

# **Safety in Mines Research Advisory Committee**

**Final Project Report (Revised)**

## **Seismology for rockburst prediction**

**W de Beer**

**Research agency : ISS International**  
**Project number : GAP 409**  
**Date : February 2000**

**Document No.: GAP409-REP-005-01**

# Executive Summary

This final report on the SIMRAC project GAP409 presents a method (SOOTHSAY) for predicting larger mining induced seismic events in gold mines, as well as a pattern recognition algorithm (INDICATOR) for characterising the seismic response of rock to mining and inferring future behaviour. The pattern recognition algorithm codifies and quantifies previous intuitive, qualitative analyses. The predictor, based on accelerating seismic release, depends on the existence of sufficient data, a past history in an area, and power law accelerating behaviour of a seismic time series.

Literature surveys form a crucial part of the research throughout the life of the project since prediction is a fast-developing and changing field. Several experts were consulted throughout the project, either in person or via e-mail in order to ensure that the latest techniques are taken into account.

The pattern recognition algorithm has been applied to several mines in the Carletonville area, as well as in the Free State, with clear patterns shown to emerge when input boundary conditions and thresholds are tuned to an area.

SOOTHSAY, the prediction algorithm, was applied to the deepest shaft pillar extraction in world in the West Wits region, as well as to a novel shaft pillar extraction in the Free State. Defining the time series of a specific function on a catalogue as a prediction strategy, the algorithm currently has a success rate of 53% and 65%, respectively, of large events claimed as being predicted in these two cases, with uncertainties in the predicted time of occurrence of a large instability of less than a week. In both cases, the prediction strategies are very far from random, and present a significant improvement over a pessimist strategy of trying to randomly forecast a large event daily by factors of 5.7 and 3.26, respectively.

As one way of unearthing analysis artefacts, the algorithm SOOTHSAY was applied to a pseudo-random synthetic catalogue. In this case only one prediction was made, for one function, with a 0% success rate for all functions.

The pattern recognition algorithm was developed in such a way that it can be extended to a fully-fledged prediction strategy in its own right. While further rigorous, statistical testing has to be performed on the SOOTHSAY algorithm, it is recommended that the current version be implemented on an experimental basis.

# Acknowledgements

The following people contributed their time, knowledge and patience to this report:

Dr Stefanel Radu, formerly of ISS International and ISS Pacific contributed Section 4.2 and Dr A J Mendecki of ISS International had a large influence on the course of the research.

## Direct assistance:

Dragan Amidzic	- AngloGold West Wits Region
Leandro Boonzaaier	- ISS International
Ric Ferriera	- East Driefontein
Marie Geringer	- ISS International
Emrich Hamman	- Formerly of Savuka Mine
Ray Heasman	- ISS International
Gerhard Hofmann	- ISS International
Clare Ilhan	- ISS International
Gerrit Kotze	- ISS International, ISS Geophysics
Kevin Riemer	- West Driefontein
Michael Sciocatti	- Formerly of ISS International
Gerrie van Aswegen	- ISS International
Elsje van Zyl	- ISS International
Vlok Visser	- Formerly of TauTona Mine

The management of AngloGold and Gold Fields of South Africa

## Discussions:

Prof Per Bak  
Niels Bohr Institute  
Copenhagen

Prof Yehuda Ben-Zion  
Department of Earth Sciences  
University of Southern California

Prof Hendrik Geyer  
Institute for Theoretical Physics  
University of Stellenbosch

Prof V. I. Keilis-Borok  
International Institute of Earthquake Prediction Theory and Mathematical Geophysics  
Russian Academy of Sciences

Prof Leon Knopoff  
Department of Physics and Astrophysics and Institute of Geophysics and Planetary Physics  
University of California

Prof Frikkie Scholtz  
Institute for Theoretical Physics  
University of Stellenbosch

Prof Didier Sornette  
Laboratoire de Physique de la Materie Condensee  
CNRS  
Universite de Nice-Sophia Antipolis

Prof Donald Turcotte  
Department of Geological Sciences  
Cornell University

Jane Dennis of ISS International typeset and produced this report.

# Table of Contents

Executive Summary .....	ii
Acknowledgements .....	iii
Table of Contents .....	v
List of Figures .....	viii
List of Tables .....	xi
Glossary .....	xiii
<b>1 Introduction .....</b>	<b>1</b>
<b>1.1 Research problem statement .....</b>	<b>1</b>
<b>1.2 Objectives and aims of the project .....</b>	<b>2</b>
<b>1.3 Research design .....</b>	<b>2</b>
1.3.1 Context .....	2
1.3.2 Research design .....	2
<b>2 Literature survey .....</b>	<b>4</b>
2.1 Accelerated seismic release .....	4
2.2 Criticality and self-organised criticality .....	4
2.3 Pattern recognition .....	5
2.4 Statistics .....	5
2.5 General .....	6
<b>3 Research methodology .....</b>	<b>7</b>
<b>3.1 Time to failure and assessment of applicability of various measures of seismic release to predict time to failure .....</b>	<b>7</b>
3.1.1 The time to failure concept .....	7
3.1.2 Comparative studies of phenomenological models that describe the kinematics of failure .....	8
3.1.3 Establish the predictive power, applicability and efficiency for each model .....	9
3.1.3.1 Non-parametric statistics .....	9
3.1.3.2 Time to failure prediction algorithm .....	9
3.1.4 Testing for deterministic components of time series of interest, noise reduction .....	11
<b>3.2 Renormalisation group theory, inversion and neural network techniques .....</b>	<b>12</b>
3.2.1 Improvement of phenomenological models by renormalisation group theory - larger seismic events as critical phenomena .....	12
3.2.2 Inversion of the time to failure .....	12
3.2.3 Neural networks .....	12
<b>3.3 Implementation .....</b>	<b>12</b>
<b>4. Results .....</b>	<b>13</b>

4.1	Pattern recognition: INDICATOR	13
4.1.1	Output	13
4.1.2	Decision 1: Relevance of questions	15
4.1.3	Decision 2: Hamming distance	15
4.1.4	Results	17
4.2	Deterministic components of time series, denoising and limits of predictability	32
4.3	Power law behaviour	40
4.3.1	Critical volumes and time windows	40
4.3.1.1	Data	41
4.3.1.2	Critical volume and time frame	45
4.3.1.3	Discussion: Accelerated seismic release, critical volumes and time windows	48
4.3.2	Critical behaviour without volume optimisation	48
4.3.2.1	The real data	49
4.3.2.2	Random catalogue	49
4.3.2.3	Case studies	49
4.3.2.4	Main events included in sequence	50
4.3.2.5	Forecast window	55
4.4	Predictor: Time to failure inversion	58
4.4.1	Free State shaft pillar extraction	58
4.4.2	West Wits pillar extraction	63
4.4.3	Random catalogue	67
5	Conclusions and recommendations	68
5.1	Conclusions	68
5.1.1	Primary output	68
5.1.2	Pattern recognition	68
5.1.3	Non linear dynamics	68
5.2	Recommendations	69
5.2.1	Pattern recognition	69
5.2.2	Time to failure predictor: shortcomings	69
5.2.3	Time to failure predictor: implementation	69
	References	70
Appendix 1	Pattern recognition: Non-parametric statistics questionnaire	73
Appendix 2	Prediction learning phase and error analysis	75
A2.1	$\text{Log}_{10}C$ vs $\text{Log}_{10}(\text{Release})$	75
A2.2	Characteristic release	76
A2.3	Error diagrams	77
Appendix 3	Deterministic component of seismic time series, noise reduction and limits of predictability	79
A3.1	False nearest strands	79
A3.2	Noise reduction	79
A3.3	Limits of predictability	81
Appendix 4	Larger seismic events as critical phenomena	82
A4.1	Criticality: A common example	82

A4.1.1	Correlation length	82
A4.1.2	Driving (tuning) parameter	82
A4.1.3	Order parameter	84
A4.1.4	Phase transitions	84
A4.1.4.1	Discontinuous phase transitions	84
A4.1.4.2	Continuous phase transitions	84
A4.1.5	Critical point: scale invariance	84
A4.1.6	Example: Ferromagnetic continuous phase transition	85
A4.2	Seismicity and scaling	85
A4.2.1	Gutenberg-Richter	85
A4.2.2	Time to failure	87
A4.2.2.1	Scale invariant processes and power laws	87
A4.2.2.2	Scale invariance in time	87
A4.2.2.3	What does scale invariance in time mean?	88
A4.3	Renormalisation group theory	88
Appendix 5	Contractual project details	90
Appendix 6	Extended literature review	94
A6.1	Project Methodologies	94
A6.1.1	Time to failure (accelerated seismic release) and refinements	94
A6.1.2	Criticality and self organised criticality	101
A6.1.3	Pattern recognition	104
A6.1.4	Statistics and error analysis	106
A6.2	Sceptics	108
A6.3	Seismicity in mines	115
A6.4	General: Earthquakes and earthquake prediction	118

# List of Figures

**Figure 3.1.1** Graphical representation of critical accelerating behaviour by a variable (power law behaviour) ..... 7

**Figure 4.1.3** Concept map ..... 17

**Figure 4.2.1a-e** An example of phase plots of cross correlation functions log moment analysis for Area 3 (4a). C1-i (horizontal axis) is the cross correlation function between the original time series and itself. C1-j (vertical axis) is the cross correlation function between the original signal and the denoised signal where the SVD required at least  $j$  neighbours in hyperspheres around states in order for it to be applied. As the minimum requirement approaches five neighbours the spreading becomes significant and the denoised signal no longer retains enough of the original signal's information content. .... 36

**Figure 4.2.2a-b** The level of the random component versus the minimum number of neighbours required to perform the SVD reduction. Different lines correspond to the 11 different areas where the analysis was performed. Notice that the gain in noise reduction is predominant when the number of neighbours is below 4 while virtually negligible above 4-5. This is also the limit beyond which the signal no longer retains its original information content. .... 37

**Figure 4.2.3** The effect of SVD noise reduction on (a) limits of predictability and (b) the deterministic content of time series of logE for five data sets on the West Rand. For two of the sets the results for (b) essentially coincide. .... 38

**Figure 4.2.4** The effect of SVD noise reduction on (a) limits of predictability and (b) the deterministic content of time series of logM for five data sets on the West Rand. For two of the sets the results for (b) essentially coincide. .... 39

**Figure 4.3.1.1(a)** ..... 42

**Figure 4.3.1.1(b)** ..... 43

**Figure 4.3.1.1(c)** ..... 44

**Figure 4.3.1.1(d)** ..... 44

**Figure 4.3.1.1(e)** ..... 45

**Figure 4.3.1.2(a)** ..... 46

**Figure 4.3.1.2(b)** ..... 46

**Figure 4.3.2.4 (a)** Critical time windows and exponents for  $\log M \geq 12.0$  events. The final source parameters were included in the analysis in order to establish pattern recognition benchmarks. Note the period of stable critical time windows (solid arrow) characterising a period of fast rockmass response during a particular type of mining. Preparation time and critical time window are equivalent concepts. Note also in this and following figures, that the x-axis does not represent time, but individual events. The solid arrow spans a period of 6 months and the dashed arrow only 3 months. .... 51

**Figure 4.3.2.4 (b)** Critical time windows and exponents for  $\log M \geq 12.0$  events. The final source parameters were included in the analysis in order to establish pattern recognition benchmarks. Note the period of stable critical time windows (solid arrow) characterising a period of fast rockmass response during a particular type of mining. Preparation time and critical time window are equivalent concepts. Note also in this and following figures, that the x-axis does not represent time, but individual events. The solid arrow spans a period of 6 months and the dashed arrow only 3 months. .... 52

**Figure 4.3.2.4 (c)** Critical time windows and exponents for  $\log M \geq 12.0$  events. The final source parameters were included in the analysis in order to establish pattern recognition benchmarks. Note the period of stable critical time windows (solid



arrow) characterising a period of fast rockmass response during a particular type of mining. Preparation time and critical time window are equivalent concepts. Note also in this and following figures, that the x-axis does not represent time, but individual events. The solid arrow spans a period of 6 months and the dashed arrow only 3 months. . . . . 53

Figure 4.3.2.4 (d) Same procedure as Figures (a) to (c), but for random data. Note an expected increase in time window as the length of the data set increased. This is absent in real data, since real seismicity has deterministic fixed points. . . . . 54

Figure 4.3.2.4 (e) Same as Figure 4.3.2.4(d), but for a random “Seismic Energy”. . . . . 54

Figure 4.3.2.4 (f) Exactly the same procedure as before, for random data. Note the absence of unique structure in the critical time windows. . . . . 55

Figure 4.4.1(a) Free State pillar extraction. Mining is concentrated at about 2000m depth, on the northern and southern sides of the shaft barrel. . . . . 58

Figure 4.4.1(b) Average error made in the prediction of the time to failure in a Free State pillar extraction by using the SOOTHSAY algorithm and using time series of cumulative apparent volume. . . . . 60

Figure 4.4.1(c) Median error made in the prediction of the characteristic cumulative apparent volume in a Free State pillar extraction by using the SOOTHSAY algorithm. . . . . 61

Figure 4.4.1(d) Error diagram for the prediction strategies in section 3.1.2 applied to the seismicity accompanying a Free State shaft pillar extraction. Note the good performance of the cumulative apparent volume, even though it still is not optimal - more experimentation with spatial selection is required. . . . . 63

Figure 4.4.2(a) West Wits pillar extraction. Mining is concentrated at about 3000m depth, on the south-eastern corner of the shaft pillar. The reef dips at about 22° to the south-south-west. . . . . 63

Figure 4.4.2(b) Median error made in the prediction of the time to failure in a West Wits pillar extraction by using the SOOTHSAY algorithm and using time series of cumulative apparent volume. . . . . 64

Figure 4.4.2(c) Median error made in the prediction of the characteristic cumulative apparent volume in a West Wits pillar extraction by using the SOOTHSAY algorithm. . . . . 65

Figure 4.4.2(d) Error diagram for the prediction strategies in section 3.1.2 applied to the seismicity accompanying a West Wits shaft pillar extraction. Note the good performance of the cumulative apparent volume, even though it tends towards an optimist strategy. . . . . 66

Figure A1.1 . . . . . 74

Figure A2.1 The constant C in the time to failure equation fitted to characteristic events with known power law precursory time series. . . . . 75

Figure A2.2 Cumulative characteristic seismic release in a shaft pillar of a West Wits mine by four events with  $\log_{10}M_0 > 11.5$  fitted to a straight line. The outer lines represent two standard deviations either side of the empirical (fitted) mean release. . . . . 77

Figure A2.3 Generic error diagram with points A (optimist strategy), B (pessimist strategy), AB (random guess). . . . . 78

Figure A4.1 Phase diagram of a typical liquid. (a): A slice through the - T plane. The lines are values of P and T for which the liquid undergoes discontinuous (in the order parameter) phase transitions. The point C is the critical point beyond which it is possible to move continuously from a liquid to a gas and back. (b):

*A slice through the P-T plane. The curve is the vapour pressure curve, representing values of the densities of the coexisting liquid and gas (after Yeomans 1992).* ..... 83

*Figure A4.2.1* ..... 86

*Figure A4.3.1* ..... 89

# List of Tables

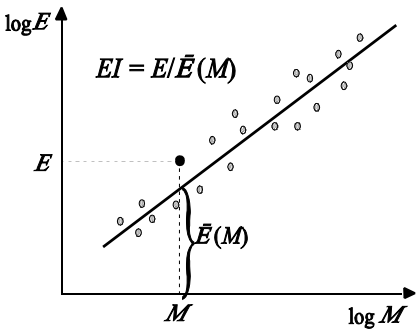
<b>Table 4.1.1</b>	<i>Example: INDICATOR output</i>	14
<b>Table 4.1.4.1(a)</b>	<i>Reduced INDICATOR output for Area 1A, with the forecast point 6 minutes before an event</i>	19
<b>Table 4.1.4.1(b)</b>	<i>Reduced INDICATOR output for Area 1A, with the forecast point 6 hours before an event</i>	20
<b>Table 4.1.4.2(a)</b>	<i>Reduced INDICATOR output for Area 1C(a), with the forecast point 6 minutes before an event</i>	21
<b>Table 4.1.4.2(b)</b>	<i>Reduced INDICATOR output for Area 1C(a), with the forecast point 6 hours before an event</i>	22
<b>Table 4.1.4.3(a)</b>	<i>Reduced INDICATOR output for Area 1C(b), with the forecast point 6 minutes before an event</i>	23
<b>Table 4.1.4.3(b)</b>	<i>Reduced INDICATOR output for Area 1C(b), with the forecast point 6 hours before an event</i>	24
<b>Table 4.1.4.4</b>	<i>Kernel output for Area 1A</i>	25
<b>Table 4.1.4.5(a)</b>	<i>Kernel output for Area 1A, with the forecast point 6 minutes before an event</i>	26
<b>Table 4.1.4.5(b)</b>	<i>Kernel output for Area 1A, with the forecast point 6 hours before an event</i>	27
<b>Table 4.1.4.6(a)</b>	<i>Kernel output for Area 1C(a), with the forecast point 6 minutes before an event</i>	28
<b>Table 4.1.4.6(b)</b>	<i>Kernel output for Area 1C(a), with the forecast point 6 hours before an event</i>	29
<b>Table 4.1.4.7(a)</b>	<i>Kernel output for Area 1C(b), with the forecast point 6 minutes before an event</i>	30
<b>Table 4.1.4.7(b)</b>	<i>Kernel output for Area 1C(b), with the forecast point 6 hours before an event</i>	31
<b>Table 4.2.1</b>	<i>Noise reduction in seismic time series: logE analysis</i>	34
<b>Table 4.2.2</b>	<i>Noise reduction in seismic time series: logM analysis</i>	35
<b>Table 4.3.1.2</b>	<i>Critical exponents, radii and time windows</i>	47
<b>Table 4.3.2.4 (a)</b>		50
<b>Table 4.3.2.4 (b)</b>		53
<b>Table 4.3.2.5 (a)</b>		55

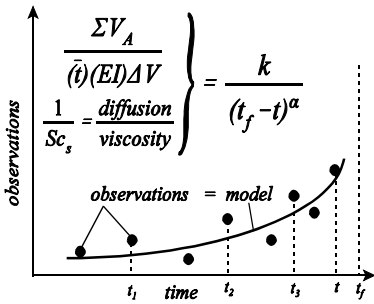
<b>Table 4.3.2.5 (b)</b> .....	55
<b>Table 4.3.2.5 (c)</b> .....	56
<b>Table 4.3.2.5 (d)</b> .....	56
<b>Table 4.3.2.5 (e)</b> .....	57
<b>Table 4.3.2.5 (f)</b> .....	57
<b>Table 4.4.1 Pattern summary</b> .....	62
<b>Table 4.4.2 Pattern summary</b> .....	66

# Glossary

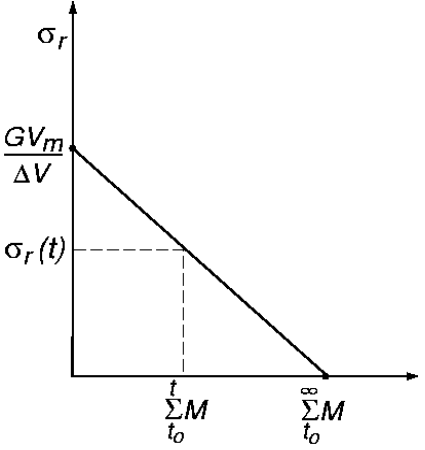
Parameter, relevant formula	Description
<p><b>Magnitude, <math>m</math></b>  <math>m = \log(A/T) + C</math>  <math>A/T</math> - the maximum displacement over associated period in the P- or S-wave group  <math>C</math> - corrections for path effects, site response and source region</p>	<p>Magnitude is a relative measure of the strength of a seismic event based on measurements of maximum ground displacement at a given frequency at multiple seismic sites. A unit increase in magnitude corresponds to a 10-fold increase in amplitude of ground displacement. Gutenberg and Richter related seismic energy and magnitude derived from P-waves recorded at large distances from the source at 1sec period as  <math>\log E(\text{ergs}) = 2.4m + 5.8</math></p>
<p><b>Seismic moment, <math>M</math>, [Nm]</b>  and  <b>Moment-magnitude, <math>m</math></b>  <math>m = 2/3 \log M - 6.1</math></p>	<p>A scalar that measures the coseismic inelastic deformation at the source. Since seismic moment is proportional to the integral of the far field displacement pulse it can easily be derived from recorded waveforms. A relation that scales seismic moment into magnitude of a seismic event is called moment-magnitude.</p>
<p><b>Seismic moment tensor</b>  <math display="block">M_{ij} = \int_V c_{ijkl} \Delta \varepsilon_{kl} dV = \int_V \Delta \sigma_{ij} dV</math> where  <math>c_{ijkl}</math> - elastic constants  <math>\Delta \varepsilon_k</math> - strain change at the source  <math>\Delta \sigma_{ij}</math> - stress change or change in moment per unit volume</p> <p><math>\Delta \theta = \text{tr}(M_{ij}) / (3\lambda + 2G)</math>, where  <math>\lambda</math> - the second Lamé constant  <math>G</math> - rigidity</p>	<p>The most general description of the processes at the seismic source <math>V</math> is by the distribution of forces or moments equivalent to the inelastic deformation. One can describe the inelastic processes at the source as the stress-free change of size and shape of an elastic body without alteration of the elastic properties of the region. If change in size and shape can be expressed as a change in strain <math>\Delta \varepsilon_{kl}</math>, then the equivalent stress change, or change in moment per unit volume is proportional to the strain change. The total moment integrated over the source volume is the seismic moment tensor, <math>M_{ij}</math>. For long waves compared to the source size, the whole source volume <math>V</math> can be considered to be a system of couples located at, say, the centre of <math>V</math>, and the moment tensor components can be defined by the equation at left. The moment tensor measures the inelastic deformation at the source during the seismic event and its value at the end of the source process measures the permanent inelastic strain produced by the event.</p> <p>The seismic moment tensor can be decomposed into isotropic (or volume change) and deviatoric components providing an additional insight into the nature of the coseismic strain drop.</p> <p>For a homogeneous body, the coseismic volumetric change, <math>\Delta \theta</math>, can be calculated from the second equation at left.</p> <p>The eigenvalues and corresponding eigenvectors of the deviatoric component of the seismic moment tensor describe the magnitude and orientation, respectively, of the principal moment axes (neglecting gravity) acting at the source. These principal moment axes are uniquely determined by moment tensor inversion. Principal moment orientation data can</p>

Parameter, relevant formula	Description
	provide sufficient information to find the best stress tensor.
<b>Radiated seismic energy, <math>E</math>, [J]</b>	The portion of the energy released or work done at the source that is radiated as seismic waves. Seismic energy is proportional to the integral of the squared velocity spectrum in the far field and can be derived from recorded waveforms. Radiated seismic energy increases with stress drop, seismic moment and with the traction rate i.e., stress oscillations at the source.
<b>Corner frequency, <math>f_0</math>, [Hz]</b> and <b>Source size, <math>l</math>, [m]</b> $l = c_1 / f_0$ $c_1 \approx 2500$ for S-wave in hard rock	The frequency at which a source radiates the most seismic energy observed as the maximum on the source velocity spectrum or as the point at which a constant low frequency trend and a high frequency asymptote on the recorded source displacement spectrum intersect. The corner frequency is inversely proportional to the characteristic size of the source.
<b>Stress drop, <math>\Delta\sigma</math>, [Pa]</b> $\Delta\sigma = c_2 M f_0^3$ $c_2 \approx 1.8 \times 10^{-10}$ for S-waves in hard rock $\Delta\sigma = G\Delta\varepsilon$ , and $\Delta\varepsilon$ - strain drop	Stress drop estimates the stress release at the seismic source. Although it is model dependent it provides reasonable estimates and a fair comparison amongst different sources from the same region recorded by the same seismic system.
<b>Source area, [m<sup>2</sup>]</b> $A = M / (Gu)$ $u$ - average displacement at the source.	The area of coseismic inelastic deformation over the planar source.
<b>Source volume, [m<sup>3</sup>]</b> $V = M / \Delta\sigma$	The volume of coseismic inelastic deformation of the order of $\Delta\sigma/G$ .
<b>Apparent stress, [Pa]</b> $\sigma_A = GE / M = E / (\Delta\varepsilon V)$ or $\sigma_A = E / (uA)$ .	Apparent stress is recognised as a model independent measure of the stress change at the seismic source.
<b>Apparent volume, [m<sup>3</sup>]</b> $V_A = M / (c_3 \sigma_A) = M^2 / (c_3 GE)$ $c_3$ - scaling factor $\approx 2$ .	The apparent volume scales the volume of rock with coseismic inelastic strain of an order of apparent stress over rigidity. The apparent volume $V_A$ is less model dependent than the source volume $V$ .

Parameter, relevant formula	Description
<p><b>Energy index, EI</b></p> 	<p>The notion of comparing the radiated energies of seismic events of similar moments can be translated into a practical tool called Energy Index (<math>EI</math>) – the ratio of the radiated energy of a given event (<math>E</math>) to the energy <math>\bar{E}(M)</math> derived from the regional <math>\log E</math> vs <math>\log M</math> relation for a given moment <math>M</math>.</p> <p>Since <math>\log \bar{E}(M) = c + d \log M</math>, then</p> $\bar{E}(M) = 10^{c+d \log M}$ <p>where <math>c</math> and <math>d</math> are constant for a given <math>\Delta V</math> and <math>\Delta t</math>. In general <math>d</math>-value increases with the system's stiffness and <math>c</math>, for a given <math>d</math>, increases with stress. A small or moderate event with <math>EI &gt; 1</math> suggests a higher than average shear stress at its location. The opposite applies to the <math>EI &lt; 1</math> case.</p>
<p><b>Seismic strain,</b>  <math>\varepsilon_s(\Delta V, \Delta t) = \sum M / (2G\Delta V)</math> and  <b>Seismic strain rate, [s<sup>-1</sup>]</b>  <math>\dot{\varepsilon}_s(\Delta V, \Delta t) = \varepsilon_s / \Delta t</math></p>	<p>Seismic strain measures strain due to cumulative coseismic deformations within the volume <math>\Delta V</math> over the period <math>\Delta t</math>. Its rate is measured by <math>\dot{\varepsilon}_s</math>.</p>
<p><b>Seismic stress, [Pa]</b>  <math>\dot{\sigma}(\Delta V, \Delta t) = 2G \sum E / \sum M</math></p>	<p>Seismic stress measures stress changes due to seismicity.</p>
<p><b>Seismic stiffness modulus, <math>K_s</math> [Pa]</b>  <math>K_s(\Delta V, \Delta t) = \sigma_s / \varepsilon_s</math>  <math>\varepsilon_s = 4G^2 \Delta V \sum E / (\sum M)^2</math></p>	<p>Seismic stiffness measures the ability of the system to resist seismic deformation with increasing stress. The stiffer systems limit both the frequency and the magnitude of intermediate and large events but have time-of-day distribution with larger statistical dispersion, thus are less time predictable.</p>
<p><b>Seismic viscosity, [Pa · s]</b>  <math>\eta_s(\Delta V, \Delta t) = \sigma_s / \dot{\varepsilon}_s</math></p>	<p>Seismic viscosity characterises the statistical properties of the seismic deformation process. Lower seismic viscosity implies easier flow of seismic inelastic deformation or greater stress transfer due to seismicity.</p>
<p><b>Seismic relaxation time, [s]</b>  <math>\tau_s(\Delta V, \Delta t) = \eta_s / G</math></p>	<p>Seismic relaxation time quantifies the rate of change of seismic stress during seismic deformation processes and it separates the low frequency response from the high frequency response of the system under consideration. It also defines the usefulness of past data and the predictability of the flow of rock. The lower the relaxation time, the shorter the time span of useful past data and the less predictable the process of seismic deformation.</p>

Parameter, relevant formula	Description
<p><b>Seismic Deborah number</b>  <math>De_s(\Delta V, \Delta t) = \tau_s / \text{flowtime}</math>            where <i>flowtime</i> is a design parameter not necessarily equal to <math>\Delta t</math>.</p>	<p>Seismic Deborah number measures the ratio of elastic to viscous forces in the process of seismic deformation and has successfully been used as a criterion to delineate volumes of rockmass softened by seismic activity (soft clusters). The lower the Deborah number the less stable is the process or the structure over the design <i>flowtime</i> - what may be stable over a short period of time (large <math>De_s</math>) may not be stable over a longer time (lower <math>De_s</math>).</p>
<p><b>Seismic diffusivity, [m<sup>2</sup>/s]</b>  <math>D_s(\Delta V, \Delta t) = (\Delta V)^{2/3} / \tau_s</math>,            or in a statistical sense  <math>d_s = (\dot{X})^2 / \bar{t}</math>.</p>	<p>Seismic diffusivity can be used to quantify the magnitude, direction, velocity and acceleration of the migration of seismic activity and associated transfer of stresses in space and time. There is an inverse relationship between the diffusivity <math>D_s</math> and the friction parameters.</p>
<p><b>Seismic Schmidt number</b>  <math>Sc_{sd}(\Delta V, \Delta t) = \eta_s / (\rho D_s)</math> or  <math>Sc_{sd} = \eta_s / (\rho d_s)</math>            where <math>\rho</math> is rock density.</p>	<p>Seismic Schmidt number measures the degree of complexity in space and time (the degree of turbulence) of the seismic flow of rock. Note that seismic Schmidt number <math>Sc_{sd}</math>, encompasses all four independent parameters describing seismicity:  <math>\bar{t}, \bar{X}, \Sigma M, \Sigma E</math>.</p>
<p><b>Time to failure, (<math>t_f - t</math>)</b>  <math>d\Omega / dt = k(t_f - t)^m</math>  <math>\Omega</math> - measurable quantity  <math>t</math> - current time  <math>t_f</math> - time of failure  <math>k, m</math> - constants</p> 	<p>This concept describes the behaviour of materials in the terminal stages of failure. It views instability as a critical point, then precursors should follow characteristic power laws in which the rate of strain or other observable, measurable, quantity <math>\Omega</math> is proportional to the inverse power of remaining time to failure. Observed oscillations in <math>\Omega</math> of an increasing frequency as the failure approaches are part of the solution to time-to-failure equation with a complex exponent, where the imaginary part relates to discrete scale transformation and introduces log-periodic oscillations decorating the asymptotic power law. The observations <math>\Omega</math> can be a combination of different seismic parameters that would exhibit power law type increase before failure. For well behaved data sets the time at failure <math>t_f</math> can be estimated from the times of three successive maxima (<math>t_1, t_2, t_3</math>) of the observed process <math>t_f = (t_2^2 - t_1 t_3) / (2t_2 - t_1 - t_3)</math>            Note that, in theory, <math>t_3 - t_2 &lt; t_2 - t_1</math></p>



Parameter, relevant formula	Description
<p><b>Seismic moments, volume mined and relative stress</b></p> 	<p>If a volume of rock, <math>V_m</math>, is mined out at time <math>t_0</math> and if the altered stress and strain field can readjust to an equilibrium state through seismic movements only, the sum of seismic moments released within a given period of time would be proportional to the excavation closure and in the long term at <math>t = t_\infty</math>, <math>\sum_{t_0}^{\infty} M \approx GV_m</math></p> <p>where <math>M</math> is the scalar seismic moment. The relative stress level at the time, <math>t</math>, in a given volume of rock <math>\Delta V</math> surrounding the excavation, can be calculated from the difference between <math>GV_m</math> and the cumulative moments released to date:</p> $\sigma_r(t) = (GV_m - \sum_{t_0}^t M) / \Delta V.$
<p><b>Seismic moments and volume of elastic convergence</b>  <math>\sum M = \gamma GV_e</math></p>	<p>The amount of strain energy stored when mining in elastic rock is directly proportional to the volume of elastic convergence, <math>V_e</math>. It has been found that the total amount of seismic moment resulting from mining within a large area and time period is related to the change in elastic convergence <math>V_e</math>. The proportional constant gamma, <math>\gamma</math>, has been found to vary between about 0.03 and 1.0. There is some evidence that <math>\gamma</math> is a function of the geotechnical area being mined.</p>

# 1 Introduction

This is the only way it clicks...  
It doesn't make any sense otherwise...  
But this is just a theory, isn't it?...  
Call it any name you like. It's good enough for me.

- Dashiell Hammett

The prediction of seismic events, be they tectonic earthquakes or induced by human activity, has been called the ultimate aim of all of seismology and earthquake science. This is a divergent and contentious field, with as many proponents of the feasibility of prediction as sceptics and detractors. However, from the current perspective it is clear that the bulk of evidence points to seismicity not being a random, unpredictable phenomenon. Consistent precursory patterns exist and have been documented extensively, and a physical theory, within the framework of statistical physics of complex phenomena is emerging. A theory is only useful if it has predictive capabilities, and as such this framework is being pursued precisely because of the predictability it shows.

The central output of this project is a prediction algorithm which automatically tests time series of nine functions quantifying seismic release for power law acceleration, learns from the history of an area and predicts larger seismic events with a success rate that varies from 53% to 65% of events successfully predicted within an uncertainty in time that varies from three days to six days, quantifying times of increasing probability of occurrence.

To date, prediction has been practised by very few people in the South African mining industry for two main reasons:

1. **The ostrich mentality: avoidance of the facts.** Prediction is a politically and economically risky issue, since the fact that precursory patterns exist creates high expectations from official, public and labour quarters. False alarms and failures to predict therefore carries a high cost in terms of trust and damages.

Better to avoid the issue, sweep facts under the carpet and sit out repeated inquiries where precursory patterns are retrospectively pointed out.

2. **The Windows mentality: point and click.** To learn to recognise precursory patterns takes a long time and involves a lot of painstaking effort. Application of this learned intuition time and again, in an iterative process, without getting distracted by initial failures, takes a lot of energy and commitment.

Better to say that in the modern fast business environment it is unreasonable to expect anything unless it can be created by pointing a mouse and clicking for an answer.

## 1.1 Research problem statement

*Mendecki et al (1996)* defined a new methodology for quantifying rock mass response to mining and introduced time series of new functions on catalogues which traced this response. On the basis of such time series and visualisation of the spatio-temporal

distribution and evolution of the functions, qualitative interpretations can be made. Quantitative levels of warning of impending rock mass instability can be established (*Spottiswoode and de Beer, 1998*) by a lot of careful experimentation and cooperation with mine rock engineering personnel.

Following this report (*Mendecki et al, 1996*), the need was identified for a more objective, automated methodology (see “the *Windows* mentality” above).

The current project aims to provide such an objective method which can be automated. In order to be more than statistics, the method has to be based in the physics of rock mass response to mining, and thus has, to a large extent, to investigate the current thinking in fracture mechanics, seismicity and earthquake dynamics.

In this report then, three problems were focussed on:

- quantification of precursory patterns in the functions defined in *Mendecki et al (1996)*
- quantitative determination of limits of predictability and observational evidence for the hypothesis that any seismic event is a critical point of a process characterised by smaller events
- the development of an automated predictor based on the kinematics of failure (still in the framework of criticality) as encoded in the time to failure accelerating seismic release curve.

## **1.2 Objectives and aims of the project**

The primary output required of this project was a method for the prediction of rock mass instability based on kinematics of failure, nonlinear dynamics (chaos theory) and neural networks, using data supplied by the modern seismic networks in South African mines.

## **1.3 Research design**

### **1.3.1 Context**

The research is conducted from the perspective of physics, focussing on the rock mass as a complex system approaching criticality. Retreat from criticality is effected by larger instabilities normalising the stress regime for a time as they release energy through large seismic events.

### **1.3.2 Research design**

#### **Literature study**

Chapter 2 highlights some of the main publications utilised for the design of a quantitative pattern recognition tool and the predictor. Seismic prediction is a dynamic, continually developing field, and literature study is an ongoing process. The resources are grouped chiefly according to the three points of the research problem statement.

## **Research methodology**

In Chapter 3 the contractually agreed methodology is fleshed out by detailing the procedures followed to establish a quantitative pattern recognition algorithm, investigate critical behaviour in rock mass response to mining and design a forward predictor.

## **Results**

The results of applying the methodologies in Chapter 3 to several case studies are presented in Chapter 4. It is shown that patterns of seismic behaviour can be quantified fairly straightforwardly. A predictor that is based on the kinematics of failure accelerated release curve performs adequately, when combined with a learning phase and subsequent feedback loop. The predictor learning phase and feedback loop, together with the pattern recognition algorithm replaces an earlier proposal to construct neural network architectures. A pseudo-random synthetic catalogue is also subjected to the prediction algorithm.

## **Conclusions and recommendations**

Chapter 5 concludes the report with comments on the three main thrusts of enquiry in the project, pointing out that

1. Patterns in seismicity can be quantified.
2. Quantitative evidence for criticality and deterministic, reasonable limits of predictability exists.
3. A predictor which predicts the occurrence of larger seismic events in a better than random fashion, with acceptable uncertainty in the order of three days has been designed and implemented in two case studies.

## 2 Literature survey

The body of literature on the subject of earthquake prediction is big and contentious. In this chapter, the main sources which informed the research are presented and summarised. During the three year duration of this project, in excess of two hundred publications were consulted, and the principal author benefited from many private communications via e-mail as well as discussions at the *Scale Invariance and Beyond* workshop at Les Houches in 1997 and the *Fifth Workshop on Non-Linear Dynamics and Earthquake Prediction*, Trieste, 1999. Here, publications which were found to be most aligned to the project outputs and, accordingly, used extensively, are highlighted. The procedure is to categorise the summary by the main lines of enquiry pursued.

### 2.1 Accelerated seismic release

*Varnes, 1987, 1989*

Synopsis: The author empirically fitted the cumulative square roots of the seismic energy release and the seismic moment prior to several large earthquakes to the time to failure curve and found extremely good fits. Recall that the time to failure curve, for a cumulative measure of seismic release, quantifies accelerating release.

*Brehm and Braile 1998a, 1998b, 1999*

Synopsis: The predictor SOOTHSAY developed in this project is based to a large extent on these three papers. The existence of a characteristic largest event size in an area or on a fault segment enables one to fit the final ("failure") release to a straight line. Once the appropriate size is found, the fit is remarkably good. A log - log relationship between the power-law constant and final release is exploited to constrain the inversion for the time to failure. They further claim an empirical relation between the seismic energy release by a large event and the power law exponent. In this project the latter relationship was found to have very little applicability - it simply does not appear often and, if it does, the fit is poor.

*Ashby and Dyson 1986*

*Glasstone et. al. 1941*

*Saito 1969*

*Servi and Grant 1951*

*Leckie and Hayhurst 1977*

Synopsis: These form a group of articles which serve to point out the ubiquity of the time to failure power law in various systems in nature, from creep to slope failure to rate processes in chemical reactions.

### 2.2 Criticality and self-organised criticality

*Bowman, Ouillon, Sammis, Sornette and Sornette 1998*

*Saleur, Sammis and Sornette 1996*

*Newman, Turcotte and Gabrielov 1996*

*Newman et al 1994*

*Bak 1996*  
*Turcotte 1999*  
*Boonzaaier 2000*

Synopsis: The question raised by the existence of the power law variation of a measure of seismic release with the time to failure is of course its origin. Although perhaps not the first articles on the subject, *Bowman et al* and *Saleur et al* wrote down in encapsulated format an answer in terms of a seismic event as a critical rupture, in the sense of a phase change. They employed mainly statistical physics concepts and renormalisation group methods. A further result from their work is the possible existence of log-periodic decorations of the standard time to failure curve. The papers by *Newman* and co-workers took the renormalisation group concept further, explicitly endowing the fracture space of rock with a hierarchical structure to reproduce time to failure power laws. This is a concept which underlies the ideas of self-organised criticality and fractality. The book by *Bak* serves as a comprehensive, easy-to-read layman's introduction to this topic, and for that reason the whole book is included here as a reference. Closer to home, *Boonzaaier* has shown that there are parallels between self-organised critical models and mining-induced seismicity.

## 2.3 Pattern recognition

*Gelfand et al 1976*  
*Keilis-Borok and Lichtman 1993*  
*Romachkova et al 1998*  
*Keilis-Borok and Soloviev 1999*  
*Kossobokov et al 1999*

Synopsis: *Gelfand et al* attempted to establish patterns in the geology and earthquake history of regions, learning to separate earthquake hypocentres from other places. The methodology was to draw up a questionnaire querying the hypocentre of a large earthquake on these parameters in such a way that the answers are of a clear "yes-no" type. The "Russian group" (*Keilis-Borok*, *Kossobov*, *Soloviev* and others) expanded the methodology into a pattern recognition prediction algorithm, M8. This methodology is the basis of the INDICATOR pattern recognition algorithm developed during this project.

