



HAL
open science

Performance of the 'Material Failure Forecast Method' in real-time situations: a Bayesian approach applied on effusive and explosive eruptions

Anaïs Boué, Philippe Lesage, Guillermo Cortés, Bernard Valette, G.
Reyes-Dávila, R. Arámbula-Mendoza, A. Budi-Santoso

► To cite this version:

Anaïs Boué, Philippe Lesage, Guillermo Cortés, Bernard Valette, G. Reyes-Dávila, et al.. Performance of the 'Material Failure Forecast Method' in real-time situations: a Bayesian approach applied on effusive and explosive eruptions. *Journal of Volcanology and Geothermal Research*, 2016, 327, pp.622-633. 10.1016/j.jvolgeores.2016.10.002 . hal-01726569

HAL Id: hal-01726569

<https://hal.science/hal-01726569>

Submitted on 8 Mar 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Performance of the 'Material Failure Forecast Method' in real-time situations: a Bayesian approach applied on effusive and explosive eruptions.

Boué A.^a, Lesage P.^a, Cortés G.^b, Valette B.^{a,c}, Reyes-Dávila G.^d, Arámbula-Mendoza R.^d and Budi-Santoso A.^e

a ISTERre, CNRS, Université Savoie Mont Blanc, Le Bourget-du-Lac, France

b Departamento de Teoría de la Señal, Telemática y Comunicaciones, E. T. S. de Ingenierías Informática y de Telecomunicación. Universidad de Granada, Granada, Spain

c ISTERre, IRD 219, Université Savoie Mont Blanc, Le Bourget-du-Lac, France

d Centro Universitario de Estudios e Investigaciones en Vulcanología, Colima, Mexico.

e Badan Geologi, Jalan Diponegoro No. 57, 40122 Bandung, Indonesia

Abstract

Most attempts of deterministic eruption forecasting are based on the material Failure Forecast Method (FFM). This method assumes that a precursory observable, such as the rate of seismic activity, can be described by a simple power law which presents a singularity at a time close to the eruption onset. Up to now, this method has been applied only in a small number of cases, generally for forecasts in hindsight, i.e. using the complete time series of precursor. In this paper, a rigorous Bayesian approach of the FFM designed for real-time applications is applied. Using an automatic recognition system, seismo-volcanic events are detected and classified according to their physical mechanism and time series of probability distributions of the rates of events are calculated. At each time of observation, a Bayesian inversion provides estimations of the exponent of the power law and of the time of eruption, together with their probability density functions. Two criteria are defined in order to evaluate the quality and reliability of the forecasts.

Thanks to the automatic procedure, long continuous seismic recordings are analysed: 13 years from Volcán de Colima, Mexico, 10 years from Piton de la Fournaise, Reunion Island, France, and several months from Merapi volcano, Java, Indonesia. The new forecasting approach is applied to 64 pre-eruptive sequences which present various types of dominant seismic activity (volcano-tectonic or long-period events) and patterns of seismicity with different level of complexity. This allows us to test the FFM assumptions, to determine in which conditions the method can be applied, and to quantify the success rate of the forecasts. 62% of the precursory sequences analysed in this study are suitable for the application of FFM and half of the total number of eruptions are successfully forecast in hindsight. In real-time, the method allows for the successful forecast of 36% of all the eruptions considered. Nevertheless, real-time forecasts are successful for 83% of the cases that fulfil the reliability criteria. Therefore, good confidence on the method is obtained when the reliability criteria are met.

Keywords: Eruption forecasting, Statistical analysis, Real-time, Volcanic-Seismology, Time-series analysis, Automatic recognition system

1. Introduction

Defining strategies for forecasting eruption is one of the most important goals in volcanology. Eruptions are commonly preceded by an increase in seismic activity and deformation of the edifice (McNutt, 1996). Thus the level of activity or its acceleration are the basic observations used in the forecasting procedures that follow probabilistic or deterministic methodologies. Probabilistic approaches have been applied for short and long-term forecasts over time intervals that are shorter and larger than inter-eruptive times, respectively (Marzocchi and Bebbington, 2012). They are usually based on the historical

activity of a particular volcano (Connor et al., 2003) or on pattern recognition techniques for identifying reliable precursors during emergency unrest (Schmid et al., 2012). Deterministic forecasts combine empirical and physical models to estimate the time of eruption. They are based on laboratory experiments and field observations at volcanoes and unstable slopes which show that the acceleration in the number, energy or amplitude of seismic signals prior to eruptions or slope failures can be described by an empirical power law relating the rate of change of a given precursor $\dot{\Omega}$ to its acceleration $\ddot{\Omega}$ (Fukuzuno, 1985; Voight, 1988, 1989) as

$$\ddot{\Omega} = A\dot{\Omega}^\alpha \quad (1)$$

where the coefficients α and A are empirical constants that determine how the rate changes with time. When $\alpha > 1$, solutions of equation 1 are power laws that involve a singularity at a finite time, which can be interpreted as the material failure or the opening of the magma conduit toward the surface before an eruption. The time when the rate of change becomes infinite can be estimated by fitting a power law to the observations. This application of Voight's law is known as the materials failure forecast method (FFM – Cornelius and Scott, 1993; Main, 1999; Kilburn, 2003; Helmstetter et al., 2004).

Equation 1 can be rewritten in a practical form as:

$$\dot{\Omega}(t) = k \left(1 - \frac{t}{t_f} \right)^{-p} \quad (2)$$

where t_f is the time of rupture, $p = \frac{1}{\alpha - 1}$ and k is a constant with the dimension of $\dot{\Omega}$.

When defining a procedure for forecasting eruptions with the FFM, several issues must be considered:

1. An easy way of using the FFM consists in setting the exponent $\alpha = 2$ which corresponds to a hyperbolic law. In this case, the solution is obtained by linear regression of the inverse rate against time. However, experimental evidences suggest that the exponent α may take other values than 2. For instance, Voight and Cornelius (1991) found typical values near 1.5 for precursory phenomena at Mount St. Helens (USA) from 1980 to 1986. Smith and Kilburn (2010) obtained estimations of α generally between 1 and 2 for the 1991 Mount Pinatubo eruption (Philippines). Consequently, the assumption of setting $\alpha = 2$ is not always well-founded and can be unreliable. Furthermore, its physical basis is not well established and the natural variability of α values is rarely taken into account.
2. The observables commonly used as precursors are the rate of events, the mean amplitude (RSAM) and the mean energy (RSEM) of the raw seismic signal or of the filtered signal (SSAM and SSEM). The results of forecast may depend on the choice of the observable. Furthermore, on volcanoes, the seismic observables used as precursors usually mix together numerous types of seismic event (VT, LP, tremor...) that are associated with different physical mechanisms. Volcano-tectonic (VT) earthquakes are related to the failure of rock surrounding a pressurized body of magma. The acceleration of this type of seismic activity is related with rock fracturing or interaction of pre-existing cracks due to dyke propagation (Main, 1999; Kilburn, 2003; 2012). Precursory activities have also been observed in the rate of long-period (LP) events, which are associated with magma fracture (Neuberg et al., 2006; Lavallée et al., 2008) or with resonances in fluid-filled cavities (Chouet, 1996; Chouet and Matoza, 2013). While several source models of LP event have been proposed, the physical processes that generate the acceleration of this type of seismicity before an eruption are not clearly understood. Nevertheless, to clearly identify acceleration sequences associated with single physical process, it is of paramount importance to analyse the different types of event separately.
3. The data used for the FFM can be either cumulative or non-cumulative values of the observable. Each option has theoretical or practical advantages and drawbacks. For example, cumulative values

produce monotonically increasing time series that are easier to fit with a power law than non-cumulative data, especially in the case of complex pre-eruptive patterns. Furthermore, they do not require binning, which is a practical advantage (Bell et al., 2013). However, when using cumulative values, each data depends on the previous ones. Thus, if the correlation between data are not taken into account, some bias can be introduced in the inversion process and in the estimation of the time of eruption (Greenhough et al., 2009).

4. Although it would be a highly valuable information for decision-makers during crises, the uncertainty on the predicted time of eruption is not provided in most of the studies or it is only roughly approximated. This calculation requires estimation of the errors on the observable and relies on hypothesis on the data structure. For this purpose, Bell et al. (2011; 2013) assumed that earthquake occurrences follow a Poisson distribution (Ogata, 1999). On the other hand, Boué et al. (2015) showed that a Gaussian distribution better describes the data structure and they proposed a method for calculating the data uncertainty.

Taking the above issues into account, Boué et al. (2015) proposed a new Bayesian approach of the FFM for real-time eruption forecasting. They first use an automatic recognition system for the classification of the different types of seismic event. Using the success rates of this classification tool, they compute the probability density function (*pdf*) of the event rates. At each time of observation, the posterior *pdf* of the time of forecast t_f and exponent p (equation 2) are estimated using a Bayesian inversion (Tarantola, 2005). Then the quality of the forecast is evaluated by using the uncertainty on the predicted time t_f , which is derived from the corresponding *pdf*, and the temporal stability of parameter t_f . This approach of the FFM is thus convenient for its integration in volcano monitoring systems.

In this paper, we take advantage of the automatic recognition system to process many years of continuous seismic recordings from three volcanoes. In this large data set, we identify 64 precursory sequences of seismic activity dominantly composed of either VT or LP events. The patterns of precursory seismic unrest present a large diversity of behaviours, with single or multiple accelerations of the level of activity and cases of quiescence of different durations before the onset of the eruptions. Thanks to the large number and diversity of cases, we carry out a statistical analysis of the results in order to evaluate the forecast potential of the FFM. We quantify the number of cases for which the method is applicable, i.e. for which the pattern of seismicity is suitable for reliable forecast. For these cases, we estimate the performance and accuracy of the forecasts based on our approach of the FFM and we compare them to similar published studies that used different types of observable and other implementations of FFM.

2. Methodology

The forecast method is based on the assumption that the precursory seismic activity increases in an accelerated manner and can be described by a simple power law. Therefore, in the cases when the seismicity is constant or increases linearly or exponentially (case $\alpha = 1$) with time, the forecast cannot be made. A second assumption is that the singularity of the power law coincides with the time of eruption. These assumptions will be tested on many cases in the following sections. The data used are the probability density functions (*pdf*) of the non-cumulative rate of seismic events of a given type estimated from the output of an automatic classification system (Benítez et al., 2007; Ibáñez et al., 2009; Cortés et al., 2009). The pdf are calculated in windows centred on each time of observation t_{obs} . The width Δt of these windows, as well as the beginning time t_0 of the power law which is adjusted to the data, must be set. Because their optimal values are not known, several calculations of these parameters are carried out in parallel with different combinations of these parameters. The values that produce the clearest patterns and the most stable results for the time of forecast are selected.

When the complete sequence of pre-eruptive seismic activity is available (i.e. after the eruption), the time of eruption can be estimated by fitting a power law (equation 2) to the data. We call this process

‘forecast in hindsight’. In real-time conditions, this estimation is carried out by fitting the power law to the incomplete sequence. More precisely, at each observation time t_{obs} , i.e. every time a new window of event rates is obtained, the *pdf* of the time of eruption and of the exponent of the power law are estimated by Bayesian inversion. The estimation of parameters t_f and ρ is given by the maximum likelihood of the corresponding *pdf*. Their temporal evolution allows us to define criteria for evaluating the reliability of the forecast. They are based on the assumption that the model parameters would stabilize if a sustained physical process takes place in the volcano and that the uncertainties would decrease when approaching the eruption. Two criteria are thus defined: 1) the stability of the estimated t_f described by the number of data windows which yield close values of t_f , and 2) the decrease of the uncertainty on t_f as advancing towards the eruption. This uncertainty is related with the spreading of the *pdf* that can be quantified by the Shannon index (or Shannon entropy, Shannon, 1948) at each time of observation:

$$I(t_{obs}) = - \int_{t_f} \rho_{post}(t_f) \times \log[\rho_{post}(t_f)] dt_f \quad (3)$$

The lower the Shannon index is, the more relevant the *pdf* of t_f is at the corresponding time of observation. Thus a reliable forecast would be characterized by a stable value of t_f and a decreasing Shannon index before the eruption. Note that the absolute value of this index is not considered here.

3. Precursory seismic activity

The volcanoes Piton de la Fournaise (PdIF), Merapi and Colima have been selected for case studies, because they provide examples of precursory seismicity (1) dominated by VT events, (2) characterised by mixed VT, hybrid and LP events, and (3) dominated by LP events, respectively.

Piton de la Fournaise is a basaltic shield volcano on the French island of La Réunion in the western Indian Ocean. It is monitored by 24 seismic stations and a network of permanent GPS and tiltmeters. About 83 volcanic episodes occurred since 1985: 54 eruptions, two pit crater collapses at the summit, 26 seismic crises that were not followed by an eruption, and one caldera collapse (Peltier, 2007). In this study, we focus on precursory trends recorded during the interval between 2000 and 2010, during which thirty eruptions, two collapses and twenty intrusions were recorded (Roult et al., 2012). All the eruptions were preceded by seismic crises. Their durations are variable, from hours to months, and they do not especially correlate with the duration or the intensity of the precursory seismicity. In addition, many seismic crises were not followed by an eruption and are associated with magmatic intrusions. Thus, if seismic crises had been used alone to forecast an eruption, 46% (25/55) would have triggered a false alarm. At this volcano, relative seismic quiescences of varying duration between the end of the precursory increase of activity and the eruption, followed by a short swarm just before the eruption onset (Roult et al., 2012) are common observations.

