Seismic pattern recognition techniques to predict large eruptions at the Popocatépetl, Mexico, volcano

D.A. Novelo-Casanova *, C. Valdés-González

Instituto de Geofísica, Universidad Nacional Autónoma de México, México D.F., 04510, México

**Abstract**

Using pattern recognition techniques, we formulate a simple prediction rule for a retrospective prediction of the three last largest eruptions of the Popocatépetl, Mexico, volcano that occurred on 23 April–30 June 1997 (Eruption 1; VEI~2–3); 11 December 2000–23 January 2001 (Eruption 2; VEI~3–4) and 7 June–4 September 2002 (Eruption 3; explosive dome extrusion and destruction phase). Times of Increased Probability (TIP) were estimated from the seismicity recorded by the local seismic network from 1 January 1995 to 31 December 2005. A TIP is issued when a cluster of seismic events occurs under our algorithm considerations in a temporal window several days (or weeks) prior to large volcanic activity providing sufficient time to organize an effective alert strategy. The best predictions of the three analyzed eruptions were obtained when averaging seismicity rate over a 5-day window with a threshold value of 12 events and declaring an alarm for 45 days. A TIP was issued about six weeks before Eruption 1. TIPs were detected about one and four weeks before Eruptions 2 and 3, respectively. According to our objectives, in all cases, the observed TIPs would have allowed the development of an effective civil protection strategy. Although, under our model considerations the three eruptive events were successfully predicted, one false alarm was also issued by our algorithm. An analysis of the epicentral and depth distribution of the local seismicity used by our prediction rule reveals that successful TIPs were issued from microearthquakes that took place below and towards SE of the crater. On the contrary, the seismicity that issued the observed false alarm was concentrated below the summit of the volcano. We conclude that recording of precursory seismicity below and SE of the crater together with detection of TIPs as described here, could become an important tool to predict future large eruptions at Popocatépetl. Although our model worked well for events that occurred in the past, it is necessary to verify the real capability of the model for future eruptive events.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

Popocatépetl volcano (Popo) is located in central Mexico (19.023°N, 98.622°W), 5452 m above sea level (hereafter all elevations will be referenced to sea level), about 71 km south-east of downtown Mexico City and 40 km west of the city of Puebla (Fig. 1). The modern cone of Popo consists of interlayered andesitic to dacitic lava flows and pyroclastic deposits erupted in the last 23 ka (Robin, 1984). This volcanism followed a cataclysmic eruption that was accompanied by partial collapse of the edifice and generation of voluminous debris avalanche (Siebe and Macías, 2004). Over the last 23 ka, Popo has generated at least seven Plinian eruptions that produced extensive pumice-fall and ash flow deposits. The last three Plinian eruptions took place 5000, 2100, and 1100 years BP when humans had already settled in the area as testified by archaeological remains buried by ash beds and pottery shards incorporated in ash-flows and lahar deposits (Siebe et al., 1996).

The morphology of the elliptical crater, E–W oriented (820×650 m) was mainly inherited from the last major Plinian eruption that left an approximately 300 m-deep crater with steep inner walls. The lowest point of the crater is at an altitude of 5180 m on the eastern rim, whereas the summit is on the opposite western rim (Macías and Siebe, 2005).

Historical activity of Popo is reported since 3195–2830 B.C. (Siebe et al., 1997). On 24 December 1994, after approximately 67 years of quiescence, the volcano reinitiated its activity with an emission of vigorous ash plumes. On 29 March 1996, the activity peaked with a strong0377-0273/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.jvolgeores.2008.05.005
located in the states of Puebla and Morelos. Since its reactivation, Popo has shown a significant increase in seismic, fumarolic and ash emission activity as well as minor and medium eruptions. Explosive phases have formed and destroyed domes at the bottom of the crater (Arciniega-Ceballos et al., 1999). Most of the recorded tectonic earthquakes associated to the volcanic activity are located between the depth range of 2 and ~2 km (Valdés-González and González-Pomposo, 1999).