## 2.4 Statistics

*Molchan 1999*  
*Sornette and Knopoff 1997*  
*JGI 1997*  
*Mendecki, Mountfort and van Aswegen 1999*

Synopsis: *Molchan* has developed a statistical model of earthquake prediction taking into account the cost of an alert and the randomness of a prediction strategy. The evaluation of the predictor developed in GAP409 follows these lines. The whole issue of the statistics of earthquake recurrence times is discussed in depth by *Sornette and Knopoff*. In *JGI*, papers delivered at the infamous "London" meeting in 1996, (where several prediction strategies were challenged) are presented, in order to balance the argument, as well as *Mendecki, Mountfort and van Aswegen, 1999*.

## 2.5 General

The books by *Cardy (1999)* and *Yeomans (1992)* give excellent introductions to statistical physics, criticality, renormalisation group theory and phase transitions. *Lomnitz (1997)* gives a comprehensive overview of earthquake prediction methodologies up to 1997. For a comprehensive discussion of seismic monitoring as practised in South African mines, the ISS International book (*Mendecki 1997*) is recommended.

### 3 Research methodology

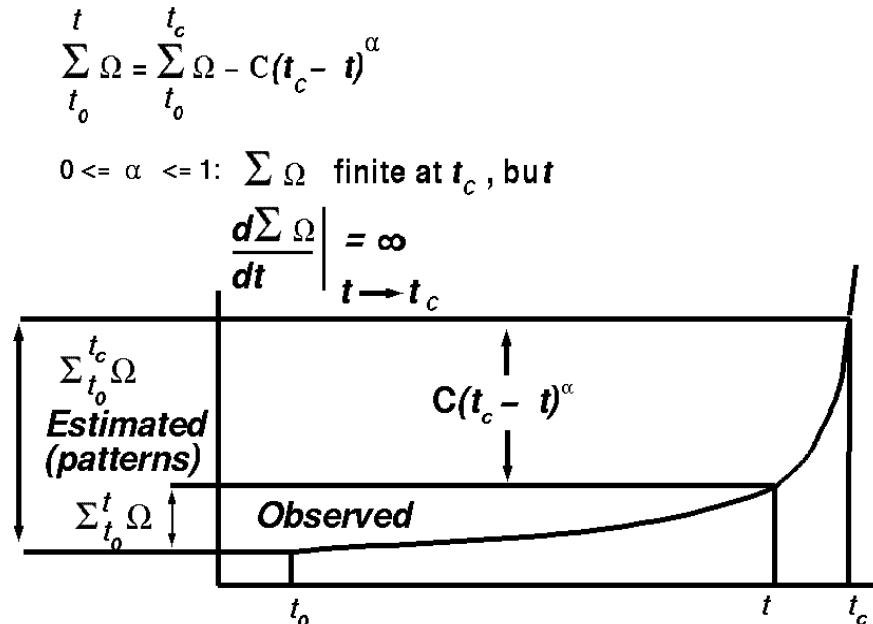
#### 3.1 Time to failure and assessment of applicability of various measures of seismic release to predict time to failure

##### 3.1.1 The time to failure concept

The function used for inversion of seismic parameters of interest is the power law in the time remaining to failure,

$$\sum_{t'=t_0}^t \Omega_{t'} = \sum_{t'=t_0}^{t_c} \Omega_{t'} - C(t_c - t)^\alpha \quad (3.1.1)$$

This function is exhibited in Figure 3.1.1. Cumulative quantities are used since one is interested in the cumulative behaviour of the rockmass as it approaches criticality. Furthermore, cumulative quantities makes the power law fit easier, since oscillations are smoothed. Intuitively, it is also easier to see at a glance that a linear sequence represents a stationary state of the rockmass, where each event contributes the same response as its predecessor. This is to be contrasted with the acceleration curve above, which clearly represents increasing instability.



**Figure 3.1.1 Graphical representation of critical accelerating behaviour by a variable (power law behaviour)**

Please note that a value of the critical exponent between zero and one implies that the release is finite, even at the moment of the characteristic event, but that the rate of change



of the seismic release diverges at the time of failure,  $t_c$  (*de Beer and Mendecki (1998)*).

In eq. (3.1.1), the final release,  $\Omega_{t_c}$ , the constant,  $C$ , the critical exponent,  $\alpha$  and the time of failure,  $t_c$  (or, more precisely, the time when the total cumulative release,  $\sum_{t'=t_0}^{t_c} \Omega_{t'}$  is reached), are inverted for.

The left hand side of equation (3.1.1) represents the observed seismic parameters (see Section 3.1.2 below).

The inversion is performed by searching for the minimum of the function

$$\min F = \left[ \sum_{t'=t_0}^t \Omega_{t'} - \sum_{t'=t_0}^{t_c-1} \Omega_{t'} - \Omega_{t_c} + C(t_c - t)^\alpha \right]^{L_1} \quad (3.1.2)$$

### 3.1.2 Comparative studies of phenomenological models that describe the kinematics of failure

Throughout the project, nine different models (functions on catalogues or measures of seismic release) were studied. Specifically, the functions were

1. The cumulative square root of the scalar seismic moment,  $\Sigma\sqrt{M}$ . This quantity, up to some prefactors which depend on stress drop and elastic properties of the medium, is known as the *Benioff strain* in the tectonic earthquake community. Note that, should one choose to just cumulate the scalar seismic moment, as opposed to the square root, the sequence is overly sensitive to the largest, rarest events. It is also possible to cumulate  $M^{1/3}$ , a quantity proportional to the linear dimension of the rupturing rock volume. The results discussed below remain robust if this function is used (*Ouillon & Sornette 1998*). Taking  $M^{2/3}$ , a quantity proportional to a rupture plane area, again gives too much weight to larger events.
2. The cumulative square root of the radiated seismic energy,  $\Sigma\sqrt{E}$ . In this case, taking the square root is a smoothing device to inhibit domination of a sequence by large events.
3. Cumulative Apparent Volume,  $\Sigma V_A$  (see glossary), measuring the total coseismic inelastic deformation during the approach of failure.
4. Apparent volume scaled by the inverse median energy index, cumulated,  $\Sigma(V_A/(EI_{Med}))$ . Since Energy Index, a measure of the average shear stress at a seismic source, is frequently observed to drop during the nucleation of an instability, (*Mendecki 1997, 1999*) while increasing apparent volumes of precursory events reflect accelerating deformation in the critical volume, this variable is tested for the power-law behaviour in Figure 3.1.1.
5. Cumulative inverse seismic Schmidt number,  $\Sigma 1/Sc_s$ . An increase in this function signals increased turbulence in the seismic flow of rock.
6. Cumulative inverse of the Seismic Stress,  $\Sigma 1/\sigma_s$ . As mentioned in the Glossary,

easier stress transfer due to seismicity implies a drop in seismic stress during nucleation. Thus the inverse of this quantity is tested for non-stationary (accelerating) behaviour.

7. Cumulative seismic strain rate,  $\Sigma \dot{\epsilon}$ , a logical candidate for accelerating behaviour.
8. Cumulative seismic diffusion,  $\Sigma(\bar{X})^2/\bar{t}$ , testing for accelerated migration of seismic activity.
9. Cumulative inverse seismic viscosity,  $\Sigma 1/\eta_s$ . Since a lower seismic viscosity implies easier flow of seismic inelastic deformation, and thus increasing instability, the inverse is expected to increase during precursory periods.

### 3.1.3 Establish the predictive power, applicability and efficiency for each model

Recall that two prediction philosophies are followed. First, time series in a catalogue are tested for power law behaviour. If found, the future release and time of release are estimated. If, however, the result of a training exercise on a catalogue is inconclusive as far as power law behaviour is concerned, non parametric statistics is employed in a pattern recognition algorithm.

#### 3.1.3.1 Non-parametric statistics

The nine functions introduced in Section 3.1.2 creates a quandary: Given that, in a particular area, a subset of these functions predict the occurrence of a large instability, without agreeing on the time or size of the final release, while another subset may not predict anything at all, a strategy to decide on which function to believe has to be devised.

The INDICATOR algorithm was developed to address this issue. The idea for it is rooted in the work of *Gelfand et al (1976)*, *Knopoff (1997)* and *Keilis-Borok and Lichtman (1993)*, who attempted to use nonparametric statistics to determine the probability that a given factor is a precursor to or a harbinger of a major seismic event. Details of the implementation of the algorithm are provided in Appendix 1.

#### 3.1.3.2 Time to failure prediction algorithm

The time to failure prediction algorithm (eq. 3.1.2) has to predict the final release,  $\Omega_{t_c}$ , the constant,  $C$ , the critical exponent,  $\alpha$  and the time of failure,  $t_c$ . It is obvious that the fact that four parameters are being inverted for in eq. (3.1.2) will almost inevitably lead to the inversion routine being trapped in a local minimum of  $F$ . The problem is tackled by establishing a positive feedback loop in which the algorithm learns the characteristics of the system. The methodology is as follows (*Brehm and Braile 1998, 1999*):

##### (a) Learning phase

1. An empirical, log-linear, relationship exists (Figure A2.1) between the constant  $C$  in eq. 3.1.2 and the final release  $\Omega_{t_c}$ , given by

$$\log C = F + E \log (\text{Release}).$$

Thus an initial learning set of large events are used to establish the

relationship. The large events are included in the power law sequences, if they exist.

2. Large, system-spanning episodes of cumulative release are related in a simple linear fashion in time (Figure A2.2). This relationship is established. The large events are included in the power law sequences, if they exist.
3. The periods between the set of characteristic large events are searched for power law sequences which exclude the characteristic events. The average and median times between the last event before a large event (the time to failure), as well as the relevant standard deviations are computed.
4. The average and median critical exponent  $\alpha$  and higher moments are established.
5. The average and median final release and higher moments are established.

**(b) Learning to predict phase**

For a further set of characteristic events

1. The constant  $C$  is replaced in eq. 3.1.2 by the to-be-predicted final release through the relation established in a(1), in the learning phase.
2. The linear relation between episodes of characteristic release is extrapolated to provide a first estimate to the inversion of the cumulative release at time  $t_c$ .
3. The moments of the exponent and time to failure (as defined above) are used as first estimates for the inversion.
4. Starting from six events after a characteristic event, the inversion is performed if an event is the termination of a power law sequence.
5. If the predicted release and time to failure falls within two standard deviations of the means computed above, a prediction is issued.
6. The average and median *errors* and higher moments in predicted quantities are computed.

**(c) Prediction phase**

1. The algorithm continues as before. Statistics continue to be collected in a sliding time window. This takes into account and quantifies the fact that, in the mining-induced seismicity, the driving force is not constant. Different rates of mining, different mining layouts and different geology introduces considerable variability.
2. However, now the results of the inversion is perturbed by the errors in prediction calculated before. These errors quantify the average over / under estimation by the algorithm.

3. ALERTS are issued when the predicted release and time to failure falls within the bounds in b(5) above.
4. Account is kept of the total duration of ALERTS.
5. When a characteristic event occurs, it is determined whether the event took place with a release and time to failure within the bounds of the last ALERT issued and it is declared whether the prediction of the event was
  - (1) **Successful prediction** -  
the event happened within the bounds above,
  - (2) **False positive prediction** -  
Also known as a false alarm. ALERTS were issued, but the event did not take place within the prescribed bounds,
  - (3) **False negative prediction** -  
Also known as a failure to predict. No ALERTS were issued - no predictions were made.
6. The event is added to the global learning set and the sliding learning set moves forward by one characteristic event.

**(d) Error analysis**

Appendix 2 contains a short explanation of the method used in this report to evaluate prediction strategies. In this report, a prediction strategy is defined to be the inversion of the power law time to failure of one of the nine functions in Section 3.1.2. Comparison of the rate of false negative alarms and the rate of time alarms is employed to decide which strategy, in a given area, is “least random”.

### **3.1.4 Testing for deterministic components of time series of interest, noise reduction**

Series were constructed from eleven catalogues from the West Wits region. These polygons were carefully selected in connection with the project GAP303 (*Mendecki and van Aswegen, (1997)*). Since these catalogues were not chosen specifically for the prediction of large events, they were used to calculate deterministic components and reduce the random component of the  $\log E$  and  $\log M$  time series. A short outline of the theory is presented in Appendix 3.

All these catalogues span about a year, and all contain between 300 and 3000 events. Care was taken to establish  $M_{min}$  for each region (i.e. the sensitivity of the relevant seismic network) and to subsequently filter out events below  $M_{min}$ .

The false nearest strands technique (fns) was used to assess the residual random component in a nonlinear time series.

## 3.2 Renormalisation group theory, inversion and neural network techniques

### 3.2.1 Improvement of phenomenological models by renormalisation group theory - larger seismic events as critical phenomena

The Gutenberg–Richter magnitude frequency distribution, described up to a roll-off point by

$$\log_{10} N_{>m_M} = a_{10} - b_{10} m_M \quad (3.2.1)$$

is an example of *scaling* in seismicity. Scale invariance is often a signature of criticality, and the fact that most large events are preceded by accelerating seismic release as described by the time to failure equation (3.1.1) is viewed as a sign that the critical state of the rock mass is about to be perturbed away from criticality. Recall that the Gutenberg-Richter relation (Figure A4.2.1 in Appendix 4) is a cumulative frequency distribution. The scaling region, where equation 3.2.1 applies, of the distribution implies that events of all sizes (up to roll-off) are possible, albeit with different probabilities. For some systems with phase spaces of dimensionality larger than one, such scaling is a result of criticality. Appendix 4 details some of the theory behind critical phenomena and its application to seismicity.

### 3.2.2 Inversion of the time to failure

The algorithm for this procedure is detailed in section 3.1.3.2, and the results appear in Chapter 4.

### 3.2.3 Neural networks

It was decided not to pursue this avenue of investigation, since the pattern recognition algorithm detailed in section 3.1.3.1, utilising non parametric statistics, combined with the learning strategy built into the time to failure algorithm was felt to be more appropriate.

## 3.3 Implementation

An algorithm (section 3.1.3.2) was developed and implemented. It was applied to several study sites, and the results are presented in Chapter 4.

## 4. Results

### 4.1 Pattern recognition: INDICATOR

Prediction is always based on a learning phase: decisions, such as which model to fit to a time to failure curve, have to be made. In the learning phase, experience of a given area is codified to assist in this decision. This is done in two phases: optimum time windows and time filters for computing *discrete* changes in 19 variables, as well as the occurrence of possible foreshocks, are determined. These time windows are then used to generate patterns.

The learning phase of a pattern recognition algorithm consists of the following:

- A dataset is split into two populations. In this case, catalogues are split into events of a given minimum local magnitude and above (population 'Yea') and a second population ('Nay') consisting of events of local magnitude less than the criterion above.
- A series of questions are asked about each event in the two populations; essentially, a questionnaire is constructed which asks questions, the answers to which are either "Yes" or "No". A "Yes" answer results in "1" being assigned to an event for a given factor, while a "No" results in a "-1". In this fashion, a bi-polar cipher can be built up which characterises an event. With enough events from a given area it may be possible to construct a cipher characterising the seismic rockmass response *at a given location*. Please refer to Appendix 1 for an example of the questionnaire.

#### 4.1.1 Output

Table 4.1.1 contains a sample of INDICATOR output for a 'Yea' population. In this table, eight datasets in a given area were considered. Without implying any chronological order, two events of local magnitude 2.0 or larger occurred in set 1, one of which was preceded by a local magnitude 1.7 event during the 150 hours preceding the event. There was only one local magnitude 2.0 or larger event in set 2, and, for example, three in dataset 5. Please note that this table only contains the 'Yea' population, consisting of 13 events in Area 1A for which the answer to the question, 'Did an event of  $M_L \geq 2.0$  happen?' was 'Yea!'. A similar table for events of magnitude  $M_L < 2.0$  is generated ('Nay') and the *same* 20 questions asked of the events in the 'Nay' population. Given that this population consisted of some 1800 events, it was not practical to reproduce the corresponding table here.

**Table 4.1.1**  
**Example: INDICATOR output**

Indica tor	$M_L \geq 1.7$	$\Delta V_A \geq 0.003$	$ \Delta \log EI  \geq 0.25$	$\Delta \log EI \leq -0.25$	$ \Delta \log EI_M  \geq 0.25$	$\Delta \log EI_M \leq -0.25$	Any $ \Delta \log EI  \geq 0.25$	Any $\Delta \log EI \leq -0.225$	Any $ \Delta \log EI_M  \geq 0.25$	Any $\Delta \log EI_M \leq -0.25$	$\Delta \log Sc_s \leq -1$	Any $\Delta \log Sc_s \leq -1$	$\Delta \epsilon_s > 0$	$\Delta \sigma_s < 0$	$\Delta X > 0$	$\Delta \bar{t} < 0$	Any $\Delta \epsilon_s > 0$	Any $\Delta \sigma_s < 0$	Any $\Delta X > 0$	Any $\Delta \bar{t} < 0$	Threshold	
Time before event	150h	78h	18h	18h	18h	18h	18h	18h	18h	18h	132h	132h	132h	132h	132h	132h	132h	132h	132h	132h		
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20		
Data Sets	1 1 2 3 4 4 5 5 6 6 7 8	1 -1 1 -1 -1 -1 -1 -1 -1 -1 -1 -1	1 -1 -1 -1 0 -1 -1 -1 -1 -1 -1 -1	-1 1 1 -1 1 -1 1 1 1 1 1 1	-1 1 -1 -1 1 -1 1 -1 -1 -1 -1 -1	-1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1	1 1 1 1 1 1 1 1 1 1 1 1	1 1 1 1 1 1 1 1 1 1 1 1	-1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1	-1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1	-1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1	-1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1	1 1 -1 1 1 1 1 1 1 1 1 1	1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1	1 1 1 1 -1 -1 -1 -1 -1 -1 -1 -1	1 1 1 1 1 1 1 1 1 1 1 1	1 1 1 1 1 1 1 1 1 1 1 1	1 1 1 1 1 1 1 1 1 1 1 1	1 1 1 1 1 1 1 1 1 1 1 1	1 1 1 1 1 1 1 1 1 1 1 1	1 1 1 1 1 1 1 1 1 1 1 1	8 9 7 7 8 9 7 8 9 8 12 9 10 3

Number of events of local magnitude  $M_L \geq 2$  : 13  
 Median Energy Index filter = 162h, 5 events.  
 Schmidt number filter = 30h, 0 events  
 Forecast point = 6min before event

## 4.1.2 Decision 1: Relevance of questions

It is not always necessary or desirable to ask as many questions as possible. To characterise an area, it is important to establish the correct questions. These questions comprise the 'kernel', and it is determined simply by considering the difference

$$K_{(i)} = \frac{n\left(\frac{i}{Y}\right)}{n(Y)} - \frac{n\left(\frac{i}{N}\right)}{n(N)} \quad (4.1.2.1)$$

where  $i = 1, 2, 3, \dots, 20$ ,  $n(Y)$  is the size of the 'Yea' population,  $n(i/Y)/n(Y)$  is the frequency of 'Yes' answers to question (i) in the 'Yea' population, and  $n(i/N)/n(N)$  are the same for the 'Nay' population.

For example, in Table 4.1.1, for question 1, asking about foreshocks,  $n(1/Y) = 4$  (four yes answers),  $n(Y) = 13$ , giving  $n(i/Y)/n(Y) = 0.3077$  or 30.77 per cent. For the 'Nay' population, i.e., for those events in Area 1A of  $M_L < 2.0$ ,  $n(1/N)/n(N) = 35.03$  per cent, i.e., 35.03 per cent events smaller than magnitude 2.0 were preceded (during 150 hours before the event) by an event of  $M_L \geq 1.7$  (Table 4.1.4.1(a)).

A question is in the kernel (i.e., it is relevant and can be used for recognition) if  $|K(i)| \geq k$ , that is,  $K(i) \geq k$  or  $K(i) \leq -k$  where  $k$  is a positive numerical threshold.

Moreover, if a question (i) does belong to the kernel, and  $K(i) \geq k$ , the answer to the question should be 'Yes' in the ideal 'Yea' population, while, if  $K(i) \leq -k$ , a 'Nay' answer is appropriate. Note that the kernel **quantifies** seismicity patterns in an area.

## 4.1.3 Decision 2: Hamming distance

In a forward-prediction scenario, the Hamming distance is a useful criterion for deciding whether the answers to the questions are indicating the imminence of an event or, alternatively, the onset of critical instability. The Hamming distance is chosen as criterion because of its applicability to associative memory in neural nets.

Recall that the answers to the diagnostic questions for a given event in the 'Yea' group will form a sequence of 1s, -1s and 0s. For example, if three questions were posed, the answers might be yes, yes, no or (1, 1, -1). Similarly, for an event in the 'Nay' group the cipher might be (-1, 1, -1). The Euclidean distance  $d_E$  between two vectors  $(x_1, x_2, x_3)$  and  $(w_1, w_2, w_3)$  is

$$d_E = \sqrt{(x_1 - w_1)^2 + (x_2 - w_2)^2 + (x_3 - w_3)^2} \quad (4.1.3.1)$$



Since  $x_i, w_i \in \{1, 0, -1\}$

$$\begin{aligned} (x_i - w_i)^2 &= 0 \text{ if } x_i = w_i \\ (x_i - w_i)^2 &= 4 \text{ if } x_i \neq w_i \text{ and } x_i \neq 0, w_i \neq 0 \\ (x_i - w_i)^2 &= 1 \text{ if either } x_i \text{ or } w_i = 0 \end{aligned} \quad (4.1.3.2)$$

Thus

$$d_E = \sqrt{4(\text{number of mismatched components} - \text{number of null components in either})} \quad (4.1.3.3)$$

The Hamming distance is defined as

$$h = \frac{d_E^2}{4} \quad (4.1.3.4)$$

In our example,  $(x_1, x_2, x_3) = (1, 1, -1)$  and  $(w_1, w_2, w_3) = (-1, 1, -1)$ . Thus

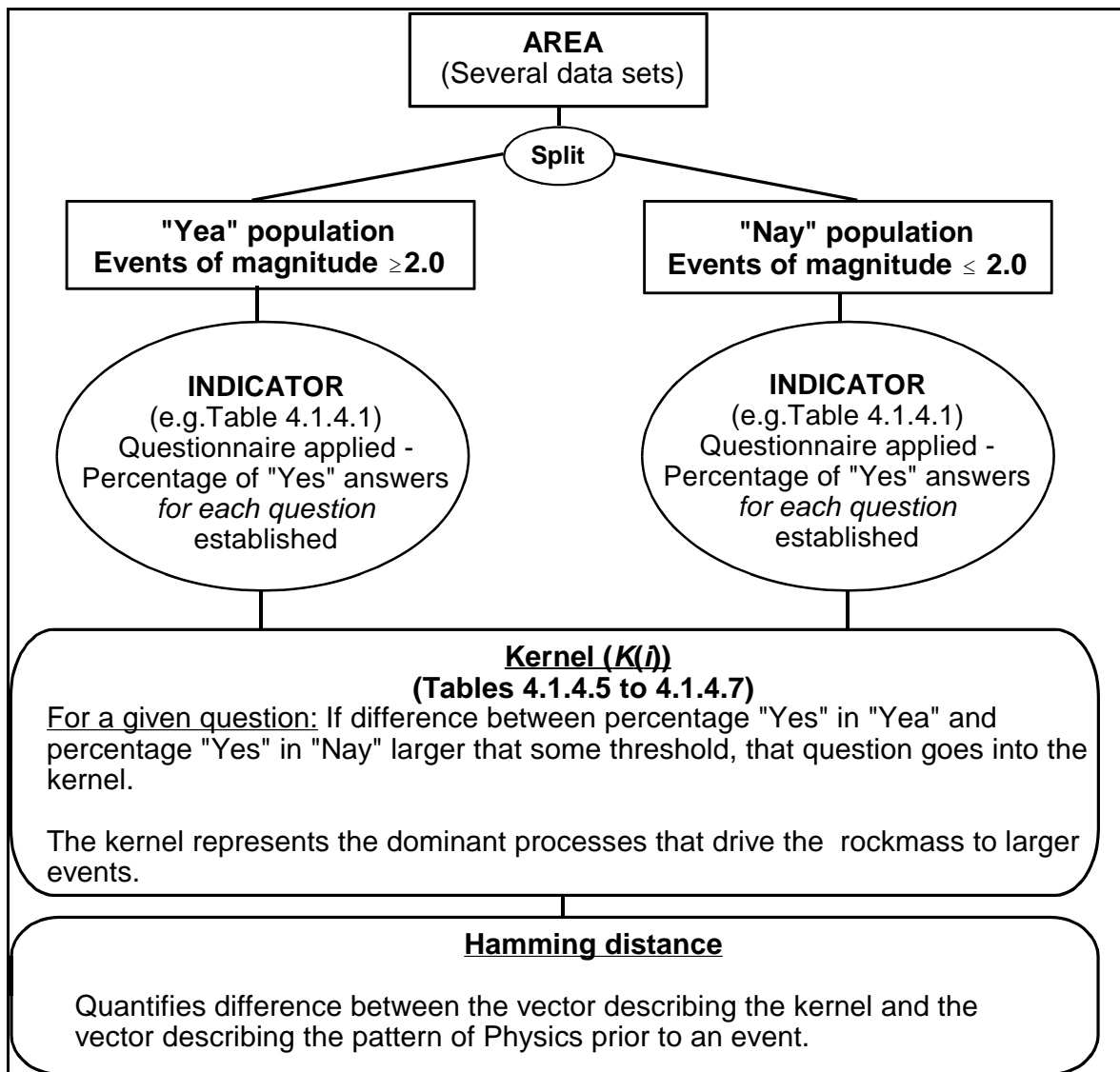
$$\begin{aligned} h &= \frac{d_E^2}{4} \\ &= (\text{number of mismatched components} - \text{null}) \\ &= 1 \end{aligned} \quad (4.1.3.5)$$

Note that

$$\begin{aligned} d_E &= \sqrt{(2)^2 + 0 + 0} = 2 \\ d_E^2 &= 4 \\ \therefore h &= 1 \end{aligned} \quad (4.1.3.6)$$

The procedure is then the following: For a given area, the questionnaire is applied, and for the thresholds used, the kernel established.

The kernel is of principal interest, since it is the kernel that enables the physics of a region to be established for the learning phase. It is also the kernel that determines which model is appropriate for a given region or mining method.



**Figure 4.1.3 Concept map**

## 4.1.4 Results

Tables 4.1.4.1 to 4.1.4.2 present the reduced output from INDICATOR i.e., the questionnaire reduced to the kernel. Note the following: for each dataset, two results are presented, being an analysis of patterns up to 6 minutes before each event in the set (Tables (a)) and an analysis up to 6 hours before each event (Tables (b)). These time intervals are called forecast points. For now, for the purpose of this section, note the *quantitative* difference between firstly, *areas* (area 1A comprises longwall mining workplaces, while 1C(a) and 1C(b) is a shaft pillar in the same geographical region) and secondly, *selecting data from the whole shaft pillar where it intersects the reef* (with 500m of the hanging wall and 1000m of the footwall included (Area 1C(a)) and selecting, carefully, two active areas in the pillar regions with the same hanging and foot walls (Area 1C(b)). These differences become even clearer upon examination of Tables 4.1.4.5 to 4.1.4.7, where the threshold for using a question for recognition has been set at  $|k| = 10\%$  (see section 4.1.2), thus filtering out only some diagnostic questions.

Example: Area 1A. Based on the threshold  $|k| = 10\%$ , questions 3, 4, 12, 13, 14, 15 and 16 are used for recognition, since, for them,  $k \geq 10\%$  or  $k \leq -10\%$  (Table 4.1.4.1(a)). These questions are shown in Tables 4.1.4.4 and 4.1.4.5(a).  $k(3)$ ,  $k(4)$ ,  $k(13)$ ,  $k(14)$ ,  $k(15)$  and  $k(16)$  are all larger than 10 per cent, and are thus each worth a '1' in  $K$ , the kernel.  $k(12) < -10\%$ , and thus leads to a '-1'. There are seven components to  $k$ , thus the maximum number of components that a pattern before an event can produce which are *different* from  $k$  is  $h_{max} = 7$ .

**Table 4.1.4.1(a)**  
**Reduced INDICATOR output for Area 1A, with the forecast point 6 minutes before an event**

Indicator	$M_L \geq 1.7$	$\Delta V_A \geq 0.003$	$ \Delta \log EI  \geq 0.25$	$\Delta \log EI \leq -0.25$	$ \Delta \log EI_M  \geq 0.25$	$\Delta \log EI_M \leq -0.25$	Any $ \Delta \log EI  \geq 0.25$	Any $\Delta \log EI \leq -0.25$	Any $ \Delta \log EI_M  \geq 0.25$	Any $\Delta \log EI_M \leq -0.25$	$\Delta \log Sc_s \leq -1$	Any $\Delta \log Sc_s \leq -1$	$\Delta \varepsilon_s > 0$	$\Delta \sigma_s < 0$	$\Delta \bar{X} > 0$	$\Delta \bar{t} < 0$	Any $\Delta \varepsilon_s > 0$	Any $\Delta \sigma_s < 0$	Any $\Delta \bar{X} > 0$	Any $\Delta \bar{t} < 0$
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Time before forecast point (hours)	150	78	18	18	18	18	18	18	18	18	132	132	132	132	132	132	132	132	132	132
'Yea': Percentage yes	30.77	7.69	76.92	61.54	7.69	7.69	100	100	7.69	7.69	7.69	69.23	84.62	46.15	69.23	84.62	100	100	100	100
'Nay': Percentage yes	35.03	3.19	55.57	28.04	4.92	3.02	99.27	94.52	6.10	3.75	8.23	82.82	70.62	27.87	51.09	59.32	100	99.83	99.94	100
K(i)	<b>-4.26</b>	<b>4.50</b>	<b>21.36</b>	<b>33.50</b>	<b>2.77</b>	<b>4.67</b>	<b>0.73</b>	<b>5.48</b>	<b>1.59</b>	<b>3.94</b>	<b>-0.53</b>	<b>-13.59</b>	<b>13.99</b>	<b>18.29</b>	<b>18.14</b>	<b>25.30</b>	<b>0.00</b>	<b>0.17</b>	<b>0.06</b>	<b>0.00</b>

Number of events of local magnitude  $M_L \geq 2$  : 13

Number of events of  $M_L$  smaller than 2: 1787

Median Energy Index filter = 162h, 5 events

Schmidt number filter = 30h, 0 events

Forecast point = 0.1h before event

**Table 4.1.4.1(b)**  
**Reduced INDICATOR output for Area 1A, with the forecast point 6 hours before an event**

Indicator	$M_L \geq 1.7$	$\Delta V_A \geq 0.003$	$ \Delta \log EI  \geq 0.25$	$\Delta \log EI \leq -0.25$	$ \Delta \log EI_M  \geq 0.25$	$\Delta \log EI_M \leq -0.25$	Any $ \Delta \log EI  \geq 0.25$	Any $\Delta \log EI \leq -0.25$	Any $ \Delta \log EI_M  \geq 0.25$	Any $\Delta \log EI_M \leq -0.25$	$\Delta \log Sc_s \leq -1$	Any $\Delta \log Sc_s \leq -1$	$\Delta \epsilon_s > 0$	$\Delta \sigma_s < 0$	$\Delta X > 0$	$\Delta \bar{t} < 0$	Any $\Delta \epsilon_s > 0$	Any $\Delta \sigma_s < 0$	Any $\Delta X > 0$	Any $\Delta \bar{t} < 0$
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Time before forecast point (hours)	144	168	114	114	114	114	114	114	114	114	144	144	144	144	144	144	144	144	144	144
'Yea': Percentage yes	23.08	7.69	76.92	61.54	15.38	15.38	100	100	23.08	23.08	7.69	76.92	84.62	38.46	53.85	69.23	100	100	100	100
'Nay': Percentage yes	33.18	6.32	61.44	38.11	10.24	6.66	100	99.50	14.61	10.52	13.37	85.17	59.65	29.77	51.15	54.73	100	99.89	99.94	100
K(i)	-10.11	1.37	15.48	23.43	5.14	8.73	0.00	0.50	8.47	12.56	-5.68	-8.25	24.96	8.69	2.70	14.50	0.00	0.11	0.06	0.00

Number of events of local magnitude  $M_L \geq 2$  : 13

Number of events of  $M_L$  smaller than 2: 1787

Median Energy Index filter = 162h, 5 events

Schmidt number filter = 30h, 0 events

Forecast point = 6h before event

**Table 4.1.4.2(a)**  
**Reduced INDICATOR output for Area 1C(a), with the forecast point 6 minutes before an event**

Indicator	$M_L \geq 1.7$	$\Delta V_A \geq 0.003$	$ \Delta \log EI  \geq 0.25$	$\Delta \log EI \leq -0.25$	$ \Delta \log EI_M  \geq 0.25$	$\Delta \log EI_M \leq -0.25$	Any $ \Delta \log EI  \geq 0.25$	Any $\Delta \log EI \leq -0.25$	Any $ \Delta \log EI_M  \geq 0.25$	Any $\Delta \log EI_M \leq -0.25$	$\Delta \log S_{c_s} \leq -1$	Any $\Delta \log S_{c_s} \leq -1$	$\Delta \epsilon_s > 0$	$\Delta \sigma_s < 0$	$\Delta X > 0$	$\Delta \bar{t} < 0$	Any $\Delta \epsilon_s > 0$	Any $\Delta \sigma_s < 0$	Any $\Delta X > 0$	Any $\Delta \bar{t} < 0$
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Time before forecast point (hours)	54	120	114	114	114	114	114	114	114	114	42	42	42	42	42	42	42	42	42	42
'Yea': Percentage yes	35.71	71.43	71.43	35.71	0.00	0.00	100	100	14.29	14.29	14.29	21.43	71.43	35.71	28.57	78.57	100	85.71	100	100
'Nay': Percentage yes	28.97	50.70	60.76	32.60	11.07	6.04	100	98.99	21.73	17.51	14.08	38.23	55.73	31.39	51.71	50.50	100	93.96	96.78	99.40
K(i)	<b>6.74</b>	<b>20.72</b>	<b>10.66</b>	<b>3.12</b>	<b>-11.07</b>	<b>-6.04</b>	<b>0.00</b>	<b>1.01</b>	<b>-7.44</b>	<b>-3.22</b>	<b>0.20</b>	<b>-16.80</b>	<b>15.69</b>	<b>4.33</b>	<b>-23.14</b>	<b>28.07</b>	<b>0.00</b>	<b>-8.25</b>	<b>3.22</b>	<b>0.60</b>

Number of events of local magnitude  $M_L \geq 2$  : 14

Number of events of  $M_L$  smaller than 2: 497

Median Energy Index filter = 48h, 5 events

Schmidt number filter = 42h, 0 events

Forecast point = 6min before event

**Table 4.1.4.2(b)**  
**Reduced INDICATOR output for Area 1C(a), with the forecast point 6 hours before an event**

Indicator	$M_L \geq 1.7$	$\Delta V_A \geq 0.003$	$ \Delta \log EI  \geq 0.25$	$\Delta \log EI \leq -0.25$	$ \Delta \log EI_M  \geq 0.25$	$\Delta \log EI_M \leq -0.25$	Any $ \Delta \log EI  \geq 0.25$	Any $\Delta \log EI \leq -0.25$	Any $ \Delta \log EI_M  \geq 0.25$	Any $\Delta \log EI_M \leq -0.25$	$\Delta \log Sc_s \leq -1$	Any $\Delta \log Sc_s \leq -1$	$\Delta \epsilon_s > 0$	$\Delta \sigma_s < 0$	$\Delta X > 0$	$\Delta \bar{t} < 0$	Any $\Delta \epsilon_s > 0$	Any $\Delta \sigma_s < 0$	Any $\Delta X > 0$	Any $\Delta \bar{t} < 0$
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Time before forecast point (hours)	54	120	150	150	150	150	150	150	150	150	78	78	78	78	78	78	78	78	78	78
'Yea': Percentage yes	35.71	71.43	78.57	35.71	0.00	0.00	100	100	28.57	28.57	14.29	28.57	64.29	57.14	28.57	57.14	100	100	100	100
'Nay': Percentage yes	28.97	48.89	60.60	36.22	9.46	4.23	100	99.4	25.15	20.32	19.92	36.62	56.34	37.22	54.33	53.32	99.80	98.49	97.18	99.40
K(i)	<b>6.74</b>	<b>22.54</b>	<b>11.97</b>	<b>-0.50</b>	<b>-9.49</b>	<b>-4.23</b>	<b>0.00</b>	<b>0.60</b>	<b>3.42</b>	<b>8.25</b>	<b>-5.63</b>	<b>-8.05</b>	<b>7.95</b>	<b>19.92</b>	<b>-25.75</b>	<b>3.82</b>	<b>0.20</b>	<b>1.01</b>	<b>2.82</b>	<b>0.60</b>

Number of events of local magnitude  $M_L \geq 2$  : 14

Number of events of  $M_L$  smaller than 2: 497

Median Energy Index filter = 48h, 5 events

Schmidt number filter = 72h, 0 events

Forecast point = 6h before event

**Table 4.1.4.3(a)**  
**Reduced INDICATOR output for Area 1C(b), with the forecast point 6 minutes before an event**

Indicator	$M_L \geq 1.7$	$\Delta V_A \geq 0.003$	$ \Delta \log EI  \geq 0.25$	$\Delta \log EI \leq -0.25$	$ \Delta \log EI_M  \geq 0.25$	$\Delta \log EI_M \leq -0.25$	Any $ \Delta \log EI  \geq 0.25$	Any $\Delta \log EI \leq -0.25$	Any $ \Delta \log EI_M  \geq 0.25$	Any $\Delta \log EI_M \leq -0.25$	$\Delta \log Sc_s \leq -1$	Any $\Delta \log Sc_s \leq -1$	$\Delta \epsilon_s > 0$	$\Delta \sigma_s < 0$	$\Delta X > 0$	$\Delta \bar{t} < 0$	Any $\Delta \epsilon_s > 0$	Any $\Delta \sigma_s < 0$	Any $\Delta X > 0$	Any $\Delta \bar{t} < 0$
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Time before forecast point (hours)	120	84	114	114	114	114	114	114	114	114	102	102	102	102	102	102	102	102	102	102
'Yea': Percentage yes	36.36	81.82	63.64	27.27	45.45	31.82	100	100	77.27	63.64	4.55	40.91	59.09	40.91	31.82	77.27	100	95.45	100	100
'Nay': Percentage yes	40.86	95.07	66.75	32.28	41.53	26.43	100	99.57	76.61	65.29	11.21	44.40	59.14	42.08	50.24	60.84	100	99.27	99.45	100
K(i)	<b>-4.50</b>	<b>-13.25</b>	<b>-3.11</b>	<b>-5.00</b>	<b>3.92</b>	<b>5.39</b>	<b>0.00</b>	<b>0.43</b>	<b>0.66</b>	<b>-1.65</b>	<b>-6.66</b>	<b>-3.49</b>	<b>-0.04</b>	<b>-1.17</b>	<b>-18.43</b>	<b>16.43</b>	<b>0.00</b>	<b>-3.81</b>	<b>0.55</b>	<b>0.00</b>

Number of events of local magnitude  $M_L \geq 2$  : 22

Number of events of  $M_L$  smaller than 2: 1642

Median Energy Index filter = 48h, 5 events

Schmidt number filter = 60h, 0 events

Forecast point = 6min before event



**Table 4.1.4.3(b)**  
**Reduced INDICATOR output for Area 1C(b), with the forecast point 6 hours before an event**

Indicator	$M_L \geq 1.7$	$\Delta V_A \geq 0.003$	$ \Delta \log EI  \geq 0.25$	$\Delta \log EI \leq -0.25$	$ \Delta \log EI_M  \geq 0.25$	$\Delta \log EI_M \leq -0.25$	Any $ \Delta \log EI  \geq 0.25$	Any $\Delta \log EI \leq -0.25$	Any $ \Delta \log EI_M  \geq 0.25$	Any $\Delta \log EI_M \leq -0.25$	$\Delta \log Sc_s \leq -1$	Any $\Delta \log Sc_s \leq -1$	$\Delta \epsilon_s > 0$	$\Delta \sigma_s < 0$	$\Delta X > 0$	$\Delta \bar{t} < 0$	Any $\Delta \epsilon_s > 0$	Any $\Delta \sigma_s < 0$	Any $\Delta X > 0$	Any $\Delta \bar{t} < 0$
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Time before forecast point (hours)	126	96	90	90	90	90	90	90	90	90	96	96	96	96	96	96	96	96	96	96
'Yea': Percentage yes	40.91	81.82	54.55	31.82	45.45	40.91	100	100	54.55	50.00	9.09	36.36	59.09	40.91	31.82	72.73	100	95.45	100	100
'Nay': Percentage yes	41.14	95.05	68.28	33.01	47.37	25.86	100	99.45	68.95	45.90	12.29	41.56	56.23	43.89	51.47	60.33	100	99.14	99.39	100
K(i)	<b>-0.23</b>	<b>-13.23</b>	<b>-13.73</b>	<b>-1.19</b>	<b>-1.92</b>	<b>15.05</b>	<b>0.00</b>	<b>0.55</b>	<b>-14.40</b>	<b>4.10</b>	<b>-3.20</b>	<b>-5.20</b>	<b>2.86</b>	<b>-2.98</b>	<b>-19.65</b>	<b>12.40</b>	<b>0.00</b>	<b>-3.69</b>	<b>0.61</b>	<b>0.00</b>

