Metropolis and Ulam*, 1949; *Metropolis et al.*, 1953) generates a sequence of draws from this distribution as follows: (1) start with an initial value $\mathbf{m}_0$ satisfying $L(\mathbf{m}_0) > 0$; (2) using the current $\mathbf{m}_t$ value, sample a candidate point $\mathbf{m}^*$ from the proposal (symmetric) distribution $\rho(\mathbf{m})$; (3) given the candidate point $\mathbf{m}^*$, calculate the ratio (α) of the density at the candidate ($\mathbf{m}^*$) and current ($\mathbf{m}_t$) points:

$$\alpha = \frac{L(\mathbf{m}^*)}{L(\mathbf{m}_t)} \tag{2.14}$$

note that the normalizing constant κ cancels out. If the move increases the density (i.e. $\alpha > 1$) then the candidate point is accepted (then $\mathbf{m}_{t+1} = \mathbf{m}^*$) and return to step 2. If the jump decreases the density (i.e. $\alpha < 1$), then with probability α accept or reject randomly the candidate point and return to step 2. This generates a Markov chain $\{\mathbf{m}_0, \mathbf{m}_1, \ldots, \mathbf{m}_\nu, \ldots\}$, as the transition probabilities (from $\mathbf{m}_t$ to $\mathbf{m}_{t+1}$) depends only on $\mathbf{m}_t$. Following a sufficient burn-in period (e.g. ν steps), the chain approaches its stationary distribution (as the number of iterations goes to infinity) and then the samples $\{\mathbf{m}_{\nu+1}, \mathbf{m}_{\nu+2}, \ldots, \mathbf{m}_{\nu+i}\}$ are samples from the target distribution $\sigma_M(\mathbf{m})$ (e.g., *Tierney*, 1994; *Sambridge and Mosegaard*, 2002). In our application, we implemented the Metropolis algorithm in a single-component configuration, which means that just one component of $\mathbf{m}$ is updated at a time; thus each updating step produces a move in the direction of a coordinate axis if the candidate is accepted.

In the algorithm to solve the problem, we have also included an optimization technique based on a *Simulated Annealing* process (*Kirkpatrik et al.*, 1983). This approach is often used for finding the maximum (or minima) of complex functions with multiple peaks. The idea is that when the space of model parameters is sampled, a reasonable number of *down-hill* moves are accepted; it consents to perform a wide exploration of the space during the first iterations. As the number of iterations increases, the probability of accepting down-hill movements is smaller up to a limit beyond which it behaves as the Metropolis algorithm. The analogy (and

hence the term) is the annealing of a crystal as temperature decreases: when the *temperature* parameter $T(n)$ is gradually lowered, it is simulated an annealing process in which the system is taken from a disordered high-temperature state to a well-ordered low temperature state. Simulated annealing is very closely related to Metropolis sampling, differing only in that the probability α of a move is given by

$$\alpha_{SA} = \min\left[1, \left(\frac{L(\mathbf{m}^*)}{L(\mathbf{m}_t)}\right)^{1/T(n)}\right] \tag{2.15}$$

where the function $T(n)$ is called the *cooling schedule* and typically is a function with geometric decay, and $n = 1, 2, 3, \ldots$ is the number of the iteration. For $T(n)$ we use a function which starts at $T_0 > 1$ and cools down to $T_f = 1$ (in order to get the Metropolis algorithm) after n_c iterations, as follows (e.g. *Kirkpatrik et al.*, 1983):

$$T(n) = \max\left(T_0\left(\frac{T_f}{T_0}\right)^{n/n_c}, T_f\right) \tag{2.16}$$

A key issue in the implementation of a MCMC sampler is the number of runs (steps) until the chain approaches the stationary distribution (i.e. the length of the burn-in period). In our tests using real signals, the first 10000 to 15000 elements were typically discharged as burn-in period, and chains generally ran up to about 30000 to 40000 steps. The velocity at which chains converge depends on many factors. For example, a poor choice of starting values and/or proposal distribution, can increase the required burn-in time. A brief discussion about the convergence assessment is performed in section 2.4.2.

2.4.2   Computing central and dispersion estimators

Strictly speaking, the general solution of the inverse problem following this probabilistic approach is the posterior distribution $\sigma_M(\mathbf{m})$ in the model space. Since the distribution of samples obtained by the MCMC algorithm will indeed converge to $\sigma_M(\mathbf{m})$ as the number of iterations goes to infinity, the most complete solution that we can provide about each specific model parameter are the probability distributions (e.g. empirical CDF or PDF) inferred from the samples of the Markov chains, which should approach the marginal distributions of $\sigma_M(\mathbf{m})$. However, if we are interested on using the information of some of the estimated model parameters (for instance, the frequency, f, and the growth rate, g) for volcano monitoring purposes, we can compute some central estimators and estimators of dispersion that characterize the probability distributions. Since we are not constrained by any particular assumption, we can calculate any kind of estimators (mean, median) to indicate our *best guess* for the model parameter. Then, in order to characterize the distribution of each parameter of the model, we arbitrarily use the median of the distribution as the best guess, while an error interval is defined using the 10th (lower bound) and 90th (upper bound) percentiles.

Before using a MCMC sequence for estimating information about parameters of interest, it is generally necessary to perform convergence tests to assess whether stationarity has indeed been reached. To do it, we follow the next approach: (1) first, we define a burn-in period, which corresponds with the first part of the chain to be discharged; (2) we try to reduce the possible autocorrelation of the chains; in general, we can expect adjacent members from a Metropolis sequence to be positively correlated, and we can quantify the nature of this correlation by using an autocorrelation function. To reduce autocorrelation of the chain, we perform a thinning procedure, which means to re-sample the chain taking one point every τ points from the output after the burn-in period. To select the optimum value of τ, we plot the autocorrelation as a function of the time lag, which should show a geometric decay as the time lag increases. (2) To test the convergence of the sequences, there are different diagnostic checks (e.g., *Geweke*, 1992; *Geyer*, 1992; *Gelman and Rubin*, 1992); we use the simple test proposed by *Geweke* (1992) in which the sequence (after removing a burn-in period) is sampled in two parts (e.g. the first 30% and the last 50%). If the chain is at stationarity, the statistics of both samples should be similar. For this scope, it is often used the z-test: a value larger that 2 indicates that the mean of the series is still drifting (it is often referred as the Geweke-z-score) and then a longer burn-in is required.

2.5   Model selection and minimum number of wave elements to fit the observed data

The forward operator defined in equation 2.10 includes the parameter n_w which should be also determined. It defines the number of wave elements to be used in order to obtain the best fit of the data. In general we do not know a priori the number of wave elements that conform a given data set, and we do not have any physical constrain to determine this parameter; this is a difficulty often found in statistical inference and inverse problems. To select a particular model out of a set of candidate models, it is often used the so called Akaike's Information Criteria - AIC (*Akaike*, 1974), which states that model selection should be done in terms of unbiased estimate of the maximum log-likelihood of the model, penalizing this value with the number of independently adjustable parameters. However, due to different factors as e.g., the fact that noise overlapped to the real signals may not be exactly random (which means that it may be reasonably autocorrelated and could be modeled in terms of realizations of some dynamical system which is not the target physical system generating the signal of interest), it is reasonably expected that AIC indicate a minimum at a relatively high number of wave elements respect to those affectively required to describe the *true* signal of interest, resulting in the coexistence of several *spurious* wave elements describing the noise component (*Kumazawa et al.*, 1990). In the Sompi method (*Kumazawa et al.*, 1990; *Hori et al.*, 1989) such a problem has been addressed in a somewhat empirical way; after some numerical experiments, the authors have found that after some minimal AR order, estimates of characteristic frequencies for *true* signals tend to be stationary with respect to the change

in AR order (the AIC often takes also a stationary value or gradually and slowly decreases). Therefore, the authors state that the characteristic solutions which are more likely the *true* ones can be determined by plotting cumulatively the estimated characteristic frequencies of all near equi-AIC models in a $f - g$ diagram (plot of the complex frequencies on a 2D-plane with (real) frequency and growth rate axes).

We propose to calculate a residual value (Δ) between the original and calculated (for a given model) time series (since we do not perform maximum likelihood estimations of model parameters). In our numerical tests we have observed that the residual Δ has a similar behavior as AIC (see section 2.6). The residual can be defined as:

$$\Delta(\mathbf{m}) = \frac{\sum_{j=1}^{s} (\mathbf{d(m)}_{cal}^{j} - \mathbf{d}_{obs}^{j})^2}{\sum_{j=1}^{s} (\mathbf{d}_{obs}^{j})^2} \qquad (2.17)$$

where $\mathbf{d(m)}_{cal}^{j}$ and $\mathbf{d}_{obs}^{j}$ are the $j - th$ samples of the synthetic and observed waveforms, respectively, and $\mathbf{m}$ is the specific model considered (e.g., *Ohminato et al.*, 1998). What we can see is that as the number of wave elements (n_w) increases, the residual calculated with equation 2.17 decreases, first with a steep slope up to a limit (in general, when the residual value is lesser than about 0.1) beyond which it continues whether to decrease with a much more gentle slope or it oscillates around low residual values (e.g. see Figs 2.1a and c).

This behavior is an important observation, since it may give further empirical constrains in order to define an interval of *optimum* values for n_w, a kind of boundary in the number of wave elements to accurately describe the signal, below which the signal is probably sparsely reconstructed, and above which an abnormal number of spurious wave elements could be included, which probably improve the fit of the recorded signal but mainly contributing in the fit of auto-correlated seismic noise. To obtain the residual it is necessary to build a synthetic waveform (to compare with the original signal) which may also provide an additional tool to check how well a given model is describing the data. In section 2.6 there are some examples performed with synthetic signals, and further discussion about the definition of this *optimum* number of wave elements; later, in section 2.7, there is a discussion of a practical problem where the concept of the *optimal number of wave elements* may be useful for both the monitoring of characteristic complex frequencies of LP events in a given time interval, and also to speed up calculations.

2.6   Numerical test using synthetic waveforms

We now illustrate the methodology described in this chapter analyzing some synthetic waveforms. Two synthetic waveforms were created using known model parameters, and the
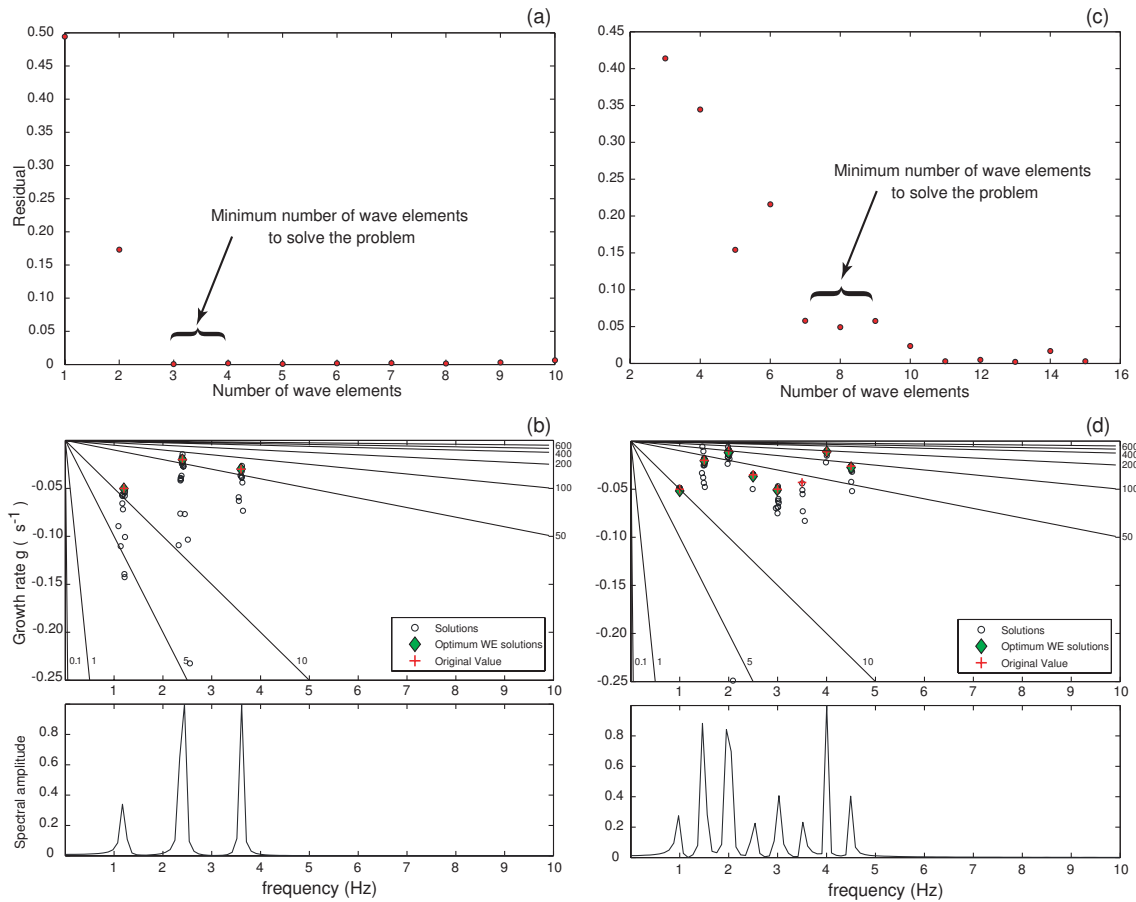
Figure 2.1: Solutions (frequency and growth rate) of synthetic examples SYNA and SYNB. Example SYNA (3 wave elements): (a) residuals in function of the number of wave elements of the model; (b) stacked solutions (circles) of models with n_w from 1 to 10 in a *f-g* diagram. Black lines represent iso-values of Q. Original values of the parameters are plotted with plus marks, and the solutions with selected optimum n_w are plotted with diamond marks. Example SYNB (8 wave elements): (c) residuals in function of the number of wave elements of the model; (d) stacked solutions (circles) of models with n_w from 1 to 16 in a *f-g* diagram. Symbols are the same as for example SYNA.

methodology described in previous chapters was used to determine those model parameters. Tables 2.1 and 2.2 summarize the parameter values used to generate the synthetic data sets SYNA and SYNB, respectively.

Table 2.1: Wave elements used to create synthetic data SYNA. Each wave element is specified by four real parameters: (real) frequency, growth rate, amplitude and phase.

frequency (Hz)	growth rate (s^{-1})	amplitude	phase (rad)
1.200	-0.050	10.900	0.900
2.400	-0.020	12.000	1.500
3.600	-0.030	11.500	2.100

In general, as a previous step before the inversion process, the time series are resampled and band-pass filtered using a second-order Butterworth filter. Then, model parameters are determined using the methodology described before in an iterative process in which the number of wave elements is increased in one at each step. Fig. 2.1 shows the results obtained for the SYNA (2.1a and 2.1b) and SYNB (2.1c and 2.1d) data sets, respectively. The simple example SYNA has been created using three wave elements (Table 2.1) with overlapped Gaussian noise. Fig. 2.1(a) is a plot of the residuals calculated using equation (2.17) against the number of wave elements. It is evident that the residual value monotonically decays up to reach the original number of wave elements. For solutions with a higher number of wave elements, the residual remains in a low value, suggesting that the point around the original number of wave elements represents an evident change point in the slope of the residual values. The first panel in Fig. 2.1(b) is a cumulative plot of all the solutions (for different number of wave elements) in a *f-g* diagram (error bars have been omitted for simplicity) while the second panel is the FFT spectrum. Circles correspond to all overlapped solutions (for $n_w = 1, 2, \ldots, 10$), diamonds are the solutions for $n_w = 3$ (selected as the optimum n_w), and plus-marks are the original values (Table 2.1) used to create the signal. The overlapped solutions show a higher density of points (recurrent solutions) around the original (*true*) parameter values. Note that if we just consider a solution with n_w around the change point in the residual plot (e.g. $n_w = 3$ or 4), most of the high density areas in the *f-g* plot (and then the more likely solutions) are successfully sampled.

Similar observations can be done in the solutions for the example SYNB, which has been created using eight wave elements (Table 2.2) plus some overlapped Gaussian noise. Fig. 2.1(c) is a plot of the residuals calculated using equation 2.17 against the number of wave elements, and Fig. 2.1(d) is a plot of the frequency and growth rate solutions in a *f-g* diagram. Symbols are the same as in previous example. In this case, we can define an interval around the correct number of wave elements (e.g. [7, 10]) to define the optimum number of wave elements to correctly describe the data, and whose solutions successfully sample most of the areas where the *true* values are located.

Table 2.2: Wave elements used to create synthetic data SYNB. Each wave element is specified by four real parameters: (real) frequency, growth rate, amplitude and phase.

frequency (Hz)	growth rate (s^{-1})	amplitude	phase (rad)
1.000	-0.050	10.900	0.900
1.500	-0.020	12.000	1.500
2.000	-0.010	11.500	0.100
2.500	-0.035	9.500	1.100
3.000	-0.050	13.500	0.500
3.500	-0.043	8.500	0.700
4.000	-0.011	9.500	1.900
4.500	-0.026	7.500	0.200

To give some numerical values as results of model parameters estimated for examples SYNA and SYNB, Tables 2.3 and 2.4 summarize, respectively, the values of central estimators (median) and errors (defined using the 10th and 90th percentiles) of the empirical distributions associated with each parameter of the model, as calculated for the *optimal* number of wave elements selected for each example (3 and 8, respectively).

The optimal number of wave elements defined in this way, provide us insights about the range of plausible models which could be used to successfully explain the data. The wave elements that make up the solution are a combination of highly likely signal components and some (hopefully few) random wave elements describing autocorrelated noise. The accuracy of the optimum solution may be also evaluated plotting the observed data and the synthetic waveform calculated (equation 2.10) using the estimated parameters of the optimum n_w solution selected. In Fig. 2.2 are plotted the observed (original) and synthetic (reconstructed) waveforms of the examples SYNA (2.2a) and SYNB (2.2b). Using this plot it is possible to inspect how well our model describes the data.

Additionally, and due to the non-linear relation between observed data and model parameters, a histogram of accepted solutions and/or a plot of the empirical cumulative distributions (CDF) for each model parameter may help in the evaluation of the quality of a specific solution and/or the possibility of multiple solutions (i.e. if the distribution is multimodal). Examples of the frequency histograms for each model parameter of example SYNA are plotted in Fig. 2.3, and the empirical CDFs for the same data set are plotted in Fig. 2.4.

Those two synthetic examples provide some important information about the efficiency of the solutions obtained using the probabilistic approach proposed in this work. On the other hand, it is possible to get some insights to address (at least empirically) the difficult problem

Table 2.3: Results of spectral analysis for the synthetic data SYNA. The *best guess* is defined using the median of the distribution, while an error interval is defined using the 10th and 90th percentiles

frequency (Hz)			growth rate (s^{-1})			amplitude			phase (rad)		
median	10th.	90th.	median	10th.	90th.	median	10th.	90th.	phase	10th.	90th.
1.200	1.200	1.203	**-0.050**	-0.050	-0.051	**10.984**	10.945	11.098	**0.900**	0.877	0.902
2.399	2.399	2.400	**-0.020**	-0.020	-0.021	**12.021**	12.015	12.197	**1.535**	1.513	1.539
3.599	3.583	3.599	**-0.030**	-0.030	-0.035	**11.488**	11.470	11.690	**2.125**	2.114	2.247

Table 2.4: Results of spectral analysis for the synthetic data SYNB. The *best guess* is defined using the median of the distribution, while an error interval is defined using the 10th and 90th percentiles

frequency (Hz)			growth rate (s^{-1})			amplitude			phase (rad)		
median	10th.	90th.	median	10th.	90th.	median	10th.	90th.	phase	10th.	90th.
0.000	0.000	0.377	-0.299	-0.196	-0.299	1.133	0.964	1.466	0.654	-1.297	1.182
1.000	1.000	1.004	-0.053	-0.052	-0.054	11.445	9.770	11.686	0.893	0.874	0.912
1.500	1.491	1.500	-0.021	-0.015	-0.021	12.311	6.821	12.334	1.509	1.505	1.976
1.999	1.994	2.000	-0.010	-0.010	-0.010	11.541	11.541	11.708	0.155	0.127	0.395
2.500	2.474	2.500	-0.038	-0.038	-0.062	10.423	10.351	11.693	1.145	1.130	1.356
2.990	1.560	2.996	-0.051	-0.050	-0.051	13.404	13.104	13.452	0.729	0.670	0.800
4.000	4.000	4.028	-0.010	-0.010	-0.018	8.737	8.541	8.789	1.873	1.132	1.877
4.503	4.503	4.504	-0.028	-0.028	-0.031	8.067	8.014	8.841	0.099	0.091	0.100
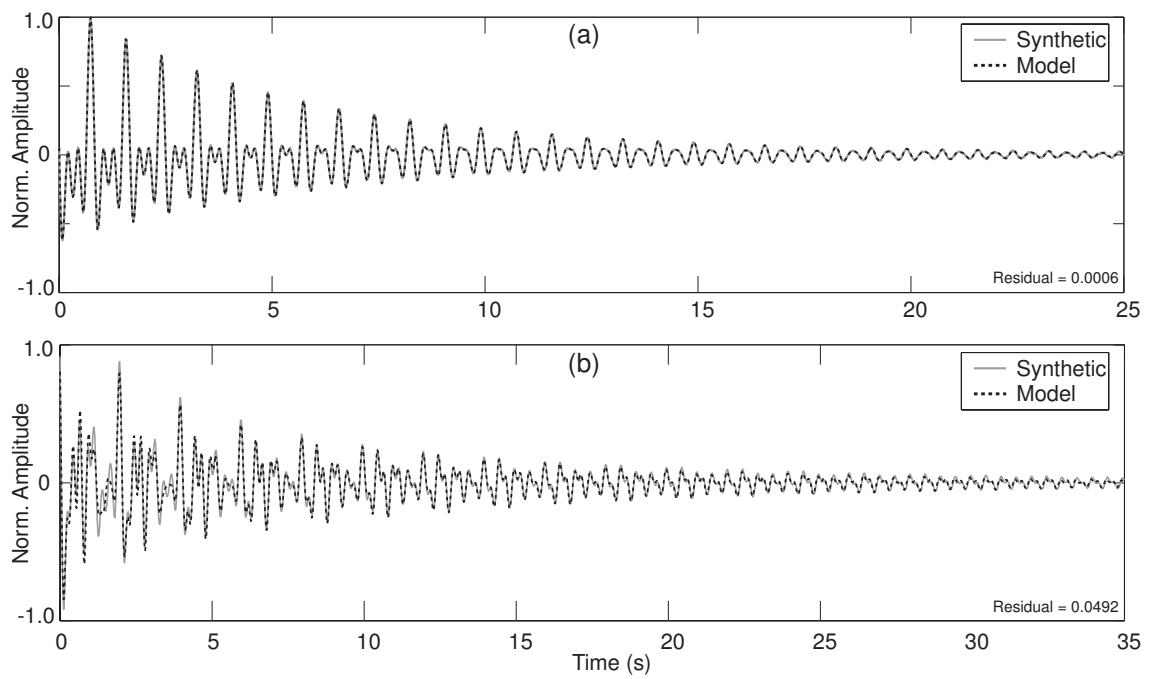
Figure 2.2: Numerical tests: waveforms of the original examples (solid lines) and the reconstructed signals using the optimum n_w models (dashed lines), for the synthetic examples SYNA (a) and SYNB (b).
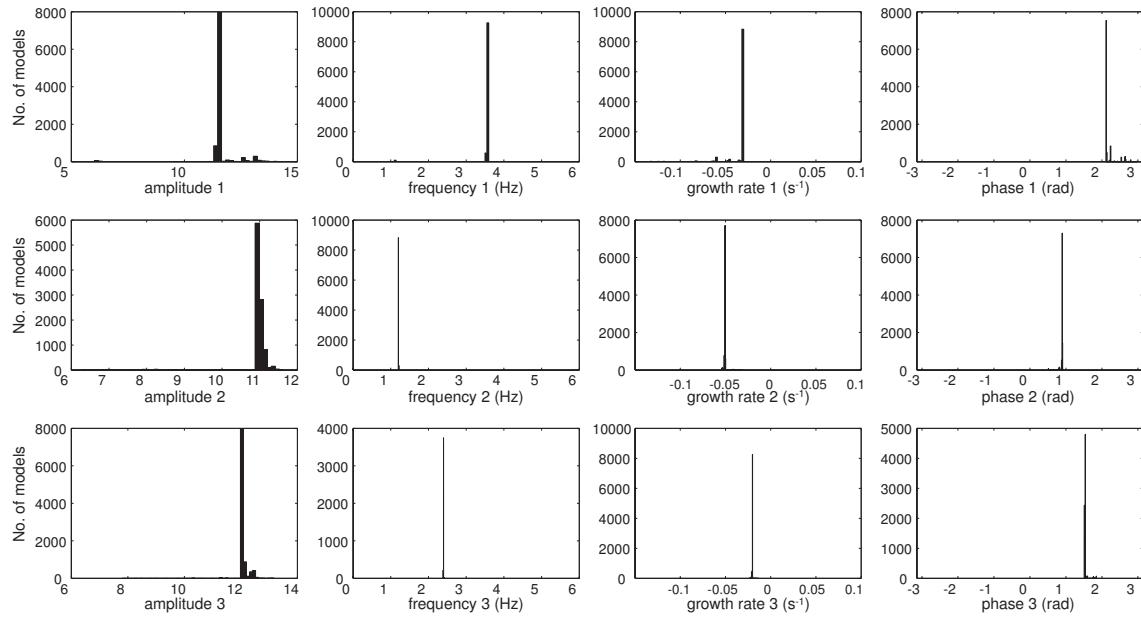
Figure 2.3: Frequency histograms of solutions for the model parameters (amplitude, frequency growth rate and phase angle) for the (three) wave elements of the solutions of example SYNA. The frequency histograms are plotted using the MCMC data, after removing the burn-in period and after the thinning procedure.

of selecting a kind of *optimum* number of wave elements to accurately describe the data. In section 2.7, further discussion about the importance of this *optimum* number of wave elements is performed, in order to propose a possible use of this information for monitoring LP activity in active volcanoes.

2.7 Strategy for routinely use characteristic complex frequencies of LP events for monitoring volcanoes

In recent years, characteristic complex frequencies (CF) of LP events became popular in volcano seismology due to its direct link to physical properties of the source resonator. After the works of *Kumagai and Chouet* (1999, 2000, 2001), its main use have been oriented toward the inference of some physical properties of fluids at the source of particular events.

It is important to consider here that the general theory at the source of this kind of time series analysis is based on the estimation of parameters linked to a dynamical system (which hypothetically generated the time series) in the most simple configuration described by a linear homogeneous differential equation. This may limit the kind of time series data suitable to be analyzed using this approach; for example, when the excitation function is present at
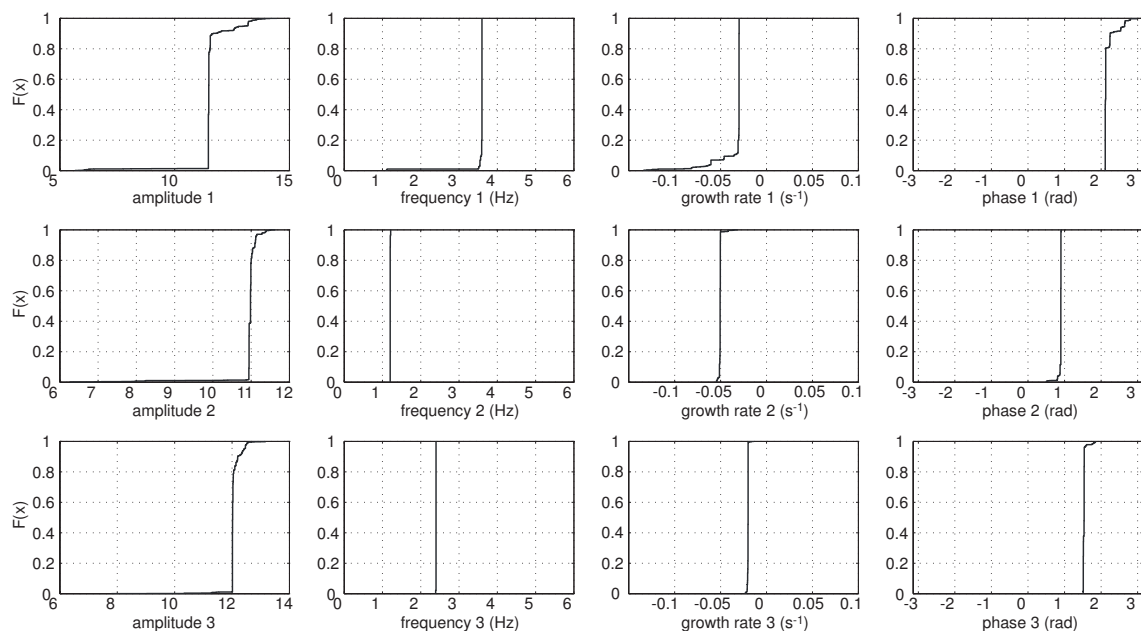
Figure 2.4: Empirical Cumulative Distributions (CDF) of the model parameters (amplitude, frequency growth rate and phase angle) for the (three) wave elements of the solutions of example SYNA. The CDFs are calculated from the MCMC data, after removing the burn-in period and after the thinning procedure.

all times in the signal, the analogous linear system should be described by an inhomogeneous differential equation and then its realization will not be represented simply by a linear combination of coherent wave elements (e.g., *Kumazawa et al.*, 1990); in this case an alternative method should be applied (e.g., *Yokoyama et al.*, 1997; *Nakano et al.*, 1998). This may become more critical in the case we want to use this information to make some inferences about fluid properties. For this reason we assume that the excitation is a time-localized function present only during a short time interval at the beginning of the event, and then the analysis should be performed in the tails of the events, to minimize the possibility of including the excitation function.

Beyond the valuable application of characteristic complex frequencies of LP events to explore fluid composition, which have provided an extraordinary tool for quantitative volcano seismology research, we consider that it could be possible to make use of LP complex frequencies in a more general way, specially oriented to quantitatively provide monitoring information in order to improve our ability to detect generic changes in the volcanic system. Let's consider the volcano as a system, where each pre-existent fracture able to contain or to drive fluids may be considered as a potential source of LP signals. This set as a whole represents all the possible structures that could generate events (Fig. 2.5a). During a given interval of time $\Delta\tau$ (hours, days, months, etc.), just some specific structures (one or more) of that set are effectively excited and generate signals (Fig. 2.5b), then, if we consider the subset of LP events occurring during the time interval $\Delta\tau$ and cumulatively plot their characteristic complex frequencies in a *f-g* diagram, then we get a map of points which will represent the characteristic *f-g* values from signals generated by the subset of sources active during that specific time interval. Areas with high density of points represent more frequent solutions; for example, let's suppose that there exist just one structure (e.g. a crack, or a cylindrical conduit, etc.) with unmodified dimensions and fluid properties (i.e. density and acoustic velocity), which is repeatedly excited generating a subset of individual LP events, then the cumulative plot of the characteristic complex frequencies estimated for that subset of LP events should be represented by a few, well constrained, high-density (of points) regions. On the other hand, if during the time interval considered there are many different sources (different geometrical configurations, dimensions, fluid compositions, etc), then we could observe many different areas with significant concentration of solutions. This kind of plots may be used to perform a selective filtering of the solutions to generate time series of f, g and Q, which would represent the time evolution of the most likely characteristic solutions for the LP seismicity occurring in a given period of time. To pursue this objective we could use the stacked solutions of each event (as in Sompi analysis), or we could use a single solution (for each event) selected from the defined optimum n_w interval. This second option may be useful since it would prevent fill the space with spurious points, and also because the time required to process the data would be optimized.

Using this approach it is possible to systematically monitor the most significant characteristic complex frequencies and then get samples of the most active sources triggering LP events dur-
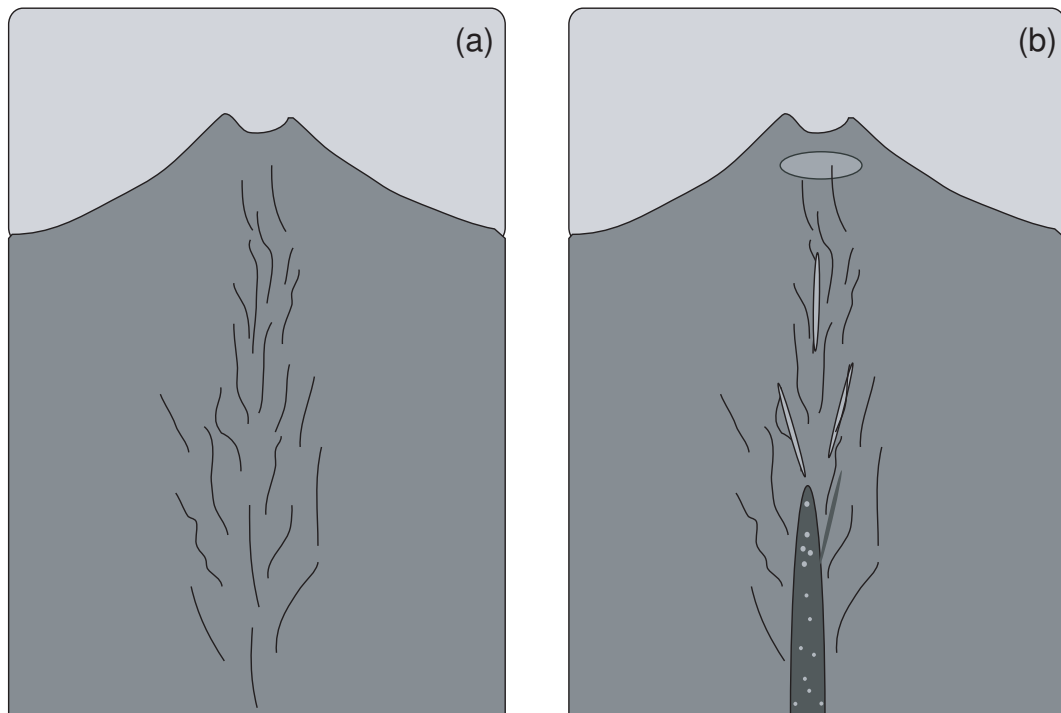
Figure 2.5: Schematic view of a volcanic system. (a) All existent structures which potentially could contain or drive fluids. (b) During a given time interval $\Delta\tau$, just some specific structures are effectively excited and generate LP signals.

ing a given time interval. In this way it would be possible to comprehensively identify changes in the characteristic complex frequencies of LP seismicity which may indicate generic changes in the physics of the source. Note that the objective of this analysis is completely different of the usual application oriented toward the inference of fluid properties. Our main interest is to find significant changes in the characteristic values of frequency (f), growth rate (g), and then the value of Q calculated using the former parameters (equation 2.4). To illustrate how this strategy could work, in section 2.8 we apply the overall methodology and this alternative monitoring strategy to real data from two active volcanoes: (1) Cotopaxi (Ecuador) and (2) Miyakejima (Japan) volcanoes.

2.8   Application to LP data

We applied the described methodology to LP data from two volcanoes: Cotopaxi (Ecuador) and Miyakejima (Japan). Cotopaxi is a huge, ice-capped, stratovolcano which had an important unrest in 2001-2002 characterized by swarms of both VT earthquakes and LP events, accompanied also by a small deformation localized in the NNE flank (*Molina et al.*, 2008). On the other hand, Miyakejima is a basaltic volcano located in the Izu islands, Japan, which in 2000 had an eruption characterized by strong deformation accompanied by VT and VLP seismic swarms. The eruption was also characterized by the migration of magmatic dikes toward the NW of the volcano, and by the formation of a $0.6 km^3$ collapse caldera in the summit of the volcano (e.g., *Nakada et al.*, 2005; *Ueda et al.*, 2005; *Kobayashi et al.*, 2009).