Forecast of volcanic hazards are essential for risk assessment. Volcanic hazards are generally assessed at short terms (days and months) when monitoring active volcanoes; or at mid-term (months to years) to long term (years to hundreds of thousands years) when dealing with sitting issues, for example, for industrial facilities (Jacket and Carniell, 2006). The Materials Failure Method has been used to track precursory volcanic activity and to forecast forthcoming activity at short terms by fitting data according to an empirical rate-acceleration relation (Voight, 1988, 1989). This method uses numerical or graphical rate extrapolation toward an expected failure rate to forecast time windows within which eruptions are expected (Cornelius and Voight, 1994, 1995). Tárraga et al. (2006) analyzed the low-frequency seismic noise at Teide-Pico Viejo complex, Canary Islands, Spain, to investigate the feasibility of using it to forecast, via the materials failure forecast method, the time of occurrence of discrete events on the basis of their relationship to volcanic activity. To avoid subjectivity in the forecast procedure, Tárraga et al. (2006) also developed an automatic program to generate “forecast”, validated by Bayes theorem. Because uncertainty is mainly related to imperfect knowledge of nonlinear volcanic processes and to a limited amount of monitoring information, the forecasting of volcanic eruptions needs to be based on a probabilistic formalism (Sparks, 2003). Wickman (1966) and Reyment (1969) proposed a stochastic model for the study of volcanic eruption patterns on specific volcanoes. More recently, for a short term forecasting, a probabilistic approach was proposed by Connor et al. (2003). It consists of applying logistic survivor functions to suggest a physical basis with competing effects for using a specific probabilistic model. This model was applied to Vulcanian explosions that occurred at the Soufriere Hills volcano, Montserrat. Another probabilistic

Fig. 1. a) Epicenters and b) W–E projection of hypocenter distribution of the January 1995–December 2005 micro-seismicity at Popo volcano. The triangles indicate station locations of the Popo’s seismic network.
method is the event tree approach, which was applied for the estimation of volcanic hazards at Vesuvius (Marzocchi et al., 2004). Jacket and Carniell (2006) summarized some theoretical aspects of geostatistical methods and models for the analysis and estimation of volcanic hazards at short and long terms and presented case studies carried out at the Stromboli volcano in Italy, the Soufrière, Hills volcano of Montserrat and the Osteifel volcanic region in Germany.

Collombet et al. (2003) using a superposed epoch analysis for all of the 15 eruptions in the 1988–2001 at Piton de la Fournaise (PdlF) volcano, observed a power law accelerating seismicity rate two weeks before the eruption time. The accelerating pattern is also recovered by stacking seismicity before the peaks of the background seismicity. Collombet et al. (2003) proposed that the average acceleration of the PdlF volcano seismicity rate before the eruptions, the complexity of each single pre-eruptive seismicity sequence, and the accelerations within the background seismicity result from two interacting processes: the fluid rock interactions that are driven by volcanic dynamics and the earthquake interactions that are driven by the generic statistics laws of earthquakes. McGuire and Kilburn (1997) and Kilburn and Voight (1998) identified slow rock cracking as an important precursor to major volcanic eruptions that triggers seismic events and the increase of seismic energy acceleration before a volcanic eruption.

Pattern recognition techniques in geophysics have been predominantly used in the time domain to predict the occurrence of large future earthquakes applying the CN and M8 codes (Keilis-Borok et al., 1988; Keilis-Borok and Kossobokov, 1990; Kossobokov and Shebalin, 2003). Marzocchi et al. (2003), however, providing insights to better formally define the earthquake-forecasting problem, both in setting up and in testing the validity of the forecasting model, found that the forecasting capability of the M8 and CN algorithms is very likely overestimated.