Merapi is an andesitic-dacitic volcano located 30 km north of Yogyakarta in Java, Indonesia. Its eruptive behaviour is characterised by the effusion at intervals of 4-6 years of lava domes that collapsed to feed pyroclastic density currents. The effusive behaviour ended in November 2010 with a VEI 4 explosive eruption. The explosion was preceded by about one year of edifice inflation and two months of elevated seismicity and increased emission of gas, which led to the evacuation of some 400,000 people (Surono et al., 2012). The VT hypocenters were clustered between 2.5 km and 5 km below the summit before 17 October and, after then, at depths of less than 1.5 km (Budi-Santoso et al., 2013; 2016). The decrease in depths is associated with the upward migration of a large volume of magma that produced damage in the surrounding medium (Budi-Santoso et al., 2013). The eruption began with an explosion on 26 October, followed by two days of relative seismic quiescence before small explosive events until 1 November which led to rapid dome growth and the main explosion on 5 November. Eruptive activity finally ceased on 8 November. Merapi’s volcanic seismic activity has been subdivided into five main categories: VT events, multi-phase (MP) or hybrid events, LP events, guguran or rockfalls or collapses

(COL), and tremors (T). Their rates of occurrence are shown in figure 1 from August 2010 until the November eruption. The VT event rate shows two clear accelerations, between ~4 October and ~17 October and during 20-26 October. LP seismicity is negligible until 16 October, after which a clear increase takes place with a peak rate on 24 October. The occurrence of rockfalls remained approximately constant at about 100 events per day for about one month, and accelerated between 20 and 26 October. The numbers of multi-phase events remained close to 100 per day for about a month until 12 October, after which they increased irregularly to 600 on 23 October and then declined until the explosion on 26 October.

Volcán de Colima is an andesitic strato-volcano in the western part of the Mexican Volcanic Belt. It is the most active volcano in Mexico and displays a wide spectrum of eruptive styles, including small phreatic explosions, major block-lava effusions and large explosive events (Bretón-González et al., 2002). In this study, we have focused on seismicity before vulcanian explosions of moderate to major size, namely those that have been reported by the civil authorities and the University of Colima and have sent ballistic ejecta to distances of 5 km. During the period 1998-2011, thirty-eight explosions were recorded, of which two occurred when seismic stations were not operating. Among the remaining 36 explosions, 17 were preceded by single accelerations in seismic LP event rate, 7 by multiple accelerations, 7 by a linear increase of event rate, and three occurred without detected seismic precursors. Although the duration of the precursory LP activity varies from 4 h to almost 2 days, most of the durations remain in a range of 4 to 20 h. In many cases, the sequence of increasing seismic activity is followed by a period of relative quiescence preceding the eruption.

4. Examples of real-time forecast

We first show the results of forecasts obtained with simple patterns of event rates acceleration before considering examples of more complex precursory trends. The displayed results include the prior *pdf* of the rates of seismic events, the marginal posterior *pdf* ($\rho_{post}(t_f)$) of the forecast time t_f with its corresponding Shannon index, and the marginal posterior *pdf* ($\rho_{post}(p)$) of the exponent p (equation 2), all as a function of the time of observation t_{obs} . A summary of results for all the eruptions studied is given in tables 1&2.

The eruption of 30 May 2003 of Piton de la Fournaise was preceded by two days of increasing VT event rates (Figure 2a). The acceleration was followed by 8 h of seismic quiescence and a VT swarm that occurred immediately prior the eruption. The optimal window width Δt and starting time t_0 were determined by the method of Boué et al. (2015) at 2h and 2.5 days before the time of eruption t_e , respectively. Accurate forecast was possible since Day 1.58, with a 99% confidence of an eruption between Days 1.8 and 4 (Figure 2c). After that, the maximum likelihood of the forecast time remained close to Day 2.35 with a confidence interval getting narrower until the time of deceleration on Day 2.20, also indicated by the decreasing Shannon index (Figure 2d). During this interval of 16 h, the forecast time of eruption is quite stable and the estimated p -values are of about 2 (Figure 2b), which corresponds to an exponent α of 1.5. An accurate and precise real time forecast, based on the stability of the forecasts and the decreasing Shannon index, could have been made by Day 1.6, i.e. more than half a day before the eruption, with the 99% confidence interval between Days 2 and 2.5.

The October-November 2010 eruption of Merapi was preceded by an accelerating VT event rate for approximately 50 days ($t_e = 51$; Figure 3a). The forecast eruption times are unstable until Day 45 (Figure 3c). After this time, the maximum likelihood eruption time stabilised around Day 55, with an accompanying decrease in the Shannon index (Figure 3d). The 99% confidence interval extends between Days 52 and 59. In real-time, therefore, a forecast could have been made on Day 48, three days before the eruption, although the preferred time was one day later than the true eruption time. It might be possible that the forecast time was delayed by the burst of activity followed by a marked decrease around Day 40. The estimated p -value is close to 1 ($\alpha = 2$; Figure 3b).

The explosion of 5 June 2005 at Volcán de Colima was preceded by a sequence of accelerating LP rates of approximately twenty hours (Figure 4a) with a time of eruption at $t_e = 28$ h. A deceleration of the activity is observed 0.5 h before the explosion. The accelerating sequence of LP events begins at Hour 21 approximately (Figure 4c). The first forecast that seems satisfying when compared with the true eruption time is obtained at Hour 23, i.e. 5 h before the explosion. However its uncertainty is large with a 99% confidence interval ranging between Hours 26 and 60. The estimated value of t_f remains stable until Hour 26 while its precision is increasing, as indicated by the Shannon index (Figure 4d). Moreover, from Hour 24 until the eruption time, the lower bound of the 99% confidence intervals of t_f is equal to the time of observation t_{obs} while the upper bound fluctuates between Hours 30 and 34. Therefore, a forecast could have been made 5 hours before the eruption, with a maximum likelihood around Hour 28 which is very close to the real time of eruption. The maximum likelihood of the p -value varies with time and is close but not equal to 1 in the last part of the acceleration (Figure 4b).

The 2 May 2004 eruption at PdlF volcano was preceded by three days of VT acceleration followed by three days of deceleration in seismic event rate (Figure 5a). In this case, the forecast time never clearly stabilizes (Figure 5c). When the Shannon index decreases, at Day 5.5, the estimated time of eruption is close to $t_f = 6$ days, with the 99% confidence interval between $t_f = t_{obs}$ and $t_f = 10$ days. Thus, the true time of eruption ($t_e = 10.5$ days) is slightly outside the interval. In this case, the forecast would have failed.

The explosion of 8 May 2005 at Volcán de Colima (Figure 6a) was preceded by 15 h of acceleration of LP events rate and by approximately 10 h of deceleration until the explosion (at $t_e = 55$ h). Since Hour 42, the maximum likelihood of the forecast time stabilises for three hours at values of $t_f = 50$ h, with 99% confidence intervals between $t_f = 48$ h and $t_f = 52$ h, while the Shannon index displays a decreasing trend. Thus a forecast would have been done at Hour 45 (i.e. 10 h before the explosion) and the true time of eruption would have fallen 3 h after the upper bound of the confidence interval. Therefore, the real-time forecast would have only given a rough estimation of the time of eruption, although the error of forecast is small in comparison with the time remaining before the explosion.

The explosion that occurred on 27 July 2005 at Colima was preceded by a sequence including three phases of acceleration during more than 70 h in total ($t_e = 72$ h, Figure 7a). The first acceleration phase lasted approximately 30 h and ended at Hour 28. It was followed by approximately 10 h of constant LP activity until Hour 40. Then, another short phase of acceleration and deceleration took place until Hour 60. This complicated pattern ended with a sharp increase of the seismicity over 10 hours, followed by a constant activity of 4 h before the explosion. During the first phase of the acceleration, relatively stable maximum likelihood of t_f between 33 and 37 h is obtained during 4 h, with upper 99% of confidence around $t_f = 49$ h (Figure 7c). Figure 7b shows that p -values stabilise close to 2 ($\alpha = 1.5$). Although the posterior pdf of t_f is widely spread (Figure 7d), an eruption could have been expected until $t_f = 50$ h. The FFM cannot be applied during the phases of deceleration and fluctuation of the seismicity and thus the results are meaningless from Hour 30 to 40. For t_{obs} between 40 and 45 h, i.e. when the second phase of acceleration is included, it is difficult to fit a simple power law to the whole complex sequence. Alternatively a trial was done by setting the origin time t_0 at the beginning of the second phase (i.e. taking $t_0 = 35$ h) and by adjusting the power law to this sub-sequence only. However, the model is poorly constrained, as the data set is too small. The third accelerating phase is longer than the second one and the forecast can be performed by setting the beginning of the fitting window at $t_0 = 50$ h. Even if the acceleration phase is short, the maximum likelihood of t_f stabilises around 70 h for t_{obs} ranging from 65.5 to 68 h (Figure 7f), with a decreasing Shannon index (Figure 7g). The method is thus successful but the stabilisation criterion would have been difficult to use in real-time because of the short duration of the acceleration phase. Note that the p -values are now much less than 1 (Figure 7e). In addition, the time lag between the forecast made with the first acceleration part and the true time of eruption is approximately 35 hours, a value close to the duration of the phases of decrease and fluctuations of seismicity, which appears to have delayed the onset of the eruption.

Precursory patterns of short duration were also observed before some vulcanian explosions at PdlF and Volcán de Colima. In these cases, the number of data points was too small to ensure a well-constrained

and stable fit to the observations. Figure 8a shows the precursory LP pattern of the explosion of 13 March 2005 ($t_e = 24$ h) at Colima. The sequence of acceleration of LP events rate lasts approximately 3 h. The uncertainty of the forecasts becomes acceptable only 1 h approximately before the eruption with $t_f = 24$ h (Figure 8c), leaving insufficient time for strategic decisions for crisis management. On the other hand, the forecast in hindsight appears to be successful.

5. Systematic application and statistical performance

5.1. Global applicability of the method

Tables 1 and 2 summarise the results of forecasts made at Piton de la Fournaise, Volcán de Colima and Merapi. They also report the type of precursory sequence: 'single' for single acceleration, 'multiple' for multiple acceleration, 'increase' for patterns with no clear acceleration, 'constant' for constant anomalous seismicity and 'short swarm' for burst of seismicity. They indicate the events for which hindsight forecast could be made (using the whole accelerating sequence) and those for which forecast could have been carried out in real-time according to our criteria (stable forecasts and decreasing Shannon index). Forecasts were not carried out for the types of precursory seismicity 'constant', 'increase', and 'short swarm'. For five cases, the estimated value of exponent p is equal to or larger than 4, which corresponds to α close to one. This indicates that the corresponding pattern has an exponential trend and that no reliable forecast can be made.

Hindsight forecasts at PdIF have been carried out for 20 among 30 eruptions (67%). In most of the 10 remaining cases, the precursory swarm was too short or its level too low for the FFM to be applied. Only 12 precursory sequences could have been used for real-time forecast, which represents 40% of the eruptions and 60% of the sequences suitable for hindsight forecast. Note that the value $p = 1$, which corresponds to a hyperbolic law ($\alpha = 2$), is comprised in the 99% confidence interval for 9 over 20 (45%) acceleration precursory patterns. Among the 33 explosions analysed at Volcán de Colima, 20 (61%) displayed precursory patterns that were suitable for forecast in hindsight. Real-time forecasts could be performed for 9 eruptions (27% of the total number of cases and 45% of cases suitable for hindsight forecast). Application of the FFM was prevented when swarm durations were too short and when multiple accelerations occurred. Only 6 over 20 (30%) of the suitable sequences are consistent with a p -value of one. Finally, real-time forecasting could have been carried out successfully for the unique eruption which is analysed for Merapi volcano. In this case, the p -value of the power law that describes the precursory sequence is close to one.

In summary for the 3 volcanoes, 41 over 64 eruptions (64%) were preceded by a sequence suitable for hindsight application of the FFM and among them, 22 (54%) could have been forecast in real-time. The statistical performance of the method is similar for Colima and PdIF volcanoes. For over half of the cases analysed, the p -value clearly differs from one. Thus, the use of the classical inverse linear method, in which $p = 1$ is assumed, would have led to biased forecasts.

5.2. Accuracy of the forecasts

We consider a forecast as *successful* when the true time of eruption is encased within the 99% confidence interval of the posterior pdf of t_f ; as *missed* when the reliability criteria were not encountered; and as *informative* when the 99% confidence interval is shorter than the remaining time before eruption.

The reliability of the real-time forecasts depends on their stability and uncertainty. Conditions are *stable* when at least N consecutive forecasts are close together. Three values of N , from 3 to 5 points of stability, are considered here and the closeness of forecast time is defined arbitrarily. If this criterion is fulfilled and the Shannon index decreases, then the forecast is considered as reliable (Tables 1 & 2). Figure 9 shows the effective error of forecasts $t_f - t_e$ as a function of the remaining time before eruption,

each time the criteria are fulfilled. At PdIF and Merapi, a 3-point stability criterion yields 15/16 uncertainty intervals of real-time forecast that contain the true time of eruption (Figure 9, left). For 8/16 cases, the uncertainties are smaller than the remaining time before eruption. Thus, while almost all real-time forecasts would have been successful, half of them were only informative. For a 4-point criterion, 9/16 forecast intervals contain the eruption time while for a 5-point criterion, half of the eruptions are successfully forecast with only one displaying a large uncertainty. At Colima, a 3-point stability criterion yields 8/12 cases with the true time of explosion within the uncertainty interval, of which only 4 had an uncertainty smaller than the remaining time before explosion, 4- and 5- point criteria yield corresponding values of 9/12 and 5/9, and 7/9 and 3/7, respectively. Hence, as the stability criterion becomes more demanding, a smaller proportion of eruption times are correctly forecast, but the correct forecast time become more accurate. Finally, 62% to 75% of the real-time forecasts could have been carried out at least two days before the eruption at PdIF and Merapi and almost all before two hours at Colima. Figure 10 shows the minimum amount of data required to apply deterministic forecast. For a 3-point stability criterion, only 4/28 successful forecasts were made with 60% of the precursory sequence while more than half of the successful forecasts (15/28) were possible with less than 80% of the sequences. For 4- and 5-point stability criteria, respectively 10/19 and 10/16 successful forecasts were possible with less than 80% of the sequences. Figure 11 displays the hindsight forecasts for the 41 accelerating precursory sequences observed at the three volcanoes. 33/41 tested cases gave the true time of eruption within the 99% confidence interval. Only 4/20 forecasts were wrong for PdIF volcano and 3/20 for Colima. The forecast time estimated for the 2010 eruption of Merapi volcano is one day after the true eruption onset.