2.8.1   Case study 1: Cotopaxi Volcano (2001 - 2002)

Cotopaxi volcano (lat. 0.684S; long. 78.437W) is located in the Eastern Cordillera of the Ecuadorian Andes, 60 km south of Quito. This 5879 m high active volcano is notable for its relief (2000 - 3000 m), conical form, massive size (22-km diameter), and its glacier-clad steep flanks. Cotopaxi is instrumentally monitored since 1977 by the Instituto Geofísico - Escuela Politécnica Nacional (Quito), that provided us the waveforms. Renewed seismic activity of Cotopaxi began in January 2001 with the increased number of LP events, followed by a swarm of VT earthquakes in November 2001. In late June 2002, VLP events accompanying LP signals began beneath the volcano (*Molina et al.*, 2008). We analyzed the most important LP events occurred from December 2001 up to November 2002. We used the waveform data recorded at station COV1 featuring a Mark Products L-4C-3D three-component seismometer with a natural frequency of 1 Hz. This station is the closest to the summit crater of the volcano (1.7 km from the summit), and provided the best quality waveform data in the permanent seismic network operating at that time.

Fig. 2.6 shows some examples of solutions obtained for LP data from Cotopaxi volcano. Fig. 2.6(a) shows four waveforms (solid lines) and their synthetic model fits (dashed lines).

Figure 2.6: (a) Recorded waveforms (solid line) and the synthetic signal reconstructed using the optimum n_w models (dashed line), for four LP signals from Cotopaxi volcano. (b) *f-g* diagram with the stacked solutions (LP event from September 24, 2002, 16h44 GMT) for models with n_w between 3 and 20 (circles). The selected optimum n_w solutions (14) are shown with diamond marks, and error bars correspond to the 10th and 90th percentile of the distribution of the growth rate parameters. (c) Fourier spectra of the signal.

Using this kind of plots it is possible to control how well a specific model explains the recorded data. The residual for each fit is also reported. The synthetic waveforms plotted are generated using the model with the optimum number of wave elements selected for each signal. Fig. 2.6(b) is an example of a *f-g* diagram for the third waveform in Fig. 2.6(a), where the solutions for the models with $w_n = 3$ to 20 are plotted (solid circles). The solutions for the optimum number of wave elements ($n_w = 14$) are plotted using diamond symbols, with the respective error bars. Fig. 2.6(b) shows how the model with the selected optimum number of wave elements successfully samples the areas with higher density of points in the *f-g* diagram. In order to provide a further tool to control the quality of the frequency solution, a Fourier spectra is also plotted (Fig. 2.6c). In this way we did calculate *f-g* solutions and the respective optimum models for 617 LP events recorded in Cotopaxi volcano between December 22, 2001 and November 14, 2002.

Following the approach described in section 2.7, optimum solutions were stacked (using windows of 50 events) in order to find the most recurrent values and then to identify the most probable solutions indicating the most active sources during a given time interval. Fig. 2.7 shows the cumulative *f-g* diagrams for six different time intervals within the time period of available data. Fig. 2.7(a) shows the cumulative solutions (solid points) for the time interval

December 22, 2001 - March 23, 2002. Contours indicate the areas with higher density of points (warmer colors). The areas with higher concentration of points in this plot indicates that the most frequent solutions for LP events in this period are in the frequency interval from 2 to 4 Hz, and growth rate from about -0.05 to -0.01 s^{-1} (S1). These parameters indicate Q values which range from 50 to 100 approximately. It is noteworthy that these solutions are present almost all the time in the whole period of analysis, indicating one or more sources of LP events which are probably active almost permanently. Figs 2.7(b) and (c) show another cluster of solutions (S2) in a frequency range 0.8 - 1.5 Hz approximately and growth rate from about -0.15 up to -0.07 s^{-1}, present in a time interval from June 28 up to the end of August 2002. During this period, VLP signals (0.5 to 0.9 Hz) accompanying LP events were recorded at Cotopaxi, and solutions S2 are some evidence of that as seen in the short period seismometers. Analyzes of those VLP events from broad band records were performed by *Molina et al.* (2008) and have been interpreted as volumetric changes of a sub-vertical crack located beneath the NNE flank of the volcano. A very interesting set of solutions (S4) can be seen in Figs 2.7(e) and (f). This set of solutions corresponds with long-lasting coda LP events (known in volcanological literature as tornillo-type) which were recorded at Cotopaxi between September and November 2002. The solutions are in a frequency interval between 3 and 4 Hz and with growth rate > -0.02 s^{-1}, corresponding to signals with Q values in the interval 100 to 400. This subset of solutions corresponds to the LP events analyzed by *Molina et al.* (2008) using Sompi method (i.e. see Fig. 8 in *Molina et al.* (2008)), which based on the model proposed by *Kumagai and Chouet* (1999, 2000, 2001), were interpreted as the resonance of a crack above the magmatic system triggered by the release of particle-laden gas.

Figs 2.7(a to f) are examples of how the *f-g* parameters and derived Q values could be systematically used to monitor changes in some properties of the LP seismicity recorded in a specific volcano, during a given time interval. The *f-g* diagrams may be useful to filter the most likely solutions in order to generate time series of frequency, growth rate and Q values, as in Fig. 2.8. Subsets of solutions S1, S2, S3, and S4 highlighted in Fig. 2.7 have been also included.

Once some events of interest are identified, conventional techniques and models could be used to make physical interpretations about the source properties. It is important to point out here that if this information is going to be used to make inferences about the source properties (e.g. properties of fluids), it is necessary to correctly analyze signals from resonant systems, and avoid taking in consideration effects from the triggering mechanisms (i.e. the excitation function). An analytical explanation of this point may be found in *Kumazawa et al.* (1990), while an approach to simultaneously inversions of both characteristic complex frequencies and the source time function can be found in *Nakano et al.* (1998).
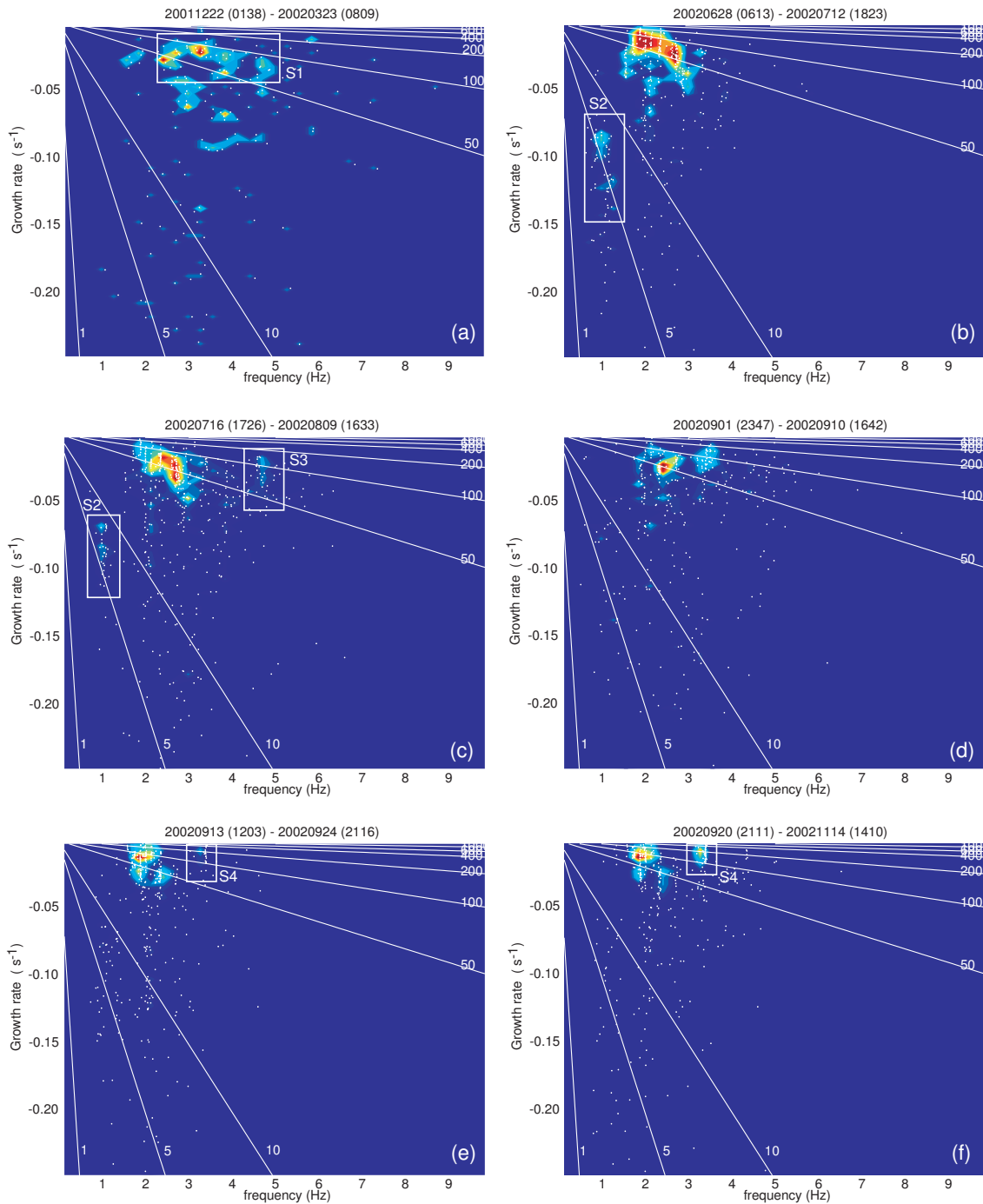
Figure 2.7: Stacked solutions of characteristic complex frequencies of LP events from Cotopaxi volcano (white points) plotted in *f-g* diagrams for six different time intervals from the period December 2001 - November 2002. Contours enclose areas with high density of points (warmer colors). Lines in the diagram correspond with iso-values of Q.
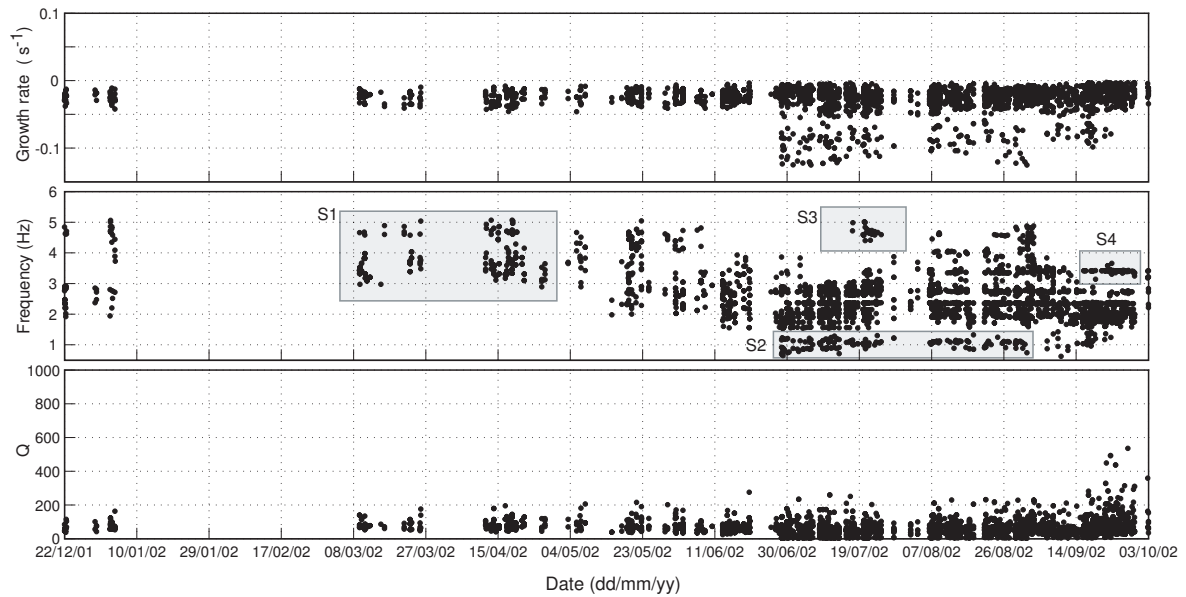
Figure 2.8: Time series of characteristic growth rate, frequency and *Q* parameters of LP events from Cotopaxi volcano, resulting from filtering points from high density areas in *f-g* diagrams from different time intervals.

2.8.2   Case study 2: Miyakejima Volcano (2000)

Miyakejima island (lat. 34.079N; long. 139.529E), located about 200 km south of Tokyo, Japan, is one of the most active basaltic volcanoes in Japan. The latest eruption started in 2000 and lasts for more than 9 years by now. A collapse caldera at the summit of the volcano was generated during this eruptive cycle. On the basis of surface phenomena observed, many authors have divided this eruptive period in at least four stages (e.g., *Nakada et al.*, 2005; *Ueda et al.*, 2005): (1) magmatic intrusion, (2) summit subsidence, (3) Explosions, and (4) degassing stage. For our test, we did analyze the waveforms of LP and VLP events occurred from 27 June to 07 July, 2000, which covers the magmatic intrusion stage and the beginning of the summit subsidence. The data used was provided by the National Research Institute for Earth Sciences and Disaster Prevention (NIED), Japan, from a seismic network composed by 4 Broad-Band observation points. The reference station chosen for data analysis is MKTB, featuring a STS-2 seismometer.

The main features of the VLP events analyzed have been summarized by *Kobayashi et al.* (2009). The VLP events are constituted of two parts: the initial part, an impulsive pulse which is dominant in the horizontal components, and the later part, a very long period oscillation with 0.2 or 0.4 Hz in dominant frequency. The source mechanisms of those VLP events have been studied by different authors. *Kobayashi et al.* (2009) analyzed the source mechanism of VLP events for the same period as our analysis. They performed waveform inversions and obtained both the location and the source time functions (STF), finding that

the initial part of the signals can be explained by a northward single force of about 1.5×10^8 N working at a depth of 2 km beneath the summit, while the later part is produced by a moment tensor solution at about 5 km depth and 2 km SW of the summit, with about 10^{12} Nm. Their interpretation has been that the single force was generated when magma containing rock blocks suddenly began to move in a choked subsurface magma path, and the resultant pressure waves propagated and excited a resonance oscillation of an axially symmetric structure; for this reason the shallow impulsive source is thought to trigger the deep oscillatory source.

Fig. 2.9 shows some examples of solutions obtained for VLP/LP data from Miyakejima volcano. Fig. 2.9(a) shows five waveforms (solid lines) and their synthetic model fits (dashed lines). Using this kind of plots it is possible to control how well a specific model explains the recorded data. The residual for each fit is also reported. The synthetic waveforms plotted are generated using the model with the optimum number of wave elements selected for each signal. Fig. 2.9(b) is an example of a *f-g* diagram for the fourth waveform in Fig. 2.9(a), where the solutions for the models with $w_n = 3$ to 20 are plotted (solid circles). The solutions for the optimum number of wave elements ($n_w = 9$) are plotted using diamond symbols, with the respective error bars. Fig. 2.9(b) shows how the model with the selected optimum number of wave elements successfully samples the areas with higher density of points in the *f-g* diagram. In order to provide a further tool to control the quality of the frequency solution, a Fourier spectra is also plotted (Fig. 2.9c). In this way we did calculate *f-g* solutions and the respective optimum models for 240 VLP and LP events recorded in Miyakejima volcano from June 27 to July 08, 2000.

Following the approach described in section 2.7, optimum solutions were stacked (using windows of 50 events) in order to find the most recurrent values and then to identify the most probable solutions indicating the most active sources in a given time interval. Fig. 2.10 shows the cumulative *f-g* diagrams for six different time intervals within the time period of available data. Figs 2.7(a to d) shows the cumulative solutions (solid points) for different periods in the time interval June 27 to July 07, 2000. Contours indicate the areas with higher density of points (warmer colors). The areas with higher concentration of points in this plot indicates that the most frequent solutions for LP events in this period are in the frequency interval from 0.2 to 0.5 Hz, and growth rate from about -0.04 to -0.001 s^{-1} (S1). These values indicate Q values lower than 50. In the sequence from Figs 2.10(d to f), it is possible to see how the solutions S1 vanishes while a subset of higher-frequency solutions appears (S2). S2 solutions are in the frequency interval 1 to 3 Hz, while the growth rates range between -0.03 and -0.01 s^{-1}, which corresponds to Q values from about 50 to 300. Again in this example, it is possible to see how the characteristic complex frequencies of LP activity can be monitored to detect changes in the characteristics of the signals. The time evolution of the solutions, as filtered selecting solutions from high density areas, can be better seen in Fig. 2.11.
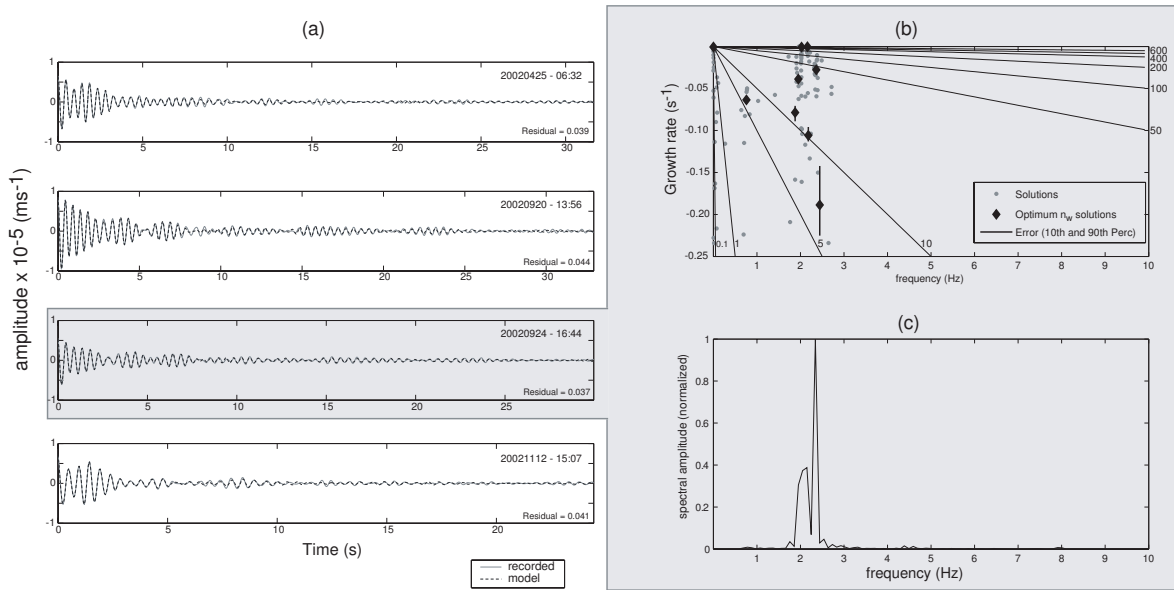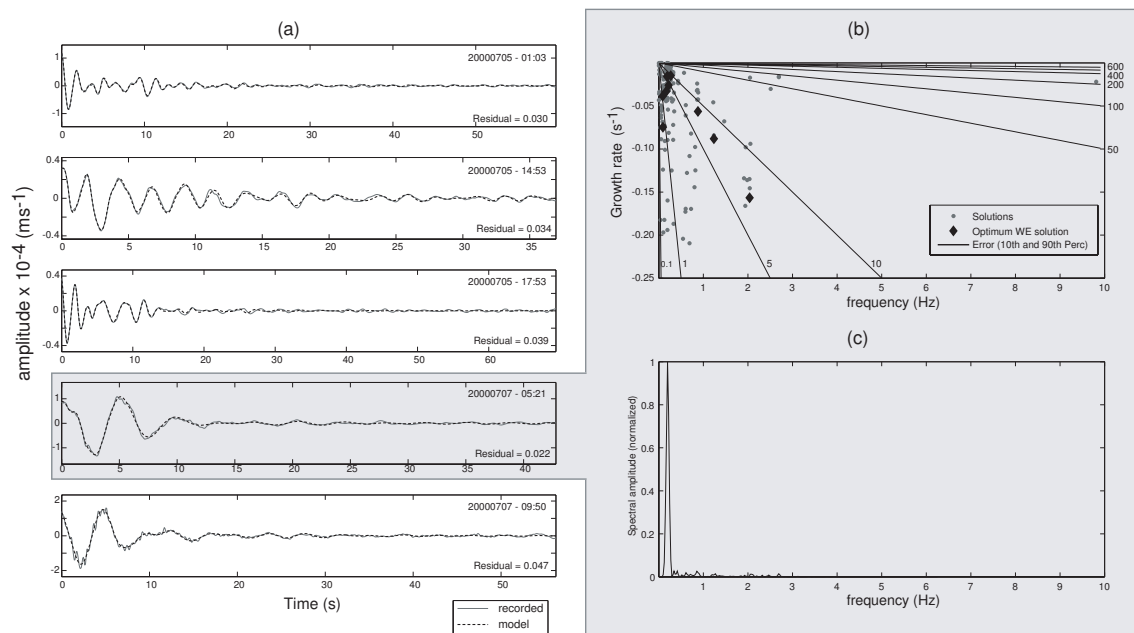
Figure 2.9: (a) Recorded waveforms (solid line) and the synthetic signal reconstructed using the optimum n_w models (dashed line), for four LP signals from Miyakejima volcano. (b) *f-g* diagram with the stacked solutions (LP event from July 07, 2000, 05h21 L.T.) for models with n_w between 3 and 20 (circles). The selected optimum n_w solutions (9) are shown with diamond marks, and error bars correspond to the 10th and 90th percentile of the distribution of the growth rate parameters. (c) Fourier spectra of the signal.
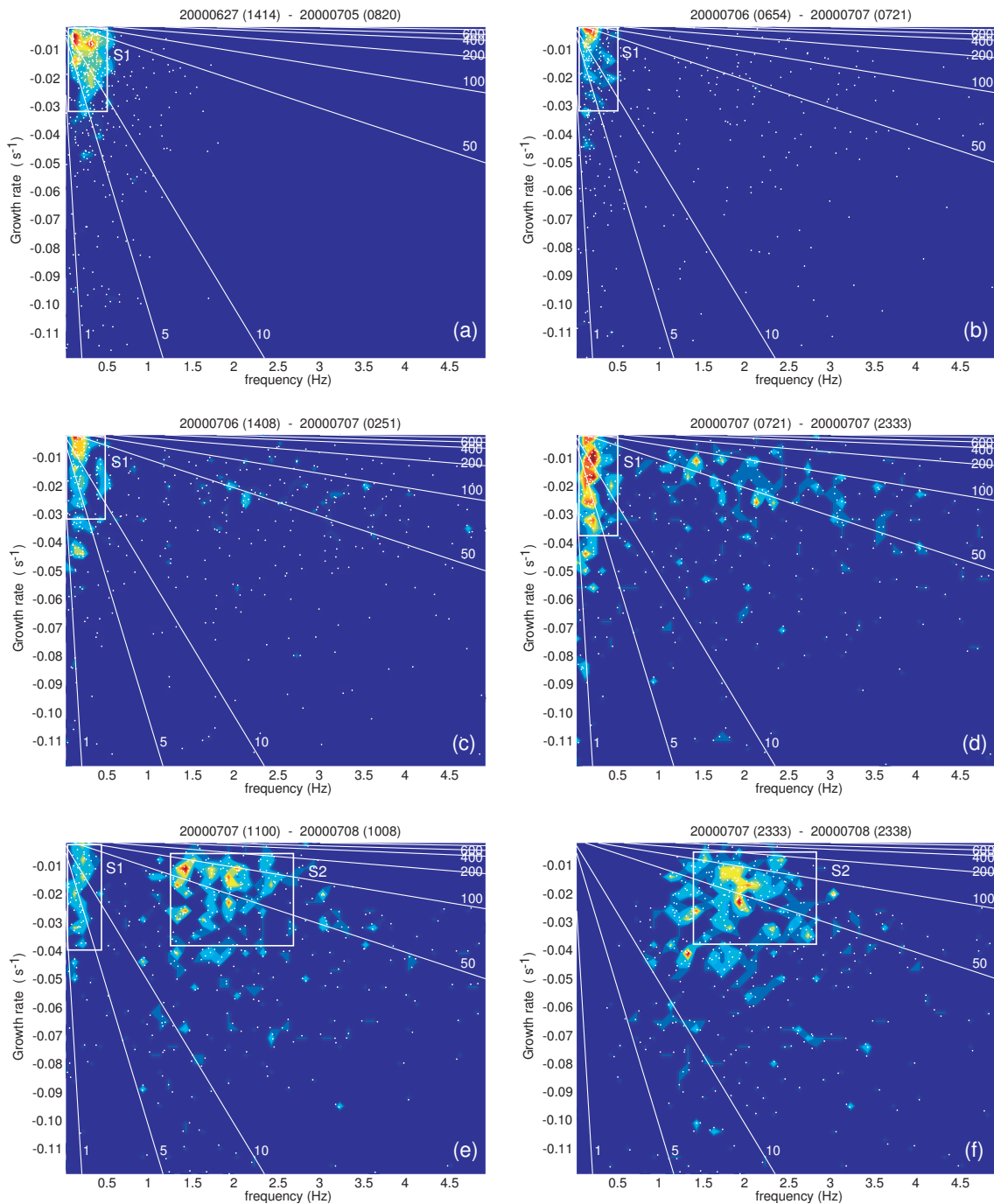
Figure 2.10: Stacked solutions of characteristic complex frequencies of LP events from Miyakejima volcano (white points) plotted in *f-g* diagrams for six different time intervals from the period June 27 - July 08, 2000. Contours enclose areas with high density of points (warmer colors). Lines in the diagram correspond with iso-values of Q.
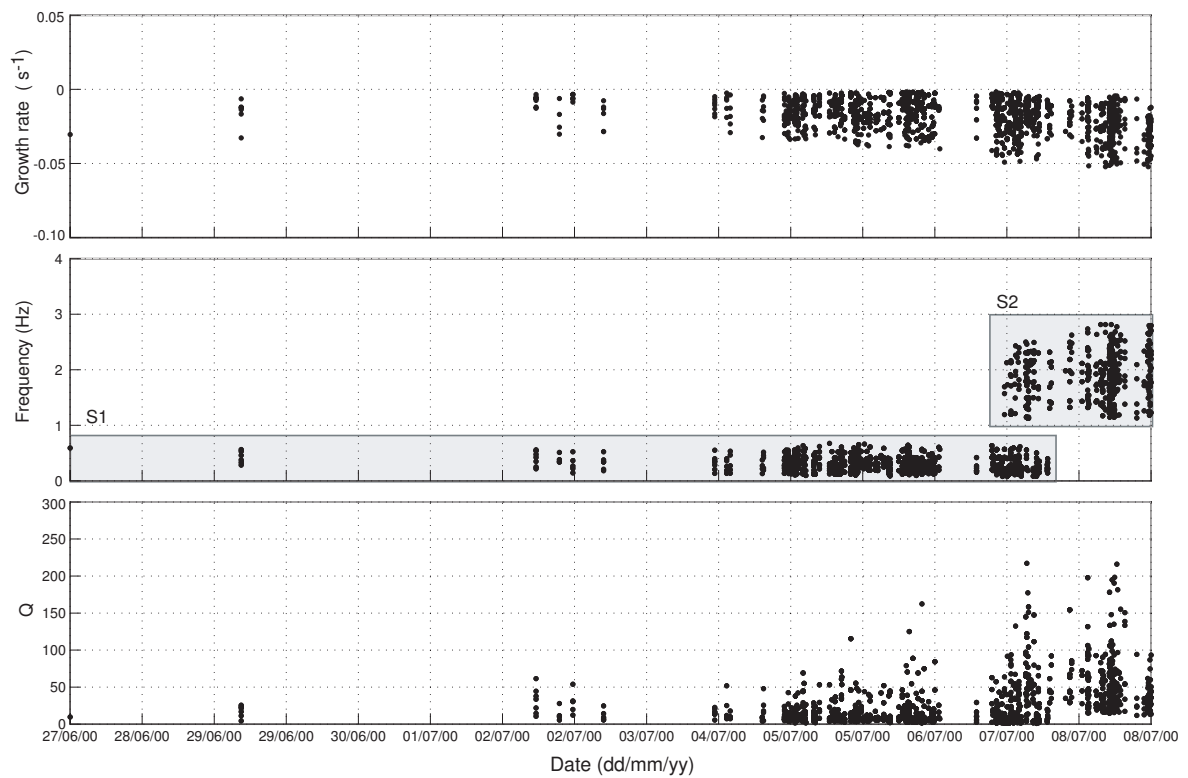
Figure 2.11: Time series of characteristic growth rate, frequency and Q parameters of LP events from Miyakejima volcano, resulting from filtering points from high density areas in *f-g* diagrams from different time intervals. Error bars correspond with 10th and 90th percentiles of distributions for each single parameter.

2.9   Discussion and concluding remarks

In this chapter we present an alternative method to estimate the parameters of a simple conceptual model presented by *Kumazawa et al.* (1990) as a method for physically-based spectral analysis. The basic idea of this method originates on the physical concept of the characteristic properties of a linear dynamic system described by a linear differential equation; for instance, the time series modeling using this approach consists essentially of estimating the governing differential equation of the hypothetical linear dynamic system that has yielded the given time series (*Kumazawa et al.*, 1990). This method takes the form of the familiar autoregressive (AR) method, however, the basic concept of the AR methods and the exact formulation based on the maximum likelihood principle have led to *Kumazawa et al.* (1990) to a model estimation algorithm different from previous AR methods.

To estimate the parameters of the model, we performed a probabilistic formulation of the problem. This kind of formulation leads to the definition of a (posterior) probability distribution in the model space $\sigma(\mathbf{m})$ which results of combining prior information (on data and model parameters) with new information obtained by measurements (the time series). When the theory linking the observable data with the parameters of the model is nonlinear, the posterior probability distribution in the model space may be a complex function difficult to describe; in this case a maximum likelihood solution may be insufficient and sometimes unreliable, and often we may not have information about the resolution power of the data.

The nature of the problem that we analyze in this chapter involves the definition of a considerable number of model parameters, so the posterior distribution is defined on a high-dimensional space; in this case an inspection of the marginal probability densities of interest may be impractical. This fact pushes us to apply Monte Carlo methods to explore $\sigma(\mathbf{m})$, which implies that a large collection of models should be pseudo-randomly generated with the support of the posterior probability distribution. To do this task we have implemented a simple Metropolis algorithm; the input to the algorithm consists of random models generated according to the prior distribution $\rho(\mathbf{m})$ and the corresponding values of the likelihood function that carries information from measurements and the theory linking observed data and model parameters; the output from the algorithm are pseudo-random realizations of the posterior distribution $\sigma(\mathbf{m})$ that contains all information about the parameterized model. From this output we can get information about model parameters: even if the most general solution would be the definition of an empirical probability distribution (CDF or PDF) for each model parameter, it may be possible (where possible) to compute any kind of central and dispersion estimator to produce both best-guess values and associated uncertainties (however, note that computing standard deviations or covariances may be meaningless if the posterior probability density is far from Gaussian, which may be the case in many nonlinear problems; in such a case it may be useful to adopt cluster analysis techniques to assess the nature of the possible solutions).

In this way we have set the problem in a fully probabilistic formulation providing a frame-

work in which restrictive assumptions are avoided and not-unique solutions are possible. We consider that this kind of formulation may be helpful for different kind of problems, specially when it exists a nonlinear relationship between the observed data and the parameters of the model; the alternative may be an uncertain parameter estimation and of course an uncertain estimation of uncertainties.

Using our formulation we have performed some numerical tests with synthetic time series, where it was possible to assess the resolution power and the reliability of the solutions (e.g., see Figs 2.1 and 2.2). An important factor in the model is the definition of the number of wave elements to properly describe the data. Unfortunately there are not physical constrains to establish the "true" number of wave elements that successfully describes the data, then an empirical approach must be used. In our approach, we propose to define an "optimum" number (or interval) of wave elements which may be determined based on a residual function that measures the misfit between the original time series and the created by the specific model; for instance, we have observed that as the number of wave elements increases, the residual value decreases with a strong slope, and after a generally sharp change point in the slope, it remains more or less constant (or may continue to decrease with a much more gentle slope, e.g., see Figs 2.1a and c). We consider that the number of wave elements associated with the change point in the slope of the residuals may indicate an "optimum" number of components to describe the data; fits with a smaller number of wave elements would produce high residuals and poor explanation of the data; on the other hand, adding more wave elements does not substantially improve the solution which results in over-fitting the observations; in this way spurious wave elements are included in the solution.

We did apply this method to analyze the waveforms of long-period events from two volcanoes: Cotopaxi (Ecuador) and Miyakejima (Japan). We did perform the analysis to each single LP event (617 for Cotopaxi, and 240 for Miyakejima) and the solutions for the optimum number of wave elements were extracted and stacked for different time intervals. This procedure may be an important tool for objectively monitor the characteristic complex frequencies of the LP events (e.g., see Figs 2.7 and 2.10); in this way, changes in the volcanic activity may be highlighted and families of events of particular interest could be identified for supplementary analysis. The information obtained using this methodology could be also used to feed probabilistic models as BET; for instance, in Chapter 4 (section 4.3), we discuss a possible strategy to integrate the information obtained from LP events using this approach into the Bayesian Event Tree model for eruption forecasting (BET_EF) of *Marzocchi et al.* (2008). Further this possible application for monitoring, our approach can also be used to analyze individual events as an alternative to the existent methodologies.

Chapter 3

Characterizing Volcano-Tectonic Seismic Swarms with non-stationary, Epidemic-Type Aftershock Sequences (ETAS) Modeling

"We have to remember that what we observe is not nature herself, but nature exposed to our method of questioning."

Werner Heisenberg
(Physics and Philosophy, 1958)

Abstract

We perform an analysis of volcano-tectonic (VT) seismic swarms using a non-stationary, epidemic-type, aftershock sequences (ETAS) model, in which two parameters of the model (the background seismicity, λ_0, and the p-value) can vary through time. Our results have evidenced some important characteristics of the VT swarms associated with magmatic intrusions, in particular, they present remarkably high p-values respect to VT swarms in cases in which no evident shallow magmatic activity has been observed, and also respect to earthquake swarms in purely tectonic environments. We consider that the stochastic modeling of VT swarms using a non-stationary ETAS model may be used for a quick analysis of VT activity in active volcanoes in order to identify and discriminate swarms which may be more likely associated with the physical response of the volcanic edifice to a magmatic intrusions.

3.1   Introduction

Nearly every recorded volcanic eruption has been preceded and/or accompanied by an increase in earthquake activity beneath or near the volcano; for this reason seismology is one of the most important and effective tools for monitoring volcanoes and can provide valuable information for short-term eruption forecasting. Active volcanoes are the source of a great variety of seismic signals, this is because volcanic seismic sources involve different physical processes whose mechanisms may include shear failure, tensile failure, and both active and passive fluid involvement (e.g., *Aki*, 1992; *McNutt*, 1996). Two basic families of processes can usually be established to classify volcanic seismicity (e.g., *Chouet*, 1996): the first family consist of volumetric sources in which the fluids play an active role in the generation of elastic waves (that in a rough generalization we call Long-Period (LP) family), and the second consists of shear or tensile sources involving brittle rock failure (that hereinafter we call Volcano-Tectonic (VT) family).