Number of events of local magnitude  $M_L \geq 2$  : 22

Number of events of  $M_L$  smaller than 2: 1636

Median Energy Index filter = 96h, 5 events

Schmidt number filter = 60h, 0 events

Forecast point = 6h before event

**Table 4.1.4.4**  
**Kernel output for Area 1A**

Indicator	$ \Delta \log EI  \geq 0.25$	$\Delta \log EI \leq -0.25$	Any $\Delta \log S_{C_s} \leq -1$	$\Delta s > 0$	$\Delta \sigma_s < 0$	$\Delta X > 0$	$\Delta \bar{t} < 0$	Hamming Distance
Time before event	18h	18h	132h	132h	132h	132h	132h	
	3	4	12	13	14	15	16	
1	-1	-1	-1	1	1	1	1	2
1	1	1	-1	1	-1	1	1	1
2	1	1	1	-1	1	1	-1	1
3	1	1	1	1	-1	1	1	1
4	1	-1	1	1	1	1	1	1
4	1	-1	1	1	-1	-1	1	2
5	1	1	1	1	1	1	1	0
5	1	1	1	1	-1	-1	1	1
5	1	1	1	1	1	-1	1	0
6	-1	-1	-1	1	-1	-1	1	3
6	1	1	-1	-1	-1	1	-1	2
7	-1	-1	1	1	-1	1	1	3
8	1	1	1	1	1	1	1	0

Number of events of local magnitude  $M_L \geq 2$  : 13  
Median Energy Index filter = 162h, 5 events.  
Schmidt number filter = 30h, 0 events  
Forecast point = 6min before event

**Table 4.1.4.5(a)**

**Kernel output for Area 1A, with the forecast point 6 minutes before an event**

Indicator	$ \Delta \log EI  \geq 0.25$	$\Delta \log EI \leq -0.25$	Any $\Delta \log Sc_s \leq -1$	$\Delta s > 0$	$\Delta \sigma_s < 0$	$\Delta X > 0$	$\Delta \bar{t} < 0$
	3	4	12	13	14	15	16
Time before forecast point (hours)	18	18	132	132	132	132	132
'Yea': Percentage yes	76.92	61.54	69.23	84.62	46.15	69.23	84.62
'Nay': Percentage yes	55.57	28.04	82.82	70.62	27.87	51.09	59.32
K(i)	<b>21.36</b>	<b>33.50</b>	<b>-13.59</b>	<b>13.99</b>	<b>18.29</b>	<b>18.14</b>	<b>25.30</b>

Number of events of local magnitude  $M_L \geq 2$  : 13

Number of events of  $M_L$  smaller than 2: 1787

Median Energy Index filter = 162h, 5 events

Schmidt number filter = 30h, 0 events

Forecast point = 6min before event

$K = (1, 1, -1, 1, 1, 1, 1)$

$\vec{h} = (2, 1, 1, 1, 1, 2, 0, 1, 0, 3, 2, 3, 0)$

$h_{MAX} = 7$

**Table 4.1.4.5(b)**  
**Kernel output for Area 1A, with the forecast point 6 hours before an event**

Indicator	$M_L \geq 1.7$	$ \Delta \log EI  \geq 0.25$	$\Delta \log EI \leq -0.25$	Any $\Delta \log EI_M \leq -0.25$	$\Delta s > 0$	$\Delta \bar{t} < 0$
	1	3	4	10	13	16
Time before forecast point (hours)	144	114	114	114	144	144
'Yea': Percentage yes	23.08	76.92	61.54	23.08	84.62	69.23
'Nay': Percentage yes	33.18	61.44	38.11	10.52	59.65	54.73
K(i)	<b>-10.11</b>	<b>15.48</b>	<b>23.43</b>	<b>12.56</b>	<b>24.96</b>	<b>14.50</b>

Number of events of local magnitude  $M_L \geq 2$  : 13

Number of events of  $M_L$  smaller than 2: 1787

Median Energy Index filter = 162h, 5 events

Schmidt number filter = 30h, 0 events

Forecast point = 6h before event

$\vec{K} = (-1, 1, 1, 1, 1, 1)$

$\vec{h} = (1, 1, 2, 3, 1, 3, 2, 1, 5, 1, 4, 2, 0)$

$h_{MAX} = 6$

**Table 4.1.4.6(a)**  
**Kernel output for Area 1C(a), with the forecast point 6 minutes before an event**

Indicator	$\Delta V_A \geq 0.003$	$ \Delta \log EI  \geq 0.25$	$ \Delta \log EI_M  \geq 0.25$	Any $\Delta \log Sc_s \leq -1$	$\Delta s > 0$	$\Delta X > 0$	$\Delta T < 0$
	2	3	5	12	13	15	16
Time before forecast point (hours)	120	114	114	42	42	42	42
'Yea': Percentage yes	71.43	71.43	0.00	21.43	71.43	28.57	78.57
'Nay': Percentage yes	50.70	60.76	11.07	38.23	55.73	51.71	50.50
K(i)	<b>20.72</b>	<b>10.66</b>	<b>-11.07</b>	<b>-16.80</b>	<b>15.69</b>	<b>-23.14</b>	<b>28.07</b>

Number of events of local magnitude  $M_L \geq 2$  : 14

Number of events of  $M_L$  smaller than 2: 497

Median Energy Index filter = 48h, 5 events

Schmidt number filter = 42h, 0 events

Forecast point = 6min before event

$K = (1, 1, -1, -1, 1, -1, 1)$

$\vec{h} = (2, 4, 0, 1, 3, 1, 3, 0, 0, 3, 2, 1, 1, 1)$

$h_{MAX} = 7$

**Table 4.1.4.6(b)**

**Kernel output for Area 1C(a), with the forecast point 6 hours before an event**

Indicator	$\Delta V_A \geq 0.003$	$ \Delta \log EI  \geq 0.25$	$\Delta \sigma_s < 0$	$\Delta X > 0$
	2	3	14	15
Time before forecast point (hours)	120	150	78	78
'Yea': Percentage yes	71.43	78.57	57.14	28.57
'Nay': Percentage yes	48.89	60.60	37.22	54.33
K(i)	<b>22.54</b>	<b>11.97</b>	<b>19.92</b>	<b>-25.75</b>

Number of events of local magnitude  $M_L \geq 2$  : 14

Number of events of  $M_L$  smaller than 2: 497

Median Energy Index filter = 48h, 5 events

Schmidt number filter = 72h, 0 events

Forecast point = 6h before event

$\vec{K} = (1, 1, 1, -1)$

$\vec{h} = (2, 1, 0, 1, 0, 2, 1, 1, 1, 1, 3, 1, 1, 2)$

$h_{MAX} = 4$

**Table 4.1.4.7(a)**  
**Kernel output for Area 1C(b), with the forecast point 6 minutes before an event**

Indicator	$\Delta V_A \geq 0.003$	$\Delta X > 0$	$\Delta \bar{t} < 0$
	2	15	16
Time before forecast point (hours)	84	102	102
'Yea': Percentage yes	81.82	31.82	77.27
'Nay': Percentage yes	95.07	50.24	60.84
K(i)	<b>-13.25</b>	<b>-18.43</b>	<b>16.43</b>

Number of events of local magnitude  $M_L \geq 2$  : 22

Number of events of  $M_L$  smaller than 2: 1642

Median Energy Index filter = 48h, 5 events

Schmidt number filter = 60h, 0 events

Forecast point = 6min before event

$K = (-1, -1, 1)$

$\vec{h} = (1, 0, 1, 1, 2, 2, 1, 2, 1, 1, 1, 2, 1, 1, 2, 1, 2, 2, 1, 2, 2)$

$h_{MAX} = 3$

**Table 4.1.4.7(b)**  
**Kernel output for Area 1C(b), with the forecast point 6 hours before an event**

Indicator	$\Delta V_A \geq 0.003$	$ \Delta \log EI  \geq 0.25$	$\Delta \log EI_M \leq -0.25$	Any $ \Delta \log EI_M  \geq 0.25$	$\Delta X > 0$	$\Delta f < 0$
	2	3	6	9	15	16
Time before forecast point (hours)	96	90	90	90	96	96
'Yea': Percentage yes	81.82	54.55	40.91	54.55	31.82	72.73
'Nay': Percentage yes	95.05	68.28	25.86	68.95	51.47	60.33
K(i)	<b>-13.23</b>	<b>-13.73</b>	<b>15.05</b>	<b>-14.40</b>	<b>-19.65</b>	<b>12.40</b>

Number of events of local magnitude  $M_L \geq 2$  : 22

Number of events of  $M_L$  smaller than 2: 1636

Median Energy Index filter = 96h, 5 events

Schmidt number filter = 60h, 0 events

Forecast point = 6h before event

$\vec{K} = (-1, -1, 1, -1, -1, 1)$

$\vec{h} = (3, 1, 2, 3, 3, 3, 4, 4, 3, 3, 2, 2, 4, 2, 2, 4, 3, 3, 5, 3, 4, 5)$

$h_{MAX} = 6$



Special note should be taken of the Hamming distance for each of the events ( $h$ ), which is always less than  $h_{max}$  for the 'Yea' population. Concerning the questions in the kernel, note, in the  $K(i)$  row, that the predominant feature of the premonitory pattern is one of softening in Area 1A (Tables 4.1.4.5(a) and (b)) as encoded by the energy index variables. Clearly, in this area, increasing seismic strain and increasing activity rate ( $\Delta\dot{t}$ ) are strong members of the kernel.

This is to be contrasted with Tables 4.1.4.6(a) and (b), the whole of the shaft pillar in a neighbouring mine where the predominant mechanism at the levels of thresholds used in this case, is increasing coseismic deformation ( $\Delta V_A$ ) and, again, on a shorter term forecast, increasing activity ( $\Delta\dot{t}$ ) in Table 4.1.4.6(a)).

Comparing Tables 4.1.4.6 and 4.1.4.7 highlights *yet again the fundamental role played by the selection of the volume of interest*. Remember that these two cases are the *same pillar, for the same span*, but Table 4.1.4.7 represents the pattern of response when the pillar is looked at as two, separated, datasets.

Interestingly, although negative  $K(i)$ s are a little counterintuitive, there is more physical information in the longer-term (6 hour) forecast than in the 6 minute one (note questions 3, 6 and 4).

In these tables,  $h_{max}$  is the maximum possible Hamming distance (the maximum possible distance between a vector of ones, minus ones and zeroes and the representative kernel vector  $K$ ).

## 4.2 Deterministic components of time series, denoising and limits of predictability

After applying the false nearest strands (fns) technique to a number of areas, the random component - as measured by the fns - ranged between 39 per cent and 44 per cent which in turn indicates that it is likely that the deterministic component is not above 61 per cent. To deal with this apparent problem, a noise reduction in the signals of interest which would preserve as much as possible of the information content of the original signal, could enhance the deterministic component. Hence the current prediction techniques (e.g. time to failure) may become more successful if applied to the "denoised" signals instead of the original ones.

Tables 4.2.1 and 4.2.2 summarize the results obtained after applying the above denoising technique to the time series of  $\log E$  and  $\log M$ . Please note that the areas are subsets of Area 3. In each case the random component (as measured by the false nearest strands) has decreased significantly after the noise reduction technique was applied. Results are shown for an increasing number of minimum neighbours contained in a hypersphere of radius 0.01 the size of the attractor. The random component decreases in magnitude as the minimum number approaches five but resulting signals differ more and more from the initial signal. This behaviour was noticed by phase plotting the cross correlation function of the original signal  $X(t)$  and itself versus the correlation function of the original signal with its denoised counterparts. An increase in the spreading is noticed as the minimum number of neighbours is increased (see Figure 4.2.1).

Thus there appears to be an upper limit to how much one may wish to denoise the original time series without destroying the underlying structure. In this case it appears that according to the examples analysed one should not count on denoised signals where the minimum

number of neighbours used to perform the SVD-noise reduction exceeds four. Nonetheless, if one considers Tables 4.2.1(d) and 4.2.2(d), "Noise reduction with four neighbours", the denoised signal exhibits a level of random component between 21 per cent and 33 per cent (in the case of the moment analysis) and 21 per cent and 40 per cent (in the case of the energy analysis). If one excludes area 7 in the energy analysis the above figures become 21 per cent to 30 per cent. Therefore a success rate of up to 79 per cent *could* in principle be attained if one uses deterministic techniques on parameters derived using the denoised moment and energy time series.

In Figures 4.2.3 and 4.2.4 the effect of SVD noise reduction on the spatio-temporal limits of predictability and deterministic content for  $LogE$  and  $LogM$  is displayed.

**Table 4.2.1**  
**Noise reduction in seismic time series: logE analysis**

area	(a)					(b)					(c)				
	Before noise reduction					Noise reduction with 2 neighbours $r = 0.01$					Noise reduction with 3 neighbours $r = 0.01$				
	Delay	Embedding dimension	Lyapunov exponent	$d_A$	Noise level (%)	Delay	Embedding dimension	Lyapunov exponent	$d_A$	Noise level (%)	Delay	Embedding dimension	Lyapunov exponent	$d_A$	Noise level (%)
area 1a	23	10	2.07	7.86	<b>43</b>	27	8	0.47	5.88	<b>40</b>	1	7	0.48	4.73	<b>31</b>
area 2a	33	9	2.73	5.78	<b>41</b>	36	8	3.22	4.92	<b>39</b>	1	7	0.91	4.39	<b>33</b>
area 3	30	10	2.91	7.36	<b>41</b>	33	10	3.21	5.33	<b>36</b>	1	7	0.74	4.39	<b>26</b>
area 4a	31	10	6.16	7.86	<b>45</b>	1	9	1.41	6.35	<b>41</b>	1	9	1.15	5.75	<b>31</b>
area 4b	27	9	8.06	4.74	<b>41</b>	1	7	1.49	3.93	<b>38</b>	1	7	1.83	3.24	<b>31</b>
area 5b	36	10	3.63	7.87	<b>41</b>	37	9	3.82	5.72	<b>36</b>	1	8	0.99	4.52	<b>29</b>
area 6b	23	10	3.70	7.22	<b>42</b>	1	9	0.63	5.96	<b>39</b>	1	9	0.70	4.98	<b>30</b>
area 7	30	9	2.75	8.22	<b>41</b>	1	8	0.77	5.58	<b>41</b>	1	6	0.81	4.71	<b>39</b>
area 8	32	9	3.37	6.35	<b>42</b>	1	8	0.77	5.84	<b>40</b>	1	8	0.79	5.36	<b>31</b>
area 9a	30	9	2.80	6.50	<b>44</b>	32	8	2.65	6.05	<b>43</b>	1	8	0.44	5.31	<b>35</b>
area 9b	29	9	n/a	5.52	<b>43</b>	1	7	0.45	4.71	<b>40</b>	1	7	0.53	4.46	<b>28</b>

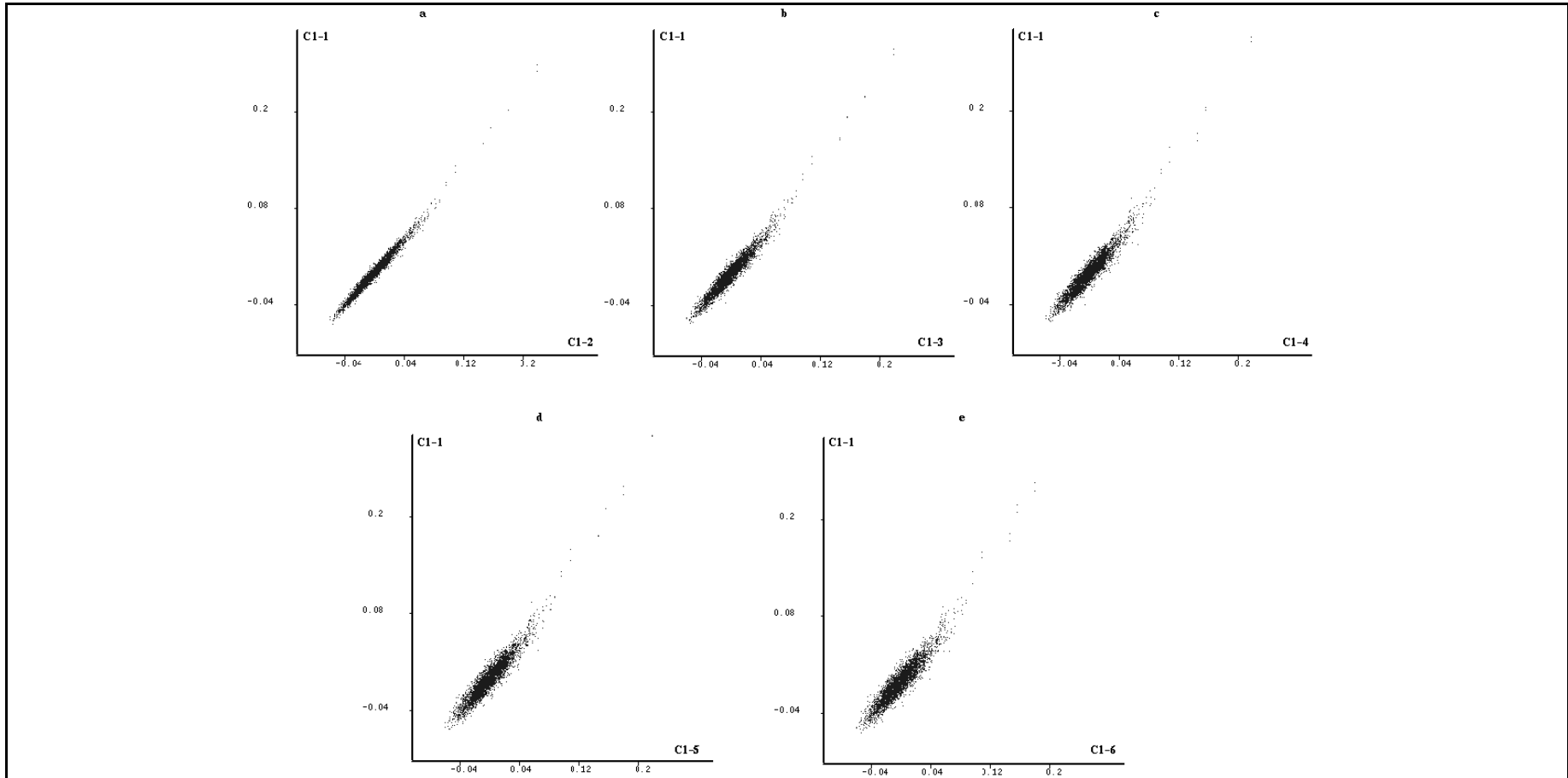
area	(d)					(e)					(f)				
	Noise reduction with 4 neighbours $r = 0.01$					Noise reduction with 5 neighbours $r = 0.01$					Noise reduction with 6 neighbours $r = 0.01$				
	Delay	Embedding dimension	Lyapunov exponent	$d_A$	Noise level (%)	Delay	Embedding dimension	Lyapunov exponent	$d_A$	Noise level (%)	Delay	Embedding dimension	Lyapunov exponent	$d_A$	Noise level (%)
area 1a	1	5	0.51	3.79	<b>27</b>	1	5	0.47	3.32	<b>23</b>	1	5	0.55	3.30	<b>17</b>
area 2a	1	7	1.81	4.28	<b>30</b>	1	7	0.77	3.98	<b>26</b>	1	7	0.85	3.79	<b>23</b>
area 3	1	7	0.70	3.83	<b>21</b>	1	6	0.57	3.29	<b>16</b>	1	6	0.80	3.23	<b>16</b>
area 4a	1	7	1.29	5.07	<b>30</b>	1	7	1.31	4.67	<b>26</b>	1	7	1.30	4.45	<b>26</b>
area 4b	1	7	1.02	2.96	<b>26</b>	1	6	1.61	3.00	<b>26</b>	1	5	0.94	2.88	<b>23</b>
area 5b	1	7	0.76	3.78	<b>22</b>	1	5	0.83	3.51	<b>20</b>	1	5	0.87	3.29	<b>16</b>
area 6b	1	7	0.80	4.29	<b>24</b>	1	6	0.90	4.10	<b>21</b>	1	6	0.98	4.09	<b>18</b>
area 7	1	6	0.78	4.19	<b>40</b>	1	6	0.45	3.95	<b>21</b>	1	6	0.82	4.25	<b>20</b>
area 8	1	8	1.16	6.05	<b>30</b>	1	8	1.24	6.16	<b>24</b>	1	8	1.46	6.28	<b>24</b>
area 9a	1	8	0.69	4.84	<b>25</b>	1	6	0.71	4.51	<b>20</b>	1	6	0.89	4.40	<b>22</b>
area 9b	1	7	0.51	4.28	<b>28</b>	1	6	0.55	4.10	<b>22</b>	1	6	3.85	3.85	<b>22</b>

**Table 4.2.2**  
**Noise reduction in seismic time series: logM analysis**

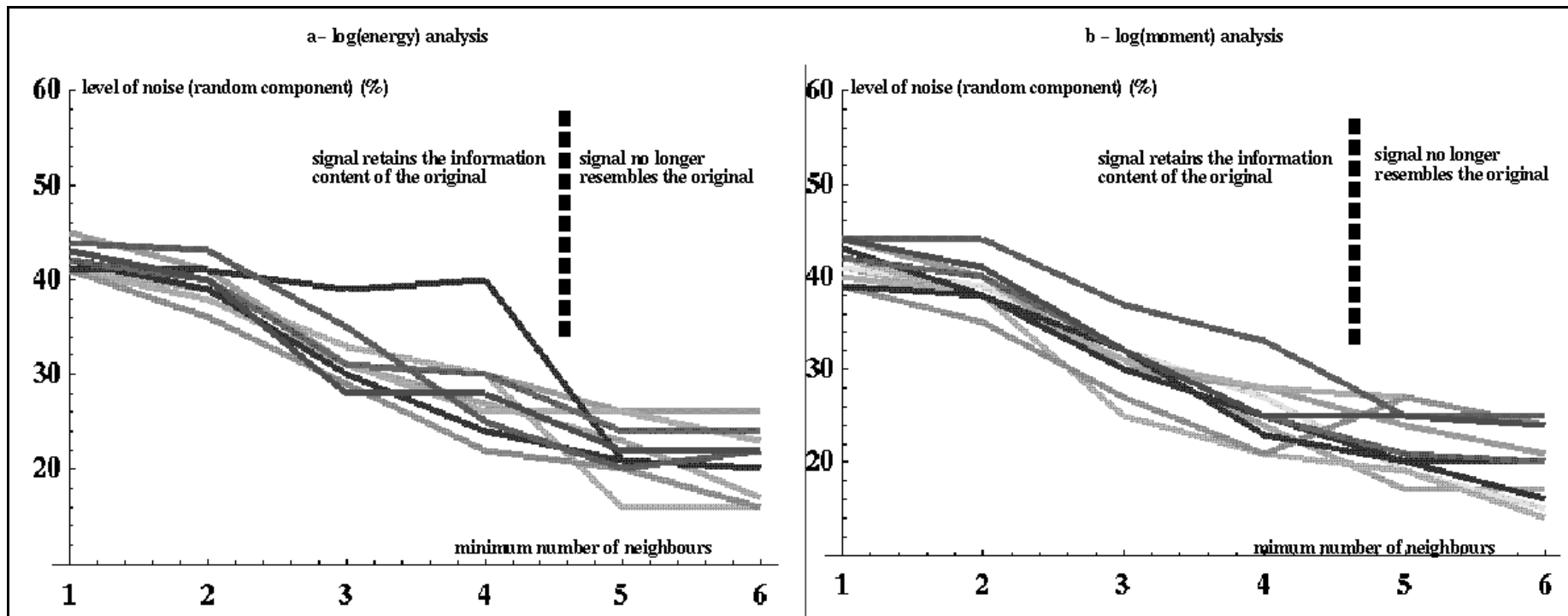
area	<b>(a)</b> Before noise reduction					<b>(b)</b> Noise reduction with 2 neighbours r = 0.01					<b>(c)</b> Noise reduction with 3 neighbours r = 0.01				
	Delay	Embedding dimension	Lyapunov exponent	$d_A$	Noise level (%)	Delay	Embedding dimension	Lyapunov exponent	$d_A$	Noise level (%)	Delay	Embedding dimension	Lyapunov exponent	$d_A$	Noise level (%)
area 1a	18	9	2.05	6.56	<b>41</b>	18	9	2.18	5.41	<b>38</b>	1	7	0.46	4.50	<b>31</b>
area 2a	18	8	2.81	5.81	<b>41</b>	16	7	3.29	4.51	<b>39</b>	1	7	0.62	4.23	<b>32</b>
area 3	39	9	2.63	7.27	<b>42</b>	35	9	2.87	5.58	<b>38</b>	1	7	0.62	4.83	<b>25</b>
area 4a	24	10	6.51	7.07	<b>44</b>	1	9	1.22	5.89	<b>40</b>	1	9	1.27	5.25	<b>30</b>
area 4b	32	6	9.63	3.40	<b>40</b>	1	5	1.53	3.24	<b>38</b>	1	5	1.89	3.07	<b>31</b>
area 5b	32	9	4.06	6.73	<b>39</b>	32	9	4.08	5.42	<b>35</b>	1	8	0.84	4.70	<b>27</b>
area 6b	18	10	3.74	7.70	<b>43</b>	1	8	0.71	5.49	<b>38</b>	1	8	0.75	4.69	<b>30</b>
area 7	40	10	2.17	7.23	<b>39</b>	1	9	0.47	5.33	<b>38</b>	1	7	0.50	4.56	<b>32</b>
area 8	29	9	3.42	6.62	<b>42</b>	1	9	0.73	5.56	<b>40</b>	1	8	0.55	4.74	<b>32</b>
area 9a	15	9	2.87	6.51	<b>44</b>	1	9	0.60	6.00	<b>44</b>	1	8	0.78	5.34	<b>37</b>
area 9b	22	7	1.76	4.94	<b>44</b>	1	7	0.52	4.71	<b>41</b>	1	6	0.63	4.53	<b>32</b>

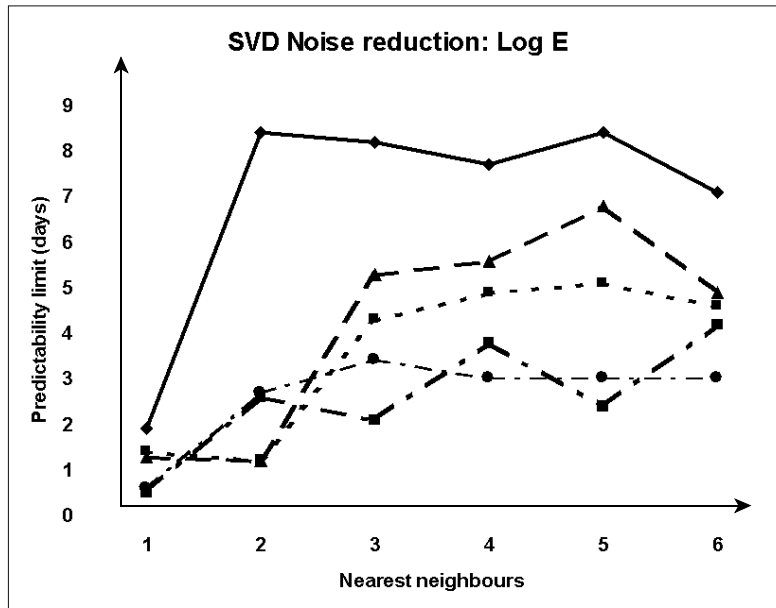
area	<b>(d)</b> Noise reduction with 4 neighbours r = 0.01					<b>(e)</b> Noise reduction with 5 neighbours r = 0.01					<b>(f)</b> Noise reduction with 6 neighbours r = 0.01				
	Delay	Embedding dimension	Lyapunov exponent	$d_A$	Noise level (%)	Delay	Embedding dimension	Lyapunov exponent	$d_A$	Noise level (%)	Delay	Embedding dimension	Lyapunov exponent	$d_A$	Noise level (%)
area 1a	1	5	0.47	3.80	<b>24</b>	1	5	0.43	3.33	<b>17</b>	1	5	0.52	3.37	<b>17</b>
area 2a	1	7	0.73	3.98	<b>27</b>	1	7	0.55	3.94	<b>19</b>	1	6	0.85	3.51	<b>15</b>
area 3	1	7	0.76	4.03	<b>21</b>	1	7	0.58	3.68	<b>19</b>	1	6	0.49	3.34	<b>14</b>
area 4a	1	8	1.19	4.77	<b>28</b>	1	8	1.25	4.61	<b>24</b>	1	8	1.24	4.31	<b>21</b>
area 4b	1	5	2.03	2.92	<b>28</b>	1	5	2.41	2.84	<b>27</b>	1	5	2.19	2.86	<b>24</b>
area 5b	1	8	0.95	4.61	<b>21</b>	1	8	0.97	4.89	<b>18</b>	1	8	0.91	4.77	<b>18</b>
area 6b	1	8	0.93	4.10	<b>25</b>	1	7	0.91	3.89	<b>20</b>	1	7	0.78	3.81	<b>16</b>
area 7	1	7	0.57	3.92	<b>23</b>	1	6	0.70	3.32	<b>20</b>	1	6	0.71	3.25	<b>20</b>
area 8	1	7	0.68	4.61	<b>25</b>	1	7	0.64	4.14	<b>21</b>	1	7	0.72	4.06	<b>20</b>
area 9a	1	8	0.75	5.09	<b>33</b>	1	7	0.74	4.78	<b>25</b>	1	7	0.71	4.59	<b>25</b>
area 9b	1	6	0.69	4.26	<b>25</b>	1	6	0.68	4.28	<b>25</b>	1	6	0.81	4.02	<b>24</b>



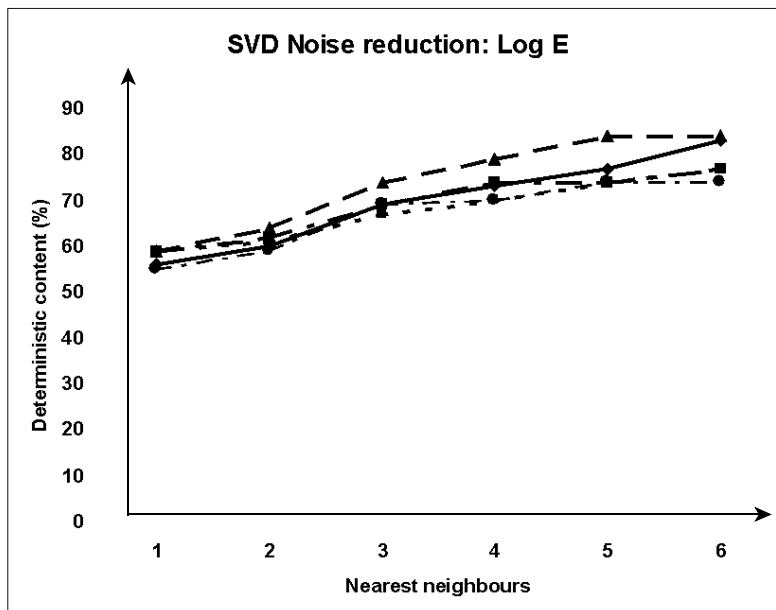
**Figure 4.2.1a-e** An example of phase plots of cross correlation functions log moment analysis for Area 3 (4a).  $C1-i$  (horizontal axis) is the cross correlation function between the original time series and itself.  $C1-j$  (vertical axis) is the cross correlation function between the original signal and the denoised signal where the SVD required at least  $j$  neighbours in hyperspheres around states in order for it to be applied. As the minimum requirement approaches five neighbours the spreading becomes significant and the denoised signal no longer retains enough of the original signal's information content.



**Figure 4.2.2a-b** The level of the random component versus the minimum number of neighbours required to perform the SVD reduction. Different lines correspond to the 11 different areas where the analysis was performed. Notice that the gain in noise reduction is predominant when the number of neighbours is below 4 while virtually negligible above 4-5. This is also the limit beyond which the signal no longer retains its original information content.

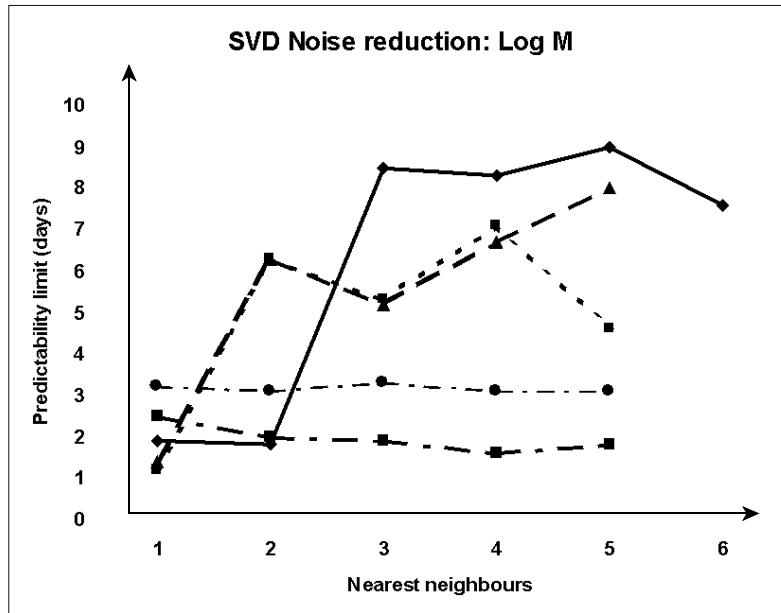


(a)

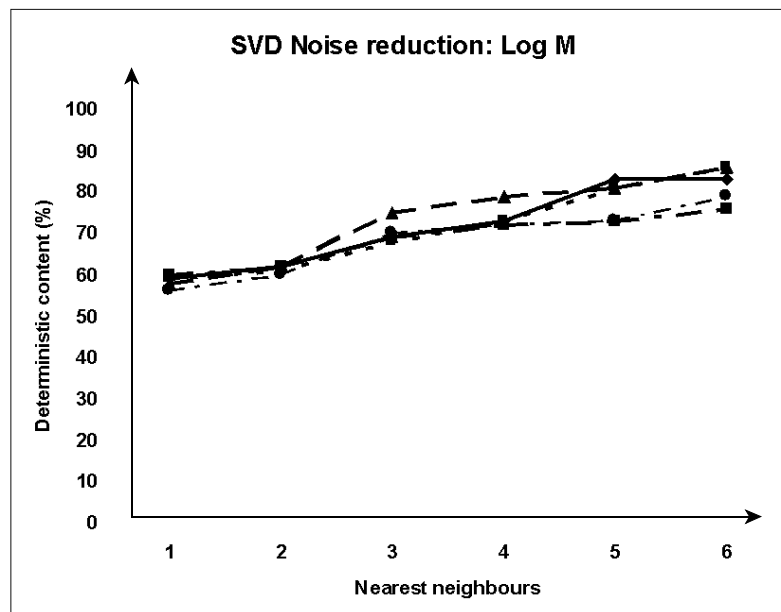


(b)

**Figure 4.2.3** The effect of SVD noise reduction on (a) limits of predictability and (b) the deterministic content of time series of logE for five data sets on the West Rand. For two of the sets the results for (b) essentially coincide.



(a)



(b)

**Figure 4.2.4** The effect of SVD noise reduction on (a) limits of predictability and (b) the deterministic content of time series of  $\log M$  for five data sets on the West Rand. For two of the sets the results for (b) essentially coincide.



## 4.3 Power law behaviour

The first order of business towards applying the time to failure algorithm was to determine to which extent acceleration in various measures of seismic release is observed. We present here the results of two experiments: First, a primitive method for determining the size of the critical volume (a measure of the correlation length) and the time window over which events are correlated is presented. Secondly, a volume constructed by a mine seismologist on the basis of clustering is analysed. In other words, in the second case, the volume is not optimised directly.

### 4.3.1 Critical volumes and time windows

A catalogue consisting of some 30 000 events ( $m_{min} = -1$ ) spanning about a year was analysed. The data comes from a West Rand mine and the first nine events with  $\log_{10} M \geq 12$  are studied here.

The procedure is as follows: In the spirit of back-analysis, the time of failure,  $t_c$  is known. The behaviour of two variables, the cumulative square roots of seismic energy release and seismic moment, is studied, as well as the cumulative apparent volume. The square roots of energy and moment are chosen since this minimises the difference between "hard" and "soft" events and also between events with drastically different amounts of coseismic deformation, leading to smoother variations. Cumulative quantities are preferred at this stage since they give an immediately easy way to extract significantly scale-invariant critical behaviour and a stationary state.

The quantities  $\sum_{t'=t_0}^t \sqrt{E_{t'}}$ ,  $\sum_{t'=t_0}^t \sqrt{M_{t'}}$  and  $\sum_{t'=t_0}^t V_{A_{t'}}$  are computed.

They are each fitted to the power law,

$$\tilde{S}_\rho(t) = C(t_c - t)^\alpha$$

and a linear hypothesis,

$$\tilde{S}_l(t) = C + Bt$$

The empirical variances

$$V_\rho = \sum_{k=1}^c \left[ \tilde{S}_\rho(t_k) - \sum_{t'=t_0}^{t_k} \Omega_{t'} \right]^2$$

and

$$V_l = \sum_{k=1}^c \left[ \tilde{S}_l(t_k) - \sum_{t'=t_0}^{t_k} \Omega_{t'} \right]^2$$

(minimised in the fits) are compared. Here,

$$\Omega_{t'} = \sqrt{E_{t'}}, \sqrt{M_{t'}} \text{ or } V_{A_{t'}}.$$

$V_{A_{t'}}$  is the apparent volume of an event occurring at time  $t'$ . Note that a linear hypothesis is an extreme case in which each event is the same as the one before.

Computing  $r = V_p / V_l$  gives a measure of “how much more (or less)” critical than stationary the volume is.

#### **4.3.1.1 Data**

Figures 4.3.1.1(a), (b), (c), (d) and (e) present a visual representation of the “critical” events, and the clustering of events around long walls. Of interest are Figures 4.3.1.1(c) and 4.3.1.1(d). Notice the dome-shaped distribution of events. This is probably an artifact of the seismic network geometry, perforce constrained to reflect the largely two-dimensional, tabular, geometry of the reef.

For this reason, the determination of critical volumes and time intervals  $t_c - t_0$  was performed without regard to the z-locations of events (i.e. in two dimensions).