6. Discussion

Although the material failure forecast method was proposed more than 25 years ago, to our knowledge, only a few tens of cases of eruption forecast were published since then. Thanks to our Bayesian approach of the FFM and to the automatic data processing that allowed very long time series to be analysed, qualitative and quantitative steps forward have been obtained in the present work. The large number of cases presented here gives the opportunity to evaluate the feasibility and the performance of the eruption forecast method using several types of seismic activity in various volcanic contexts. They can also be compared with some published eruption forecastings at Colima and Merapi volcanoes using different approaches and observables.

The precursory seismicity of most eruptions at PdIF consists of single or, more frequently, multiple accelerating patterns followed, after a relative quiescence of varying duration, by a swarm preceding the eruption (Figures 2&5). The acceleration of VT seismicity is thought to be related to the damage of the surrounding rock (Kilburn, 2003, 2012; Main, 2000; Amitrano and Helmstetter, 2006) under the stress induced by magma pressure. The forecast target of the FFM is thus more likely the failure of the edifice rather than the eruption itself. When a large failure occurs, pressure may be released which can result in the seismic quiescence between the acceleration episode and the last VT swarm (Carrier et al., 2015). This terminal swarm is associated with the final propagation of a dyke (Taisne et al., 2011). In these cases, the forecast times are a few hours or days before the eruptions itself. Its accuracy is better when the quiescent period is short and is generally lower for multiple acceleration patterns than for single ones.

Budi-Santoso et al. (2013) carried out a real-time forecast of the 2010 eruption of Merapi volcano using the FFM with the classical linear inverse approach. Their observables are cumulative band-pass filtered RSAM (SSAM). Their better results were produced by using the frequency range from 1 to 3 Hz, which contains mainly VT signals. In this case, they obtained forecast time $t_f = t_e \pm 4$ hours for 5 days before the eruption. Furthermore, when they estimated the exponent α together with t_f , they obtained values close to 2 ($p \approx 1$). For the sake of comparison, we applied the Bayesian approach by setting $p = 1$ (Figure 12) instead of using a variable p -value (Figure 3). Its leads to more stable, although slightly biased, forecast. Indeed, the maximum likelihood of t_f is closer to the true time of eruption t_e when setting $p =$

1, but the 99% confidence interval, which is narrower than with a variable p , does not contain t_e . Thus, the comparison of the results of Budi-Santoso et al. (2013) with those obtained in the present work, using variable or not p -values, emphasizes the influence of the choice of the observable and of the fitting procedure on the estimation of the eruption time. On the other hand, the seismic activity of Merapi is clearly dominated by the rates of rockfalls and multi-phase events in the days preceding the eruption (Figure 1). The MP activity strongly increased since 12 October 2010 and then decreased in the last days before the first explosion. The number of rockfalls presented a marked acceleration during the last week of the pre-eruptive period. This activity may have been mainly triggered by the strong deformation of the crater wall (Suroño et al., 2012). Therefore, the application of the FFM to the total rate of seismic events (Figure 1, bottom panel) appears to be poorly founded and would have probably given unreliable forecasts. These remarks underline the usefulness of signal recognition and the importance of selecting the most significant type of seismo-volcanic event as observable for eruption forecasting.

Arámbula-Mendoza et al. (2011) performed forecasts in hindsight of the vulcanian explosions that occurred in 2005 at Volcán de Colima. They applied the classical linear inverse FFM, with the whole sequence of acceleration as fitting window, and they used as observable the Root Mean Square of the energy of the seismic signal filtered in the interval 1 to 3 Hz (SSEM), which contains mainly LP signals. Among the 11 events that were forecast in both studies, Arámbula-Mendoza et al. (2011) obtained more accurate eruption times in 4 cases while the results of the present work are better for 7 cases. However, Arámbula-Mendoza et al. (2011) did not estimate the uncertainty of their forecasts, which prevents more detailed comparison. The precursory LP swarms at this volcano present both single and multiple acceleration patterns and they are often followed by a relative seismic quiescence. Despite this behaviour, the estimated forecast times coincide in many cases with the time of the eruption onset. However, an important issue for crisis management at Colima is the short duration of the precursory sequences.

Acceleration patterns of LP activity are relatively common prior to vulcanian explosion on volcanoes such as Galeras (Colombia, Gil Cruz and Chouet, 1997), Tungurahua (Ecuador, Molina et al., 2004), Sakurajima (Japan, Maryanto et al., 2008), Ubinas (Peru, Traversa et al., 2011) and Volcán de Colima. However, the source mechanisms of LP events and the physical processes involved in the acceleration of their activity are still not well understood. Several source models of LP events, including oscillations of fluid-filled cavities, brittle fractures within magmas or slow ruptures have been proposed (Neuberg et al., 2006; Chouet and Matoza, 2013; Bean et al., 2013). Geological observations (Tuffen et al., 2003; Tuffen and Dingwell, 2005), laboratory experiments (Tuffen et al., 2008; Lavallée et al., 2011), models of magma conduits (Neuberg et al., 2006; Goto, 1999), and theoretical developments (Ichihara and Rubin, 2010) suggest that, in the case of viscous magma intrusions, LP events can be generated by brittle fracturing of the ascending magma due to large strain rates close to the conduit walls. Laboratory experiments on lavas from Colima showed that complete sample failure can be forecast using the FFM and acoustic emissions (Lavallée et al., 2008; 2008). The acceleration of seismicity can result from mechanisms of strain localization and reduction in the friction between the ascending magma and the solid conduit walls (Hale and Muhlhaus, 2007). On the other hand, following the model of Holland et al. (2011), brittle fractures of magmas can generate a network of cracks that progressively become interconnected. An explosion may then occur when gases filling the network of fractures decompress on approaching the surface. A better understanding of these complex mechanisms would help interpreting the single or multiple acceleration patterns of LP activity, as well as the deceleration and quiescence observed before some explosions.

The Bayesian approach proposed by Boué et al. (2015) provides a rigorous way of implementing the FFM. Among the 64 precursory sequences studied in the present paper, 35 cases (55%) lead to correct forecast in hindsight, i.e. using the complete sequence, while 23 (36%) would yield successful forecast in real-time, i.e. before the eruption. Moreover, when the reliability criteria are fulfilled, the success rate is quite high (about 83% according to the stability criteria used) which implies only few wrong forecasts.

The forecast method is based on several assumptions described in §2. When one of the assumptions fails, the FFM cannot be applied or produces wrong forecast. This occurs when the seismicity pattern is

not correctly described by a power law, for example for multiple acceleration patterns, when there is a long quiescence between the estimated failure time and the eruption time, or when no temporal stability of t_f is obtained. The use of complex precursory patterns for eruption forecasting would require to understand better the physical processes involved in the VT or LP seismic activity and their relationship with the triggering of eruptions. Another condition for a forecast to be effective is that the pre-eruptive sequence is long enough and, more specifically, that the delay between the moment when the forecast is delivered and the estimated time of eruption is sufficient for civil protection management. This condition was not always fulfilled at Volcán de Colima in 2005. Finally, the forecasting algorithm presented in this paper must not be applied as a black box. Operators in volcano observatories must be aware of the limitations of the FFM and must keep a permanent look at the observations, including raw and processed data.

Although the present work has considered the largest set of precursory sequences ever used for testing eruption forecasting methods, a more statistically significant evaluation of the global performance of the FFM still remains to be done. This would require to process the pre-eruptive seismic activity of a large variety of eruptions from various types of volcano. Nevertheless, the forecasting system presented here was designed as a tool for supporting decision-makers and it can be integrated in volcano monitoring systems in complement of other approaches such as the Bayesian event trees (Marzocchi. et al., 2008). Thus, it still has to be tested in ‘true real-time conditions’ during eruptive crises. Improved architectures of recognition system, such as parallel implementation (Cortés et al., 2014), and unsupervised classification are promising approaches that should also be explored, especially in the cases of dormant volcano unrest.

7. Conclusions

Thanks to the automatic signal processing tools described in the present paper and in Boué et al. (2015), very long continuous seismic recordings – 13 years at Colima, 10 years at Piton de la Fournaise and several months at Merapi – have been analysed. A Bayesian approach of the material Failure Forecast Method has been tested on the precursory seismic activity of 64 eruptions. The studied cases include different patterns of seismicity, seismic events of various types (VT and LP), as well as basaltic and andesitic volcanoes. This large diversity of situations allowed to extract statistics on the success rate of the forecasts and to draw the following conclusions.

1. Only 64% of the precursory sequences analysed were suitable for applying the FFM. This limited proportion of suitable sequences is due to the complexity of the precursory patterns which reflects the variability of the precursory processes before eruptions.
2. Reference cases with single acceleration yielded successful real-time forecasts several hours or days before the eruption, according to the volcano and the duration of the precursory sequence, with uncertainties smaller than the remaining time. However, these simple cases represented only 6/64 eruptions and so are not representative of general precursory conditions.
3. Half of the studied eruptions were successfully forecast in hindsight but only 36% in real-time. Hence, hindsight forecasts cannot be used to evaluate the performance of the method in real time.
4. When the criteria of stability and decreasing Shannon index were fulfilled, 83% of the forecasts were successful in real time.
5. Successful forecasts of about one third of the vulcanian explosions of Volcán de Colima could be carried out using LP events, including cases of decelerating LP rates some hours before the explosion.
6. The best-fit power laws to precursory accelerations commonly have p -values different from 1 ($\alpha \neq 2$). Forecasting methods should therefore take into account the non-linear behaviour in inverse event rate with time.

7. It is important to separate and to use as observable the most significant type of seismic activity for each volcano, discarding signals of different origins. For this task, the automatic classification systems are valuable tools that can process signals in real-time, even in case of seismic crisis.

Acknowledgements.

We thank the Piton de La Fournaise Observatory (OVPF) team for making their data available and maintaining the seismic network. We acknowledge Carlos Ariel Ramírez, Alejandro Martínez Fierros and Miguel González Amuezca from RESCO (Red Sísmica del Estado de Colima) for their support in keeping the seismic network working. The continuous seismic data have been provided by RESCO. We thank AXA Research Funds for the PhD fellowship funding. This work has also been partially supported by the ANR-12-BS06-0012-01 DOMERAPI project of the Agence Nationale de la Recherche, by the ECOS-Nord program, by the Spanish projects Ephestos, CGL2011-470 29499-C02-01 and APASVO (TC2012-31551), and by the EU project EC-FP7 MEDiterranean SUPersite Volcanoes (MED471 SUV). Finally, we are particularly grateful to the reviewers for their constructive comments that greatly improved the manuscript.

References

- Amitrano, D., Helmstetter, A., 2006. Brittle creep, damage, and time to failure in rocks. *Journal of Geophysical Research* 111, B11201, doi: 10.1029/2005JB004252.
- Arámbula-Mendoza, R., Lesage, P., Valdés-González, C., Varley, N.R., Reyes-Davila, G. and Navarro, C., 2011. Seismic activity that accompanied the effusive and explosive eruptions during the 2004-2005 period at Volcán de Colima, Mexico. *J. Volcanol. Geotherm. Res.*, 205: 30-46.
- Bean, C.J., De Barros, L., Lokmer, I., Métaxian, J.-P., O'Brien, G.S. and Murphy, S., 2013. Long-period seismicity in the shallow volcanic edifice formed from slow-rupture earthquakes. *Nature Geoscience*, 7: 71-75.
- Bell, A.F., Greenhough, J., Heap, M.J. and Main, I.G., 2011. Challenges for forecasting based on accelerating rates of earthquakes at volcanoes and laboratory analogues. *Geophys. J. Int.*, 185(2): 718-723.
- Bell, A.F., Naylor, M. and Main, I.G., 2013. The limits of predictability of volcanic eruptions from accelerating rates of earthquakes. *Geophys. J. Int.*, doi: 10.1093/gji/ggt191.
- Benítez, C., Ramírez, J., Segura, J.C., Ibáñez, J.M., Almendros, J., García-Yeguas, A. and Cortés, G., 2007. Continuous HMM-based seismic event classification at Deception Island, Antarctica. *IEEE Transactions On Geoscience and Remote Sensing*, 45: 138-146, doi 10.1109/TGRS.2006.882264.
- Boué, A., Lesage, P., Cortés, G., Valette, B. and Reyes-Davila, G., 2015. Real-time eruption forecasting using the material Failure Forecast Method with a Bayesian approach. *J. Geophys. Res.*, 120: 2143-2161.
- Bretón-González, M., Ramirez, J.J. and Navarro, C., 2002. Summary of the historical eruptive activity of Volcán De Colima, Mexico 1519–2000. *J. Volcanol. Geotherm. Res.*, 117(1-2): 21-46.
- Budi-Santoso, A., Lesage, P., Dwiyono, S., Sumarti, S., Subandriyo, Surono, Jousset, P. and Metaxian, J.-P., 2013. Analysis of the seismic activity associated with the 2010 eruption of Merapi Volcano, Java. *J. Volcanol. Geotherm. Res.*, 261: 153-170.
- Budi-Santoso, A. and Lesage, P., 2016. Velocity variations associated with the large 2010 eruption of Merapi volcano, Java, retrieved from seismic multiplets and ambient noise cross-correlation. *Geophys. J. Int.*, doi: 10.1093/gji/ggw145.
- Carrier, A., Got, J.-L., Peltier, A., Ferrazzini, V., Staudacher, T., Kowalski, P. and Boissier, P., 2015. A damage model for volcanic edifices: Implications for edifice strength, magma pressure, and eruptive processes. *J. Geophys. Res.*, 120: 567-583.
- Chouet, B., 1996. Long-period volcano seismicity: its source and use in eruption forecasting. *Nature*, 380: 309-316.
- Chouet, B.A. and Matoza, R.S., 2013. A multi-decadal view of seismic methods for detecting precursors of magma movement and eruption. *J. Volcanol. Geotherm. Res.*, 252: 108-175.
- Connor, C.B., Sparks, R.S.J., Mason, R.M., Bonadonna, C. and Young, S.R., 2003. Exploring links between physical and probabilistic models of volcanic eruptions: The Soufrière Hills volcano, Montserrat. *Geophys. Res. Lett.*, 30(13): 1701, doi:10.1029/2003GL017384.
- Cornelius, R.R. and Scott, P.A., 1993. A materials failure relation of accelerating creep as empirical description of damage accumulation. *Rock Mechanics and Rock Engineering*, 26(3): 233-252.