A volcanic unrest has a complex nature, involving different interactive processes; some authors have tried to identify common pre-eruptive patterns during volcanic unrest analyzing data from different worldwide volcanoes (e.g., *McNutt*, 1996; *Sandri et al.*, 2004). Following these authors, we can roughly consider that a seismic crisis during a volcanic unrest may be composed of some or all of the following components: (1) background activity; (2) swarms of VT events; (3) relative quiescence after the peak rate of seismicity; (4) swarms of LP and VT events; (5) volcanic tremor; (6) eruption; and (7) deep earthquakes following eruption. Each component may reflect a different stage during the volcanic activity, and deep understanding of the physical phenomena that they are reflecting should be the basis to feed probabilistic models (e.g., *Marzocchi et al.*, 2008) for short-term forecasting of eruptions. We are interested in exploring some characteristics of VT swarms; in general, VT swarms are often one of the earliest detectable precursors to volcanic activity, but often they are also present during the intrusive period before or accompanying the eruptive stage. This duality make them difficult to interpret and their use for forecasting purposes may lead to a high false alarm rate. For this reason it is necessary to improve our knowledge about the characteristics of VT swarms in order to better characterize their origin and to produce better interpretations.

Most VT events are thought to be caused by shear failure or slip on faults and fractures, and differ from purely tectonic activity mainly in their patterns of occurrence, which, at volcanoes, are typically in swarms rather than a clear mainshock-aftershock sequence. For instance, we can define a seismic swarm as a group of many earthquakes of similar magnitude characterized by the lack of clear mainshock signatures. In general, seismic swarms feature complex spatiotemporal patterns probably linked to the contribution of different sources that in general may be related to the presence of magma/fluids and to seismic interactions (*Lombardi et al.*, 2006).

Some conceptual models have been proposed to explain the generation of earthquake swarms in volcanic environments; *Hill* (1977) proposed a model in which VT earthquakes occur on

shear planes connecting the edges of offset and inflating dikes; *Ukawa and Tsukahara* (1996) proposed a model in which VT earthquakes occur in a zone of inflation-induced tension ahead of the tip of a propagating dike; *Toda et al.* (2002) described the spatiotemporal evolution of a swarm at the Izu Islands through the co-seismic stress variations induced by a constantly growing dike emplaced at the beginning of the swarm; *Roman* (2005) proposed that VT earthquakes occur in the walls (away from the tips and edges) of a dike inflating in the direction of regional minimum compressive stress. Some characteristics of those models are confirmed by some analytical and numerical models (e.g., *Bonafede and Danesi*, 1997).

All those models provide valid descriptions of physical processes that can generate VT events during dike intrusions, starting from the basis of an ongoing intrusion. However, we want to go one step behind and start from the question, given that we have a generic swarm of earthquakes in a given volcanic region, is it directly associated with dike migration processes or is it driven by other (e.g. tectonic) factors? to answer this question we explore some characteristics of seismic swarms based on a non-stationary epidemic-type, aftershock sequences (ETAS) modeling.

ETAS (*Ogata*, 1988) is a stochastic model that describes the short-time clustering features of earthquakes as a superposition of modified Omori functions (*Utsu*, 1961) shifted in time. At its origins, ETAS was applied to analyze typically tectonic mainshock-aftershock sequences, but a generalization of this model that allows time variations of one or more parameters proposed by *Hainzl and Ogata* (2005) and *Lombardi et al.* (2006), has opened a new window to characterize complex seismic swarms.

In this chapter we apply a general stochastic non-stationary ETAS modeling to analyze complex seismic swarms in both tectonic and volcanic environments, in order to extract information useful to characterize the swarms directly associated with dike migration. The final goal is to get insights for a quantitative tool useful to interpret VT swarms during monitoring procedures, yielding physical constrains of the driving processes.

3.2 Stochastic modeling of a seismic swarm

The epidemic-type aftershock sequences (ETAS) model is a stochastic point process in which each earthquake has some magnitude-dependent ability to trigger its own Omori law type aftershocks (*Ogata*, 1988; *Ogata et al.*, 1993; *Ogata*, 1998). The total occurrence rate can be described, in time, as the superposition of a background uncorrelated seismicity λ_0 and the events triggered by another earthquake:

$$\lambda(t) = \lambda_0 + \sum_{i:t_i<t} \lambda_i(t) \tag{3.1}$$

where $\lambda_i(t)$ is the rate of aftershocks induced by an event occurred at time t_i with magnitude M_i, defined as (for details see *Ogata*, 1988)

$$\lambda_i(t) = \frac{\kappa}{(c + t - t_i)^p} e^{\alpha(M_i - M_c)} \tag{3.2}$$

for $t > t_i$. The parameter κ measures the productivity of the aftershock activity; α defines the relation between triggering capability and magnitude m_i of a triggering event; c measures the incompleteness of the catalog in the earliest part of each cluster; the parameter p controls the temporal decay of triggered events; M_c is the completeness magnitude of the catalog.

We use a generalized version of the ETAS model by considering a non-stationary behavior (understood as a process/model whose parameters vary through time) of some model parameters. In particular, we consider the time variations of λ_0 and p ($\lambda_0(t_j)$ and $p(t_j)$, respectively, for the j-th time window) as suggested by *Lombardi et al.* (2006), but using overlapping windows. Those parameters are thought to be the most clearly linked to variations of the physical system due to dike intrusions; specifically, time variations in the $\lambda_0(t_j)$ parameter have been interpreted as indicator of seismicity associated with fluid flow which initiate seismicity that cannot be explained without considering complex patterns resulting from both pore pressure variations and earthquake-connected stress field changes (e.g., *Hainzl and Ogata*, 2005; *Lombardi et al.*, 2006); on the other hand, the $p(t_j)$ values have been found to be positively correlated with crustal temperature, which controls stress release and therefore aftershocks decay (e.g., *Mogi*, 1967; *Kisslinger and Jones*, 1991; *Lombardi et al.*, 2006).

To explore possible characteristic properties of VT swarms during dike intrusions, we analyze different earthquake swarms selected from three general groups: (1) swarms from purely tectonic environments, (2) swarms occurred in volcanic areas but that did not preceded nor accompany eruptive activity , and (3) swarms in volcanic areas during an eruptive process.

The estimation of model parameters $\{\kappa,\ c,\ \alpha,\ \lambda_0(t_j),\ p(t_j)\}$ is carried out by the Maximum Likelihood method (*Dalay and Vere-Jones*, 1988), and we use the Davidon-Fletcher-Powell optimization procedure (*Fletcher and Powell*, 1963), which provides also a numerical approximation of errors (*Lombardi et al.*, 2006). We fit the ETAS model in a moving overlapping time window of length τ (in days), the length of which is selected based on two basic requirements: (1) the need to include enough data for the calculations, and (2) to have the minimum possible time window to follow the details of the time evolution of the process. The time window is moved forward in time at steps of 1 day. At each time interval j we estimate the updated value of the parameters allowed to change with time (i.e. $\lambda_0(t_j)$ and $p(t_j)$), setting all the other parameters to the values estimated for the whole sequence. In computing the model parameters for each time window we take in account all past occurrence history, to include the possibility that an event has been triggered by an earthquake occurred at a time before the start time of the considered j-th window.

3.3   Data

We analyze different earthquake swarms selected from three general groups as: (1) VT swarms in volcanic areas before/during an eruptive process; (2) VT swarms occurred in volcanic areas where no eruption happened, and (3) Earthquake swarms from purely tectonic environments.

3.3.1   Sequences from erupting volcanic areas

In this group we have selected data from three volcanic regions: Guagua Pichincha (Ecuador), Miyakejima (Japan), and Pinatubo (Philippines), which have experienced VT seismic swarms preceding and/or accompanying volcanic eruptions (see table 3.1).

(1) Guagua Pichincha volcano is a dacitic stratovolcano located 12 km west of Quito, Ecuador. Holocene eruption styles at Guagua Pichincha have been dominantly explosive and include phreatic explosions, dome growth and collapse, Vulcanian eruptions, Plinian ash falls, and pyroclastic flows (e.g., *Hall*, 1977; *Barberi et al.*, 1992). The most recent unrest period began in 1981 with small seismic swarms and episodes of phreatic activity that included small phreatic explosions and gas emissions. A second phase of activity initiated on August 1998 with the onset of larger phreatic explosions preceded and accompanied by volcanic tremor and swarms of LP and VT seismic events. Intense VT and LP swarms occurred before and during the emplacement of dacitic lava domes within the calderic crater (*Garcia-Aristizabal et al.*, 2007). The database, provided by the Instituto Geofísico - Escuela Politécnica Nacional (Quito, Ecuador), consists of more than 650 VT events recorded from January 1998 to November 1999; the catalog is considered complete for magnitudes larger than $M1.6$.

(2) Miyakejima is an active basaltic volcano located in the northern Izu islands, Japan, that produced a noteworthy eruption in 2000 after an interval of 2,500 years without erupting a large quantity of magma (e.g., *Nakada et al.*, 2005; *Amma-Miyasaka et al.*, 2005). During the initial stage of the eruption, magma moved in a dike system westward from the volcano (e.g. *Fujita and Ukawa*, 2002) and 12 days after the intrusion, the summit of the volcano began to collapse to from a new caldera (e.g., *Nakada et al.*, 2005; *Amma-Miyasaka et al.*, 2005). The volcanic unrest started on 26 June 2000 (around 18:00 local time) with an intense seismic swarm. The earthquakes initially occurred beneath the SW flank of the volcano, but at around 21:30 they jumped to the W flank and continued an initially westward and, later, northwestward migration (*Uhira et al.*, 2005). Epicenters moved progressively away from the island, and a submarine flank eruption occurred near the west coast of Miyakejima next morning (e.g., *Uhira et al.*, 2005). The hypocenters continued to migrate northwestward and intense earthquake swarms occurred intermittently between Miyakejima and Kozushima until September 2000, including five $M6.0$ earthquakes. New swarms occurred beneath the volcanic edifice on 4 July, which have been interpreted as caused by instability of the volcanic edifice, and associated with the caldera formation process (*Uhira et al.*, 2005). We analyzed a data set provided by the National Research Institute for Earth Science and Disaster Preven-

tion, Japan, and from *Toda et al.* (2002), which consists of more than 5000 events recorded from June to December 2000; the catalog is considered complete for magnitudes larger than $M3.0$ (*Lombardi et al.*, 2006).

(3) Mount Pinatubo is a part of the Luzon arc, whose volcanism is related to the activity of the Manila trench, located about 120 km West of the volcano. The NW-trending, left-lateral Philippine fault passes NE of Mount Pinatubo, and on July, 16, 1990, a 125-km-long segment of this fault ruptured and produced a $M_s7.8$ earthquake. The epicenter of this earthquake was about 100 km NE of Mount Pinatubo (*Bautista et al.*, 1996). The cataclysmic eruption of Mount Pinatubo on June 15, 1991, was preceded by at least 10 weeks of unrest character- ized by increasing seismic activity (a complex sequence comprising VT earthquakes and LP events), high SO_2 emissions, emplacement of a lava dome, and numerous smaller explosive eruptions (e.g., *Power et al.*, 1996). We analyzed a data set distributed as annexed material in *Newhall and Punongbayan* (1996), which consists of more than 380 events recorded from May to June, 1991 (before the cataclysmic eruption on June 15), and more than 900 events recorded from June to August, 1991. The catalogs are considered complete for magnitudes larger than $M0.6$ for the period before the main eruption, and larger than $M0.8$ for for the period post cataclysmic eruption.

Table 3.1: Summary of the seismic sequences analyzed in this work: (1) VT sequences from erupting volcanoes, (2) VT sequences from non-erupting volcanoes, and (3) earthquake se- quences from tectonic environments.

Sequence	Period of time	Completeness Magnitude	τ (days)
1. Erupting volcanoes			
Guagua Pichincha	1998-1999	1.6	20
Miyakejima	June-December 2000	3.0	5
Pinatubo - before	May-June 1991	0.6	2
Pinatubo - after	June-August 1991	0.8	2
2. Non-erupting volcanoes			
Campi Flegrei	1982-1984	0.4	5
Cotopaxi	1996-2007	3.2	10
Mt. Fuji	1997-2006	0.2	150
Long Valley	1997-1998	1.5	10
3. Tectonic			
Landers	1992-1993	3.0	5
Umbria-Marche	1997-1998	2.5	10

3.3.2 Sequences from non-erupting volcanic areas

In this group we have selected data from four volcanic regions: Campi Flegrei (Italy), Cotopaxi (Ecuador), Mt. Fuji (Japan), and Long Valley caldera (California), which have experienced VT seismic activity that not preceded, accompanied or concluded with a volcanic eruption (see table 3.1). (1) At Campi Flegrei, the 1982-84 unrest episode is the most significant activity recorded in recent times, and the recorded seismicity has been widely studied in literature (e.g., *De Natale et al.*, 1984; *De Natale and Zollo*, 1986; *De Natale et al.*, 1995; *Troise et al.*, 1997; *Orsi et al.*, 1999; *Marzocchi et al.*, 2001; *Troise et al.*, 2003). The seismic database consists of more than 13,000 events that were recorded during a seismic crisis occurred from January 1983 to December 1984 (*De Natale and Zollo*, 1986; *Marzocchi et al.*, 2001); the catalog is considered complete for magnitudes larger than 0.4 (*Marzocchi et al.*, 2001). (2) At Cotopaxi, we have considered the VT seismic swarms occurred between 2001 and 2005 during the most important seismic unrest occurred since the volcano is instrumentally monitored. Renewed seismic activity of Cotopaxi began in January 2001 with an increased number of LP events which were followed by a swarm of VT earthquakes in November 2001 (*Molina et al.*, 2008). After that, the rate of VT events showed high levels (respect to the background activity) for about three years. The database, provided by the Instituto Geofísico - Escuela Politécnica Nacional (Quito, Ecuador), consists of more than 1800 events recorded from January 1996 to December 2007; the catalog is considered complete for magnitudes larger than $M3.2$.

(3) Mount Fuji is the largest active stratovolcano in Japan, and in recent times its seismic activity has been relatively low; the most significant seismic activity occurred in 2001-2002 which was mainly characterized by swarm-like mid-crustal low-frequency (DLF) earthquakes (e.g., *Ukawa*, 2005). The DLF events have been located at 2-4 km NE of the summit and in the depth range 10-20 km, and in the active period of DLF events, tectonic-like of mainly strike slip mechanisms earthquake activity in the Mt. Fuji area increased slightly (*Ukawa*, 2005). We extract the information of high-frequency seismicity around Mt. Fuji from a database provided by the National Research Institute for Earth Science and Disaster Prevention, Japan (*Ukawa*, 2005), which consists of more than 340 events recorded from January 1997 to December 2006; the catalog is considered complete for magnitudes larger than $M0.2$. (4) For the Long Valley caldera, we considered 9111 seismic events occurred in the July 1997 - March 1998 swarm, which is one of the most energetic seismic swarm in recent years (*Hill*, 2006). After nearly a year of relative quiescence within the caldera, seismicity gradually increased in mid 1997 with earthquake-swarm activity concentrated at depths between 3 and 8 km beneath a broad 15-km-long zone, and it included more than 12000 $M > 1.2$, 120 $M > 3.0$, and 8 $M > 4.0$ recorded earthquakes. The 1997 seismic-moment release was dominated by right-lateral slip along the WNW-trending fault zone beneath the south moat (*Hill*, 2006). The catalog is considered complete for magnitudes larger than $M1.5$ (*Marzocchi et al.*, 2001).

3.3.3   Sequences from purely tectonic environments

We have selected two well known seismic sequences occurred in tectonic areas: The 1992-1993 Landers (California) sequence, and the 1997-1998 Umbria-Marche (Italy) sequence. (1) The Landers earthquake sequence began on April 23, 1992 with a $M_W 6.1$ event (Joshua Tree preshock); this sequence ruptured almost 100 km of both superficial and concealed faults and caused aftershocks over an area 100 km wide by 180 km long (*Hauksson et al.*, 1993). The $M_W 7.3$ Landers mainshock occurred on June 28, 1992, its focal mechanism showed a right-lateral strike-slip faulting, and was preceded for 12 hours by 25 small $M \leq 3$ earthquakes at the mainshock epicenter (*Hauksson et al.*, 1993); mapping showed that the rupture occurred on five overlapping, curved fault segments with the cumulative length of 85 km (*Sieh et al.*, 1993); Numerous $M > 4$ aftershocks occurred near the mainshock rupture zone. We analyzed the data set of *Hauksson and Shearer* (2005) which consists of more than 1800 events recorded from April 1992 to December 1993 (see table 3.1); the catalog is considered complete for magnitudes larger than $M 3.2$. (2) The Umbria-Marche is a long seismic sequence which started on September 3, 1997 with a $M_W 4.5$ foreshock followed by several small events in the following two weeks. The three largest shocks occurred on September 26 close to Colfiorito ($M_W 5.7$ and 6.0) and on October 14 a $M_W 5.6$ in the Sellano area (e.g., *Murru et al.*, 2004). The seismogenic structure consist of a NW-SE elongated fault zone extending for about 40 km and the focal mechanism of the largest shock reveal normal faulting with NE-SW extension perpendicular to the trend of the Apennines, consistently with the Quaternary tectonic setting revealed by previous earthquakes in adjacent regions (e.g., *Amato et al.*, 1998); for instance, stress tensor solutions obtained for the six main shock focal mechanisms of the sequence are in agreement with the tectonic stress active in the inner chain of the Apennines, and also the aftershock focal mechanisms show a consistent extensional kinematics (e.g., *Chiaraluce et al.*, 2003). We analyzed a data set from the CSI1.1 (Catalogo della Sismicità Italiana, *Castello et al.*, 2005, 2007) which consists of more than 800 events recorded from May 1997 to December 1998; the catalog is considered complete for magnitudes larger than $M 2.5$.

3.4   Results

3.4.1   VT swarms in erupting volcanoes

We fit the non-stationary ETAS model with both $\lambda_0(t_j)$ and $p(t_j)$ varying through time for all the cases described in section 3.3. In this section we analyze some VT swarms that preceded and/or accompanied eruptions in 3 different volcanic areas: Guagua Pichincha, Miyakejima, and Pinatubo.

In Fig. 3.1 we report the time evolution of λ_0 (3.1a), p (3.1b) and the number of events/window (3.1c, $\tau = 20$ days), for the seismic activity at Guagua Pichincha volcano, from February to September 1999. This period includes 3 seismic swarms: the first one from February to March 1999, characterized by a peak in the background seismicity and p-values around 1.3 ($\pm 10\%$)

which tends to drop at the end of the swarm; the second swarm is characterized by a peak in the background seismicity which tends to decrease with time, as the p values show an opposite trend, with the lower values at the beginning of the swarm (but with big uncertainties in the parameter determination), growing up to a value around 1.5 ($\pm 6\%$) at the end of May and during the first weeks of June 1999. This swarm was characterized by fewer events and larger magnitudes ($M > 3.0$) than those in the first swarm in February-March (*Garcia-Aristizabal et al.*, 2007). This swarm was also accompanied by ash-rich explosions, which at the moment was an important observation if we consider that in the months before the activity at the crater was characterized by vapor-rich phreatic emissions. This may indicate that an important perturbation of the volcanic system was going on. The third swarm (August-September 1999) is characterized by a clear increase in the background seismicity (Fig. 3.1a) up to a maximum (about two times respect to the swarms before) at the end of August, which was followed by a trend toward lower values until the end of the swarm. On the other hand, the p-values at the beginning of the swarm range around 1.5 ($\pm 5\%$) which is about the same value as the p-values at the end of the second swarm on June. In September 1999 the most intense VT seismicity took place; this period was characterized by the highest p-values measured during the whole unrest period, reaching values of 1.7 ($\pm 2\%$) at the end of the VT swarm. The maximum of the VT seismicity rate occurred on September 23, and on September 25 the first intense LP swarm occurred in association with the intrusion of the first dacitic lava dome; for instance, following an explosion containing juvenile magma on September 26, the first dome grew into the crater, the LP swarm almost ceased on September 30, and the first dome was destroyed by Vulcanian-type magmatic explosions that occurred on October 5 and 7, 1999 (*Garcia-Aristizabal et al.*, 2007).

In Fig. 3.2 we report the time evolution of λ_0 (3.2a), p (3.2b) and the number of events/window (3.2c, $\tau = 5$ days), for the seismic activity at Miyakejima volcano, in the period June - August 2000. The non-stationary ETAS modeling was applied by *Lombardi et al.* (2006) to this swarm and somehow this chapter may be considered as an extension of that work. Our results show a trend similar to the trend found by *Lombardi et al.* (2006); for instance, The background activity (λ_0) shows two important peaks: the first one at the beginning of the sequence (around June 26 - July 4) and the second about a week later; these two peaks of λ_0 reach values remarkably high respect to the cases described before. Two smaller peaks can be distinguished at the end of July and beginning August. On the other hand, p-values for this sequence show values around 2.4 ($\pm 1\%$) without evident changes up to July 29 (which corresponds with a change point identified by *Lombardi et al.* (2006) using a two-sample Kolmogorov-Smirnov test (*Mulargia and Tinti*, 1985; *Mulargia et al.*, 1987)); after that, the p-value slightly decreases and shows some long-period oscillations up to the end of the sequence.

In summary, the 2000 Miyakejima sequence is characterized by wide oscillations of λ_0 at the beginning of the sequence, accompanied by a rather stable and remarkably high value of p.

Figure 3.1: Time evolution of the non-stationary parameters of the ETAS model: Background activity $\lambda_0(t_j)$ (a), p-value (b), and number of events/window (c; $\tau = 20$ days), during the 1999 unrest and eruption of Guagua Pichincha volcano, Ecuador; grey areas show $\pm 1\sigma$ error interval in all plots.

Figure 3.2: Time evolution of the non-stationary parameters of the ETAS model: Background activity $\lambda_0(t_j)$ (a), p-value (b), and number of events/window (c; $\tau = 5$ days), during the 2000 unrest and eruption of Miyakejima volcano, Japan; grey areas show $\pm 1\sigma$ error interval in all plots.

The first and most important peak (in number of events/day) in λ_0 corresponds with the intrusive phase (June 26 - July 8) defined in *Nakada et al.* (2005). The intrusive phase was characterized by a high seismicity rate around Miyakejima; tilt changes suggest that horizontal propagation was dominant in magma migration at this phase (e.g., *Fujita et al.*, 2004; *Ueda et al.*, 2005). The second peak in λ_0 (July 8 to 15) corresponds with the start of the summit subsidence stage. The collapse of the summit area began suddenly on the evening of July 8 accompanied by the first (and rather small) explosion. During the subsidence, two phreatomagmatic explosions occurred on July 14 and 15, which were preceded by a slight inflation of the volcano, as observed in GPS data (*Nakada et al.*, 2005). During the summit subsidence stage, the p-value was at the same levels (about 2.4) as during the intrusive stage. Subsidence of the summit area continued through early August, and after August 10 the main explosive phase started and continued intermittently through August 29 (*Nakada et al.*, 2005). The main explosive phase was characterized by an evident reduction in λ_0 and oscillations in the p-values, which in general show a decreasing trend.

Some authors (e.g., *Nakada et al.*, 2005) suggest that phreatic explosions of middle July may have been triggered by interaction of magma with hydrothermal fluids. In the model proposed, roof rocks gravitationally unstable above the foamy magma collapsed into the reservoir as magma migrates through dikes, generating caldera subsidence. Cavities formed became reservoirs of hydrothermal fluids, into which magmatic gases were absorbed (*Nakada et al.*, 2005). Then, magma ascent was accelerated after the end of the caldera subsidence and the most explosive magmatic event occurred on August 18; a cool pyroclastic surge was generated on August 29, probably due to boiling of hydrothermal fluids sealed inside the vent after August 18. The explosive stage may imply a significant stress release in the volcanic edifice, and phreatomagmatic explosions may be an efficient mechanism of heat transport to cool down the system; those processes may be coherent respectively with both the decrease in the λ_0 values, and the general decreasing trend of p-values during this phase.

Fig. 3.3 shows the results for the data set of Mt. Pinatubo volcano. Time evolution of λ_0 (3.3a and d), p (3.3b and e) and the number of events/window (3.3c and f, $\tau = 2$ days) are plotted for both a period (3 to 10 June) before the cataclysmic eruption on June 15, 1991 (a, b, and c) and also for a period from June 29 to August 17, 1991, after that eruptive phase (d,e, and f). Even if the analyzed data set of VT activity before the main eruptive phase contains information from about a week (up to the 10th of June), the volcano was seriously restless for at least 2 months before (instrumental monitoring started on April 5, for details see *Harlow et al.* (1996)). For instance, Between April and May VT swarms (many events felt by local population) took place 5 km NW of the summit and 5 km deep (e.g. *Newhall and Punongbayan*, 1995). Some authors divide the period April 5 to June 15, prior to the climactic eruption, into five phases based on the level, type, character, and intensity of the seismic activity (e.g., *Harlow et al.*, 1996); the precursory phases are (1) Meta-stable: through May 31; (2) Pre-dome: June 1 to 7, (3) Pre-explosive Buildup: June 8 to 12, (4) Long-Period Buildup: June 12 to 14, and (5) Pre-climactic: June 14 to 15 (for details see, e.g., *Harlow*

Figure 3.3: Time evolution of the non-stationary parameters of the ETAS model: Background activity $\lambda_0(t_j)$ (a and d), p-value (b and e), and number of events/window (c and f; $\tau = 2$ days), before (a, b, c) and after (d,e,f) the cataclysmic eruption of Mt. Pinatubo, Philippines, on June 15, 1991; grey areas show $\pm 1\sigma$ error interval in all plots.

et al., 1996). It means that the pre-eruption catalog that we analyze here (Fig. 3.3a, b, and c) corresponds with the predome period. The time evolution of λ_0 shows a peak around June 4-5, while the p-value shows a rather stable values around 2.1 to 2.2 ($\pm 1\%$). Distinct changes in seismic activity began in early June that included an increase in the number of localized VT earthquakes beneath the active fumaroles, an increase in small explosions, and an increase in the intensity and durations of episodes of tremor (*Harlow et al.*, 1996). Seismic activity during this phase eventually evolved into an intense swarm of shallow VT earthquakes that heralded the beginning of dome growth on the northwest flank of the volcano (*Hoblitt et al.*, 1996).

The results for the data after the main eruptive period (June 29 - August 17, 1991) show a still high rate of background seismicity (Fig. 3.3d), but p-values have decreased respect to the pre-eruptive period: for instance, at the end of June the p-value was around 1.4 ($\pm 5\%$) and up to the end of the period it showed a decreasing trend. In summary, the most remarkable observation of these results is that even if the background seismicity shows still high rates (i.e. comparable with the activity before the main eruptive period), the p-values show an evident drop respect to the values found during the predome phase.

Figure 3.4: Time evolution of the non-stationary parameters of the ETAS model: Background activity $\lambda_0(t_j)$ (top panels), p-value (central panels), and number of events/window (bottom panels), during the 1997 Umbria-Marche (a, b, and c; $\tau = 10$ days), and 1992 Landers (d, e, and f; $\tau = 5$ days) seismic sequences; grey areas show $\pm 1\sigma$ error interval in all plots.

3.4.2   Tectonic sequences and VT swarms in non-erupting volcanoes

In this section we analyze both the tectonic sequences and the VT swarms occurred in non-erupting volcanoes. In Fig. 3.4 we report the time evolution of the model parameters for the seismic sequences of tectonic origin considered. The time evolution of λ_0, p and the seismicity rate (number of events per time window τ) for the Umbria-Marche sequence are reported in Fig. 3.4a, b and c, respectively, while results for the Landers sequence are reported, in the same order, in Fig. 3.4d, e, and f.

Both sequences have some similar features; for instance, an increase in the background seismicity (λ_0) is observed at the same time or just after the mainshock event of the sequence; then, it decreases gently showing some oscillations which in general are contained within the error intervals (see Fig. 3.4a and d). Note that the increase in λ_0 observed for the Landers sequence is less significant if we consider the error in the parameter estimation. On the other hand, the p parameter does not show any evident variation neither at the beginning of the sequence nor during the development of the aftershocks. Another remarkable feature is the value of p of around 1.2 ($\pm 4\%$) in both sequences for the period around the mainshock (June 28, 1992 for Landers, and September 26, 1997 for the Umbria-Marche, see Fig. 3.4b and e, respectively). Hereinafter, uncertainties in parameter estimations are relative errors equivalent to $\pm 1\sigma$.

During few weeks after the mainshocks, low frequency oscillations in the p-value are observed, but those variations are accompanied by larger error intervals suggesting that they could be probably caused by instabilities in the parameter estimation.

Now we analyze the VT swarms occurred on volcanic areas in periods in which no eruptive activity was observed. In most cases, the origin of this kind of unrest activity remains unknown or not well understood; for instance, their origin may range from deep magmatic activity that succeed in disturbing the shallower volcanic system, or it may just be the evidence of mechanical response of the volcanic edifice to other (e.g. tectonic, gravity, etc.) factors. What they have in common is that neither during nor after the VT activity the volcano experienced an eruption. The results for the Campi Flegrei unrest are plotted in Fig. 3.5. The temporal behavior of the background activity (λ_0, Fig. 3.5a) shows variations both at high frequencies (i.e. in quite short times), and at low frequencies (i.e. its overall trend has a bell shape, that means a smooth onset as well as a smooth end of the energy release). The p-value (Fig. 3.5b), despite it has small oscillations close to 1, is fairly constant all over the crisis (there is no any clear trend). The lower values at the beginning of the swarm are affected by strong errors due to the small number of events.

The results for the Cotopaxi unrest are summarized in Fig. 3.6. The most intense VT activity in this data set took place from November 2001 to January 2002, and at least two peaks on λ_0 can be identified within this time interval: one from November to mid-December 2001, and the second during the first week of January 2002. At the beginning of the crisis, the p-value is around 1 ($\pm 5\%$), then it reach a value around 1.2 ($\pm 5\%$) at the end of the first peak of λ_0. The second peak of λ_0 (January 2002) is accompanied by a decrease in the p-value, as the seismicity rate also decrease to lower levels. The VT activity in the following two months was characterized by some low-frequency oscillations in the temporal behavior of λ_0, but reaching peaks with lower amplitude respect to those observed at the beginning of the crisis. At the same time, the p-value showed small variations around 1. In summary, the period of most intense VT activity at Cotopaxi is more correlated to increases in the background activity rather than evident changes in the p-value; this period of intense VT swarms occurred about 10 months after the start of the volcanic unrest (which was mainly characterized by high rates of LP activity) and preceded a period of important LP and VLP activity started about 5 months latter and never observed before during the unrest (e.g., *Molina et al.*, 2008).

