

Ergodicity and seismicity clustering with applications in statistical seismology

(Spline title: Ergodic Applications to Seismicity Clustering)

(Thesis format: Integrated-Article)

by

Nelson F. Cho

Graduate Program in
Geophysics

A thesis submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

The School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada
August, 2011

© Nelson F. Cho, 2011

THE UNIVERSITY OF WESTERN ONTARIO
THE SCHOOL OF GRADUATE AND POSTDOCTORAL STUDIES
CERTIFICATE OF EXAMINATION

Supervisor:

Examiners:

Dr. Kristy Tiampo

Dr. Ilya Zaliapin

Dr. Lalu Mansinha

Dr. Paul Wiegert

Dr. Gerhard Pratt

The thesis by

Nelson F. Cho

entitled:

**Ergodicity and seismicity clustering with applications in statistical
seismology**

is accepted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

Date

Chair of the Thesis Examination Board

Abstract

The implications of understanding fault networks as ergodic systems are addressed here by using the Thirumalai-Mountain metric that identifies effective ergodic periods, when long but finite time intervals are considered. This framework was proven to be useful in statistical seismology studies. Initially, it was established that that the metric can be used to quantify seismicity clustering. Periods of effective ergodicity were characterized by the occurrence of declustered seismicity. This interpretation was implemented for synthetic and seismic data from southern California and Canadian mines.

Next, methods used in the estimation of regional seismic hazard were applied to mining induced seismicity. The interpretation of this metric as a measurement of seismicity clustering was fundamental to the adaptation of these techniques to mining seismicity. The latter also provided a better understanding of the mechanism for increased efficiency of hazard assessment methods based on seismic patterns during these ergodic periods.

In addition, this metric was employed in order to develop a robust seismic declustering technique that does not depend on a large number of parameters. Optimization methods and the Gutenberg-Richter law were used as constraints to identify clustered events in a given dataset. This method was applied to a synthetic catalog and seismic data from regions with different tectonic settings: southern California, Taiwan, Switzerland and the Gibraltar arc. Comparison to other declustering methods applied to the datasets did not show the same success due to their parameter dependence.

The effects of location errors on a particular pattern-based seismic hazard assessment technique was also studied. Perturbed catalogs were generated from the southern Californian dataset by adding noise to epicenter locations. Seismicity trends identified by the metric did not change with the increase in noise levels. A combination of the latter and the large number of small events offset the effects of location errors in the performance of the considered method in retrospective forecasts, where no systematic degradation was found. This indicates that these uncertainties do not affect the technique significantly. The same occurred for smaller catalogs.

Finally, remarks on the advantages and limitation of the framework are discussed along with suggestions of future work.

Keywords

Ergodic hypothesis, earthquake clustering, seismicity declustering, seismic hazard assessment, mining seismicity, southern California

Co-Authorship

This thesis is prepared in integrated-article format with the following manuscripts written by Nelson Cho:

1. N.F. Cho, K.F. Tiampo, S.D. McKinnon, J.A. Vallejos, W. Klein and D. Dominguez, A Simple Metric to Quantify Seismicity Clustering, *Nonlinear Processes in Geophysics*, 2010, 17, 293302.
2. N.F. Cho, K.F. Tiampo, J.A. Vallejos and S.D. McKinnon, Resubmitted to *Geophysical Journal International*.
3. N.F. Cho, K.F. Tiampo, P. Bhattacharya, R. Shcherbakov and W. Klein, Declustering seismicity using the TM metric, submitted to *BSSA*.
2. N.F. Cho, K.F. Tiampo, submitted to *Pure and Applied Geophysics*.

The work for these projects was completed under supervision and financial support of NSERC and WSIB. The Taiwanese catalog was provided by the Central Weather Bureau of Taiwan. The southern Californian dataset was obtained from ANSS and NCSN. The Spanish catalog was obtained from IGN. The Swiss catalog was obtained from the Swiss Seismological Service. The mining data was obtained under the WSIB grant.

Epigraphy

“Para quem está se afogando, jacaré é tronco.”

popular saying

Acknowledgments

How often can lightning strike the same target? How do people come up with 1/29387631 odds for these phenomena? I have been extremely fortunate over the years with the opportunity to meet good people from all over the world. Thank you to God and my family, especially mother and grandpa. You have always given me everything one could ask for.

A special thank you to Dr. Tiampo and her guidance. From being intimidated by my supervisor in the beginning of this journey to considering myself lucky to have had the opportunity to know her at the end, this has been an interesting experience. More than an excellent supervisor and researcher, she is an incredible person of great character.

Thank you to all the people that have somehow contributed to this work, let it be in a pub or at the great outdoors. In special the co-authors of the different papers that compose this thesis, JB, Roman, Zaheer, Patti, Caitlin, Pablo, Sergey, Hsien-chi, Pathikrit (and his long talks), Ranjeet, Jalpa, Ayumi, Dr. Mansinha, Dr. Shcherbakov, Javier, Dr. McKinnon, my labmates, Bernie, Dr. Atkinson, and the professors/staff in the department.

It is impossible to thank all the good people I have met over the years individually. It is a long list and I never had a respectable memory, so thank you to all the soccer buddies, beer buddies, camping buddies, department buddies, buddies from home, and all the other buddies out there.

How often can lightning strike the same target? Well, I don't know. But from the

list of good people above, I am pretty lucky. These are not the most beautiful words, there might be some English mistakes roaming around and that initial question made much more sense before, but my gratitude is sincere. Thank you all... ”obrigado pelo futebol!”.

Contents

Title page	i
Certificate of Examination	ii
Abstract and Keywords	iii
Co-Authorship	v
Epigraphy	vi
Acknowledgments	vii
Table of Contents	ix
List of Figures	xiii
List of Tables	xxiv
1 General Introduction	1
1.1 Background	3
1.1.1 Ergodic Hypothesis	3
1.1.2 Seismicity Patterns	6
1.1.3 Particle Swarm Optimization	8
References	15
2 A Simple Metric to Quantify Seismicity Clustering	22
2.1 Introduction	22

2.2 The Thirumalai-Mountain Metric	24
2.3 Results	28
2.3.1 Synthetic Catalog	28
2.3.2 Southern California	32
2.3.3 Mining Seismicity	38
2.4 Conclusions	41
2.5 Appendix	43
References	45
3 Pattern Informatics in Mining Induced Seismicity and its Applications to Rockburst Hazard Assessment	49
3.1 Introduction	49
3.2 Data and Methods	52
3.2.1 Mining Data	52
3.2.2 The Pattern Informatics and Relative Intensity	55
3.2.3 The Heidke Skill Score	57
3.2.4 The Thirumalai-Mountain Metric	60
3.3 The PI in mining seismicity	61
3.3.1 The Magnitude of Completeness	63
3.3.2 Spatial/Temporal Discretization	66
3.3.3 Forecast Magnitude	70

3.3.4 PI and RI in Mining Induced Seismicity	72
3.4 Conclusion	79
References	81
4 Declustering seismicity using the TM metric	88
4.1 Introduction	88
4.2 The TM Metric and Clustering	91
4.3 Declustering Method	93
4.4 Results	96
4.4.1 Synthetic Catalog	98
4.4.2 Southern California	104
4.4.3 Taiwan	115
4.4.4 Switzerland	121
4.4.5 Southern Spain and northern Africa	128
4.5 Conclusion	136
References	140
5 Effects of location errors in the Pattern Informatics	147
5.1 Introduction	147
5.2 Methodology	149
5.2.1 The Pattern Informatics	149

5.2.2	The Heidke and Pierce Skill Scores	151
5.2.3	The TM Metric and Clustering	152
5.2.4	Generation of Perturbed Catalogs	154
5.3	Results	155
5.4	Conclusion	167
	References	170
6	Conclusion	174
6.1	Future Work	177
	References	179
A	Computer code	181
A.1	decluster.m	181
A.2	ergodicity.m	184
A.3	PSO.m	185
A.4	locations.m	188
A.5	pi_score.m	189
A.6	moore_box_update.m	191
VITA		195

List of Figures

2.1	<i>Plot of the two terms in Eq. (2.3) for four different cases. Figure 2.1a represents the case with 10000 events randomly distributed in time and space, Figure 2.1b obtained from clustering 600 and 1500 out of the 10000 events at $t = 0.2$ and 0.7 respectively, Figure 2.1c is obtained by clustering 600 and 1500 out of the 10000 events spatially with the same latitudes and longitudes and Fig. 2.1d is obtained from the combination of the latter two scenarios.</i>	29
2.2	<i>Inverse TM metric for the different scenarios described in Fig. 2.1. From the uppermost to the lowermost solid line: random catalog, temporal clustering only, spatial clustering only, and the combination of both temporal and spatial clustering.</i>	30
2.3	<i>Plot of Eq. (2.4) for the southern California dataset from 1932 to 2006. Dash-dot lines are linear regressions for the three effective ergodic periods from 1933 to 1951, from 1955 to 1978, and from 1980 to 1991. Upside triangles locate the years of the occurrences of large earthquakes: the 1952 Kern County, the 1979 Imperial Valley, the 1992 Landers and the 1999 Hector Mine.</i>	33