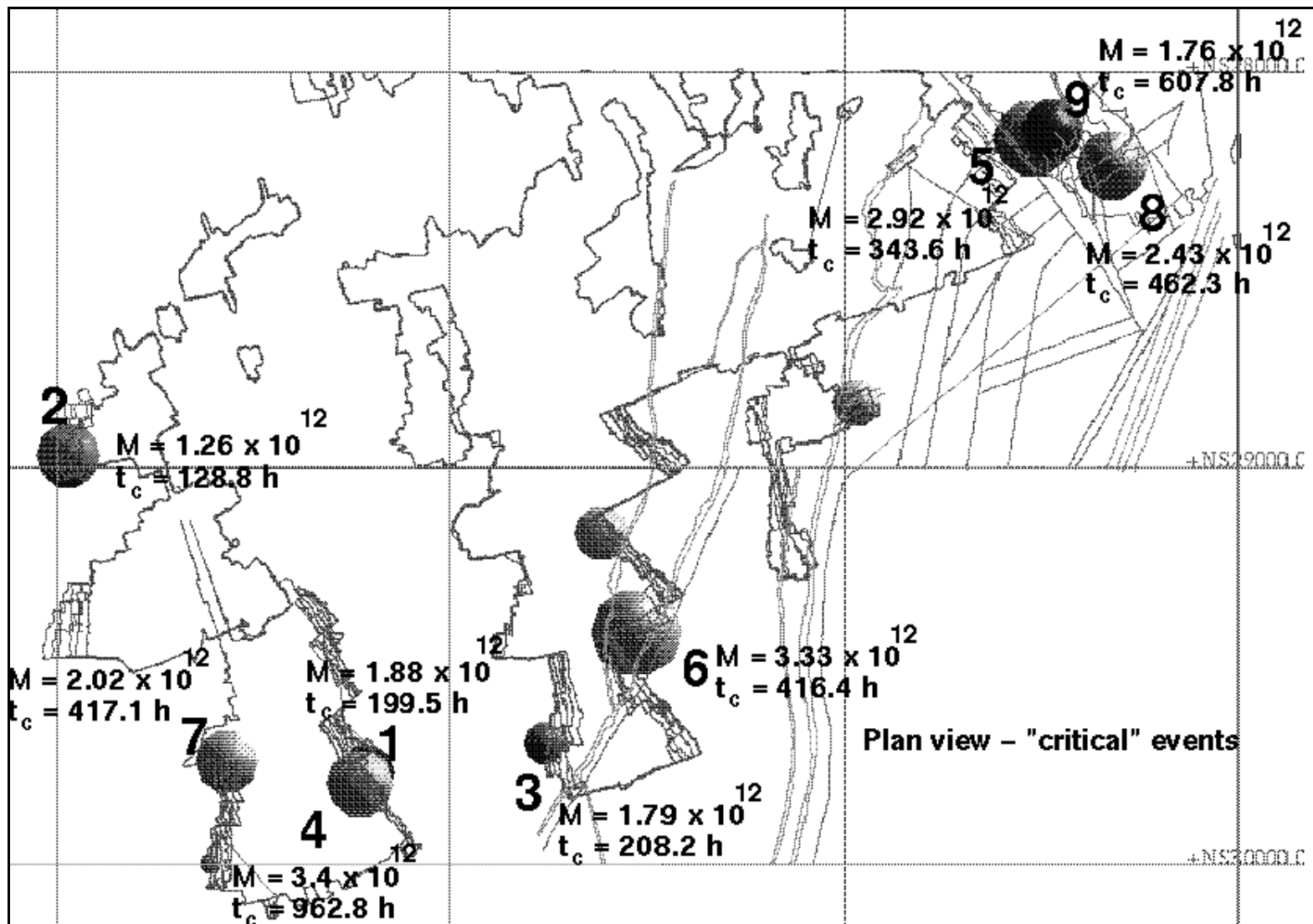


Figure 4.3.1.1(a)

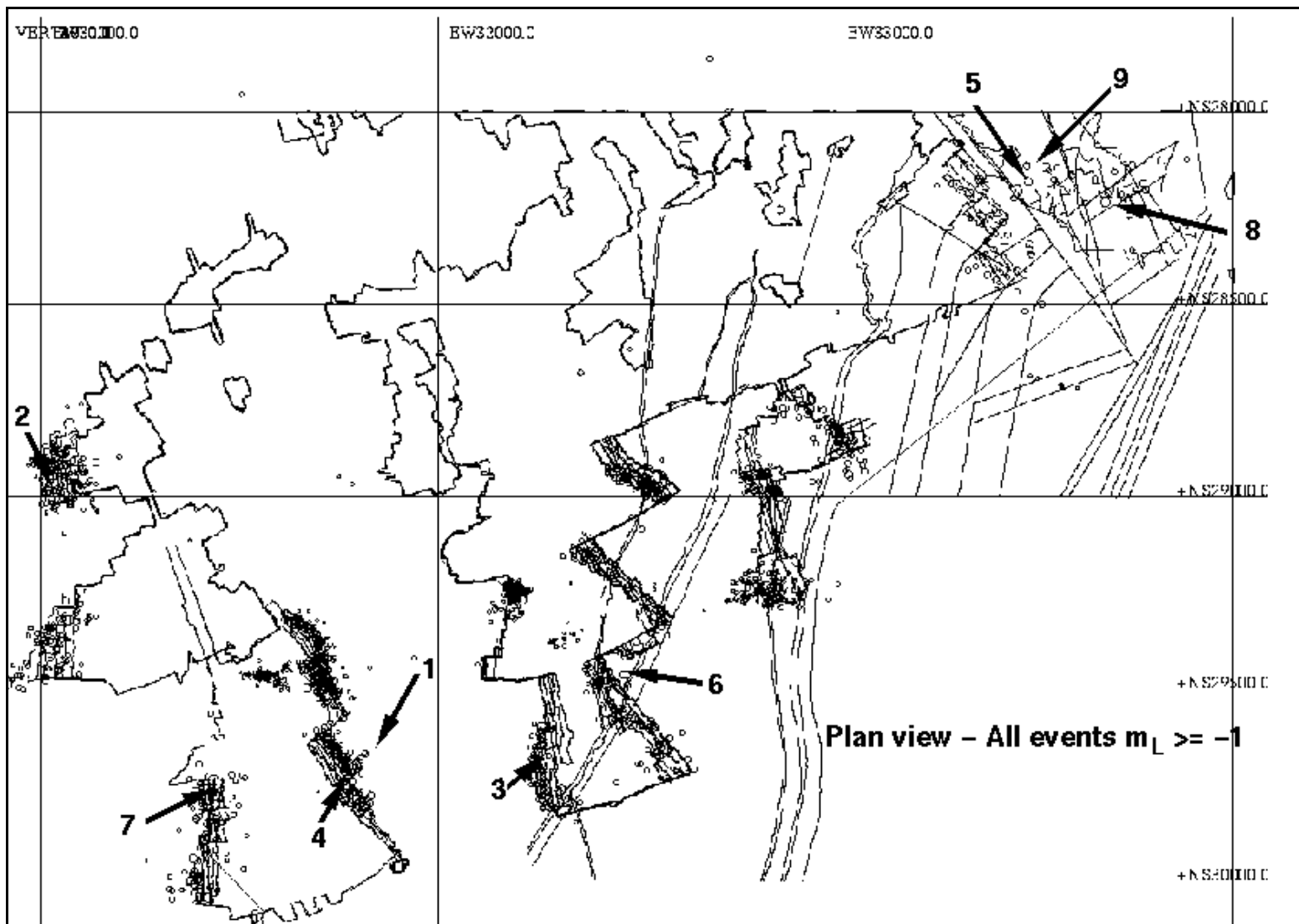
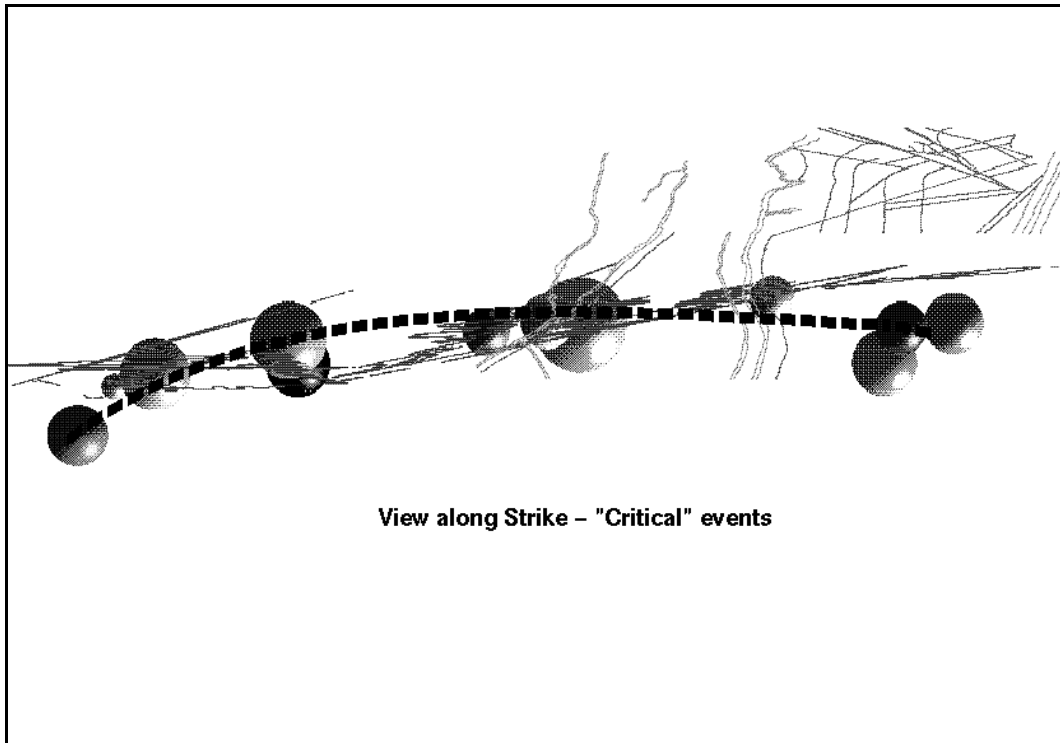
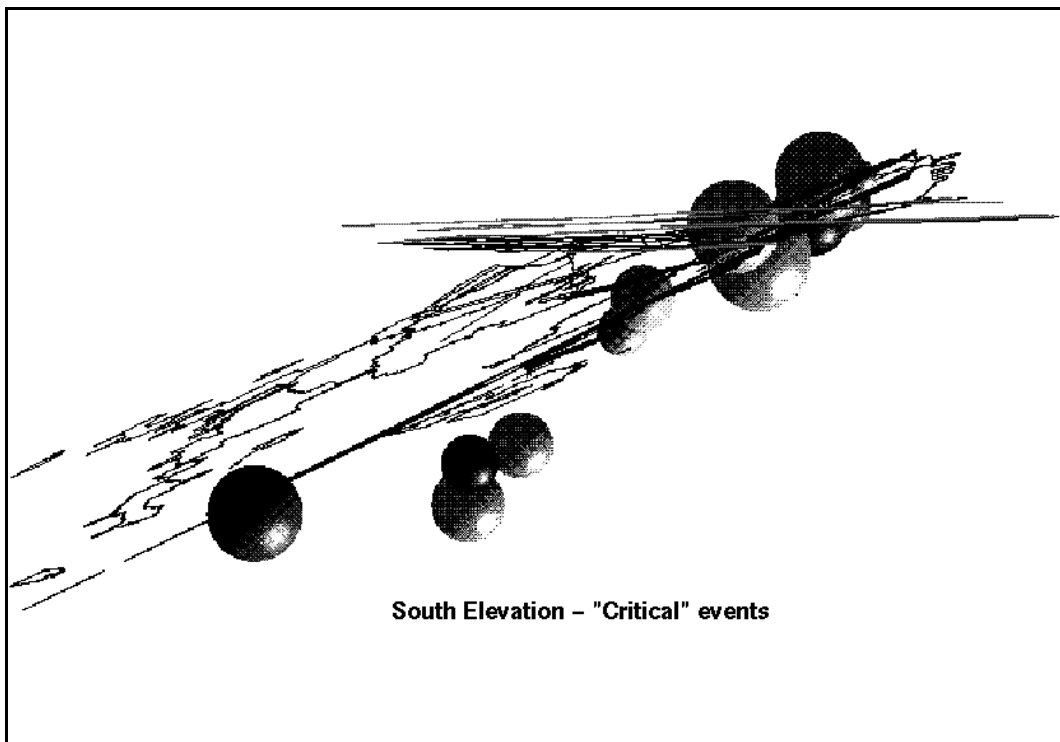


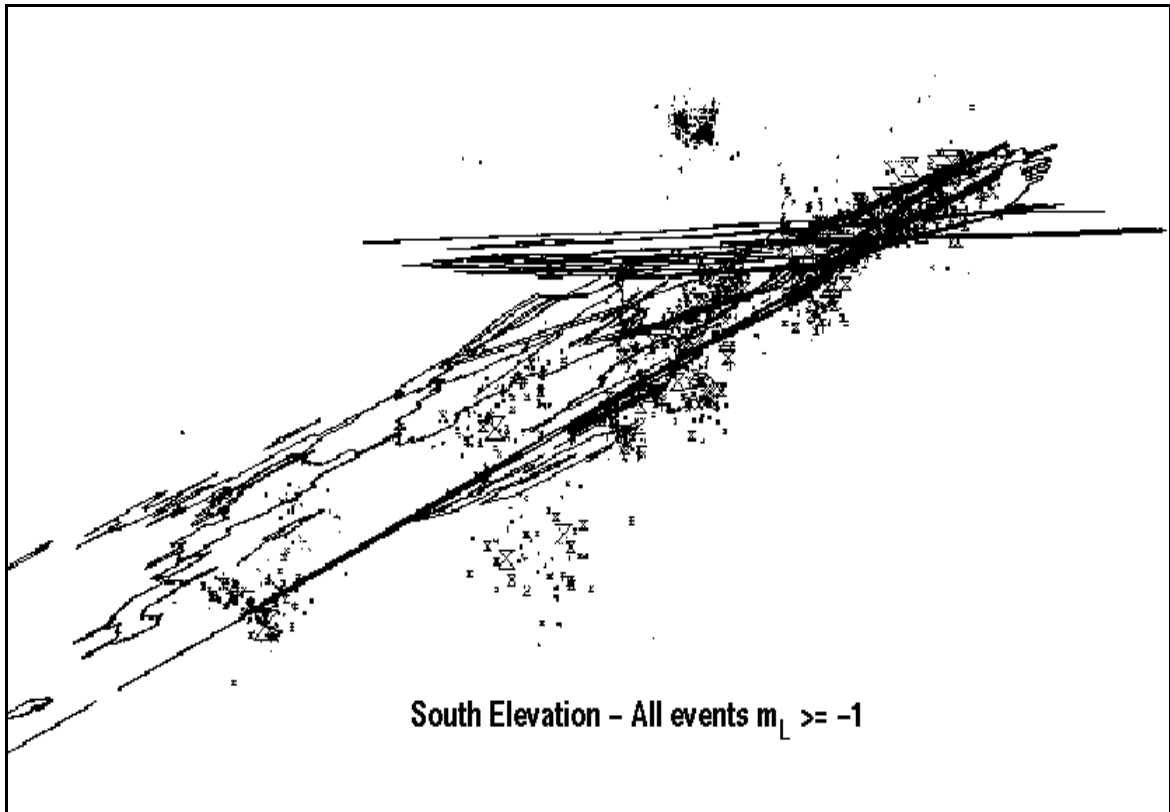
Figure 4.3.1.1(b)



**Figure 4.3.1.1(c)**



**Figure 4.3.1.1(d)**



**Figure 4.3.1.1(e)**

#### 4.3.1.2 Critical volume and time frame

Starting from a smallest spherical or circular region around a given “critical” event,  $V_p$ ,  $V_l$  and  $r$  were calculated for time periods of  $\Delta t = t_c - t_0 = 24\text{h}, 48\text{h}, 72\text{h}, \dots$ , up to the start of the catalogue, before  $t_c$ . The minimum value of  $r$ , the corresponding exponent  $\alpha$  and constant  $C$  were determined.

This process was repeated for a larger volume of radius 100m more, until the spatial limit of the catalogue was reached.

Finally, the minimum ratio for each range of time windows was plotted against the radius of the effective volume in which it was computed.

A sample plot of  $r$  versus radius, for three dimensions, is displayed in Figure 4.3.1.2(a), while Figure 4.3.1.2(b) presents the power law fit to the data points. Power law fits were only performed for cases where  $r_{min} < 0.5$ .

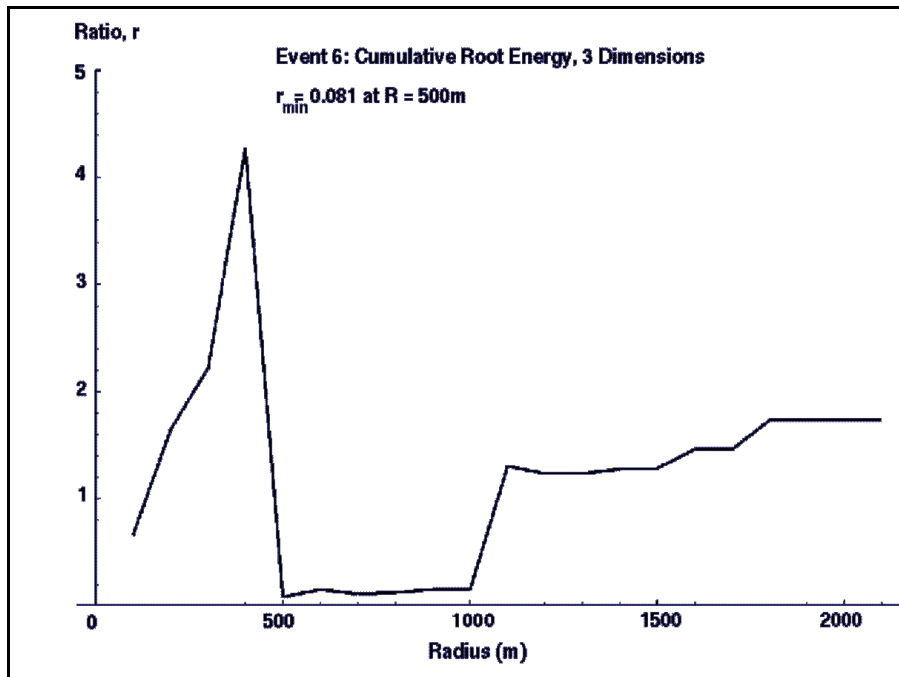


Figure 4.3.1.2(a)

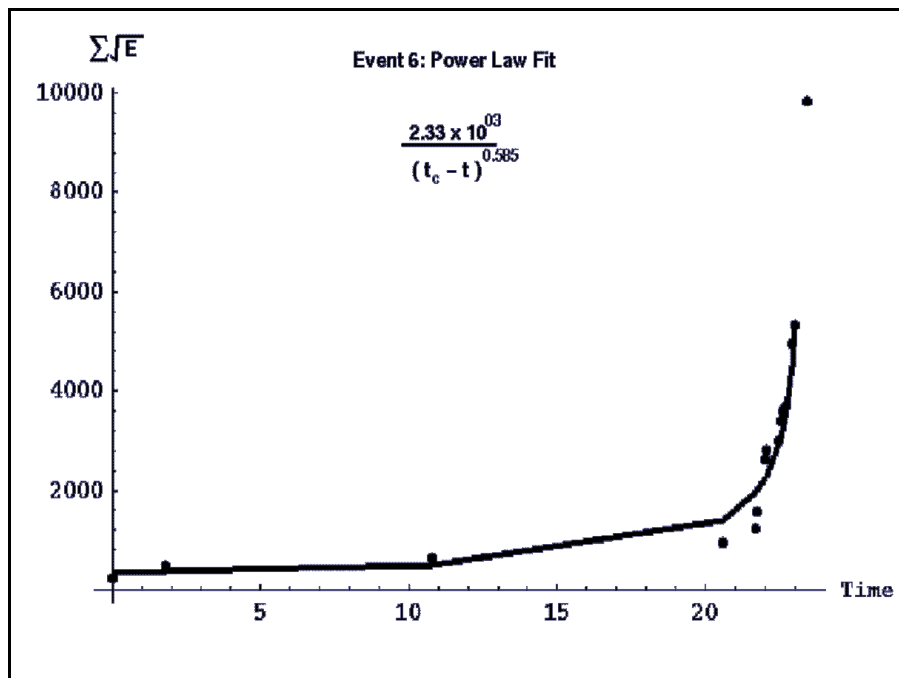


Figure 4.3.1.2(b)

The following table summarises the results, simultaneously contrasting the “2D” and “3D” cases.

**Table 4.3.1.2**  
**Critical exponents, radii and time windows**

Event	Variable	$\alpha$		$r$		$R_c$		$\Delta t(h)$	
		3D	2D	3D	2D	3D	2D	3D	2D
1	$\sum \sqrt{E}$	0.232	0.232	0.211	0.211	600	600	3.75	3.75
1	$\sum \sqrt{M}$	0.228	0.228	0.234	0.234	300	300	3.75	3.75
1	$\sum V_A$	-	0.189	-	0.436	-	300	-	3.75
2	$\sum \sqrt{E}$	-	-	-	-	-	-	-	-
2	$\sum \sqrt{M}$	-	-	-	-	-	-	-	-
2	$\sum V_A$	-	-	-	-	-	-	-	-
3	$\sum \sqrt{E}$	-	-	-	-	-	-	-	-
3	$\sum \sqrt{M}$	-	-	-	-	-	-	-	-
3	$\sum V_A$	5.04	5.04	0.468	0.468	300	300	9.5	9
4	$\sum \sqrt{E}$	-	0.627	-	0.230	-	100	-	25
4	$\sum \sqrt{M}$	-	0.521	-	0.199	-	100	-	25
4	$\sum V_A$	-	-	-	-	-	-	-	-
5	$\sum \sqrt{E}$	-	0.57	-	0.000	-	100	-	295
5	$\sum \sqrt{M}$	-	0.054	-	0.000	-	100	-	295
5	$\sum V_A$	-	-	-	-	-	-	-	295
6	$\sum \sqrt{E}$	0.585	0.670	0.081	0.107	500	600	23	23.5
6	$\sum \sqrt{M}$	0.612	0.628	0.141	0.202	500	800	23	23.5
6	$\sum V_A$	-	0.472	-	0.429	-	600	-	53
7	$\sum \sqrt{E}$	0.286	0.286	0.047	0.047	600	500	19.5	19
7	$\sum \sqrt{M}$	0.364	0.365	0.048	0.049	500	500	19.5	19
7	$\sum V_A$	-	0.731	-	0.117	-	500	-	7.5
8	$\sum \sqrt{E}$	0.087	0.118	0.251	0.117	600	900	180	190
8	$\sum \sqrt{M}$	0.054	0.327	0.297	0.000	600	300	460	462.5
8	$\sum V_A$	-	-	-	-	-	-	-	-
9	$\sum \sqrt{E}$	-	-	-	-	-	-	-	-
9	$\sum \sqrt{M}$	-	-	-	-	-	-	-	-
9	$\sum V_A$	-	1.54	-	0.000	-	100	-	600



#### 4.3.1.3 Discussion: Accelerated seismic release, critical volumes and time windows

The effect of “switching off” the difference in vertical distance between events is obvious. Recall that the 2D-projection of events was done because of very obvious artifacts of tabular network configuration and poor event processing. Events 4, 5 and 9, which showed no evidence of criticality when z-location was considered, suddenly show this in the projection. In the cases of events 1, 6 and 7, extra information has been gained by ignoring z-location.

The question may rightly be asked: “Has information been gained, or has the analysis not been obscured further by z neglect?” The answer, it is suggested, lies in the values of the critical exponent  $\alpha$ : with very few exceptions,  $\alpha$  changes very little in going from three to two dimensions.

This is extremely important. The *type* of physical process at work is characterised by the exponent  $\alpha$  in a power law or scale invariant process. One says that  $\alpha$  characterises the universality class of the process. Again, this is not unknown to the mine seismicity and rock engineering community. It has been used before in seismic hazard correlations with production rates (see for example *Mendecki and van Aswegen (1997)*), where the exponent at issue is the slope of the Gutenberg-Richter relation).

Thus, since the *type of process* (exponent) is robust under the change of dimensions, one can be fairly confident that ignoring z-location, at least for these nine events, indeed induces an improvement. In fact, z-location of events here probably just introduces noise - it is, in any event, only the 2-D dynamics that are properly captured.

The r-values for event 2 indicate that in both the 2- and 3-dimensional calculations the likelihood that the fit is linear is almost double that of a power law fit. This event lies on the edge of a cluster of events which in turn lies at the edge of the database. Such a boundary event introduces all sorts of problems relating to network sensitivity and the fact that, in this analysis, symmetric (spherical and circular) volumes were searched. In a case such as this the search volume will not be well-filled.

Events 3 and 9 present a different kind of problem again: In spite of plots that appear similar, there are probably separate explanations for events 3 and 9. Notice that event 9 has more in common with 5 and 8 than with 3. It is closer spatio-temporally to them, and a characteristic is the sparseness of events in the group containing 9. All three of the events 5, 8 and 9 have large critical time windows and this would seem to indicate something different from 3: the power law fits that the processes of which 5 and 8 are part of were interrupted by large jumps, which spoil the plots but seem to drive the system.

Event 3 locates between two longwalls and essentially shares in two elongated clusters. It appears as if this event is an outsider: it does *not* seem to belong with the rest of its cluster. Rather, it appears as if it interrupted a clearly critical nucleation associated with development.

#### 4.3.2 Critical behaviour without volume optimisation

In what follows a data set was searched for patterns of critical (accelerating) behaviour in measures of seismic release. The physical volume of rock mass was defined by a mine seismologist. In addition, a synthetic catalogue with randomly generated events considered. Very clear differences between real, deterministic and predictable data, as opposed to

random, unpredictable quantities, are pointed out.

#### 4.3.2.1 The real data

Time series of cumulative Apparent Volume, cumulative root Moment and cumulative root Energy preceding fourteen events of  $\log_{10}M \geq 12.0$  ( $m_L \sim 2.0$  and larger) and thirty - nine events of  $11.0 \leq \log_{10}M < 11.6$  ( $1.3 \leq m_L \leq 1.6$ ) were analysed for power law behaviour in the time remaining to these events. The data comes from a shaft pillar extraction on a West Wits mine.

#### 4.3.2.2 Random catalogue

Within a hypothetical rock mass consisting of a cube of sides 1000 m a data set consisting of 3000 *randomly* occurring events with randomly assigned energies, moments and locations were simulated. The same analyses as for the real data were performed and the results compared.

#### 4.3.2.3 Case studies

Four cases were considered:

- 1 **Area characterisation.** Time windows which include main events  $\log_{10}M \geq 12.0$ . Such windows contain information about processes within which the main events are embedded. It was found that the majority of event sequences exhibited power law behaviour for all three variables in this case, within reasonable time windows of, on average, three to four weeks. The random data exhibited very few cases of critical behaviour.
- 2 **Forecast window.** Time windows which terminated four hours before the main events. Such windows may form a subset of the larger processes (in time and space) above and therefore predict the processes in (1) to some extent. The seismic moment time series showed the most promising results in this case, 50% of the periods of about a month before the main events showing power law (accelerating) behaviour. Once again, the random data set behaved as expected. Not only did very few cases show indicative behaviour, but the percentage was roughly the same as in case 1 above.
- 3 **False positive predictions:** Critical behaviour for smaller events. Time windows which include main events  $11.0 \leq \log_{10}M < 11.6$ . Such windows contain information about processes within which these events are embedded and the idea is to determine whether these processes differ from those in which larger events are embedded. The majority of events again exhibited power law behaviour, over an average time window of about two weeks, which supports the criticality (scale invariance) hypothesis on which most of the work in this project is based. The same comments as before apply to the random data.
- 4 **False positive predictions: Forecast window for smaller events.** One would like to predict larger events rather than disrupt production by mis-predicting the magnitude of an impending event. Time windows which terminated four hours before main events of  $11.0 \leq \log_{10}M < 11.6$ . Such windows may form a subset of the larger processes (in time and space) above and therefore predict the processes in case 2 to some extent. At this stage one starts to probe a different evolution of the rock

mass, and this is reflected in the very small proportion of processes showing critical behaviour.

The four cases above are summarised by **comparing larger and smaller events**. An attempt is made to determine quantitative differences between possible precursory behaviour for events of  $\log_{10}M \geq 12.0$  and events of  $11.0 \leq \log_{10}M < 11.6$ .

1. Two phases of recurrent dynamics that may have been present in both the larger and the smaller instabilities can be distinguished. Although these phases for the larger and smaller events overlap, they do not coincide. Such states are absent for random data. These phases are recurrent states or fixed points of the seismic rock mass evolution in response to mining, and are *signatures of determinism and predictability*.
2. For the area and time span as a whole, time series of the cumulative seismic moments may exhibit *patterns significantly different during the imminence of larger instabilities as opposed to smaller events*. For random data, the differences are of the order of the initial percentages, as expected.

In more detail:

#### 4.3.2.4 Main events included in sequence

In what follows it has is important to keep in mind that spatial optimisation, which falls considerably outside of the scope of this project, has not been preformed. A polygon of data defined by a mine seismologist around a workplace has been used. This matter needs urgent attention, but is a non-trivial problem.

##### Area characterisation (real data)

The tables below contain the results of the first run in broad outlines.

**Table 4.3.2.4 (a)**

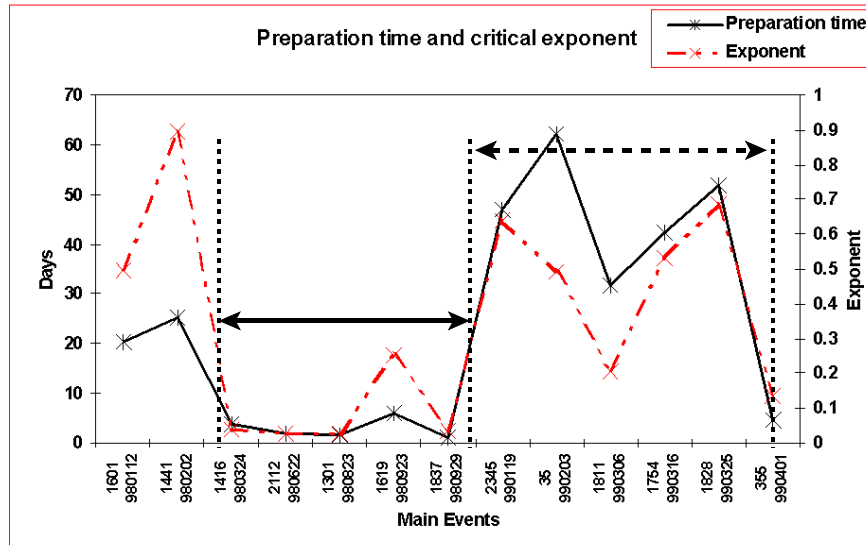
<i>Sequence</i>	<i>% Real Events Exhibiting Power Law</i>	<i>Avg. Critical Time (days)</i>	<i>Median Critical Time (days)</i>
<i>Moment</i>	86	23.56	18.25
<i>Volume</i>	93	23.07	20.41
<i>Energy</i>	79	26.29	20.41

There are, however, details which these tables, averaging the Physics as they do, cannot provide. When one studies Figures 4.3.2.4 (a) to 4.3.2.4 (c), it is clear that the period under consideration can be divided into two distinct regions:

The “deformation variables” Apparent Volume and Seismic Moment distinguish clearly between a period spanning the second and third quarters of 1998 and a period spanning the fourth quarter 1998 to the second quarter 1999. In these periods the time windows over which accelerating behaviour in the two seismic variables above took place settled into a fairly stable size. Investigation in collaboration with the mine seismologist confirmed that there was a clear correlation with the mining method / rate as well as excessive lead / lags during these periods. The “stress variable”, Seismic Energy, does not show this pattern as clearly. From a daily analysis viewpoint, this is the kind of pattern one would like to see: the more constant the time window over which acceleration takes place, the more confidence one can have in daily patterns. The system seems to have settled into a recurrent state of response during these six months.

Cumulative Apparent Volume

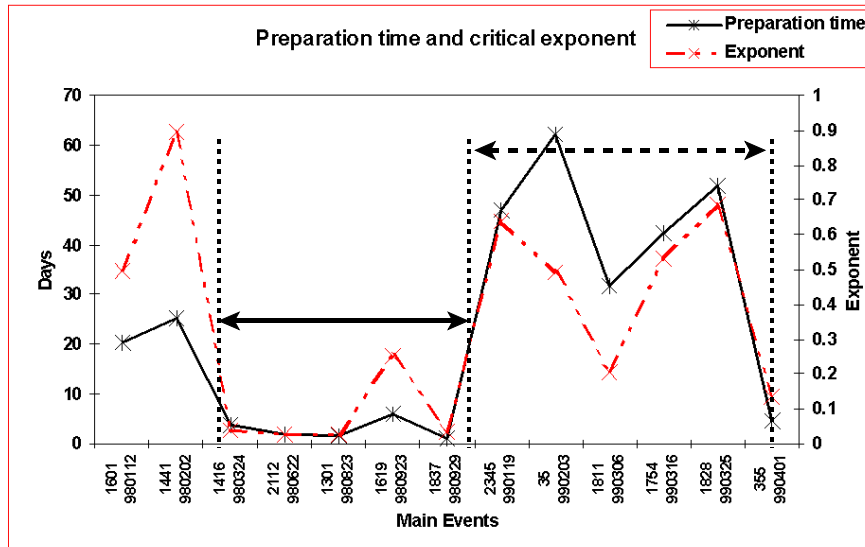
Minimum 14 points, time increment 8 hours



**Figure 4.3.2.4 (a) Critical time windows and exponents for  $\log M \geq 12.0$  events. The final source parameters were included in the analysis in order to establish pattern recognition benchmarks. Note the period of stable critical time windows (solid arrow) characterising a period of fast rockmass response during a particular type of mining. Preparation time and critical time window are equivalent concepts. Note also in this and following figures, that the x-axis does not represent time, but individual events. The solid arrow spans a period of 6 months and the dashed arrow only 3 months.**

Cumulative Apparent Volume

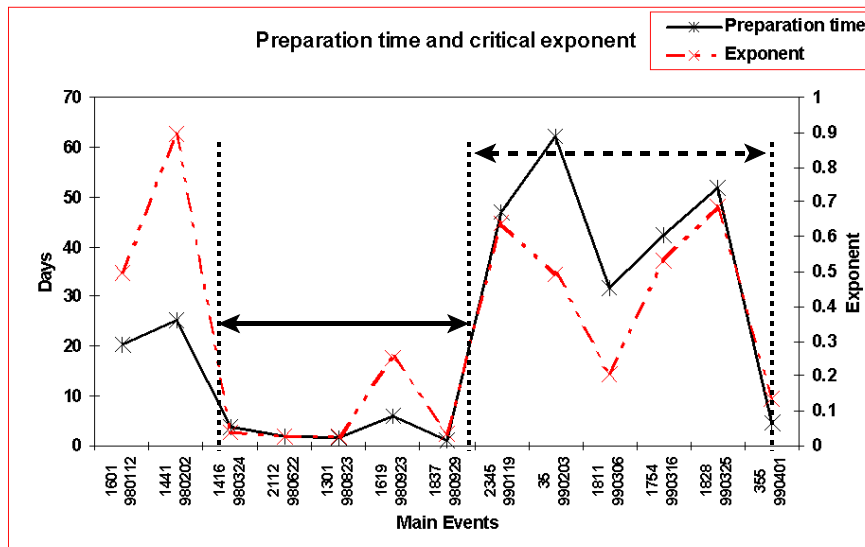
Minimum 14 points, time increment 8 hours



**Figure 4.3.2.4 (b) Critical time windows and exponents for  $\log M \geq 12.0$  events. The final source parameters were included in the analysis in order to establish pattern recognition benchmarks. Note the period of stable critical time windows (solid arrow) characterising a period of fast rockmass response during a particular type of mining. Preparation time and critical time window are equivalent concepts. Note also in this and following figures, that the x-axis does not represent time, but individual events. The solid arrow spans a period of 6 months and the dashed arrow only 3 months.**

Cumulative Apparent Volume

Minimum 14 points, time increment 8 hours



**Figure 4.3.2.4 (c) Critical time windows and exponents for  $\log M \geq 12.0$  events. The final source parameters were included in the analysis in order to establish pattern recognition benchmarks. Note the period of stable critical time windows (solid arrow) characterising a period of fast rockmass response during a particular type of mining. Preparation time and critical time window are equivalent concepts. Note also in this and following figures, that the x-axis does not represent time, but individual events. The solid arrow spans a period of 6 months and the dashed arrow only 3 months.**

**Area characterisation (random simulation)**

Recall that this exercise specifically points out the difference in behaviour between real data and random data. It is also shown that extremely primitive attempts at defining spatial critical volumes, outside the scope of the present work, significantly improves the detection of precursory patterns.

Table 4.3.2.4 (b) contrasts the percentage of random events exhibiting power law behaviour with the real data:

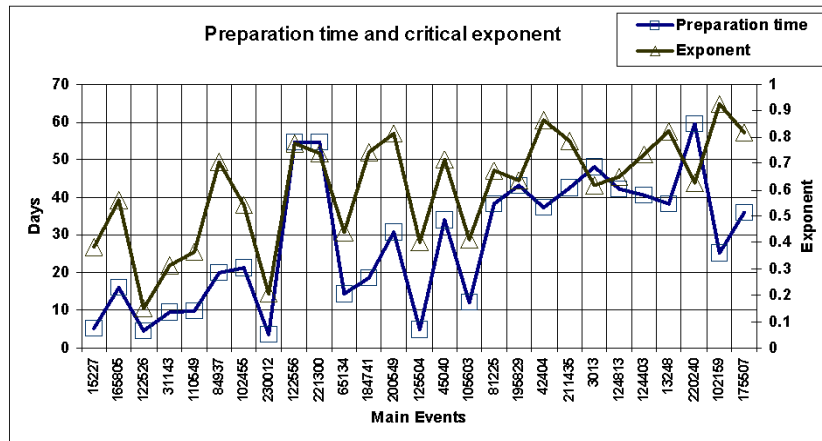
**Table 4.3.2.4 (b)**

Sequence	% Real Events Exhibiting Power Law	% Random Events Exhibiting Power Law
Moment	86	13.5
Volume	93	6.5
Energy	79	5.0

Figures 4.3.2.4 (d) to 4.3.2.4 (f) show the same variables as for the real data. Noteworthy here is Figure 4.3.2.4 (f), where a clear trend is visible for the "Moment" variable, which had

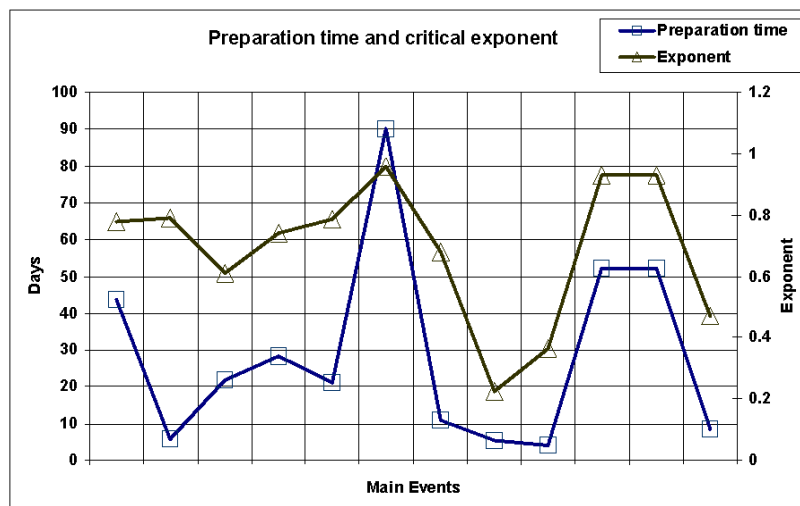
the highest percentage of power law cases. This is of course to be expected for random data - by enlarging the time window sufficiently, chances are that power law behaviour will be found eventually, albeit a poor power law - note how large the exponent is (the larger the exponent, the closer one is to a straight line, stationary situation). Small as they are, the random percentages actually appear quite high, until one remembers that the size of the final event is included in the analysis. Hence the next test.

Random data: Cumulative Root Moment, events  $\log M \geq 12.0$



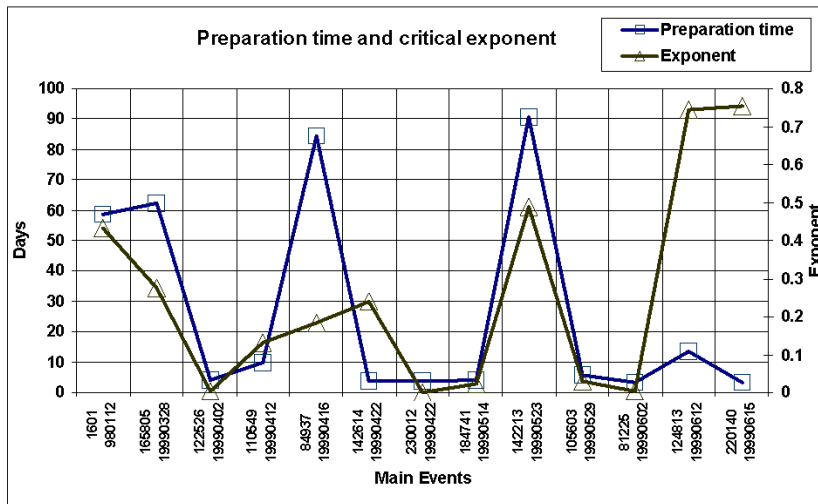
**Figure 4.3.2.4 (d) Same procedure as Figures (a) to (c), but for random data. Note an expected increase in time window as the length of the data set increased. This is absent in real data, since real seismicity has deterministic fixed points.**

Random data: Cumulative Root Energy, events  $\log M \geq 12.0$



**Figure 4.3.2.4 (e) Same as Figure 4.3.2.4(d), but for a random "Seismic Energy".**

Random data: Cumulative Apparent Volume, events  $\log M \geq 12.0$



**Figure 4.3.2.4 (f)** Exactly the same procedure as before, for random data. Note the absence of unique structure in the critical time windows.