Novelo-Casanova and Alvarez-Moctezuma (1995) used pattern recognition techniques and analyzed historical seismicity and conditional probabilities for the recurrence of large earthquakes applying the CN and M8 codes (Keilis-Borok et al., 1988; Keilis-Borok and Kossobokov, 1990; Kossobokov and Shebalin, 2003). Marzocchi et al. (2003), however, providing insights to better formally define the earthquake-forecasting problem, both in setting up and in testing the validity of the forecasting model, found that the forecasting capability of the M8 and CN algorithms is very likely overestimated.

Novelo-Casanova and Alvarez-Moctezuma (1995) used pattern recognition techniques and analyzed historical seismicity and conditional probabilities for the recurrence of large earthquakes applying the CN and M8 codes (Keilis-Borok et al., 1988; Keilis-Borok and Kossobokov, 1990; Kossobokov and Shebalin, 2003). Marzocchi et al. (2003), however, providing insights to better formally define the earthquake-forecasting problem, both in setting up and in testing the validity of the forecasting model, found that the forecasting capability of the M8 and CN algorithms is very likely overestimated.

Pattern recognition techniques in geophysics have been predominantly used in the time domain to predict the occurrence of large future earthquakes applying the CN and M8 codes (Keilis-Borok et al., 1988; Keilis-Borok and Kossobokov, 1990; Kossobokov and Shebalin, 2003). Marzocchi et al. (2003), however, providing insights to better formally define the earthquake-forecasting problem, both in setting up and in testing the validity of the forecasting model, found that the forecasting capability of the M8 and CN algorithms is very likely overestimated.

Novelo-Casanova and Alvarez-Moctezuma (1995) used pattern recognition techniques and analyzed historical seismicity and conditional probabilities for the recurrence of large earthquakes applying the CN and M8 codes (Keilis-Borok et al., 1988; Keilis-Borok and Kossobokov, 1990; Kossobokov and Shebalin, 2003). Marzocchi et al. (2003), however, providing insights to better formally define the earthquake-forecasting problem, both in setting up and in testing the validity of the forecasting model, found that the forecasting capability of the M8 and CN algorithms is very likely overestimated.

2. Data and procedures

Here, we use the local seismicity recorded from 1 January 1995 to 31 December 2005 by the Popo’s seismic network (Fig. 1). This network is operated jointly by the Centro Nacional de Prevención de Desastres (CENAPRED) and the Instituto de Geofísica (IGF) of the Universidad Nacional Autónoma de México (UNAM). Details about seismic instrument characteristics are given in Quass et al. (1995). The distribution of the small aperture network, allowed us to detect all the local volcano-tectonic (VT) events above magnitude 1.0 located in the surroundings of the volcano edifice, thus ensuring a complete and uniform seismic catalogue for this study.

During the studied period, a total of 2500 events were located in the vicinity of the volcano with about 96% of them concentrated at depths between 2 and –2 km below the crater (Fig. 1). Magnitudes of microearthquakes ranged from 1.2 to 3.8 and were determined from coda wave duration (Lee et al., 1997). Only 29 of these events had magnitudes larger or equal to 3.0. The quality of the locations and the seismic velocity model are discussed in Valdés-González et al. (1997).

Some characteristics of the recorded seismograms are presented in Novelo-Casanova et al. (2006).

TIPs were calculated for a retrospective prediction of the three largest volcanic eruptions that occurred during the studied period: 1) 23 April–30 June 1997; 11 December 2000–23 January 2001; and 7 June–4 September 2002. The characteristics of these eruptions are described in the following section.

3. Main characteristics of Popocatepetl’s largest volcanic eruptions

3.1. The 23 April–30 June 1997 eruption

On 23 April 1997, background seismicity at Popo increased and on the next day, a volcanic explosion reached about 4 km above its summit. On April 29, an eruption sent incandescent fragments on the northwest flank (Meli et al., 1997a). Large ash emissions occurred on May 11, 14, 15, 24 and 27, which also produced incandescent fragments on the vicinity of the crater, accompanied by high-frequency (HF) tremors on May 24. On June 11, ash emissions covered an area of about 1500 km² reaching distances between 6 and 8 km from the crater. The plume that occurred on June 14 was visible from Mexico City and Puebla at a distance of about 60 and 40 km, respectively (Meli et al., 1997b).