- Cortés, G., Arámbula R., Gutiérrez L., Benítez C., Ibañez J. and Lesage P., 2009. Evaluating robustness of a HMM-based classification system of volcano-seismic events at Colima and Popocatepetl Volcanoes. IEEE International Symposium on Geoscience and Remote Sensing, Cape Town, South Africa.
- Cortés, G., García, L., Álvarez, I., Benítez, C., De la Torre, A. and Ibañez, J., 2014. Parallel System Architecture (PSA): An efficient approach for automatic recognition of volcano-seismic events. *J. Volcanol. Geotherm. Res.*, 271: 1-10.
- Fukuzuno, T., 1985. A method to predict the time of slope failure caused by rainfall using the increase number of velocity of surface displacement. *Journal of Japan Landslide Society*, 22(2): 8-14.
- Gil Cruz, F. and Chouet, B.A., 1997. Long-period events, the most characteristic seismicity accompanying the emplacement and extrusion of a lava flow in Galeras volcano, Colombia, in 1991. *J. Volcanol. Geotherm. Res.*, 77: 121-158.
- Goto, A., 1999. A new model for volcanic earthquake at Unzen volcano: Melt rupture model. *Geophys. Res. Lett.*, 26: 2541-2544.
- Greenough, J., Bell, A., Main, I., 2009. Comment on "relationship between accelerating seismicity and quiescence, two precursors to large earthquakes" by Arnaud Mignan and Rita Di Giovambattista. *Geophys. Res. Lett.* 36, 10.1029/2009GL039846.
- Hale, A.J. and Muhlhaus, H.-B., 2007. Modelling shear bands in a volcanic conduit: Implication for over-pressures and extrusion rates. *Earth and Planetary Science Letters*, 263: 74-87.
- Helmstetter, A., Sornette, D., Grasso, J.-R., Andersen, J., Gluzman, S., Pisarenko, V., 2004. Slider block friction model for landslides: Application to Vaiont and la Clapière landslides. *J. Geophys. Res.*, 109, B02409, doi: 10.1029/2002JB002160.
- Holland, A.P., Watson, I.M., Phillips, J.C., Caricchi, L. and Dalton, M.P., 2011. Degassing process during lava dome growth: Insight from Santiaguito lava dome, Guatemala. *J. Volcanol. Geotherm. Res.*, 202: 153-166.
- Ibañez, J.M., Benítez, C., Gutiérrez, L.A., Cortés, G., García-Yeguas, A. and Alguacil, G., 2009. The classification of seismo-volcanic signals using Hidden Markov Models as applied to the Stromboli and Etna volcanoes. *J. Volcanol. Geotherm. Res.*, 187(3-4): 218-226.
- Ichihara, M. and Rubin, M.B., 2010. Brittleness of fracture in flowing magma. *J. Geophys. Res. Solid Earth*, 115(B12): B12202.
- Kilburn, C., 2003. Multiscale fracturing as a key to forecasting volcanic eruptions. *J. Volcanol. Geotherm. Res.*, 125: 271-289.
- Kilburn, C., 2012. Precursory deformation and fracture before brittle rock failure and potential application to volcanic unrest. *J. Geophys. Res.*, 117(B2): B02211.
- Lavallée, Y., Meredith, P.G., Dingwell, D.B., Hess, K.-U., Wassermann, J., Cordonnier, B., Gerik, A. and Kruhl, J.H., 2008. Seismogenic lavas and explosive eruption forecasting. *Nature*, 453: 507-510.
- Lavallée, Y., Benson, P., Heap, M.J., Flaws, A., Hess, K.-U. and Dingwell, D.B., 2011. Volcanic conduit failure as a trigger to magma fragmentation. *Bull. Volcanol.*: DOI 10.1007/s00445-00011-00544-00442.
- Main, I.G., 1999. Applicability of time-to-failure analysis to accelerated strain before earthquakes and volcanic eruptions. *Geophys. J. Int.*, 139: F1-F6.
- Main, I.G., 2000. A damage mechanics model for power-law creep and earthquake aftershock and foreshock sequences. *Geophys. J. Int.*, 142: 151-161.
- Maryanto, S., Iguchi, M. and Tameguri, T., 2008. Constraints on the source mechanism of harmonic tremors based on seismological, ground deformation, and visual observations at Sakurajima volcano, Japan. *J. Volcanol. Geotherm. Res.*, 170(3-4): 198-217.
- Marzocchi, W., Sandri, L. and Selva, J., 2008. BET_EF: a probabilistic tool for long- and short-term eruption forecasting. *Bull. Volcanol.*, 70: 623-632.
- Marzocchi, W. and Bebbington, M., 2012. Probabilistic eruption forecasting at short and long time scales. *Bull. Volcanol.*, 74: 1777-1805.
- McNutt, S.R., 1996. Seismic monitoring and eruption forecasting of volcanoes: A review of the State-of-the-Art and case histories. In: R. Scarpa and R.I. Tilling (Editors), *Monitoring and mitigation of volcano hazards*. Springer-Verlag, Berlin, pp. 99-146
- Molina, I., Kumagai, H. and Yepes, H., 2004. Resonances of a volcanic conduit triggered by repetitive injections of an ash-laden gas. *Geophys. Res. Lett.*, 31(3): L03603 10.1029/2003GL018934.
- Neuberg, J., Tuffen, H., Collier, L., Green, D.N., Powell, T.W. and Dingwell, D.B., 2006. The trigger mechanism of low-frequency earthquakes on Montserrat. *J. Volcanol. Geotherm. Res.*, 153: 37-50.
- Ogata, Y., 1999. Seismicity analysis through point-process modeling: A review. *Pure Applied Geophysics*, 155, 471-507
- Peltier, A., 2007. Suivi, modélisation et évolution des processus d'injections magmatiques au Piton de la Fournaise (Réunion). Ph.D. thesis, Institut de Physique du Globe de Paris, St Denis, France.

- Roult, G., Peltier, A., Taisne, B., Staudacher, T., Ferrazzini, V., Di Muro, A., and the OVPF Team, 2012. A new comprehensive classification of the Piton de la Fournaise activity spanning the 1985-2010 period. Search and analysis of short-term precursors from broad-band seismological station. *J. Volcanol. Geotherm. Res.*, 241-242, 78-104.
- Schmid, A., Grasso, J. R., Clarke, D., Ferrazzini, V., Bachelery, P., Staudacher, T., 2012. Eruption forerunners from multiparameter monitoring and application for eruptions time predictability (Piton de la Fournaise). *J. Volcanol. Geotherm. Res.*, 117, B11203, doi: 10.1029/2012JB009167.
- Shannon, C., 1948. A mathematical theory of communication. *The Bell System Technical Journal* 27, 379-423.
- Smith, R. and Kilburn, C., 2010. Forecasting eruptions after long repose intervals from accelerating rates of rock fracture: The June 1991 eruption of Mount Pinatubo, Philippines. *J. Volcanol. Geotherm. Res.*, 191: 129-136.
- Surono, Jousset, P., Pallister, J., Boichu, M., Buongiorno, M.F., Budisantoso, A., Costa, F., Andreastuti, S., Prata, F., Schneider, D., Clarisse, L., Humaida, H., Sumarti, S., Bignami, C., Griswold, J., Carn, S., Oppenheimer, C. and Lavigne, F., 2012. The 2010 explosive eruption of Java's Merapi volcano—A '100-year' event. *J. Volcanol. Geotherm. Res.*, 241-242: 121-135.
- Taisne, B., Brenguier, F., Shapiro, N.M. and Ferrazzini, V., 2011. Imaging the dynamics of magma propagation using radiated seismic intensity. *Geophys. Res. Lett.*, 38(4): L04304.
- Tarantola, A., 2005. *Inverse Problem Theory and Methods for Model Parameter Estimation*. Society for Industrial and Applied Mathematics, Philadelphia.
- Traversa, P., Lengliné, O., Macedo, O., Métaixian, J.-P., Grasso, J.-R., Inza, A. and Taipei, E., 2011. Short term forecasting of explosions at Ubinas volcano, Peru. *J. Geophys. Res.*, 116: B11301, doi:11310.11029/12010JB008180.
- Tuffen, H., Dingwell, D.B. and Pinkerton, H., 2003. Repeated fracture and healing of silicic magma generate flow banding and earthquakes? *Geol. Soc. Am.*, 31(12): 1089-1092.
- Tuffen, H. and Dingwell, D.B., 2005. Fault textures in volcanic conduits: evidence for seismic trigger mechanisms during silicic eruptions. *Bull. Volcanol.*, 67: 370-387.
- Tuffen, H., Smith, R. and Sammonds, P., 2008. Evidence for seismogenic fracture of silicic magma. *Nature*, 453(22): 511-514.
- Voight, B., 1989. A relation to describe rate-dependent material failure. *Science*, 243: 200-203.
- Voight, B., 1988. A method for prediction of volcanic eruptions. *Nature*, 332: 125-130.
- Voight, B. and Cornelius, R.R., 1991. Prospects for eruption prediction in near real-time. *Nature*, 350: 695-698.

Figure captions

Figure 1: Histograms of seismicity obtained by automatic classification of volcano-seismic events for the 2010 eruption of Merapi volcano. From top to bottom: number of seismic events per day as a function of time, observed at station PUS, for the classes of VT, LP, rockfalls (COL), MP, and all events together, respectively. The red lines represent the main explosions and the grey area the period of eruptive tremor.

Figure 2: a) Prior probability density functions (pdf) of the rate of events (grey scale) for the 30 May 2003 eruption at Piton de la Fournaise (PdIF) volcano, as a function of the observation time t_{obs} . The true time of eruption is $t_e = 2.4$ days (red dashed line). The black dashed line indicates the time of deceleration. b) Posterior marginal pdfs of exponent p . The red line is the maximum likelihood of the pdfs and the yellow, green, and blue lines indicate the 85%, 95% and 99% intervals of confidence, respectively. c) Posterior marginal pdfs of the forecast time t_f with its maximum likelihood (red line) and the 85%, 95% and 99% intervals of confidence (yellow, green, and blue lines, respectively). The black line corresponds to $t_f = t_{obs}$. d) Shannon's index of the marginal pdfs of t_f .

Figure 3: Same as figure 2 for the 26 October 2010 eruption of Merapi volcano. The true time of eruption is $t_e = 51$ days (red dashed line).

Figure 4: Same as figure 2 for the 5 June 2005 explosion at Volcán de Colima. The true time of eruption is $t_e = 28$ h.

Figure 5: Same as figure 2 for the 2 May 2004 eruption at PdIF volcano. The true time of eruption is $t_e = 10.25$ days.

Figure 6: Same as figure 2 for the 8 May 2005 explosion at Colima volcano. $t_e = 55$ h.

Figure 7: Panels a-d: same as figure 2 for the 27 July 2005 explosion at Colima volcano. The true time of eruption is $t_e = 72$ h. Panels e-g: same as b-d for a beginning of the fitting window at $t_0 = 50$ h.

Figure 8: Same as figure 2 for the 13 March 2005 explosion at Colima volcano. $t_e = 24$ h.

Figure 9: Results of real-time forecasts showing the error $t_f - t_e$ as a function of the time remaining before the eruption. The results are reported for three different stability criteria, considering 3-, 4- or 5-point stability. Eruptions for which the criteria are fulfilled only one time are represented in black and eruptions for which stability is reached several times are denoted in color. Black dashed lines represent $t_{obs} = t_f$. Left panel: results obtained for PdIF (dot) and Merapi (square). Right panel: results obtained for Volcán de Colima.

Figure 10: Relative error of forecast $(t_f - t_e)/t_e$ as a function of the proportion of precursory sequence used to make the real-time forecast, for all the eruptions studied from the three volcanoes. The results are reported for three stability criteria, using 3-, 4- or 5-point stability. Eruptions for which the criteria are fulfilled only one time are represented in black and eruptions for which stability is reached several times are denoted in color. Black dashed lines represent $t_{obs} = t_f$.

Figure 11: Results of forecast carried out in hindsight with the whole precursory sequence, for all the studied eruptions. The error of forecast $(t_f - t_e)$ is represented for each case. Error bars correspond to the limits of the 99% confidence intervals.

Figure 12: Same as figure 3 for the 26 October 2010 eruption at Merapi volcano, with the p-value set to one.

Table 1: Results of eruption forecasts carried out on Piton de la Fournaise volcano. The columns correspond to: date of eruption; type of forecast (one and two stars indicate hindsight only or both hindsight and real-time forecasts, respectively); type of pattern (see main text §5); duration of precursors; duration of quiescence (delay between end of acceleration and eruption); lag between hindsight forecast time t_f and true eruption time t_e (maximum likelihood, lower and upper bounds of 99% confidence interval); p -value of sequence with 99% confidence interval; window width Δt used for event counting; duration of stable forecast (in unit of Δt) and fulfilment (yes or not) of reliability criteria. All times in days.