2.4	<i>Distribution of events before the 1952 Kern County earthquake before the mainshock (top left), after the mainshock (top right) and up until the mainshock since 1932 (bottom right). The approximation in Eq. (2.4) with the time of occurrence of the 1952 Kern County event as the dashed line are also plotted (bottom left).</i>	35
2.5	<i>Distribution of events before the 1992 Landers earthquake before the mainshock (top left), after the mainshock (top right) and up until the mainshock since 1932 (bottom right). The approximation in Eq. (2.4) with the time of occurrence of the 1992 Landers event as the dashed line are also plotted (bottom left).</i>	36
2.6	<i>Distribution of events around the 1989 Loma Prieta earthquake: before the mainshock (top left), after the mainshock (top right) and up until the mainshock since 1932 (bottom right). The approximation in Eq. (2.4) with the time of occurrence of the 1989 Loma Prieta event as the dashed line are also plotted (bottom left).</i>	37
2.7	<i>Plot of Eq. (2.4) for Kidd Creek D (solid line) and the number of events recorded per time bin t (dashed lines).</i>	39
2.8	<i>Plot of Eq. (2.4) for Macassa mine (solid line) and the number of events recorded per time bin t (dashed lines)</i>	39
3.1	<i>Histogram of blasts for Kidd Creek D (a) and Macassa (b) mines that highlight the difference in blasting activity between the mines. The first mine has frequent low-magnitude blasts whereas the second has rarer high-magnitude blasts.</i>	53

3.2	<i>Time series of blasts for Kidd Creek D (a) and Macassa (b) mines per week. At Kidd Creek D mine, there is always blasting activity in the period considered whereas Macassa mine has periods in which blasting is halted.</i>	54
3.3	<i>Cumulative frequency-magnitude distribution of events for Kidd Creek D (a) and Macassa (b) mines. Solid lines represent the power law.</i>	64
3.4	<i>Inverse TM metric and seismicity rates for a 3D mesh at Kidd Creek D mine considering: 7-day time series with cube edge of 10m (a) and 30-day time series with cube edge of 20m (b). Most features of the inverse TM metric are roughly the same for the different space/time configurations.</i>	67
3.5	<i>Inverse TM metric and seismicity rates for a 3D mesh at Macassa mine considering: 7-day time series with cube edge of 10m (a) and 30-day time series with cube edge of 20m (b). Most features of the inverse TM metric are roughly the same for the different space/time configurations.</i>	68
3.6	<i>Inverse TM metric for Kidd Creed D (a) and Macassa (b) mines considering events with magnitude greater than $-1.5M$ and $-1.07M$ respectively. The same effective ergodic period highlighted in Figure 3.4 is observed for Kidd Creek D mine and a new effective ergodic period is observed for Macassa mine from the 66th onwards.</i>	71

3.7	<i>Plot of the maximum values of the Heidke score for a 7-day time series with cube edge of 10m for both PI and RI at Kidd Creek D mine for $\Delta t = 2$ (a) and 4 weeks (b). A magnitude forecast of $-1.25\mathbf{M}$ was considered in this case.</i>	74
3.8	<i>Plot of the maximum values of the Heidke score for a 7-day time series with cube edge of 10m for both PI and RI at Macassa mine for $\Delta t = 2$ (a) and 4 weeks (b). A magnitude forecast of $-1.07\mathbf{M}$ was considered in this case.</i>	75
3.9	<i>Plot of the maximum values for the Heidke score for a 7-day time series with cube edge of 10m for both PI and RI at Kidd Creek D (a) and at Macassa (b) mine for $\Delta t = 4$. Forecast magnitudes of $-0.73\mathbf{M}$ and $-0.7\mathbf{M}$ was considered for Kidd Creek D and Macassa mines, respectively.</i>	76
4.1	<i>Variance of $n_j(t)$ for (a) the entire catalog along with the reference line. (b) Enlargement of the reference line and the variance of both the true declustered portion of the catalog and the events identified as declustered by the TM method. (c) Plot of the seismicity rates of the entire catalog, the original declustered portion and the identified declustered seismicity.</i>	99
4.2	<i>Frequency-magnitude distribution of the original synthetic dataset. . .</i>	100
4.3	<i>Frequency-magnitude distribution for the synthetic (a) clustered and (b) declustered events, along with the distributions for the (c) clustered and (d) declustered events identified by the proposed method.</i>	101
4.4	<i>Spatial distribution of the (a) original clustered synthetic seismicity along with the (b) identified clustered seismicity by the TM metric. . .</i>	102

4.5	<i>(a) Spatial distribution and (b) rate of occurrence of events recorded with magnitudes $3M$ or greater from 1933 to 2006 in southern California.</i>	106
4.6	<i>Plot of the reference line to perform the declustering and the variances of $n_j(t)$ for the original catalog and the declustered seismicity obtained from the TM and Reasenber’s method from southern California for (a) events recorded from 1933 to 1971 and (b) an enlargement of the variances for the declustered catalogs. The same plots for events recorded in the same region from 1972 onwards are shown in (c) and (d), respectively.</i>	107
4.7	<i>Frequency-magnitude distribution for the entire catalog of southern California for (a) the entire period, (b) 1933-1971 and (c) 1972-2006.</i>	108
4.8	<i>Spatial distribution of (a) clustered and (b) declustered events identified by the TM metric along with (c) clustered and (d) declustered events identified by Reasenber’s method for events recorded in the 1940-1971 period.</i>	109
4.9	<i>Spatial distribution of (a) clustered and (b) declustered events identified by the TM metric along with (c) clustered and (d) declustered events identified by Reasenber’s method for events recorded in the 1975-2006 period.</i>	110

4.10	<i>Plot of seismicity rates for the (a) declustered and (b) clustered seismicity identified by the method based on the TM metric and Reasenberg's algorithm for the period 1940-1971. The same plots for the identified (c) declustered and (d) clustered events for the 1975-2006 period from both methods.</i>	111
4.11	<i>Frequency-magnitude distribution for (a) all clustered events identified by the TM method from 1940 until 1971 and 1975 until 2006 and (b) aftershock sequences extracted from the latter for the 1992 Landers earthquake. A square area of $1.25^\circ \times 1.25^\circ$ and events recorded in the year of the mainshock and the following year were considered.</i>	112
4.12	<i>Spatial distribution of the entire Taiwanese catalog for events with magnitude $3.4M$ or greater recorded from 1987 to 2008.</i>	116
4.13	<i>(a) Plot of the declustering reference line and the variances of $n_j(t)$ for the original catalog and the declustered seismicity obtained from the TM and Reasenberg's method for Taiwan, from 1987 to 2008. (b) Enlargement of the variances of $n_j(t)$ for the Taiwanese catalog along with its declustered portion and the reference line for the TM declustering method.</i>	117
4.14	<i>Yearly seismicity rate for the entire Taiwanese catalog from 1987 to 2008 along with declustered portions obtained from the TM and Reasen-berg's method from 1990 to 2008.</i>	118

4.15	<i>Frequency-magnitude distribution for (a) the original Taiwanese catalog from 1987 until 2008, (b) its clustered portion as identified by the TM metric from 1990 until 2008, (c) the declustered portion for the same period and (d) the seismicity associated with the 1999 Chi-chi earthquake. For the latter, the seismicity identified as clustered recorded in 1999 and 2000 within a $1^\circ \times 1^\circ$ square centered at the mainshock was considered in the aftershock sequence.</i>	119
4.16	<i>Spatial distribution of the (a) clustered and (b) declustered seismicity identified by the TM metric and (c) clustered and (d) declustered seismicity identified by Reasenberg's method in Taiwan from 1990 to 2008.</i>	120
4.17	<i>(a) Spatial distribution of the entire Swiss catalog for events with magnitude $1.6M$ or greater from 1983 until 2008. (b) The frequency-magnitude distribution for the entire Swiss catalog for the same time period.</i>	122
4.18	<i>(a) Plot of the reference line to perform the declustering and the variances of $n_j(t)$ for the original catalog and the declustered seismicity obtained from the TM and the GK method from Switzerland for events recorded from 1988 to 2008 and (b) enlargement of the variances for the declustered catalogs.</i>	123
4.19	<i>Seismicity rates for the entire Swiss catalog from 1988 until 2008 and the declustered seismicity identified by the TM metric and the GK method from 1991 until 2008.</i>	124