The June 30 volcanic eruption had a VEI between 2 and 3 and was the largest eruption registered at Popo since its activity was reinitiated on December 1994. This eruptive event generated the formation of domes, fast lava emission and fragmentation that originated a large ash cloud and the collapse of lava material leaving a funnel type structure (Meli et al., 1997b).

3.2. The 11 December 2000–23 January 2001 eruption

This eruptive period was a rapid dome growth and intense eruptive phase. On 12–15 December 2000, ash-bearing exhalations, incandescent fragments originating in the crater and ejection of hot debris were observed. Episodes of high amplitude, low-frequency (LF) harmonic tremor lasting from a few minutes to several hours were also detected. Up to present, the seismic energy released by the LF tremor during this four-day period has exceeded the energy accumulated by LF tremors during any other year for which record exist. From aerial photographs, the total volume of fresh lava accumulated within the crater of Popo in December 18, is estimated to be between 15 and 19 million cubic meters, exceeding the combined lava volume released during the formation of previous domes.

On January 22, a very large explosive event took place with an ash column of more than 10 km and pyroclastic flows descending from the N, SE, SW and NW flanks of the volcano. Ash fall out was reported toward SE up to a 200 km distance. The estimated VEI for this volcanic event was 3–4, and it is considered the largest eruptive period that has taken place at Popo since its reactivation on December 1994 (Valdés-González et al., 2000).

3.3. The 7 June–4 September 2002 eruption

This eruptive phase was characterized by intense formation and destruction of lava domes. Episodes of LP activity were registered on July 23 and August 15–17, as well as 20 h of harmonic tremor on August 10. The VT activity during this stage showed a narrow cluster of
earthquakes between 2 and ~2 km below the crater. A total of 48 VT microearthquakes were recorded on July.

A summary of the main characteristics of the three eruptions are shown in Table 1.

### 4. Prediction scheme

The capability to predict with at least a few days prior to the onset of a large eruption may have a high practical relevance for civil protection measures. Here, by a quantification of the daily seismicity, we attempt to identify precursory patterns to the three largest eruptions that occurred at Popo during the 1995–2005 period. A TIP is issued when a cluster of seismic events occurs under our algorithm conditions in a temporal window several days (or weeks) prior to large volcanic activity providing sufficient time to organize an effective alert strategy. Our general approach is reduced to the following three steps of data analysis (Grasso and Zaliapin, 2004). First, we define a cluster of seismic events by considering the sequence of VT event occurrence times \( C = \{ t_e; e = 1, 2, \ldots, E; t_e \leq t_{e+1} \} \) at Popo volcano for the studied period. Note that we use neither magnitude nor location of events. Second, on the sequence \( C \) we define a function \( N(t,s) \) as the number of earthquakes in the time window \([t-s, t]\), \(s\) being a numerical parameter. This function is calculated for the time interval considered with different values of the numerical parameter \( s \). Third, a TIP is triggered when the function \( N(t,s) \) exceeds a predefined threshold \( N_o \). The threshold \( N_o \) is usually chosen as a certain percentile of the distribution function \( N(t,s) \). We consider that when a TIP is declared active for a time interval \( \Delta \), the alarm is terminated until an eruption occurs during this time \( \Delta \) or this period expires, whichever comes first. Thus, the prediction scheme depends on three parameters: time window \( s \), threshold \( N_o \), and duration \( \Delta \) of alarms (Kossobokov and Shebalin, 2003).