Table 2. Same as table 1 for Volcán de Colima and Merapi volcanoes. All times in hours for Colima, in days for Merapi.

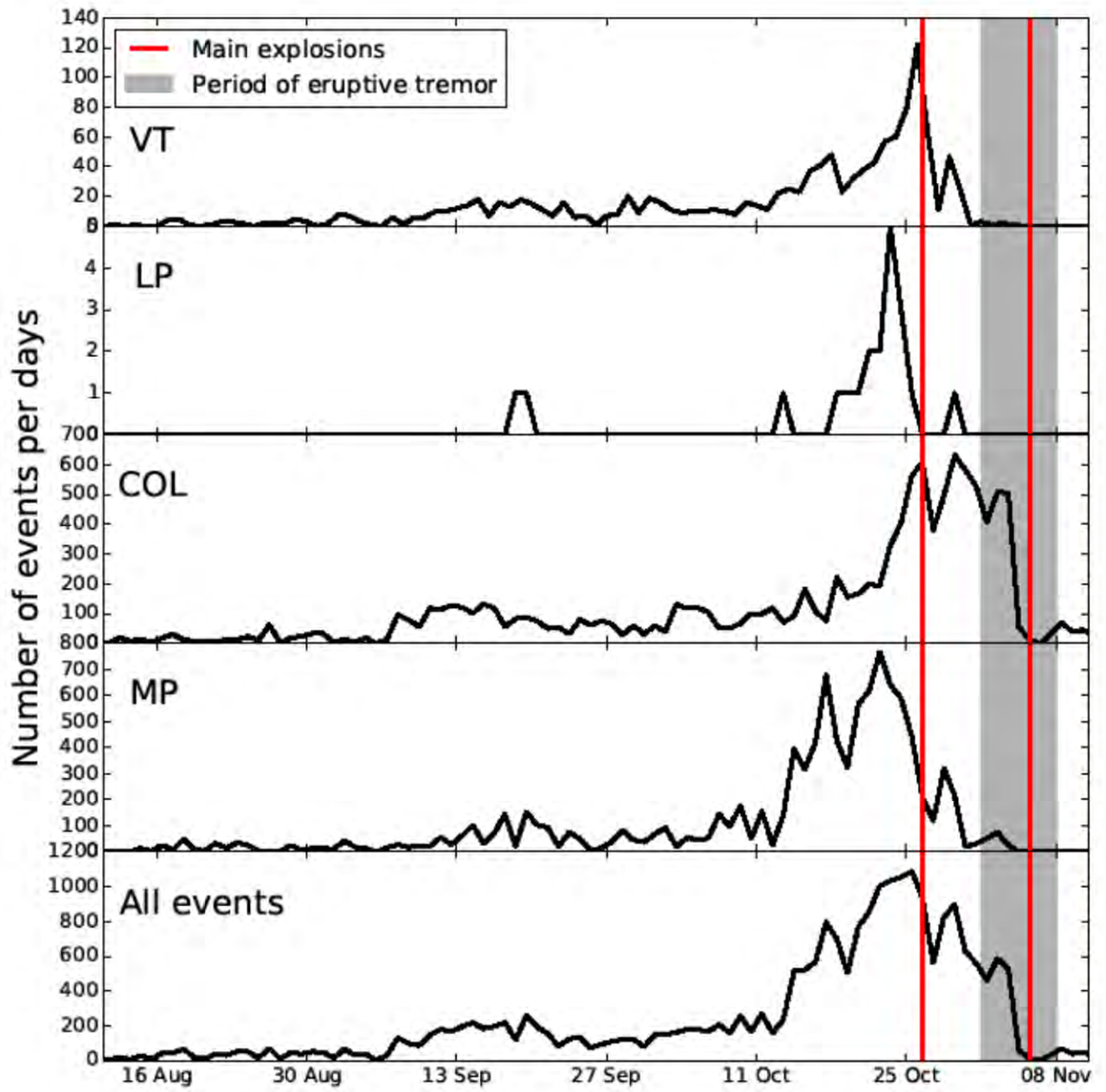


Figure 1

PdIF, 30 May 2003

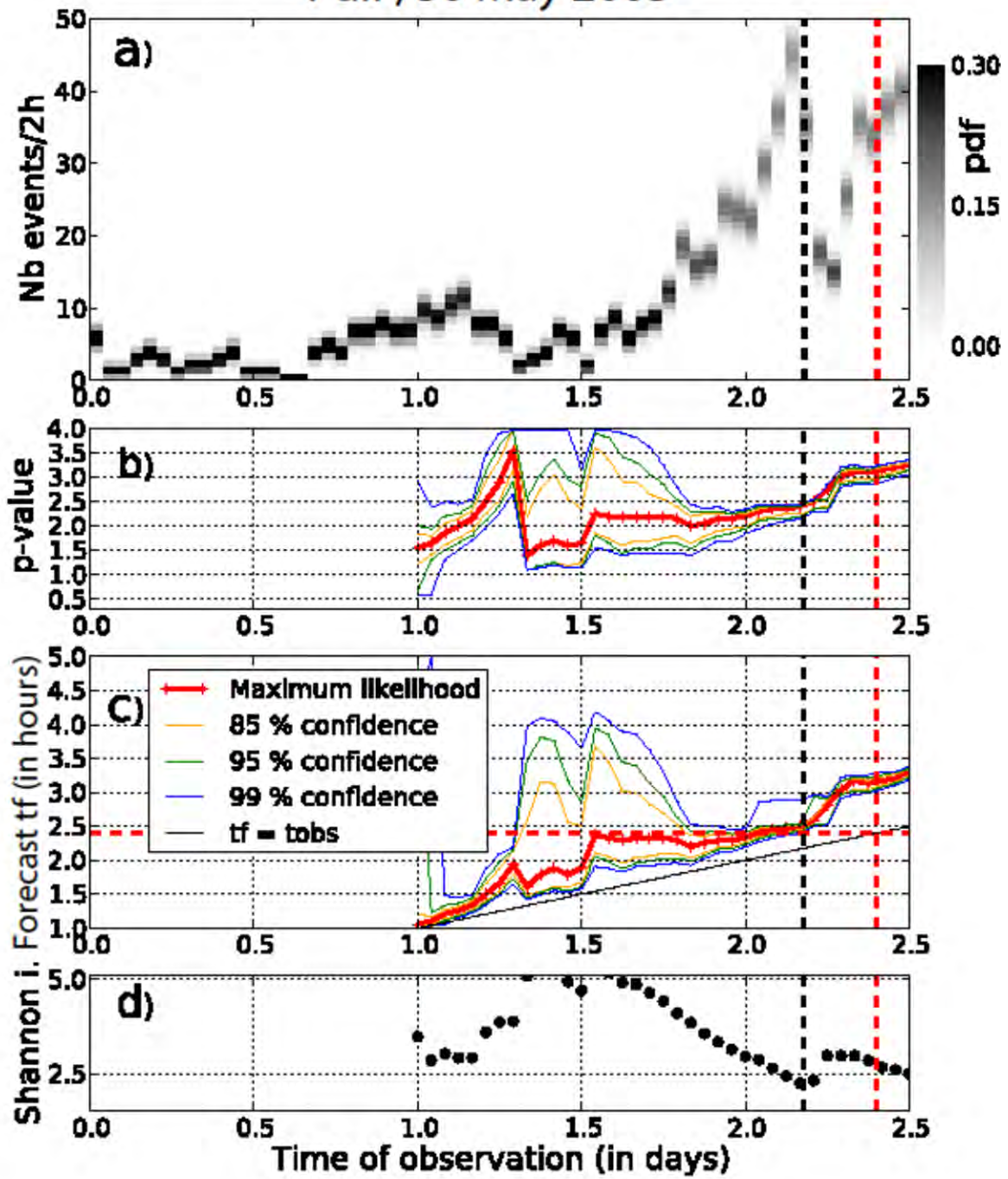


Figure 2

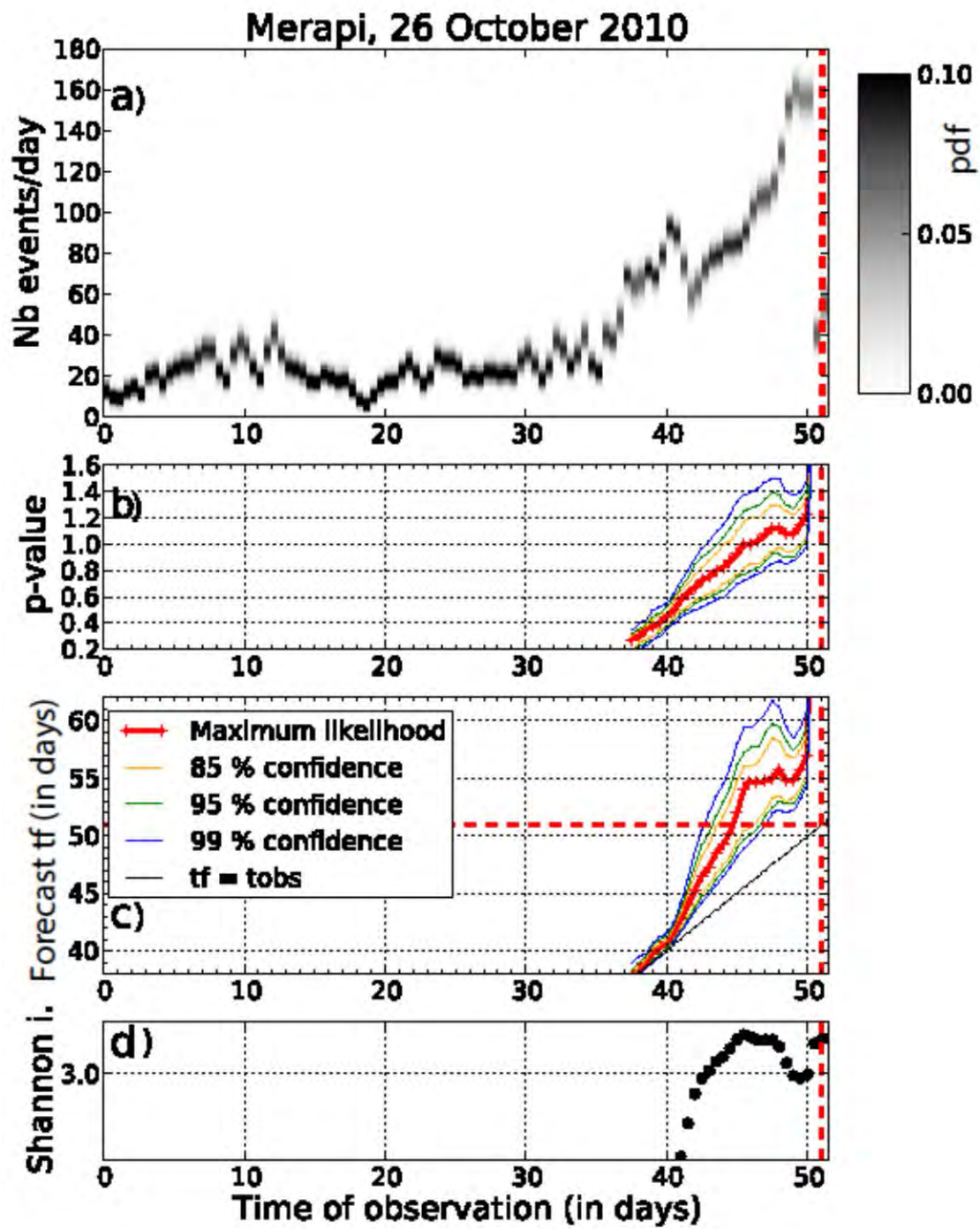


Figure 3

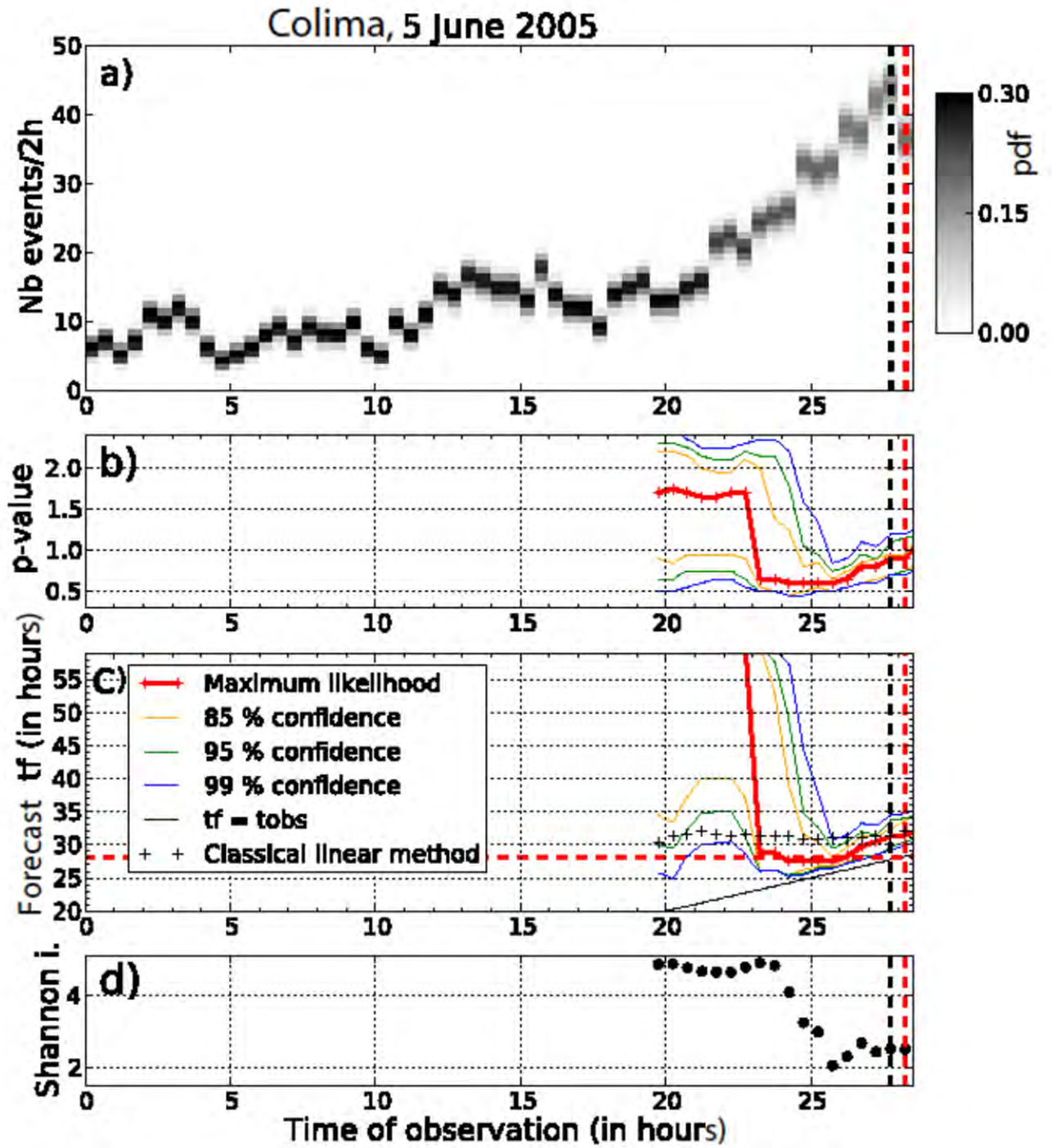


Figure 4

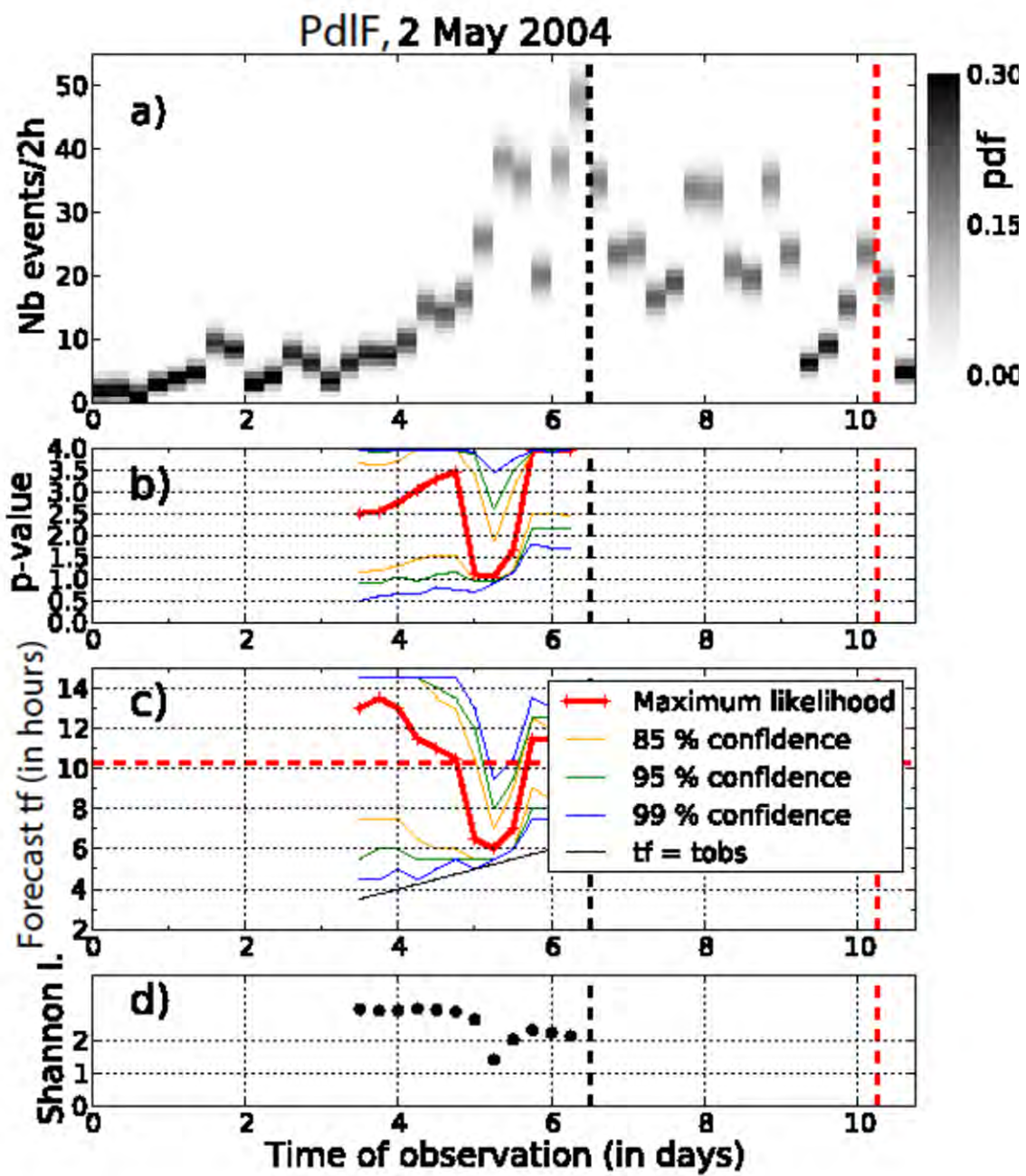


Figure 5

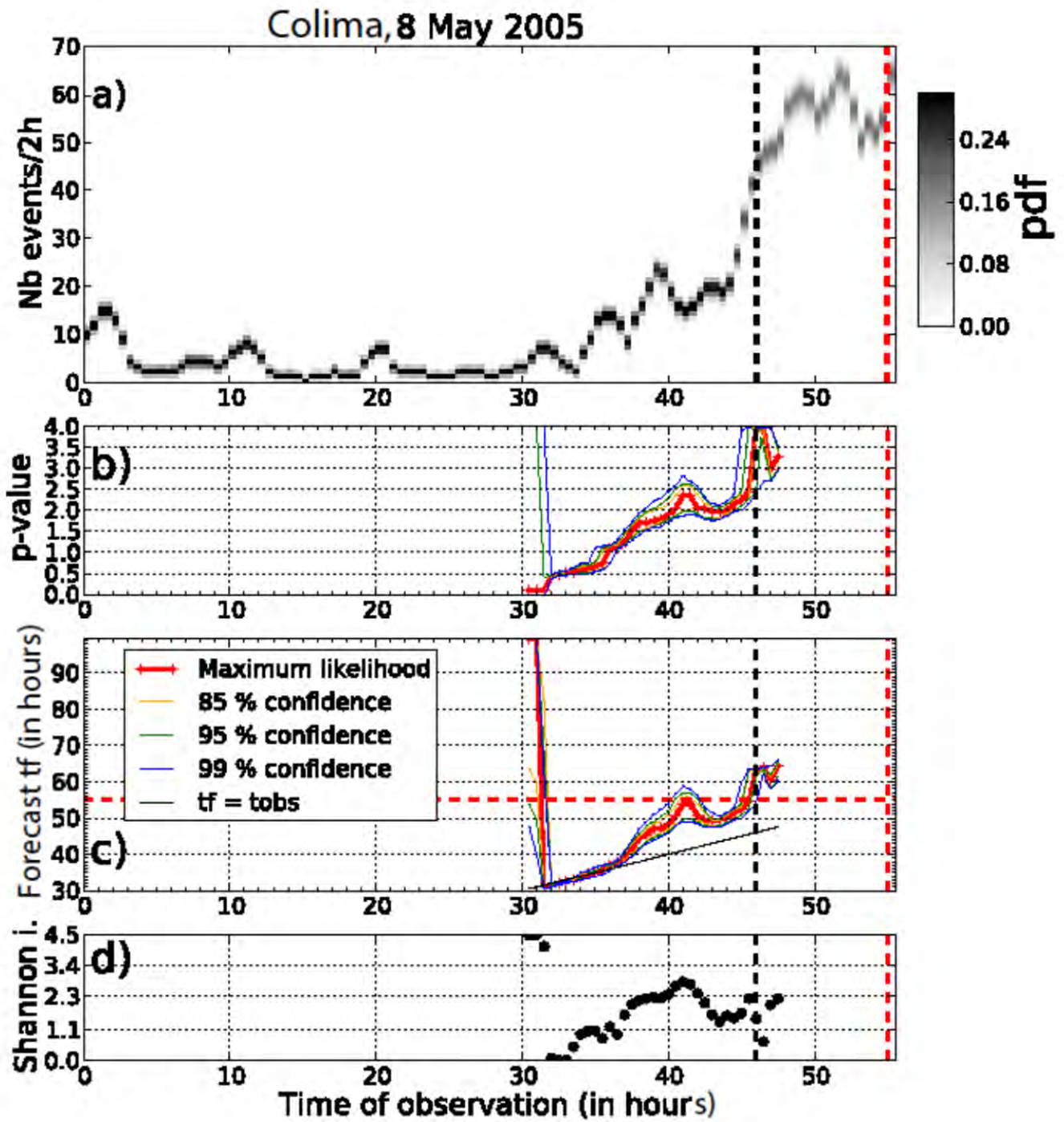


Figure 6

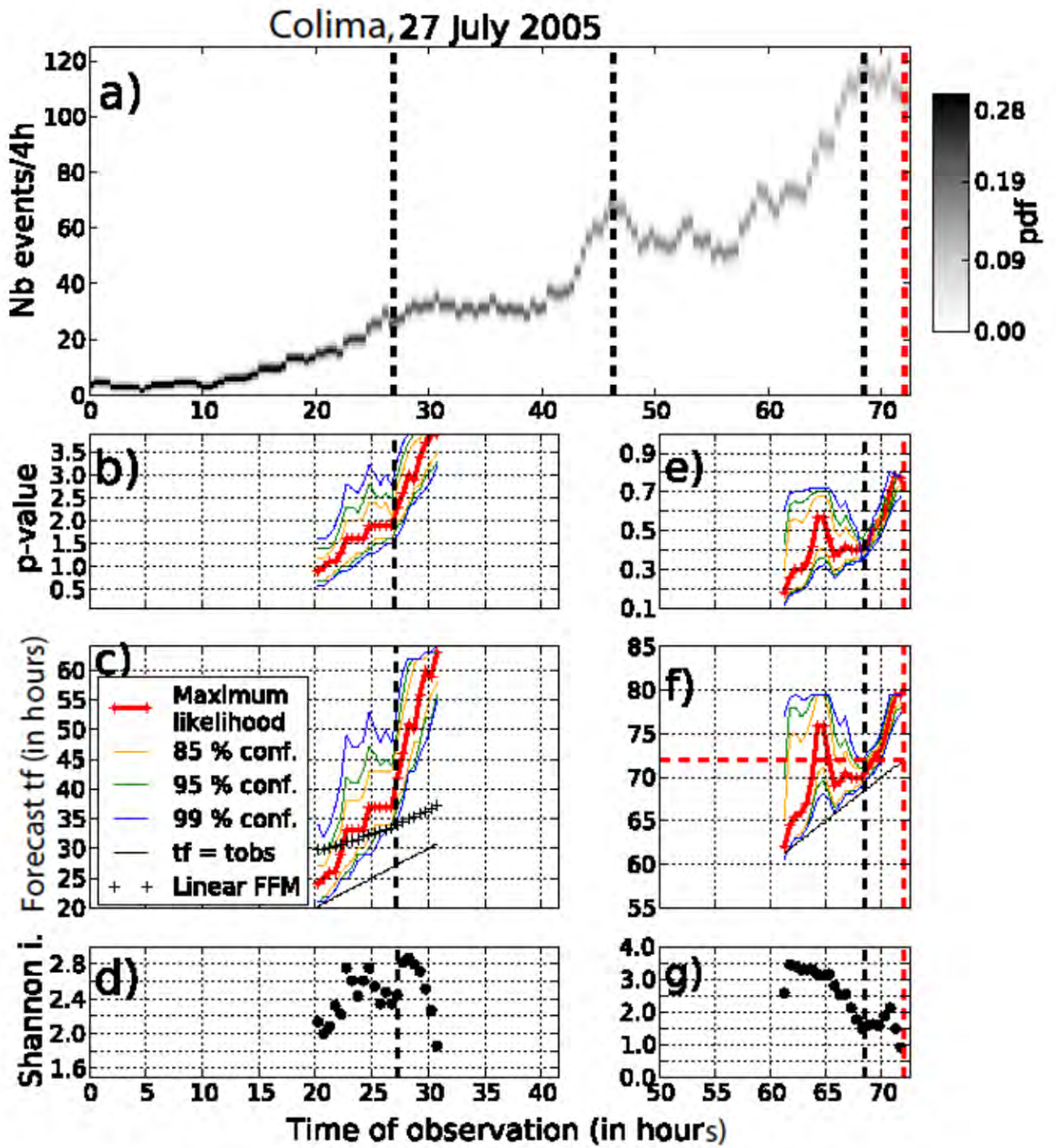


Figure 7

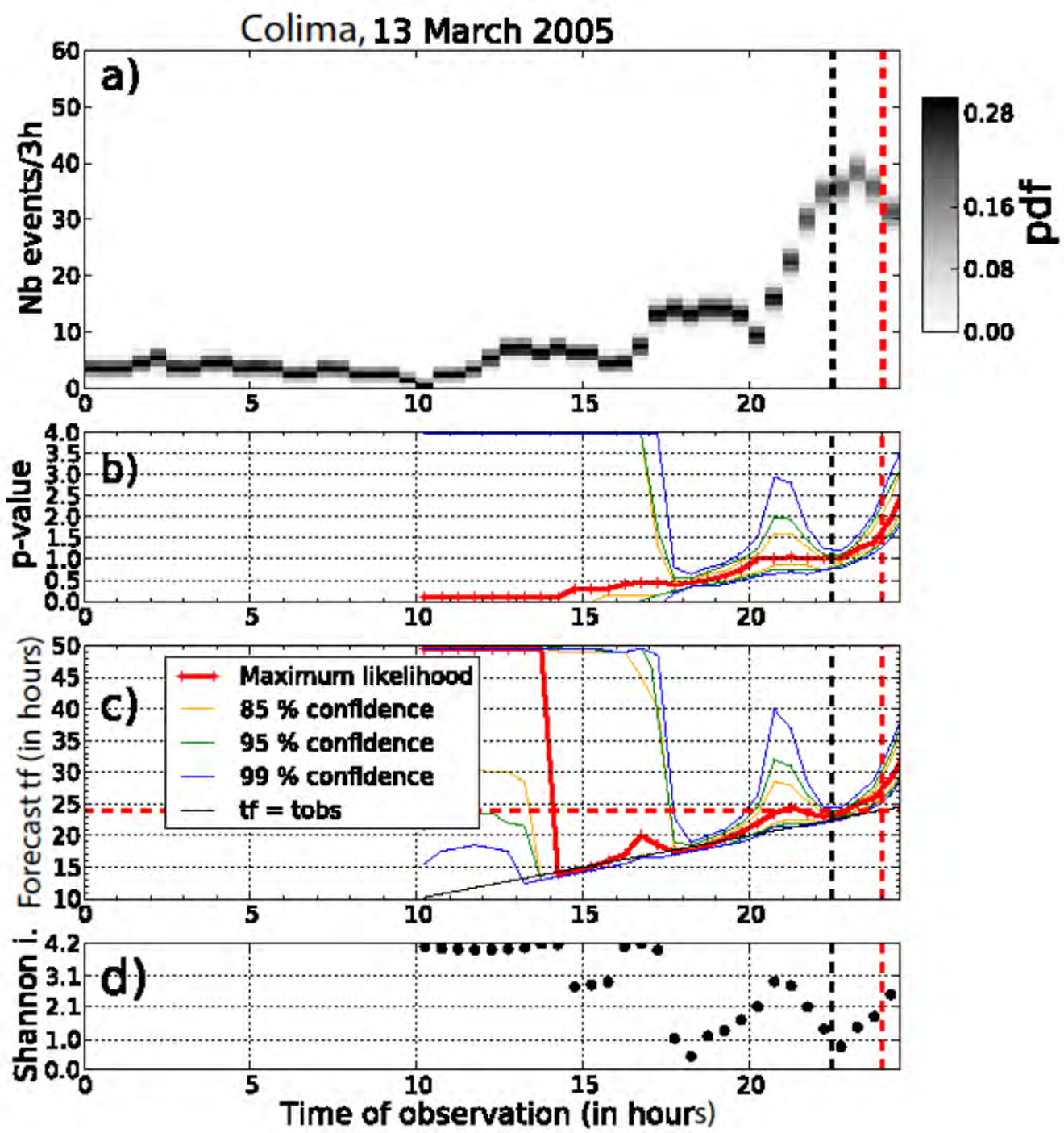


Figure 8

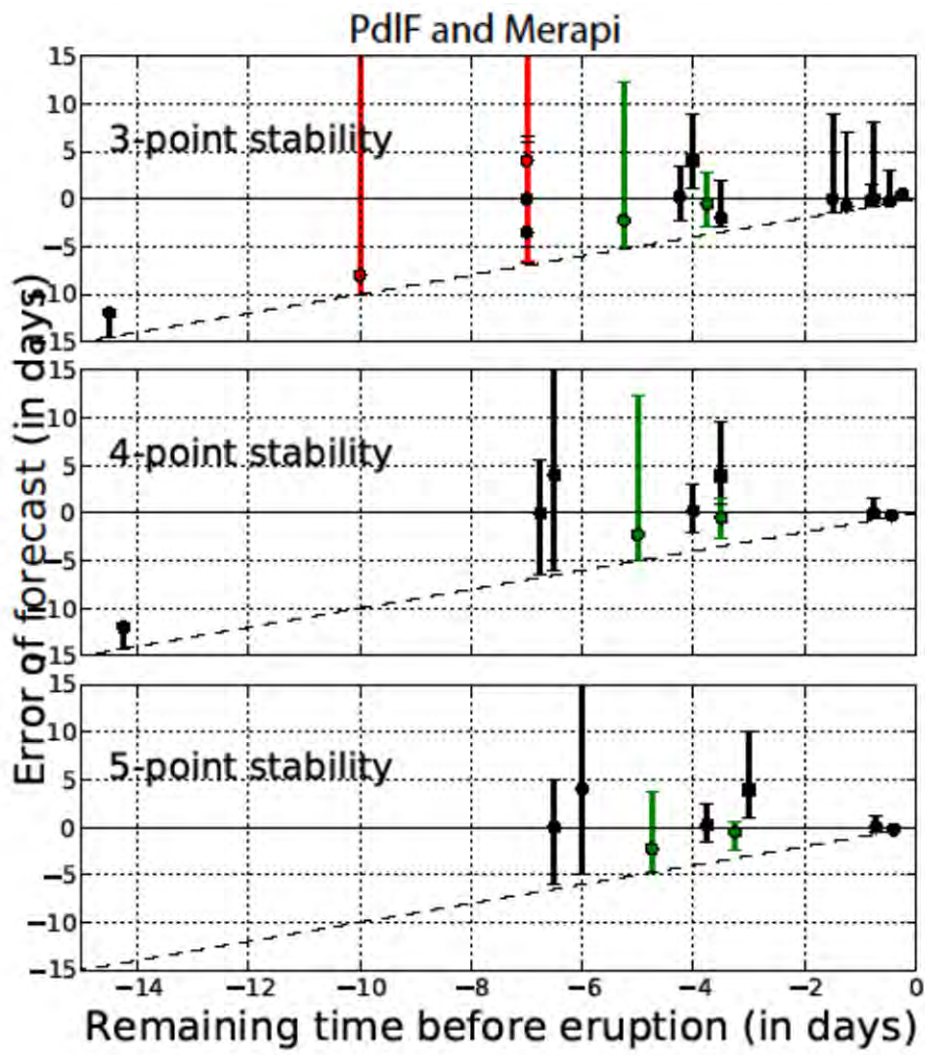


Figure 9a

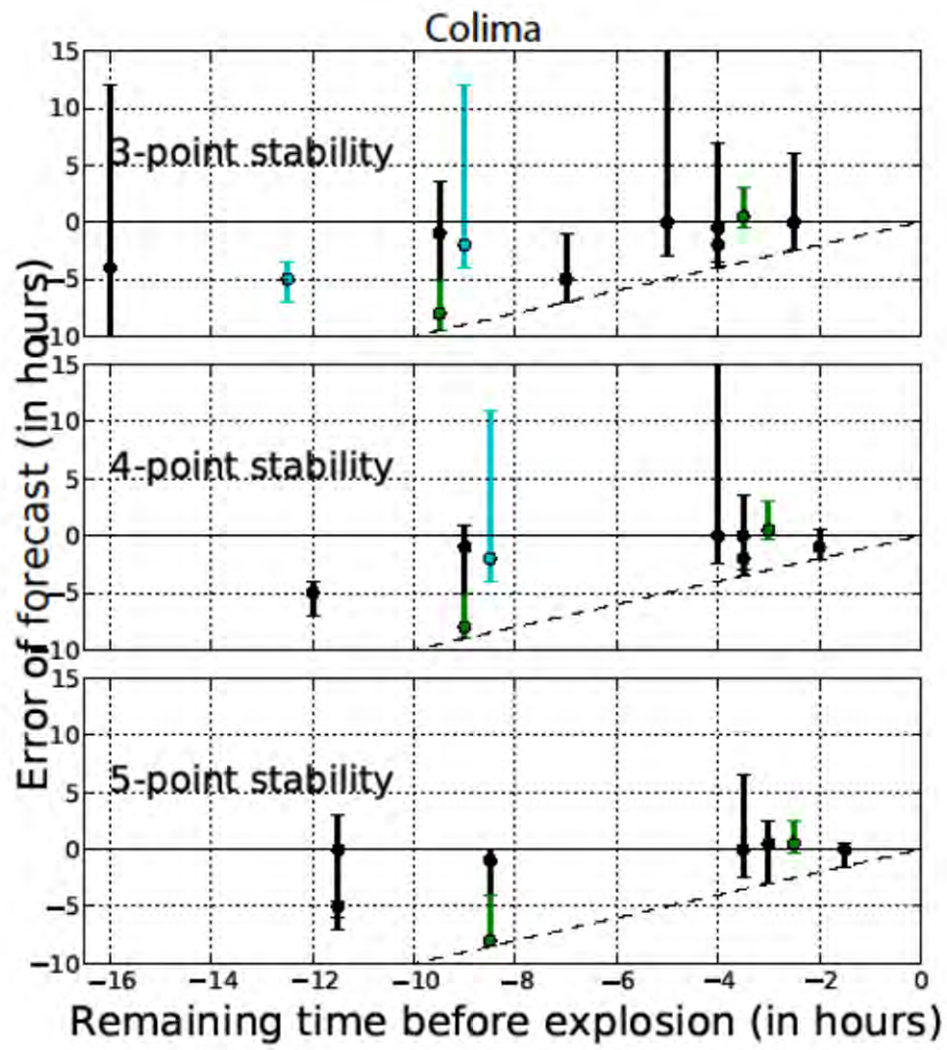


Figure 9b

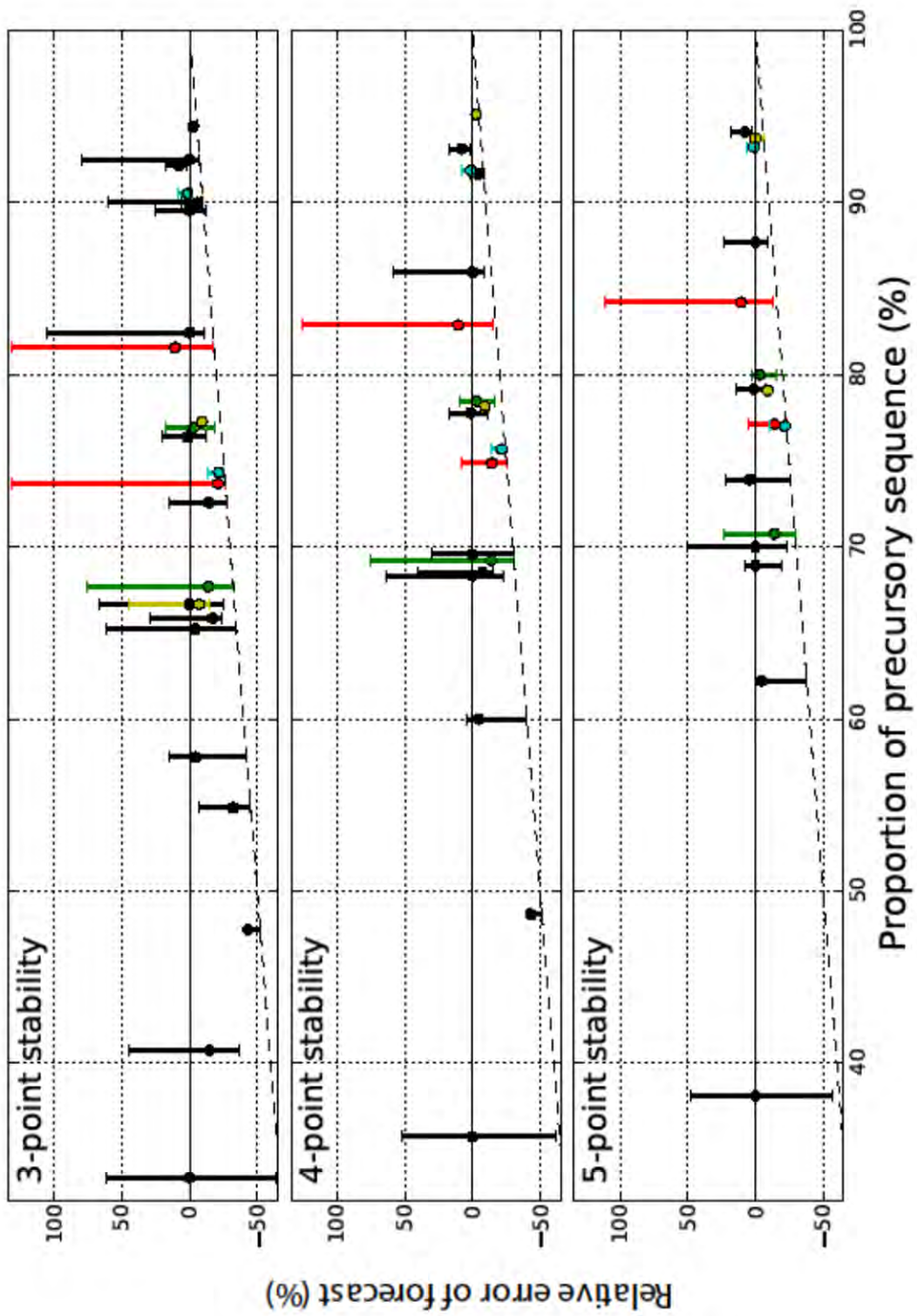


Figure 10

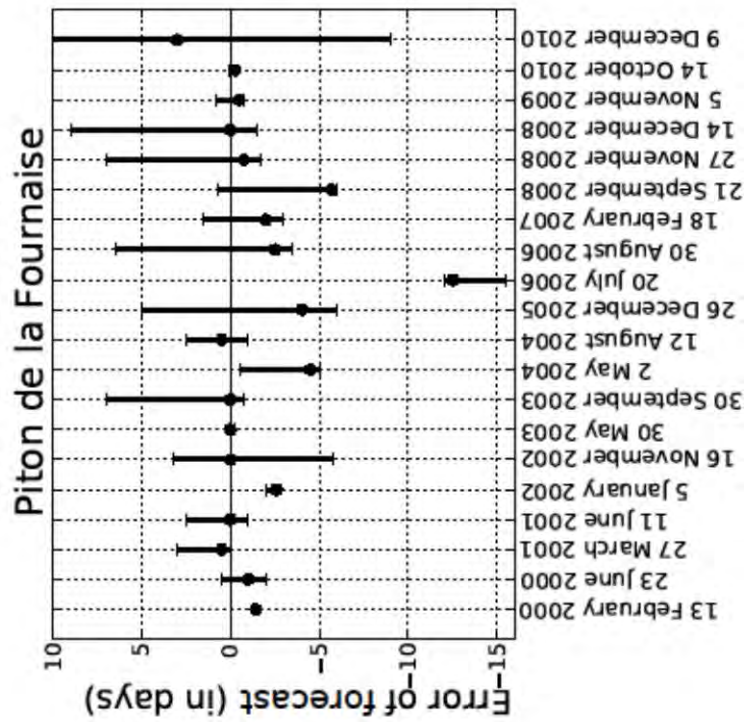
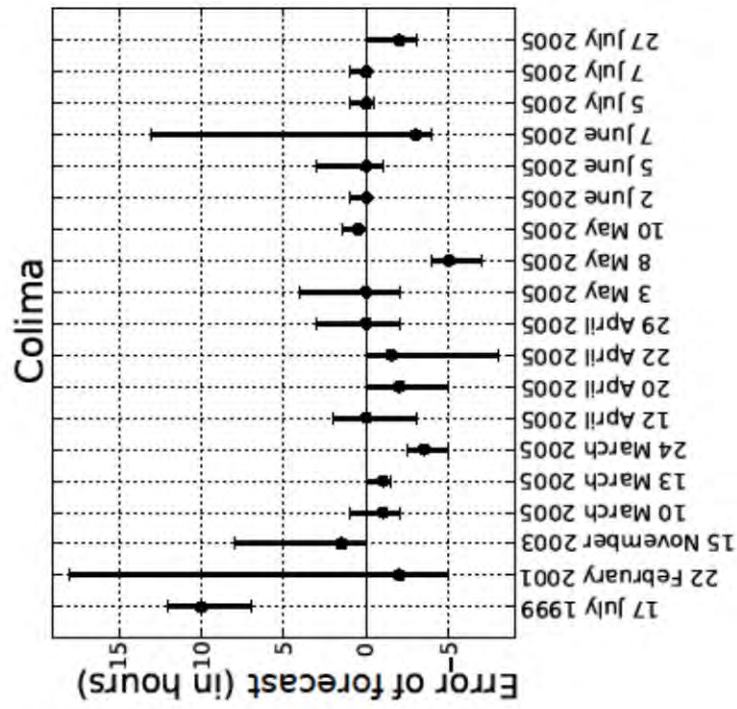


Figure 11

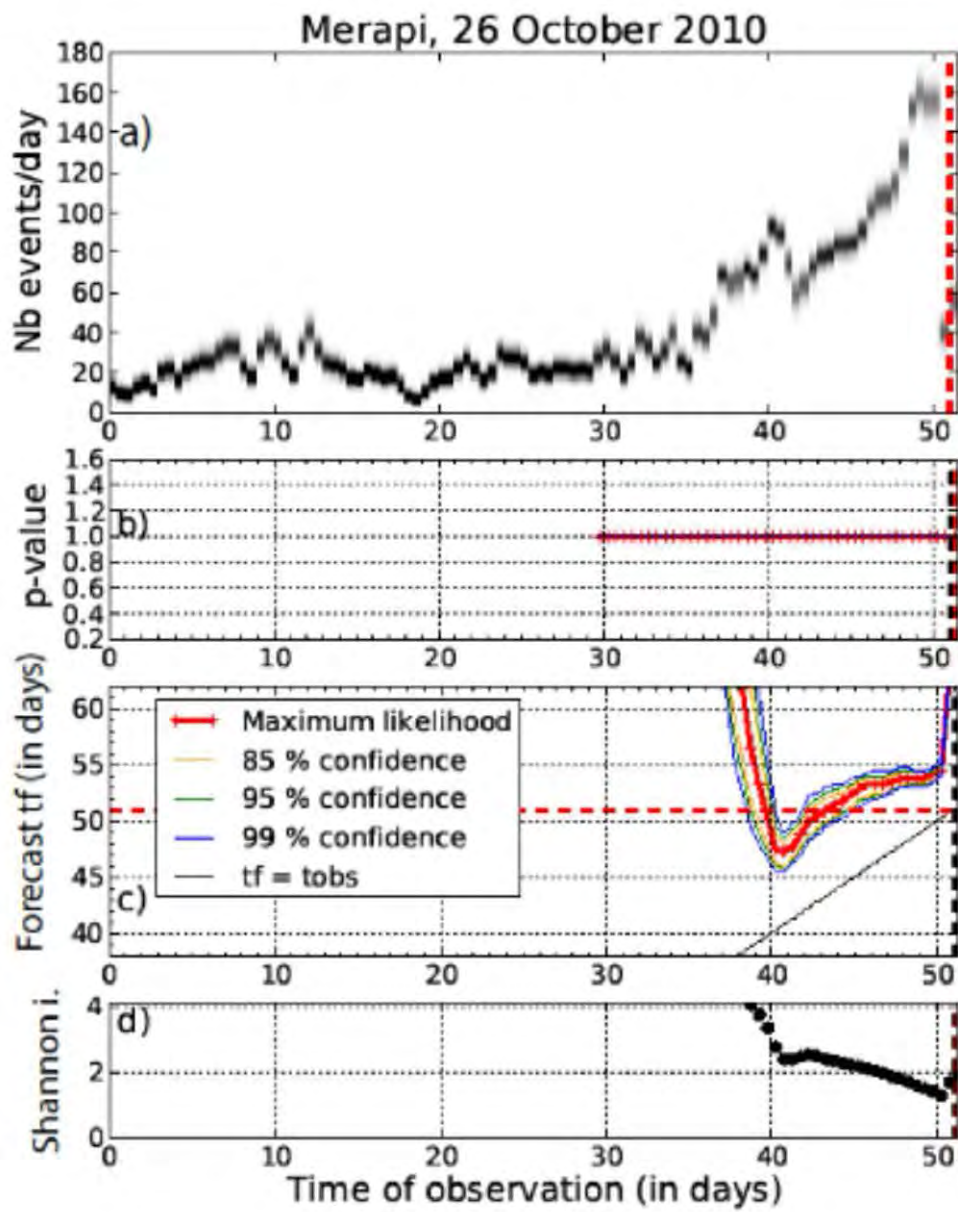


Figure 12

Date of eruption	Forecast	Pattern	Duration of precursors (in days)	Duration of quiescence (in days)	Hindsight prediction t_f-t_e (lower/ML/upper) (in days)	Estimated p-value (lower/ML/upper)	Δt	Real-time criteria Stab./Shannon
13/02/2000	*	Multiple	2	1.26	-1.53/-1.43/-1.33	1.4/2.5/3	3h	0 / No
