Nonlinear Dynamics (2006) 44: 341–349
DOI: 10.1007/s11071-006-2018-1

Earthquake Forecasting Using Neural Networks: Results and Future Work

E. IVO ALVES
Centro de Geofísica da Universidade de Coimbra, Av. Dias da Silva, 3000-134 Coimbra, Portugal; (e-mail: livo@ci.uc.pt; fax: +351-239-793-428)

(Received: 10 September 2004; accepted: 20 April 2005)

Abstract. In 1994, a new earthquake forecasting method was developed, that integrated in a neural network several forecasting tools that had been originally developed for financial analysis. This method was tested with the seismicity of the Azores, predicting the July, 1998, and the January, 2004, earthquakes, albeit within very wide time and location windows. Work is beginning to integrate physical precursors in the neural network, in order to narrow the forecasting windows.

Key words: Azores, Earthquake prediction, financial technical analysis, forecasting, neural networks, Portugal, seismicity

Abbreviations: ANN – Artificial neural network; HL – Hidden (middle) layer; IL – Input layer; MA – Moving average; MACD – Moving Averages' Convergence-Divergence; MMI – Modified Mercalli intensity; MOM – Momentum; ODI – Optimised decision index; OL – Output layer; RM – Real-modulated Index; RSI – Relative strength index; SO – Stochastic oscillator

1. Introduction

When looking for analytical tools designed to forecast time series, one finds that the practical domain in which they are most commonly used is financial forecasting – especially for the prediction of trends in the stocks, bonds, commodities and currencies markets [1] – where technical analysts name these tools "financial oscillators".

Time-series of earthquake parameters are also of a chaotic nature [2], with comparable scaling laws [3, 4], and testing those financial oscillators on seismicity was a natural step.

The transposition of scenery appears to be fairly straightforward. Instead of analysing a series of quotes in order to predict the next quote, one has to analyse a series of seismic parameters, namely those related to time, energy and epicentral locations, in order to predict the next time, effect and local.

A natural doubt can arise: if a predictive model must be related to the process it tries to predict, and there is no relationship between the fluctuation mechanisms of markets and seismic mechanisms, why and how should this approach work? Indeed, there is an important relationship because both are deterministic non-periodic processes – chaotic processes.

But there is still a more important reason for using these "financial" tools: they reduce the degree of chaos in the analysed sequence, making it more predictable. This happens because the fractal dimension of a system is a measure of its degree of chaoticity. The movement of a particle in the real plane can take dimensions between 1 (completely deterministic) and 2, in the case of Brownian movement (completely random). We expect that a sequence of, say, earthquake magnitudes, will have a fractal dimension between 1 and 2. If we calculate that dimension and then the dimension of the sequence of moving averages (the simplest oscillator, on which most others are based) we find such a reduction of fractal dimension – a reduction of chaos [5].

The evaluation of the effects of an earthquake can be made resorting to seismic magnitudes or intensities. The first one has the natural advantage of being quantitative and proportional to the liberated energy but has a shortcoming: the available data-series only begin in the twentieth century. If one is to forecast damaging earthquakes, one needs to study very long time series – historical seismicity – hence resort to semi-quantitative intensity data.

Also, in the specific case of the Azores, where this method was tested, although there are instrumental records since 1902, these were obtained by a Milne seismograph located in São Miguel Island, the only one until 1950 [6], and magnitude estimations based on this equipment's records are not reliable.

This, and the need to predict those earthquakes that are effectively damaging, led to work with Modified Mercalli Intensities [7] and to try and find a method that would deal with semi-quantitative data.

2. The Oscillators

Since these are intended to be periodic functions, one must first choose a time series, x_i, $i = 1, 2, \ldots, n$ (events), and a period P.

The following formulas intend to represent the developed computational algorithms, rather than being a rigorous mathematical description of each function. The chosen oscillators were.

2.1. Moving Averages (MA)

$$\mathrm{MA}(P)_i = (x_i + x_{i-1} + \ldots + x_{i-P} + 1)/P \tag{1}$$

Notice that, unlike what is the common practice, the attribution point for the MA is not the mid-point for the considered period but rather the last one, which was found to yield the best results [8].