Because of the low average daily seismicity rate at Popo (~two events per day), we varied each adjustable parameter of the algorithm independently of the others, within a broad range as follows: \( 1 \leq s \leq 15 \) days, \( 1 \leq N_o \leq 15 \) events per \( s \) days, \( 1 \leq \Delta \leq 60 \) days. Raising the threshold \( N_o \) will reduce the number of alarms \( A \) but may increase the number of failures to predict. Raising \( \Delta \), on the other hand, will increase the duration alarms \( D \) but may reduce the number of failures to predict \( N_f \).

### 5. Results and discussion

Following our prediction scheme, we found that the best predictions of the three last largest eruptions at Popo are obtained when averaging seismicity rate over a 5-day window \((s=5)\) with a threshold value of 12 events \((N_o=12)\) and declaring an alarm for 45 days \((\Delta=45)\) (Fig. 2). In other words, a TIP or seismic cluster

![Fig. 2. Successful (solid squares) and false alarm (open square) TIPs detected from VT events recorded at Popo from 1 January 1995 to 31 December 2005. The length of the squares represents a time interval of 45 days (see text). The vertical dotted lines indicate the duration of the eruptive periods. On 14 September 1996, a false alarm was issued (see text).](image-url)
precurory to a Popo’s eruption is defined by a group of 12 or more events which occur in the Popo region within a 5-day period 45 days or less before the eruption. We consider that forecasting an eruption at Popo within a 45-day period would be important for establishing an adequate civil protection strategy. The choice of safety measures depends on two factors: (1) expected damage and (2) reliability of the prediction probabilities of false alarm and failure to predict and duration of the alarm (Zaliapin and Keilis-Borok, 2002). Here, we concentrate our work in point (2).

Under the best parameter conditions of our prediction scheme, the three studied large eruptions are preceded by TIPs (Fig. 2). For the first analyzed eruptive period, a TIP is issued about six weeks before the eruption (Fig. 3). TIPs are also detected about one and four weeks before the second and third eruptions, respectively (Fig. 3). Notice that in some cases TIPs overlap extending the period within which an eruption is forecasted. TIPs were also issued within the periods of occurrence of Eruptions 2 and 3 and by the end of Eruption 3 (Fig. 3). We disregard these TIPs as false alarms because they were clearly issued by the increase of the volcano volcano-tectonic seismicity generated by these eruptions. Thus, we consider that the three analyzed events were successfully predicted by our algorithm.

Our prediction algorithm issued one false alarm on 14 September 1996, because no large volcanic activity followed 45 days after this TIP (Fig. 2). This false alarm could be associated to the September 14–21, 1996 volcanic activity at Popo. During this period several local VT events were recorded by the Popo’s seismic network. The largest number of VT per day was 8 and occurred on September 20, 1996. This increase of seismicity may have been the result of magma material pushing the walls of the conduit with excess pressure because of the presence of a dome previously formed on the crater. Although no large eruption was reported, we consider that this important VT activity was detected by our algorithm triggering the TIP (Fig. 2).

To understand better why three TIPs were successful predictions and one was a false alarm we proceeded to analyze the epicentral and depth distribution of the events used by our prediction algorithm. The three successful TIPs were issued from microearthquakes that took place below and towards SE of the crater (Fig. 4a–c). On the contrary, the seismicity that issued the false alarm occurred only below the summit of the volcano (Fig. 4d). The number of VT events towards SE was 14, 3, and 51 during the first, second, and third studied large eruptions, respectively (Fig. 4a–c). No VT seismicity towards SE was recorded during the 14–21 September 1996, volcano activity (Fig. 4d). Thus, we conclude that recording of seismicity below and SE of Popo’s crater at the same time period together with detection of TIPs using our prediction scheme may provide indications of a future large eruption at this volcano within a 45-day period following the time the TIP is issued.