4.20	<i>Spatial distribution of (a) clustered and (b) declustered events identified by the TM metric along with (c) clustered and (d) declustered events identified by the Gardner-Knopoff method for the Swiss catalog from 1991 until 2008.</i>	125
4.21	<i>Frequency-magnitude distribution for the (a) clustered and (b) declustered portions of the Swiss catalog identified by the TM metric from 1991 until 2008.</i>	126
4.22	<i>(a) Spatial distribution of the entire Spanish catalog for events with magnitude $3M_b$ or greater from 1990 until 2010. (b) The frequency-magnitude distribution for the entire Spanish catalog for the same time period.</i>	129
4.23	<i>(a) Plot of the reference line to perform the declustering and the variances of $n_j(t)$ for the original catalog and the declustered seismicity obtained from the TM and Reasenbergs method from Spain for events recorded from 1990 to 2010 and (b) enlargement of the variances for the declustered catalogs.</i>	130
4.24	<i>Seismicity rates for the entire Spanish catalog from 1990 until 2010 and the declustered seismicity identified by the TM metric and Reasenbergs methods from 1993 until 2010.</i>	131
4.25	<i>Spatial distribution of (a) clustered and (b) declustered events identified by the TM metric along with (c) clustered and (d) declustered events identified by Resenberg’s method from 1993 until 2010.</i>	132

4.26	<i>Frequency-magnitude distribution for the (a) clustered and (b) declustered portions of the Spanish catalog identified by the TM metric from 1993 until 2010.</i>	133
5.1	<i>Spatial distribution of results from equation (5.4) for the training period Δt_1 considering (a) no noise, (b) $\sigma_n = 0.007^\circ$, (c) $\sigma_n = 0.05^\circ$ and (d) $\sigma_n = 0.1^\circ$. Events with magnitude M_f or greater that occurred during the forecast period are represented by open blue circles. Color scale represents $\log_{10}(PI/PI_{max})$.</i>	157
5.2	<i>Inverse TM metric of seismicity recorded from 1932 until 2000 along with plots for the datasets obtained from different noise levels σ_n. For these plots, $D_e = 1$ for a better comparison of clustering intensity of the different catalogs.</i>	158
5.3	<i>Plot of the (a) HSS and (b) PSS for Δt_1 and different threshold levels.</i>	159
5.4	<i>Distribution of maximum HSS from (a) to (d) and PSS from (e) to (h) for Δt_1 and $\sigma_n = 0.007^\circ$, 0.05°, 0.1° and 0.2°, respectively. Red lines represent normal distributions generated from the means and standard deviations σ_{SS} of the score values in each case, while the dot-lines represent the $\pm 2\sigma_{SS}$. Green lines represent the maximum score obtained from the unperturbed catalog.</i>	160

5.5	<p><i>Distribution of maximum HSS from (a) to (d) and PSS from (e) to (h) and Δt_2 for $\sigma_n = 0.007^\circ$, 0.05°, 0.1° and 0.2°, respectively. Red lines represent normal distributions generated from the means and standard deviations σ_{SS} of the score values in each case, while the dot-lines represent the $\pm 2\sigma_{SS}$. Green lines represent the maximum score obtained from the unperturbed catalog.</i></p>	161
5.6	<p><i>Spatial distribution of results from equation (5.4) considering the cropped catalog, where $N_r = 22500$ events are removed, for the training period Δt_1 considering (a) no noise, (b) $\sigma_n = 0.007^\circ$, (c) $\sigma_n = 0.1^\circ$ and (d) $\sigma_n = 0.2^\circ$. Events with magnitude M_f or greater that occurred during the forecast period are represented by open blue circles. Color scale represents $\log_{10}(PI/PI_{max})$.</i></p>	162
5.7	<p><i>Inverse TM metric of seismicity of part of the seismicity recorded from 1932 until 2000. In this case, more than 20000 events were randomly removed from the original catalog. The inverse TM metric is also plotted for the datasets obtained from the addition of different noise levels σ_n to this cropped dataset. For these plots, $D_e = 1$ for a better comparison of clustering intensity of the different catalogs.</i></p>	163

5.8 *Distribution of maximum HSS from (a) to (c) and PSS from (d) to (f) for Δt_1 and $\sigma_n = 0.007^\circ, 0.1^\circ$ and 0.2° , respectively. Events were randomly removed in this case to total 5359 events from 1932 until 2002. Red lines represent normal distributions generated from the means and standard deviations σ_{SS} of the score values in each case, while the dot-lines represent the $\pm\sigma_{SS}$. Green lines represent the maximum score obtained from the unperturbed catalog. 166*

List of Tables

1.1	<i>List of optimized parameters that were used in this work for the PSO.</i>	
	<i>The problem dimensions stand for the number of parameters in the benchmark function, w is the inertia factor, and the parameters c_1 and c_2 are the learning factors. Modified and reprinted from Pedersen (2010).</i>	11
3.1	<i>Example of a contingency table: a, b, c and d represent correct forecasts, false alarms, misses and correct negatives respectively.</i>	59
5.1	<i>Example of a contingency table: a, b, c and d represent correct forecasts, false alarms, misses and correct negatives respectively</i>	152

Chapter 1

General Introduction

One of the long-sought goals in seismology is to effectively forecast earthquakes in order to mitigate their effects when they occur in inhabited areas. Several attempts have been made to accomplish this goal by using different phenomena as precursory signals for large events over the years (Kanamori, 1981; Petersen et al., 2008; Cicerone et al., 2009), but no method is currently accepted as a standard given the complexity of the physical processes involved in earthquake faulting. Anomalous electromagnetic activity, changes in groundwater levels and surface deformation are a few examples of these precursory signals. In particular, historic seismic activity has been extensively used in different studies (Field, 2007; Wiemer et al., 2009). The Pattern Informatics (PI) (Rundle et al., 2002; Tiampo et al., 2002) and the Relative Intensity (RI) (Holliday et al., 2005) are two methods that rely on past seismicity for hazard assessment: the first quantifies changes in seismicity rates whereas the second evaluates the long-term seismic activity.

The statistical analysis of seismicity, or statistical seismology, may offer valuable information, particularly in conjunction with the constant improvement in both the available data and computational capacity. Empirical distributions that describe the occurrence of earthquakes based on their magnitudes, known as the Gutenberg-Richter (GR) law (Gutenberg and Richter, 1954), or on their time of occurrence from a given initial time, the Omori Law (Omori, 1894), have been extensively used over the years for seismic hazard assessment. These distributions also have been used

as constraints for models developed to simulate the dynamics of earthquake systems based on population growth, such as the Epidemic Type Aftershock Sequence (ETAS) (Ogata, 1988) and the self-similar Branching Aftershock Sequence (BASS) (Holliday et al., 2007). Here one event directly triggers other events with given spatial and temporal distributions based upon these statistical seismology relationships. The main goal of these approaches is to offer different tools to better understand and/or improve results in topics such as seismic hazard assessment, seismicity clustering, and other branches of seismology.

It has been argued that one of the main reasons for the development of statistical seismology was to obtain a similar success to that obtained in statistical mechanics, or at least to obtain improvements in our understanding of different geophysical phenomena (Vere-Jones et al., 2005). Restricted random walks have been used to simulate the slip distribution of earthquakes, and they reproduced the GR distribution as well as the mean slip for a given fault length (Ward, 2004). A few other examples of the use of well-established statistical approaches applied in seismology include the characterization of the distribution of waiting times and the locations of events in a given region using a Lévy flight representation (Sotolongo-Costa et al., 2000), and the study of the interoccurrence time interval between earthquakes using a non-homogeneous Poisson process model (Shcherbakov et al., 2005).

The GR and Omori laws suggest that earthquake faulting systems may be understood in terms of Self-Organized Criticality (SOC) (Sornette and Sornette, 1989), where the dynamics of a given system is in a critical state (Bak et al., 1987). Forecasting earthquakes in this framework becomes a delicate matter, given that any small

perturbation may trigger large events in the system. However, Chen et al. (1991) has argued that uncertainties in self-organized critical states are weakly chaotic, meaning that there is a gradual loss in predictability but at a much slower rate than a chaotic system. This suggests that the ability to forecast large events in this context may still be possible.

1.1 Background

In this section, the basic concepts behind the work presented are discussed. The ergodic hypothesis and its applicability to real earthquake fault systems is considered. This is a fundamental part for the entire framework here developed. Optimization problems are introduced along with the method employed in Chapter 4.

1.1.1 Ergodic Hypothesis

The goal of statistical mechanics is to offer a framework to study systems with a large number N_p of particles. This is done by relating individual properties of particles to measurable bulk properties of the considered system in a macroscopic level. One of the possible definitions for the ergodic hypothesis states that, for an invariant physical quantity f ,

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T f(t) dt = \langle f \rangle \quad (1.1)$$

where the left side of (1.1) is the temporal average of f for a given particle and $\langle f \rangle$ is the ensemble average of f . This means that, over a long period of time, a single particle of the system samples all possible configurations of the phase space uniformly.

As it stands, the ergodic hypothesis is a perfect candidate to be used as a cornerstone to the development of statistical mechanics because it offers a link between microscopic states and macroscopic properties of a given system. A simple ensemble average could be used to calculate the time averages of microscopic properties with complex dynamics. However, it has been argued that (1.1) does not hold for all scenarios, such as hydrodynamics or electrodynamics (Balescu, 1975). This is due to the fact that $\langle C_i \rangle$ is time-independent, which means that (1.1) is most likely only valid for a system in equilibrium. One of the main points of debate regarding the validity of (1.1) is that a system of particles cannot cover the entire phase space from a topological standpoint (Isihara, 1971). As a result, the use of the ergodic hypothesis in statistical mechanics is avoided by some (Landau and Lifshitz, 1971; Balescu, 1975).