2.2. MA Convergence-Divergence (MACD)

Let there be two MAs of different periods, P and Q, such that $P > Q$. Then,

$$\mathrm{MACD}(P,\ Q)i = \mathrm{MA}(P)_i - \mathrm{MA}(Q)_i \tag{2}$$

2.3. Relative Strength Index (RSI)

Let u be the average of positive variations in the considered period P and let d be that of negative variations. Then,

$$\mathrm{RSI}(P) = 100 - (100/(1 + (u/d))) \tag{3}$$

2.4. Real-Modulated Index (RM)

$$\mathrm{RM}(P)_i = x_i/\mathrm{MA}(P)_i \tag{4}$$

2.5. Optimised Decision Index (ODI)

$$\mathrm{ODI}(P)_i = [\mathrm{RM}(P)_i + \mathrm{RM}(P)_{i-1} + \cdots + \mathrm{RM}(P)_{i-P+1}]/P \tag{5}$$

which is a moving average of the RMs.

2.6. Stochastic Oscillator (SO)

Let m be the minimum and M the maximum of x_i in the considered period P. Then,

$$\mathrm{SO}(P)_i = (x_i - m)/(M - m) \tag{6}$$

2.7. Momentum (MOM)

$$\mathrm{MOM}(P)_i = x_i - x_{i-P+1} \tag{7}$$

2.8. Pattern Matching

This was another very simple analytical tool that was added to the six "financial" oscillators to calculate the minimal algebraic difference pattern-matching with the last sequence of 20 values.

Let there be a sequence x_i; $i = 1, 2, \ldots, n$; we intend to find the most probable value for x_{n+1}.

Take the vector of the last 20 values,

$$[x_n,\ x_{n-1},\ x_{n-2}, \ldots, x_{n-19}]$$

and the immediately previous one,

$$[x_{n-1},\ x_{n-2}, x_{n-3}, \ldots,\ x_{n-20}]$$

and add their absolute-term-wise differences:

$$S_1 = |x_n - x_{n-1}| + |x_{n-1} - x_{n-2}| + \cdots + |x_{n-19} - x_{n-20}| \tag{8a}$$

Iterate the process,

$$S_2 = |x_n - x_{n-2}| + |x_{n-1} - x_{n-3}| + \cdots + |x_{n-19} - x_{n-21}| \tag{8b}$$

$$\ldots$$

$$S_{n-20} = |x_n - x_{20}| + |x_{n-1} - x_{19}| + \cdots + |x_{n-19} - x_1| \tag{8c}$$

The most similar vector will be the one for which that sum is minimal, for instance

$$S_m = |x_n - x_{n-m}| + |x_{n-1} - x_{n-m-1}| + \cdots + |x_{n-19} - x_{n-m-19}| \tag{8d}$$

and so the next most likely value will be

$$x_{n+1} = x_{n-m+1} \tag{9}$$

As was shown elsewhere [8] there are periods *P* which yield optimal results in the analysis of seismic sequences. These are of four occurrences for RSI and RM, seven for SO, 21 for ODI and 28 for MOM.

Since the computational algorithm for ODI21 implies computing RM21, and the computation of RM21 that of MA21, we have to compute in all nine oscillators, namely, MA4, MA21, MACD21-4, RSI4, RM4, RM21, ODI21, SO7 and MOM28. Of these, only RSI4, RM4, it MACD21-4, it SO7, ODI21 and MOM28 will be used as input to the neural net. Tests showed that the extended training time would not be compensated by a better accuracy if one was to use all the computed oscillators [8].

The first tests of these tools for seismic forecasting [5], though encouraging, had two shortcomings: first, the outputs of the oscillators are qualitative, since they only indicate if a trend is rising, declining or stable ("buy", "sell" or "hold" in financial applications); then, when we apply several oscillators to the same sequence, the results are not always consistent.

Both quantitative output and consistency were achieved by integrating the oscillators in an artificial neural network (ANN).

3. The Neural Network

ANN's are software emulators of the nervous system and seemed adequate for earthquake forecasting because of their mathematical universality, their fault-tolerance, and their ability to deal with semi-quantitative data such as the modified Mercalli intensity (MMI).

In [9] it was shown that a neural network can behave as a model for the seismic process. That, together with the extensive use of neural nets in prediction tasks – mainly, again, in finance – indicated that these tools could be successful in integrating the oscillators to produce quantitative results.

Neural networks are distributed parallel processing systems (software programs) that emulate the behaviour of natural neurones in the animal nervous system. Figure 1 illustrates the particular kind of neurodes (artificial neurones) chosen for this task, of the many available – Rosenblatt's perceptron [10].