Fig. 3.7 shows the results for the data set from Mount Fuji volcano. VT activity in Mt. Fuji is remarkably low, reaching few events a day in the period of most intense activity (around February 2001). We explore the behavior of ETAS parameters for this period and find a remarkable stability in both λ_0 (Fig. 3.7a) and p (Fig. 3.7b) parameters. Note that this stability may be influenced by the long temporal window used for the parameter estimation. In con-
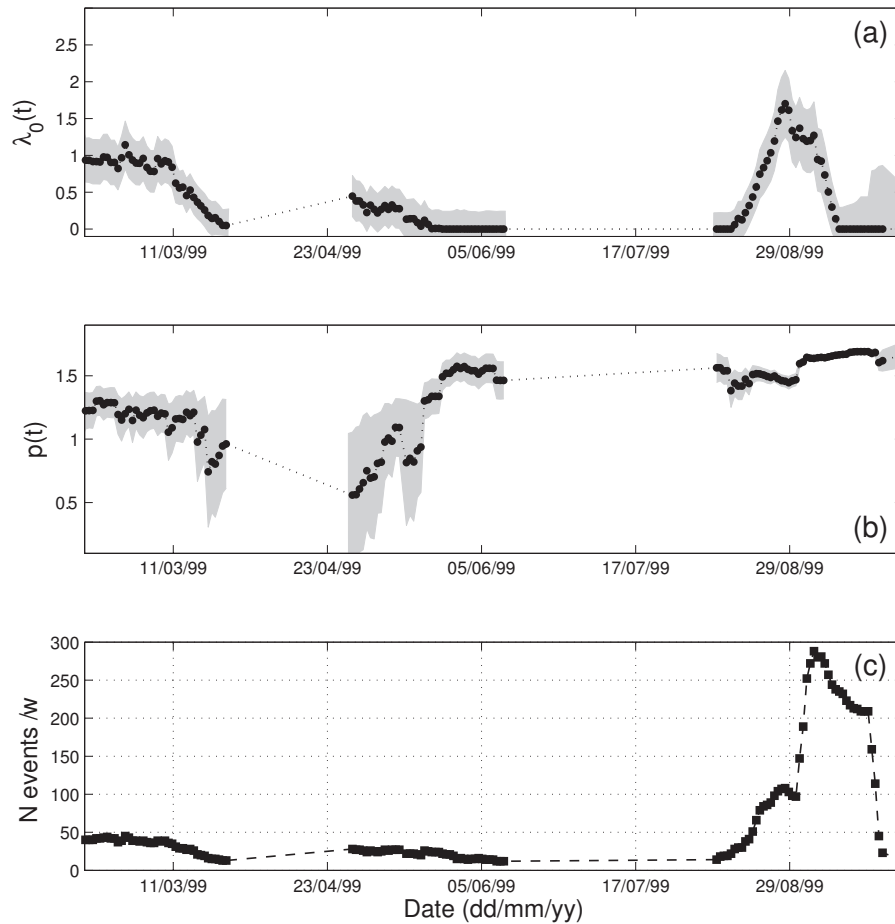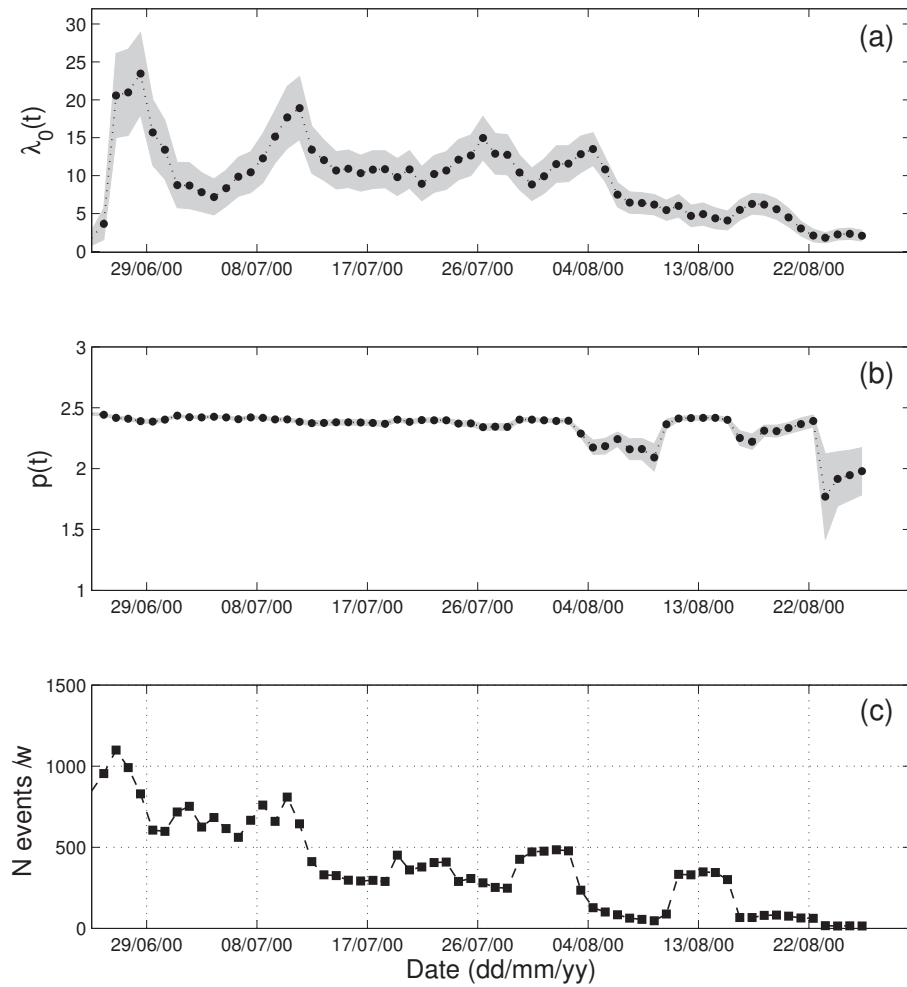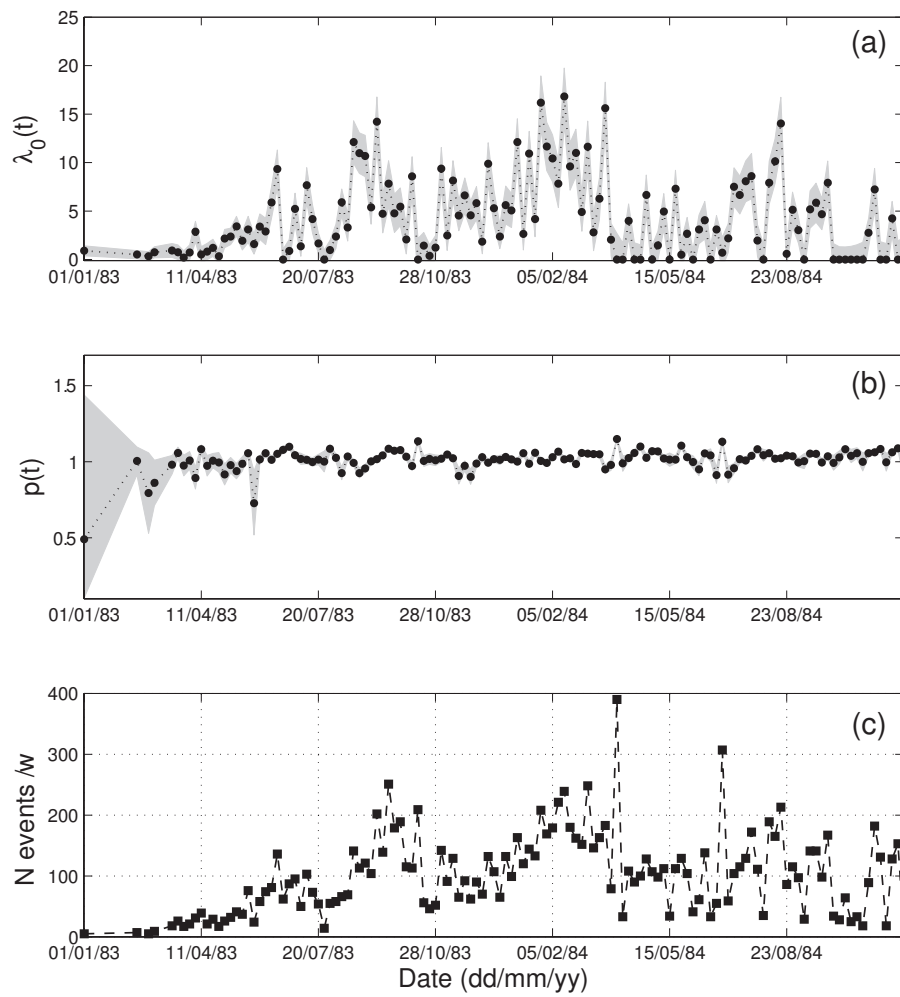
Figure 3.5: Time evolution of the non-stationary parameters of the ETAS model: Background activity $\lambda_0(t_j)$ (a), p-value (b), and number of events/window (c; $\tau = 5$ days), during the 1983-1984 unrest at Campi Flegrei caldera, Italy; grey areas show $\pm 1\sigma$ error interval in all plots.
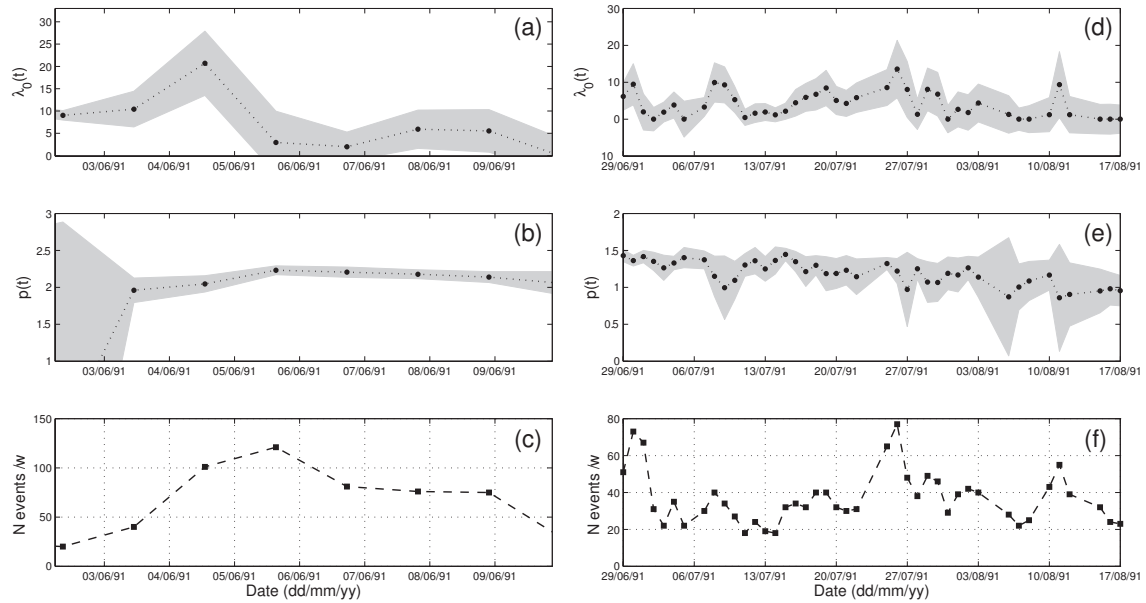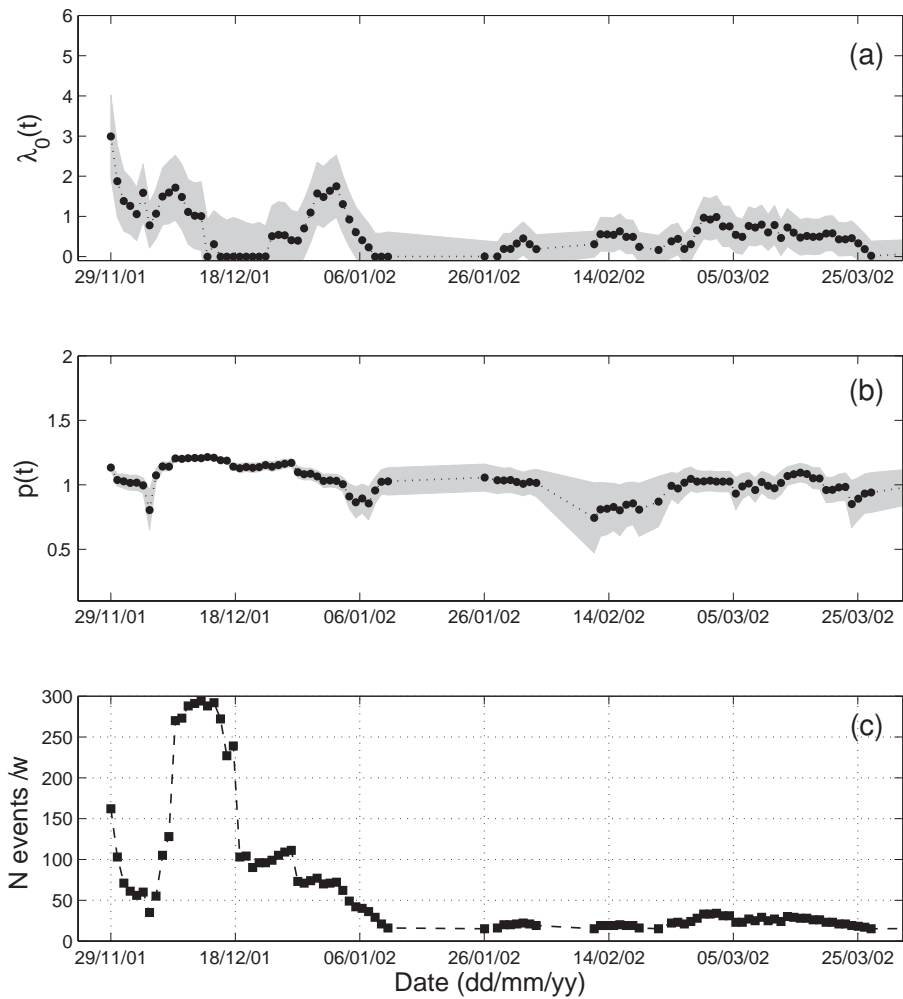
Figure 3.6: Time evolution of the non-stationary parameters of the ETAS model: Background activity $\lambda_0(t_j)$ (a), p-value (b), and number of events/window (c; $\tau = 10$ days), during the 2001-2002 unrest at Cotopaxi volcano, Ecuador; grey areas show $\pm 1\sigma$ error interval in all plots.

trast, the most important seismic activity recorded at Fuji since it is instrumentally monitored is characterized by deep low frequency events, which increased sharply in number in the period from October 2000 to May 2001. Some authors (e.g., *Ukawa*, 2005) suggest that the sharp increase of DLF activity at Mount Fuji began soon after magma discharge and intrusion events in the Miyakejima and Kozu-shima region in July and August 2000; then, these events may have modified the state of the deep magmatic system beneath Mount Fuji, thus triggering the DLF earthquake swarm (*Ukawa*, 2005). Nevertheless, no volcanic tremors, large earthquake swarm activity or abnormal crustal deformations were detected in or around Mt. Fuji in this period, for this reason we agree that the VT events probably indicate minor stress changes beneath the volcano and that abnormal activity was limited to the mid-crustal depth beneath the volcano, as suggested by some authors (e.g., *Ukawa*, 2005); then, no magma migrated upwards into the shallower crust, which is coherent with the stability in λ_0 and also the quite constant (close to 1 $\pm 4\%$) p-values.

In Fig. 3.8 we report the results for the Long Valley caldera unrest. The background activity λ_0 shows some variations with at least three major peaks which are correlated with periods of intense seismicity: two peaks in λ_0 occur between the end of October and the beginning of November 1997, and a third peak is located between the end of December 1997 and January 1998. On the other hand, no evident changes in the p-value are observed during the changes in λ_0. A very-low frequency trend is observed in the p-value, which shows a rather stationary value (around 1.3) during the first two peaks of high background activity (October-November), and a decreasing trend during the third peak (December 1997 - January 1998) toward a value around 1. In summary, the first period of more intense seismicity is characterized by the most intense background activity λ_0, and corresponds with no evident changes in the p-value; this period is also correlated with the greatest velocity of uplift observed in the caldera, and strong degassing processes (*Dreger et al.*, 2000; *Hill*, 2006). It is interesting to observe that the second seismicity peak at the end of December 1997 corresponds with low λ_0 values, suggesting that the bulk of that activity could probably be conformed by aftershock events.

3.5   Discussion

We have analyzed different seismic sequences through a stochastic non-stationary ETAS modeling. in particular, the non-stationary ETAS model with both background seismicity and p-value varying through time, provides a satisfactory description of complex seismic swarms, as pointed out by *Lombardi et al.* (2006). In our analysis we compare the time evolution of these two parameters for different seismic swarms which we group in 3 cases: VT swarms in volcanic areas before and/or during an eruptive process, VT swarms in non-erupting volcanic areas, and seismic swarms in tectonic areas.
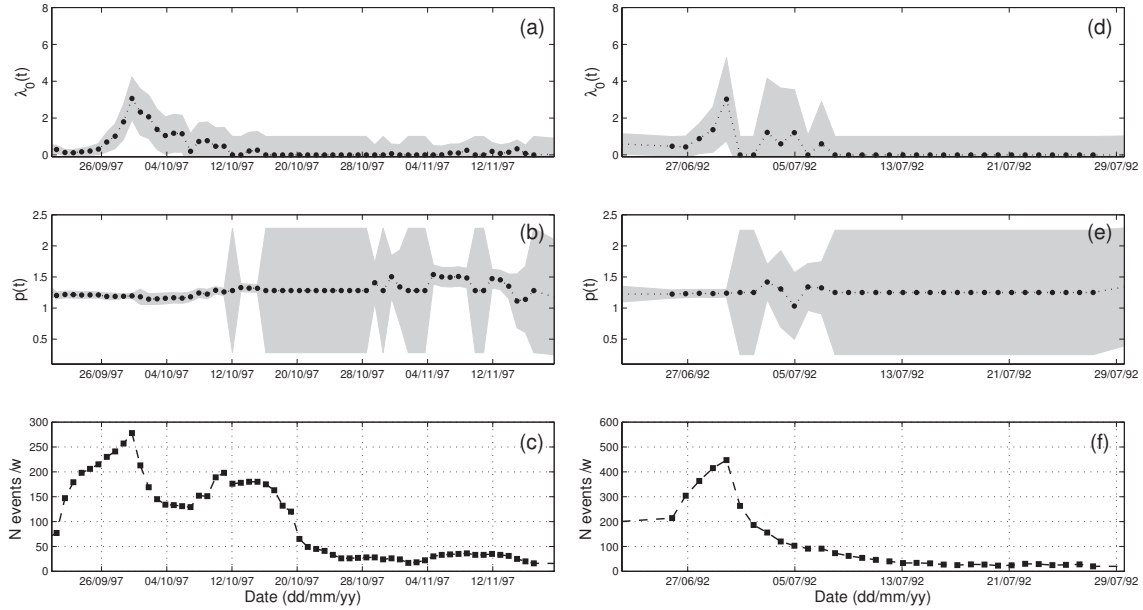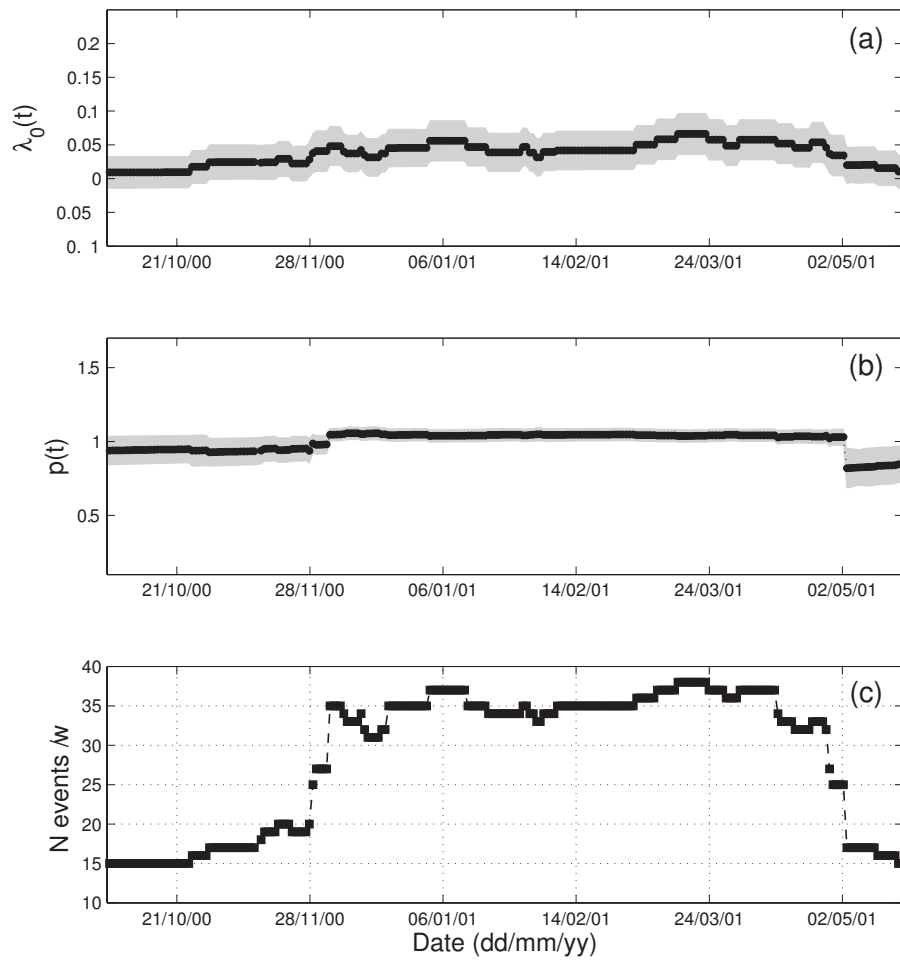
Figure 3.7: Time evolution of the non-stationary parameters of the ETAS model: Background activity $\lambda_0(t_j)$ (a), p-value (b), and number of events/window (c; $\tau = 150$ days), Mount Fuji volcano, Japan; grey areas show $\pm 1\sigma$ error interval in all plots.
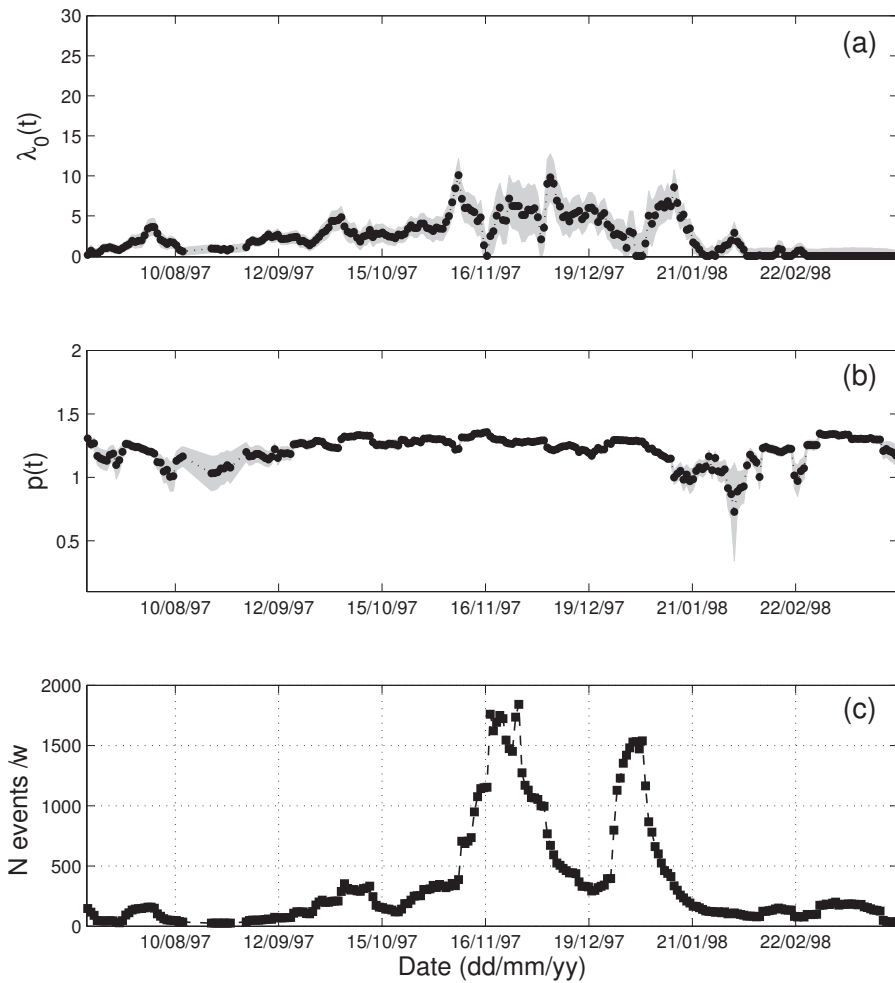
Figure 3.8: Time evolution of the non-stationary parameters of the ETAS model: Background activity $\lambda_0(t_j)$ (a), p-value (b), and number of events/window (c; $\tau = 10$ days), during the 1997-1998 unrest at Long Valley caldera, California; grey areas show $\pm 1\sigma$ error interval in all plots.

The studied VT swarms that have preceded or accompanied eruptive activity have some re-markable common features that have not been observed in the other two groups (i.e. tectonic and swarms in non-erupting volcanoes). In the case of Guagua Pichincha volcano, the period of most intense VT seismicity (September 1999) was characterized by the higher p-values measured during the whole unrest period, reaching values of 1.7 ($\pm 2\%$) at the end of the VT swarm. The period of high p-values is correlated with a peak in the background seismic-ity. The maximum of the VT seismicity rate occurred on September 23, and on September 25 the first intense LP swarm occurred in association with the intrusion of the first dacitic lava dome; for instance, following an explosion containing juvenile magma on September 26, the first dome grew into the crater up to October 5 and 7, 1999, when it was destroyed by Vulcanian-type magmatic explosions (*Garcia-Aristizabal et al.*, 2007). In the case of Miyake-jima volcano, the seismic sequence is characterized by wide oscillations of λ_0 at the beginning of the sequence (intrusion stage), accompanied by a rather stable, high value of p (around 2.4), as was already pointed out by *Lombardi et al.* (2006). Finally, the pre-eruption catalog that we analyze for Pinatubo volcano corresponds with the predome period defined by *Har-low et al.* (1996). The time evolution of λ_0 shows a peak around June 4-5, while the p-value shows a rather stable values around 2.1 to 2.2 ($\pm 1\%$). The seismic activity during this phase preceded the beginning of dome growth on the northwest flank of the volcano (e.g., *Hoblitt et al.*, 1996). In summary, the analyzed swarms of VT events that preceded or accompanied volcanic eruptions indicate that (1) they are characterized by high p-values (respect to both tectonic sequences and VT swarms in non-erupting volcanoes); in the analyzed cases the p-values range from 1.7 to 2.4; (2) the background seismicity generally shows wide variations during the period of maximum seismicity rate and the maximum peaks normally corresponds with the periods of maximum values of p.

To assess if a non-stationary ETAS model with both λ_0 and p-values is indeed the best model describing the observed data, we calculate the Akaike information criterion values (AIC, *Akaike*, 1974) for this model and also for an alternative model with just λ_0 as free parameter. The AIC statistic is defined by $AIC(\kappa) = 2\kappa - 2LogL$, where κ is the number of parameters, and $LogL$ is the log-likelihood of the model computed for the best parameters. In comparing models with different number of parameters, the quantity 2κ roughly compensates for the additional flexibility that the extra parameters provide. The lower values of AIC identifies the model that better represents the data. The results obtained for the sequences from erupting volcanoes (Guagua Pichincha, Miyakejima, and Pinatubo), are summarized on table 3.2, in-dicating that the model with both $\{\lambda_0(\tau)$ and $p(\tau)\}$ varying through time is the best one to describe the data.

Conversely, from the tectonic sequences analysis we could conclude that (1) an increase in the background seismicity may occur contemporarily with the period in which the main-shocks take place, and (2) no evident temporal changes in the p value are observed during the seismic swarms. The p-value is stationary around 1.2 in both Landers (1992-1993) and Umbria-Marche (1997-1998) sequences analyzed in this study. Further analysis should be

Table 3.2: AIC values for two non-stationary ETAS models with 1 free parameter ($\lambda_0(\tau)$, model 1), and 2 free parameters ($\lambda_0(\tau), p(\tau)$, model 2), for the erupting volcanoes data sets.

Data set	AIC - Model 1 $\lambda_0(\tau)$	AIC - Model 2 $\lambda_0(\tau),\ p(\tau)$
Guagua Pichincha	1983.6	1581.1
Miyakejima	-8800.4	-8821.8
Pinatubo (before)	-940.6	-1125.5

performed to establish the origin of the observed increase in λ_0 in these cases, since it could also be a bias due to incompleteness of the catalog (for small events) just after the mainshock events.

On the other hand, VT swarms in non-erupting volcanoes show a little more complex behavior of the analyzed parameters. In the case of Cotopaxi volcano, the period of most intense VT activity corresponds with increases in the background activity and no evident changes in the p-value; the maximum p-value found during this crisis was around 1.2. In the case of Mt. Fuji activity, some authors (e.g., *Ukawa*, 2005) suggest that the VT events analyzed probably indicate minor stress changes beneath the volcano and that abnormal activity was limited to the mid-crustal depth beneath the volcano, then no magma migrated upwards into the shallower crust, which is coherent with the quite constant (close to $1 \pm 4\%$) p-values. In the Long Valley case, the period of more intense seismicity is characterized by the most intense background activity λ_0, and corresponds with no evident changes in the p-value (which is rather stationary around a value of 1.3). In summary, the analyzed swarms of VT events that occurred on volcanic areas in absence of eruptive activity indicate that (1) the periods of more intense seismicity are mainly correlated with increases in the background seismicity (λ_0), even if the temporal behavior of λ_0 is not necessarily a proxy of the seismicity rate; (2) the p-value, which for the analyzed cases range from 1 to 1.3, does not show evident changes correlated with variations of λ_0.

Those observations can give us some insights on the mechanics of the seismic swarms on volcanic areas. The overall impression is that changes in the background seismicity (λ_0) may be observed in all the cases, even in tectonic environments. However, the p-value have some substantial differences that may be a helpful tool to characterize VT swarms. For example, the tectonic sequences analyzed show in general a rather stationary value of p, which in general is ≤ 1.2 in the analyzed cases; the VT sequences in non-erupting volcanic areas have some similar features as the tectonic cases, with stable values or showing small, high and low frequency oscillations. In the analyzed cases the p-value was always $p \leq 1.3$. Conversely, the VT swarms during eruptive episodes have p-values that in general are higher respect to the cases described before, being in general $p \geq 1.7$.

The results of this stochastic non-stationary ETAS model yield important clues to interpret the temporal evolution of complex seismic swarms. The phenomenological description of the temporal behavior of λ_0 and p parameters may be accompanied by a more detailed interpretation of the results in terms of the physical processes at the source of the swarms. Despite analytical definition of the physical meaning associated with λ_0 and p is not possible, some insights of the processes driving the changes may be deduced, specially from the relative changes in the parameter values and their possible correlation with other geophysical observations or with the eruptive activity itself. In general terms, VT seismic activity may be controlled both by fluid intrusion and stress triggering, and then it is possible to observe mixed occurrence of mainshock-aftershocks sequences and magma-related swarms. As consequence, background (λ_0) as well as induced activity could show evident non-stationarities.

Time variations of the background seismicity, λ_0, are usually associated to the time evolution of the energy source (e.g., see *Hainzl and Ogata*, 2005; *Lombardi et al.*, 2006). In particular, *Hainzl and Ogata* (2005) analyzed a large earthquake swarm that occurred in 2000 in Vogtland Bohemia, central Europe, and by fitting a non-stationary ETAS model (where λ_0 was the only free parameter) and modeling the 3D elastic stress changes in the crust, they found that stress triggering is dominant in creating the observed seismicity patterns, but external forcing, identified with pore pressure changes due to fluid intrusions, is found to directly trigger part of the activity, specially during the first days of the swarm.

It is possible to find fluids in both tectonic and volcanic regions, for this reason it is possible to find time variations of λ_0 in both purely tectonic and volcano-tectonic seismic sequences. Physical constrains to understand temporal changes in λ_0 associated with fluid activity may be explored, as highlighted by *Hainzl and Ogata* (2005); for example, stress triggering, a mechanism based on the well-known Coulomb failure stress (CFS) criterion ($CFS \equiv \tau - \mu(\sigma - P') \geq 0$, e.g., *Harris*, 1998), has been identified as a possible mechanism for aftershock sequences (e.g., *Stein*, 1999). The CFS criterion states that a positive Coulomb failure stress could promote failures; here, τ defines the shear and σ the normal stress on the failure plane (positive for compression), μ is the coefficient of friction, and P' is the pore pressure. In presence of fluids, the pore pressure has to be taken in account; for instance, in regions where fluids may migrate through fractures (i.e. volcanoes), the effective normal stress ($\sigma - P'$) may decrease during fluid injections (as pore pressure increases) and then a positive variation of CFS may promote failure. In this case, the background rate λ_0 changes may be related to activity forced by pore pressure changes.

In the other hand, the p-value, which controls the temporal decay of triggered events, is found to be positively correlated with crustal temperature; for instance it controls stress release and therefore aftershocks decay (e.g., *Mogi*, 1967; *Kisslinger and Jones*, 1991; *Lombardi et al.*, 2006). Physical constrains to interpret time variations of p are more difficult to define quantitatively. The studied cases indicate that VT swarms that precede or accompany dike intrusions are characterized by relatively high p-values ($p \geq 1.7$ in the cases analyzed here).

High p-values, that indicate a sharp decaying aftershock activity, in general are correlated to maximum peaks in the background seismicity, which in turn may be stimulated by fluid migration. Fluid migration may be an effective mechanism of heat transfer by transport, then even if characteristic times associated with dike intrusion may be relatively short respect to the characteristic times for heat transfer by conduction (e.g. the Miyakejima case), hyper-heated fluids that escape from magma as it decompresses during the ascent toward the surface may effectively transport and transfer heat to the country rock and shallower hydrothermal fluids.

3.6   Concluding remarks

We perform a retrospective analysis of volcano-tectonic seismic swarms using a non-stationary, epidemic-type, aftershock sequences model in which both the background seismicity and the p-value can vary through time, which has evidenced some important characteristics of the VT swarms associated with magmatic intrusions. In particular, swarms of VT events that precede or accompany magmatic intrusions (i.e. before or during eruptive activity) present remarkably high p-values respect to both, VT swarms in cases with no evident shallow magmatic activity, and swarms in purely tectonic environments. From a practical point of view, the stochastic modeling of VT swarms using a non-stationary ETAS model may be potentially used for a quick analysis of VT swarms to identify and discriminate those which are more likely associated with the physical response of the volcanic edifice to a magmatic intrusion. Routinely monitoring of both λ_0 and p parameters using a moving and overlapping temporal window as used in this work may produce a valuable seismic parameter for both volcano monitoring and for short-term eruption forecasting purposes. In Chapter 4 (section 4.4), we discuss a possible strategy to integrate this information into the Bayesian Event Tree model for eruption forecasting (BET_EF) of *Marzocchi et al.* (2008).

Chapter 4

Integration of Parameters in a Bayesian Event Tree model for Eruption Forecasting

"The more we learn about the world, and the deeper our learning, the more conscious, clear, and well-defined will be our knowledge of what we do not know, our knowledge of our ignorance. The main source of our ignorance lies in the fact that our knowledge can only be finite, while our ignorance must necessarily be infinite."

Karl Popper
(Conjectures and Refutations: The
Growth of Scientific Knowledge, 1963)

4.1   Introduction

In previous Chapters, some specific analysis have been discussed in order to provide further tools for the short- and long-term eruption forecasting (EF) assessment. The main objective of this Chapter is to provide a general methodology to integrate the most important results obtained in Chapters 1 (for long-term EF), 2 and 3 (for short-term EF) to improve the setup of the Bayesian Event Tree model for eruption forecasting (BET_EF, e.g., *Marzocchi et al.*, 2008).

In particular, we are interested in considering the next topics:

- Integrate, into the non-monitoring part of BET_EF, information derived from time-dependent or time-independent, long-term eruption forecasting models based on any

renewal process, as for example the Brownian passage-time described in Chapter 1.

- Integrate, into the monitoring part of BET_EF, information derived from geophysical measurements; in particular, we are interested in provide some insights for the interpretation and further inclusion in BET of the specific analysis performed in this work (i.e. spectral characteristics of LP events and characteristics of VT swarms)

- Provide a general approach for the quantitative determination of thresholds of anomalous seismicity rates.

The Bayesian Event Tree (BET) is a flexible tool to provide probabilities of any specific event which we are interested in, considering both volcanological and monitoring information; furthermore, it may be dynamically used for long- and short-term eruption forecasting (for a general overview of the mathematical background, see electronic supplementary material in *Marzocchi et al.* (2008), and Appendix A). The non monitoring part of BET applied to EF (BET_EF) uses mainly information from the past unrest episodes, eruptive history of the volcano, theoretical and empirical models based on similar volcanoes, etc. All the temporal considerations in BET are implicit at the Node 1 (e.g. see Appendix A), where the probability of a generic unrest in a given time window τ is defined. However, available catalogs of eruptions are more common than catalogs of unrest, then in some cases it can be easier to construct a probabilistic model based on past eruptive activity than based on past unrest episodes. For instance, as seen in Chapter 1, when a catalog of eruptions is available, time-dependent (or time-independent) probabilistic models for long-term eruption forecasting may be constructed; in section 4.2 of this chapter we propose a methodology to consider any time-dependent or time-independent, long-term, eruption forecasting model based on the eruptive history of the volcano (as for example the Brownian passage-time model described on Chapter 1) into the BET model structure.

On the other hand, monitoring information is used in BET for the short-term eruption forecasting assessment. At each node (see electronic supplementary material in *Marzocchi et al.* (2008), and/or Appendix A for details), strong monitoring parameters should be defined to detect changes in the volcanic system that may potentially indicate generic unrest (node 1), magmatic unrest (node 2), and eruption (node 3). The monitoring parameters defined to feed the event tree are fundamental tools to identify and characterize unrest episodes, and contribute information that BET translates in short-term probabilities of a given event of interest. However, this task is not easy and the parameters should be carefully defined. Volcanoes are monitored using a wide range of geophysical and geochemical techniques (e.g. seismology, geodesy, acoustic, thermal (IR), gas and water chemistry, etc.), each of them providing information which could be potentially considered. In general terms, multidisciplinary parameters should be identified for each node, since information from a single discipline or a single measurement may be meaningless if considered alone for forecasting purposes.

Two main problems should be considered to define monitoring parameters in BET: (1) the definition of the physical measurement to be considered (and in which node); and (2) the

thresholds (generally two, for the Fuzzy approach) to define the *anomaly* values of the parameter. Of course, there is not a *magic* setting of BET and many parameters and thresholds are specific for each particular case. In BET applications to some volcanic areas as Mt. Vesuvius (*Marzocchi et al.*, 2004; *Sandri et al.*, 2009) and the Aukland Volcanic Field (*Lindsay et al.*, 2010), the monitoring parameters have been established mainly using elicitation procedures; in this way experiences of past unrest episodes (from the same volcano or a similar one) and expert opinion (based on basic research on each specific area) are the main information sources to define both parameters and thresholds.