The employment of the ergodic hypothesis in the description of a physical system should then be performed with caution. A *quasi-ergodic hypothesis* (Isihara, 1971) may be considered, in which a particle does not cover all the phase space, but covers it to the extent that (1.1) holds true. A suggestion that an ergodic framework may be applied to earthquakes can be found in Rundle et al. (1995), where statistical properties of mean field slider block models used to describe fault systems were studied. These models consist of mass-less blocks connected to nearest neighbors by springs of constant K_C and to a loader plate that translates at a velocity V by springs with constant K_L . In the mean field approximation, when $K_c \gg K_L$, they were shown to be characterized as systems in equilibrium if they display small fluctuations around a fixed internal energy and enough noise is available to allow them to explore the phase space.

A system in equilibrium is a positive indication that the ergodic hypothesis may be applicable but it does not guarantee it. Nevertheless, the results in Rundle et al. (1995) suggest that the framework of ergodicity may be applicable to that slider block model. Its properties were further examined in Ferguson et al. (1999) by employing a metric developed by Thirumalai and Mountain (the TM metric) (Thirumalai et al., 1989; Mountain and Thirumalai, 1989) to demonstrate that they display effective ergodicity when $V \rightarrow 0$ and the range of interaction increases. The limit when $V \rightarrow 0$ refers to a slowly-driven system.

The TM metric was originally developed to study effective ergodicity in liquids and glasses. For a system of N particles and an observable G , the TM metric is given by

$$\Omega_G(t) = \frac{1}{N} \sum_{i=1}^N [g_i(t) - \langle g \rangle]^2 \quad (1.2)$$

where $g_i = \frac{1}{t} \int_0^t G_i(t) dt$ is the time average of $G_i(t)$ until t , $G_i(t)$ is an observable for particle i , and $\langle g \rangle = \frac{1}{N} \sum_i g_i(t)$ is the average of g_i over all particles i . The term effective ergodicity relates to periods of time that are finite but long enough to ensure that the phase space is being sampled evenly. This metric is also the spatial variance of g_i calculated for different instances of time. Effective ergodic periods are identified when $\frac{1}{\Omega_G(t)} \propto t$, which relates the convergence of temporal means of a single particle and the ensemble average in (1.2) to a random process (Mountain and Thirumalai, 1989).

Relating models that may describe earthquake faulting to systems in equilibrium, or a metastable state of equilibrium, using an ergodic framework can offer valuable

insights into the understanding of the physics of earthquakes and/or the description of phenomena related to these natural processes. The work done in Tiampo et al. (2003) and Tiampo et al. (2007) explore these possibilities by applying the TM metric to earthquake systems, considering earthquakes as point-source processes and as the physical quantity under consideration. This application has resulted in the interesting observation of a correlation between the improvement of the performance of the PI and the occurrence of periods of effective ergodicity (Tiampo et al., 2010). The goal of this work is to offer a deeper understanding of the implications and advantages of studying earthquake fault networks in terms of an ergodic dynamical system and, from this knowledge, profit on the description of different phenomena related to earthquakes.

1.1.2 Seismicity Patterns

An important aspect of statistical seismology is the study of seismicity patterns, with useful applications to declustering and hazard assessment. Declustering focuses on separating the background portion of a catalog from the events that are correlated (foreshocks, aftershocks and swarms). The first stands for the declustered seismicity and the second is the clustered portion of the catalog. The use of the background seismicity allows for an unbiased estimation of long-term seismic activities, and thus a more reliable estimation of long-term seismic hazard (Wiemer et al., 2009; Petersen et al., 2008). Several declustering methods are available (Gardner and Knopoff, 1974; Reasenberg, 1985; Marsan and Lengliné, 2008; Zaliapin et al., 2008; Wu, 2010) and no current standard practice is available.

It is a common practice to consider declustered seismicity as a set of events that

can be described as a Poissonian process in time. However, there is no absolute definition for this type of events and declustering techniques usually consider different features regarding the distances between hypocenters and the interoccurrence times between events. Simple space/time windows for different mainshock magnitudes can be used (Gardner and Knopoff, 1974), or more elaborate criteria such as using a time-dependent Poisson model to describe aftershocks (Reasenber, 1985). These methods are highly dependent on the choice of parameters and adaptive measures for their implementation to different region are required (Wiemer et al., 2009). Non-parametric methods have been also been proposed (Baiesi and Paczuski, 2004; Marsan and Lengliné, 2008; Zaliapin et al., 2008) to account for such limitations. The availability of different declustering methods leads to a high variability in the estimation of background seismicity that affect the long-term seismic hazard assessment in a region (van Stiphout et al., 2011).

Regarding hazard assessment, different seismicity patterns also have been observed prior to large earthquakes distributed worldwide: foreshocks, seismic gaps, quiescence periods, precursory swarms and doughnut patterns (Kanamori, 1981; Mogi, 1985). This suggests that the use of seismic patterns as precursory signals in earthquake forecasting is feasible, and different methods have been developed (Bowman and King, 2001; Keilis-Borok, 2002; Rundle et al., 2002). Among these is the PI method, a technique that quantifies seismic activation or quiescence levels for a region in a given time period compared to historical seismic activity. This method was applied to different regions in the world for seismic hazard assessment (Chen et al., 2005; Nanjo et al., 2005) and has also been applied to study stress shadows, regions with a

decrease in static stress (Tiampo et al., 2006).

In most analyses using statistical seismology, earthquakes are considered point processes that are described by their location (hypocenter), occurrence time, and magnitude. Their results can be affected by the uncertainties related to these measurements. Magnitude uncertainties have been shown to follow a Laplace distribution, and they affected the performance of forecasts and estimations of seismic rate in a model with simple seismicity clustering (Werner and Sornette, 2008). It has been also shown that epicenter location uncertainties may also affect the determination of background seismicity levels, but to a lesser extent when compared to influence of the choice of declustering method used (van Stiphout et al., 2011). For seismic hazard assessment methods that analyze patterns of moderate seismicity, such as the PI, the effects of location errors should be attenuated due to the large number of events considered and the maintenance of the seismic patterns.

1.1.3 Particle Swarm Optimization

Optimization is an important discipline in mathematics and computer science with a broad range of applications to problems involving decision-making. It can be thought of as an implementation of the variational principle to find the extreme value of a function, for which the Hamiltonian formalism for classical mechanics and the Fermat's principle are good examples (Goldstein et al., 2000). The first consists in obtaining the equations of motion for a particle from minimizing an observable defined as *action*, whereas the second states that light takes the path which requires the shortest time when travelling from one point to another.

There are several different methods to perform optimizations, with some inspired by phenomena observed in nature. Genetic Algorithms (GAs) (Holland, 1992) are an example of the latter, in which a fitness function is optimized by considering possible solutions to be individuals that evolve over generations. This fitness function is the measurement of how good each individual is given the problem in hand. The population of possible solutions undergoes constant changes inspired by biological evolution such as mutation, recombination, crossover and elitism to optimize a given fitness function. The best individual yields the optimum value of the fitness function and is considered the solution of the problem. These methods do not require the computation of gradients but they can be computationally expensive for a large number of parameters. They also tend to oversample probable states based upon the selection criteria, which can be both an advantage to operate a better local search or a disadvantage if this oversampling occurs near local minima.

The Particle Swarm Optimization (PSO) (Kennedy and Eberhart, 1995) is a stochastic optimization method that is based on the social behavior of animals. This framework considers a set of possible solutions that optimizes a fitness function h to be the position of birds in the search space. Each bird i is associated with a position $X_i^{(t)} = \{x_{i1}^{(t)}, \dots, x_{ik}^{(t)}\}$, where k is the number of parameters to be considered in the optimization of h , and t is the iteration in which the system is in. The best known position of bird i up until t is given by $P_i^{(t)} = \{p_{i1}^{(t)}, \dots, p_{ik}^{(t)}\}$ and the best known position of the entire flock up until t is given by $G^{(t)} = \{g_1^{(t)}, \dots, g_k^{(t)}\}$. Both $P_i^{(t)}$ and $G^{(t)}$ are determined from the values obtained for h given a set of parameters $X_i(t)$. Depending on the values of h sought after, a maximization (highest values of h) or

a minimization (lowest values of h) may be performed. The update of the positions of all birds from iteration t to $t + 1$ is performed by computing $v_{ik}^{(t+1)}$, the velocity of the parameter k of particle i , as:

$$v_{ik}^{(t+1)} = v_{ik}^{(t)} + c_1 \text{rand}() (p_{ik}^{(t)} - x_{ik}^{(t)}) + c_2 \text{Rand}() (g_k^{(t)} - x_{ik}^{(t)}) \quad (1.3)$$