The input of a neurode is the sum *S* of the products of the outputs of other neurodes x_i by the so-called connection-weights, w_i:

$$S = \sum_{i=1}^{N} w_i x_i \tag{10}$$

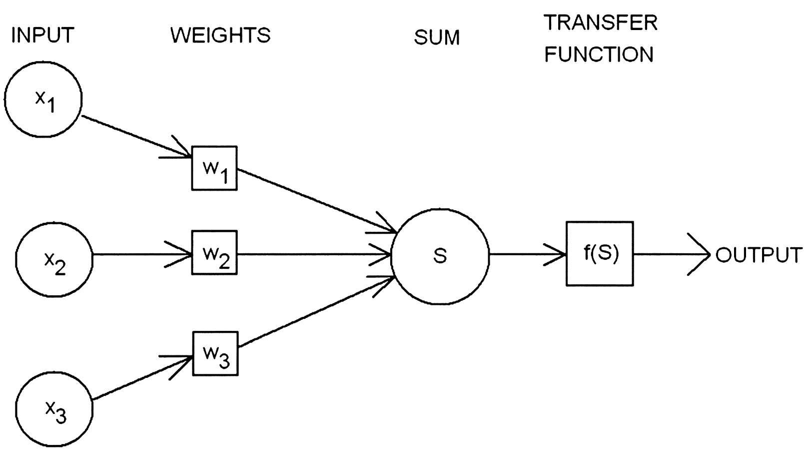

Figure 1. Rosenblatt's perceptron.

The output of a neurode is a function *f*(*S*) of that sum; in this case, the sigmoid function was chosen:

$$f(S) = \frac{1}{(1 + e^{-S})} \tag{11}$$

The complete neural network is a set of neurodes, arranged in layers, three in this case: an input layer (IL), a middle, or hidden, layer (HL) and an output layer (OL).

Neurodes are arranged in such a way that the ones in the same layer are not connected among themselves, but are connected to all the neurodes in the previous and next layers – the network is said, thus, to be partially connected.

The learning process is simple. The network is fed with an array of three vectors (time, intensity and location) and as many records as there are recorded seismic events. It would have been possible to convert intensities to magnitudes, using one of the several available energy-intensity relationships, following the work Shebalin [11] began in 1955, but this would only burden the ANN with a further degree of complexity, with a negligible improvement in accuracy.

The input layer neurodes are the input data: the oscillators. Since we intend to predict three variables (time, intensity and location), we shall have seven oscillators times three, 21 input neurodes.

The best size for the hidden layer was found to be of four neurodes.

The output layer will have as many neurodes as the variables to be predicted: three.

Then there are in all $(21 \times 4) + (4 \times 3) = 96$ connections. This architecture is illustrated in Figure 2.

The program computes the oscillators for each variable until the 29th occurrence (remember that one oscillator, MOM28, only yields results from the 29th value onwards) and, beginning with random weights, propagates the weighed sums to the hidden layer and from there to the output layer, where the output values are compared with the real values for the next (30th) event.

Using the algorithm of backpropagation of errors [12], the error is used to adjust the weights OL-HL and then HL-IL. The process is repeated for all the data and that concludes what is called one training epoch. Then, all these steps are repeated for as many epochs as are necessary to obtain an acceptable minimum training error – in the present case, 2% – when the net is considered to be trained. The implementation of the training algorithm was adapted from [13].

The net being trained, we can use the last set of connection weights – the one that yields minimal errors – and the oscillators that were calculated for the last observation, and propagate the weighed sums from IL to HL and from there to OL where we get the prediction of the time, intensity and location of the next, unknown occurrence.

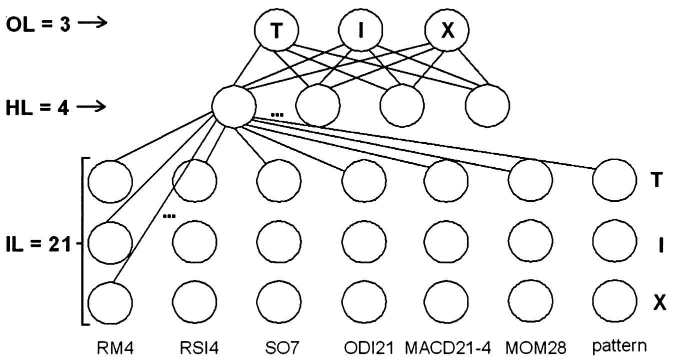

Figure 2. Architecture of the neural network. IL: input layer. HL: hidden layer. OL: output layer. T: times of pause. I: intensities. X: locations. Abbreviations for the oscillators are in the text.

4. Results

A computer program was developed that computes the oscillators, trains the neural network and predicts the parameters for the next event using the computed connection weights [8].