Within this context, it is clear that geophysical research oriented to measure and understand physical processes during volcanic unrest episodes, magma movement, etc., is a source of information from which strong and reliable monitoring parameters may be discovered and used to feed probabilistic models for eruption forecasting (as BET). In this work, some topics from volcano seismology have been analyzed in order to explore further seismological parameters that could provide valuable information to feed some nodes of BET model, specially the node 2 (magmatic unrest), that is probably the most important source of epistemic uncertainty in the model. In Chapter 2 we have briefly described some characteristics of volcanic long-period (LP) seismicity, defined an alternative methodology for LP spectral analysis, and proposed a new approach to use it for volcano monitoring; likewise, in Chapter 3 we analyzed some characteristics of the volcano-tectonic earthquake swarms by non-stationary ETAS modeling. In this chapter (sections 4.3 and 4.4) we make a general discussion of the physical implications of the information extracted using the described methodologies, and we suggest some procedures to include this information as parameters to feed the monitoring part for short-term eruption forecasting in the BET model.

The last section of this Chapter is a brief discussion about the analysis of *background* levels of activity; in particular, one of the most *natural* and *intuitive* seismological parameters often used to identify a volcanic unrest is the seismicity rate. Anomalous levels of seismicity respect to a background level are often one of the first indications of the beginning of an unrest episode. However, the determination of thresholds of anomalous seismicity may be an empirical task in which a complete database and the opinion of experienced volcano-seismologist (on the activity of a specific volcano) are normally the source of information. In an attempt to give some general guidelines for a quantitative determination of the thresholds of anomalous seismicity rates, in section 4.5 we discuss a generic approach to model the background seismicity, which can be straightforward extended to other kinds of available data.

4.2   Setting BET parameters according to the results of long-term, time-dependent or time-independent eruption forecasting models

The temporal analysis in BET is done at node 1 (probability of unrest) and it is defined through a static setting of its parameters, i.e., the average and the equivalent number of data of the prior model (Θ_1 and Λ_1 respectively), and past data $y_i^{(1)}$ (for details see *Marzocchi et al.*, 2008). For the definition of the *equivalent number of data* concept, see Appendix A: sections A.0.5 and A.0.6, and equations A.10 and A.11. Once defined the parameters of BET at node 1, the probability estimation is set and does not change through time. In general terms, the BET_EF model is neither completely time-dependent nor completely time-independent; in fact, the statistical assessment simply refers to the next time window defined (τ), and both prior model(s) and past data may change through time, and making consequently change also the BET statistical assessments. In this way, time-dependent analysis may be represented with BET through a series of repeated static analysis whose parameters change through time. In practice, the results of a time-dependent model assessing the probability of a specific event may enter into the BET model through a continuous update of the BET parameters. For this reason, in this section we will refer to the analysis done at a generic time t and, being this method repeatable at any given time, the results will be valid for both time-dependent and time-independent models.

Mainly for short-term practical utility, the first node of the Event Tree defined in the BET_EF model defines the probability of unrest in the next time window τ. Since, as mentioned above, data about past unrest episodes are available almost only for the last decades, most of long-term assessments are based on the analysis of past eruptions (e.g. see Chapter 1). This means that a possible output variable of such models may be the absolute probability of eruption in the next time window τ, i.e., θ_E. The goal of this section is to write the parameters of the BET model in order to fit the values of θ_E as obtained by a given external long-term model (that hereinafter we call $\theta_E^{(model)}$). This allows including the information provided by a given model into the more general framework of eruption forecasting provided by BET, in order (1) to extend forward the model's forecast to vent location and eruption size forecasts, and (2) to infer backward forecasts about unrest and magmatic unrest episodes.

Referring to BET symbols (for details see *Marzocchi et al.* (2008), or the Appendix A), the distribution relative to the absolute probability of eruption in the next time window τ is:

$$[\theta_E] = [\theta_1][\theta_2][\theta_3] \tag{4.1}$$

where θ_1 represents the probability of an unrest episode in τ, θ_2 the probability of a magmatic unrest, given an unrest episode, and θ_3 the probability of eruption, given a magmatic unrest, and the brackets $[a]$ means the probability distribution of the variable a. Therefore, the absolute probability of eruption simultaneously depends on the probability distributions at nodes 1, 2 and 3.

In principle, the $[\theta_2]$ and $[\theta_3]$ distributions must be assumed valid only for the time window τ, but their parameters are usually set using all past data (*Marzocchi et al.*, 2004, 2008), implying that all the time-dependence of the probabilistic assessment is relied to node 1. This simply means that the relative proportion of eruptions from magmatic unrest (node 3) and the proportion of magmatic unrest from generic unrest (node 2) is considered constant over time, and based on time independent considerations. Here, we will follow the same philosophy, assuming the distribution $[\theta_2]$ and $[\theta_3]$ as constant through time. Therefore, the BET parameters to be set from $\theta_E^{(model)}$ are the parameters of the $[\theta_1]$ distribution, i.e., its average Θ_1 and its variance through Λ_1. On the other hand, the distributions $[\theta_2]$ and $[\theta_3]$ and their parameters will be considered here as known.

Most of probabilistic models does not provide the analytic form of $[\theta_E^{(model)}]$; instead, only the best guess value of $\theta_E^{(model)}$ is evaluated, i.e., $E([\theta_E^{(model)}])$. Sometimes, also other parameters defining $[\theta_E]$ are provided, i.e., its variance or several percentiles, based on statistical analysis and/or bootstrap procedures. On the other hand, in the BET model, the statistical distributions $[\theta_1]$, $[\theta_2]$ and $[\theta_3]$ are Beta distributions, while the statistical distribution of θ_E is generally not set or is obtained numerically. This imply that in most of the applications, the analytical form of $[\theta_E]$ will not be known, and the inversion must be done numerically. However, this also implies that equation 4.1 cannot be simply inverted for $[\theta_1]$, since each realization of θ_E strictly depends on specific (and unknown) realizations of θ_2 and θ_3.

The mean of $[\theta_E]$ in BET can be expressed in terms of the means of $[\theta_1]$, $[\theta_2]$ and $[\theta_3]$, i.e., Θ_1, Θ_2 and Θ_3 respectively, being $[\theta_1]$, $[\theta_2]$ and $[\theta_3]$ Beta distributions:

$$E([\theta_E]) = E([\theta_1])E([\theta_2])E([\theta_3]) = \Theta_1\Theta_2\Theta_3 \tag{4.2}$$

where Θ_2 and Θ_3 are known. Therefore, we can set Θ_1 using our best guess value of $E([\theta_E])$, that is $E([\theta_E^{(model)}])$. In practice, we set

$$\Theta_1 = \frac{E([\theta_E^{(model)}])}{\Theta_2\Theta_3} \tag{4.3}$$

Since Θ_1 must be smaller than 1, we have that

$$\Theta_2\Theta_3 > E([\theta_E^{(model)})]) \tag{4.4}$$

that is always true for well defined assessments, since the probability that an unrest episode leads to an eruption ($\Theta_2\Theta_3$) must be greater than the absolute probability of eruption.

To set the other free parameter, the equivalent number of data at node 1 (Λ_1), we must do several considerations. First, we must have an explicit estimate on the variance of $[\theta_E^{(model)}]$ (or, equivalently, a confidence interval through percentiles), which in most of the applications is not evaluated; in this case, Λ_1 may be set according to the (subjective) evaluation of the reliability of the model. Second, when the variance is estimated, we must consider whether

such variance correctly represents the entire epistemic uncertainty related to the physical process.

If the variance of $[\theta_E^{(model)}]$ is provided and it is assumed be representative of the entire epistemic uncertainty, Λ_1 can be set using this information. Unfortunately, an analytic relationship like the one for averages (Eq. 4.2) does not exist; therefore, Λ_1 must be inferred using an inverse procedure, that is, evaluate $[\theta_E]$ for different values of Λ_1 (forward modeling with BET) and compare its variance (or percentiles) with the variance (or percentiles) of $[\theta_E^{(model)}]$. A possible procedure might be:

$$V_i = \left\| variance\left([\theta_E^{(model)}]\right) - variance\left([\theta_E]^{(\lambda_i)}\right) \right\| \tag{4.5}$$

in which $[\theta_E]^{(\lambda_i)}$ is obtained using equation 4.1, where $[\theta_1]$ is a Beta distribution whose parameters are set with Θ_1 from Eq. 4.3, and $\Lambda_1 = \lambda_i$, for $\{\lambda_i = 1, 2, \ldots, \lambda_i \in \mathbb{N}\}$. In this case, the best guess value for Λ_1 will be the value λ_i (i.e. the equivalent number of data) that minimize V_i.

With the selected values of Θ_1 and Λ_1, the parameters of the BET model will be completely defined, and its long-term probability of eruption assessment will be coherent with the forecast provided by external probabilistic models based on eruptive catalogs (see Eq. 4.1). Moreover, this procedure intrinsically allows the estimation of the probability of unrest ($[\theta_1)]$), that, in principle is more testable (i.e. verifiable/falsifiable) than the probability of eruptions, being unrest episodes more frequent than eruptions. Moreover, through BET, it is possible to assess the probability of occurrence of any possible path within the event tree, i.e., the probability of an eruption of a given size and with vent in a given location (*Marzocchi et al.*, 2004, 2008).

4.3   Potential use of Complex Frequencies of LP events as input monitoring parameter in BET

In recent years, there has been considerable interest in the origin of LP events and volcanic tremor, and some quantitative models to describe physical mechanisms of their source have been proposed. Several attempts have been made to explain the tremor signals involving source or path effects; for example, *Omer* (1950) in the 1950s attributed tremor to the vibration of shallow volcanic layers caused by surging magma, while *Shima* (1958) thought that the dominant tremor frequency was due to the spherically symmetric, fundamental mode of vibration of a liquid sphere. In the 1960s it was considered to be an eigen-frequency of a vibrating cylindrical magma column (*Shimozuru*, 1961; *Steinberg*, 1965). In the seventies, *Steinberg and Steinberg* (1975) took a different view and associated the tremor with the self-oscillations of gases in the volcanic vent generated by the transition of subsonic to supersonic flow. However, these models failed to provide a unified theory accounting for the driving force in the fluid and the elastic radiation from the source, and failed to explain not only

the observed frequency but also the amplitude as well as the duration of the tremor (*Chouet*, 1981).

Aki et al. (1977) proposed three possible models of a fluid-filled tensile crack driven by excess fluid pressure to explain shallow tremor in terms of physical parameters related to the source, offering at the same time an adequate geometry for mass transport. In these models the sudden extension of the crack or opening of the channel is the source of the seismic waves. *Chouet* (1981) presented a model of a fluid-filled crack driven by excess fluid pressure and investigated the effects of depth of burial, fluid characteristics, and medium structure on the ground response in the near-field of the crack. Sustained tremor can be viewed as resulting from the superposition of many individual sources triggered randomly in time. The spatial and temporal characteristics of tremor and in particular the occurrence of spasmodic tremor can be directly related to fluid supply and excess pressure in the model (*Chouet*, 1981). A model of the resonance of a fluid-driven crack induced by an impulsive pressure transient was found to explain many of the observed characteristics associated with long-period (or low-frequency) events and harmonic tremor. The main feature of this model is the existence of a very slow wave in the fluid-filled crack (what he called "crack wave"), the wave speed of which decreases with increasing wavelength and with increasing values of the crack stiffness parameter (C, e.g., *Aki et al.*, 1977). This crack wave provides a harmonious explanation for the observation of very long period tremor without requiring an unrealistically large magma container (*Chouet*, 1988). The fluid-driven crack models fully analyze the coupling between the fluid and the solid, leading to quantitative results that allow a direct comparison with seismic observations. In these models, the equations of mass and momentum transfer in the fluid are solved simultaneously with the equations of elastodynamics in the solid using a finite-difference approach (*Chouet*, 1988, 1992). *Ferrazzini and Aki* (1987) followed an elegant analytical approach and derived the dispersion relation of what they called "slow waves", concluding that low-frequency resonances can be generated in fluid-filled containers such as cracks, dykes and conduits, if the width of the container is small relative to the seismic wavelength.

Both spherical and cylindrical geometries have been also considered as possible source geometries: *Fujita et al.* (1995) and *Fujita and Ida* (2003) analyzed the eigen-oscillations of a fluid sphere and defined a model for the source of harmonic volcanic tremor and LP events; on the other hand, using a 2D finite difference method, *Neuberg et al.* (2000) modeled major features of low-frequency seismic signatures on Montserrat volcano and compared them with observations, discussing some features of both fluid (magma) properties and geometrical effects of the conduit.

Beside the specific model characteristics, the common feature of most proposed models is that we can consider that the source mechanisms associated with LP events and volcanic tremor are intimately associated with fluids contained in a solid rock matrix. Both the frequency content and the characteristic long-lasting coda of these seismic signals are of fundamental

importance since they represent evidences of characteristic properties (e.g. geometrical and fluid physical properties) of the source, as have been highlighted by many authors (e.g., *Aki et al.*, 1977; *Chouet*, 1981, 1982, 1985, 1988; *Ferrazzini and Aki*, 1987; *Fujita et al.*, 1995; *Fujita and Ida*, 2003; *Neuberg et al.*, 2000; *Morrissey and Chouet*, 2001).

Spectral analysis in the complex frequency domain opened a window to try to understand some of the source properties. Despite of the importance of this argument, not so many papers have been produced trying to analyze and quantify the link between the spectral characteristics (i.e. frequency, growth rate, and Q factor, for details see Chapter 2) and source properties of LP events. *Aki et al.* (1977) derived a formula for Q for the fundamental radial mode of a sphere; further developments of the fluid-filled crack model accounting for various mixtures of liquid, gas, and solids have been carried out by *Kumagai and Chouet* (1999, 2000, 2001), in which a discussion of the acoustic properties of a crack containing magmatic or hydrothermal fluids is performed for a fixed (crack-like) geometry. Other analyses attempt to quantify both geometrical effects and acoustic properties of fluids (e.g. conduit length (*Sturton and Neuberg*, 2006), or sphere radius (*Fujita and Ida*, 1999, 2003)) and its influence on the waveforms and spectral characteristics of recorded events.

In summary, even if it seems clear that the spectral characteristics of LP events in terms of frequency, growth rate, and Q factor are key parameters for the inference of some source properties, non univocal interpretations are possible at the moment. In Chapter 2 we presented an alternative method to use this information for the systematic monitoring of possible changes in these values. However, if we were interested on the interpretation of source properties and to perform inferences about fluids at the source, then a quantitative exercise can be done using the results of some papers (those with more detailed results) just as a first order approximation to define some intervals for the spectral parameters; in this way, we can provide the possible interpretations up to the current knowledge, considering always that those values and interpretations are merely approximative and subject to significant epistemic uncertainties.

Assuming that the source of LP events consists of a time-localized excitation (present only during a short time interval at the beginning of the event) and a subsequent resonance of a (crack-like) volcanic fluid system in the source region, and that the recorded waveform may be described by a superposition of simple decaying sinusoids, we can use the results of *Kumagai and Chouet* (1999, 2000, 2001) to give some insights for the interpretation of the results of the spectral analysis in the complex frequency domain.

The observed Q of the LP event may be expressed as $Q^{-1} = Q_r^{-1} + Q_i^{-1}$, where Q_r^{-1} and Q_i^{-1} represent the radiation and intrinsic losses, respectively (*Aki*, 1984; *Kumagai and Chouet*, 1999, 2000, 2001). In the simulations performed by *Kumagai and Chouet* (1999, 2000, 2001) using the crack model of *Chouet* (1986, 1988, 1992), predictions of the expected Q_r^{-1} were obtained for different values of α/a and ρ_f/ρ_s, where α is the compressional wave velocity

in the rock matrix, a is the acoustic velocity in the fluid, and ρ_f and ρ_s are the fluid and rock matrix densities, respectively.

Kumagai and Chouet (1999, 2000, 2001) found that Q_r almost monotonically increases with increasing α/a and slightly increases with decreasing ρ_f/ρ_s; therefore, Q_r increases with the impedance contrast Z (where $Z = \alpha\rho_s/a\rho_f$), as suggested by *Aki et al.* (1977), but they argue that Q_r more strongly depends on α/a and on the geometry of the resonator than on Z (e.g., *Kumagai and Chouet*, 2000). On the other hand, the dimensionless frequency ν (defined as fL/α, where f is the frequency and L is the crack length), decreases with increasing both α/a and ρ_f/ρ_s and depends equally on both of these parameters.

To interpret these properties in terms of a crack containing hydrothermal fluids, *Kumagai and Chouet* (2000) examined various mixtures of gas, liquid and solid including gas-gas mixtures ($H_2O - CO_2$ and $H_2O - SO_2$), liquid-gas mixtures (water$-H_2O$ and basalt$-H_2O$), and *dusty* and *misty* gases (ash-SO_2 and water droplet-H_2O); the acoustic properties of some of these fluid mixtures are summarized in *Morrissey and Chouet* (2001). The results of the model proposed by *Kumagai and Chouet* (2000) are summarized as follows: (1) Q_r in the gas-gas mixtures ranges from 10 to 90 and Q_i^{-1} in such mixtures is negligibly small compared to Q_r^{-1}; (2) Q_r may range from nearly unity in pure liquid water and pure liquid basalt to 70 and 110 in bubbly water and bubbly basalt, respectively, at a gas-volume fraction of 10%, but in this kind of bubbly fluids the Q_i^{-1} factors strongly depend on the bubble radius (for instance, Q_i^{-1} in bubbly fluids becomes comparable to or larger than Q_r^{-1} in presence of bubbles whose radii are larger than 1mm); (3) Q_r in liquid-gas foams is larger than Q_r in bubbly fluids and ranges up to 140 in water-gas mixtures, and 210 in basalt-gas mixtures; (4) Q_r in dusty and misty gases ranges up to several hundred due to the higher velocity contrast between fluid and rock matrix; in general, Q_i^{-1} is negligibly small compared to Q_r^{-1} for small-size particles ($\leq 1\mu m$), but it can be larger than Q_r^{-1} for large-size particles ($\geq 100\mu m$).

As can be seen, a crack filled with different types of fluids may produce similar values of Q, so multiple interpretations of possible fluid types are plausible based on the observed Q, specially for values of Q in the range $10 \leq Q \leq 100$. The expected Q values and the most likely fluid mixtures (according to the model of *Kumagai and Chouet* (2000)) are summarized in Fig. 4.1. Four general intervals of Q values have been subjectively created (I_1 to I_4, see Fig. 4.1) in order to provide some general guidelines for the interpretation of the analysis of LP data; the likely fluid mixtures for each Q interval are summarized in table 4.1.

Our main interest is to individuate the cases in which the LP data could be more informative to identify activity associated with magma; this information may be potentially used as monitoring parameter in a probabilistic model (e.g., node 2 in BET, *Marzocchi et al.*, 2004, 2008). Interpretations based just in the Q value may be manifold, but in some cases the depth of the LP event may provide a further constrain to identify cases in which the most likely fluid composition may be associated with magma; for instance, liquid water is present in some of

Table 4.1: Summary of expected values of Q_r for different mixtures of fluids at the source of LP events, following the model and proposed values of *Kumagai and Chouet* (2000): intervals arbitrarily selected, possible fluid mixtures, and discriminant criteria using the depth of the event (for details, see the text).

Interval name	Range of Q values	Likely fluid mixtures	Event depth $z < \zeta_c$	Event depth $z \geq \zeta_c$
I_1	$Q \leq 10$	- Pure liquid water	*	
		- Pure liquid basalt	*	*
I_2	$10 < Q \leq 100$	- Gas-Gas mixtures (H_2O-CO_2 and H_2O-SO_2)	*	*
		- Pure liquid or bubbly $H_2 0$ [1]	*	
		- Water-gas foam	*	
		- Pure liquid or bubbly basalt [1]	*	*
		[1] $\leq 10\%$ *gas volume*		
I_3	$100 < Q \leq 150$	- Water-Gas foams	*	
		- Bubbly basalt or basalt-Gas foams	*	*
		- Misty (water droplet) gases	*	
		- Dusty (ash) gases [2]	*	*
I_4	$Q > 150$	- Basalt-gas foam	*	*
		- Dusty (ash) gases [2]	*	*
		- Misty (water droplet) gases	*	
		[2] *small−size particles*		
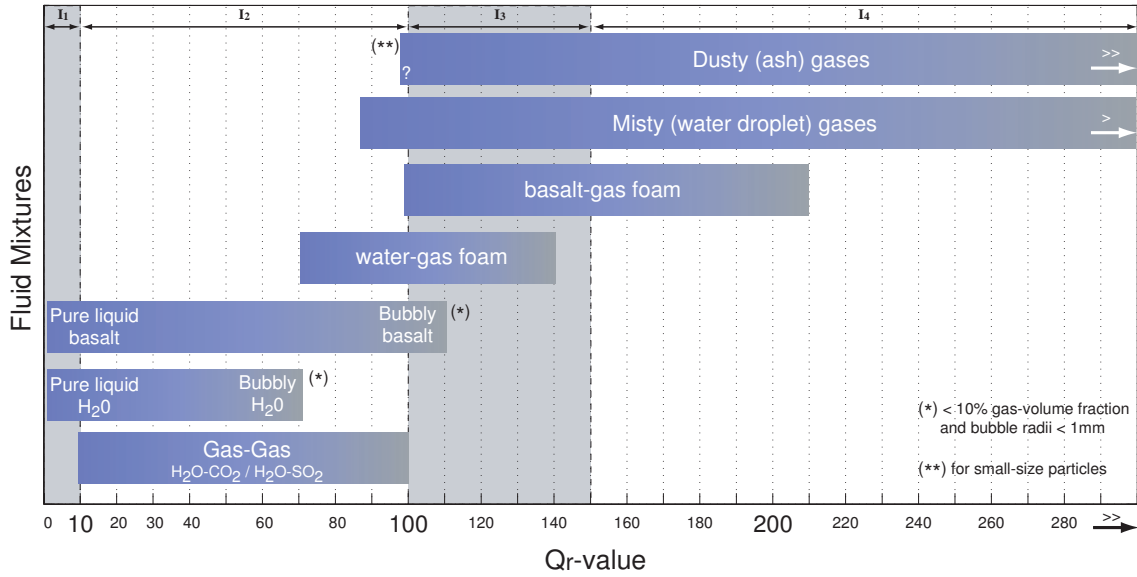
Figure 4.1: Intervals of expected values of Q_r for different mixtures of fluids at the source of LP events, following the model and proposed values of *Kumagai and Chouet* (2000).

the possible fluid mixtures (i.e. pure liquid water, bubbly H_2O, water-gas foams, and misty gas). If we consider the critical point for the water in terms of temperature and pressure (647°K and 22MPa, respectively, see, e.g., *Palmer et al.* (2004)), then a critical depth (ζ_c) may be defined in a specific volcanic zone, in function of both the geothermal gradient and the lithostatic pressure. Unfortunately, it is nor easy to find detailed quantitative information about the geothermal gradient for specific areas, but in a first order approximation we can consider that values ranging from ~30°K/km up to ~200°K/km or higher may be found in volcanic regions (e.g., *Ingebritsen and Sanford*, 1998; *Plummer et al.*, 2001; *Turcotte and Schubert*, 2002; *Arnorsson et al.*, 2008); roughly, these values may indicate a range of about 2 to 12 km for ζ_c. For example, in the Campi Flegrei caldera, Italy, temperatures of ~ 670°K have been reported at about 3 km depth (*Chelini and Sbrana*, 1987; *Agip*, 1987).

ζ_c would represent a limit value: LP events localized deeper than the critical depth will be less likely to contain liquid water in the fluid mixture at the source, and then this information combined with the Q estimates could help to identify LP events that are more informative to identify magmatic activity. For example, a LP event with $Q < 10$ and located at depth $z >> \zeta_c$, would be more likely interpreted as the resonance of a structure containing magmatic fluids.

To combine both Depth (z) of LP source and Q values in a single parameter useful in BET (specifically for the monitoring part in node 2), we can define a combined parameter $\mathbf{P}_{(z,Q)}$ whose value will depend on both the input values of z and Q, and on certain logic rules. Specifically, a flexible approach based on Fuzzy logic (*Zadeh*, 1965) may be applied to this kind of problems: this is an approach often used to model and deal with imprecise information such as inexact measurements, imprecise concepts, imprecise dependencies, modeling of

expert knowledge, or representation of information extracted from inherent imprecise data (e.g. *Berthold and Hand*, 2003).

Fuzzy sets, introduced by *Zadeh* (1965) as an extension of the classical notion of set, are sets whose elements have degrees of membership. In classical set theory, the membership of elements in a set is assessed in boolean terms according to a bivalent condition (an element either belongs or does not belong to the set); by contrast, Fuzzy set theory permits the gradual assessment of the membership of elements in a set, which is described with the aid of a *membership function* valued in the real interval $[0, 1]$.

The membership function is a representation of the magnitude of participation of each input. It associates a weighting with each of the inputs that are processed (defining the degree of membership - DOM), and ultimately determines an output response (e.g., *Zadeh*, 1965; *Klir and Folger*, 1987). For the combined analysis of z and Q parameters, two membership functions may be created: the first membership function, μ_z, mapping information from the depth of the LP event, and the second membership function, μ_Q, mapping the information of the quality factor Q. To discriminate LP events that can potentially be associated with magmatic activity (or magmatic-induced activity), we can define basic logic rules based on z and Q values in order to consider the most informative combinations indicating the possibility of magmatic fluids at the source. In general terms, we consider that the most informative combination (for magmatic fluids at the source) of z and Q values of LP events is (1) events localized at depth $z > \zeta_c$, and (2) Q values in the range $Q \lessapprox 10$ or $Q \gtrapprox 100$. For instance, deep LP events (i.e. $z > \zeta_c$) with low Q values $Q \lessapprox 10$ may be more likely interpreted as perturbations on magma-filled structures; on the other hand, deep events with $Q \gtrapprox 100$ may be interpreted as magma-gas foams, or dusty (ash) gases.

Using these indicative critical values, we define the membership function μ_z for the depth of the LP event as:

$$
\mu_z(z) = \begin{cases} 0 & \text{if } z < \zeta_c^l \\ \frac{1}{2}\left\{\sin\left[\pi\left(\frac{z-\zeta_c^l}{\zeta_c^u-\zeta_c^l}\right) - \frac{\pi}{2}\right] + 1\right\} & \text{if } \zeta_c^l \le z < \zeta_c^u \\ 1 & \text{if } z \ge \zeta_l^u \end{cases} \qquad (4.6)
$$

where ζ_c^l and ζ_c^u define a confidence interval around the critical depth ζ_c; note that between ζ_c^l and ζ_c^u a sinusoidal function has been defined in order to reproduce a smooth transition, however, a linear function may be also defined (as often found in Fuzzy logic applications).

On the other hand, for the membership function for the Q values, two critical points Q_{c1} and Q_{c2} must be defined; if we consider uncertainty intervals $[Q_{c1}^l, Q_{c1}^u]$ for the critical Q_{c1}, and $[Q_{c2}^l, Q_{c2}^u]$ for the critical Q_{c2}, the membership function may be defined as:

$$\mu_Q(Q) = \begin{cases} 1 & \text{if } Q < Q_{c1}^l \text{ or } Q > Q_{c2}^u \\ \frac{1}{2}\left\{\sin\left[\pi\left(\frac{Q-Q_{c1}^l}{Q_{c1}^u-Q_{c1}^l}\right) + \frac{\pi}{2}\right] + 1\right\} & \text{if } Q_{c1}^l \leq Q < Q_{c1}^u \\ 0 & \text{if } Q_{c1}^u \leq Q < Q_{c2}^l \\ \frac{1}{2}\left\{\sin\left[\pi\left(\frac{Q-Q_{c2}^l}{Q_{c2}^u-Q_{c2}^l}\right) - \frac{\pi}{2}\right] + 1\right\} & \text{if } Q_{c2}^l \leq Q < Q_{c2}^u \end{cases} \tag{4.7}$$

The functional form of both $\mu_z(z)$ and $\mu_Q(Q)$ are plotted on Fig. 4.2a and b, respectively. The critical values ζ_c^l and ζ_c^u in Fig. 4.2a can be defined as the depth at which liquid water cannot exits in the fluids and may be estimated for each volcanic region considering the most likely geothermal gradient for the zone. The two critical values Q_{c1} and Q_{c2} (and corresponding uncertainty intervals) in Fig. 4.2b have been selected (considering the results of *Kumagai and Chouet* (2000)) as $Q_{c1} \approx 10, \pm 10\%$ and $Q_{c2} \approx 100, \pm 10\%$.

In practice, for a given LP data analyzed, input values z_j and Q_j are provided; then, the degree of membership may be defined as $\text{DOM}_z = \mu_z(z_j)$, and $\text{DOM}_Q = \mu_Q(Q_j)$ for both z and Q sets, respectively. The inputs are combined logically using the conjunction ($\wedge$) operator to produce output values of the combined parameter $\mathbf{P}_{(z,Q)} = \{\mu_z(z_j) \wedge \mu_Q(Q_j)\}$. In the Fuzzy logic framework, the conjunction operation may be performed using two approaches: (1) the Min/Max-norm, in which $\mathbf{P}_{\{z_j \wedge Q_j\}} = \min\left[\mu_z(z_j), \mu_Q(Q_j)\right]$ (e.g., *Mamdani*, 1974), implying the most significant variable influence the results; and (2) the Product/Bounded-sum, in which $\mathbf{P}_{\{z_j \wedge Q_j\}} = \left[\mu_z(z_j) \cdot \mu_Q(Q_j)\right]$ (e.g., *Mizumoto*, 1995)), implying that all input variables can influence the results. Fig. 4.3 shows bidimensional contour plots of $\mathbf{P}_{\{z_j \wedge Q_j\}}$ generated using both Min/Max-norm (Fig. 4.3a) and Product/Bounded-sum (Fig. 4.3b).

$\mathbf{P}_{\{z_j \wedge Q_j\}}$ is a function valued in the real interval $[0, 1]$ that maps all possible input values (measures) of Q and z. Events with values of z and Q for which $\mathbf{P}_{\{z_j \wedge Q_j\}}$ is close or equal to one are events likely belonging to the set of interest (in our case, magmatic or magmatic-induced LP activity). To integrate $\mathbf{P}_{\{z_j \wedge Q_j\}}$ as a monitoring parameter in BET (see electronic supplementary material in *Marzocchi et al.* (2008), and Appendix A for details), one or two thresholds (i.e. t, or t_l and t_u) should be selected in order to measure and quantify the information provided by the parameter; for instance, BET uses t_l and t_u to construct a membership function $\mu^{\{LP\}}$ and quantify the degree of anomaly indicated by the parameter. For example, Fig. 4.4 is a representation of the membership function $\mu^{\{LP\}}$ in BET, in which $t_l = 0.8$ and $t_u = 0.95$.

In a generic example, we apply this scheme described to the LP solutions obtained in Chapter 2 for the LP events of Miyakejima volcano. We did combine the Q solutions plotted in Fig. 2.10 with the source locations (depth) obtained and provided by the National Research Institute for Earth Sciences and Disaster Prevention (NIED), Japan. The results obtained are plotted in Fig. 4.5; for this analysis, the critical depth ζ_c has been set (subjectively) to 2km. No references were found for the geothermal gradient in Miyakejima; the only information found for reference comes from the magnetic measurements of *Sasai et al.* (e.g., 2001), who
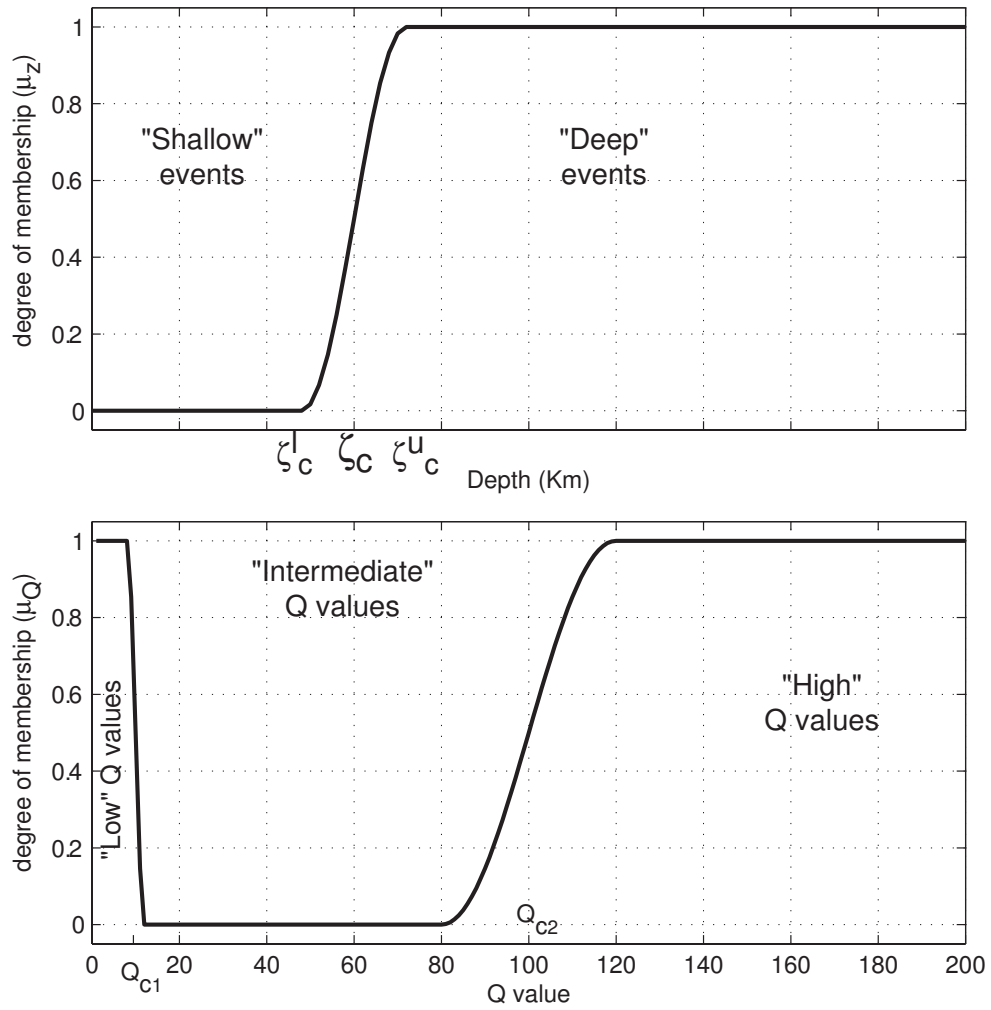
Figure 4.2: Membership functions of both Depth (z) and Q-values, to measure the degree of information bearing about magmatic unrest.
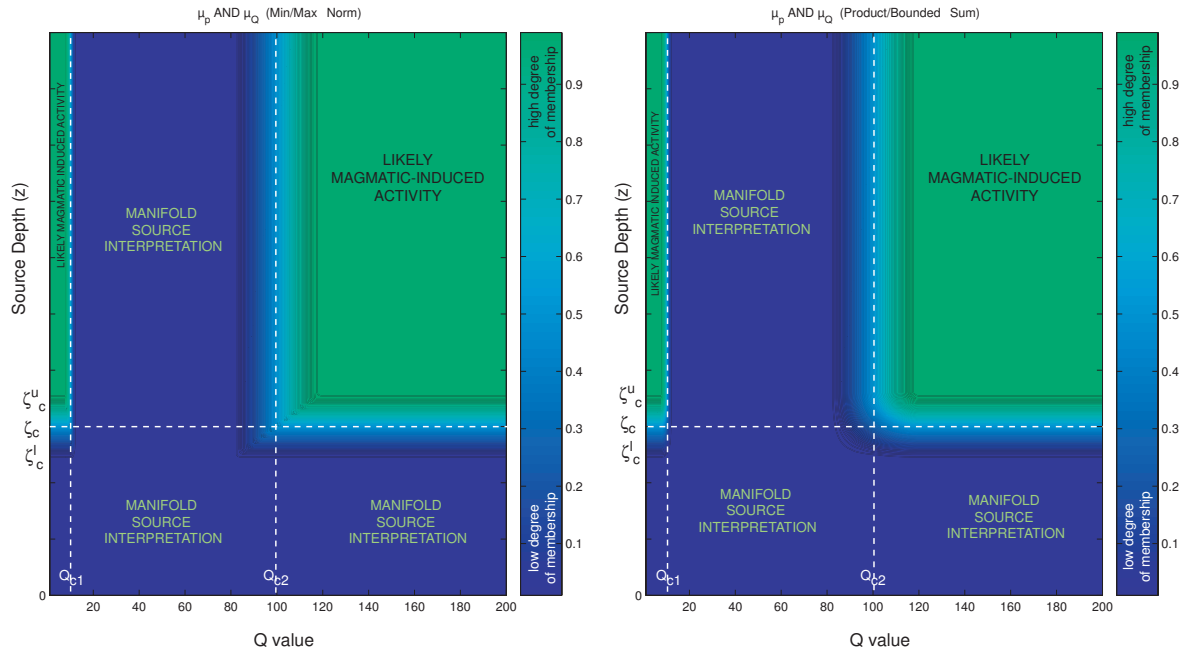
Figure 4.3: Bidimensional plot of the conjunction $\{\mu_z \wedge \mu_Q\}$ using (a) the Min/Max Norm, and (b) the Product/Bounded Sum; green zones (light colors) show areas where the resulting parameter ($\mathbf{P}_{\{\mu_z \wedge \mu_Q\}}$) is informative for magmatic-induced activity.

detected a shallow thermally demagnetized area before the 2000 eruption. Fig. 4.5a shows the values of the membership function μ_z mapping the depth information; Fig. 4.5b shows the values of μ_Q obtained using the Q-values obtained (Chapter 2, section 2.8); and Fig. 4.5c shows the values of $\mathbf{P}_{(z, Q)}$ after the conjunction $\{\mu_z \wedge \mu_Q\}$. In a BET_EF scheme, the values $\mathbf{P}_{(z, Q)}$ could be used as monitoring parameter which BET would integrate, for example, through a membership function (e.g. $\mu^{\{LP\}}$) as that shown in Fig. 4.4. To obtain a single value of the parameter for BET (in a given time window), values of $\mathbf{P}_{(z, Q)}$ may be grouped for fixed time intervals, and the empirical CDF of the values in the time interval may be obtained; then, a selected percentile (e.g. 95th percentile) may be defined to represent the solutions of that time interval. For example, Fig. 4.5d shows the 95th percentile of the CDF of the $\mathbf{P}_{(z, Q)}$ values obtained in the Miyakejima example, grouped in time intervals of two days.