#### 4.3.2.5 Forecast window

##### Real data

A further test was to determine whether possible precursory sequences could be identified for the main events. Here, sequences of the variables of interest were analysed up to 4 hours prior to the main events and power law behaviour identified if it existed. Abbreviated results are tabulated below.

**Table 4.3.2.5 (a)**

Sequence up to 4 hours prior to main events	% Sequences Exhibiting Power Law	Average Critical time (days)	Median Critical time (days)
<b>Moment</b>	50	28.51	17.85
<b>Volume</b>	36	28.38	24.09
<b>Energy</b>	36	34.9	36.26

The issue of objectively determining critical spatial volumes of interest is a crucial one to rock mechanics practitioners. It comes up repeatedly in conversations and consultations. That this is a relevant problem can be seen in Table 4.3.2.5 (b). In a crude experiment, circular spatial volumes of various sizes around the large events were applied to the real precursory sequences (the same ones as in Table 4.3.2.5 (a)).

**Table 4.3.2.5 (b)**

Sequence up to 4 hours prior to main events	% Sequences Exhibiting Power Law
<b>Moment</b>	50
<b>Volume</b>	43
<b>Energy</b>	50



Please note the increase in apparent volume and energy signatures. It is regrettable that this line of inquiry could not be pursued, as it falls outside the ambit of GAP409.

A further note is that these figures **do not mean** that the main events were only 50% predictable - the automated predictor picks up **all power law behaviour** not only the fastest, as here, and it does this events by event. These figures are meant to be used to characterise an area and assist in deciding which prediction is statistically more significant.

Furthermore, more data in the window always improved these percentages drastically - remember that we are using cumulative variables describing a **process**, not single precursors!

The critical time windows tend to be long, the ratio of power law to linear variance large and exponents large: this indicates that the behaviour is still close to linear, and that one is still in the tail of the power law. It is probably a signature of a system as stiff as a pillar that the onset of instability happens late in the process, over a short time

## Random data

The random data were also tested for “precursory” signatures (to use an oxymoron). Note in Table 4.3.2.5 (c) that the percentages of random sequences exhibiting power law behaviour did not change much from the full sequences in Table 4.3.2.5 (c). This is, of course, to be expected.

**Table 4.3.2.5 (c)**

<i>Sequence up to 4 hours prior to main events</i>	<i>% Real Sequences Exhibiting Power Law</i>	<i>% Random Sequences Exhibiting Power Law</i>
<i>Moment</i>	50	4.0
<i>Volume</i>	36	2.5
<i>Energy</i>	36	6.0

## False positive predictions

An attempt must be made to establish to which extent  $\log_{10}M \geq 12.0$  processes can be forecast without confusing them with, say,  $11.0 \leq \log_{10}M < 11.6$  processes. This is known as a **false positive prediction** problem.

Four tests were performed:

### **Critical behaviour for smaller events (real data)**

Sequences of the variables of interest are analysed for events of  $11.0 \leq \log_{10}M < 11.6$  ( $1.3 \leq m_L \leq 1.6$ ) in order to identify critical behaviour for smaller events.

**Table 4.3.2.5 (d)**

<i>Sequence including events <math>11.0 \leq \log_{10}M &lt; 11.6</math></i>	<i>% Sequences Exhibiting Power Law</i>	<i>Average Critical time (days)</i>	<i>Median Critical time (days)</i>
<i>Moment</i>	77	15.34	4.8
<i>Volume</i>	90	14.95	6.07
<i>Energy</i>	92	14.78	5.54

Of note in the detail was

1. again, a clear recurrent period in the third and fourth quarters 1998 (up to end January 1999), and
2. the possibility of distinguishing two populations (high and low exponent) in all three cases.

Please note that the high percentage of smaller events exhibiting scaling in the time to failure is not necessarily a cause for concern - it merely contributes to the body of evidence suggesting that there is Physics behind mining induced seismicity, namely the framework of critical, complex phenomena. In other words, mine seismicity is **not** a random process, unpredictable by nature.

However, once again the hypothesis is tested against the random data set in order to bring to light any artifacts in the algorithm or bias on the part of the research contractor.

***Critical behaviour for smaller events (random data)***

An indication of the answer to the possible indictment above may be found by perusing the results of a similar run on the random data set. Table 4.3.2.5 (e) shows the contrast between the number of smaller events exhibiting power law behaviour for the random, as opposed to the real, data. Note again that the percentages for random data remain largely unaffected by the magnitude of an event in a random sequence.

**Table 4.3.2.5 (e)**

<i>Sequence up to 4 hours prior to main events</i>	<i>% Real Sequences Exhibiting Power Law</i>	<i>% Random Sequences Exhibiting Power Law</i>
<i>Moment</i>	77	4.0
<i>Volume</i>	90	0.0
<i>Energy</i>	92	4.5

***False positive predictions: Forecast window for smaller events.***

In Table 4.3.2.5 (f) the search for precursory sequences (recall that a precursory sequence is one that terminates, in this case, four hours before the event being forecast) for smaller events is summarised. Here we are considering smaller events in order to see whether they can be distinguished from the larger ones while predicting.

**Table 4.3.2.5 (f)**

<i>Sequence up to 4 hours prior to main events</i>	<i>% Real Sequences Exhibiting Power Law</i>	<i>% Random Sequences Exhibiting Power Law</i>
<i>Moment</i>	21	5.3
<i>Volume</i>	26	2.6
<i>Energy</i>	18	6.7

By now the results should speak for themselves. Far fewer small events have distinguishable precursory behaviours than do large ones. Moreover, once again, the random sequence not only has very few such patterns, but the numbers do not differ greatly between small and large “events”.

## 4.4 Predictor: Time to failure inversion

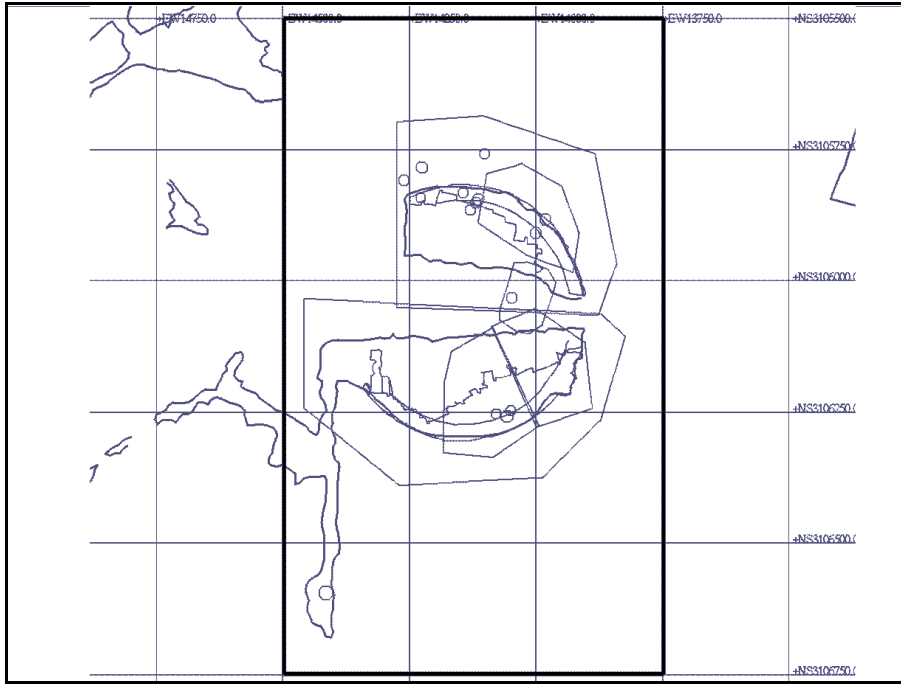
The learning and prediction algorithm detailed in section 3.1.3.2 was applied to the West Wits pillar extraction above, as well as to a Free State pillar extraction. In everything that follows, the following criteria must be borne in mind:

1. The algorithm learns from the immediate past by collecting the time intervals between the events terminating power law sequences and characteristic events. Such time intervals will be called the “real time to failure”. In the same way, track is kept of the characteristic release. An ALERT is issued if the predicted time to failure falls within two standard deviations of the average “real” time to failure and the predicted release falls within two standard deviations of the average real release in the past.
2. An ALERT is cancelled if the event immediately after the ALERT was issued is of a size less than the predicted release.
3. When a characteristic event occurs, the ALERTS issued are analysed to see whether the event had fallen within any of their time and size windows. If it does, the prediction is declared successful. If at least one of the predictions is successful, the event is declared a successfully predicted event.

It should be noted that, after an initial training set of about eight characteristic events (the number depending on the area), the algorithm operates in a forward prediction mode.

### 4.4.1 Free State shaft pillar extraction

This catalogue consists of 2621 events generated over fourteen months of shaft pillar extraction in the Free State. Mining takes place at a depth of about 2000m, and the system is very soft: the shaft pillar is the only remnant in a large mined out area. The catalogue contains 40 characteristic events. The case is of interest for two reasons: First, remnant extraction in general is of great importance to the gold mining industry and secondly this particular instance is novel, in the sense that the pillar is mined from the inside out, leaving a 30m annulus for stability. Figure 4.4.1(a) gives an impression of the mining and some larger events.



**Figure 4.4.1(a) Free State pillar extraction. Mining is concentrated at about 2000m depth, on the northern and southern sides of the shaft barrel.**

The predictor algorithm SOOTHSAY described in section 3.1.3.2 was applied to this dataset with, after some initial experimentation, the following parameters:

The highest characteristic release with the best linear fit was taken as events with  $\log_{10}M_0 \geq 11.6$ . In principle one is interested in events of moment-magnitude 2.0 ( $\log_{10}M_0 \geq 12.0$ ) or larger. However, empirically it is found that (Figure A2.2) the “large” release in a given volume of rock mass fits a straight line, which is handy, since one is provided with a definition for a “large” event or episode of events. The linear relationship is a reflection of the fact that the seismic release which moves the rock mass away from criticality, or “resets” the system, is truly a physical characteristic of the system of interest. One can also view these episodes as transient states, i.e., a dynamic macrostate of the system that resets it, whereupon it self-organises again to a critical point, which is then reset, and so on.

In the case of mining-induced seismicity, there are two types of characteristic release: large seismic events on structures, which are releases triggered by long-term evolution of the rock mass driven by mining. A further characteristic release, which the principal author of this report believes to be relevant, as opposed to being treated as something to be excised from the data, is post-blasting rockmass response. This contention is borne out by the fact that lowering the threshold for identification of characteristic episodes to include the larger of these transient responses leads to a better linear fit of the characteristic release (Figure A2.2). Here, a sliding window of the minimum threshold of four characteristic events (the smallest number usually taken as giving a statistically significant fit to a straight line) was used for the fit. Increasing this number was counter-productive. Mining-induced seismicity, as opposed to tectonically driven seismicity, is generated by a varying driving force, so that the linear fit of characteristic release as well as the log-log fit power-law constants and

characteristic release (Figure A2.1) changes from time to time. In practice, this number is a function of the size and shape of the polygon within which the catalogue is constructed, as well as the local geology and mining method.

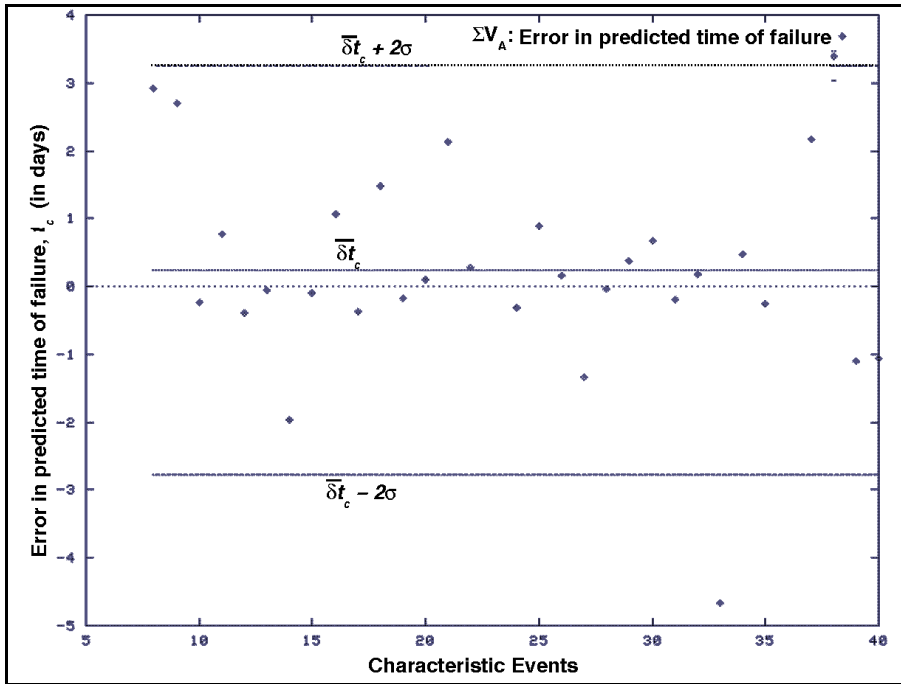
As an overall result for the Free State pillar extraction, the Cumulative Apparent Volume was by far the most successful prediction strategy. For this strategy, the algorithm utilised 11 events as an initial training set, attempted to invert for the time of occurrence of 27 events (two had insufficient precursory events for the algorithm to use), and issued alerts for 25.

Of the 29 characteristic episodes, 65.5% were declared successfully predicted, as per the criteria above) with 27.6%, or eight of the 29, failures to predict, i.e., these events were false negative predictions.

Of the 371 predictions made, 23.45% were successful and 76.550% false positive (false alarms). This figure is highly misleading and naive, though:

1. It should be noted that not all these predictions led to extended ALERT periods: in fact, most of them were cancelled fairly soon: *of the total time span of the catalogue of 403.923 days (14 months), only 35.949 days, or 8.9% were spent under ALERT.*
2. The following simple exercise in logic should bring home the improvement in prediction accuracy: If, on every day of the life of the catalogue, a prediction was issued that a characteristic event would happen within twenty-four hours, the probability of success would have been 7.2%, so that, in terms of the number of predictions, the algorithm is a an improvement over essentially a random situation by a factor of 3.26.

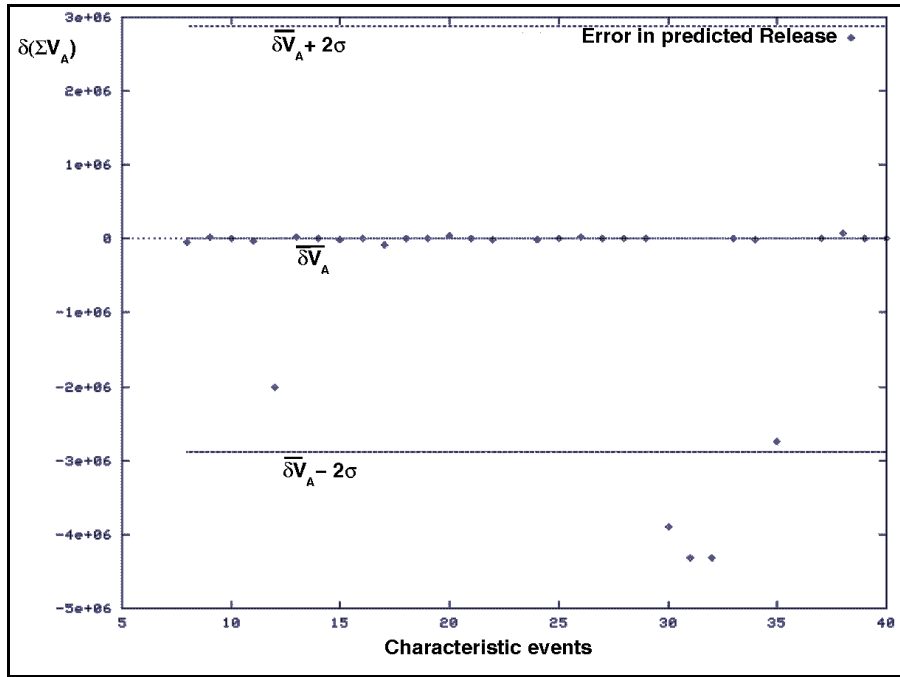
Focussing attention on Figure 4.4.1(b), the average prediction error was underestimation of 0.241 days +/- 3.02 days, i.e., the error made in the prediction of the time to failure ranged from an overestimation of 77.304 hours to an underestimation of 78.264 hours. Thus, at the current point of submission of this report, an ALERT should be interpreted as a period of increased probability of occurrence and operating procedures should be devised specifically for such periods. Such a system is already followed by ISS International in Welkom and several mines in the area (*Spottiswoode and de Beer, 1998*).



**Figure 4.4.1(b) Average error made in the prediction of the time to failure in a Free State pillar extraction by using the SOOTHSAY algorithm and using time series of cumulative apparent volume.**

Figure 4.4.1(c) analyses the performance of the algorithm, using the cumulative apparent volume, in predicting the size of characteristic events.

The median error in predicting the release (Figure 4.4.1(c)), is negligible, although large outliers exist.

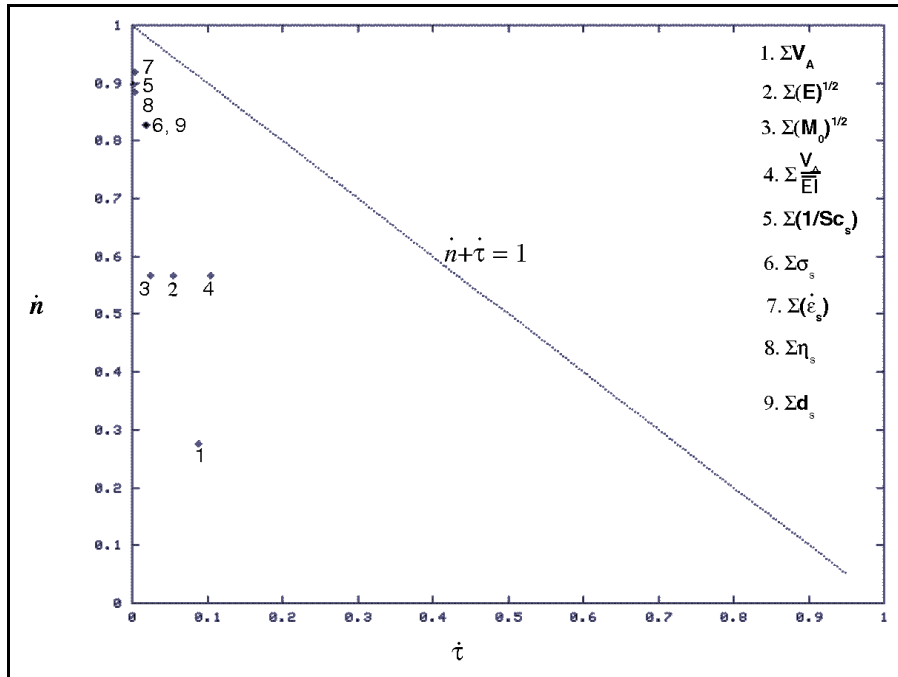


**Figure 4.4.1(c) Median error made in the prediction of the characteristic cumulative apparent volume in a Free State pillar extraction by using the SOOTHSAY algorithm.**

Finally, let us turn to an error analysis of the prediction strategies employed. Of the remaining strategies, the cumulative square root of the moment, the cumulative square root of the energy (the so-called “Bennioff” strain) and the cumulative quotient of the apparent volume and the energy index were the only noteworthy performers, all claiming a success rate of 36.667%. The rest are displayed in table 4.4.1. Of more interest is the error diagram (Figure 4.4.1(d)), which clearly shows the non-random nature of the cumulative apparent volume predictor. Recall that the line  $\dot{n} + \dot{\tau} = 1$ , where  $\dot{n}$  is the rate of failures to predict (missed events) and  $\dot{\tau}$  the rate of time alarms (fraction of the whole catalogue duration spent under ALERT), the more random the prediction strategy can be considered to be. The best performers (Table 4.4.1) all lie safely away from this line, although clearly the cumulative apparent volume is the most satisfactory.

**Table 4.4.1 Pattern summary**

Pattern summary									
	AppVol	SeisEn	SeisMom	AppV/EI	SeisSch	SeisSig	SeisEps	SeisDiff	SeisVisc
<b>Percentages in terms of numbers of predictions</b>									
Success	23.450	26.695	36.667	20.286	7.407	10.072	0.000	85.714	19.355
False Alarm	76.550	73.305	63.333	79.714	92.593	89.928	100.000	14.286	80.645
<b>Percentages in terms of numbers of events</b>									
Success	65.517	36.667	36.667	36.667	3.448	10.345	0.000	3.846	10.345



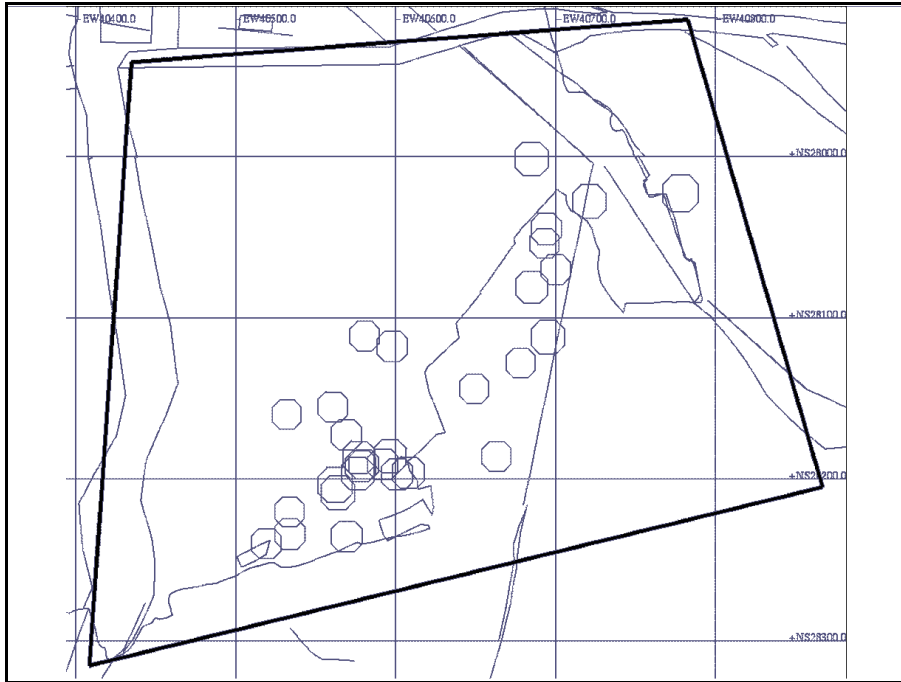
**Figure 4.4.1(d) Error diagram for the prediction strategies in section 3.1.2 applied to the seismicity accompanying a Free State shaft pillar extraction. Note the good performance of the cumulative apparent volume, even though it still is not optimal - more experimentation with spatial selection is required.**

#### 4.4.2 West Wits pillar extraction

The West Wits case consisted of 10204 events spanning eighteen months. The general outlay and the largest events are indicated in Figure 4.4.2(a). This is a stiff system, a large remnant.

The predictor algorithm SOOTHSAY was applied to this dataset with the same parameters as above.





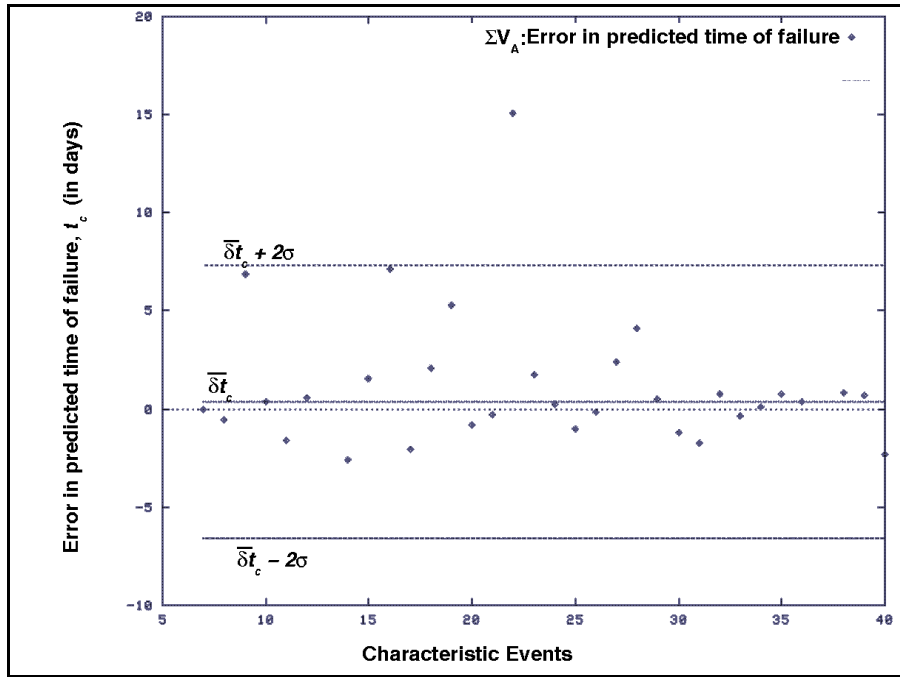
**Figure 4.4.2(a) West Wits pillar extraction. Mining is concentrated at about 3000m depth, on the south-eastern corner of the shaft pillar. The reef dips at about  $22^\circ$  to the south-south-west.**

The Cumulative Apparent Volume was again the most successful prediction strategy. For this strategy, the algorithm utilised 11 events as an initial training set, attempted to invert for the time of occurrence of 18 events and issued alerts for 16.

Of the 30 characteristic episodes, 53.3% were declared successfully predicted.

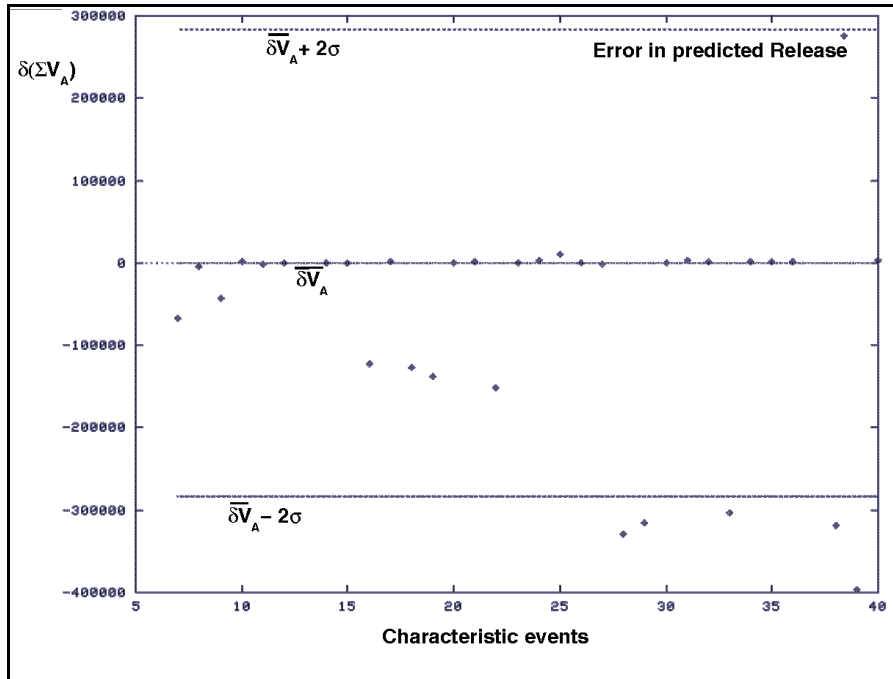
Of the 290 predictions made, 33.3% were successful and 66.7% false positive (false alarms). Again it should be noted that this figure includes cancelled alerts: *of the total time span of the catalogue of 518.3 days (15 months), only 22.255 days, or 4.3% were spent under ALERT.* Furthermore, in the same exercise as before: If, on every day of the life of the catalogue, a prediction was issued that a characteristic event would happen within twenty-four hours, the probability of success would have been 5.8% of success, so that, in terms of the number of predictions, the algorithm is an improvement over essentially a random situation by a factor of 5.7.

Focussing attention on Figure 4.4.2(b), the median prediction error was an underestimation of 0.378 days 6.978 days, i.e., the error made in the prediction of the time to failure ranged from an overestimation of 158.4 hours to an underestimation of 176.568 hours.



**Figure 4.4.2(b) Median error made in the prediction of the time to failure in a West Wits pillar extraction by using the SOOTHSAY algorithm and using time series of cumulative apparent volume.**

Figure 4.4.2(c) analyses the performance of the algorithm, using the cumulative apparent volume, in predicting the size of characteristic events. The median overestimation of the final release is negligible.

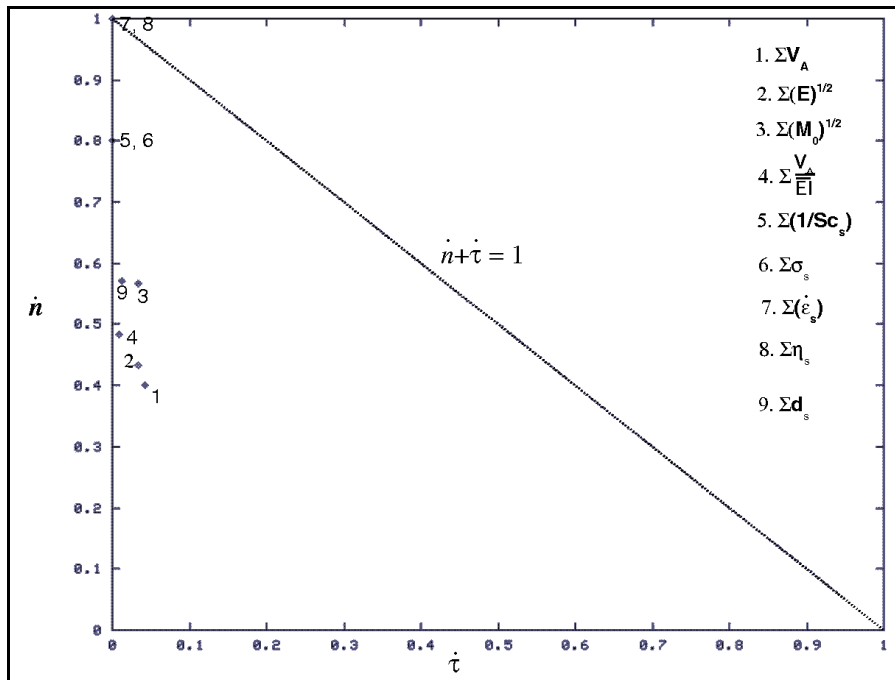


**Figure 4.4.2(c) Median error made in the prediction of the characteristic cumulative apparent volume in a West Wits pillar extraction by using the SOOTHSAY algorithm.**

The error diagram (Figure 4.4.1(d)) shows again that most of the prediction strategies were non-random, although not in optimal mode. Of the remaining strategies, the cumulative square root of the moment, the cumulative square root of the energy (the so-called “Benioff” strain) and the cumulative quotient of the apparent volume and the energy index (Table 4.4.2.) were the only noteworthy performers. Of special note is the Benioff strain, which, with 50% of events claimed as successfully predicted, also shows that there were more successful predictions than false positives issued. All the strategies require optimisation.

**Table 4.4.2 Pattern summary**

Pattern summary									
	AppVol	SeisEn	SeisMom	AppV/EI	SeisSch	SeisSig	SeisEps	SeisDiff	SeisVisc
<b>Percentages in terms of numbers of predictions</b>									
Success	33.333	56.633	40.823	32.463	0.000	0.000	0.000	0.000	10.976
False Alarm	66.667	44.367	59.177	67.537	100.000	100.000	100.000	100.000	89.024
<b>Percentages in terms of numbers of events</b>									
Success	53.333	50.000	36.337	44.828	0.000	0.000	0.000	0.000	28.571



**Figure 4.4.2(d) Error diagram for the prediction strategies in section 3.1.2 applied to the seismicity accompanying a West Wits shaft pillar extraction. Note the good performance of the cumulative apparent volume, even though it tends towards an optimist strategy.**

### 4.4.3 Random catalogue

SOOTHSAY was applied to the same pseudo-random catalogue as the one used in section 4.3. For only one strategy, the cumulative root seismic moment, was a single prediction issued. For all others no predictions were made and hence no successes claimed for any of the 1122 events of " $\text{Log}M_0 \geq 11.6$ ".

## 5 Conclusions and recommendations

### 5.1 Conclusions

There is a serious lack of quantitative pattern recognition and prediction methodologies in the analysis of seismicity in South African gold mines. The goal of this project is to move closer to filling the gap between quantitative monitoring and automated interpretation for prediction, currently performed intuitively and within the limitation of human frailty. This chapter details the extent to which this goal has been achieved, points out serious deficiencies and recommends action to alleviate these deficiencies.

#### 5.1.1 Primary output

**Method for the prediction of rock mass instability based on kinematics of failure, nonlinear dynamics (chaos theory) and neural networks, using data supplied by the modern seismic networks in South African mines.**

A literature study which spanned the whole duration of the project, as well as many private discussions with some of the best minds alive in earthquake science informed the design of the method. There is no doubt that, where enough data exists, power law in the time to failure of a measure of seismic release is a predictor of higher-level rupture. The ideal predictor is a combination between pattern recognition (in cases of sparse catalogues) and the time to failure inversion.

A forward predictor, SOOTHSAY, utilising the time to failure equation, was designed and performs with remarkable, if not optimal, success. Optimising the predictor is mainly a matter of tuning it for a given area. A shortcoming is the lack of objective, time-effective determination of critical volumes. The preliminary tests determining the performance of the predictor indicate that it performs at a level better than random for the better prediction strategies. It has also been demonstrated that there is a significant gain in the confidence with which a SOOTHSAY prediction should be weighed over a pessimist strategy of issuing daily warnings. It has been demonstrated that the predictor is solidly based on current thinking in the physics of complex and disordered systems.

This project moves prediction ever further away from art, superstition and statistical / numerological wizardry, forwards, towards a hard, quantitative science.

#### 5.1.2 Pattern recognition

It has been shown that quantitative pattern recognition can be applied to mining induced seismicity. The pattern recognition algorithm INDICATOR gives significant results which can be expanded to a full prediction strategy using the appropriate statistical techniques. Alternatively, the algorithm can be used to supplement the deterministic time to failure predictor.

#### 5.1.3 Non linear dynamics

Techniques from non linear dynamics such as false nearest strands to determine the random component of a seismic time series and singular value decomposition noise reduction can

be applied to time series of seismic variables. Currently the problem is that, in order to obtain significant results, such series have to span too long a time interval to be of practical use. However, it was demonstrated that significant increases in spatio-temporal predictability horizons can possibly be effected by application of these methods.

## 5.2 Recommendations

### 5.2.1 Pattern recognition

The pattern recognition algorithm INDICATOR can be converted to a fully fledged prediction algorithm like the well-known M8. The latter rests mainly on magnitude and activity rate to quantify seismicity and periods of increased probability of instability. INDICATOR is based on far richer functions. Motivation for this action lies in the fact that seismicity in an area may not always exhibit power law behaviour.

**Recommendation:** A year long project to perform the conversion, test it and implement it on a daily basis in industry.

### 5.2.2 Time to failure predictor: shortcomings

The predictor SOOTHSAY suffers from a serious shortcoming, which could not be addressed in GAP409. Techniques for the spatial prediction of areas of instability are in their infancy. Several physical models, such as anomalous diffusion, Markov point processes and contouring of historical critical volumes using the current SOOTHSAY methodology should be tested for efficiency and applicability in this regard.

**Recommendation:** A one to two - year long project, concurrent with the one recommended in section 5.2.1. This type of project is ideally suited to the initial stages of a PhD degree in Physics, thus combining the needs of industry with the current imperative to train high-level manpower.

### 5.2.3 Time to failure predictor: implementation

Little attention has been paid up to now in the practical implementation of the SOOTHSAY algorithm. Guidelines for the learning phase of the algorithm and the setting of appropriate thresholds for ALERT conditions need to be developed.

**Recommendation:** A one-year project in close collaboration with several mines to develop such guidelines, if necessary concurrent with the recommendations in 5.2.1 and 5.2.2.

## References

- Ashby, M. F. and B. F. Dyson 1986.** Creep damage mechanics and micromechanisms. *Advances in fracture research*, 1: 3-30.
- Bak, P. 1996.** How nature works - the science of self-organized criticality. New York: Springer-Verlag, 198 pp.
- Bowman, D. D., G. Ouillon, C. G. Sammis, A. Sornette and D. Sornette 1998.** An observational test of the critical earthquake concept. *Journal of Geophysical Research*, 103: 24359-24372.
- Boonzaaier, L. 2000.** Application of models of self-organised criticality to mining-induced seismicity. *M. Sc. dissertation*. Stellenbosch: University of Stellenbosch, ???p. In preparation.
- Brehm, D. J. and L. W. Braile 1998a.** Intermediate term earthquake prediction using precursory events in the New Madrid seismic zone. *Bulletin of the Seismological Society of America*, 88: 564 – 580.
- Brehm, D. J. and L. W. Braile. 1998b.** Intermediate-term earthquake prediction using the modified time-to-failure method in the New Madrid seismic zone. *Bulletin of the Seismological Society of America* 88. pp 564 -580.
- Brehm, D. J. and L. W. Braile. 1999.** Intermediate-term earthquake prediction using the modified time-to-failure method in Southern California. *Bulletin of the Seismological Society of America* 89. pp 275 - 293.
- Brummer, R. 1999.** Simple Truths about Rockbursts. *SARES '99, Keynote Address*.
- Cardy, J. 1996.** Scaling and renormalisation in Statistical Physics. Cambridge: Cambridge University Press, 238 pp.
- De Beer, W and A. J. Mendecki 1997.** Seismology for rockburst prediction. *SIMRAC year-end report GAP409*. Pretoria. Department of Minerals and Energy.
- De Beer, W. 1999.** Seismology for rockburst prediction. *SIMRAC first quarter report GAP 409*. Pretoria. Department of Minerals and Energy.
- De Beer, W and A. J. Mendecki 1997, 1998.** Seismology for rockburst prediction. *SIMRAC interim reports GAP409*. Pretoria. Department of Minerals and Energy.
- Freeman, J. A. and D. M. Skapura 1991.** Neural networks: Algorithms, applications and programming techniques. Reading, Massachusetts: Addison-Wesley, p128-130.
- Gelfand, I. M., SH. A. Guberman, V. I. Keilis-Borok, L. Knopoff, F. Press, E. Ya. Ranzman, I. M. Rotwain and A. M. Sadovsky 1976.** Pattern recognition applied to earthquake epicenters in California. *Physics of the earth and planetary interiors*, 11:227-283.
- Glasstone, S. et al 1941.** The theory of rate processes - Kinetics of chemical reactions, viscosity, diffusion and electrochemical phenomena. New York: McGraw-Hill, 611p.

**JGI 1997.** Special section - Assessment of schemes for earthquake. *Geophysical Journal International*, 131: 413-533.

**Kagan, Y. Y. 1996.** Statistical aspects of Parkfield earthquake sequence and Parkfield prediction experiment. *Tectonophysics*, 270: 207-219.

**Keilis-Borok, V. I. and A. J. Lichtman 1993.** The self-organisation of American society in presidential and senatorial elections. (In: *Kravtsov, Y. A. Limits of predictability*. Springer-Verlag, Berlin, 1993, p223-237).

**Keilis-Borok, V. I. and A. Soloviev 1999.** *Pattern recognition: Algorithms and Applications. Lecture and tutorial notes of the Fifth Workshop on Non-Linear Dynamics and Earthquake Prediction - document H4.SMR/1150-6*. Trieste, 4-22 October 1999: The Abus Salaam International Centre for Theoretical Physics.

**Knopoff, L. 1997.** Personal communication on the use of non parametric statistics in seismology. Institute of Geophysics and Planetary Physics, University of California, California, USA.

**Kossobokov, V. G., L. L. Romashkova, V. I. Keilis-Borok and J. H. Healy 1999.** Testing earthquake prediction algorithms: statistically significant advance prediction of the largest earthquakes in the Circum-Pacific, 1992-1997. *Physics of the earth and planetary interiors*, 111:187-196.