Since it would not be realistic to try and predict the exact geographical co-ordinates of the next epicentres, these were divided in major groups of 1° longitude, between 24° W and 31° W, so that epicentral location, $X =$ (W longitude) $- 23$.

In order to optimise both characteristics of neural networks, memory and generalisation, training sets with 100 examples were used. This number was chosen to be roughly equal to the number of connections (96), so that the network would be neither strongly overdetermined nor strongly underdetermined.

The neural network was trained with historical and instrumental seismic data from the Azores, between 1912 and 1993, for earthquakes with MMI $\geq$ V [14, 15]. It forecasted an earthquake in the Azores Central Group ($27 \leq X \leq 29$) with VI $\leq$ MMI $\leq$ VII in February 1998 $\pm$ 5 months [16].

Table 1. Summary of the results.

Forecasts					Occurrences		
Time		MMI					
Min	Max	Min	Max	Location	Time	MMI	Location
Sep 1997	Jul 1998	VI	VII	Central Group	Jul 1998	VIII	Central Group
Aug 2003	Aug 2004	VI	VII	Central Group	Jan 2004	V	Central Group

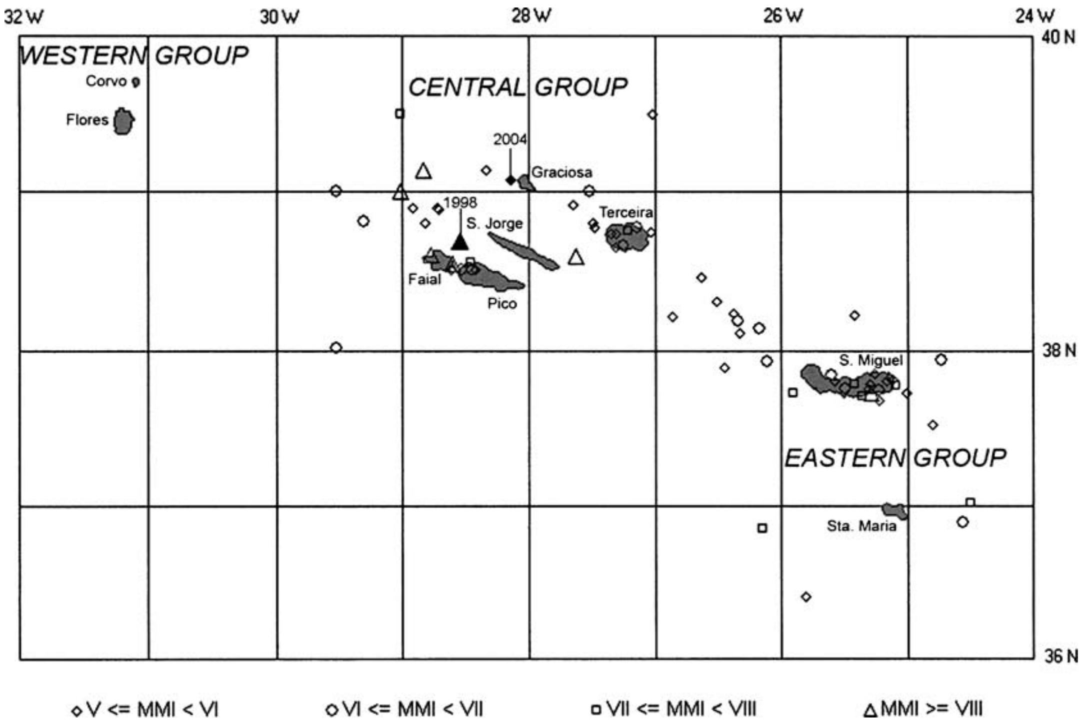

Figure 3. Location of epicentres of the 98 earthquakes with MMI $\geq$ V occurred in the Azores between 1912 and 2004. The two predicted earthquakes are marked (1998 and 2004). Data [6].

In July 9, 1998, an earthquake struck the Azores, being most destructive in the island of Faial, in the Central Group (MMI = VIII, 10 deaths).

When this last earthquake was included in the ANN training set, it forecasted an earthquake to be felt in the Azores Central Group with VI ≤ MMI ≤ VII in February 2004±6 months [17].

Between December 2003 and February 2004 the Azores experienced an earthquake swarm. The maximum intensity (MMI = V) was felt in the island of Graciosa, Central Group, on January 28, 2004 [18], the strongest since 1998. There was no major earthquake in the Azores between the 1998 and 2004 events.