4.4   Potential use of non-stationary ETAS modeling of VT swarms as input monitoring parameter in BET

In Chapter 3 we have applied a general stochastic non-stationary ETAS modeling to analyze complex seismic swarms in both tectonic and volcanic environments, in order to extract information useful to characterize the swarms directly associated with migration of magmatic dikes. In our analysis we compare the time evolution of both background seismicity λ_0 and p-value (as suggested by *Lombardi et al.*, 2006) for different seismic swarms grouped in 3
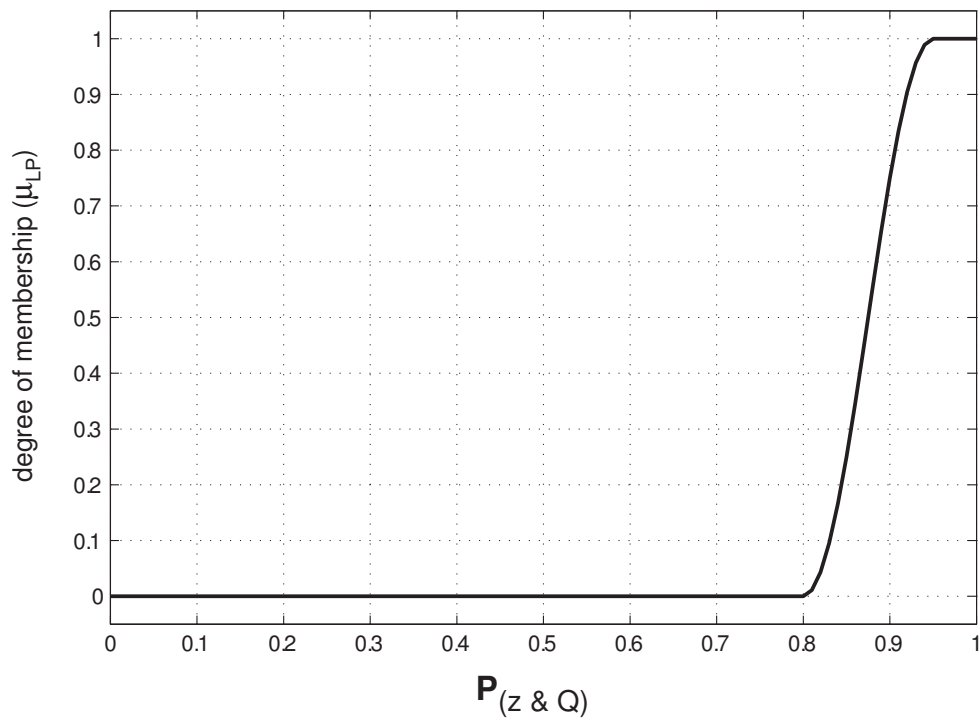
Figure 4.4: Membership function $\mu^{\{LP\}}$ in BET, to consider both the depth (z) of the source and the spectral analysis of LP events (in terms of f, growth rate, and Q) as parameters for short-term eruption forecasting.
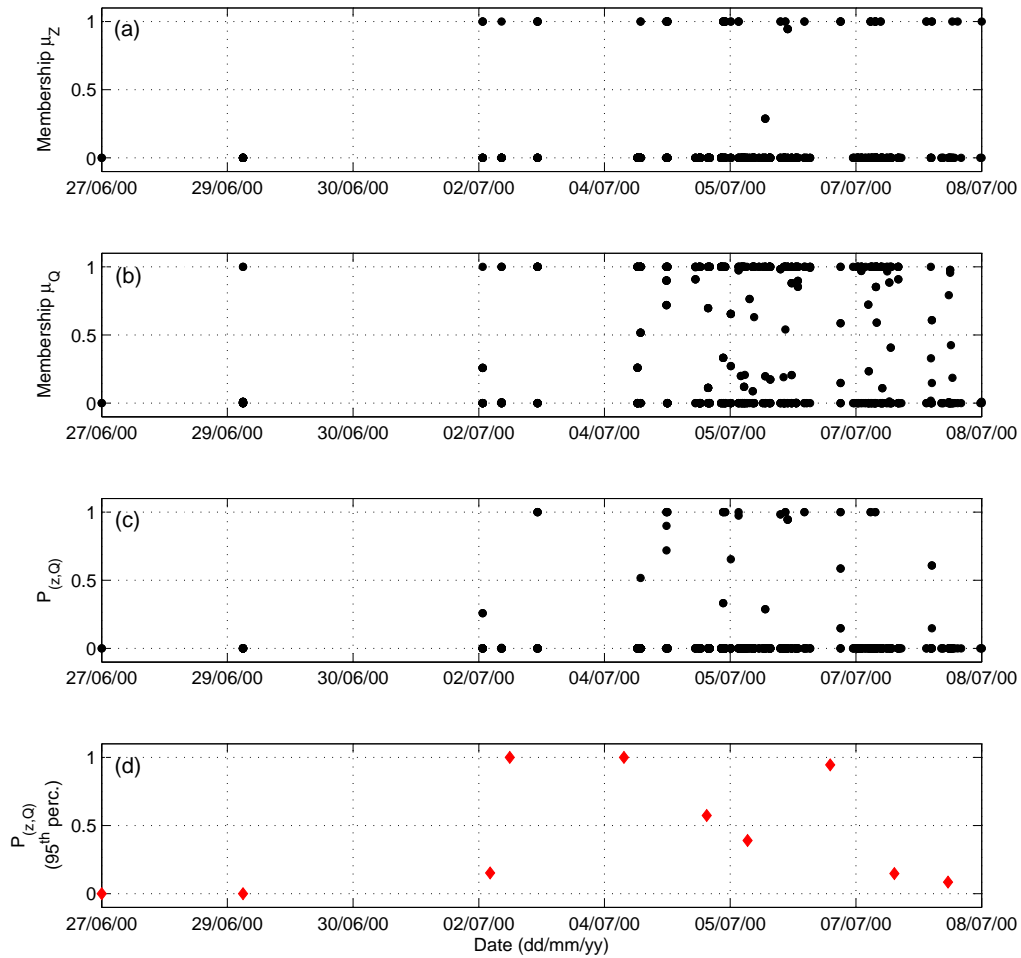
Figure 4.5: Membership function $\mu^{\{LP\}}$ evaluated for the LP events from Miyakejima volcano, Japan, analyzed in Chapter 2. (a) Membership function for the depth, μ_z; (b) Membership function for the Q-values, μ_Q; (c) values of $\mathbf{P}_{(z,\,Q)}$ after the conjunction $\{\mu_z \wedge \mu_Q\}$; (d) 95th percentile of the empirical CDF of the solutions obtained, grouped in intervals of two days.

cases: (1) seismic swarms in tectonic areas, (2) VT swarms in non-erupting volcanic areas, and (3) VT swarms in volcanic areas before and/or during an eruptive process.

The obtained results of the stochastic non-stationary ETAS model can give us some insights on the mechanics of the seismic swarms on volcanic areas. As discussed on Chapter 3, the overall impression is that changes in the background seismicity (λ_0) may be observed in all the cases, even in tectonic environments. However, the p-value have some substantial differences that may be used as a tool to characterize VT swarms. The results obtained in Chapter 3 may be summarized as: (1) the tectonic sequences analyzed show in general a rather stationary value of p, which in general is ≤ 1.2 in the analyzed cases; (2) the VT sequences in non-erupting volcanic areas have some similar features as the tectonic cases, with stable values or showing small, high and low frequency oscillations. In the analyzed cases the p-value was always $p \leq 1.3$; (3) the VT swarms during eruptive episodes have p-values that in general are higher respect to the cases described before, being in general $p \geq 1.7$, and normally positively correlated to increased background seismicity rates λ_0.

The results of the stochastic non-stationary ETAS model yield important clues to interpret the temporal evolution of complex seismic swarms. The phenomenological description of the temporal behavior of λ_0 and p parameters may be accompanied by a more detailed interpretation of the results in terms of the physical processes at the source of the swarms (for details see Chapter 3). We can use this information to construct a monitoring parameter to feed the monitoring part of node 2 (magmatic unrest) in BET. Again as in section 4.3, Fuzzy logic may provide a valid framework to build both a decision scheme and a quantitative parameter to translate the information of non-stationary ETAS modeling in a monitoring tool to feed probabilistic eruption forecasting models as BET (e.g., *Marzocchi et al.*, 2004, 2008). Based on the results summarized in the previous paragraph, we can conclude that to detect swarms associated with magma migration, increased values of λ_0 is a not informative parameter; just p-values may bear some information: for instance, VT swarms directly associated with shallow dike intrusions are directly correlated with high p values; in the cases analyzed in Chapter 3 we have found that p-values < 1.3 were associated with both tectonic and volcano-tectonic seismic swarms, the last without eruptive activity associated, while p-values > 1.7 were found for VT sequences just before and/or during eruptive episodes. Then, to construct a membership function we can use those values as reference for the anomaly thresholds: for example, we select $t_l = 1.5$ and $t_u = 1.7$ as lower and upper thresholds, respectively, and then, the membership function $\mu^{\{\text{N-ETAS}\}}$ in BET (Fig. 4.6) would be of the form (see electronic supplementary material in *Marzocchi et al.* (2008), and Appendix A for details):

$$\mu^{\{\text{N-ETAS}\}}(p\text{-value}) = \begin{cases} 0 & \text{if } p\text{-value} < 1.5 \\ \frac{1}{2}\left\{\sin\left[\pi\left(\frac{p\text{-value}-1.5}{0.2}\right) - \frac{\pi}{2}\right] + 1\right\} & \text{if } 1.5 \leq p\text{-value} < 1.7 \\ 1 & \text{if } p\text{-value} \geq 1.7 \end{cases} \quad (4.8)$$
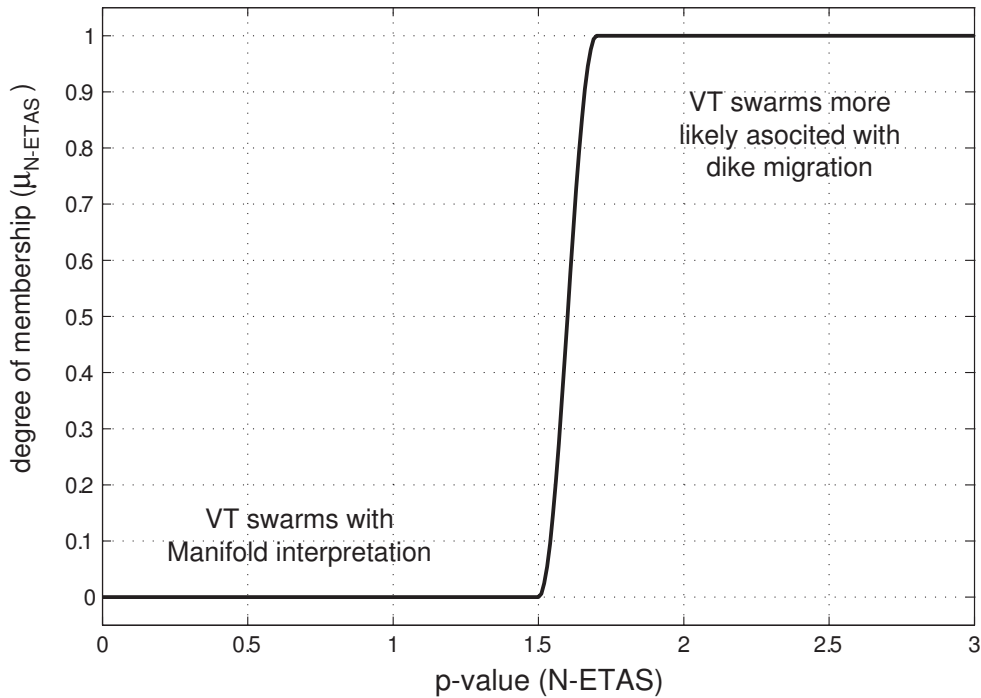
Figure 4.6: Membership function $\mu^{\{\text{N-ETAS}\}}$ in BET, to consider the non-stationary ETAS modeling as monitoring parameter for short-term eruption forecasting.

4.5 Guidelines for quantitative determination of seismicity-rate thresholds for BET

Volcanoes with permanent monitoring systems have often databases containing information of many years of geophysical observations in which a record of the *background* activity level (i.e. repose periods) may be determined. The definition of anomaly thresholds may be quite controversial and even non univocal values may be selected. In this section we analyze the case of background seismicity-rate databases, and try to define some simple guidelines to quantitatively establish the thresholds required in BET.

As a general approach to quantitatively define the thresholds (1 or 2) required in BET, the next procedure is proposed:

1. Select the data from repose periods in the data set (background seismic activity);

2. Divide the background activity in time intervals of length τ, and create a data set of number of events/τ;

3. Fit the best statistical distribution describing the data set, and test the hypothesis (e.g. using a 1 sample Kolmogorov-Smirnov (K.S.) test);

4. use the Cumulative Distribution Function (CDF) of the fitted distribution to select the

required threshold values (t_l and t_u); this may be done selecting (subjectively) two percentiles: in our applications we consider the 90th and 95th percentiles of the CDF. Note that if $t_l = t_u$, a boolean logic is reproduced.

In the case of seismicity catalogs, one of the probability distributions that better fits seismicity rate data is the Poisson distribution; however, volcano seismicity (VT, or LP) is generally clustered and then the use of a Generalized Poisson model may provide a more accurate description of the observed data.

4.5.1   Brief introduction to the Generalized Poisson Process (GPP)

The Generalized Poisson model assumes that the events occur grouped in time (clusters) and that the number of clusters follows a Poisson distribution. In other words, this approach intends to solve the problem of estimating a cluster size distribution, assuming some model for the occurrence of these clusters, to derive the distribution of the number of events in a time interval in terms of the cluster size distribution.

The GPP distribution is characterized by two parameters: K and E (for details, see *Shlien and Toksoz* (1970); *Zollo et al.* (2002). It is assumed that the clusters are distributed in time according to a Poisson process with (mean) rate K, while E is a free parameter which describes the degree of clustering of the sequence.

The GPP distribution parameters are derived by maximizing the likelihood ratio l:

$$l = W^W \prod_{i=1}^{W} \frac{p_q(n_i, \tau)}{f(n_i, \tau)} \tag{4.9}$$

where $p_q(n_i, \tau)$ is a theoretical probability (see below), $f(n_i, \tau)$ is the observed frequency of occurrence on n_i events in the time window τ, n_i is the number of events in the i-th time window, and W is the total number of time windows.

The theoretical probability $p_q(n_i, \tau)$ is calculated with the recursive formula

$$p_q(n, \tau) = \frac{K\tau}{n} \sum_{j=0}^{n-1} (n - j) \, q \, (n - j) \, p_q \, (j, \tau) \tag{4.10}$$

starting with

$$p_q(0, \tau) = e^{-K\tau} \tag{4.11}$$

In equation (4.10), $q(n)$ is the **Z** distribution or discrete Pareto distribution; it represents the theoretical distribution of the number of grouped events (which is generally observed to have an inverse power form) and is defined as

$$q(n) = \frac{1}{\zeta(E)} n^{-E} \tag{4.12}$$

where E describes the degree of clustering of the sequence, and $\zeta(E)$ is the Zeta-Riemann function, a normalization factor.

To find the parameters K and E of the GPP, we created a code to perform a grid search in the $\{E, \ K\}$ space to find the maximum value of the likelihood defined in equation 4.9. Then, using the estimated values of K and E, the generalized Poisson distribution is constructed and tested using a K.S. 1-sample test.

4.5.2 Example 1: Modeling the distribution of the monthly number of VT events during repose periods at Cotopaxi volcano, and threshold determination

The determination of the theoretical (generalized Poisson) distribution for the VT catalog is used to define the thresholds necessary to feed the BET (generally at node 1 or 2). The analysis was performed using the events of the background activity grouped in intervals of $\tau = 1$ month; the background activity was select using a Change Point analysis (e.g., *Mulargia et al.*, 1985, 1987). The time series of monthly number of VT events (Fig. 4.7a), is successfully described by a generalized Poisson distribution with parameters (Fig. 4.8):

$$
\begin{aligned}
K &= 7.8 \ (\pm 0.3) \\
E &= 2.5 \ (\pm 0.1)
\end{aligned}
\tag{4.13}
$$

$K = 7.8$ means that the mean rate of clusters of VT events is about 7.8 cluster/month. The K.S. test result is that the null hypothesis (data fits the theoretical distribution) cannot be rejected, at a significance level of 0.01.

The next step is to define the *thresholds* (one or two values) of number of events to be used as the limits to determine if the seismicity is indicating a rest/unrest period. These threshold values help BET in determining if a selected monitoring measure (in this case VT seismicity) can be considered anomalous and are used using the Fuzzy set theory (*Zadeh*, 1965) that quantitatively emulates the expert-like flexible judgment of the anomalous state of a monitoring parameter (for details see *Marzocchi et al.* (2008)). For this reason, it is necessary to define two thresholds for each monitoring parameter: a lower t_l and a upper t_u threshold. Note that if $t_l = t_u$, a boolean login is reproduced.

The thresholds are selected as the 90th and 95th percentiles of the distribution for t_l and t_u, respectively. This means that we consider that under t_l the parameter indicates *non-unrest*; over t_u it indicates *unrest*, and for values between t_l and t_u there is a relative degree of anomaly calculated by a membership function $\mu^{\{\text{VT-rate}\}}$ (for details see *Marzocchi et al.* (2008), and Appendix A).
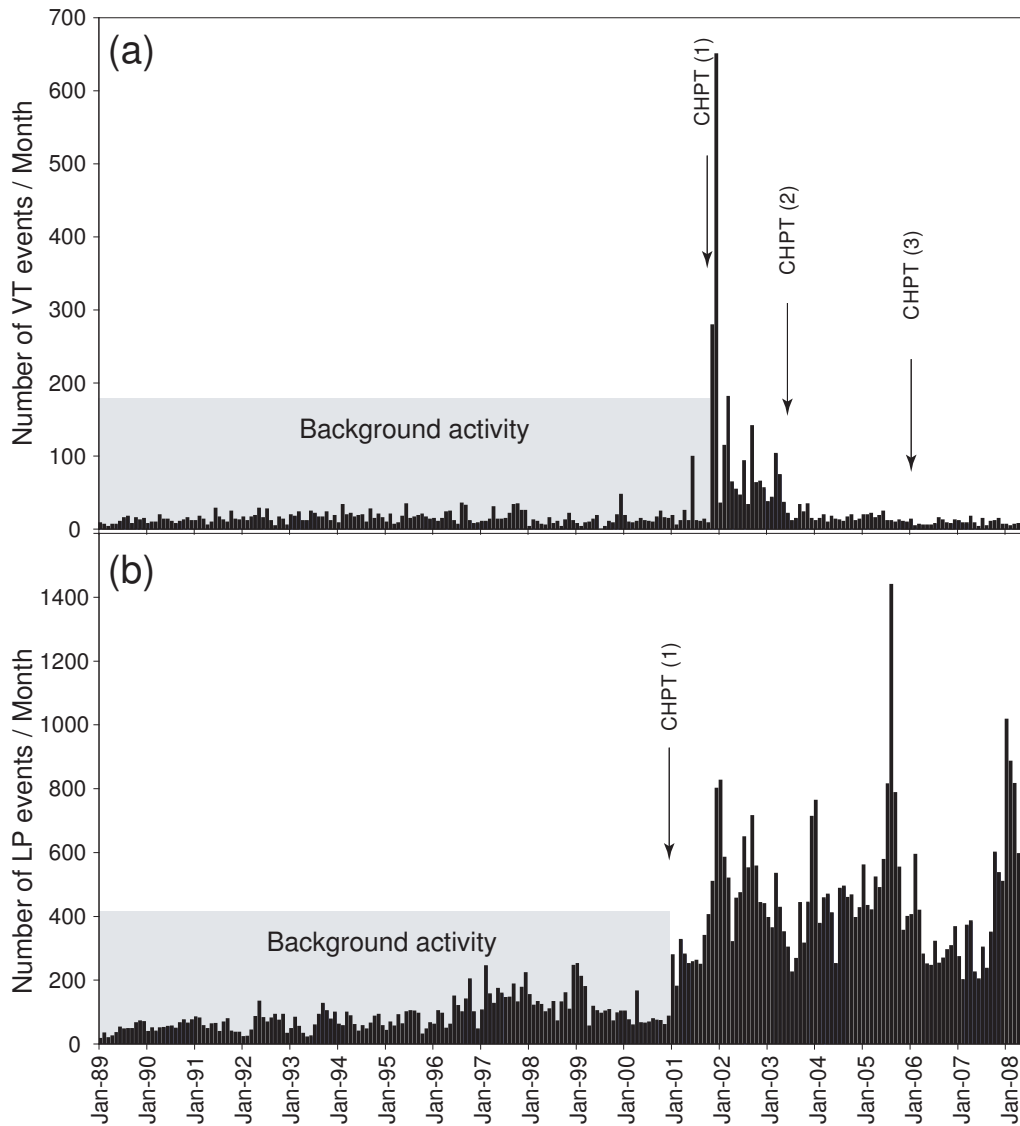
Figure 4.7: Monthly VT (a) and LP (b) seismicity at Cotopaxi volcano, Ecuador, from January 1989 to April 2008 (data from Instituto Geofísico - EPN, Quito, Ecuador), and Change Point analysis to identify background activity before the unrest started on January 2001.
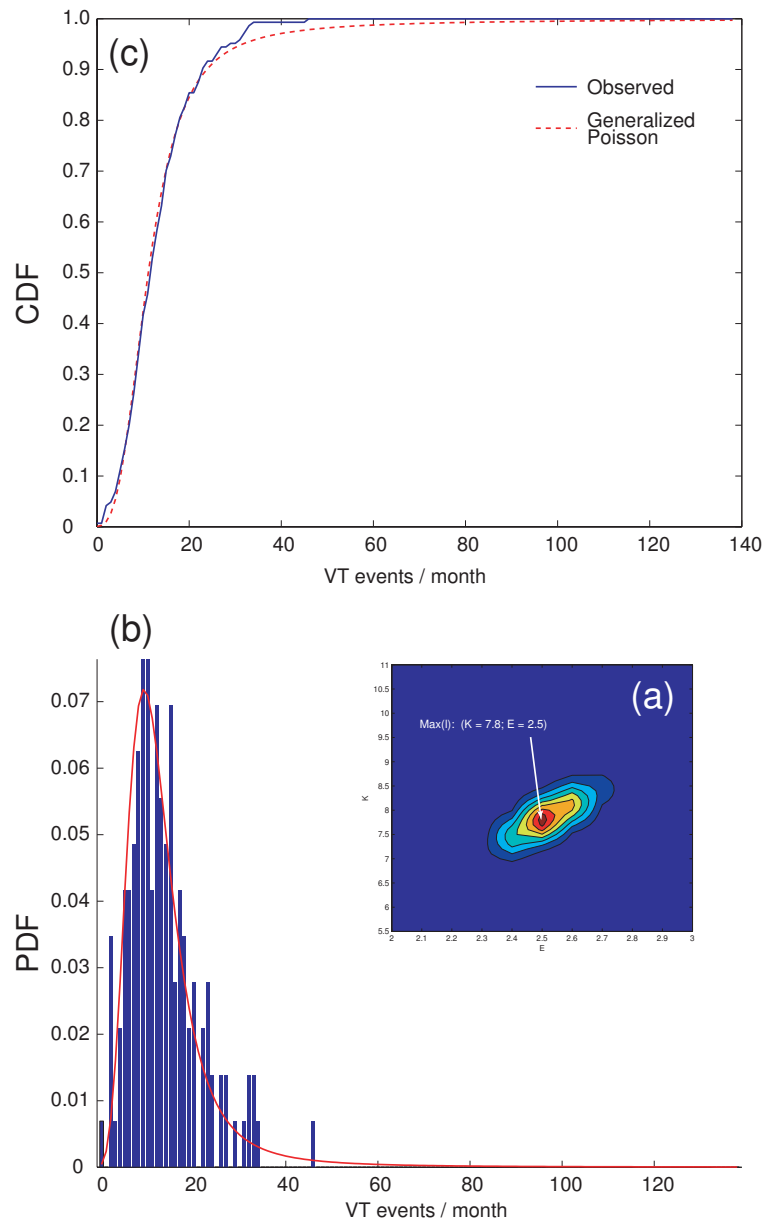
Figure 4.8: Generalized Poisson distribution describing VT background seismicity at Cotopaxi volcano, Ecuador; (a) grid search for model parameters, (b) probability density function (PDF); and (c) cumulative distribution function (CDF).

For the VT events, the thresholds (obtained through random sampling) defined by the 90th and 95th percentiles of the distribution are:

$$\begin{aligned} t_l &= 24 \text{ VT events/month} \\ t_u &= 32 \text{ VT events/month} \end{aligned}$$

$$(4.14)$$

4.5.3 Example 2: Modeling the distribution of the monthly number of LP events during the repose periods at Cotopaxi volcano, and threshold determination

The determination of the theoretical (generalized Poisson) distribution of the LP seismicity is used to define the thresholds which will be used in BET (generally at node 1). The time series of monthly number of LP events (Fig. 4.7b) during the period previous to the start of the unrest defined, is successfully described by a generalized Poisson distribution with parameters (Fig. 4.9):

$$\begin{aligned} K &= 23.0 \ (\pm 1.0) \\ E &= 2.0 \ (\pm 0.1) \end{aligned}$$

$$(4.15)$$

$K = 23$ means that the mean rate of clusters of LP events is about 23 cluster/month. The K.S. test result is that the null hypothesis (data fits the theoretical distribution) cannot be rejected, with a significance level of 0.01. The next step is to define the *thresholds* (two values: t_l and t_u) of number of events to be used as the limits to determine if the seismicity is indicating a rest/unrest period.

The thresholds are selected as the 90th and 95th percentiles of the distribution for t_l and t_u, respectively, as in the previous example. For the LP events, the thresholds (obtained through random sampling) defined by the 90th and 95th percentiles of the distribution are:

$$\begin{aligned} t_l &= 205 \text{ LP events/month} \\ t_u &= 335 \text{ LP events/month} \end{aligned}$$

$$(4.16)$$
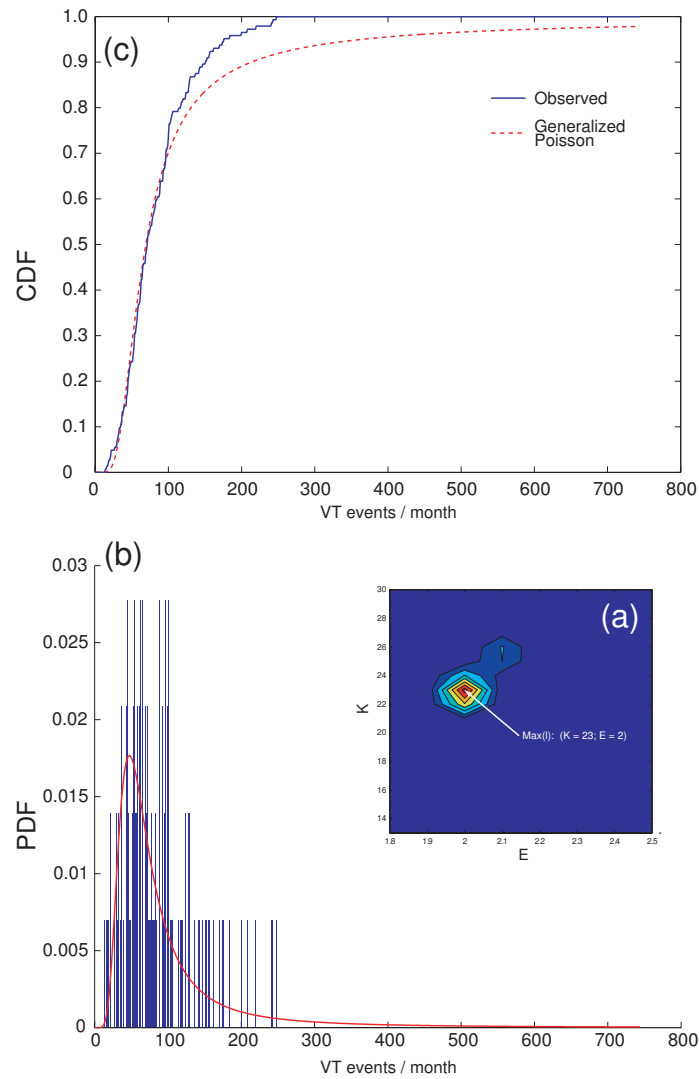
Figure 4.9: Generalized Poisson distribution describing LP background seismicity at Cotopaxi volcano, Ecuador; (a) grid search for model parameters, (b) probability density function (PDF); and (c) cumulative distribution function (CDF).

Chapter 5

General Conclusions

This Ph.D. thesis explores the past eruptive activity and the pre-eruptive phase of volcanoes, using both empirical and physical methods. Such a modeling provides new insights on understanding the physics of pre- and sin-eruptive processes and it can be used to improve short- and long-term eruption forecasting (EF) assessments. For this purpose, the results obtained have been integrated in a homogeneous scheme for EF provided by the Bayesian Event Tree model for Eruption Forecasting (BET EF, e.g., *Marzocchi et al.*, 2008). Specifically, guidelines to use our results to improve the setup of BET EF have been proposed, highlighting how new models may be successfully integrated in BET for a comprehensive analysis of all the available information existing for a volcano.

Beyond this general objective achieved, the most remarkable concluding remarks from each specific analysis performed can be summarized as follows:

1. For long-term eruption forecasting:

   (a) The intrinsic complexity of volcanic systems motivates the definition of probability models as mathematical structures to describe the response of the considered systems. This kind of models may be the basis for time-dependent or time-independent, long-term, eruption forecasting models based on the past eruptive activity of a volcano.

   (b) Renewal processes characterized by five different probabilistic models, plus a TPM and a SPM, have been applied to analyze the repose times between eruptive episodes of Miyakejima volcano. From our analysis we conclude that the probabilistic model that better explains the observed data is a Brownian passage-time (BPT). This model is based upon a simple physical model resembling the characteristic cycles of volcanic activity, the Brownian relaxation oscillator, and is parameterized by the mean rate of event occurrence, μ, and the aperiodicity about the mean, α.

   (c) The Brownian passage-time family differs from other usual candidate distributions for long-term eruption forecasting in that it may reflect physical properties of the

macro-mechanics of volcanic processes. The Brownian relaxation oscillator and Brownian passage-time distribution connect together physical notions of unobservable loading and failure processes of a point process with observable response statistics (i.e. event recurrence in time).

(d) In the BPT model, μ is an scale parameter that provides information about the typical frequency at which events occur, and α is a dimensionless parameter that measures the aperiodicity of the mean response of the system, and for this reason this parameter may be useful to compare the behavior of different volcanoes, spanning from periodic-like to Poisson-like systems; for instance, the more periodic the process, the more α approaches zero. The value $\alpha = 0.51 \pm 0.01$ found in this work for the aperiodicity in Miyakejima volcano indicates probably for the first time a clear recurrent behavior in a volcanic system.

(e) The BPT model provides some insights for time-dependent, long-term eruption forecasting; for instance, if we consider its hazard function, some noteworthy properties can be defined: the probability of having renewed eruptive activity just after an eruptive cycle is very low, then it increases steadily from about zero to a finite maximum near the mean recurrence time. Finally, for times greater than the mean recurrence time the hazard function tends to a finite constant value, indicating that for long repose times the system tends to behave as a Poisson process. These are unique properties among the set of candidate models considered, and provide a more realistic asymptotic behavior of the failure rate.

2. For short-term eruption forecasting:

(a) Volcano Seismology can provide useful monitoring parameters to feed probabilistic models for short-term eruption forecasting. Both long-period (LP) and volcano-tectonic (VT) seismic events may be analyzed to track some physical processes inside active volcanoes.

(b) Spectral analysis of LP events may provide information about the source properties of this kind of seismic events often observed on active volcanoes. We propose a probabilistic formulation of the physically-based, spectral analysis method of *Kumazawa et al.* (1990), which leads to the definition of a (posterior) probability distribution in the model space $\sigma(\mathbf{m})$ that results of combining prior information (on data and model parameters) with the new information obtained by measurements (the time series).

(c) We implemented a Monte Carlo method to explore $\sigma(\mathbf{m})$, in which a large collection of models are pseudo-randomly generated with the support of the posterior probability distribution. To do this task we have implemented a simple Metropolis algorithm; the input to the algorithm consists of random models generated according to the prior distribution $\rho(\mathbf{m})$ and the corresponding values of the likelihood function (that carries information from measurements and the theory, linking

observed data and model parameters); the output from the algorithm are pseudo-random realizations of the posterior distribution $\sigma(\mathbf{m})$ that contains all information about the parameterized model. From this output we get information about model parameters: even if the most general solution would be the definition of an empirical probability distribution (CDF or PDF) for each model parameter, it is possible (where possible) to compute any kind of central and dispersion estimators to produce both best-guess values and associated uncertainties; it is also possible to adopt cluster analysis techniques to assess the nature of the possible solution(s).