**Kostrov, B. V. and S. Das 1988.** Principles of earthquake source mechanics. Cambridge: Cambridge University Press, 274p.

**Kravtsov, Y. A. 1993.** Fundamental and practical limits of predictability. (In: *Kravtsov, Y. A. Limits of predictability*. Springer-Verlag, Berlin, 1993, p.173-204 ).

**Leckie, F. A. and D. R. Hayhurst 1977.** Constitutive equations for creep rupture. *Acta Metallurgica*, 25:1059-1070.

**Lomnitz, C. 1997.** Fundamentals of earthquake prediction. New York: John Wiley and Sons, 326 p.

**Molchan, G. 1999.** Some statistical problems in earthquake prediction and seismicity. *Fifth Workshop on Non-Linear Dynamics and Earthquake Prediction*. Trieste, 4-22 October 1999: The Abus Salam International Centre for Theoretical Physics.

**Mendecki, A. J. 1996.** Quantitative seismology and rockmass stability. (In: *Mendecki, A. J. Seismic Monitoring in Mines*. London: Chapman and Hall, London, 1997, p178-219).

**Mendecki, A. J. 1996.** Seismology for rockburst prevention, control and prediction. *SIMRAC report GAP017*. Pretoria. Department of Minerals and Energy.

**Mendecki, A. J. and G. van Aswegen 1997.** Mine layout, geological features and seismic hazard. *SIMRAC mid year report GAP303*. Pretoria. Department of Minerals and Energy.

**Mendecki, A. J. and G. van Aswegen 1997, 1998.** Mine layout, geological features and seismic hazard. *SIMRAC mid year report GAP303*. Pretoria. Department of Minerals and Energy.



**Mendecki, A. J., G. van Aswegen and P. Mountford. 1999.** A guide to routine seismic monitoring in mines. (In: *Jager A. J. and J. A. Ryder. A handbook on rock engineering practice for tabular hard rock mines.* Safety in Mines Research Advisory Committee, Johannesburg. Cape Town: Creda Communications, 1999, p287-309).

**Newman, W. I., D. L. Turcotte and A. Gabrielov 1996.** A Hierarchical model for precursory seismic activity. (In: *Rundle, Turcotte and Klein. Reduction and predictability of natural disasters, SFI studies in the sciences of complexity XXV.* Addison-Wesley, 1996, p.243-260).

**Newman, W. I., A. M. Gabrielov, T. A. Durand, S. L. Phoenix, D. L. Turcotte 1994.** An exact renormalisation model for earthquakes and material failure. Statics and dynamics. *Physica D*:200-216.

**Ouillon, G. and D. Sornette 1998.** Testing the degree of predictability of mine rockbursts from improved time-to-failure analysis. *Report commissioned by ISS International Limited.* Welkom: ISS International Limited.

**Romachkova, L. L., V. G. Kossobokov, G. F. Panza and G. Costa 1998.** Intermediate-term predictions of earthquakes in Italy: Algorithm M8, *Pure and Applied Geophysics* 152: 37-55.

**Saito, M. 1969.** Forecasting time of slope failure by tertiary creep. *Proceedings of the 7th International Conference on Soil Mechanics and Foundation Engineering*, 2:677-683.

**Saleur, H., C. G. Sammis and D. Sornette 1996.** Discrete scale invariance, complex fractal dimensions and log-periodic fluctuations in seismicity. *Journal of Geophysical Research*, 101:17661-17677.

**Servi, I. S. and N. J. Grant 1951.** Creep and stress rupture behaviour of aluminium as a function of purity. *Transactions of the American Institute of Mining and Metallurgical Engineers*, 191:909-916.

**Sornette, D. and L. Knopoff 1997.** The paradox of the expected time until the next earthquake. *Bulletin of the Seismological Society of America*, 87: 789-798.

**Spottiswoode, S. M. and W. de Beer. 1999.** A review of current seismic prediction and hazard assessment practice and capacity. *Deepmine final report task 5.4.1.* Johannesburg. Deepmine Consortium.

**Turcotte, D. L. 1999.** Self-organized criticality. *Reports on Progress in Physics*, 62:1377-1429.

**Varnes, D. J. 1989.** Predicting earthquakes by analyzing accelerating precursory seismic activity. *Pure and Applied Geophysics*, 130:662-686.

**Varnes, D. J. 1987.** Foreshock seismic energy release functions: Tools for estimating time and magnitude of main shocks. *U. S. Geological Survey Open-File Report 87-429*, 39p.

**Yeomans, J. M. 1992.** Statistical Mechanics of Phase Transitions. Oxford: Oxford Science Publications, 1993, p22-25.

## **Appendix 1 Pattern recognition: Non-parametric statistics questionnaire**

An example of the initial questionnaire is supplied in Figure A1.1, with numerical thresholds referring to a particular area.

Note that question 1 refers to general instability of an area, as evidenced by foreshock activity. Questions 2 asks about accelerated coseismic deformation, while 3 questions the (change in) stress regime, with 4 referring to softening. Questions 5 and 6 apply a median filter to the energy index to highlight the role, if any, of outlying events. Questions 7, 8, 9 and 10 consider fluctuations in the stress environment. The Seismic Schmidt number and its components, seismic strain, stress, inter-event distance and inter-event time, are questioned in questions 11-20. Note that the average inter-event distance properly normalised, would be an indicator of clustering, while inter-event time measures activity rate.

Energy index A and B calculated using 4Di Energy Index option

Events with local magnitude 2

- 1 = Magnitude  $\geq 1.7$  event 54 hours before forecast point?  
Minimum number of events = 5
- 2 = Overall change in Cumulative Apparent Volume  $> 0.003 \text{ km}^3$  during 120 hours before forecast point, given existing background of  $0.003 \text{ km}^3$ ?  
Minimum number of events = 5
- 3 = Overall change in absolute value of  $\text{Log}_{10}\text{EI} \geq 0.25$  during 114 hours before forecast point?  
Minimum number of events = 5
- 4 = Drop in value of  $\text{Log}_{10}\text{EI}$  114 hours before forecast point, when change in absolute value of  $\text{Log}_{10}\text{EI} \geq 0.25$ ?
- 5 = Overall change in absolute value of  $\text{Log}_{10}$  MEDIAN EI  $\geq 0.25$  during 114 hours before forecast point, with median EI calculated with 48 hour, 5 event filter?
- 6 = Drop in value of  $\text{Log}_{10}\text{EI}$  114 hours before forecast point, when change in absolute value of  $\text{Log}_{10}$  MEDIAN EI  $\geq 0.25$ ?  
Filter as above.
- 7 = Any change in absolute value for  $\text{Log}_{10}\text{EI} \geq 0.25$  during 114 hours before forecast point?  
Minimum number of events = 5
- 8 = Any drop in value of  $\text{Log}_{10}\text{EI}$  114 hours before forecast point when change in absolute value of  $\text{Log}_{10}\text{EI} \geq 0.25$ ?
- 9 = Any drop in value of  $\text{Log}_{10}$  MEDIAN EI during 114 hours before forecast point with median EI calculated with 48 hour, 5 event filter?
- 10 = Any drop in value of  $\text{Log}_{10}$  MEDIAN EI 114 hours before forecast point, when change in absolute value of  $\text{Log}_{10}$  MEDIAN EI  $\geq 0.25$ ?  
Filter as above for EI-.
- 11 = Overall drop in  $\text{Log}_{10}$  Seismic Schmidt number  $\geq 1$  during 42 hours before forecast point?  
Minimum number of events = 5
- 12 = Any drop in  $\text{Log}_{10}$  Seismic Schmidt number  $\geq 1$  during 42 hours before forecast point?
- 13 = Overall increase in Seismic Strain  $\geq 0$  during 42 hours before point of forecast?
- 14 = Overall decrease in Seismic Strain  $\geq 0$  during 42 hours before point of forecast?
- 15 = Overall increase in average inter-event distance  $\geq 0$  during 42 hours before point of forecast?
- 16 = Overall decrease in average inter-event distance  $\geq 0$  during 42 hours before point of forecast?
- 17 = Any increase in Seismic Strain  $\geq 0$  during 42 hours before event point of forecast?
- 18 = Any decrease in Seismic Strain  $\geq 0$  during 42 hours before point of forecast?
- 19 = Any increase in average inter-event distance  $\geq 0$  during 42 hours before point of forecast?
- 20 = Any decrease in average inter-event distance  $\geq 0$  during 42 hours before point of forecast?

\*\*\*\*\*FILTERS\*\*\*\*\*

Energy Index: 48 hours, 5 events  
Schmidt number, strain rate and viscosity: 42 hours, 0 events  
In all cases, 1 = 'yes', -1 = 'no', 0 = cannot compute (usually insufficient data for background)

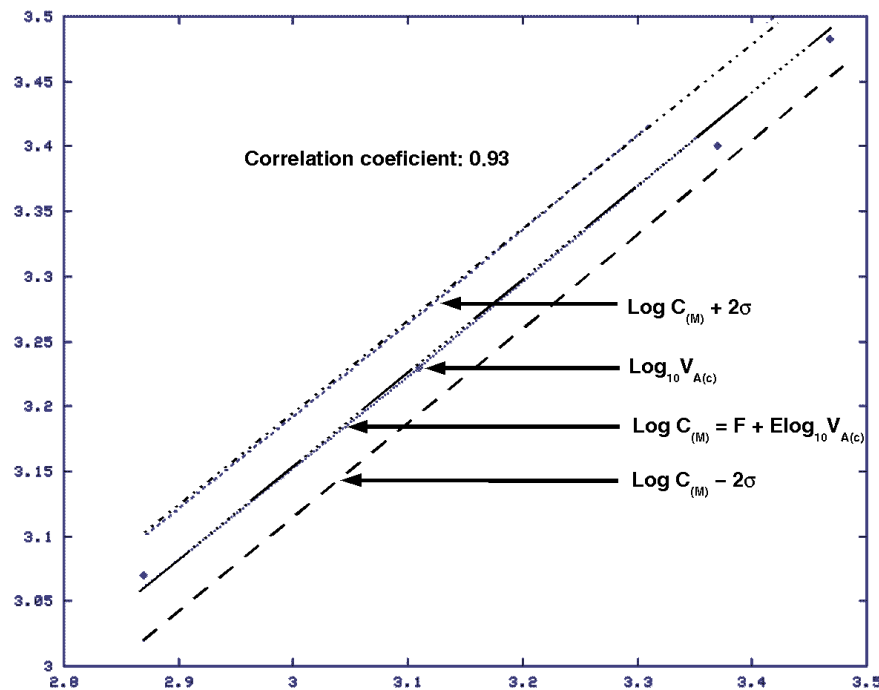
**Figure A1.1**

## Appendix 2 Prediction learning phase and error analysis

### A2.1 $\text{Log}_{10}C$ vs $\text{Log}_{10}(\text{Release})$

An empirical relationship exists between the constant  $C$  and the final release in the power law time to failure equation (A2.1.1). This linear relation is exhibited for the apparent volume in a West Wits pillar extraction in figure A2.1.

$$\sum_{t'=t_0}^t \Omega_{t'} = \sum_{t'=t_0}^{t_c} \Omega_{t'} - C(t_c - t)^\alpha \quad (\text{A2.1.1})$$



**Figure A2.1** The constant  $C$  in the time to failure equation fitted to characteristic events with known power law precursory time series.

Since this empirical relationship enables one to estimate  $C$  in terms of  $\Omega_{t_c}$  by

$$C = 10^{F + E \text{Log}_{10} \Omega_{t_c}}$$

the number of unknowns in the inversion are reduced to three. The function  $F$  to be minimised then becomes

$$\min F = \left[ \sum_{t'=t_0}^t \Omega_{t'} - \sum_{t'=t_0}^{t_c-1} \Omega_{t'} - \Omega_{t_c} + 10^F 10^{\beta \log_{10} \Omega_{t_c}} (t_c - t)^\alpha \right]^{L_1} \quad (\text{A2.1.2})$$

## A2.2 Characteristic release

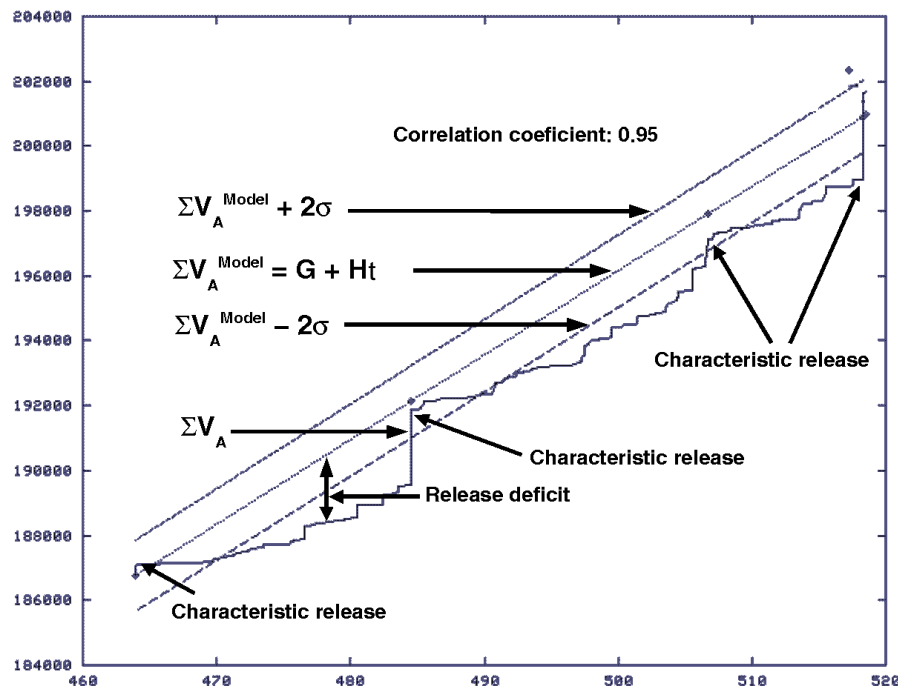
A further simplification is introduced by the concept of characteristic release in a given volume of rock. It should be noted that, in the context of mining induced seismicity, this project found that the characteristic release is not necessarily quantified by the largest single events in an area of interest. Since cumulative functions are used, the rockmass response immediately after blasting can be sufficiently large when cumulated to qualify as “characteristic”. Figure A2.2 below illustrates the concept further. The proposed explanation is the following: the driving force (excavation) is applied on a time scale which is immensely large (days and weeks) compared to the time scale of a seismic event (microseconds). This is a situation which leads to self organisation in the dynamic of rock mass response, exemplified by the Gutenberg-Richter relation. The essential idea to grasp is that the scaling regime (straight line section) of the Gutenberg - Richter relation is just a phase in the evolution of the rock mass under the influence of a slow driving force. Also evident in the Gutenberg-Richter relation is a roll-off at high magnitudes (Figure A4.2.1) which is usually taken to be a signature of either a finite size effect, i.e., there is a dearth of events which involve the deformation of a volume of rock comparable to the physical size of the volume under consideration, or of an incomplete catalogue.

An alternative explanation currently popular in the literature (*Turcotte 1999*) is that the system properties and dynamics themselves play a role in the location of the roll-off point. In a complex, disordered material, the stress distribution is not uniformly perturbed by the driving force. Initially a random distribution of small concentrations of excess stress (clusters of stress excess) are created and energy release from these clusters is small. However, the release itself is not uniform. This leads to the scaling regime of the Gutenberg-Richter relation, in which events of all sizes (up to an intermediate maximum, where the roll-off initiates) are possible at all times. This can be viewed as a self-organising phase during which, if the driving force is slow enough, stress is transferred to larger and larger clusters, a well understood and quantified process in models of self-organised criticality such as forest-fire and sandpile models (*Turcotte 1999, Boonzaaier 2000*). It takes time for a sufficiently large “release deficit”, which can only be normalised by large events, to build up. It has been shown for numerical models (*Turcotte 1999*) that a percolation threshold exists where, suddenly, the response of the system changes to the creation of system-spanning clusters of built-up stress, released by large events. This transition takes place at the start of the roll-off of the Gutenberg-Richter distribution. The events that now appear are “characteristic” events, since the stress clusters which trigger them are characteristic of the material properties and size of the volume of interest.

The existence of episodes of system-wide cascades of stress-release, punctuating periods of self-organisation is good news: it endows the seismic rock mass response with a measure of predictability, since these episodes wipe the slate clean, destroying the essentially unpredictable self-organised, perhaps critical, state. There are indications that the periods between characteristic releases are, indeed critical: In most cases (see chapter 4) the cumulative release follows a *power law in the time to failure* (failure defined as the occurrence of a characteristic episode). Figure A2.2 illustrates (for real data from a West Wits pillar

extraction) several characteristic releases, connected by periods when the system self-organises to the point of no return. Note that, clearly, a “release deficit” builds up in the phase when energy that is put into the system by the driving force is utilised for internal self-organisation into larger and larger clusters of stress. When the system saturates, most of the energy is released in a characteristic release.

Apart from illustrating the current thinking on the underlying physics of rock mass response to mining, the concept of power law self-organisation punctuated by characteristic release also provides an avenue towards enhancing a predictor. It is possible, if a good enough fit inspires sufficient confidence, to extrapolate the release deficit in order to give the inversion routine a first estimate of the final (“failure”) release.



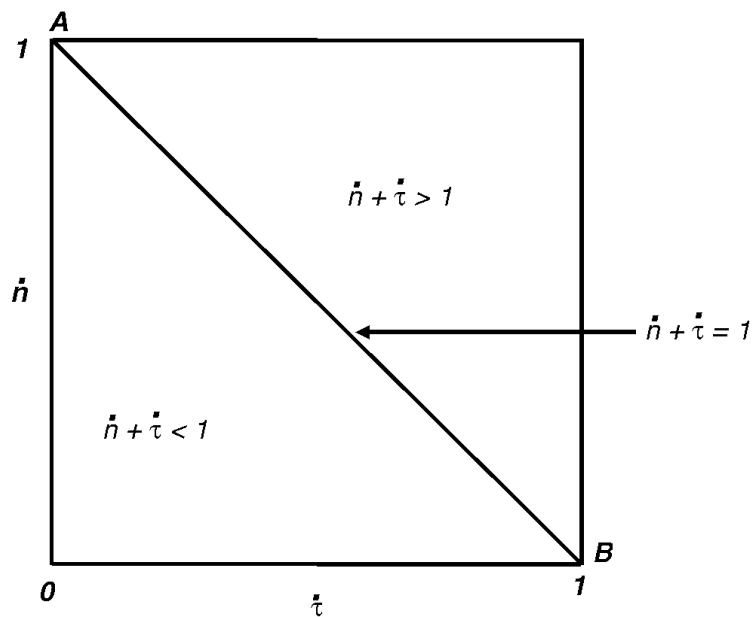
**Figure A2.2 Cumulative characteristic seismic release in a shaft pillar of a West Wits mine by four events with  $\log_{10}M_0 > 11.5$  fitted to a straight line. The outer lines represent two standard deviations either side of the empirical (fitted) mean release.**

## A2.3 Error diagrams

A sequence of strong seismic events (or seismic release episodes) can be viewed as stationary point processes (Molchan 1999). This author used two prediction parameters or errors to compare various prediction strategies:

1. **The rate of failures-to-predict,  $\hat{n}$ :** the number of false negative predictions divided by the total number of characteristic events in the time interval under consideration;
2. **The rate of time alarms,  $\hat{\tau}$ :** the total time under ALERT, divided by the time interval under consideration.

Keeping track of these two parameters, one can plot them, for a given prediction strategy, on an error diagram (Figure A2.3). Denote  $G = \{(\dot{h}, \dot{\tau})_{\pi}\}$  by the set of errors corresponding to various prediction strategies  $\pi$ . The set  $G$  belongs to the unit square ( $0 \leq \dot{h}, \dot{\tau} \leq 1$ ). The point  $A = (0,1)$  is known as the “optimist strategy”, i.e., no alerts are declared, while the point  $B = (1,0)$  is known as the “pessimist strategy” (the whole time span of the catalogue is under alert). Points on the line  $AB$  or  $\dot{h} + \dot{\tau} = 1$  correspond to a random guesses in which alarms are declared with probabilities independent of the available information. All points with  $\dot{h} + \dot{\tau} < 1$  correspond to non-trivial strategies, better than random.



**Figure A2.3** Generic error diagram with points  $A$  (optimist strategy),  $B$  (pessimist strategy),  $AB$  (random guess).

## Appendix 3 Deterministic component of seismic time series, noise reduction and limits of predictability

### A3.1 False nearest strands

The false nearest strands technique (fns) can assess the residual random component in a nonlinear time series. After applying the fns technique to a number of areas, the random component - as measured by the fns - ranged between 39 per cent and 44 per cent which in turn indicates that it is likely that the deterministic component is not above 61 per cent. To deal with this apparent problem, a noise reduction in the signals of interest which would preserve as much as possible of the information content of the original signal, could enhance the deterministic component. Hence the current prediction techniques (i.e. time to failure) may become more successful if applied to the “denoised” signals instead of the original ones.

### A3.2 Noise reduction

Denote by  $X(t)$  the time series of a seismic parameter (intermittent), for example  $\log E$ . Instead of expanding in Fourier modes - as it is often done in traditional signal processing - we expand  $X(t)$  in terms of another complete set of functions. This set is established by analysing the distribution of states in the phase space and is not imposed from the outside, as in the Fourier series representation. So we can write:

$$X(t) = \sum_{m>0} u_m(t) \quad (\text{A3.2.1})$$

where the functions  $u_m(t)$  form a complete orthogonal set. We can now define

$$P_m = \left( \int X(t) u_m(t) \right)^2 \quad (\text{A3.2.2})$$

as the probability of the system of being in the state  $u_m$ . Now, as in the case of Fourier modes expansion, we can anticipate that there may be situations where  $X(t)$  can be sufficiently accurately represented by only a *finite* number of functions  $u_m$ . Then, just as a few peaks in the usual power spectrum indicate the possibility of an expansion in a small number of Fourier modes, the existence of a few dominant values for  $P_m$  indicates the possibility of a useful expansion in a small number of the  $u_m$  functions.

SVD is a method for the generation of an orthogonal set of functions. The essentials of the method are as follows. Once the embedding of the original signal  $X(t)$  was done (using false nearest strands) and the dimension of the attractor is established, trajectories can be expanded locally: Let  $X_n = X(t_n)$ . A set of M-dimensional vectors  $V_e$  defined by

$$V_e = \{X_e, X_{e+1}, \dots, X_{M+e-1}\} \quad (\text{A3.2.3})$$

is constructed, and then the autocorrelation function defined by



$$C(n) = \sum_{e=1}^n X_e X_{e+n} \quad (\text{A3.2.4})$$

Finally, the symmetric  $M \times M$  correlation matrix  $W$  with elements  $W_{ep} = C(|e-p|)$  is constructed.

The matrix  $W$  has  $M$  eigenvalues  $w_m$  and corresponding eigenfunctions  $v_m$ . The function  $u_m$  is defined by

$$u_m(t_n) = v_m \cdot V_n \quad (\text{A3.2.5})$$

The functions are orthogonal and normalized such that

$$\frac{1}{N} \sum_{e=1}^N u_m^2(t_e) = w_m \quad (\text{A3.2.6})$$

This also means that

$$P_m = w_m^2 \quad (\text{A3.2.7})$$

in the discrete limit. In this way,  $w_m$  is a measure of the importance of the mode  $u_m$  in the expansion of the signal  $X(t)$ . By analogy with the usual procedure of expanding in a finite number of Fourier modes, expand in a finite number, say  $d$ , of values for  $u_m$  and select those functions  $u_m$  that correspond to the  $d$  largest values of  $w_m$ . Note that this really comes down to picking out the most likely states for the system in  $\{u_m\}$  space. Thus, as an approximation to the original signal  $X(t)$ , one can write:

$$X(t) \text{ approx. } \sum_{m=1}^d u_m(t) \quad (\text{A3.2.8})$$

Now  $d$  has to be determined. If the embedding dimension is known,  $d$  must always be a number smaller than the embedding dimension, desirably the local dimension. Here,  $d$  has been taken as the dimension of the largest sphere which could wholly be contained by the attractor in the phase space. Another requirement is that, when a given state in the phase space is visited, a small sphere of radius equalling one per cent of the attractor's size is constructed around that state. If a prescribed number of neighbouring states fall within that sphere (e.g. a minimum of three neighbours) the SVD procedure is applied, otherwise one skips that state in the phase space. In other words, if a state has less than a prescribed minimum neighbours in the sphere, it is rejected as noise, or an outlier.

Once the signal has been SVD-denoised it is subjected to the fns algorithm to be re-embedded. The dimension of the "new" attractor is recalculated and the new level of the random component observed.

### A3.3 Limits of predictability

The limits of predictability were calculated using cspX. Once the temporal horizon is determined, the spatial horizon can be calculated using the following formula:

$$\text{spatial horizon} = \left( \frac{\text{temporal horizon}}{\text{average time between events}} \right) \text{average distance between events} \quad (\text{A3.3.1})$$

## Appendix 4 Larger seismic events as critical phenomena

### A4.1 Criticality: A common example

The example of a critical system most often quoted is a ferromagnet in an external magnetic field (Yeomans 1992). What has to be born in mind in the following description is that the ferromagnet undergoes a continuous phase transition from zero magnetisation to a magnetised phase. This is to be distinguished from a discontinuous phase transition as is found at the boiling / condensation point of a liquid / gas (the liquid / gas phase transition) or the freezing / melting point of a solid / liquid (solid / liquid phase transition).

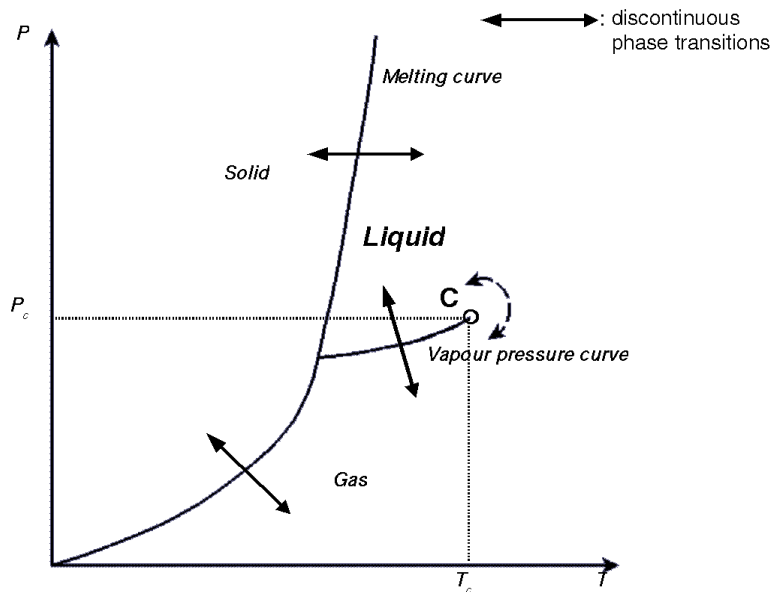
Let us first introduce a few terms:

#### A4.1.1 Correlation length

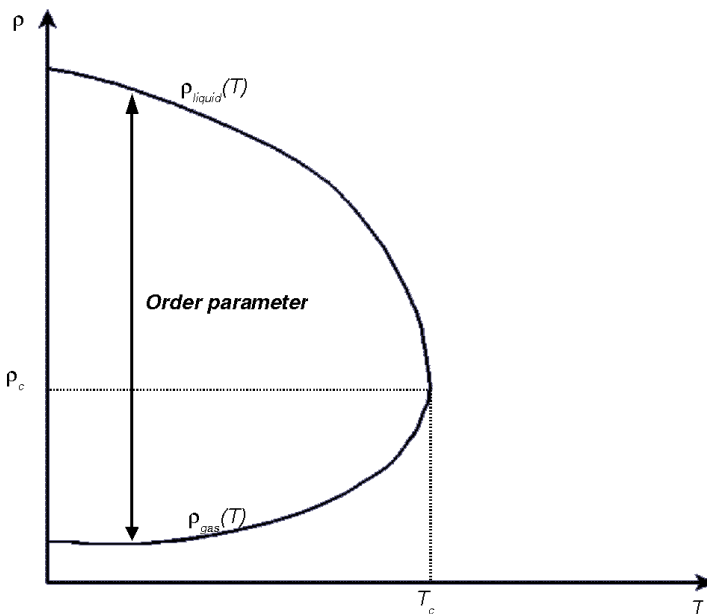
The length scale at which the properties of a subdivision of a material begin to differ from those of a bulk sample gives a measure of the *correlation length* of the material. For example, far from its boiling point, the molecules of a liquid perform thermal vibrations on the scale of a few molecule diameters. If one were to zoom in to this scale, the liquid would appear extremely agitated. Zooming out, a point would be reached at which the visible portion of the liquid would appear to be at rest. Think of a cup of coffee, where the thermal motion of the molecules are not visible when one looks at the cup as a whole. As external conditions are varied (say the temperature starts approaching boiling point), the correlation length increases, until it encompasses large regions, visible to the eye. Again, think of, say, water where, close to the boiling point, individual, unconnected, regions may be boiling (and in fact this is visible as bubbles of vapour rise to the surface) while others are not. In these bubbles, the molecules' motion is correlated and they act in concert. At the boiling point, the correlation length becomes "effectively infinite", in other words, spanning the whole system (cup of water). The motion of the molecules are correlated all through the container.

#### A4.1.2 Driving (tuning) parameter

A driving parameter is the variable which determines whether a material moves to or from a phase transition. In the case of both a liquid and a ferromagnet, the driving parameter is the temperature. In mining-induced seismicity, the removal of rock perturbs the stress state.



(a)



(b)

**Figure A4.1 Phase diagram of a typical liquid. (a): A slice through the  $\rho$ - $T$  plane. The lines are values of  $P$  and  $T$  for which the liquid undergoes discontinuous (in the order parameter) phase transitions. The point  $C$  is the critical point beyond which it is possible to move continuously from a liquid to a gas and back. (b): A slice through the  $P$ - $T$  plane. The curve is the vapour pressure curve, representing values of the densities of the coexisting liquid and gas (after Yeomans 1992).**

### A4.1.3 Order parameter

An order parameter characterises the phase a material is in. In the case of a liquid, the order parameter is the difference  $\rho_{\text{liquid}}(T) - \rho_{\text{gas}}(T)$  between the densities of the liquid and gas, while for a ferromagnet it would be the magnetisation. The order parameter also characterises the type of phase transition a material experiences.

### A4.1.4 Phase transitions

Two types of phase transition will be briefly touched on here. Both derive their names from the behaviour of the order parameter.

#### A4.1.4.1 Discontinuous phase transitions

Figure 4.1 depicts the familiar phase diagram of a fluid. Recall that it is usual to show a three-dimensional diagram with the pressure (P), volume (V) (alternatively, density ( $\rho$ ) and temperature (T) on the mutually orthogonal axes. In figure 4.1(a), only the PT plane is shown. The lines represent values of temperature and pressure for a given constant density along which the familiar *discontinuous phase transitions* (boiling, condensation, freezing, sublimation) take place. These phase transitions are called *discontinuous* because the order parameter, density, changes abruptly (infinitely quickly) when the material is taken from one phase to the other.

#### A4.1.4.2 Continuous phase transitions

The point C in figure 4.1 is known as the critical point of the material. Beyond C, the material can be taken from a liquid to a gas and back while the order parameter (the density difference between the liquid and the gas) changes continuously (smoothly). Note that the critical point corresponds to very specific values of the temperature, pressure and density,  $\rho_c, P_c, T_c$ .

At the point C the material is in a critical state.

### A4.1.5 Critical point: scale invariance

The critical point is characterised by a property which is of crucial interest in this report: the degrees of freedom of the system are correlated at all scales (in the case of a liquid, the macroscopic behaviour on a scale of centimetres is correlated with the furious motion at interatomic ( $10^{-8}\text{cm}$ ) scale, at all locations in the bulk of the liquid).

As the temperature of the material approaches the critical temperature,  $T_c$ , correlation lengths and times grow until, at the critical point, they span the system. The system-spanning correlation length at the critical point means that, as one subdivides the system, or looks at it with increasing magnification, the observed process does not change. For a liquid, the vigorous motion of molecules is the same at all places in the bulk, while the collective motion of a large quantity of molecules reflects this at all scales. This is to be contrasted with the case far from criticality, where the behaviour of molecules is not reflected in the collective behaviour of the liquid.

At the critical point, the behaviours at different scales are larger or smaller (however measured) copies of behaviour at other scales. This phenomenon is called scaling, or the scale invariance of the process.

## A4.1.6 Example: Ferromagnetic continuous phase transition

The magnetic field of a material in a material is the sum of the magnetic dipoles of its electrons, induced by their spin. What precisely a magnetic dipole is, is not of particular importance here. What is important is that a material has its own magnetic field only when more than half of its electron spins “point” in the same direction. At temperatures far above the critical point, the spins in a potentially magnetic material have random orientations, since there is a dominating thermal fluctuation of electrons and atoms, adding up to an overall zero field. As the temperature is lowered, the effects of interactions between spins become more significant so that they begin to align themselves with each other. Initially isolated clusters of parallel, correlated, spins appear. These clusters grow as the temperature is lowered. The size of the largest clusters is a measure of the correlation length. It is important to realise that the fluctuations of spins on smaller scales still remain important - there are correlated regions of spins on all length scales up to the correlation length.

At the critical temperature itself, the correlation length becomes infinite. Of course, in practice, the correlation length, at the critical point, spans the sample, with ordered structures existing on every length scale. Below the critical temperature the sample will have a non-zero magnetisation, that is, more spins will point in one direction than in the other. One says that long-range order now exists in the sample. At zero temperature the all the spins are aligned in the same direction.

This process can be repeated in the opposite direction: increasing the temperature unfreezes the long-range order and now the correlation length measures the size of the biggest clusters of disorder. Of particular interest to this part of the report is the result that the correlation length,  $\xi$ , behaves as  $\xi \sim |T - T_c|^{-\nu}$  where  $\nu > 0$ , i.e., stating that the correlation length *scales* with the temperature as the critical temperature is approached. To see this, note that if the system is initially at a temperature  $T_1$ , the correlation length is  $\xi_1 \sim |T - T_c|^{-\nu} \equiv |\Delta T_1|^{-\nu}$ . At a different temperature,  $T_2$ , the correlation length becomes  $\xi_2 \sim |T_2 - T_c|^{-\nu} \equiv |\Delta T_2|^{-\nu} = |\lambda \Delta T_1|^{-\nu} = (\lambda)^{-\nu} \xi_1$ . In other words,  $\xi_2$  differs from  $\xi_1$  by a scale factor  $(\lambda)^{-\nu}$ . The physical processes at the two scales, are, however still the same.

## A4.2 Seismicity and scaling

### A4.2.1 Gutenberg-Richter

The Gutenberg–Richter magnitude frequency distribution

$$\log_{10} N_{> m_M} = a_{10} - b_{10} m_M \quad (\text{A4.2.1})$$

is the best-known example of such a scaling phenomenon. It is worth the time revisiting this relationship, casting the light of criticality on it.

Since  $m_M = \frac{2}{3} \log_{10} M - 6.1$ , where  $M$  is the seismic moment of an event of moment-magnitude  $m_M$ ,

$$\log_{10} N_{>m_M} = [a_{10} + (6.1)b_{10}] - \frac{2b_{10}}{3} \log_{10} M$$

or

$$N_{>m_M} = 10^{a_{10} + (6.1)b_{10}} M^{-\alpha} \quad (\text{A4.2.2})$$

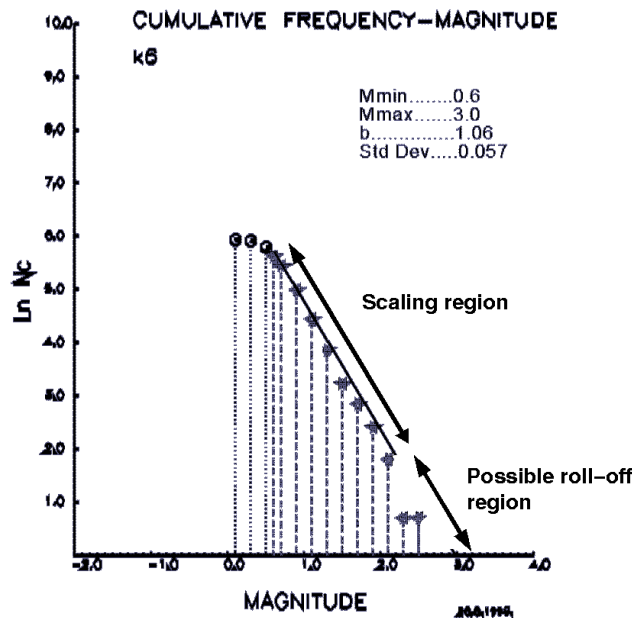
where  $\alpha \equiv (2b_{10})/3$ .

Note that  $\alpha$ , a so-called critical exponent, and  $a_{10} + (6.1)b_{10}$ , are characteristic of the area and time interval in which moments  $M$  were measured.

The Gutenberg–Richter relation is an incredibly strong statement:

It claims that all one has to do to establish the expected number of events  $N_{>m_{M2}}$  with moment-magnitude larger than or equal to  $m_{M2}$  in an area, given that  $N_{>m_{M1}}$  is known, is to take the ratio  $M_2/M_1$  (that is, of course, up to the roll-off point). In other words, the cumulative frequencies of events of different scales or magnitudes are *correlated*.

To test this statement, consider Figure A4.2.1, which shows a typical Gutenberg-Richter plot. Note that the  $b$ -value of 1.06 refers to a  $\log_{10} N_{>m_M}$  graph, whereas  $\log_e N_{>m_M}$  is plotted for the sake of a clearer picture.



**Figure A4.2.1**

## A4.2.2 Time to failure

### A4.2.2.1 Scale invariant processes and power laws

Let the patterns in square 3 in Figure 4.3.1 represent some “physical process” which is a function of the length,  $l$ . In other words, the process is some power of  $l$ :

$$\text{"Process"} = Kl^n \quad (\text{A4.2.2.1})$$

here  $K$  is a constant (which can be one).

Scale invariance of the process means that there exists a constant,  $\mu$ , such that the process remains the same if one magnifies the square, in the sense that one still sees the same process, just more of it. Recall the example of boiling milk, for visualization purposes.

Thus,  $l$  is magnified to  $\lambda l$ , and the “NEW PROCESS” is  $\frac{K}{\mu} \lambda^n l^n$ . Scale invariance means  $\lambda^n / \mu = 1$ . The quantity  $\mu$  is called a scaling factor.

Equation (A4.2.2.1) is called a power law, and is the only form of mathematical relation that can exhibit scale invariance.

### A4.2.2.2 Scale invariance in time

*Varnes (1987, 1989)* fitted a measure of seismic energy release in seismogenic regions to a time to failure curve of the form where  $\dot{\Omega}_t$  is the instantaneous rate of change of a measure of seismic release at time  $t$ . The arguments for the applicability of (A4.2.2.1) can be found in *Varnes (1987, 1989)* and *de Beer and Mendecki (1997)*.

$$\dot{\Omega}_t = C(t_c - t)^{-z} \quad (\text{A4.2.2.2})$$

The time to failure equation used in this report is, up to a counter term, equivalent to the integrated form of (A4.2.2.2.1).

*Scale invariance* in time of the underlying process  $Q$  giving equation (A4.2.2.2.1) means that, upon rescaling the time to failure by, say,  $\lambda$  (where  $\lambda$  is a number,  $0 < \lambda \leq \infty$ ) there exists a number,  $\mu$ , such that

$$\sum_{t'=t_0}^t \Omega_{t'} \equiv Q(t_f - t) = \frac{1}{\mu} Q(\lambda(t_c - t))$$

or

$$\frac{K}{(t_c - t)^\alpha} = \frac{1}{\mu} \frac{K}{(\lambda(t_c - t))^\alpha}$$

$$\sum_{t'=t_0}^t \Omega_{t'} = \frac{K}{(t_c - t)^\alpha} \quad (\text{A4.2.2.2.2})$$



so that

$$1 = \frac{1}{\mu\lambda^\alpha} \quad \text{or} \quad \mu = \lambda^{-\alpha}$$

or

$$\alpha = -\frac{\log\mu}{\log\lambda} \quad (\alpha \text{ real}) \quad \text{(A4.2.2.3)}$$

The parallels with the case of the Gutenberg-Richter law should be obvious.