where $\text{rand}()$ and $\text{Rand}()$ are random numbers between 0 and 1, c_1 and c_2 are constants usually set to be equal to 2. This velocity represents the update in the coordinate $x_{ik}^{(t)}$ to $t + 1$. The updated position is then obtained as:

$$x_{ik}^{(t+1)} = x_{ik}^{(t)} + v_{ik}^{(t+1)}. \quad (1.4)$$

A modified version for the PSO (Shi and Eberhart, 1998) has been proposed in which (1.3) is re-written as

$$v_{ik}^{(t+1)} = w v_{ik}^{(t)} + c_1 \text{rand}() (p_{ik}^{(t)} - x_{ik}^{(t)}) + c_2 \text{Rand}() (g_k^{(t)} - x_{ik}^{(t)}). \quad (1.5)$$

where $w \in [0.9, 1.2]$. The difference between the original and this modified version of the PSO is the introduction of the parameter w in the first term of the right-side of (1.5) that controls the local/global aspect of the search. The position updates of the birds in (1.4) are directly proportional to w , meaning that the greater the value of w , the farther the particle can move from iteration t to $t + 1$ in the phase space and vice-versa. This is analogous to the effects of the weight of the bird when flying, and thus this parameter is known as the *inertia factor*. The second term in (1.5) is the cognitive term for its dependence on $P_i^{(t)}$, whereas the third term is deemed a social term for its dependence on $G^{(t)}$. Because of these parallels, the constants c_1 and c_2 are called learning factors.

Problem Dimensions	PSO Parameters		
	w	c_1	c_2
20	-0.4438	-0.2699	3.395
30	-0.6031	-0.6485	2.6475
50	-0.2256	-0.1564	3.8876
100	-0.2089	-0.0787	3.7637

Table 1.1: List of optimized parameters that were used in this work for the PSO. The problem dimensions stand for the number of parameters in the benchmark function, w is the inertia factor, and the parameters c_1 and c_2 are the learning factors. Modified and reprinted from Pedersen (2010).

This method does not require the explicit computation of gradients or a large number of parameters, but it can be inefficient in local searches. Improvements to the PSO aimed at the mitigating this disadvantage have been proposed (Qin et al., 2006; Nickabadi et al., 2010), where an adaptive w that is dependent on the number of interactions in the PSO is employed. This ensures that a global search is executed in the early stages of the optimization, and a local search is performed towards its end. The improvement of the performance of the PSO was also studied by executing optimizations on the PSO parameters themselves instead of the parameters in h to determine their best values when fitting different benchmark functions (Pedersen, 2010). Table 1.1 displays a part of the results obtained in the previous study, where the Problem Dimension stands for the number of parameters in h .

In Chapter 4, the modified PSO version in (1.5) and (1.4) along with the different values of w , c_1 and c_2 in Table 1.1 was employed to decluster seismic catalogs to improve the computational performance. These parameters were chosen based on

the Problem Dimension that was the closest to the number of parameters in the optimization. The PSO was then applied as following:

1. define the fitness function h , the number of particles to be considered, and the range in which the parameters are allowed to vary.
2. define the optimal parameters w , c_1 and c_2 from Table 1.1 to be used.
3. define the initial set of $X_i^{(t^*)}$ and $v_{ik}^{(t^*)}$ for $t^* = 0$ from a uniform random distribution.
4. define $P_i^{(t^*)} \leftarrow X_i^{(t^*)}$ and $G^{(t^*)}$ as the best $X_i^{(t^*)}$.
5. compute (1.5) and then (1.4) to obtain $v_{ik}^{(t^*+1)}$ and then the updated $X_i^{(t^*+1)}$
6. determine if the new positions $X_i^{(t^*+1)}$ perform better than the previous ones. If so, update $P_i^{(t^*+1)}$.
7. determine if any of the new positions perform better than $G^{(t^*)}$. If so, update $G^{(t^*+1)}$.
8. repeat items 4-6 for $t^* \leftarrow t^* + 1$ until the convergence criteria is met.

In this work, the TM metric will be used to study seismicity patterns. In Chapter 2, the physical meaning of the TM metric is analyzed when applied to earthquake fault systems. It is shown that this metric can be used to quantify seismicity clustering and that this measurement is dependent on past seismicity. The occurrence of declustered seismicity is related to the periods of effective ergodicity, providing a

physical interpretation for such phenomena. The understanding of this metric as a measurement of clustering is applied to different applications to statistical seismology.

Chapter 3 consists of the analysis of seismic hazard in mines located in northern Ontario, Canada, with two methods originally applied to regional seismicity: the PI and the RI. The interpretation of the TM metric as a measurement of clustering was crucial to the adaptation of both methods to a mining seismicity scale. A higher performance of the PI was observed during periods of effective ergodicity, as noted by previous studies on regional seismicity (Tiampo et al., 2010). This was attributed to the declustered nature of seismicity during effective ergodic periods, allowing for a better identification of the quiescence/activation signals that the PI is sensitive to. The identification of effective ergodicity in these small scales show that metastable states can be obtained when blasting activity is regular. This suggests that seismicity in smaller scales may still maintain some features observed in regional seismicity in particular cases.

Chapter 4 focuses on the development of a new declustering technique for seismicity based on the TM metric. It consists of removing events so that the remaining catalogs display a long effective ergodic period, during which seismicity is considered declustered. The removal of events was managed by an optimization method, the PSO. This choice was based on its simple implementation and the better performance compared to GAs in preliminary tests. The GR distribution was used as a constraint to constrain the particular choice of events to be removed from a given catalog. The method was applied to a synthetic catalog and four regions with different shear strain rates (Kreemer et al., 2003): southern California, Taiwan, Switzerland

and a combination of southern Spain and northern Africa. The proposed method performed well for all different regions with no major change in its implementation, displaying a relative lack of sensitivity to regional tectonics.

Finally, the effects of location errors in the performance of the PI method were analyzed in Chapter 5. The southern Californian dataset was used to generate perturbed catalogs with the addition of different levels σ_n of normally distributed random noise to epicenter locations. No systematic degradation of the efficiency of the PI was observed as the noise levels increased, but two isolated instances of decline in the PI performance were identified for $\sigma_n = 0.007^\circ$ and 0.2° . These were likely due to the nature of the methods used to quantify the PI performance. The effects of the total number of events was also considered by randomly removing events from the original catalog, with no significant decrease in the performance for the PI. The TM metric did not vary considerably from the clustering trends of the original dataset in all cases, indicating that the clustering features remain relatively invariant.

References

- Baiesi, M. and Paczuski, M. (2004). Scale-free networks of earthquakes and aftershocks. *Physical Review E*, 69(066106).
- Bak, P., Tang, C., and Wiesenfeld, K. (1987). Self-organized criticality: an examination of the $1/f$ noise. *Physical Review Letters*, 59(4):381–384.
- Balescu, R. (1975). *Equilibrium and nonequilibrium statistical mechanics*. John Wiley and Sons.
- Bowman, D. D. and King, G. C. P. (2001). Accelerating seismicity and stress accumulation before large earthquakes. *Geophysical Research Letters*, 28(21):4039–4042.
- Chen, C. C., Rundle, J. B., Holliday, J. R., Nanjo, K. Z., Turcotte, D. L., Li, S. C., and Tiampo, K. F. (2005). The 1999 chi-chi, taiwan, earthquake as a typical example of seismic activation and quiescence. *Geophysical Research Letters*, 32(L22315):1–4.
- Chen, K., Bak, P., and Obukhov, S. P. (1991). Self-organized criticality in a crack-propagation model of earthquakes. *Physical Review A*, 43(2):625–630.
- Cicerone, R. D., Ebel, J. E., and Britton, J. (2009). A systematic compilation of earthquake precursors. *Tectonophysics*, 476:371–396.
- Ferguson, C. D., Klein, W., and Rundle, J. B. (1999). Spinodals, scaling, and ergodicity in a threshold model with long-range stress transfer. *Physical Review E*, 60:1359–1373.

- Field, E. H. (2007). Overview of the working group for the development of regional earthquake likelihood models (relm). *Seismological Research Letters*, 78:7–16.
- Gardner, J. K. and Knopoff, L. (1974). Is the sequence of earthquakes in southern California, with aftershocks removed, poissonian? *Bulletin of the Seismological Society of America*, 64(5):1363–1367.
- Goldstein, H., Poole, C., and Safko, J. (2000). *Classical mechanics*. Addison Wesley.
- Gutenberg, B. and Richter, C. F. (1954). *Seismicity of the earth and associated phenomena*. New York, Hafner Pub. Co.
- Holland, J. (1992). *Adaptation in natural and artificial systems : an introductory analysis with applications to biology, control, and artificial intelligence*. MIT Press, Cambridge, Mass.
- Holliday, J. R., Nanjo, K. Z., Tiampo, K. F., Rundle, J. B., and Turcotte, D. L. (2005). Earthquake forecasting and its verification. *Nonlinear Processes in Geophysics*, 12:965–977.
- Holliday, J. R., Turcotte, D. L., and Rundle, J. B. (2007). Self-similar branching of aftershock sequences. *Physica A*, 387:933–943.
- Isihara, A. (1971). *Statistical physics*. Academic Press.
- Kanamori, H. (1981). *The nature of seismicity patterns before large earthquakes, in earthquake prediction: an International Review*. American Geophysical Union.