(d) We have set the problem in a fully probabilistic formulation, providing a framework in which restrictive assumptions are avoided and not-unique solutions are possible. We consider that this kind of formulation may be helpful for different kind of problems, specially when it exists a nonlinear relationship between the observed data and the parameters of the model.

(e) We did apply this method to analyze LP waveforms from two volcanoes: Cotopaxi (Ecuador) and Miyakejima (Japan). We did perform the analysis to each single LP event (617 for Cotopaxi, and 240 for Miyakejima) and the solutions for the optimum number of wave elements where extracted and stacked for different time intervals. This procedure may be an important tool for objectively monitor the characteristic complex frequencies of the LP events; in this way, changes in the volcanic activity may be highlighted and families of events of particular interest could be identified for supplementary analysis. Further this possible application for monitoring, our approach can also be used to analyze individual events as an alternative to the existent methodologies.

(f) We perform a retrospective analysis of volcano-tectonic seismic swarms using a non-stationary, epidemic-type, aftershock sequences (ETAS) model in which both the background seismicity and the p-value can vary through time, which has evidenced some important characteristics of the VT swarms associated with magmatic intrusions. In particular, swarms of VT events that precede or accompany magmatic intrusions (i.e. before or during eruptive activity) present remarkably high p-values respect to both, VT swarms in cases with no evident shallow magmatic activity, and swarms in purely tectonic environments. From a practical point of view, the stochastic modeling of VT swarms using a non-stationary ETAS model may be used for a quick and almost automatic analysis of VT swarms to identify and discriminate those which are more likely associated with the physical response of the volcanic edifice to a magmatic intrusion.

(g) Fuzzy logic, provides a valid framework to quantitatively consider information from geophysical parameters whose physical interpretation may be manifold, or in which no clear or sharp anomaly thresholds may be defined. We did create monitoring parameters based on the analysis performed on both LP and VT events, which may be straightforward integrated into the BET model.

Appendix A

Brief introduction to the Bayesian Event Tree Model for Probabilistic Eruption Forecasting (BET_EF)

As for any generic complex system (intrinsically unpredictable from a deterministic point of view), the description of eruptive processes in terms of probability is particularly a suitable procedure. Beyond the extreme complexity, nonlinearities, and the large number of degrees of freedom of a volcanic system (aleatory uncertainty), also our still limited knowledge of the processes involved (epistemic uncertainty) make deterministic prediction of the evolution of volcanic processes practically impossible (*Marzocchi et al.*, 2008).

Bayesian statistics is a suitable framework for producing an EF in a rational, probabilistic form. The Bayesian approach (BET) used by *Marzocchi et al.* (2004, 2008) starts from modeling the statistical distribution using our basic knowledge (or complete ignorance), and then it refines the distribution as long as new information come in as in a sort of data assimilation procedure *Marzocchi et al.* (2008).

Basically, BET translates volcanological input (it means any kind of relevant information such as theoretical/empirical models, geological and historical data, and/or monitoring information, etc.) into probability of any possible volcano-related event. Such probabilities represent an homogeneous and quantitative synthesis of the present knowledge about the volcano.

A.0.4 Event Tree Scheme

BET is based on a branching scheme in which individual branches are alternative steps from a general prior event, state, or condition, and which evolve into increasingly specific subsequent events (intermediate outcomes) up to a final outcome. This is called an event tree (ET), and in this way it shows all relevant possible outcomes of volcanic unrest at progressively higher degrees of detail. The points on the graph where new branches are created are

referred to as nodes (*Newhall and Hoblitt*, 2002; *Marzocchi et al.*, 2004, 2008). At each node we have the following states:

- **Node 1**: (*Unrest / No unrest*) in the time interval $(t_0, t_0 + \tau]$, where t_0 is the present time and τ is the time window considered.

- **Node 2**: (*Origin*) the Unrest is due to *magma* or *other causes* (e.g. hydrothermal, tectonics, etc.), given that an unrest has been detected.

- **Node 3**: (*Outcome*) the magma will reach the surface (i.e. the volcano will erupt), or not, in the time interval τ, provided that the unrest has a magmatic origin.

- **Node 4**: (*Location*) the eruption will occur in a specific location (e.g. crater, a flank, etc.), provided that there is an eruption.

- **Node 5**: (*Magnitude*) the eruption will be of a certain magnitude/size (e.g. VEI), provided that there is an eruption in a certain location.

For a further quantification of volcanic hazard (VH) there exist other nodes:

- **Node 6**: (*Phenomena*) the occurrence of a specific threatening event (pyroclastic flows, lahars, tephra fall, lava flows, etc.), provided that there is an eruption of certain size and in a certain location.

- **Node 7**: (*Area*) the threatening event reaches an specific area, given that the phenomenon has occurred.

- **Node 8**: (*Overcoming threshold*) the overcoming of a threshold related to a certain threatening event in a certain area, given the threatening event has reached this area.

- **Node 9**: (*Exposure*) an individual or a building will be present at the specified sector reached by a phenomenon.

- **Node 10**: (*Vulnerability*) a hypothetical individual will be killed/injured or a building will be destroyed/damaged at the specific sector, given that the phenomenon arrives there and that persons and/or buildings are present.

At each of these nodes a probability density function (PDF) is assigned; the use of these probability functions (characteristic of the Bayesian approach) allows BET to estimate aleatory and epistemic uncertainties.

θ_k^j is the probability of the conditional event j at the k-th node, then for each one of the nodes we have: $[\theta_1^{(unrest)}]$, $[\theta_2^{(magma)}]$, $[\theta_3^{(eruption)}]$, $[\theta_4^{(location)}]$, $[\theta_5^{(size)}]$, $[\theta_6^{(threat)}]$, $[\theta_7^{(area)}]$, $[\theta_8^{(threshold)}]$, where the square brackets stand for a generic PDF.

Given all the PDF at each node, BET combines them in order to obtain the absolute probability of each event which we are interested in. For example, the PDF ϕ_1 of the probability of having an eruption of type m in the time interval $(t_0, t_0+\tau]$ at the n-th vent location would be:

$$[\phi_1] = [\theta_1^{(unrest)}] \, [\theta_2^{(magma)}] \, [\theta_3^{(eruption)}] \, [\theta_4^{(n)}] \, [\theta_5^{(m)}]$$

The functional form of $[\phi]$ is not determined analytically, but through a Monte Carlo simulation (*Marzocchi et al.*, 2008). Each PDF is sampled 1000 times and the calculation is performed using each sample, obtaining 1000 values of $[\phi]$; then, using these values, a functional form of $[\phi]$ is determined numerically. In this way, both aleatory and epistemic uncertainties are properly propagated at all nodes and the best guess (i.e. the average) and errors (the standard deviation) of the absolute probability (of any possible vent) are estimated (*Marzocchi et al.*, 2008).

A.0.5   Estimating the probabilities at the nodes: (1) General aspects

To estimate the PDF at each node is the main technical problem in BET but the most important part of the procedure since it translates volcanological information into probability. Estimating a value for a probability always involves a subjective element; for this reason a Bayesian approach, in which a probability distribution is associated with each event of a node, is adopted (rather than a single value). In this way it is possible to estimate uncertainties of these event probabilities (*Marzocchi et al.*, 2004).

The distribution $[\theta_k^j]$ has to be *unimodal* and with domain [0,1] (since the random variable is a probability). A suitable distribution with these requirements is the *Beta* distribution. The PDF associated with θ_k^j is calculated following three steps: (1) design of a generic Bayesian Event Tree, (2) estimate the conditional probability at each node, and (3) combine the probabilities of each node to obtain probability distribution of any relevant event (*Marzocchi et al.*,

2004, 2008).

The Dirichlet distribution and the Beta distribution

In the Bayesian model of *Marzocchi et al.* (2004, 2008), it is assumed that $[\theta_k]$ is a random variable with *Dirichlet distribution*:

$$[\theta_k] = \frac{\Gamma(\alpha_1 + \ldots + \alpha_m)}{\Gamma(\alpha_1) \ldots \Gamma(\alpha_m)} \left[\theta_k^{(1)}\right]^{\alpha_1 - 1} \ldots \left[\theta_k^{(m)}\right]^{\alpha_m - 1} \tag{A.1}$$

where $\alpha_j > 0$ $(j = 1, 2, \ldots, m)$, $\theta_k^{(1)}, \ldots, \theta_k^{(m)} > 0$, $\sum_{j=1}^m \theta_k^{(j)} = 1$, and m is the number of possible mutually exclusive and exhaustive events.

The first two moments of the distribution (mean and variance) are:

$$E\left[\theta_k^{(j)}\right] = \frac{\alpha_j}{\alpha_0} \tag{A.2}$$

$$Var\left[\theta_k^{(j)}\right] = \frac{\alpha_j(\alpha_0 - \alpha_j)}{\alpha_0^2(\alpha_0 + 1)} \tag{A.3}$$

where $\alpha_0 = \sum_{j=1}^k \alpha_j$.

In the case $m = 2$ and for events mutually exclusive (e.g. magma or not), the *Dirichlet distribution* becomes a *Beta distribution*:

$$[\theta_k] = Beta(\alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta_k^{\alpha - 1} (1 - \theta_k)^{\beta - 1} \tag{A.4}$$

where $\alpha, \beta > 0$, and a sufficient condition to have a finite PDF is $\alpha, \beta \geq 1$ (e.g., *Gelman et al.*, 1995; *Marzocchi et al.*, 2004).

Often (but not necessarily) α and β are integers related to the number of available data: in this case, $\alpha - 1$ is the number of successes and $\beta - 1$ is the number of failures.

The first two moments of the Beta distribution are:

$$E\left[\theta_k\right] = \frac{\alpha}{\alpha + \beta} \tag{A.5}$$

$$E\left[\theta_k^2\right] = Var\left[\theta_k\right] = \frac{\alpha\beta}{(\alpha+\beta)^2 \, (\alpha+\beta+1)}$$

$$= E\left[\theta_k\right](1 - E\left[\theta_k\right]) \frac{1}{\alpha+\beta+1} \tag{A.6}$$

In general, the first moment of the Beta distribution represents an estimation of the aleatory uncertainty, i.e., the intrinsic random variability due to the complexity of the process. The dispersion around the central value (i.e. the variance) instead represents an estimation of the

epistemic uncertainty due to our limited knowledge of the process.

The Beta distribution has two important limiting cases. The case of $\alpha = \beta = 1$ represents the *Uniform distribution*, and in this case the mean and variance are $E[\theta_k] = 1/2$ and $Var[\theta_k] = 1/12$, respectively. On the other hand, when α (or β) $\rightarrow \infty$, then $Var[\theta_k] \rightarrow 0$ which means that the Beta distribution tends to a Dirac's δ distribution centered around the average. In practice, this means that when we have a large amount of information the epistemic uncertainty becomes negligible.

Accounting for monitoring and non-monitoring information

In order to perform Eruption Forecasting, there are two broad classes of information to consider: (1) measurements from monitoring (data set $\mathcal{M}$) and (2) all the other kinds of data/information (data set $\bar{\mathcal{M}}$). This subdivision is due because usually these two types of information have different weights in different states of a volcano; for example, monitoring data may be the most relevant information for short-term EF purposes, while non-monitoring information may be more useful for long-term EF (*Marzocchi et al.*, 2008). In BET the PDF of the j-th event at the generic k-th node $\left[\theta_k^{(j)}\right]$ is a linear combination of the probabilities based on these two types of knowledge:

$$\left[\theta_k^{(j)}\right] = \gamma_k \left[\theta_k^{(j)\mathcal{M}}\right] + (1 - \gamma_k) \left[\theta_k^{(j)\bar{\mathcal{M}}}\right] \tag{A.7}$$

where γ_k, a variable in the interval $[0, 1]$, sets the degree at which monitoring data control the posterior probabilities with respect to the non-monitoring part, $\left[\theta_k^{(j)\mathcal{M}}\right]$ and $\left[\theta_k^{(j)\bar{\mathcal{M}}}\right]$ have the same meaning as $\left[\theta_k^{(j)}\right]$ but they are defined using only monitoring information and all the other kind of information, respectively. Through γ_k BET switches dynamically from long-term (when volcano is found to be at rest) to short-term (during unrest) probabilities.

For the nodes where monitoring parameters are informative, γ_k is a function of the "state of information" η which in turn, is a *fuzzy* parameter (*Zadeh*, 1965) in the interval $[0, 1]$ that indicates the degree at which unrest is detected by the monitoring observations at $t = t_0$.

The monitoring term in equationA.7 is the leading term in short-term probability evaluation and is determined through Bayes' rule by combining estimated probabilities from monitoring measures at time t_0 and monitoring measurements from past episodes of unrest (if any). On the other hand, the non-monitoring part in equation A.7 is the leading term in long-term probability evaluation and is determined through Bayes' rule by combining estimated probabilities from all our knowledge based on theoretical models and/or beliefs, and past data.

The estimation of these three unknowns requires the use of two concepts: (1) the *Bayesian Inference* and the *Fuzzy approach*. For details on these concepts, see *Marzocchi et al.* (2008)

(electronic supplementary material).

A.0.6 Estimating the probabilities at the nodes: (2) Specific calculations

In this part, the details for the conditional probability assessment for each node are explained following the methodology described by *Marzocchi et al.* (2008) (electronic supplementary material). More details and some examples can be found in *Marzocchi et al.* (2004) and *Marzocchi et al.* (2006).

Node 1: *Rest/Unrest* - $[\theta_1]$

The first node has two possible outcomes: the *presence or not* of unrest in a time interval $(t_0, t_0 + \tau]$. The definition of the state of *unrest* is still a very subjective matter, as is the choice of the parameters used to define it. In general, any specific case (i.e. each volcano) has its own, not fixed, behavior, and then the parameters and information chosen is not a rule but will be defined according with the particular case.

- Parameter γ_1

At this node, $\gamma_1 = \eta$, where η is the so called *state of unrest*, which defines the degree at which the volcano is found in unrest at time $t = t_0$ (for details, see *Marzocchi et al.* (2008), electronic supplementary material). This means that monitoring measurements are considered fully informative during an unrest.

- The posterior distribution $\left[\theta_1^{\mathcal{M}}\right]$ **(monitoring part)**

The distribution $\left[\theta_1^{\mathcal{M}}\right]$ is the PDF of the posterior probability to have an unrest in the time interval $(t_0, t_0 + \tau]$ using the monitoring data when, at time $t = t_0$ the volcano is in unrest (i.e. $\eta = 1$). In this case Bayes' rule is not used, but it is set that:

$$\left[\theta_1^{\mathcal{M}}\right] = \delta\left(\left[\theta_1^{\mathcal{M}}\right] - 1\right) \tag{A.8}$$

where $\delta(\cdot)$ indicates the *Dirac Function*. It means that, if at $t = t_0$ the volcano is in unrest, then it is assumed that it will also be at unrest in the time interval $(t_0, t_0 + \tau]$.

- The posterior distribution $\left[\theta_1^{\bar{\mathcal{M}}}\right]$ **(non-monitoring part)**

The prior distribution for θ_1 is defined as the Beta distribution:

$$\left[\theta_1^{\bar{\mathcal{M}}}\right] = \mathbf{Beta}\left(\alpha_1, \beta_1\right) \tag{A.9}$$

where parameters α_1 and β_1 are determined by

$$\alpha_1 \;=\; \Theta_1^{\bar{\mathcal{M}}}\left(\Lambda_1^{\bar{\mathcal{M}}}+1\right) \tag{A.10}$$

$$\beta_1 \;=\; \left(\Lambda_1^{\bar{\mathcal{M}}}+1\right)-\alpha_1 \tag{A.11}$$

where $\Theta_1^{\bar{\mathcal{M}}}$ is the central value inferred by *a priori* information (models, theoretical beliefs, etc.) and $\Lambda_1^{\bar{\mathcal{M}}}$ is the so called *equivalent number of data* (for details see *Marzocchi et al.* (2008), electronic supplementary material) and controls the confidence at which $\Theta_1^{\bar{\mathcal{M}}}$ is considered a reliable value. Both $\Theta_1^{\bar{\mathcal{M}}}$ and $\Lambda_1^{\bar{\mathcal{M}}}$ are input parameters of the BET model.

To write the *likelihood* for node 1, the two possible outcomes can be treated as *success* (unrest) and *failure* (non-unrest) using a binomial model under some specific conditions. If we define $n*$ as the total number of non-overlapping time windows (of length τ) of the available catalog of data, we obtain a sequence of *non-unrest* (let's say 0) and *unrest* (1). However, this sequence could be autocorrelated if the duration of an unrest is larger than the time window τ considered. To remove this autocorrelation and therefore to use the binomial model, *Marzocchi et al.* (2008) proposed to consider n_1 as the total number of non-overlapping time windows in a state of non-unrest, and y_1 as a variable that counts the number of (non-overlapping) time windows which contain an onset of unrest, in a set of n_1 observations. It means that each unrest episode counts as one, regardless its duration.

Using the binomial function, the likelihood function is defined as:

$$\left[y_1|\theta_1^{\bar{\mathcal{M}}}\right] = \mathbf{Bin}(y_1,n_1;\theta_1) \tag{A.12}$$

The choice of the Beta and Binomial (or Dirichlet and Multinomial in the multivariate case) distributions simplifies the computation because they are conjugate distributions (*Gelman et al.*, 1995). A Dirichlet multiplied by a Multinomial is still a Dirichlet. Then, the *posterior* distribution for $\left[\theta_1^{\bar{\mathcal{M}}}\right]$ is:

$$\left[\theta_1^{\bar{\mathcal{M}}}\right] = \left[\theta_1^{\bar{\mathcal{M}}}|y_1\right] = Beta\left(\alpha_1+y_1,\beta_1+n_1-y_1\right) \tag{A.13}$$

Node 2: *Origin: Magmatic or not-magmatic* - $[\theta_2]$

As in node 1, node 2 has two possible outcomes: Given that an unrest has occurred, is it of magmatic origin (*success*), or is not a magmatic unrest (*failure*)? Theoretical information that can be used to assign a probability at this node is practically nonexistent. Unrest can be due to different factors and it is practically impossible to a priori anticipate the origin of unrest on theoretical grounds.

- Parameter γ_2

At this node, $\gamma_2 = \eta$, which again means that monitoring measures are fully informative

during an unrest.

- The posterior distribution $\left[\theta_2^{\mathcal{M}}\right]$ (monitoring part)

At this point it is necessary to select the monitoring parameters which could be indicative of magmatic origin for the unrest. . These parameters are quite subjective and could change from volcano to volcano, and could even be a combination of different measures.

For each monitored variable $x_2^{(i)}$ selected ($i = 1, \ldots, L$), where L is the number of selected monitored variables), it is necessary to define the *membership function* (i.e. lower and upper thresholds, and trends, for details see *Marzocchi et al.* (2008)). At each monitoring variable $x_2^{(i)}$ it is possible to associate a *weighting factor* $\omega_2^{(i)}$, $\omega_2^{(\cdot)} \geq 1$ that ranks the relative importance of each parameter respect to the others.

Once the parameters have been selected, the next step is to translate this information in probabilities. First, using the *fuzzy* approach, the *degrees of anomaly* ($z_2^{(i)}$) at $t = t_0$ are estimated (for details see *Marzocchi et al.* (2008)) for all the parameters considered at this node. Using this information, it is possible to estimate a kind of *index of magmatic unrest* Z_2 which is calculated as (*Marzocchi et al.*, 2004, 2008):

$$Z_2 = \sum_{i=1}^{L} \omega_2^{(i)} \, z_2^{(i)} \tag{A.14}$$

Then, Z_2 is defined as a linear combination of the weighted membership functions, and represents the degree of anomaly of the monitored parameters as measured at time t_0. The higher value of Z_2, the higher probability to have a magmatic unrest. The next problem is to translate this information in terms of probability. Again, it is assumed that the *a priori* distribution is a Beta distribution, with mean $\Theta_2^{\mathcal{M}}$ and the maximum variance allowed (then $\Lambda_2^{\mathcal{M}} = 1$). *Marzocchi et al.* (2008) proposed a functional relationship that links Z_2 with $\Theta_2^{\mathcal{M}}$ using an exponential relationship like:

$$\Theta_2^{\mathcal{M}} = 1 - a \, e^{-bZ_2} \tag{A.15}$$

where a and b are positive parameters defined by means of Bayesian inference. The choice of this kind of function is due to fit the next requirements: (1) It should be a monotonically increasing function, so that the larger Z_2, the larger $\Theta_2^{\mathcal{M}}$; (2) the largest increase in probability mean occurs when one of the monitoring variables shows some degree of anomaly, but as more monitored variables become anomalous, the probability mean keeps on rising, but slower (this is to give a higher weight to the first change in a *strong* indicator).

Parameters a and b are assumed independent and their prior distribution is a uniform; the domain of a and b is chosen to have reliable both prior distribution and probability values: for instance, the domain of a could be $[0, 1]$, but, in order to account for the so called *damping effect* (*Marzocchi et al.*, 2004) (when $Z_2 = 0$), *Marzocchi et al.* (2008) have defined an interval

$[0.5, 1]$ as the reliable domain for parameter a. On the other hand, parameter b could assume values in the interval $[0, \infty)$ but as b grows, the *sensibility* to changes in Z_2 (for $Z_2 > 0$) is attenuated; then, parameter b has been calibrated to take values in the domain $[0, 2]$ (*Marzocchi et al.*, 2008).

The prior distribution for the parameter $\theta_2^{\mathcal{M}}$ is set as the marginal distribution of $\theta_2^{\mathcal{M}}$ by integrating over a and b the joint prior distribution of $\theta_2^{\mathcal{M}}$, a, and b:

$$[\theta_2^{\mathcal{M}}]_{prior} = \int_{0.5}^{1} da \int_0^2 db \, [\theta_2^{\mathcal{M}}, a, b]_{prior}$$

$$= \int_{0.5}^{1} da \int_0^2 db \, [\theta_2^{\mathcal{M}} \mid a, b] \, [a, b]_{prior} \tag{A.16}$$

Since a and b are assumed independent and with prior uniform distribution, then: $[a, b]_{prior} = [a]_{prior} [b]_{prior} = 2 \cdot 0.5$. On the other hand, $[\theta_2^{\mathcal{M}} \mid a, b]$ is a Beta distribution with parameters α and β (equations A.10 and A.11) $(2 - 2ae^{-bZ_2})$ and $(2ae^{-bZ_2})$, respectively. Then, from eq. A.16:

$$[\theta_2^{\mathcal{M}}]_{prior} = \int_{0.5}^{1} da \int_0^2 db \, \mathbf{Beta}(2 - 2ae^{-bZ_2}, \, 2ae^{-bZ_2}) \cdot 2 \cdot 0.5 \tag{A.17}$$

The posterior distribution of the parameters a and b is based on Bayesian inference. At node 2, it is assumed that past monitoring data can modify the prior distribution for $\theta_2^{\mathcal{M}}$ only by refining the values of a and b (*Marzocchi et al.*, 2008). In Bayesian terms, given $n_2^{\mathcal{M}}$ past monitored episodes of unrest, each with an observed Z_{2_i} and with outcome y_{2_i} (where y_2 is 1 if the unrest was magmatic, and 0 on the opposite), the joint posterior distribution of $\theta_2^{\mathcal{M}}$, a, and b, is:

$$
\begin{aligned}
\left[\theta_2^{\mathcal{M}}, a, b \mid Z_2, \{Z_{2_i}\}, \{y_{2_i}\}\right] \quad &\propto \quad \left[\theta_2^{\mathcal{M}} \mid a, b, Z_2, \{Z_{2_i}\}, \{y_{2_i}\}\right] \, \left[a, b \mid Z_2, \{Z_{2_i}\}, \{y_{2_i}\}\right] \\[2mm]
&\propto \quad \left[\theta_2^{\mathcal{M}} \mid a, b, Z_2\right] \, \left[a, b \mid \{Z_{2_i}\}, \{y_{2_i}\}\right] \\[2mm]
&\propto \quad \left[\theta_2^{\mathcal{M}} \mid a, b, Z_2\right] \, \left[a, b\right]_{prior} \, \left[\{y_{2_i}\} \mid a, b, \{Z_{2_i}\}\right]
\end{aligned}
$$

$$\tag{A.18}$$

where $\{Z_{2_i}\}$ and $\{y_{2_i}\}$ are the data sets of Z_{2_i} (from past monitored unrest episodes) and y_{2_i} (of their outcomes), respectively, for $i = 1, \ldots, n_2^{\mathcal{M}}$, and the sign $\propto$ stands for proportionality, since the term on the right side is not normalized to 1.

Assuming independent past data, the likelihood $[\{y_{2_i}\} \mid a, b, \{Z_{2_i}\}]$ is given by the product (*Marzocchi et al.*, 2008):

$$[\{y_{2_i}\} \mid a, b, \{Z_{2_i}\}] = \prod_{i=1}^{n_2^{\mathcal{M}}} \underbrace{\left(1 - a^{-bZ_{2_i}}\right)^{y_{2_i}}}_{\mathbf{A}} \underbrace{\left(a^{-bZ_{2_i}}\right)^{1-y_{2_i}}}_{\mathbf{B}} \qquad (A.19)$$

When the i-th event has been cataloged as magmatic, (i.e., if $y_{2_i} = 1$), in eq. A.19 remains only the first term (**A**). Indeed, this term represents the probability that the i-th event is of magmatic origin, as computed with a given set of $a, b, \{Z_{2_i}\}$. On the opposite, when the event has been cataloged as not-magmatic, (i.e., if $y_{2_i} = 0$), it remains only the second term (**B = 1 - A**). Indeed, this term represents the probability that the i-th event is not of magmatic origin, as computed with a given set of $a, b, \{Z_{2_i}\}$.

This likelihood resembles the so called *dose-response* relationship based on the multinomial scheme (*Gelman et al.*, 1995; *Marzocchi et al.*, 2008). It means that the most likely values of a and b are those maximizing the probability of having observed those monitoring data.

Finally, the posterior distribution of $\theta_2^{\mathcal{M}}$ is given by:

$$[\theta_2^{\mathcal{M}}] = \int_{0.5}^1 da \int_0^2 db \ [\theta_2^{\mathcal{M}}, a, b \mid Z_2, \{Z_{2_i}\}, \{y_{2_i}\}]$$

$$\propto \int_{0.5}^1 da \int_0^2 db \ \mathbf{Beta}(2 - 2ae^{-bZ_2}, \ 2ae^{-bZ_2}) \cdot 2 \cdot 0.5$$

$$\cdot \prod_{i=1}^{n_2^{\mathcal{M}}} \left(1 - a^{-bZ_{2_i}}\right)^{y_{2_i}} \left(a^{-bZ_{2_i}}\right)^{1-y_{2_i}} \qquad (A.20)$$

In BET, parameters a and b are determined by sampling randomly 1000 pairs of (a, b) from their posterior bivariate joint distribution $[a, b \mid \{Z_{2_i}\}, \{y_{2_i}\}]$ (in eq. A.18) with likelihood as in eq. A.19 and prior $[a, b]$ distribution as a bivariate uniform on the defined domains. Then, each pair is used in eq. A.16 to build up 1000 different Beta (prior) distributions. From each of them, a single random sampling is performed and the obtained 1000 samples represent the posterior $[\theta_2^{\mathcal{M}}]$ searched.

- The posterior distribution $\left[\theta_2^{\bar{\mathcal{M}}}\right]$ (non-monitoring part)

$\left[\theta_2^{\bar{\mathcal{M}}}\right]$ is the expected value of the probability for the unrest being magmatic, given that there is an unrest, provided by theoretical models and beliefs. The formulation of $\left[\theta_2^{\bar{\mathcal{M}}}\right]$ is identical to the one of $\left[\theta_1^{\bar{\mathcal{M}}}\right]$ (but obviously with index $k = 2$). $\Lambda_2^{\bar{\mathcal{M}}}$ is the number of *equivalent* data that we assign to our prior model; y_2 is the number of observed magmatic unrest episodes at the volcano in the past; n_2 is the total number of observed unrest episodes at the volcano in the past, whose source process (magmatic or not) is known. It implies that $n_2 \leq y_1$.

Then, the posterior $\left[\theta_2^{\bar{\mathcal{M}}}\right]$ is (*Marzocchi et al.*, 2008):

$$\left[\theta_2^{\bar{\mathcal{M}}}\right] = \left[\theta_2^{\bar{\mathcal{M}}} \mid y_2, y_1\right] = \mathbf{Beta}(\alpha_2 + y_2, \ \beta_2 + y_1 - y_2) \tag{A.21}$$

Node 3: *Outcome: Occurrence or not occurrence or eruption* - $[\theta_3]$

Node 3 has also two possible outcomes: the occurrence (*success*, i.e. the magma will reach the surface) or not (*failure*) of an eruption, in a time interval $(t_0, \ t_0 + \tau]$, provided that the unrest has a magmatic origin.

- Parameter γ_3
At this node, $\gamma_3 = \eta$, which again means that monitoring measures are fully informative during an unrest.

- The posterior distribution $\left[\theta_3^{\mathcal{M}}\right]$ **(monitoring part)**

The treatment of the monitored variables to compute the probability distribution for the monitoring part at node 3 is exactly the same of node 2, except that the variables used $\left(x_i^3\right)$ should be indicative of magma erupting. Then, at this node the posterior distribution of $\left[\theta_3^{\mathcal{M}}\right]$ is give by:

$$\left[\theta_3^{\mathcal{M}}\right] = \int_{0.5}^1 da \ \int_0^2 db \ \left[\theta_3^{\mathcal{M}}, a, b \mid Z_3, \{Z_{3_i}\}, \{y_{3_i}\}\right]$$

$$\propto \int_{0.5}^1 da \ \int_0^2 db \ \mathbf{Beta}(2 - 2ae^{-bZ_3}, \ 2ae^{-bZ_3}) \cdot 2 \cdot 0.5$$

$$\cdot \prod_{i=1}^{n_3^{\mathcal{M}}} \left(1 - a^{-bZ_{3_i}}\right)^{y_{3_i}} \ \left(a^{-bZ_{3_i}}\right)^{1-y_{3_i}} \tag{A.22}$$

where $n_3^{\mathcal{M}}$ is the number of past monitored episodes of magmatic unrest, $\{Z_{3_i}\}, \ (i = 1, \ldots, n_3^{\mathcal{M}})$ is the data set of Z_{3_i} from past monitored of magmatic unrest (equation A.14 with index 3); $\{y_{3_i}\}, \ (i = 1, \ldots, n_3^{\mathcal{M}})$ is the set of outcomes of the past monitored magmatic unrest (0 for no eruption, 1 for eruption), and the parameters a and b determined in the same way as in node 2 (note that their posterior distributions are different from the ones at node 2).

- The posterior distribution $\left[\theta_3^{\bar{\mathcal{M}}}\right]$ **(non-monitoring part)**

$\left[\theta_3^{\bar{\mathcal{M}}}\right]$ is the expected value of the probability of eruption, given that there is a magmatic unrest, provided by theoretical models and beliefs. The formulation of $\left[\theta_3^{\bar{\mathcal{M}}}\right]$ is also identical

to the one of $\left[\theta_1^{\bar{\mathcal{M}}}\right]$ (but obviously with index $k=3$). $\Lambda_3^{\bar{\mathcal{M}}}$ is the number of *equivalent* data that we assign to our prior model; y_3 is the number of observed eruptions at the volcano in the past; n_3 is the total number of observed magmatic unrest at the volcano in the past, whose outcome (magmatic or not) is known. It implies that $n_3 \leq y_2$.

Then, the posterior $\left[\theta_3^{\bar{\mathcal{M}}}\right]$ is (*Marzocchi et al.*, 2008):

$$\left[\theta_3^{\bar{\mathcal{M}}}\right] = \left[\theta_3^{\bar{\mathcal{M}}} \mid y_3, y_2, y_1\right] = \mathbf{Beta}(\alpha_3 + y_3, \ \beta_3 + y_2 - y_3) \tag{A.23}$$

Node 4: *Location* - $[\theta_4]$

Node 4 considers if the eruption will occur in a specific location (e.g. crater, a flank, etc.), provided that there is an eruption. If we define J_4 possible areas where the eruption can take place, then node 4 has J_4 possible outcomes, each one related to a specific location of the eruption.

- Parameter γ_4
Several phenomena observed in monitoring (i.e. apparent migration of seismic events) may lead to false interpretations to individuate possible localizations of future vents (*Marzocchi et al.*, 2008), then at node 4 the monitoring information should not completely control the probability $[\theta_4^j]$. In BET it has been established that both monitoring and non-monitoring parts should have the same weight, even in the case of a clear unrest ($\eta = 1$) (*Marzocchi et al.*, 2008). Then, γ_4 is set like:

$$\gamma_4 = min(\eta, \ 0.5) \tag{A.24}$$

and in this way, during an unrest, the probability of a specific vent opening accounts for both monitoring and non-monitoring information.

- The posterior distribution $\left[\theta_4^{(j)\mathcal{M}}\right]$ (monitoring part)

When it is possible to localize relevant monitored measurements, BET calculates the *fraction* of the measured parameter occurring in the specific J_4-th vent location. Then, for every monitored parameter, the sum of these fractions on all the vent locations defined must give 1. All the localized monitoring measures are combined to form the monitoring probability distribution.

In this case, since no other data is going to be used to compute the posterior distribution, then $\left[\theta_4^{\mathcal{M}}\right]_{posterior} \equiv \left[\theta_4^{\mathcal{M}}\right]_{prior} = \left[\theta_4^{\mathcal{M}}\right]$, which is defined as:

$$\left[\theta_4^{\mathcal{M}}\right] = \mathbf{Di}_{J_4}\left(\alpha_4^{(1)}, \alpha_4^{(2)}, \alpha_4^{(3)}, \ldots, \alpha_4^{(J_4)}\right) \tag{A.25}$$

where

$$\alpha_4^{(j)} = \frac{J_4}{L_4} \sum_{n=1}^{L_4} f_n^{(j)}, \qquad \text{for } j = 1, 2, 3, \ldots, J_4 \tag{A.26}$$

$f_n^{(j)}$ is the *fraction* of the n-th localized monitoring parameter in the j-th vent location, L_4 is the number of monitoring parameters considered, and $\mathbf{Di}_{J_4}(\cdot)$ is the *Dirichlet distribution*. Then for the j-th vent location, the j-th posterior distribution for the monitoring part is the *marginal* distribution $\left[\theta_4^{(j)\mathcal{M}}\right]$ of eq. A.25 which is:

$$\left[\theta_4^{(j)\mathcal{M}}\right] = \mathbf{Beta}\left(\alpha_4^{(j)}, \; J_4 - \alpha_4^{(j)}\right) \tag{A.27}$$

Then, the mean of the j-th vent location is

$$E\left[\theta_4^{(j)\mathcal{M}}\right] \equiv \Theta_4^{(j)\mathcal{M}} = \frac{\sum_{n=1}^{L_4} f_n^{(j)}}{L_4} \tag{A.28}$$

which means that at each location, the distribution mean is the average of the fractions of all localized monitoring measures at that vent location. On the other hand, the variance of the probability distribution is taken as the maximum allowed, then $\Lambda_4^{\mathcal{M}} = 1$.