### A4.2.2.3 What does scale invariance in time mean?

Processes become scale invariant (whether in time, space or, in the case of seismicity, in magnitude) at some critical point. In this stage of the evolution of the rock mass, its response to mining is predominantly a cooperative phenomenon, driven by repeated interactions between smaller ruptures which “phase up” (*Bowman et al 1998*) to create a macroscopic self-similar state. This is when one finds the appearance of many different scales. Since this is a hierarchical process, every large event is essentially a critical point to all the lower levels. These larger events tend to move the system away from criticality, or “resets” the system. The scale invariance in the time to failure is then the signature of the imminence of criticality, or the approach to criticality. Literally, the critical point is “smeared out” in time.

## A4.3 Renormalisation group theory

In this section, the bare bones of the machinery used to move between different scales in a critical system is sketched. The mathematics of this procedure is called “renormalisation group theory”. The name comes from the fact that one does not *normalise* the whole process to a single length and time scale, but rather normalises repeatedly (renormalises) between different scales. A prerequisite for this is scale invariance, since otherwise different scales would be uncorrelated (small seismic events have nothing to do with one another), and there would be no way to move from one scale to another.

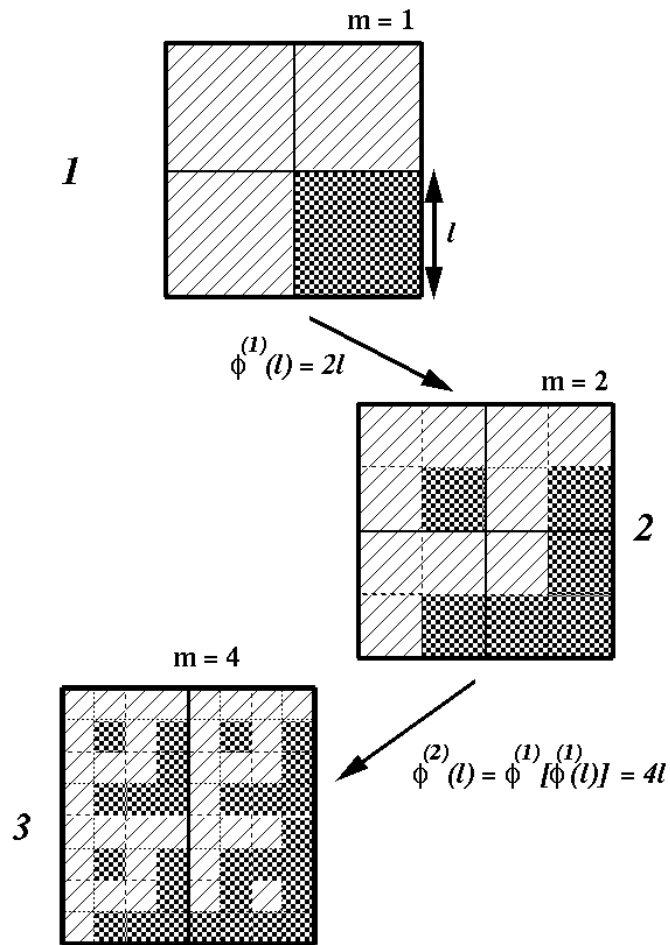
Renormalisability of a process is extremely important if inferences about large-scale behaviour are to be made from knowledge gained at smaller scales.

Consider Figure A4.3.1. The three squares represent a single object at different magnifications. A magnification is defined as a “zooming in” process which makes the object apparently double in size. This “zooming in” is called the *renormalisation group flow* and is represented in the figure by  $\Phi^{(l)}(l) = (\lambda l)l$ , where  $l$  is some chosen length scale. Due to space limitations, the square is rescaled, which is why successive magnifications appear the same in size.

Notice how, starting at magnification  $m = 1$ , the deeper one looks into the structure, the clearer the self-similar, scale invariant structure of the square becomes. The three original “light” squares resolve at  $m = 2$  into three exact, smaller copies of the large square at  $m = 1$ , while the dark square resolves into a kind of “negative” of the original.

Increasing the magnification once more, reveals that the structure of the light and dark squares again repeats itself at a smaller scale. Of course, as one moves the other way (from

$m = 4$  to  $m = 1$ , say), the square loses definition. To compensate for this, the mathematics of the process gets additional “fudge factors”. These renormalisation constants carry extremely important information about the nature of the physics (such as where critical points are).



**Figure A4.3.1**

# Appendix 5 Contractual project details

SIMRAC 1

**DEPARTMENT OF MINERALS AND ENERGY  
PROPOSAL FOR A PROJECT TO BE FUNDED IN TERMS OF THE MINERALS ACT**

- CONFIDENTIAL -

**DME REFERENCE NUMBER**

TO BE SUBMITTED BY 12:00 NOON ON THE CLOSING DATE

**(FOR OFFICE USE ONLY)**

### 3. **PROJECT SUMMARY:**

**PROJECT TITLE:** SEISMOLOGY FOR ROCKBURST PREDICTION  
(Extension of project GAP017 "Seismology for rockburst prevention, control and prediction)

**PROJECT LEADER:** Dr W de Beer

**ORGANIZATION:** ISS International

**ADDRESS:** P O Box 2083, Welkom 9460

**TELEPHONE:** 057 912-9111      **TELEFAX:** 057 912-2652      **EMAIL:**

**PRIMARY OUTPUT<sup>1</sup>:** Method for the prediction of rock mass instability based on kinematics of failure, nonlinear dynamics (chaos theory) and neural networks, using data supplied by the modern seismic networks in South African mines.

**HOW USED?<sup>2</sup>:** The developed procedures will be transferable to different seismic systems in South African mines for the real-time prediction of rockbursts in areas of interest.

**BY WHOM?<sup>3</sup>:** Rock mechanics engineers, mine seismologists and mine management.

**CRITERIA FOR USE<sup>4</sup>:** Manageable success rate of at least 50% of proposed methods in terms of predicting rock mass instabilities.

**POTENTIAL IMPACT<sup>5</sup>:** Reduction of accidents resulting from rockbursts and rockfalls.

**DURATION (YY/MM)**                      1997/01    **TO**                      1999/12

#### 4. PROJECT DETAILS

##### 2.1 PRIMARY OUTPUT<sup>1</sup>

Method for the prediction of rock mass instability based on kinematics of failure, nonlinear dynamics (chaos theory) and neural networks, using data supplied by the modern seismic networks in South African mines.

##### 2.2 OTHER OUTPUTS (deliverables)<sup>6</sup>

The developed methodology will be applicable to the monitoring of the stability of roofs and pillars in coal mines. Exposure to the new ideas of nonlinear dynamics (chaos theory) that is proving to be a new powerful tool in rock mechanics.

##### 2.3 ENABLING OUTPUTS<sup>7</sup>

NO.	ENABLING OUTPUT	MILE- STONE DATE	MAN DAYS
1	<p>Procedures to invert the TIME TO FAILURE from models of kinematics of failure applied to the observed behaviour of various measures of seismic release.</p> <p><i>The <u>time to failure</u> quantifies - for a given volume of interest - the instance when a larger rock mass instability is likely to occur.</i></p> <p>Report on the EFFICIENCY and APPLICABILITY of various measures of seismic release that are used to predict the time of failure.</p>	<p>1997/06</p> <p>1997/12</p>	
2	<p>TIME TO FAILURE - calculated more accurately by applying renormalisation group theory to the most efficient classes of kinematics of failure models.</p> <p>Early versions of PREDICTION METHODS to estimate further values of seismic parameters of interest using nonlinear modelling and neural networks.</p>	<p>1998/06</p> <p>1998/12</p>	
3	<p>ADVANCED PREDICTION TECHNIQUES based on the cooperative use of kinematics of failure, nonlinear dynamics and neural networks.</p> <p>IMPLEMENTATION of the prediction techniques into the real-time seismic monitoring at two different sites, e.g. mini-longwall and pillar.</p> <p>WORKSHOP and TRAINING on the use of rockburst prediction techniques based on kinematics of failure, nonlinear dynamics and neural networks.</p>	<p>1999/12</p> <p>1999/06</p> <p>2000/03</p>	

2.4 **METHODOLOGY<sup>3</sup>**

NO. OF ENABLING OUTPUT	STEP NO.	METHODOLOGY TO BE USED TO ACCOMPLISH THE ENABLING OUTPUT (INDICATE STEPS/ACTIVITIES)
1	1	Comparative studies of phenomenological models that describe the kinematics of failure. The models incorporate the behaviour of various measures of seismic release based on apparent volume, energy index, seismic Deborah number, Schmidt number, etc, before the occurrence of a damaging event and aim at calculating the time to failure. Fit the models of kinematics of failure to available seismic data and invert the time to failure as well as the values of seismic parameters of interest e.g. seismic moment or energy. Case studies.
	2	Establish the predictive power, applicability and efficiency for each model.
	3	Selection of time series of parameters of interest from mining areas using current techniques of monitoring. Test for the deterministic component of each time series using the method of return maps. Noise reduction by local reconstruction of the underlying dynamics. Detecting nonlinearity and chaotic behaviour using classes of generalised correlation integrals. Test for multifractality. Case studies: the determination of spatio-temporal limits of predictability.
2	1	Improving the phenomenological models using procedures of renormalisation group theory, as we consider that the occurrence of larger seismic events is a critical phenomenon that exhibits spatio-temporal scale invariance.
	2	Inverting the time to failure using an increased number of fitting parameters which result from applying the renormalisation group approach. Apply the new technique to concrete cases to establish the performance of the procedure.
	3	Determine the local linear functional forms to estimate the predictor for the time series of seismic parameters. Estimate the future values of the respective parameters and compare to data to determine the model efficiency.
	4	Design neural network architectures in order to investigate the possibility of predicting future values of seismic parameters in real-time.
3	1	Implement a suitable learning strategy for the functionality of the neural networks. Achieve a real-time prediction of seismic events.
	2	Combine the insight gained from the calculation of the time to failure using kinematics of failure models and data with the techniques of nonlinear modelling. Estimate the success rate of the prediction function.
	3	Devise an algorithm aimed at implementing neural networks methods of prediction into the cooperative use of kinematics of failure and nonlinear dynamics to obtain enhanced forecasting accuracy.
	4	Select study sites for the testing and implementation of the rockburst prediction techniques. Achieve real-time implementation of the forecasting methods.

	5	Report on "Transfer of prediction technology to the mining industry". Organise workshops and training sessions involving the use of the newly developed prediction tools.
--	---	---

**Key Facilities and Procedures to be used in the Project**

1. 17 Digital Seismic Networks covering the Regional, mine-wide and shaft/shaft pillar wide areas.
2. Digital 27 channel accelerometer and 4 strain meters network of Kyoto University (Japan), covering longwall operation at WDL South Mine (project sponsored by Kyoto University, International Association of Seismology and Physics of the Earth Interior, WDL south Mine and ISS International)
3. Two clusters of seismic stations covering scattered operations (partly sponsored by GAP017)
  - 3.1 Mining around hazardous dyke
  - 3.2 Mining in the vicinity of intersections of geological features
4. ISS International Seismological Software Library (500 000 lines of C-code)
5. 3D and Stereo Computer Graphics and large data storage facilities at ISSI

## Appendix 6 Extended literature review

The body of literature on the subject of earthquake prediction is big and contentious. In this Appendix, summaries of the contents and results of the most relevant sources (as also summarised in the quarterly reports) are presented.

The survey and discussion are organised in terms of topics, rather than alphabetically, as follows: In the first section, the research that informed the project methodology are discussed. These fall into three categories, namely time to failure accelerating seismic release, seismic events as critical phenomena, criticality and pattern recognition. This section closes with the sources used to devise ways of estimating errors and deciding on whether a forecasting strategy is random or not.

Section 2 presents the sceptics' view of earthquake and mine seismic prediction. With one exception, the articles are contained in the special edition of the Geophysical Journal International devoted to the RAS/JAG Discussion Meeting in London in 1996. Where necessary, comments from the research contractor are included, to provide perspective where it was felt that an author was too selective with examples.

Section 3 focusses on literature with direct relevance to mine seismicity. These are mainly reports and books produced by ISSI and collaborators. Finally, in Section 4, some standard textbooks of a more general nature (both in Physics and Seismology) are listed.

### A6.1 Project Methodologies

#### A6.1.1 Time to failure (accelerated seismic release) and refinements

**Ashby, M. F. and B. F. Dyson 1986.** Creep damage mechanics and micromechanisms. *Advances in fracture research*, 1: 3-30.

##### Summary

See *Leckie and Hayhurst (1977)* below.

**Bowman, D. D., G. Ouillon, C. G. Sammis, A. Sornette and D. Sornette 1998.** An observational test of the critical earthquake concept. *Journal of Geophysical Research*, 103: 24359-24372.

##### Summary

The hypothesis is that seismicity prior to a large earthquake can be understood in terms of the statistical physics of a critical phase transition. In the framework of such an hypothesis, the cumulative seismic strain release increases as a power law in the time to failure before the final event. A model of critical phase transitions predicts that the region of correlated seismicity is much greater than that predicted by a model of simple elastodynamic interactions. The authors use the algorithm employed in Section 4.3 of this report to test for the accelerating seismicity predicted by the critical point model and to identify the region approaching criticality, essentially comparing the cumulative energy (Benioff strain) release and the theoretical power law behaviour.

## Results

21. The authors claim to find critical regions before all earthquakes along the San Andreas system since 1950 with Richter magnitude  $M \geq 6.5$ .
22. The same procedure is performed on a large number of randomly generated synthetic catalogues (as is also done in this report, on mining-induced seismicity) in order to assess the statistical significance of the results. The null hypothesis is that accelerated seismic release, as observed in real earthquakes, could also appear in purely random catalogues as spurious patterns generated by the algorithm. The authors reject the null hypothesis 99.5% confidence.
23. An empirical relation between the logarithm of the critical region radius ( $R$ ) and the magnitude of the final event ( $M$ ) is found, such that  $\log R \propto 0.5M$ , suggesting that the largest probable event in a given region scales with the size of the regional fault network.

**Brehm, D. J. and L. W. Braille. 1998.** Intermediate term earthquake prediction using precursory events in the New Madrid seismic zone. *Bulletin of the Seismological Society of America*, 88: 564 – 580.

## Summary

The predictability of smaller intraplate earthquakes is explored using the time to failure method. The authors use the time-to-failure method described by *Varnes (1989)*, *Bufe and Varnes (1990)* and *Bowman et al. (1998)*. Recall that precursory events (also called foreshocks) are used to define an accelerated seismic release curve. By fitting an equation to the data, a predicted time of failure and magnitude can be calculated. Of further relevance to GAP409 is that the authors use what they term microearthquake network data from the New Madrid Seismic Zone (NMSZ), which is reasonably complete for earthquakes of magnitude  $\geq 1.5$  in the area of interest. They claim that the method has yielded predicted values of past events as small as  $m_b = 3.5$ . The relevance to mining-induced events is obvious.

The network data set used in this evaluation covers the time interval from 29 June 1974 to 20 July 1995 for the NMSZ. There were 36 earthquakes of magnitude  $\geq 3.5$  over the 21-yr period in which the network had been operating. However, not enough precursory data is available for main events that occurred before 1980, leading the authors to utilise the 26 earthquakes that occurred after 1980 with magnitudes  $\geq 3.5$ . Sixteen of the 26 main events were modelled. In most cases, the precursory sequences yielded predicted times of failure and magnitudes that were reasonably close to the actual main event values. The remaining main events, including those occurring before 1980, could not be modelled due to either not enough events to adequately define a precursory sequence, or interfering events that disrupt the accelerated energy release curve. In addition, two events were modelled from the Nuttli catalogue along with one that used a combination of both catalogues. Nineteen earthquakes with magnitudes  $\geq 3.5$  were evaluated using the time-to-failure method.



## Results

1. The first calculation using the time-to-failure method gave predictions with large error bounds and essentially no upper bound on the predicted magnitude.
2. Empirical relationships between parameters (see also Section 3.1.3.2 of this report) serves to constrain the range of the predicted magnitude and, to a lesser extent, the estimated time of failure. These relationships modify the time-to-failure equation and yield predicted values for magnitudes that have an upper limit.
3. Another empirical relationship suggests that the logarithm of the moment of the mainshock increases linearly with the logarithm of the size of the precursory event search diameter (see also *Bowman et al. (1998)*).
4. The relative seismicity of a region influences the optimum search diameter used to find precursory events.
5. An analysis of the occurrence of false-positive acceleration sequences (acceleration sequences that do not end in a main event) was conducted. The preliminary falsepositive analysis was conducted by randomly selecting potential mainshock locations. The results yielded a false-positive acceleration sequence occurrence rate of 2%.

**Brehm, D. J. and L. W. Braile. 1999a.** Refinement of the modified time-to-failure method for intermediate-term earthquake prediction. *Journal of Seismology* 3: pp 121-138.

## Summary

With the reduction in the number of unknown parameters in the modified time to failure method, a scheme of contouring predicted magnitude and time of failure is implemented to improve accuracy. The modified time-to-failure method for intermediate-term earthquake prediction utilises empirical relationships to reduce the number of unknown parameters providing a stable and unique solution set. The only unknown parameters in the modified time-to-failure method are the time and size of the impending main event. The modified time-to-failure equation is used to model the precursory events and a prediction contour diagram is constructed with the magnitude and time of failure as the axes of the diagram. The root-mean-square (rms) error is calculated for each time / magnitude point on such a prediction diagram, representing the difference between the calculated accelerating release and the actual rate of seismic release by precursory events. A small region, corresponding to the low rms region on the diagram, defines the prediction. Time to failure prediction consistently underestimate the final release and over-estimate the time of failure. This is caused by an underestimation of the seismic release in the modified time to failure equation at the very end of a sequence. An empirical correction can be applied to the predicted results to minimize this problem.

The authors developed a main event location search technique for use in conjunction with the modified time-to-failure method. The location technique is used to systematically search an earthquake catalogue and identify the centres of circular areas corresponding to precursory sequences that display accelerated seismic release. It has shown good results when applied in 'retrospective predictions', and is essential for the practical application of the modified time-to-failure method. In addition, an observed linear characteristic (see Section 3.1.3.2(a) of this report) in long-term seismic release is used to minimize false predictions.

## Results

1. The refined empirical relationships that eliminate or constrain unknown constants used in the modified time to failure method and the main event location search were used in a practical application in the New Madrid Seismic Zone (NMSZ).
2. The NMSZ is considered overdue for a magnitude 6 event according to recurrence rates. This makes the region ideal for testing the method. One location was identified in the NMSZ as a 'high risk' area for an event in the magnitude 4.5 range.

**Brehm, D. J. and L. W. Braile. 1999b.** Intermediate-term earthquake prediction using the modified time-to-failure method in Southern California. *Bulletin of the Seismological Society of America* 89. pp 275 - 293.

## Summary

Based on retrospective modelling of earthquakes from the southern California earthquake catalogue, along with previously published evaluations from the New Madrid Seismic Zone, the authors argue that the modified time-to-failure method (also employed in this report) may be used as an intermediate-term earthquake prediction technique for locating and predicting the size and time of a future earthquake. Modelling of previous main events for hypothesis development indicated to the authors that the method predicts the actual magnitude of the main event to within  $\pm 0.5$  magnitude units. The error associated with the time of failure is  $\pm 1.1$  years assuming the last precursory event is known. When the last event in the precursory sequence is not known, the predicted magnitude remains similar, but the predicted time requires refinement as additional events are added, with time, to the sequence. The main event location can also be identified within a circular region with a radius on the order of tens of kilometres for tectonic earthquakes. Criteria are provided for defining acceleration sequences and main event locations, which reduce the number of false predictions but also eliminate some main events from evaluation. Main events as small as magnitude 5.5, occurring between 1980 and 1995, are evaluated from the Southern California earthquake Catalogue (SCC). The results were used in association with previous studies to develop a method that it is claimed can be used for practical (forward prediction) applications. Again, the modified time-to-failure method is used to search the SCC for future main events occurring after 17 August 1998.

## Results

1. One region satisfied all the criteria and could be modelled by the modified time to failure method. The region likely to have a large event is claimed to be a 65-km-radius area centred at  $31.43^{\circ}$  N,  $115.47^{\circ}$  W (northern Baja California, Mexico).
2. The predicted magnitude is 6.36,  $\pm 0.55$ , and the predicted time of failure is 1998.565 (7/25/98),  $\pm 1.127$  years. The addition of future precursory events will allow refinement of the predicted values.

**Glasstone, S. et al. 1941.** The theory of rate processes - Kinetics of chemical reactions, viscosity, diffusion and electrochemical phenomena. New York: McGraw-Hill, 611p.

## Summary

Time to failure and rate process theory. The authors related viscosity and plasticity mathematically to thermal activation and forces and rates of shear at atomic and molecular levels. By relating the net rate at which molecular bonds are broken (bond repair is also possible) under applied stress,

the rate of local stress increase at molecular level was found to be inversely proportional to the time remaining to failure of a sample.

The equation, developed to explain viscosity at molecular level, has been applied to larger masses in thermally activated creep processes and fracture of brittle materials as well as stress corrosion and crack propagation in the solid earth (see references in *Varnes (1987)*).

**Leckie, F. A. and D. R. Hayhurst 1977.** Constitutive equations for creep rupture. *Acta Metallurgica*, 25:1059-1070.

### Summary

Continuum damage mechanics: Kachakov and Rabotnov first postulated relations between stress, strain, time and an internal state variable called “damage” in metals in the 1960s in the Russian literature. Leckie and Hayhurst generalised these relations to multiaxial stress states and *Ashby and Dyson (1986)* included creep damage:

$$\frac{\dot{\varepsilon}}{\dot{\varepsilon}_0} = \left(\frac{\sigma}{\sigma_0}\right)^j \frac{1}{(1 - \omega)^k} \quad (\text{A6.1})$$

$$\frac{\dot{\omega}}{\dot{\omega}_0} = \left(\frac{\sigma}{\sigma_0}\right)^p \frac{1}{(1 - \omega)^g} \quad (\text{A6.2})$$

where the dot indicated differentiation with respect to time and  $\omega$  is the damage variable ( $\omega = 0$  when the material is undamaged and  $\omega = 1$  at rupture). Assuming that the load remains constant and that  $t = 0$  when  $\omega = 0$ , and  $t = t_f$  when  $\omega = 1$ , eq. (A6.2) can be integrated:

$$1 - \omega = \left(\frac{t_f - t}{t_f}\right)^{\frac{1}{g+1}} \quad (\text{A6.3})$$

Combined with (A6.2) this gives

$$\frac{d\varepsilon}{dt} = \frac{C}{(t_f - t)^\alpha} \quad (\text{A6.4})$$

with

$$C = \varepsilon_0 \left(\frac{\sigma}{\sigma_0}\right)^j t_f^{\frac{k}{g+1}}, \quad \alpha = \frac{k}{g+1}$$

Thus the same expression as that for tertiary creep is recovered.

**Ouillon, G. and D. Sornette 1998.** Testing the degree of predictability of mine rockbursts from improved time-to-failure analysis. *Report commissioned by ISS International Limited.* Welkom: ISS International Limited.

### Summary

Tests of the critical earthquake concept performed on rockbursts in deep South African mines. The concept of an optimal time and space correlation region was tested on the eight main events in the catalogue provided by ISSI. As elsewhere in GAP409, the simplest signature of criticality in terms of a power law time-to-failure formula was searched for. Notwithstanding the fact that the search for the optimal correlation size is performed with a simple power law, evidence both for accelerated seismicity and for the presence of log-periodic behaviour, with a preferred scaling factor close to 2, was found.

**Saleur, H., C. G. Sammis and D. Sornette 1996.** Discrete scale invariance, complex fractal dimensions and log-periodic fluctuations in seismicity. *Journal of Geophysical Research*, 101:17661-17677.

### Summary

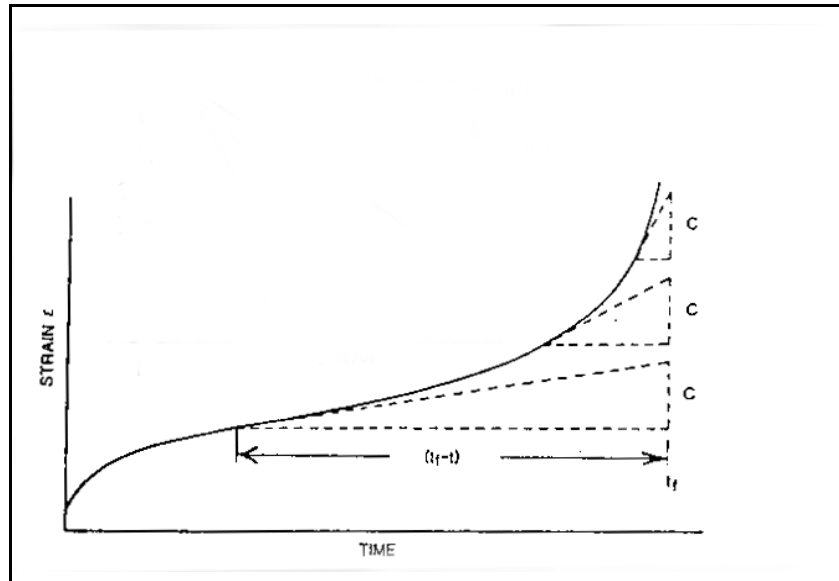
The concept of discrete scale invariance is discussed in detail and it is illustrated how it leads to complex critical exponents (see Appendix 4) and hence to the log-periodic corrections to scaling exhibited by various measures of seismic activity close to a large earthquake singularity. Discrete scale invariance is first illustrated on a geometrical fractal, the Sierpinsky gasket, which is shown to be fully described by a complex fractal dimension whose imaginary part is a simple function (inverse of the logarithm) of the discrete scaling factor (see Appendix 4). Then, a set of simple physical systems (spins and percolation (see, for example, *Boonzaaier (2000)* and *Turcotte (1999)*) on hierarchical lattices is analysed to exemplify the origin of the different terms in the discrete renormalisation group formalism (see Appendix 4.3) introduced to tackle this problem. As a more specific example of rupture relevant for earthquakes, the authors propose a solution of the hierarchical time-dependent fibre bundle of *Newman et al. (1994)* which exhibits explicitly a discrete renormalisation group from which log-periodic corrections follow.

### Results

1. Discrete scale invariance does not necessarily require an underlying geometrical hierarchical structure. A hierarchy may appear 'spontaneously' from the physics and/or the dynamics in a Euclidean (nonhierarchical) heterogeneous system.
2. A simple dynamical model of such mechanism, in terms of a random walk (or diffusion) of the seismic energy in a random heterogeneous system is discussed.

**Servi, I. S. and N. J. Grant 1951.** Creep and stress rupture behaviour of aluminium as a function of purity. *Transactions of the American Institute of Mining and Metallurgical Engineers*, 191:909-916.

## Summary



**Figure A6.1: Strain versus time creep curve with a decelerating primary part and accelerating tertiary part (modified from Varnes, 1989).**

Many materials, when subjected to sustained differential stress which ultimately leads to failure, follow creep curves (Figure A6.1). After a series of tests to failure on the same material, but at different stress levels, the authors showed that the minimum creep rate, occurring at the point of inflection on a strain - time curve, obeys a power law relation in the time to failure (Figure 3.1.1 of this report).

**Varnes, D. J. 1987.** Foreshock seismic energy release functions: Tools for estimating time and magnitude of main shocks. *U. S. Geological Survey Open-File Report 87-429*, 39p.

and

**Varnes, D. J. 1989.** Predicting earthquakes by analysing accelerating precursory seismic activity. *Pure and Applied Geophysics*, 130:662-686.

## Summary

Probably the first application of power-law time to failure ideas to earthquakes. During 11 sequences of earthquakes that were in retrospect classed as foreshocks, the accelerating rate at which seismic moment is released followed, at least in part, a simple equation,  $d(\Sigma\sqrt{M_o}) / dt = C(t_f - t)^n$ , where  $\Sigma\sqrt{M_o}$  is the cumulative sum until time,  $t$ , of the square roots of seismic moments of individual foreshocks computed from reported magnitudes.  $C$  and  $n$  are constants and  $t_f$  is a limiting time at which the rate of seismic moment accumulation becomes infinite. This paper was written by an empirical scientist, who found the possible time of a major event or main event,  $t_f$ , by the best fit of the power law or its integral, to step-like plots of  $\Sigma\sqrt{M_o}$ .

versus time using successive estimates of  $t_f$  in linearized regressions until the maximum coefficient of determination,  $r^2$ , was obtained.

Analysed examples included sequences preceding earthquakes at Cremasta, Greece, 2/5/66; Haicheng, China 2/4/75; Oaxaca, Mexico, 11/29/78; Petatlan, Mexico, 3/14/79 and Central Chile, 3/3/85.

In twenty-nine estimates of main-event time (time of failure), made as the sequences developed, the author claimed that errors in twenty of these estimates were less than one-half and in nine less than one tenth the time remaining between the time of the last data used and the main shock. Some precursory sequences, or parts of them, yielded no solution and two sequences appeared to include in their first parts the aftershocks of a previous event. Plots using the integral of the power law show that the sequences are easily separable into aftershock and foreshock segments.

Varnes constructed synthetic seismic sequences of shocks at equal time intervals to follow the power law using four values of the exponent. In each series the resulting distributions of magnitudes closely followed the linear Gutenberg-Richter relation  $\log N = a - bM$ , and the product  $n$  times  $b$  for each series was the same constant.

In various forms and for decades, power law time to failure has been used successfully to predict failure times of stressed metals and ceramics, landslides in soil and rock slopes, and volcanic eruptions. Results of experiments and theoretical studies on crack propagation, fault mechanics, and acoustic emission can be closely reproduced by the equation.

## A6.1.2 Criticality and self organised criticality

**Bak, P. 1996.** How nature works - the science of self-organized criticality. New York: Springer-Verlag, 198 pp.

### Summary

A history and explanation of self organised criticality by one of the originators of the concept. Very little mathematics, written to be accessible and entertaining to the layman.

**Boonzaaier, L. 2000.** Application of models of self-organised criticality to mining-induced seismicity. *M. Sc. dissertation*. Stellenbosch: University of Stellenbosch, ???p. In preparation.

The concept of self-organized criticality was introduced to explain the behaviour of the sandpile model. In this model, particles are randomly dropped onto a square grid or lattice of boxes. When a box accumulates four particles they are redistributed to the four adjacent boxes or lost off the edge of the grid. Redistributions can lead to further instabilities with the possibility of more particles being lost from the grid, contributing to the size of each 'avalanche'. The model 'avalanches' satisfied a Gutenberg-Richter power-law frequency-area distribution with a slope near unity, if the model is not driven too hard (i.e., the number of particles added to the grid at any given time step is small). The driving rate corresponds to the rate of stress buildup in the model. If the driving rate becomes large, the model switches from Gutenberg-Richter scaling behaviour for the size distribution to the predominance of large characteristic avalanches.

Other cellular-automata models, including the slider-block and forest-fire models, are also said to exhibit self-organized critical behaviour. It has been argued that earthquakes, landslides, forests fires and species extinctions are examples of self organised criticality in nature. In addition, wars and stock market crashes have been associated with this behaviour.

In this dissertation, a characteristic avalanche is defined as an avalanche involving at least 59% (the so-called percolation threshold (*Turcotte (1999)*) of the lattice. It is shown that, in the self organised phase (Gurtenberg-Richter scaling regime), the characteristic avalanches (in one simulation 350 of them) are mostly (85%) preceded by a sequence of avalanches increasing in size according to a power law in the time to failure. In the characteristic phase, when mostly large, characteristic, avalanches occur, they are seldom (23 out of 2138, or 1.08%) preceded by such accelerating release.

### Comment

This result serves to point out some (perhaps inadvertent) obfuscation practised by opponents of the quest for reliable forecasting.

In this simple toy model of self organisation, it was shown that there is no conflict between the scale invariance of a phenomenon and the existence of deterministic precursory behaviour in its finer structure. It may be true that in the scaling regime, no characteristic scale for seismicity (*Evans (1997)*, *Geller (1997)*, *Leary (1997)*) exists when viewed over a geological time scale. However, finer resolution (on the scale of human lifetimes or a production week) may reveal structure often glossed over or dismissed as a result of preconceived notions or poor understanding of the concepts at issue.

**Cardy, J. 1996.** Scaling and renormalisation in Statistical Physics. Cambridge: Cambridge University Press, 238 pp.

### Summary

Standard textbook.

**Newman, W. I., D. L. Turcotte and A. Gabrielov 1996.** A Hierarchical model for precursory seismic activity. (*In: Rundle, Turcotte and Klein. Reduction and predictability of natural disasters, SFI studies in the sciences of complexity XXV.* Addison-Wesley, 1996, p.243-260).

### Summary

Seismic activation has been recognized to occur before many major earthquakes. A specific case is the seismic activation that occurred in the San Francisco Bay area, particularly in the East Bay, prior to the 1906 earthquake. At the time of writing, there was serious concern that the (then recent) series of earthquakes in Southern California was seismic activation prior to a great Southern California earthquake. Seismic activation in a broad region forms a primary basis for the intermediate term earthquake prediction carried out by the Moscow group under Keilis-Borok (see the papers on pattern recognition described elsewhere in this survey). *Bufe and Varnes (1987, 1989)* quantified seismic activation prior to the Loma Prieta earthquake in terms of a power-law increase in the regional Benioff strain release prior to this earthquake. *Saleur, Sornette and Sammis (1996)* have considered this data and concluded that there is an excellent fit to a log-periodic increase in the Benioff strain release.

In the paper summarised here, Newman et al. studied a hierarchical seismic failure model. An array of stress-carrying elements was considered. Formally, such an array is a cellular automaton (see also the description of self-organised criticality in the discussion of *Bak (1996)* and *Boonzaaier (2000)* elsewhere in this survey) or lattice gas, but here more analogous to the strands of an ideal, frictionless cable. Different from *Boonzaaier (2000)*, Newman et al. endowed each element with a time-to-failure that is dependent on the stress that the element is subjected to and a statistical distribution of values. Analogous to the sandpile models, the stress on the element is transferred to a neighbouring element when an element fails. If two adjacent elements fail, stress is transferred

to two neighbouring elements; if four elements fail, stress is transferred to four adjacent elements, and so forth. When stress is transferred to an element its time to failure is reduced. The stress redistribution technique models the transfer of stress from faults of varying sizes. The reduced times to failure are consistent with an aftershock sequence following an earthquake (Omori's law).

## Results

1. The intermediate-size failure events prior to total failure each have a sequence of precursory failures, and these precursory failures each have an embedded precursory sequence of smaller failures.
2. The entire failure sequence is "fractal" - each failure appears to have a self-similar sequence of failures on a smaller scale.
3. The total failure of the array appears to be a critical point.
4. There appears to be a sequence of partial failures leading up to the total failure that resembles a log-periodic time to failure sequence.

## Comment

These results should be carefully contrasted with the statements by *Leary (1997)* and *Kagan (1997)* above. As in the case of *Boonzaaier (2000)*, the total, or characteristic failure, while forming part of a scaling law (Gutenberg-Richter) still can have a power law activation sequence.

**Newman, W. I., A. M. Gabrielov, T. A. Durand, S. L. Phoenix, D. L. Turcotte 1994.** An exact renormalisation model for earthquakes and material failure. Statics and dynamics. *Physica D*:200-216.

## Summary

Earthquake events are well-known to possess a variety of empirical scaling laws. Accordingly, renormalisation methods (see Appendix 4.3) offer some hope for understanding why earthquake statistics behave in a similar way over orders of magnitude of energy. This article reviews the progress made in the use of renormalisation methods in approaching the earthquake problem. In particular, seismic events were modelled before as hierarchically organized bundles of fibres with equal load sharing. The authors considered by computational and analytic means the failure properties of such bundles of fibres, a problem that may be treated exactly by renormalisation methods.

## Results

Independent of the specific properties of an individual fibre, the stress and time thresholds for failure of fibre bundles obey universal, albeit different, scaling laws depending on the size of the bundles. Application to fracture processes in earthquake events and in engineering materials is claimed to provide insight into some of the observed patterns and scaling, such as the apparent weakening of earthquake faults and composite materials with size and the emergence of relatively well-defined stresses and times when failure appears to be assured.



**Turcotte, D. L. 1999.** Self-organized criticality. *Reports on Progress in Physics*, 62:1377-1429.

### Summary

One of the models of self organised criticality (if the concept has a basis in reality) is the forest-fire model. A variation on the sandpile model (*Bak (1996), Boonzaaier (2000)*), it is particularly interesting in terms of its relation to the critical-point behaviour of the site percolation model.

In the basic forest-fire model, trees are randomly planted on a grid of points. Periodically in time, sparks are randomly dropped on the grid. If a spark drops on a tree, that tree and adjacent trees burn in a model fire. The flames are the 'avalanches' and they are found to satisfy power-law frequency-area distributions with slopes near unity. This forest-fire model is closely related to the site-percolation model which exhibits critical behaviour. In the forest-fire model there is an inverse cascade of trees from small clusters to large clusters and trees are lost primarily from model fires that destroy the largest clusters. This quasi steady-state cascade gives a power-law frequency-area distribution for both clusters of trees and smaller fires. The site-percolation model is equivalent to the forest-fire model without fires. In this case there is a transient cascade of trees from small to large clusters and a power-law distribution is found only at a critical density of trees.

**Yeomans, J. M. 1992.** Statistical Mechanics of Phase Transitions. Oxford: Oxford Science Publications, 1993, p22-25.

### Summary

Well-known and widely-used textbook on phase transitions. See Appendix 4.

## A6.1.3 Pattern recognition

**Freeman, J. A. and D. M. Skapura 1991.** Neural networks: Algorithms, applications and programming techniques. Reading, Massachusetts: Addison-Wesley, p128-130.

### Summary

Standard textbook.

**Gelfand, I. M., SH. A. Guberman, V. I. Keilis-Borok, L. Knopoff, F. Press, E. Ya. Ranzman, I. M. Rotwain and A. M. Sadovsky 1976.** Pattern recognition applied to earthquake epicenters in California. *Physics of the earth and planetary interiors*, 11:227-283.

### Summary

The use of quantitative pattern recognition to deduce the location of earthquake epicentres. Geological data and the earthquake history of a region, in this case California, is codified and used to construct a pattern recognition algorithm which learns how to separate earthquake epicentres from other places. The problem is formulated in several ways and control experiments are devised and applied in order to test the stability of the procedures and engender confidence in the results.

## Results

Sites of future earthquake epicentres are predicted as well as places where epicentres will not occur.

**Keilis-Borok, V. I. and A. J. Lichtman 1993.** The self-organisation of American society in presidential and senatorial elections. (*In: Kravtsov, Y. A. Limits of predictability.* Springer-Verlag, Berlin, 1993, p223-237).

## Summary

The methodology used in *Gelfand et al. (1976)* is applied to the prediction of elections. The value of this paper is pedagogical. The method is explained in detail in this report in the explanation of the INDICATOR algorithm.

**Keilis-Borok, V. I. and A. Soloviev 1999.** Pattern recognition: Algorithms and Applications. *Lecture and tutorial notes of the Fifth Workshop on Non-Linear Dynamics and Earthquake Prediction - document H4.SMR/1150-6.* Trieste, 4-22 October 1999: The Abdus Salam International Centre for Theoretical Physics.

## Summary

Some theoretical background on pattern recognition (used in the construction of the pattern recognition algorithm INDICATOR in this report) and a reason for having an alternative to fitting data to a physical model (SOOTHSAY).

When one considers a set of objects, phenomena or processes, some information is available about each element of the set, and additionally, there often is a feature shared only by a part of the elements. If the information available about an element of the set does not clearly exhibit the existence of this feature in the element, the problem becomes one of distinguishing those elements that do exhibit the trait. One can attempt to solve the problem by constructing a model on the basis of mechanical, physical, chemical or other scientific laws to explain the connection between the available information and the feature under consideration. In many cases the construction of such a model is difficult or practically impossible. In such a case a starting point is pure pattern recognition.

**Kossobokov, V. G., L. L. Romashkova, V. I. Keilis-Borok and J. H. Healy 1999.** Testing earthquake prediction algorithms: statistically significant advance prediction of the largest earthquakes in the Circum-Pacific, 1992-1997. *Physics of the earth and planetary interiors*, 111:187-196.

## Summary

Statistical significance above 99% level is established for prediction of earthquakes by the algorithms M8 and "Mendocino Scenario", MSc. This evaluation results from a real-time intermediate-term prediction of the largest earthquakes in Circum Pacific seismic belt from 1992 to 1997. The authors claim that all earthquakes of magnitude 8 or greater were predicted and that the predictions are completely reproducible, though most of them were not filed officially. These algorithms (the foundations of the INDICATOR algorithm in this report), are pattern recognition algorithms, based on seismicity patterns. Alarm periods determined by M8 extend to several years and cover areas up to ten times the length of the fault segment expected to rupture in a large earthquake. The algorithm MSc reduced such areas to between one and three times the length of the fault segment expected to rupture.