- Keilis-Borok, V. (2002). Earthquake prediction: state-of-the-art and end emerging possibilities. *Annual Review of Earth and Planetary Sciences*, 30:1–33.
- Kennedy, J. and Eberhart, R. (1995). Particle swarm optimization. *Proceedings of IEEE International Conference on Neural Networks*, IV:1942–1948.
- Kreemer, C., Holt, W. E., and Haines, A. J. (2003). An integrated global model of present-day plate motions and plate boundary deformation. *Geophysical Journal International*, 154:8–35.
- Landau, L. D. and Lifshitz, E. M. (1971). *Statistical physics, Part 1*, volume 5. Pergamon Press.
- Marsan, D. and Lengliné, O. (2008). Extending earthquakes’ reach through cascading. *Science*, 319:1076–1079.
- Mogi, K. (1985). *Earthquake prediction*. Academic Press, NY.
- Mountain, R. D. and Thirumalai, D. (1989). Measures of effective ergodic convergence in liquids. *Journal of Physical Chemistry*, 93:6975–6979.
- Nanjo, K. Z., Holliday, J. R., Chen, C. C., Rundle, J. B., and Turcotte, D. L. (2005). Application of a modified pattern informatics method to forecasting the locations of future large earthquakes in the central japan. *Tectonophysics*, 424:351–366.
- Nickabadi, A., Ebadzadeh, M., and Safabakhsh, R. (2010). A novel particle swarm optimization algorithm with adaptive inertia weight. *Applied Soft Computing Journal*, Accepted.

- Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point process. *Journal of the American Statistical Association*, 83(401):9–27.
- Omori, F. (1894). On the aftershocks of earthquakes. *The Journal of the College of Science, Imperial University of Tokyo*, 7:111–200.
- Pedersen, M. E. H. (2010). Good parameters for particle swarm optimization. *Technical Report HL1001 (Hvass Laboratories)*.
- Petersen, M. D., Frankel, A. D., Harmsen, S. C., Mueller, C. S., Haller, K. M., Wheeler, R. L., Wesson, R. L., Zeng, Y., Boyd, O. S., Perkins, D. M., Luco, N., Field, E. H., Wills, C. J., and Rukstales, K. S. (2008). Documentation for the 2008 update of the united states national seismic hazard maps. Technical report, USGS.
- Qin, Z., Yu, F., Shi, Z., and Wang, Y. (2006). Adaptive inertia weight particle swarm optimization. *Lecture Notes in Computer Science*, 4029:450–459.
- Reasenber, P. (1985). Second-order moment of central california seismicity, 1969–1982. *Journal of Geophysical Research*, 90(B7):5479–5495.
- Rundle, J. B., Klein, W., Gross, S., and Turcotte, D. L. (1995). Boltzmann fluctuations in numerical simulations of nonequilibrium lattice threshold systems. *Physical Review Letters*, 75:1658–1661.
- Rundle, J. B., Tiampo, K. F., Klein, W., and Martins, J. S. S. (2002). Self-organization in leaky threshold systems: The influence of near-mean field dynamics and its implications for earthquakes, neurobiology, and forecasting. *Proceedings of*

- the National Academy of Sciences of the United States of America*, 99(Supplement 1):2514–2521.
- Shcherbakov, R., Yakovlev, G., Turcotte, D. L., and Rundle, J. B. (2005). Model for the distribution of aftershock interoccurrence times. *Physical Review Letters*, 95(218501).
- Shi, Y. and Eberhart, R. (1998). A modified particle swarm optimizer. *Proceedings of the IEEE Congress on Evolutionary Computation. IEEE Service Center, USA*, pages 69–73.
- Sornette, A. and Sornette, D. (1989). Self-organized criticality and earthquakes. *Europhysics Letters*, 9(3):197–202.
- Sotolongo-Costa, O., Antoranz, J. C., Posadas, A., Vidal, F., and Vásquez, A. (2000). Lévy flights and earthquakes. *Geophysical Research Letters*, 27(13):1965–1968.
- Thirumalai, D., Mountain, R. D., and Kirkpatrick, T. R. (1989). Ergodic behavior in supercooled liquids and in glasses. *Physical Review A*, 39:3563–3574.
- Tiampo, K. F., Klein, W., Li, H. C., Mignan, A., Toya, Y., Kohen-Kadoshand, S. Z. L., Rundle, J. B., and Chen, C. C. (2010). Ergodicity and earthquake catalogs: forecast testing and resulting implications. *Pure and Applied Geophysics*, 167(6–7):763–782.
- Tiampo, K. F., Rundle, J. B., and Klein, W. (2006). Stress shadows determined from a phase dynamical measure of historic seismicity. *Pure and Applied Geophysics*, 263:2407–2416.

- Tiampo, K. F., Rundle, J. B., Klein, W., Holliday, J. B., Martins, J. S. S., and Ferguson, C. D. (2007). Ergodicity in natural earthquake fault networks. *Physical Review E*, 75(066107):1–15.
- Tiampo, K. F., Rundle, J. B., Klein, W., Martins, J. S. S., and Ferguson, C. D. (2003). Ergodic dynamics in a natural threshold system. *Physical Review Letters*, 91(23):1–4.
- Tiampo, K. F., Rundle, J. B., McGinnis, S., and Klein, W. (2002). Pattern dynamics and forecast methods in seismically active regions. *Pure and Applied Geophysics*, 159(10):2429–2467.
- van Stiphout, T., Schorlemmer, D., and Wiemer, S. (2011). The effect of uncertainties on estimates of background seismicity rate. *Bulletin of the Seismological Society of America*, 101(2):482–494.
- Vere-Jones, D., Ben-Zion, Y., and Zúñiga, R. (2005). Statistical seismology. *Pure and Applied Geophysics*, 162:1023–1026.
- Ward, S. N. (2004). Earthquake simulation by restricted random walks. *Bulletin of the Seismological Society of America*, 94(6):2079–2089.
- Werner, M. J. and Sornette, D. (2008). Magnitude uncertainties impact seismic rate estimates, forecasts, and prediction experiment. *Journal of Geophysical Research*, 113(B08302):1–20.
- Wiemer, S., Giardini, D., Fäh, D., Deichmann, N., and Sellami, S. (2009). Proba-

bilistic seismic hazard assessment of switzerland: best estimates and uncertainties. *Journal of Seismology*, 13(4):449–478.

Wu, Z. (2010). A hidden markov model for earthquake declustering. *Journal of Geophysical Research*, 115(B03306):24513–24528.

Zaliapin, I., Gabrielov, A., Keilis-Borok, V., and Wong, H. (2008). Clustering analysis of seismicity and aftershock identification. *Physical Review Letters*, 101(018501).

Chapter 2

A Simple Metric to Quantify Seismicity Clustering¹

2.1 Introduction

A simple inspection of the distribution of hypocenters of earthquakes indicates that they do not occur randomly in space or time. Several attempts have been made in order to better understand the cause/consequence relationship between events (Reasenberg, 1985; Dieterich, 1994; Felzer et al., 1994; Baiesi and Paczuski, 2004; Marsan and Lengliné, 2008; Zaliapin et al., 2008). Clusters of earthquakes are commonly addressed in terms of swarms, a set of events with no single predominant event, or sequences that are highly related to a mainshock of large magnitude.

Mainshocks are often considered as stationary Poisson processes with a fixed occurrence rate over time (Kagan and Jackson, 1991) and the remainder of the associated seismicity is classified as foreshock/aftershock sequences. The definition of foreshocks and aftershocks is not absolute and the studies on seismicity clustering mentioned previously use different criteria to identify clusters. Kanamori (1981) considered different seismic patterns that can be used as precursors to large events, in which foreshocks are an important type, even though their occurrence is not mandatory prior to mainshocks. Mogi (1985) classified foreshock sequences into two types: C and D. In the first, the seismic activity increases gradually towards the mainshock.

¹Modified from *Nonlinear Processes in Geophysics*, 2010, 17, 293-302

The opposite occurs in type D sequences, in which seismicity decreases towards the mainshock.

Aftershocks are an important source of information about the mechanism of earthquake triggering and they have been widely studied over the years. Mendonza and Hartzell (1988) studied the correlation between the spatial distribution of aftershocks and the coseismic slip in faults in California. The Epidemic Type Aftershock Sequence (ETAS) model (Ogata, 1988; Helmstetter and Sornette, 2002) was proposed to understand the occurrence of aftershocks as generated by a mainshock. Utsu (2002) assembled a series of studies on seismicity, including findings on the spatial distribution of aftershocks. Shcherbakov et al. (2005) studied the interoccurrence time interval between aftershocks using a nonhomogeneous Poissonian model.