- The posterior distribution $\left[\theta_4^{(j)\bar{\mathcal{M}}}\right]$ (non-monitoring part)

At this node, the prior distribution is defined as a Dirichlet distribution with J_4 parameters $\alpha_4^{(j)}$ (*Marzocchi et al.*, 2008):

$$\left[\theta_4^{\bar{\mathcal{M}}}\right]_{prior} = \mathbf{Di}_{J_4}\left(\alpha_4^{(1)}, \alpha_4^{(2)}, \ldots, \alpha_4^{(J_4)}\right) \tag{A.29}$$

where

$$\alpha_4^{(j)} = \Theta_4^{(j)\bar{\mathcal{M}}}\left(\Lambda_4^{\bar{M}} + J_4 - 1\right) \tag{A.30}$$

and where $\Theta_4^{(j)\bar{\mathcal{M}}}$ and $\Lambda_4^{\bar{M}}$ are parameters that depend on the expected value and variance of the theoretical models and/or beliefs (note that the $\alpha_4^{(j)}$) used here are different from the ones used used in the monitoring part).

For the likelihood, it is used a multinomial distribution with J_4 possible outcomes (one for each vent). Using Bayes theorem and adopting the results of the conjugate families, the next posterior distribution is obtained:

$$\left[\theta_4^{\bar{\mathcal{M}}}\right] \equiv \left[\theta_4^{\bar{\mathcal{M}}} \mid y_4, y_3, y_2, y_1\right] = \mathbf{Di}_{J_4}\left(\alpha_4^{(1)} + y_4^{(1)}, \; \alpha_4^{(2)} + y_4^{(2)}, \; \ldots, \alpha_4^{(J_4)} + y_4^{(J_4)}\right) \tag{A.31}$$

where $y_4^{(j)}$ is the number of eruptions observed in the past at the j-th location; then $\sum_{j=1}^{J_4} y_4^{(j)}$ is the total number of eruptions at the volcano with known localization, which should represent a complete eruption catalog of the period considered.

For the j-th vent location, the posterior distribution for the non-monitoring part is the *marginal* distribution $\left[\theta_4^{(j).\bar{\mathcal{M}}}\right]$ in eq. A.31:

$$\left[\theta_4^{(j).\bar{\mathcal{M}}}\right] = \textbf{Beta}\left(\alpha_4^{(j)}, \left\{\left(\sum_{i=1}^{J_4}\alpha_4^{(i)}\right) - \alpha_4^{(j)}\right\}\right), \quad \text{for } j = 1, 2, \ldots, J_4 \qquad (A.32)$$

Node 5: *Magnitude* - $[\theta_5]$

Node 5 considers the possibility that the eruption will be of a certain magnitude/size, provided that there is an eruption in a certain location. If we define J_5 classes generically defined of sizes, magnitudes, or even types of eruptions, then node 5 has J_5 possible outcomes, each one related to a specific class of eruption defined. The definition of sizes/types is a generic matter since it may depend on the goal of the specific BET application, and on the characteristics of the target volcano, and even on the available information.

Often such a size can be quantified through the *Volcanic Explosivity Index* - VEI (*Newhall and Self*, 1982) and then often is the most common information available to quantify eruption size. However, it should be noticed that is not the only possible choice.

In BET, it is assumed that the size distribution is independent from bent location. This means that in all possible vent locations the probability estimations of sizes are identical (*Marzocchi et al.*, 2008)

- Parameter γ_5
Until now there is no reliable precursor monitored parameters to infer the eruption size (*Sandri et al.*, 2004; *Marzocchi et al.*, 2008), then BET does not use monitoring information to improve the probability estimation at this node. Therefore, at this node $\gamma_5 = 0$, and then $[\theta_5^{(j).\mathcal{M}}]$ is not computed.

- The posterior distribution $\left[\theta_5^{(j).\bar{\mathcal{M}}}\right]$ **(non-monitoring part)**

$\left[\theta_5^{\bar{\mathcal{M}}}\right]$ is the expected value of the probability of eruption of j-th size, given that there is an eruption, provided by theoretical models and beliefs. The formulation of $\left[\theta_5^{\bar{\mathcal{M}}}\right]$ is identical to the one of $\left[\theta_4^{\bar{\mathcal{M}}}\right]$ (but obviously with index $k = 5$). $\Lambda_5^{\bar{\mathcal{M}}}$ is the number of *equivalent* data that we assign to our prior model and defines the variance of the distribution; $y_5^{(j)}$ is the number of observed eruptions of the j-th size at the volcano in the past; $n_5 = \sum_{j=1}^{J_5} y_5^{(j)}$ is the total number of observed eruptions at the volcano in the past with known size. Those events must represent a complete eruption catalog of the period considered.

Then, the posterior $\left[\theta_5^{\bar{\mathcal{M}}}\right]$ is (*Marzocchi et al.*, 2008):

$$\left[\theta_5^{\bar{\mathcal{M}}}\right] \equiv \left[\theta_5^{\bar{\mathcal{M}}} \mid y_5, y_3, y_2, y_1\right] = \mathbf{Di}_{J_5}\left(\alpha_5^{(1)} + y_5^{(1)}, \ \alpha_5^{(2)} + y_5^{(2)}, \ \ldots, \alpha_5^{(J_5)} + y_5^{(J_5)}\right) \quad \text{(A.33)}$$

and for the j-th class, the posterior distribution for the non-monitoring part is the marginal distribution $\left[\theta_5^{\bar{\mathcal{M}}}\right]$, from eq. A.33

$$\left[\theta_5^{(j)\bar{\mathcal{M}}}\right] = \mathbf{Beta}\left(\alpha_5^{(j)}, \ \left\{\left(\sum_{i=1}^{J_5}\alpha_5^{(i)}\right) - \alpha_5^{(j)}\right\}\right), \quad \text{for } j = 1, 2, \ldots, J_5 \quad \text{(A.34)}$$

Bibliography

Agip, Geologia e geofisica del sistema geotermico dei Campi Flegrei, Internal Report, *Tech. rep.*, Agip, Milano, Italy, 1987.

Akaike, H., A new look at the statistical model identification, *I.E.E.E. Trans. Automat. Contr.*, *AC 19*, 716–723, 1974.

Aki, K., Evidence for magma intrusion during the Mammoth Lakes of May 1980 and implications of the absence of volcanic (harmonic) tremor, *J. Geophys. Res.*, *89*, 7689–7696, 1984.

Aki, K., State of the art in volcanic seismology, in *Volcanic Seismology*, edited by P. Gasparini, R. Scarpa, and K. Aki, pp. 3–10, Springer-Verlag, Berlin New York, 1992.

Aki, K., and R. Koyanagi, Deep volcanic tremor and magma ascent mechanism under Kilauea, Hawaii, *J. Geophys. Res.*, *86*, 7095–7109, 1981.

Aki, K., M. Fehler, and S. Das, Source mechanism of volcanic tremor: fluid-driven crack models and their application to the 1963 Kilauea eruption, *J. Volcanol. Geotherm. Res.*, *2*, 259–287, 1977.

Amato, A., et al., The 1997 Umbria-Marche, Italy, earthquake sequence: a first look at the main shocks and aftershocks, *Geophys. Res. Lett.*, *25*, 2861–2864, 1998.

Amma-Miyasaka, M., M. Nakagawa, and S. Nakada, Magma plumbing system of the 2000 eruption of Miyakejima volcano, Japan, *Bull. Volcanol.*, *67*, 254–267, 2005.

Arnorsson, A., G. Axelsson, and S. Kristjan, Geothermal system in Iceland, *The Icelandic J. Earth Sc.*, *58*, 269–302, 2008.

Bain, L. J., *Statistical analysis of reliability and life-testing models, Theory and Methods*, Marcel Dekker, New York, 1978.

Barberi, F., M. Ghigliotti, G. Macedonio, H. Orellana, M. T. Pareschi, and M. Rosi, Volcanic hazard assessment of Guagua Pichincha (Ecuador) based on past behaviour and numerical models, *J. Volcanol. Geotherm. Res.*, *49*, 53–68, 1992.

Bautista, B. C., M. L. P. Bautista, R. S. Stein, E. S. Barcelona, R. S. Punongbayan, E. P. Laguerta, A. R. Rasdas, G. Ambubuyog, and E. Q. Amin, Relationship of regional and local structures to Mount Pinatubo activity, in *Fire and Mud: eruptions and lahars of Mount Pinatubo, Philippines*, edited by C. G. Newhall and R. S. Punongbayan, University of Washington Press, Seattle and London, 1996.

Bebbington, M. S., and C. D. Lai, On nonhomogeneous models for volcanic eruptions, *Math. Geol.*, *28*, 585–600, 1996a.

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

Berthold, M. R., and D. J. Hand (Eds.), *Intelligent Data Analysis, an Introduction*, 2 ed., Springer-Verlag, 2003.

Bonafede, M., and S. Danesi, Near-field modifications of stress induced by dyke injection at shallow depth, *Geophys. J. Int.*, *130*, 435–448, 1997.

Bowers, N., H. Gerber, J. Hickman, D. Jones, and C. Nesbitt, *Actuarial Mathematics*, 2 ed., Society of Actuaries, 1997.

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

Castello, B., G. Selvaggi, C. Chiarabba, and A. Amato, CSI - Catalogo della sismicità Italiana 1981-2002, versione 1.0., 2005.

Castello, B., M. Olivieri, and G. Selvaggi, Local and duration magnitude determination for the Italian Earthquake Catalog, 1981-2002, *Bull. Seism. Soc. Am.*, *97*, 128–139, 2007.

Chelini, W., and A. Sbrana, Subsurface geology, in *Phlegrean Fields*, edited by M. Rosi and A. Sbrana, vol. 114 of *Quad. Ric. Sci.*, pp. 94–102, Consiglio Nazionale delle Ricerche, Roma, 1987.

Chiaraluce, L., W. L. Ellsworth, C. Chiarabba, and M. Cocco, Imaging the complexity of an active normal fault system: The 1997 Colfiorito (central Italy) case study, *J. Geophys. Res.*, *108*, 2003.

Chouet, B. A., Ground motion in the near field of a fluid-driven crack and its interpretation in the study of shallow volcanic tremor, *J. Geophys. Res.*, *86*, 5985–6016, 1981.

Chouet, B. A., Free surface displacements in the near field of a tensile crack expanding in three dimensions, *J. Geophys. Res.*, *87*, 3868–3872, 1982.

Chouet, B. A., Excitation of a buried magmatic pipe: A seismic source model for volcanic tremor, *J. Geophys. Res.*, *90*, 1881–1893, 1985.

Chouet, B. A., Dynamics of a fluid-driven crack in three dimensions by the finite difference method, *J. Geophys. Res.*, *91*, 13,967–13,992, 1986.

Chouet, B. A., Resonance of a fluid-driven crack: Radiation properties and implications for the source of long-period events and harmonic tremor, *J. Geophys. Res.*, *93*, 4375–4400, 1988.

Chouet, B. A., A seismic model for the source of long-period events and harmonic tremor, in *Volcano Seismology*, edited by P. Gasparini, R. Scarpa, and K. Aki, pp. 133–156, Springer Berlin Heidelberg, New York, 1992.

Chouet, B. A., New methods and future trends in seismological volcano monitoring, in *Monitoring and mitigation of volcano hazards*, edited by R. Scarpa and R. Tilling, pp. 23–97, Springer-Verlag, Berlin New York, 1996.

Coles, S., and R. Sparks, Extreme value methods for modeling historical series of large volcanic magnitudes, in *Statistics in Volcanology*, edited by H. Mader, S. Coles, C. Connor, and L. Connor, IAVCEI Publications, pp. 47–56, Geological Society, London, 2006.

Cooley, ., J. W., and J. W. Tukey, An algorithm for machine calculation of complex Fourier series, *Math. Comput.*, *19*, 297–301, 1965.

Cox, D. R., and P. A. W. Lewis, *The statistical analysis of Series of Events*, Mathuen, New York, 1966.

Dalay, D. J., and D. Vere-Jones, *An introduction to the Theory of Point Processes*, Springer, New York, 1988.

De Angelis, S., Analyses of unusual long-period earthquakes with extended coda recorded at Katmai National park, Alaska, USA, *Geophys. Res. Lett.*, *33*, 2006.

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

De Natale, G., and A. Zollo, Statistical analysis and clustering features of the Phlegrean Fields earthquake sequence (May 1983 - May 1984), *Bull. Seism. Soc. Am.*, *76*, 801–814, 1986.

De Natale, G., A. Zollo, C. Del Gaudio, G. P. Ricciardi, and M. Martini, Error analysis in hypocentral locations at Phlegraean Fields, *Bull. Volcanol.*, *47*, 210–218, 1984.

De Natale, G., A. Zollo, A. Ferraro, and J. Virieux, Accurate fault mechanism determinations for a 1984 earthquake swarm at Campi Flegrei caldera (Italy) during an unrest episode: implications for volcanological research, *J. Geophys. Res.*, *100*, 24,167–24,185, 1995.

Dreger, D. S., H. Tkalcic, and M. Johnston, Dilational processes accompanying earthquakes in the Long Valley Caldera, *Science*, *288*, 122–125, 2000.

Ellsworth, W. L., M. V. Matthews, R. M. Nadeau, S. P. Nishenko, P. A. Reasenberg, and R. W. Simpson, A physically based earthquake recurrence model for estimation of long-term earthquake probabilities, *U. S. Geol. Surv. Open-File Rept.*, pp. 99–522, 1999.

Ferrazzini, V., and K. Aki, Slow waves trapped in a fluid-filled infinite crack: implications for volcanic tremor, *J. Geophys. Res.*, *92*, 9215–9223, 1987.

Fletcher, R., and M. J. D. Powell, A rapidly convergent descent method for minimization, *Comput. J.*, *6*, 163–168, 1963.

Fujita, E., and Y. Ida, Low attenuation resonance of a spherical magma chamber: Source mechanism of monotonic volcanic tremor at Asama Volcano, Japan, *Geophys. Res. Lett.*, *26*, 3221–3224, 1999.

Fujita, E., and Y. Ida, Geometrical effects and low-attenuation resonance of volcanic fluid inclusions for the source mechanism of long-period earthquakes, *J. Geophys. Res.*, *108*, 2003.

Fujita, E., and M. Ukawa, Cyclic jerky opening of magma sheet and caldera formation during the 2000 miyakejima volcano eruption, *Geophys. Res. Lett.*, *29*, 2002.

Fujita, E., Y. Ida, and J. Oikawa, Eigen oscillation of a fluid sphere and source mechanism of harmonic volcanic tremor, *J. Volcanol. Geotherm. Res.*, *69*, 365–378, 1995.

Fujita, E., M. Ukawa, and E. Yamamoto, Subsurface cyclic magma sill expansions in the 2000 Miyakejima volcano eruption: Possibility of two-phase flow oscillation, *J. Geophys. Res.*, *109*, 2004.

Garcia-Aristizabal, A., H. Kumagai, P. Samaniego, P. Mothes, H. Yepes, and M. Monzier, Seismic, petrologic, and geodetic analyses of the 1999 dome-forming eruption of Guagua Pichincha volcano, Ecuador, *J. Volcanol. Geotherm. Res.*, *161*, 333–351, 2007.

Gasparini, P., R. Scarpa, and K. Aki (Eds.), *Volcano Seismology*, Springer Berlin Heidelberg, New York, 1992.

Gelman, A., and D. Rubin, Inference from iterative simulation using multiple sequences, *Statist. Sci.*, *7*, 457–472, 1992.

Gelman, A., J. Carlin, H. Stern, and D. Rubin, *Bayesian Data Analysis*, Chapman and Hall, New York, 1995.

Geweke, J., *Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments, In: Bayesian Statistics 4, J.M. Bernardo, J.O. Berger, A.P. Dawid, and A.F. Smith (eds.)*, pp. 169–193, Oxford University Press, 1992.

Geyer, C. J., Practical Markov Chain Monte Carlo, *Statist. Sci.*, *7*, 473–483, 1992.

Hainzl, S., and Y. Ogata, Detecting fluid signals in seismicity data through statistical earthquake modeling, *J. Geophys. Res.*, *110*, 2005.

Hall, M., *El volcanismo en el Ecuador*, Instituto Panamericano de Geografía e Historia, Quito, 1977, (in spanish).

Harlow, D. H., J. A. Power, E. P. Laguerta, G. Ambubuyog, R. A. White, and R. P. Hoblitt, Precursory seismicity and forecasting of the june 15, 1991, eruption of Mount Pinatubo, in *Fire and Mud: eruptions and lahars of Mount Pinatubo, Philippines*, edited by C. G. Newhall and R. S. Punongbayan, University of Washington Press, Seattle and London, 1996.

Harris, R. A., Introduction to special section: Stress triggers, stress shadows, and implications for seismic hazard, *J. Geophys. Res.*, *103*, 24,347–24,358, 1998.

Hastings, W. K., Monte Carlo sampling methods using Markov chains and their applications, *Biometrika*, *57*, 97–109, 1970.

Hauksson, E., and P. Shearer, Southern California hypocenter relocation with waveform cross-correlation, part 1: Results using the double-difference method, *Bull. Seism. Soc. Am.*, *95*, 896–903, 2005.

Hauksson, E., L. M. Jones, K. Hutton, and D. Eberhart-Phillips, The 1992 Landers earthquake sequence: Seismological observations, *J. Geophys. Res.*, *98*, 19,835–19,858, 1993.

Hill, D. P., A model for earthquake swarms, *J. Geophys. Res.*, *82*, 1347–1352, 1977.

Hill, D. P., Unrest in Long Valley Caldera, California, 1978–2004, in *Mechanisms of Activity and Unrest at Large Calderas*, edited by C. Troise, G. De Natale, and C. R. J. Kilburn, 269, pp. 1–24, The Geological Society of London, London, 2006.

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

Ho, C. H., Volcanic time-trend analysis, *J. Volcanol. Geotherm. Res.*, *74*, 171–177, 1996.

Hoblitt, R. P., E. W. Wolfe, W. E. Scott, M. R. Couchman, J. S. Pallister, and D. Javier, The preclimactic eruptions of Mount Pinatubo, june 1991, in *Fire and Mud: eruptions and lahars of Mount Pinatubo, Philippines*, edited by C. G. Newhall and R. S. Punongbayan, University of Washington Press, Seattle and London, 1996.

Hori, S., Y. Fukao, M. Kumazawa, M. Furumoto, and A. Yamamoto, A new method of spectral analysis and its application to the Earth's free oscillations: The Sompi method, *J. Geophys. Res.*, *94*, 7535–7553, 1989.

Ingebritsen, S. E., and W. E. Sanford, *Groundwater in Geologic Processes*, Cambridge University Press, Cambridge, UK, 1998.

Jackson, D. D., The use of a priori data to resolve non-uniqueness in linear inversion, *Geophys. J. R. Astron. Soc.*, *57*, 137–157, 1979.

Kay, S. M., and S. L. Marple, Spectrum analysis - a modern perspective, *Proc. I.E.E.E.*, *69*, 1380–1419, 1981.

Kimeldorf, G., and G. Wahba, A correspondence between Bayesian estimation of stochastic processes and smooting by splines, *Ann. Math. Stat.*, *41*, 495–502, 1970.

Kirkpatrik, S., D. Gelatt, and M. Vecchi, Optimization by simulated annealing, *Science*, *220*, 671–680, 1983.

Kisslinger, C., and L. Jones, Properties of aftershock sequences in southern California, *J. Geophys. Res*, *96*, 11,947–11,958, 1991.

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

Klir, G. J., and T. A. Folger, *Fuzzy sets, uncertainty, and information*, Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1987.

Kobayashi, T., T. Ohminato, Y. Ida, and E. Fujita, Very long period seismic signals observed before the caldera formation with the 2000 Miyake-jima volcanic activity, Japan, *J. Geophys. Res.*, *114*, 2009.

Kumagai, H., and B. Chouet, Acoustic properties of a crack containing magmatic or hydrothermal fluids, *J. Geophys. Res.*, *105*, 25,493–25,512, 2000.

Kumagai, H., and B. Chouet, The dependence of acoustic properties of a crack on the resonance mode and geometry, *Geophys. Res. Lett.*, *28*, 3325–3328, 2001.

Kumagai, H., and B. A. Chouet, The complex frequencies of long-period seismic events as probes of fluid composition beneath volcanoes, *Geophys. J. Int.*, *138*, F7–F12, 1999.

Kumagai, H., B. A. Chouet, and M. Nakano, Temporal evolution of a hydrothermal system in Kusatsu-Shirane volcano, Japan, inferred from the complex frequencies of long-period events, *J. Geophys. Res.*, *107*, 2002.

Kumagai, H., B. A. Chouet, and P. B. Dawson, Source process of a long-period event at Kilauea volcano, Hawaii, *Geophys. J. Int.*, *161*, 243–254, 2005.

Kumazawa, M., Y. Imanishi, Y. Fukao, M. Furumoto, and A. Yamamoto, A theory of spectral analysis based on the characteristic property of a linear dynamic system, *Geophys. J. Int.*, *101*, 613–630, 1990.

Lindsay, J., W. Marzocchi, G. Jolly, R. Constantinescu, J. Selva, and L. Sandri, Towards real-time eruption forecasting in the Auckland Volcanic Field: application of BET_EF during the New Zealand National Disaster Exercise, *Bull. Volcanol.*, *72*, 2010.

Lombardi, A. M., W. Marzocchi, and J. Selva, Exploring the evolution of a volcanic seismic swarm: The case of the 2000 izu islands swarm, *Geophys. Res. Lett.*, *33*, 2006.

Mamdani, E. H., Applications of fuzzy algorithm for control of a simple dynamic plant, *Proc. I.E.E.E.*, *121*, 1585–1588, 1974.

Marzocchi, W., and L. Zaccarelli, A quantitative model for time-size distribution of eruptions, *J. Geophys. Res.*, *111*, 2006.

Marzocchi, W., and A. Zollo (Eds.), *Conception, verification, and application of innovative techniques to study active volcano*, Istituto Nazionale di Geofisica e Vulcanologia, 2008.

Marzocchi, W., G. Vilardo, D. P. Hill, G. P. Ricciardi, and C. Ricco, Common features and peculiarities of the seismic activity at Phlegraean Fields, Long Valley, and Vesuvius, *Bull. Seism. Soc. Am.*, *91*, 191–205, 2001.

Marzocchi, W., L. Sandri, P. Gasparini, C. Newhall, and E. Boschi, Quantifying probabilities of volcanic events: The example of volcanic hazard at Mount Vesuvius, *J. Geophys. Res.*, *109*, 2004.

Marzocchi, W., L. Sandri, and L. Furlan, A quantitative model for Volcanic Hazard assessment, in *Statistical Volcanology*, edited by H. Mader, S. Coles, C. Connor, and L. Connor, pp. 31–37, IAVCEI Publications, 2006.

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

Matsuura, T., Y. Imanishi, M. Imanari, and M. Kumazawa, Application of a new method of high-resolution spectral analysis, Sompi, for free induction decay of nuclear magnetic resonance, *Appl. Spectrosc.*, *44*, 618–626, 1990.

Matthews, M., W. Ellsworth, and P. Reasenberg, A Brownian model for recurrent earthquakes, *Bull. Seism. Soc. Am.*, *92*, 2233–2250, 2002.

McNutt, S. R., Seismic monitoring and eruption forecasting of volcanoes: A review of the state-of-the-art and case histories, in *Monitoring and Mitigation of Volcano Hazards*, edited by R. Scarpa and R. R. Tilling, pp. 99–146, Springer-Verlag, Berlin New York, 1996.

Menke, W., *Geophysical Data Analysis: Discrete Inverse Theory, Revised Edition*, vol. 45 of *International Geophysics Series*, Academic Press, Inc., New York, 1989.

Metropolis, N., and S. Ulam, The Monte Carlo method, *J. Amer. Stat. Assoc.*, *44*, 335–341, 1949.

Metropolis, N., A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller, Equation of state calculations by fast computing machines, *J. Chem. Phys.*, *21*, 1081–1092, 1953.

Mizumoto, M., Realization of pid controls by fuzzy control methods, *Fuzzy Sets and Systems*, *70*, 171–182, 1995.

Mogi, K., Earthquakes and fractures, *Tectonophysics*, *5*, 35–55, 1967.

Molina, I., H. Kumagai, and H. Yepes, Resonances of a volcanic conduit triggered by repetitive injections of an ash-laden gas, *Geophys. Res. Lett.*, *31*, 2004.

Molina, I., H. Kumagai, A. Garcia-Aristizabal, M. Nakano, and P. Mothes, Source process of very-long-period events accompanying long-period signals at Cotopaxi volcano, Ecuador, *J. Volcanol. Geotherm. Res.*, *176*, 119–133, 2008.

Morrissey, M., and B. Chouet, Trends in long-period seismicity related to magmatic fluid compositions, *J. Volcanol. Geotherm. Res.*, *108*, 265–281, 2001.

Mosegaard, K., and A. Tarantola, Probabilistic approach to inverse problems, in *International handbook of earthquakes and engineering seismology, Part A*, pp. 237–265, Academic Press., 2002.

Mulargia, F., and S. Tinti, Seismic sample areas defined from incomplete catalogs: An application to the Italian territory, *Phys. Earth Planet. Inter.*, *40*, 273–300, 1985.

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

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

Murru, M., R. Console, and A. Lisi, Seismicity and mean magnitude variations correlated to the strongest earthquakes of the 1997 Umbria-Marche sequence (central Italy), *J. Geophys. Res.*, *109*, 2004.

Nakada, S., M. Nagai, T. Kaneko, A. Nozawa, and K. Suzuki-Kamata, Chronology and products of the 2000 eruption of Miyakejima volcano, Japan, *Bull. Volcanol.*, *67*, 205–218, 2005.

Nakano, M., H. Kumagai, M. Kumazawa, K. Yamaoka, and B. A. Chouet, The excitation and characteristic frequency of the long-period volcanic event: An approach based on an inhomogeneous autoregressive model of a linear dynamic system, *J. Geophys. Res.*, *103*, 10,031–10,046, 1998.

Neuberg, J., R. Luckett, B. Baptie, and K. Olsen, Models of tremor and low-frequency earthquake swarms on Montserrat, *J. Volcanol. Geotherm. Res.*, *101*, 83–104, 2000.

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

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

Newhall, C. G., and R. S. Punongbayan, Pinatubo: Eruption and muddy aftermath of the century's second largest eruption, *Bulletin of Volcanic Eruptions*, *32*, 147–152, 1995.

Newhall, C. G., and R. S. Punongbayan (Eds.), *Fire and Mud: eruptions and lahars of Mount Pinatubo, Philippines*, University of Washington Press, Seattle and London, 1996.

O'connell, R., and B. Budiansky, Measures of dissipation in viscoelastic media, *Geophys. Res. Lett.*, *5*, 5–8, 1978.

Ogata, Y., Statistical models for earthquake occurrences and residual analysis for point processes, *J. Am. Stat. Assoc.*, *83*, 9–27, 1988.

Ogata, Y., Space-time point-process models for earthquake occurrences, *Ann. Inst. Stat. Math.*, *50*, 379–402, 1998.

Ogata, Y., Estimating the hazard of ruptire using uncertain occurrence times of paleoearthquakes, *J. Geophys. Res.*, *104*, 17,995–18,014, 1999.

Ogata, Y., R. S. Matsu'ura, and K. Katsura, Fast likelihood computation of epidemic type aftershock-sequence model, *Geophys. Res. Lett.*, *20*, 2143–2146, 1993.

Ohminato, T., B. A. Chouet, P. Dawson, and S. Kedar, Waveform inversion of very long period impulsive signals associated with magmatic injections beneath Kilauea volcano, Hawaii, *J. Geophys. Res.*, *B10*, 23,839–23,862, 1998.

Omer, G. C., Volcanic tremor, *Bull. Seism. Soc. Am.*, *40*, 175–194, 1950.

Orsi, G., L. Civetta, C. Del Gaudio, S. De Vita, M. Di Vito, R. Isaia, S. Petrazzuoli, G. Ricciardi, and C. Ricco, Short-term ground deformations and seismicity in the resurgent Campi Flegrei caldera (Italy): an example of active block-resurgence in a densely populated area, *J. Volcanol. Geotherm. Res.*, *91*, 415–451, 1999.

Palmer, D. A., R. Fernandez-Prini, and A. H. Harvey (Eds.), *Aqueous Systems at Elevated Temperatures: Physical Chemistry in Water, Steam, and Hydrothermal Solutions*, Elsevier Academic Press., London, 2004.

Plummer, C. C., D. McGeary, and D. Carlson, *Physical Geology*, 8th ed., Mc. Graw Hill, New York, 2001.

Power, J. A., T. L. Murray, J. N. Marso, and E. P. Laguerta, Preliminary observations of seismicity at Mount Pinatubo by use of the Seismic Spectral Amplitude Measurement (SSAM) system, May 13-June 18, 1991, in *Fire and Mud: eruptions and lahars of Mount Pinatubo, Philippines*, edited by C. G. Newhall and R. S. Punongbayan, University of Washington Press, Seattle and London, 1996.

Roman, D. C., Numerical models of volcanotectonic earthquake triggering on non-ideally oriented faults, *Geophys. Res. Lett*, *32*, 2005.

Sambridge, M., and K. Mosegaard, Monte Carlo methods in geophysical inverse problems, *Rev. Geophys.*, *40*, 1–29, 2002.

Sandri, L., W. Marzocchi, and L. Zaccarelli, A new perspective in identifying the precursory patterns of volcanic eruptions, *Bull. Volcanol.*, *66*, 263–275, 2004.

Sandri, L., W. Marzocchi, and P. Gasperini, Some insights on the occurrence of recent volcanic eruptions of Mount Etna volcano (Sicily, Italy), *Geophys. J. Int.*, *163*, 1203–1218, 2005.

Sandri, L., E. Guidoboni, W. Marzocchi, and J. Selva, Bayesian event tree for eruption forecasting (BET_EF) at Vesuvius, Italy: a retrospective forward application to the 1631 eruption, *Bull. Volcanol.*, *71*, 2009.

Sasai, Y., J. Zlotnicki, Y. Nishida, M. Uyeshima, P. Yvetot, Y. Tanaka, H. Watanabe, and Y. Taka-hashi, Evaluation of electric and magnetic field monitoring of miyake-jima volcano (Central Japan): 1995-1999, *Annali di Geofisica*, *44*, 239–260, 2001.

Scarpa, R., and R. Tilling (Eds.), *Monitoring and Mitigation of Volcano Hazards*, Springer-Verlag, Berlin New York, 1996.

Shima, M., On the second volcanic micro-tremor at volcano Aso, *Bull. Disaster Prev. Res. Inst. Kyoto Univ.*, *22*, 1–6, 1958.

Shimozuru, D., Volcanic microseisms-discussion on the origin, *Bull. Volcanol. Soc. Japan*, *5*, 154–162, 1961.

Shlien, S., and M. Toksoz, A clustering model for earthquake occurrence, *Bull. Seism. Soc. Am.*, *60*, 1765–1787, 1970.

Sieh, K., et al., Near-field investigations of the Landers earthquake sequence, April to July 1992, *Science*, *260*, 171–176, 1993.

Simkin, T., and L. Siebert, *Volcanoes of the World: an Illustrated Catalog of Holocene Volcanoes and their Eruptions.*, Smithsonian Institution, Global Volcanism Program Digital Information Series, GVP-3, (http://www.volcano.si.edu/world/), 2002-.

Stein, R. S., The role of stress transfer in earthquake occurrence, *Nature*, *402*, 605–609, 1999.

Steinberg, G. S., Genesis of volcanic tremors and the long-range forecasting of eruptions, *Dolk. Akad. Nauk.*, *165*, 1294–1297, 1965.

Steinberg, G. S., and A. S. Steinberg, On possible causes of volcanic tremor, *J. Geophys. Res.*, *80*, 1600–1604, 1975.

Sturton, S., and J. Neuberg, The effects of conduit length and acoustic velocity on conduit resonance: Implications for low-frequency events, *J. Volcanol. Geotherm. Res.*, *151*, 319–339, 2006.

Tarantola, A., *Inverse problem theory and methods for model parameter estimation*, SIAM (Society for industrial and applied mathematics), Philadelphia, 2005.

Tarantola, A., and B. Valette, Inverse problems = Quest for information, *J. Geophys.*, *50*, 159–170, 1982a.

Tarantola, A., and B. Valette, Generalized nonlinear inverse problems solved using the least-squares criterion, *Rev. Geophys.*, *20*, 219–232, 1982b.

Tierney, L., Markov chains for exploring posterior distributions, *Ann. Stat.*, *22*, 1701–1762, 1994.

Toda, S., R. S. Stein, and T. Sagiya, Evidence from the ad 2000 Izu islands earthquake swarm that stressing rate governs seismicity, *Nature*, *419*, 58–61, 2002.

Troise, C., G. De Natale, and F. Pingue, A model for earthquake generation during unrest episodes at Campi Flegrei and Rabaul calderas, *Geophys. Res. Lett.*, *24*, 1575–1578, 1997.

Troise, C., F. Pingue, and G. De Natale, Coulomb stress changes at calderas: Modeling the seismicity of Campi Flegrei (southern Italy), *J. Geophys. Res.*, *108*, 2003.

Tsukui, M., and Y. Suzuki, Eruptive history of Miyakejima volcano during the last 7000 years, *Bull. Volcanol. Soc. Japan*, *43*, 149–166, 1998.

Turcotte, D. L., and G. Schubert, *Geodynamics*, 2nd ed., Cambridge University Press, Cambridge, UK, 2002.

Ueda, H., E. Fujita, M. Ukawa, E. Yamamoto, M. Irwan, and F. Kimata, Magma intrusion and discharge process at the initial stage of the 2000 activity of Miyakejima, Central Japan, inferred from tilt and GPS data, *Geophys. J. Int.*, *161*, 891–906, 2005.

Uhira, K., T. Baba, H. Mori, H. Katayama, and N. Hamada, Earthquake swarms preceding the 2000 eruption of Miyakejima volcano, Japan, *Bull. Volcanol.*, *67*, 219–230, 2005.

Ukawa, M., Deep low-frequency earthquake swarm in the mid crust beneath Mount Fuji (Japan) in 2000 and 2001, *Bull. Volcanol.*, *68*, 47–56, 2005.

Ukawa, M., and H. Tsukahara, Earthquake swarms and dike intrusions off the East coast of Izu peninsula, central Japan, *Tectonophysics*, *253*, 285–303, 1996.

Utsu, T., A statistical study on the occurrence of aftershocks, *Geophys. Mag.*, *30*, 521–605, 1961.

Varanasi, M. K., and B. Aazhang, Parametric generalized Gaussian density estimation, *J. Acoust. Soc. Am.*, *86*, 1404–1415, 1989.

Wickman, F. E., *Markov models of repose-period patterns of volcanoes, in: Random Process in Geology*, pp. 135–161, Springer, New York, 1976.

Yokoyama, Y., M. Kumazawa, Y. Imanishi, and N. Mikami, A new method of non-stationary time series analysis based on inhomogeneous AR equation, *IEEE Trans. Signal Process.*, *45*, 2130–2136, 1997.

Zadeh, L., Fuzzy sets, *Inf. Control*, *8*, 338–353, 1965.

Zollo, A., W. Marzocchi, P. Capuano, A. Lomax, and G. Iannaccone, Space and time behavior of seismic activity at Mt. Vesuvius volcano, southern Italy, *Bull. Seism. Soc. Am.*, *92*, 625–640, 2002.