To detect possible correlation between peaks of seismicity rate with the occurrence of the three large eruptions, we analyzed the daily seismicity and the cumulative number of HF and LF events registered during our studied period (Fig. 5). We observe large variations of the daily seismicity and changes of the slopes of both, the cumulative number of analyzed HF and LF earthquakes. However, it is difficult to identify systematic changes associated with any of the three eruptions. These results confirm that our prediction scheme is a more effective tool to identify long-term seismicity pattern than analysis of seismicity rates. Our method, however, has been tested using past data, it is important to apply our model for future eruptive events at Popo to verify its reliability.

Arámbula-Mendoza (2007) using focal mechanism solutions of VT events located SE of Popo’s crater demonstrated that most of the microearthquakes that occur on this side of the volcano are associated to transcurrent type faulting and inferred a volcanic fault in this site oriented SE–NW. Arámbula-Mendoza (2007) concluded that this fault may act as a damping mechanism due to the compression of magmatic intrusions associated to large eruptions at this volcano. Our results also indicate that seismic activation of this volcanic fault may play an important role for prediction of imminent large eruptions at Popo.

Fig. 3. Detailed representation of successful TIPs (solid horizontal bars) issued for the three analyzed eruptions at Popo (Fig. 2).
According to our objectives, the observed successful TIPs will allow, in advance, effective civil protection actions. Marzocchi and Woo (2007), however, pointed out the importance of developing a strategy to integrate a probabilistic scheme for eruption forecasting and cost-benefit analysis. The approach of these authors incorporates: 1) A decision-analysis framework, expressed in terms of event probability, accounting for all modes of available hazard knowledge; and 2) Quantitative and transparent rules that can be tested. Since the quantitative rules are defined during a period of quiescence, it allows prior scrutiny of any scientific input into the model. We believe that detection of TIPS at Popo, as described here and used as a relevant scientific tool into a strategy as the one proposed by Marzocchi and Woo (2007), can be used by civil protection officials taking the "optimal" decision.

Although our results allowed us to establish some quantitative rules through a retrospective analysis of three eruptions at Popo, our observations may be over-fitted because our data (three volcanic events) are few compared to the number of parameters that we are considering (number of events, magnitudes, length of time window, earthquake location, seismic rate for unit time, etc.). In other words, our results, based on past data may not be a necessary condition that guarantees an "excellent" forecast of future eruptions at Popo. In cases where the data are enough, a good practice is to split the dataset in learning and testing datasets, where the learning dataset is used to set

---

**Fig. 4.** Epicentral and depth distribution of the micro-seismicity used by the prediction algorithm to issue the successful (a–c) and the false alarm (d) TIPS. Observe that for the successful TIPS, seismicity occurred below (Box 1) as well as towards SE (Box 2) of the crater. On the contrary, for the false alarm case, seismicity was constrained below the summit of the volcano. *N* indicates the number of VT events in each box.
6. Conclusions

A TIP was issued about six weeks before the 23 April–30 June 1997, eruption. TIPs were also detected about one and four weeks before the 11 December 2000–23 January 2001 and 7 June–4 September 2002, eruptive periods, respectively. The observed TIPs would have allowed the development of an effective civil protection strategy. The epicentral and depth distribution of the local seismicity used by our prediction rule reveal that successful TIPs were issued from microearthquakes that took place below and towards SE of Popo’s crater. The seismicity that issued the observed false alarm was constrained below the summit of the volcano. Recording of precursory seismicity below and SE of the crater together with detection of TIPs with a real time monitoring system may be an important tool to predict future large eruptions at Popo. Although our method worked well for eruptive events that took place in the past, it is necessary to verify the reliability of the model for future eruptive events.

Acknowledgments

The authors are grateful to two anonymous reviewers for their detailed review of this work and providing helpful comments. We also thank G. Suárez for helpful criticism and discussions. Our appreciation to the people of the volcano instrumentation group at the Centro Nacional de Prevención de Desastres (CENAPRED) as well as the team of the Instituto de Geofísica (IGF) of the Universidad Nacional Autónoma de México (UNAM) that operates the seismic network at Popocatépetl volcano.

References