In this work, the TM metric will be interpreted as a simple measure of clustering. This metric, originally developed to study liquid systems and glasses (Thirumalai et al., 1989; Thirumalai and Mountain, 1993), was applied to earthquake simulations (Ferguson et al., 1999) and to regional seismicity by Tiampo et al. (2003, 2007). The result was the identification of periods of metastable equilibrium in seismic activity. The relationship between the effective ergodic periods and certain types of seismicity patterns was also addressed in Tiampo et al. (2010). These previous studies indicate that clustering plays an important role in the results. The interpretation presented here offers a clear understanding of how seismicity clustering affects the TM metric, especially for southern California.

This article begins with a demonstration of the framework in which the TM metric can be considered a measurement of seismicity clustering. The method is then applied

to three different cases: synthetic data, the southern California dataset, and mining induced seismicity. Remarks and limitations of the method are considered in the conclusion.

2.2 The Thirumalai-Mountain Metric

The ergodic hypothesis is an important assumption for classical statistical mechanics in order to relate micro and macro states (Farquhar, 1964; de Oliveira and Werlang, 2007). A classical definition of the ergodic hypothesis states that, for an ensemble of particles, the time average of a property $f(t)$ of a single particle and the ensemble average of the same property f are related as follow

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T f(t) dt = \langle f \rangle \quad (2.1)$$

where $\langle f \rangle$ is the ensemble average of $f(t)$.

The TM metric was first developed to study the effective ergodicity in liquids and glasses (Thirumalai et al., 1989; Thirumalai and Mountain, 1993). The same framework will be used in the present work to study seismicity. For a system comprised of N particles and an observable G , the TM metric is written as

$$\Omega_G(t) = \frac{1}{N} \sum_{j=1}^N [g_j(t) - \langle g \rangle]^2 \quad (2.2)$$

where g_j is the time average of the observable G for particle j until t and $\langle g \rangle$ is the ensemble average of g_j over all particles j at time t . Equation (2.2) is simply the spatial variance of the temporal mean of g_j . Effective ergodicity arises from the fact

that the relationship between ensemble and particle time averages are addressed for a long but finite time interval to ensure that all the phase space is sampled with equal likelihood, and it occurs when $1/\Omega_G$ is linear in time (Thirumalai and Mountain, 1993). If the time average of g_j for a particle is the same as the ensemble average of g_j , the metric is null.

Tiampo et al. (2007) applied Eq. (2.2) to historic seismicity to identify periods of effective ergodicity by dividing the region of interest into a mesh of N boxes. These are considered to be the particles of the system and the cumulative number of events per box n_j was used as a proxy for the observable G : the seismic energy released. It has been argued that out of the possible proxies available for the seismic released energy, the number of events displays the longest correlations in time (Jimenez et al., 2006).

Different magnitude cutoffs were considered in the search for effective ergodic periods in southern California and it was found that the metric changes considerably for each case (Tiampo et al., 2007). The lack of effective ergodic periods for smaller events was hypothesized to be due to the stability of the catalog of southern California for the lower magnitude range. The system is stationary during these periods of effective ergodicity, meaning that the average of the studied property is constant over the considered period, and can be considered in a state of metastable equilibrium (Tiampo et al., 2003). Metastable equilibrium is a state that the system tends to occupy unless a disturbance is strong enough to propel the system to a new, but more stable, state.

Equation (2.2) can be re-written in terms of the variance of the cumulative number

of events in each box $n_j(t)$. For simplicity, a simple translation in time is considered so that the initial time considered t_0 is set to be zero and $\Delta t = t_f - t_0 = t$. The TM metric then becomes

$$\begin{aligned}\Omega_n(t) &= \frac{1}{\Delta t^2} \frac{\sum_j^N (n_j(t))^2}{N} - \frac{1}{\Delta t^2} \left(\frac{\sum_j^N n_j(t)}{N} \right)^2 \\ &= \frac{1}{t^2} (\langle n_j(t)^2 \rangle - \langle n_j(t) \rangle^2).\end{aligned}\tag{2.3}$$

Periods in which the inverse TM metric is linear with a positive slope, $\frac{1}{\Omega_n(t)} = \frac{t}{D_e}$ are considered to be effective ergodic and in a state of metastable equilibrium. From the latter and Eq. (2.3), it can be seen that the variance of $n_j(t)$ is linear in time during the effective ergodic periods. This can be interpreted as the variance of a normal diffusion processes, and the parameter D_e can be regarded as a diffusion parameter (Tiampo et al., 2007) related to the rate in which the phase space is sampled. In this case, $D_e = \frac{1}{\Omega_n(t_0)}$ where t_0 is the initial time considered.

Equation (2.3) can be examined to analyze the effects of clustering in the TM metric. Its rightmost term is proportional to the square of the sum of cumulative events per box. It is sensitive to variations in the total number of events in each time step and thus it can be used to quantify temporal clustering of events. The same cannot be said for spatial clustering: different spatial configurations of a fixed number of hypocenters will yield the same sum of cumulative events. The left term in the right side of Eq. (2.3) can be regarded as a more complete measurement of clustering (refer to Appendix A).

A simpler form of the TM metric can be obtained if $\sum n_j \ll N$ or $\sum n_j \ll \sum n_j^2$.

Under one of these assumptions, the TM metric can be written as

$$\Omega_n(t) = \frac{1}{t^2} \langle n_j(t)^2 \rangle \quad (2.4)$$

and considering the scaling in time of the TM metric for ergodic periods,

$$\begin{aligned} \Omega_n(t) &\propto t^{-1} \\ \langle n_j(t)^2 \rangle &\propto t. \end{aligned} \quad (2.5)$$

The imposed linearity in time for the effective ergodic periods can be regarded as a benchmark to compare the evolution of seismicity clustering in time. It is important to stress, however, that the values obtained by the metric are non-unique. This means that different configurations might yield the same result.

Techniques that identify seismicity pattern changes often measure variations relative to long-term averages and generally are more accurate during ergodic periods because the spatial and temporal averages are stationary and approach the same value. Tiampo et al. (2010) showed that a better effectiveness of the Pattern Informatics (Tiampo et al., 2002; Holliday et al., 2006), a method that quantifies seismicity changes, was achieved during effective ergodic periods. The interpretation of effective ergodic periods presented here is similar to the relationship between the time evolution of the metric for liquids and systems in thermal equilibrium (Mountain and Thirumalai, 1989). In this analogy, the temperature of the system is the background seismicity and noise when dealing with the cumulative number of events and these effective ergodic periods are disrupted by the aftershock sequences from large earthquakes.

2.3 Results

2.3.1 Synthetic Catalog

The method was initially tested for a simple synthetic catalog comprised of 10000 recordings for different scenarios. These scenarios are generated by distributing events both randomly and artificially clustered in space and time. Temporal and spatial coordinates were distributed in a unity interval for simplicity. The first case consists in a catalog comprised of 10000 events randomly displaced in space and time. A time clustering scenario was achieved for the second case by agglomerating 600 and 1500 events of the initial 10000 for times $t_1 = 0.2$ and $t_2 = 0.7$ respectively. For the third case, spatial clustering was obtained by giving the same latitudes and longitudes to subset of 600 and 1500 events while maintaining their randomness in time. The last case is obtained by combining the previous two scenarios into one where the clustered subsets partially overlap to produce clusters in space, time, and both.

The initial step is to verify the feasibility of the approximation for the TM metric illustrated in Eq. (2.5). Figure 2.1 displays both terms used to compute the full form of the TM metric in Eq. (2.3) for all scenarios. The random and time clustered cases are displayed in Figs. 2.1a and 2.1b respectively. In these cases, Eq. (2.5) does not approximate the TM metric since $(\Sigma n_j)^2$ is not negligible compared to Σn_j^2 . Figure 2.1c illustrates the scenario with spatial clustering alone and it shows that $(\Sigma n_j)^2$ remains the same compared to the random catalog case in Fig. 2.1a, whereas Σn_j^2 increases considerably for all t . This demonstrates that the first term does not measure spatial clustering, as noted previously, and that the approximation in Eq.