23/06/2000	**	Single	8	1	-2/-1/0.5	1.75/2.1/2.5	12h	12 / Yes
12/10/2000		Short swarm	-	-	-	-	-	0 / No
27/03/2001	**	Single	14	0.5	-0.25/0.5/3	1/1.2/1.8	12h	3 / Yes
11/06/2001	*	Single	7	0	-1/0/2.5	0.8/1.5/3	24h	0 / No
05/01/2002	*	Multiple	4	2	-2.5/-2.6/-2	1.9/1.5/3	2h	0 / No
16/11/2002	**	Multiple	10.75	2.75	-5.75/0/3.25	1.5/3.9/4	12h	8 / Yes
30/05/2003	**	Single	2.4	0.2	-0.25/0/0.1	2.25/2.4/2.45	2h	16 / Yes
22/08/2003		Constant+swarm	-	-	-	-	-	0 / No
30/09/2003	**	Single	10	1	-0.75/0/7	0.4/0.5/2	12h	3 / Yes
07/12/2003		Constant+swarm	-	-	-	-	-	0 / No
08/01/2004		Increasing steps	-	-	-	-	-	0 / No
02/05/2004	*	Multiple	10.5	4	-5/-4.5/-0.5	0.9/1/3.5	12h	0 / No
12/08/2004	**	Single	18	2.5	-1/0.5/2.5	1.9/2.4/3.3	12h	6 / Yes
17/02/2005		Constant+swarm	-	-	-	-	-	0 / No
04/10/2005		Short swarm	-	-	-	-	-	0 / No
29/11/2005		Constant+swarm	-	-	-	-	-	0 / No
26/12/2005	**	Multiple	15	4	-6/-4/5	1.5/2.2/3.7	24h	3 / Yes
20/07/2006	**	Multiple	27	12.5	-15.5/-12.5/-12	0.7/0.9/1	12h	5 / Yes
30/08/2006	*	Multiple	16	2.5	-3.5/-2.5/6.5	0.5/0.9/3.8	24h	0 / No
18/02/2007	**	Multiple	35	2	-3/-2/1.5	0.5/0.6/0.75	24h	4 / Yes
30/03/2007		Increases	-	-	-	-	-	0 / No
21/09/2008	*	Multiple	7	6	-6/-5.7/-5	0.5/0.6/0.75	12h	0 / No