## Results

- For tectonic earthquakes, large fault systems can be involved in the formation of earthquake precursors at the scale of years.
- Confirmation of the existence of premonitory seismicity patterns.
- A case for universality: the partial similarity of these phenomena in a wide range of tectonic environments.
- Speculation that the results provide empirical constraints for modelling seismicity.

see also

**Romachkova, L. L., V. G. Kossobokov, G. F. Panza and G. Costa 1998.** Intermediate-term predictions of earthquakes in Italy: Algorithm M8, *Pure and Applied Geophysics* 152: 37-55.

## A6.1.4 Statistics and error analysis

**Kagan, Y. Y. 1996.** Statistical aspects of the Parkfield earthquake sequence and Parkfield prediction experiment. *Tectonophysics*, 270: 207-219.

### Summary

This note discusses three interconnected statistical problems concerning the Parkfield sequence of moderate earthquakes and the Parkfield prediction experiment:

1. Is it possible that the quasi-periodic Parkfield sequence of characteristic earthquakes is no uncommon, specific phenomenon (the research hypothesis), but can be explained by a preferential selection from available earthquake catalogues?  
The null hypothesis is then formulated as follows: Earthquakes occur according to the Poisson process in time and their size follows the Gutenberg-Richter relation.
2. If the null hypothesis cannot be refuted, what is the probability of magnitude  $m \geq 6$  earthquake occurrence in the Parkfield region?
3. The direct goal of the Parkfield experiment was the registration of precursory phenomena prior to an  $m \geq 6$  earthquake. In the absence of the characteristic earthquake, could the experiment have resolved which of the two competing hypotheses was true in a reasonable time?

### Results

Statistical analysis was hampered by an insufficiently rigorous definition of the research model and inadequate or ambiguous data. However, the author proceeded to show that the null hypothesis could not be decisively rejected. The quasi-periodic pattern of intermediate size earthquakes in the Parkfield area is a statistical event likely to occur by chance if it had been preferentially selected from available earthquake catalogues. The observed magnitude-frequency curves for small and intermediate earthquakes in the Parkfield area agree with the theoretical distribution computed on the basis of a modified Gutenberg-Richter law (a gamma distribution), using deformation rates for the San Andreas fault.

It is then showed that the size distribution of the Parkfield characteristic earthquakes could also be attributed to selection bias. According to the null hypothesis, the yearly probability of a  $m \geq 6$

earthquake originating in the Parkfield area was less than 1%, signifying that several more decades of observation may be needed before the expected event occurs. The author concludes that, by its very design, the Parkfield experiment could be expected to yield statistically significant conclusions on the validity of the research hypothesis for many decades.

**Molchan, G. 1999.** Some statistical problems in earthquake prediction and seismicity. *Fifth Workshop on Non-Linear Dynamics and Earthquake Prediction*. Trieste, 4-22 October 1999: The Abdus Salam International Centre for Theoretical Physics.

### Summary

An interdisciplinary problem of earthquake prediction involving economics is considered. The convolution of the two disciplines is meant to aid in understanding the prediction problem as a whole and revealed additional requirements for seismostatistics. The problem is formulated as an optimal control problem: Possessing the possibility to declare several types of alerts, using various prediction strategies, it is necessary to find an optimal strategy minimizing the total expected losses (in the mining scenario, this would correspond to the loss of production, for example).

Losses include the costs for maintaining alerts and for changing alert types. Each successful prediction prevents a certain amount of losses; total expected losses are integrated over a semi-infinite time interval. A discount factor is included in the model while algorithmic and exact solutions are indicated.

This paper was based on results by *Molchan (1990, 1991, 1992)*.

### Comment

The methodology presented in this paper was used in the current project to determine, analytically, the degree of randomness in the prediction strategies followed by the SOOTHSAY algorithm.

**Sornette, D. and L. Knopoff 1997.** The paradox of the expected time until the next earthquake. *Bulletin of the Seismological Society of America*, 87: 789-798.

### Summary

An analytic exploration of the proposition that the longer it has been since the last earthquake, the longer the expected time till the next.

The authors show that the validity of the statement depends crucially on the statistics of the fluctuations in the interval times between earthquakes. For the periodic, uniform, semi-Gaussian, Rayleigh, and truncated statistical distributions of interval times, as well as Weibull distributions with exponent greater than one, the expected time to the next earthquake decreases with increasing time since the last one, for long times since the last earthquake. The lognormal and power-law distributions and the Weibull distributions with exponents smaller than one have increasing times to the next earthquake as the elapsed time since the last increases, for long elapsed times.

There is an identifiable crossover between these models, which is gauged by the rate of fall-off of the long-term tail of the distribution in comparison with an exponential fall-off. The response to the question for short elapsed times is also evaluated. The lognormal and power-law distributions give one response for short elapsed times and the opposite for long elapsed times. Even the sampling of a finite number of intervals from a Poisson distribution will lead to an increasing estimate of time to the next earthquake for increasing elapsed time since the last one.

## A6.2 Sceptics

**Brummer, R. 1999.** Simple Truths about Rockbursts. *SARES '99, Keynote Address.* Johannesburg.

### Summary

Rockbursts cannot be predicted - what are the alternatives? Consideration is given to orebody shape, rock type, in-situ stresses, statistical analysis and the requirements for operating microseismic systems. A method to estimate the Probability of Occurrence of seismic events is described in terms of a Magnitude/Hazard/Time diagram. This information can be used to estimate the risks faced by a mine, and to make optimal choices between various different options, using approaches such as Risk-Cost-Benefit Analysis.

**JGI 1997.** Special section - Assessment of schemes for earthquake prediction. *Geophysical Journal International*, 131: 413-533.

**Evans, R. 1997.** Assessment of schemes for earthquake prediction: Editor's introduction. (*In: Special section - Assessment of schemes for earthquake prediction. Geophys. J. Int. (1997) 131, 413-420*)

### Summary of the special edition.

A series of papers on assessment of schemes for earthquake prediction, originally presented at an RAS/JAG Discussion Meeting. The papers also address the social and administrative implications of earthquake prediction and the question of whether this objective is feasible in either a strictly scientific or a socially useful sense.

Contributions on the assessment process itself stressed the importance of clearly defining procedures and anticipated outcomes in advance of any prediction experiment.

It was pointed out that after-the-fact searches for correlations drastically weaken the statistical significance of any association of phenomena with precursors which is found. The record of a proposed prediction scheme can be made to appear more significant by comparing it with an inadequate or unsuitable model of seismicity. Because the distribution of earthquakes is clustered in both time and space, simple models (e.g. Poisson distribution in time) do not fully characterise it, and statistical tests which rely on such assumptions are inappropriate. The history of attempts to predict earthquakes suggests that little if any progress is being made towards the goal of specifying - in advance and within reasonably tight limits - the time, place of occurrence and the size of a forthcoming event. The view that prediction may be impossible is regaining support following a change in perception regarding the nature and origin of earthquakes and faulting. Reid's elastic rebound model and its associated conceptual framework of a near-homogeneous Earth were at the time of the meeting being superseded by a fundamentally heterogeneous model employing the concept of self-organized criticality. Within this model, both the initiation and the eventual progress of rupture on a fault depend on a multitude of factors related to both local stress and strength, which are inter-related and heterogeneous beyond practical measurement. The implication is then that the course of a fracture towards large-scale failure cannot practically be determined in advance.

This view must be contrasted with the results of *Boonzaier (2000)*.

Other contributions showed that an unusually high degree of scientific confidence is required in order for a prediction to be of significant social value. The claim is then made that, in terms of saving both lives and property, statistical approaches to hazard assessment, combined with appro-

priate programmes of risk reduction, continue to offer the most cost-effective means of reducing earthquake-related losses.

**Bernard, P, Pinettes, P M, Hatzidimitriou, P M, Scordilis, E M, Veis, G and Milas, P. 1997.** From precursors to prediction: a few recent cases from Greece. (*In: Special section - Assessment of schemes for earthquake prediction. Geophys. J. Int. (1997) 131, 467-477*).

### Summary

The two destructive earthquakes of 1995 in Greece, the May 13  $M_s=6.6$  Kozani-Grevena and the June 15  $M_s=6.2$  Aigion events, form case studies of problems related to the identification of precursors and the efficiency and usefulness of prediction.

The Kozani earthquake was preceded, within 30 minutes of the main event, by five foreshocks with magnitude greater than 3.5. These events were relocated with respect to each other, showing that they are clustered within 2 km of one another, about 5 to 10 km to the SSW of the main-shock epicentre. This size of foreshock clustering correctly fits the correlation law with the main-shock magnitude obtained by Dodge, Beroza & Ellsworth (1996) for Californian earthquakes. The foreshocks led to people leaving their houses, which explains the absence of casualties, despite the partial destruction of several villages. The possibility of issuing predictions in this area from the observation of earthquake clustering is analysed in light of the seismicity observed during the previous 15 years.

The VAN group issued, based on SES signals at a station called IOA, a prediction before this earthquake, considered by them as a success, while others dispute their claim. The SES was also recorded by a magneto-telluric station installed by IPGP, a few kilometres from IOA. An artificial origin for the SES was suggested, but could not be identified. Simple amplitude estimates show that a local, natural source such as an electrokinetic effect is unlikely, and that a remote electrokinetic source in the epicentral area can be even more confidently rejected.

Another SES on VAN's network (VOL station, 30 April 1995) led the VAN group to predict an earthquake outside the IOA sensitivity area (IOA did not record any anomaly), and to announce a success when the Aigion earthquake occurred. However, this event was located inside the IOA sensitivity area, and the prediction was hence viewed as a failure. Furthermore, at the time of this SES, neither tilt nor strain was observed above the noise level of a few  $10^{-8}$  at the IPGP/NTUA Galaxidi geophysical observatory, 20 km from the hypocentre, leading to the conclusion that the electrical source of this SES was most probably located near VOL, 100 km away, whatever its correlation with the earthquake.

**Enomoto, Y, Tsutsumi, A, Fujinawa, Y, Kasahara, M and Hashimoto, H. 1997.** Candidate precursors: pulse-like geoelectric signals possibly related to recent seismic activity in Japan. (*In: Special section - Assessment of schemes for earthquake prediction. Geophys. J. Int. (1997) 131, 485-494*).

### Summary

Telemetric network observations of pulse-like geo-electric charge signals using a buried vertical dipole were undertaken at various observation sites over a wide area in Japan. The authors attempted to identify anomalous signals, possibly related to seismic electric activity, from continuous records of signals during six months from April 1996. Special attention was paid to the elimination of noise resulting from industrial and meteorological electric activity, comparison with other electromagnetic signals in the VLF band and the relation between the precursor period and the distance from the eventual main event that occurred in Japan.

## Results

Four candidate precursor electric signals, which were not contaminated by industrial and meteorological electric activity, were tentatively identified.

- The second appeared before the Akita-ken Nairiku-nanbu earthquake swarm, for which the maximum  $M = 5.9$  occurred on 11 August 1996.
- The third and fourth possible precursor appeared before the Chiba-ken Toho-oki earthquake,  $M = 6.6$ , on 11 September 1996.
- A tentative qualitative model explaining why the candidate precursory signal is related to stress building up before an earthquake is presented in terms of the electrification of gases released from fracturing rocks immediately prior to the main event.

**Geller, R J. 1997.** Earthquake prediction: a critical review. (*In: Special section - Assessment of schemes for earthquake prediction. Geophys. J. Int. (1997) 131, 425-450*).

## Summary

Geller claims that earthquake prediction research has been conducted for over 100 years with no obvious successes while claims of breakthroughs have failed to withstand scrutiny. He states that extensive searches have failed to find reliable precursors, while theoretical work suggests that faulting is a non-linear process which is highly sensitive to unmeasurably fine details of the state of the Earth in a large volume, not just in the immediate vicinity of the hypocentre. Any small earthquake thus has some probability of cascading into a large event.

## Conclusion

Reliable issuing of alarms of imminent large earthquakes appears to be effectively impossible.

## Comment

The view of this project is that the heterogeneity and nonlinearity inherent in seismicity was poorly understood in 1997, and the results of this project as well as *Boonzaaier (2000)* provide a different perspective.

**Kagan, Y Y. 1997.** Are earthquakes predictable? (*In: Special section - Assessment of schemes for earthquake prediction. Geophys. J. Int. (1997) 131, 505-525*).

## Summary

The author discusses several definitions and possible classifications of earthquake prediction methods, considers various measures of prediction efficiency, reviews several recent examples of earthquake prediction and describes the methods that can be used to verify prediction schemes.

## Results

Kagan concludes that an empirical search for earthquake precursors that forecast the size of an impending earthquake has been fruitless. The claim is that no statistically rigorous validation of proposed precursory phenomena is available and that therefore, reported cases of precursors can be explained by random noise or by chance coincidence. The usual evidence that earthquakes are non-linear, chaotic, scale-invariant phenomena are presented, and the conclusion reached that the most probable consequence of earthquake self-similarity is a lack of earthquake predictability as

popularly defined, that is a forecast of a specific individual earthquake. Many small earthquakes occur throughout any seismic zone, demonstrating that the critical conditions for earthquake nucleation are satisfied almost everywhere.

The author speculates that, apparently, any small shock can grow into a large event (in contradiction to *Leary (1997)* in the same meeting). Then, in apparent contradiction to the previous sentence, he states that “Thus, it is likely that an earthquake has no preparatory stage”. This sceptical view of current earthquake prediction efforts is qualified by the statement that it should not be interpreted as a statement that any further attempts to mitigate the destructive effects of earthquakes are futile. The seismic-moment conservation principle (used in this project as the seismic release deficiency), when combined with geodetic deformation data, offers a new way to evaluate the seismic hazard, not only for tectonic plate boundaries, but also for areas of low seismicity, that is, the interiors of continents. The view is put forward that earthquake clustering with a power-law temporal decay (Omori's law) can be used to estimate the rate of future earthquake occurrence and that realtime seismology can facilitate relief efforts after large earthquakes and eventually provide an immediate warning of severe shaking a few seconds or tens of seconds before the shaking starts.

**Leary, P C. 1997.** Rock as a critical-point system and the inherent implausibility of reliable earthquake prediction. (*In: Special section - Assessment of schemes for earthquake prediction. Geophys. J. Int. (1997) 131, 451-466*).

## Summary

The statistical mechanics of critical-point thermodynamic systems (see Appendix 4 of this project report) as a conceptual tool with which to examine the plausibility of reliable earthquake prediction. The issue turns on the heterogeneity of rock. If the correlation range of a spatially fluctuating rock property has a characteristic length  $\xi$ , the spatial average of that property over dimensions greater than  $\xi$  may be regarded as defining an 'effective', continuum (“smeared-out”) value for that property. Then earthquake preparation in rock with an effective continuum mechanical strength can be seen as failure that is initiated at small anomalous zones and then grow to instability, fed by a continuous volume of strain energy. The dimensions and temporal evolution of the growing anomaly can then be expected to relate to the size, location and timing of the impending seismic event.

However, according to the author, borehole log evidence indicates that rock-heterogeneity correlation lengths  $\xi$  are unbounded ( $\xi \rightarrow \infty$ ), and hence that the 'effective' continuum approximation for rock must be applied with discretion. An infinite correlation length indicates a statistical mechanical critical-point system in which the correlation length is related to a singularity in a physical property. In thermodynamic critical systems, the singularity occurs at the critical system temperature  $T_c$ ,  $\xi \propto 1 / \sqrt{|T - T_c|}$  (see Appendix 4), while in rock the singularity occurs at the critical percolation probability  $P_c$ ,  $\xi \propto 1 / \sqrt{|P - P_c|}$ .

Statistical physics descriptions of critical-point spatial correlations are based on nearest-neighbour interactions at every scale length. Nearest-neighbour interactions at one scale, say that large failure structures, involve minimal coupling to structures at other scales, in particular small-scale failure structures. The mechanics of nearest-neighbour interactions as explained by the author is claimed to limit the role of small-scale failure (smaller events) or related nucleating activity in the build-up to large-scale failure.



## Conclusions

In light of the quoted evidence against rock being an effective medium supporting localized, smaller, precursory activity that grows to large-scale mechanical failure, and the history of failure to identify reliable precursors to major earthquakes, the author urges advocates of earthquake prediction schemes to evaluate their claims for precursory activity in terms of rock as a critical-point system.

## Comment

The foundation of the research methodology in this report is critical-point analysis.

**Matthews, R A J. 1997.** Decision-theoretic limits on earthquake prediction. (*In: Special section - Assessment of schemes for earthquake prediction. Geophys. J. Int.* (1997) 131, 526-529).

## Summary

Decision Theory is applied to the question of how accurate earthquake predictions must be to serve as a reliable basis for action.

## Results

Even with optimistic estimates of the parameters involved, the lower bound on the required accuracy is calculated as being over ten times higher than that of current meteorological forecasts. The author reasons that, given the abruptly self-organizing nature of earthquakes, it is extremely unlikely that precursors can attain such levels of accuracy.

**Mulargia, F. 1997.** Retrospective validation of the time association of precursors. (*In: Special section - Assessment of schemes for earthquake prediction. Geophys. J. Int.* (1997) 131, 500-504).

## Summary

This paper examines the problematic issue of retrospective association of events, a starting point of any search for precursors to geophysical phenomena. Association is analysed with respect to time, disregarding the spatial component. Two sources of bias are examined: The time lag between precursor and event series and the tolerance of that lag (also linked to the number of associated events).

Typically, in retrospective analyses, time windows are chosen *a posteriori* to minimize the probability of random association in the given series. Usually, for different input data, optimal association of the two series will be achieved for different values of time windows, so that any valid system of significance testing must include an estimation of the effect of retrospectively choosing optimal parameters rather than fixing them in advance.

## Results

The paper purports to show that, as long as at least one of the two time-series - either the precursors or the events - is Poissonian, the correction factors for the significance level can be written in a simple closed form. The correction scheme is then applied to a proposed association between seismic clusters and flank eruptions of Mount Etna. Although the association appears to be highly significant when assessed using simple standard tests, this significance is not confirmed when the effect of retrospective parameter optimization is taken into account.

The key point made is that retrospective studies can be investigated using standard techniques, but where such techniques are employed, the threshold for significance of association must be greatly lowered in order to account for the bias due to the optimal choice of parameters.

**Stark, P B. 1997.** Earthquake prediction: the null hypothesis. (*In: Special section - Assessment of schemes for earthquake prediction. Geophys. J. Int.* (1997) 131, 495-499).

### Summary

The null hypothesis in assessing earthquake predictions is often that successful predictions are random. The author makes the important point that the null hypothesis tends to be rejected not only when predictions do have merit, but also when the random model is inappropriate.

Constructing a null hypothesis more precisely requires specifying a random model for the predictions and / or the seismicity. Two approaches are considered:

1. Standard: The seismicity is taken to be random and the predictions are held fixed. 'Conditioning' on the predictions this way tends to reject the null hypothesis even when it is true, if the predictions depend on the seismic history.
2. The author's proposition: Predictions are compared with the predictions of a 'sensible' random prediction algorithm that uses seismicity up to time  $t$  to predict what will happen after time  $t$ . The null hypothesis is then that the predictions are no better than those of the random algorithm. Significance levels can be assigned to this test in a more satisfactory way, because the distribution of the success rate of the random predictions is controllable. Failure to reject the null hypothesis indicates that there is no evidence that any extra-seismic information the predictor uses (electrical signals for example) helps to predict earthquakes.

### Comment

In this report, it is pointed out that both these null hypotheses are in effect tested:

1. The algorithm SOOTHSAY was tested on a synthetic pseudo-random catalogue.
2. The success rate of SOOTHSAY (taking into account false alarms) was tested against random predictions.

In both cases, the algorithm performed startlingly better than random.

**Stiros, S C. 1997.** Costs and benefits of earthquake prediction studies in Greece. (*In: Special section - Assessment of schemes for earthquake prediction. Geophys. J. Int.* (1997) 131, 478-484).

### Summary

The author advances the hypothesis that hypothetical successful earthquake prediction in Greece would be of, at best, limited benefit to society.

- On average, less than 5 per cent of sizeable earthquakes (magnitude greater than 4.5) cause significant damage or loss of life.
- Organized evacuation of urban centres is unlikely to be successfully accomplished, because the public lacks confidence in the authorities and cannot be expected to respond promptly.

- Panic and other undesirable side-effects can be anticipated.
- The lead time between a short-term prediction and earthquake occurrence is too short for most actions aimed at reducing or eliminating primary or secondary earthquake effects; in any case, most such actions would be superfluous if appropriate longer-term preparedness plans were implemented.
- Prediction of 'an earthquake' is not an appropriate objective in an area such as Greece, which experiences complicated and long seismic sequences consisting of several destructive events, and large earthquakes with anomalous meizoseismal areas.
- The seismic death toll in Greece is relatively small (less than 10 people per year over the last 40 years), and due to recent changes in building styles and construction practices, current fatalities are mainly associated with the failure of multistory buildings.

## Conclusion

Earthquake prediction in Greece, would not be cost-effective, but alternative use of funding could be expected to save more lives with much greater certainty.

**Wys, M and Booth, D C. 1997.** The IASPEI procedure for the evaluation of earthquake precursors. (*In: Special section - Assessment of schemes for earthquake prediction. Geophys. J. Int. (1997) 131, 423-424.*)

## Summary

Full scientific evaluation of proposed earthquake precursors for earthquake prediction is a problem because independent testing is difficult or impossible. To approach this difficulty, and to assess the current state of the art of earthquake prediction research, IASPEI has devised a peer-review procedure for precursor evaluation. The procedure does not consider predictions of impending earthquakes, but evaluates case histories of proposed precursors for past events according to stated validation criteria, which are specified in terms of guidelines concerning the hypothesized physical model, data quality, anomaly definition, the rules of association of precursor with earthquake, and statistical significance.

## Results

By 1997, five precursors have been placed on a preliminary list of significant earthquake precursors, although none has satisfied the validation criteria well enough to ensure that their placement is permanent. The authors do not give particulars, beyond mentioning that three of these are based on seismicity patterns, one based on ground-water chemistry and temperature, and one on measurement of crustal deformation by ground-water levels. Exclusion of a precursor from the list does not mean it is useless, but further work is required if it is to become convincing. The main objectives in producing the list are to establish a consensus on the criteria which a precursor must satisfy to be recognized as validated, and to find case histories which satisfy these criteria. Further nominations of precursor candidates are requested for evaluation by the IASPEI procedure.

## A6.3 Seismicity in mines

**De Beer, W and A. J. Mendecki 1997.** Seismology for rockburst prediction. *SIMRAC year-end report GAP409*. Pretoria. Department of Minerals and Energy.

**De Beer, W. 1999.** Seismology for rockburst prediction. *SIMRAC first quarter report GAP 409*. Pretoria. Department of Minerals and Energy.

**De Beer, W and A. J. Mendecki 1997, 1998.** Seismology for rockburst prediction. *SIMRAC interim reports GAP409*. Pretoria. Department of Minerals and Energy.

**Mendecki, A. J. 1996.** Quantitative seismology and rockmass stability. (In: *Mendecki, A. J. Seismic Monitoring in Mines*. London: Chapman and Hall, London, 1997, p178-219).

**Mendecki, A. J. 1996.** Seismology for rockburst prevention, control and prediction. *SIMRAC report GAP017*. Pretoria. Department of Minerals and Energy.

**Mendecki, A. J. and G. van Aswegen 1997.** Mine layout, geological features and seismic hazard. *SIMRAC mid year report GAP303*. Pretoria. Department of Minerals and Energy.

### Summary

This project quantified and classified the seismic rock mass response to different mining layouts and geological features for seismic hazard assessment. The first requirement is the quantification of seismic hazard, i.e. the probability that a seismic event of a given 'size' would occur within a given time period. Objective seismic hazard assessment in deep level gold mines should be based on a combination of statistical methods and parameters which quantify the seismic rockmass response to mining in a physical sense. The tools created by GAP017 were utilised for the quantitative assessment of the rockmass response to mining under different circumstances. The variables determining the different circumstances include: mining face geometry, mining rate, depth, rock strength and the nature of geological structure. Since not all rockmass response to mining is recorded by seismic monitoring systems, it is further required to apply numerical modelling to fill the gaps. In principle, the difference between seismic output predicted by a calibrated numerical model and that measured by the seismic system could be the 'available' response posing a potential seismic hazard.

**Mendecki, A. J. and G. van Aswegen 1998.** Mine layout, geological features and seismic hazard. *SIMRAC mid year report GAP303*. Pretoria. Department of Minerals and Energy.

### Summary

This report covered

- the first attempt to take into account the role of the geological fabric of the rockmass in a quantitative way,
- progress on the modelling of mine design for seismic hazard evaluation,
- progress on the quantitative and qualitative description of seismic hazard based on seismic data and

- the first attempt towards providing a potential seismic damage map for a mining area based on the combined use of the most useful tools available.

The quantification of the size distribution of faults helps to define the macro-fabric of the rock mass. The macro-fabric, the meso-fabric (hand specimen scale to the scale of mine openings) and the lithology control the rock strength. The meso-fabric is primarily important for strata-control purposes, while the macro-fabric plays an important role in seismic hazard associated with large seismic events.

The quantification of the size distribution of faults is done in the same way as for seismic events and, as in the case of the Gutenberg-Richter analysis,

1. one parameter (*a*) is a measure of the frequency of the faults while
2. a second (*b*) is a measure of the relation between the frequency and the size of the structures.

The correlation of apparent stiffness with the *a* parameter suggests that apparent stiffness reflects aspects of the inherent strength characteristics of the rockmass, while seismic stiffness more directly measure the rockmass response to mine design and volume of production.

Numerical modelling with the objective to develop methods to compare the effects of varying mine layouts on geological structures and to develop an improved engineering basis for the assessment of seismic potential as a function of mine layouts, was initiated. The results were consistent with the interpretation that a steeper gradient of the E-M relation (empirical relation between seismic energy radiated and seismic moment for a series of events which differ in magnitude) reflects greater system stiffness. A basis was established for the development of a method to estimate the potential size distribution of seismic events for given mine design.

In an application of the concept of seismic release deficiency, seismic hazard magnitude was derived from the cumulative potential seismic moment and seismic energy release between  $M_{min}$  and  $M_{max}$  within one year as a function of the *a* and *b* values of the Gutenberg-Richter relation. A 3D scatter plot of seismic hazard magnitude (derived from traditional statistical seismology), seismic stiffness (from quantitative seismology) and apparent stress level (from the E-M relation) provided a comprehensive seismic hazard assessment in one simple graphical presentation - a possible base for standardised seismic hazard assessment in mines.

The relation between the seismic response to mining and production appears to be well represented by time history analyses of the ratio between cumulative apparent volume and volume mined (now routinely used as the SRP - Seismic Response to Production). Any significant deviation from an average value of this ratio could represent an increase in seismic hazard, either because the seismic response lags behind (energy builds up) or the system over reacts to production (system is unstable).

A first attempt to put most of the GAP303 research into practice to evaluate seismic hazard for an example case involved

1. fault geometry modelling,
2. seismic event selection around the fault surface,
3. seismic hazard evaluation on selected data,
4. spatial prediction of potential large events,

5. numerical modelling to estimate potential event mechanism (not done in the particular example discussed in the report),
6. modelling potential maximum velocities of ground motion along the ore body,
7. contouring a resultant ground motion hazard map.

**Mendecki, A. J., G. van Aswegen and P. Mountford. 1999.** A guide to routine seismic monitoring in mines. (In: Jager A. J. and J. A. Ryder. *A handbook on rock engineering practice for tabular hard rock mines*. Safety in Mines Research Advisory Committee, Johannesburg. Cape Town: Creda Communications, 1999, p287-309).

**Spottiswoode, S. M. and W. de Beer. 1999.** A review of current seismic prediction and hazard assessment practice and capacity. *Deepmine final report task 5.4.1*. Johannesburg. Deepmine Consortium.

## Summary

Seismic analysis is used as part of efforts directed at hazard identification at all mines surveyed. Seismic hazard assessment was at that time based on two complementary quantities

1. Parameters closely related to strain rate, such as event rate and Cumulative Apparent Volume and
2. Parameters that characterize individual events of any size, such as stress drop or energy index.

Short-term seismic prediction was not being routinely practised on mines because methods were unproven and unquantified. Precursors in the form of accelerated changes in event rate have been seen within identified areas, but were not perceived as ubiquitous. Conversely, accelerated rates of seismicity do not always lead to damaging seismic events. In the general case, optimal hazard assessment requires complex physical understanding and statistical data manipulation. The many case studies of larger events that were preceded by accelerating rate of seismicity and/or changing event character have not been translated into practical hazard assessment. In fact many seismologists are unconvinced that parameters that are used, even by themselves, add value to simple event rate for seismic hazard assessment.

## Recommendations

1. Research is needed to improve the statistical methods and to quantify the uncertainties involved.
2. Numerous parameters are derived from the two basic measures of a point source. Consideration could be given to other parameters that could be expected to reflect a changing rock mass, such as degree of scattering measured from coda waves and velocity tomography.
3. Inevitably control of mining by management using information on hazard levels will be decided "on the ground" in each mining area.
4. There should be open debate on rockburst hazard management. Currently many people avoid the debate as they are uncomfortable with managing the uncertainties involved.

## A6.4 General: Earthquakes and earthquake prediction

**Lomnitz, C. 1997.** Fundamentals of earthquake prediction. New York: John Wiley and Sons, 326 p.

### Summary

An overview of earthquake prediction methodologies through the ages (starting from Chinese methods in antiquity).

**Kostrov, B. V. and S. Das 1988.** Principles of earthquake source mechanics. Cambridge: Cambridge University Press, 274p.

### Summary

Standard textbook.

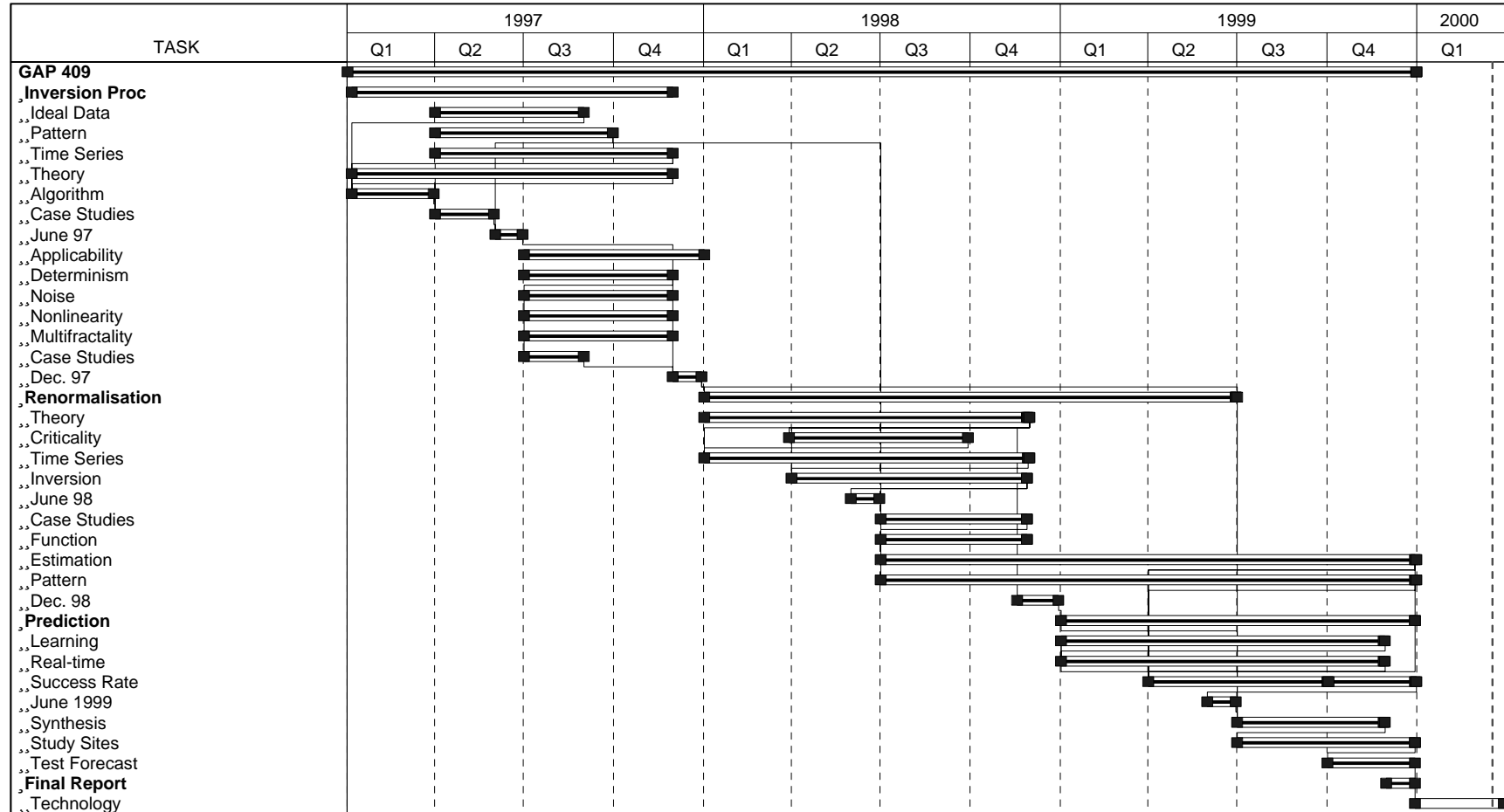
**Kravtsov, Y. A. 1993.** Fundamental and practical limits of predictability. (*In: Kravtsov, Y. A. Limits of predictability.* Springer-Verlag, Berlin, 1993, p.173-204 ).

### Summary

The author considers dynamical processes involving oscillations that obey a complex law. This example is used to extract and list factors that limit predictability, such as the fundamental gap between a physical model and reality, noise, non linearity of a process, errors in measurement, prediction costs and computation limits. An important factor is bifurcation in noisy systems. In the mining scenario, one such bifurcation is the interruption of a physical process evolving towards a large event when, in reaction to a warning, mining is changed or halted, and the rock mass relaxes aseismically or through many small events.

**GAP 409**  
Final Report - February 2000

Page 1 of 1



▬ Task

▬ Portion of task completed



# SIMRAC PROJECT GAP 409

Financial statement as at 31 December 1999	To Date	Total
<b>Income</b>	<b>1,186,392.00</b>	<b>1,482,988.28</b>
Income invoiced for		
January	296,598.00	
April	296,598.00	
July	296,598.00	
October	296,598.00	
Jan/Febr		
<b>Expenditure</b>		
<b>Project Staff expenditure</b>		
786 mandays @ 1189.98		935,324.28
728 mandays @ 1189.98	866,305.44	
<b>Operating Expenditure</b>		
Motor Vehicle expenses	25,361.45	29,837.00
Computer expenses and Maintenance	239,646.45	281,937.00
Telephone and Postage	20,591.25	24,225.00
Lectures and Presentations	18,507.90	21,774.00
Text Books and Subscriptions	18,077.80	21,268.00
Office and equipment rental	26,951.80	31,708.00
Building Maintenance	0.00	0.00
Travel and accommodation	48,377.75	56,915.00
<b>Total operating expenditure</b>	<b>397,514.40</b>	<b>467,664.00</b>
<b>Capital Expenditure</b>	<b>59,174.00</b>	<b>60,000.00</b>
<b>Sub Contracted work</b>	<b>0.00</b>	<b>20,000.00</b>
<b>Total expenditure</b>	<b>1,322,993.84</b>	<b>1,482,988.28</b>
<b>Surplus/(Shortfall)</b>	<b>(136,601.84)</b>	<b>0.00</b>

## GAP409 - SEISMOLOGY FOR ROCKBURST PREVENTION,CONTROL AND PREDICTION

ITEM DESCRIPTION	AMOUNT	Possession	Condition	In use
<b>1997 - SEISMIC EQUIPMENT</b>				
1 (ERTS @ US\$ 6000 @ R:\$ of 4.8:1)	28,800.00	Y	G	Y
1 x 3Di software	24,000.00	Y	G	Y
	52,800.00			
<b>1998 - COMPUTERS R60 000</b>				
1 x HP4050T	11,654.00	Y	G	Y
1 x High end Dell PC	34,080.00	Y	G	Y
1 x XDi SOFTWARE(US\$2500@R:\$ OF 6:1)	15,000.00	Y	G	Y
	60,734.00			
<b>1999 - COMPUTERS R60 000</b>				
1 X MATHEMATICA SOFTWARE	8,990.00	Y	G	Y
1 X SDK SOFTWARE	9,000.00	Y	G	Y
1 X DEC ALPHA + memory upgrade	40,174.00	Y	G	Y
	58,164.00			

# Safety in Mines Research Advisory Committee

## Project Summary: GAP409

<b>Project Title:</b>	Seismology for rockburst prediction		
<b>Author(s):</b>	W. de Beer	<b>Agency:</b>	ISS International Ltd
<b>Report Date:</b>	February 2000	<b>Related Projects:</b>	GAP017
<b>Category:</b>	Gold and Platinum		

### Summary

Fatalities and injuries caused directly or indirectly by a large seismic event are damaging to employee morale, shareholder confidence and public perception of the mining industry. They also diminish progress made in other areas of health and safety. As primary output, a prediction algorithm was developed which automatically quantifies power law accelerated seismic release, learns from history and predicts characteristic events with a 53% to 65% success rate.

Three problems were focussed on:

- quantifying precursory patterns of functions defined using seismic data,
- quantifying limits of predictability and evidence of criticality,
- an automated predictor based on kinematics of failure (time to failure accelerating seismic release curve).

The research perspective is physics, viewing the rock mass as a complex system approaching and retreating from criticality: larger instabilities normalise stress regimes for a time.

Furthermore, a pattern recognition algorithm for characterising the seismic response of rock to mining was developed. The algorithm quantifies previous intuitive, qualitative analyses.

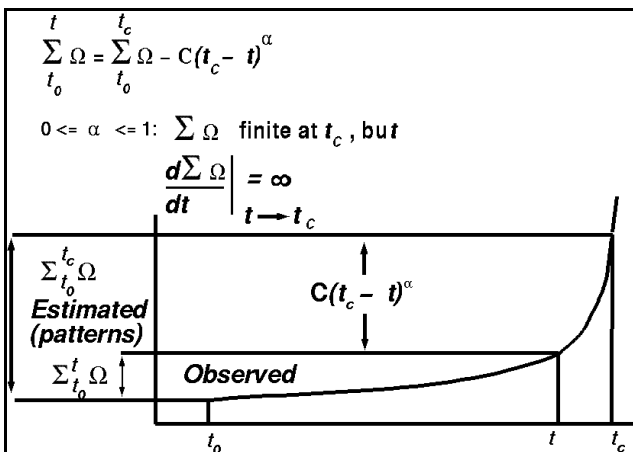


Figure 1. Graphical representation of critical accelerating behaviour by a variable (power law behaviour)

### Conclusions

- Simple, quantitative examples serve to illustrate that seismicity is not a random phenomenon and therefore not somehow inherently unpredictable.
- Temporal patterns in seismicity can be quantified.
- Abundant, quantitative evidence for criticality and deterministic, reasonable limits of predictability exists.
- The prediction algorithm can achieve uncertainties in the predicted time of occurrence of a large instability of less than a week.
- The prediction strategies followed are very far from random, and present a significant improvement, by a factor of about 4 over a pessimist random forecasting strategy.

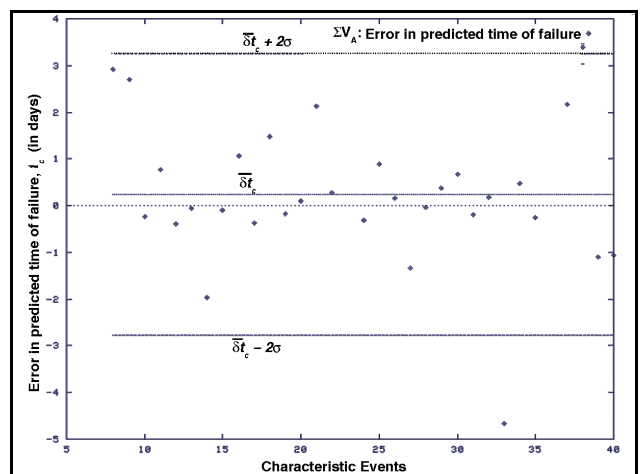


Figure 2. Average error made in the prediction of the time to failure in a Free State pillar extraction by using the SOOTHSAY algorithm and using time series of cumulative apparent volume.