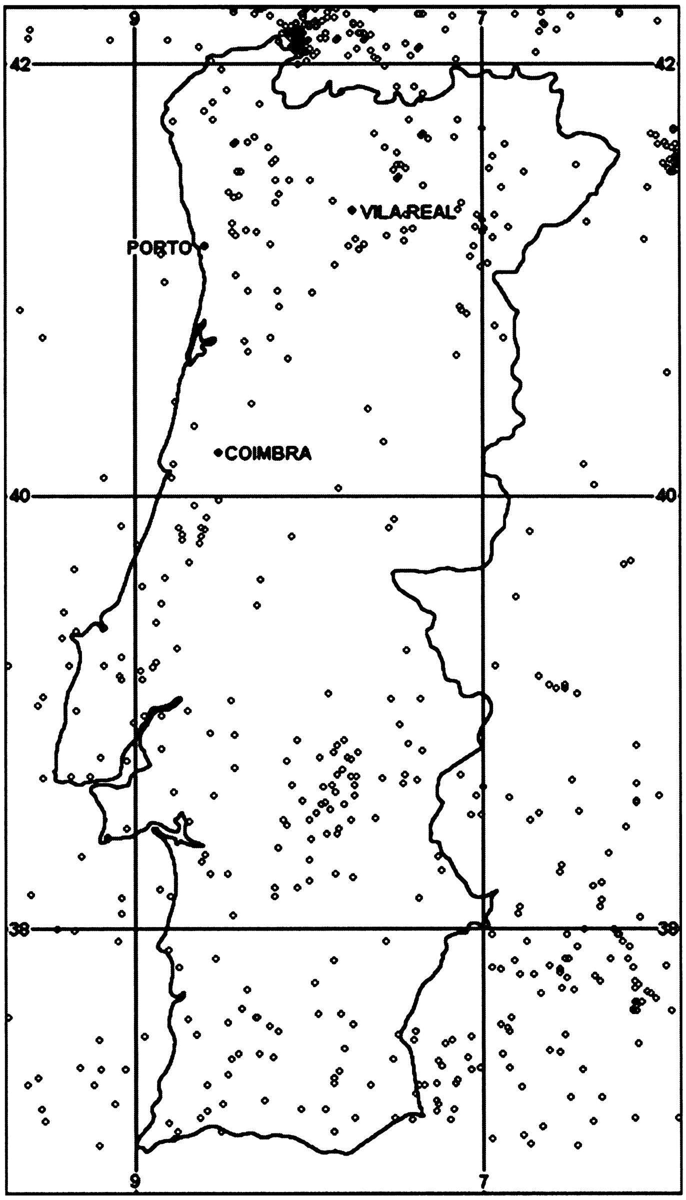

Figure 4. Location of 1090 epicentres of earthquakes with magnitude M ≥3.0 in Continental Portugal between 1974.03.11 and 2004.01.25. 71% of the epicentres are located North of latitude 40°. Data [15].

As can be seen in Table 1 the errors in both forecasts were related to the intensity. This was underestimated for the first event, although a visit to the site by members of the author's team suggested MMI = VII and the fact that [15] never mentioned an intensity for this event, the value VIII is attributed by [19]. Intensity was overestimated for the second earthquake.

It is notable that most of the Azorean seismicity before the 1998 earthquake was concentrated in and around São Miguel Island, in the Eastern Group, not in the Central Group where both predictions were accurately made (Figure 3).

The time windows for the predictions were accurate in both cases.

These results are encouraging, but more needs to be done.

5. Future Work

A group including researchers from the universities of Coimbra, Porto and Trás-os-Montes e Alto Douro (Vila Real) is beginning work on project *DESIRE* (*Dynamic Evaluation of Seismic Risk*) this time aimed at seismic forecasting in Continental Portugal. Three stations will be deployed near the three universities and monitor microseismicity, seismic waves' velocities, water piezometry, ground self-potential, EM piezoelectric emissions and magnetic susceptibility, known earthquake precursors [20, 21, 22]. This area was chosen due both to its moderate seismicity – around 35 events with $M \geq 3.0$ per year in the last 30 years, of which 71% had epicentres North of latitude 40° – and also because of the locations of the monitoring stations (Figure 4).

The collected physical data will be added as input nodes to a neural network that is similar to the one that was described above, together with the pre-processed seismic catalogue data, this time using magnitudes instead of intensities. It is expected that this mixed approach will succeed in narrowing the time forecast window – that is, to achieve the goal of mid- to short-term prediction.