27/11/2008	**	Multiple	6	1	-1.75/-0.75/7	0.5/0.6/3.5	12h	3 / Yes
14/12/2008	**	Multiple	15	1.5	-1.5/0/9	0.6/1/2.4	24h	3 / Yes
5/11/2009	*	Multiple	3	0.75	-0.75/-0.5/0	0.4/0.6/0.8	12h	0 / No
14/12/2009		Short swarm	-	-	-	-	-	0 / No
2/01/2010		Short swarm	-	-	-	-	-	0 / No
14/10/2010	**	Multiple	1.75	0.25	-0.3/-0.25/0.1	1.5/1.6/2.5	2h	5 / Yes
09/12/2010	**	Multiple	35	5	-9/3/40	0.55/1/2	24h	6 / Yes

Table 1

Date of eruption	Forecast	Type of pattern	Duration of precursors (in hours)	Duration of quiescence (in hours)	Hindsight prediction (lower/ML/upper) (in hours)	Estimated p-value (lower/ML/upper)	Δt	Real-time criteria Stab./Shannon
10/02/1999		Constant	-	-	-	-	-	-
10/05/1999		Constant	-	-	-	-	-	-
17/07/1999	*	Single	13h	0	+7/+10/+12	2.5/3.9/4	1h	0 / No
22/02/2001	*	Increase	40h	6h	-5/-2/18	0.6/1.2/3.9	2h	0 / No
17/07/2003		Constant	-	-	-	-	-	-
15/11/2003	*	Single	8h	0	0/+1.5/+8	0.6/1.2/3.5	2h	0 / Yes
10/03/2005	*	Single	10h	4h	-2/-1/1	1.9/2.7/3.5	2h	0 / No
13/03/2005	**	Single	24h	1.5h	-1.5/-1/0	0.8/1/1.2	3h	6 / Yes
24/03/2005	*	Increase	24h	0	-5/-3.5/-2.5	2.5/3.85/4	3h	0 / No
26/03/2005		Constant	-	-	-	-	-	0 / No
12/04/2005	*	Single	10h	3h	-3/0/2	1.5/3.9/4	2h	0 / Yes
20/04/2005	**	Multiple	15h	0.5h	-5/-2/0	2/4/4	2h	7 / Yes
22/04/2005	**	Multiple	21h	1h	-8/-1.5/0	0.5/0.7/1.4	3h	8 / Yes
29/04/2005	**	Single	37h	12h	-2/0/3	2.5/3/3.5	2h	9 / Yes
01/05/2005		Multiple	-	-	-	-	-	-
03/05/2005	**	Multiple	11h	0	-2/0/4	1.3/3.4/3.9	2h	5 / Yes
04/05/2005		Constant	-	-	-	-	-	-
05/05/2005		Constant	-	-	-	-	-	-
08/05/2005	**	Single	55h	10h	-7/-5/-4	1.7/2/2.1	2h	5 / Yes
10/05/2005	**	Multiple	37h	0.5h	0/+0.5/+1.5	1.15/1.2/1.25	2h	8 / Yes
16/05/2005		Constant	20h	-	-	-	-	-
24/05/2005		Multiple	40h	-	-	-	-	-
25/05/2005		Multiple	30h	-	-	-	-	-
30/05/2005	*	Multiple	70h	25h	-24/-21/-17	1.1/1.25/1.7	2h	4 / Yes
02/06/2005	*	Increase	30h	0	0/0/+1	0.1/0.3/1	4h	0 / No
05/06/2005	**	Single	28h	0.5h	-1/0/3	0.5/0.6/0.9	2h	7 / Yes
07/06/2005	**	Single	26h	9h	-4/-3/13	1/1.8/4	2h	4 / Yes
10/06/2005		Increase	10h	-	-	-	-	-
05/07/2005	*	Multiple	50h	0.5h	-0.5/0/1	0.4/0.6/0.8	3h	0 / Yes
07/07/2005	*	Multiple	27h	0h	0/0/+1	0.45/0.45/0.5	3h	0 / No
27/07/2005	**	Multiple	72h	4h	-3/-2/0	0.35/0.4/0.45	4h	5 / Yes
16/09/2005		Multiple	72h	-	-	-	-	-
27/09/2005		Constant	48h	-	-	-	-	-
26/10/2010 Merapi	**	Multiple	51 days	0.5 days	1/4/8	0.9/1.1/1.4	24h	7 / Yes

Table 2