Acknowledgements

The author is indebted to the insightful comments of one anonymous reviewer. This work was funded by project DESIRE, POCTI, FCT.

References

1. Achelis, S. B., *Technical Analysis from A to Z*, McGraw-Hill, New York, 1995.
2. Turcotte, D. L., *Fractals and Chaos in Geology and Geophysics*, Cambridge University Press, Cambridge, UK, 1993.
3. Mandelbrot, B. B., *The Fractal Geometry of Nature*, Freeman, New York, 1982.
4. Pietronero, L., Tosatti, E., Tosatti, V., and Vespignani, A., 'Explaining the uneven distribution of numbers in nature: The laws of Benford and Zipf', *Physica A* **293**, 2001, 297–304.
5. Alves, E. I., 'Notice on the predictability of earthquake occurrences', *Memórias e Notícias* **117**, 1994, 51–61.
6. Nunes, J. C., Forjaz, V. H., and Oliveira, C. S., 'Catálogo sísmico da região dos Açores, VERSÃO 1.0 (1850–1998)', CD-ROM, *Sísmica 2004, Proceedings of the 5th Congresso sobre Sismologia e Engenharia Sísmica*, Universidade do Minho, Guimarães, Portugal, 2004.
7. Richter, C. F., *Elementary Seismology*, Freeman, San Francisco, 1958.
8. Alves, E. I., *Uma Abordagem Conexionista ao Problema da Previsão Sísmica*. Ph.D. Thesis, University of Coimbra, 1995.
9. Alves, E. I., 'A simple seismic computational model', *Memórias e Notícias* **120**, 1995, 37–42.
10. Rosenblatt, F., *Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms*, Spartan Books, New York, 1962.

11. Shebalin, N. V., 'O svyazi mezhdu energey, ballnost'yu i glubinoy ochaga zemletryaseniya', *Izvestiya Akademiya Nauk SSSR, Ser. Geofiz.*, 1955, pp. 3–21 (cit. by [15]).
12. Rumelhart, D. E., Hinton, G. E., and Williams, R. J., 'Learning internal representations by error propagation', in *Parallel Distributed Processing*, Vol. 1, D. E. Rumelhart, J. L. McClelland, and the PDP Research Group (eds.), MIT Press, Cambridge, MA, 1986, pp. 318-362.
13. Rich, E. and Knight, K., *Artificial Intelligence*, McGraw-Hill, Singapore, 1991.
14. Costa Nunes, J., *Sismicidade Histórica e Instrumental do Arquipélago dos Açores. Catálogo Preliminar: 1444–1980*. Report INMG/LNEC, Lisboa, 1986.
15. USGS, Preliminary determinations of epicenters: Monthly listing – United States Department of the Interior – Geological Survey, NEIC – National Earthquake Information Centre, Denver, CO, http://neic.usgs.gov, 2005.
16. Alves, E. I., *Techniques for Predicting Natural Disasters*. 1st Cycle of Conferences on Engineering Geology, guest lecture. Institute for Water Research, Coimbra, October 8, 1996.
17. Alves, E. I., 'Earthquake prediction in the Azores: First results and forecasts', *Journal of Conference Abstracts* **4**, 1, 1999, 544.
18. Azores Regional Civil Protection and Firemen Service, Communiqué 11/2004.
19. Senos, L., Gaspar, J. L., Cruz, J., Ferreira, T., Nunes, J.C., Pacheco, J., Alves, P., Queiroz, G., Dessai, P., Coutinho, R., Vales, D., and Carrilho, F., 'O terramoto do Faial de 9 de Julho de 1998', *Proceedings of the 1° Simpósio de Meteorologia e Geofísica da APMG*, Lagos, Portugal, Nov. 23–25, 1998, pp. 61–68.
20. Sadovski, M. A., Nersesov, I. L., Nigmatullaev, S. K., Latynina, L. A., Lukk, A. A., Semenov, A. N., Simbireva, I. G., and Ulomov, V. I., 'The processes preceding strong earthquakes in some regions of Middle Asia', *Tectonophysics* **14**(3/4), 1972, 295–308.
21. Johnston, M. J. S. and Mortensen, G. E., 'Tilt precursors before earthquakes on the San Andreas Fault', *Science* **186**, 1974, 1031–1033.
22. Latoussakis, J., Stravakakis, G., Drakoupolos, J., Papanastassiou, D., and Drakatos, G., 'Temporal characteristics of some earthquake sequences in Greece', *Tectonophysics* **193**(4), 1991, 299–310.