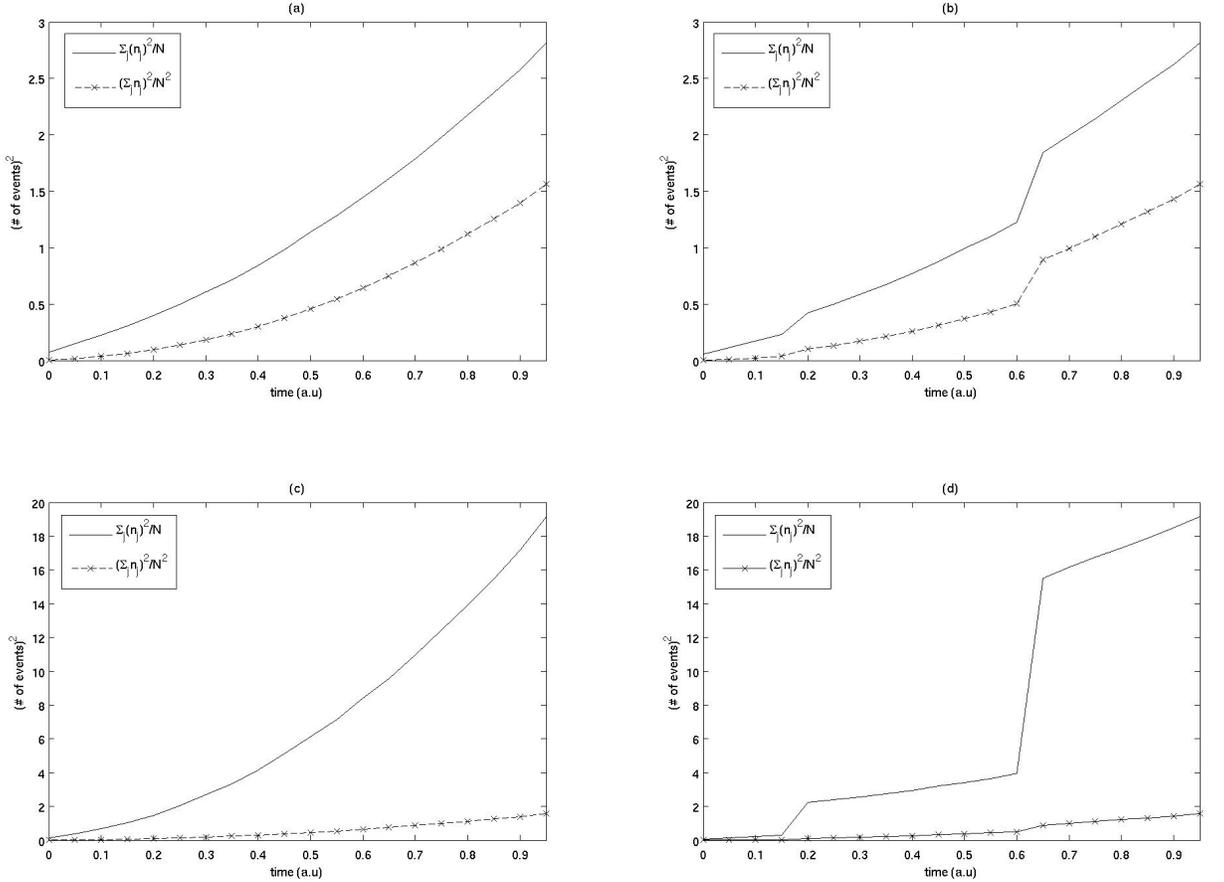


Figure 2.1: Plot of the two terms in Eq. (2.3) for four different cases. Figure 2.1a represents the case with 10000 events randomly distributed in time and space, Figure 2.1b obtained from clustering 600 and 1500 out of the 10000 events at $t = 0.2$ and 0.7 respectively, Figure 2.1c is obtained by clustering 600 and 1500 out of the 10000 events spatially with the same latitudes and longitudes and Fig. 2.1d is obtained from the combination of the latter two scenarios.

(2.5) becomes feasible in the presence of spatial clustering. The same rationale can be applied to the last scenario illustrated in Fig. 2.1d, in which temporal and spatial clustering are considered simultaneously, to validate the applicability of Eq. (2.5). In this case, $(\Sigma n_j)^2$ is the same as in Fig. 2.1b while Σn_j^2 increases substantially for all t due to the spatial clustering.

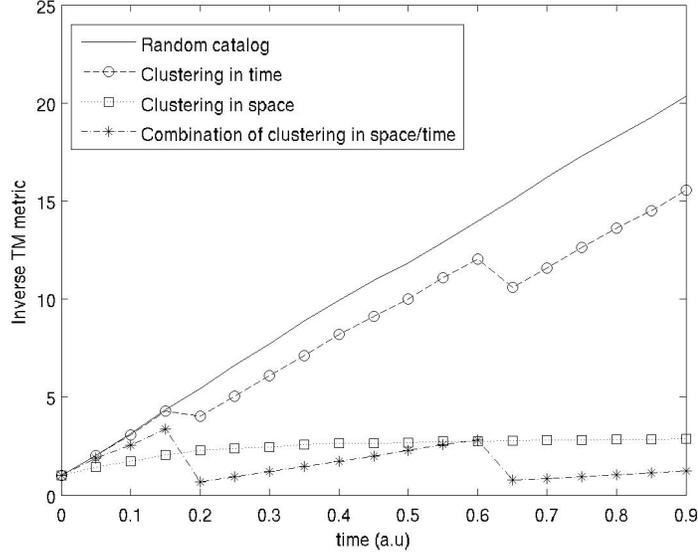


Figure 2.2: Inverse TM metric for the different scenarios described in Fig. 2.1. From the uppermost to the lowermost solid line: random catalog, temporal clustering only, spatial clustering only, and the combination of both temporal and spatial clustering.

Both terms from Eq. (2.3) display a non-linear behavior in Fig. 2.1a, scaling as t^β and t^α where $\beta, \alpha > 1$. Their combination, however, results in the linear trend of the inverse TM metric along the entire time domain observed for this case in Fig. 2.2. An examination of Eq. (2.3) and the linearity of the inverse TM metric indicate that the cumulative number of events per box n_j evolves as a normal diffusive process during this long effective ergodic period. Based on the interpretation of Tiampo et al. (2007), the period in which the inverse TM metric is linear is effectively ergodic. The original work on the TM metric in glasses and liquids relates effective ergodicity and fluids in thermal equilibrium (Mountain and Thirumalai, 1989). Considering the cumulative number of events per box n_j as a proxy for the seismic released energy and the latter statement, it can be inferred that effective ergodic periods result from configurations

in which the number of events are evenly distributed over the boxes. This suggests that these periods are characterized by non-clustered seismic activity.

Time clustering is verified from the abrupt vertical shifts of both $(\Sigma n_j)^2$ and Σn_j^2 during the instances when the imposed agglomerations occur in Figs. 2.1b and 2.1c. As discussed previously, both terms of Eq. (2.3) respond to the introduced temporal clustering. Figure 2.2 shows that the effective ergodic period displayed throughout the whole time domain for the random catalog in both time and space is disrupted during the instances when the time clustering is introduced. In between the instances when the temporal clustering is inserted, the system is in an effective ergodic state. The imposed temporal clustering in the given instants t_1 and t_2 promotes abrupt changes in the rates of seismic activity which translates to a sudden increase of the TM metric during these instances. This may be interpreted as a break in the thermal equilibrium of the system while the disruptions in the seismicity rate lasts, which means changes of the phase space subset where the system resides.

Figure 2.1c shows that the spatial clustering alone changes the scaling of Σn_j^2 to $t^{\beta'}$ while maintaining the same t^α scaling for $(\Sigma n_j)^2$ observed for the random case, where $\beta' > \beta > \alpha$. This is not a localized effect as observed for time clustering, but rather a global effect that is spread over the entire period considered. By assuming the approximation for the variance of $n_j(t)$ to be $(\Sigma n_j)^2$ and its scaling with time shown in Fig. 2.1c, it can be inferred that the cumulative number of events per time step evolves in time as a super-diffusive process. The latter differs from the normal-like diffusive behavior observed for effective ergodic periods and the result is the smooth non-linear trend observed for the inverse TM metric of the spatially clustered case

in Fig. 2.2. The Coso Geothermal Field in southeastern California is an example of a region where seismicity occurs in a swarm-like manner and might display spatial clustering with little-to-no temporal clustering (Lees, 1998).

A more realistic situation encompasses both time and space clustering. The previous individual analysis of the effects of time and space agglomeration allow for the study of their combination and the results are also illustrated in Figs. 2.1 and 2.2. The vertical shifts observed due to time clustering are enhanced for Σn_j^2 as t increases due to the spatial clustering effects. The combination of the super-diffusive nature of $n_j(t)$ due to spatial clustering and the change in the scanned phase space subset generated by time clustering results in a system that is locally effectively ergodic. The term “locally“ means that the system is effectively ergodic within the period during which temporal clustering does not occur: whenever $t \neq t_1$ and $t \neq t_2$. Once it happens, the system is reset and a different normal diffusion process takes place. This behaviour is illustrated in Fig. 2.2: from the moment when each of the two time clustering occurs, the inverse TM metric displays effective ergodic periods with different diffusion coefficients.

2.3.2 Southern California

Implementing the method to synthetic catalogs offers a good foundation to understand the effects of different seismicity patterns to Eqs. (2.4) and (2.5). However, real seismicity offers a behavior which is intrinsically more complex and a deeper understanding of the technique is required. The next step is to apply the method to a well studied dataset: the southern Californian catalog. The spatial clustering in

this data allows for the application of the approximation in Eq. (2.5) as discussed previously.

Tiampo et al. (2007) showed that events recorded from 1932 to 2006 in southern California with magnitude $M4$ or greater display long effective ergodic periods that are disrupted by large events. These periods of effective ergodicity were disrupted by some large earthquakes, but not all. The 1952 Kern County, the 1979 Imperial Valley and the 1992 Landers earthquakes but not the 1989 Loma Prieta disrupted the effective ergodic period.

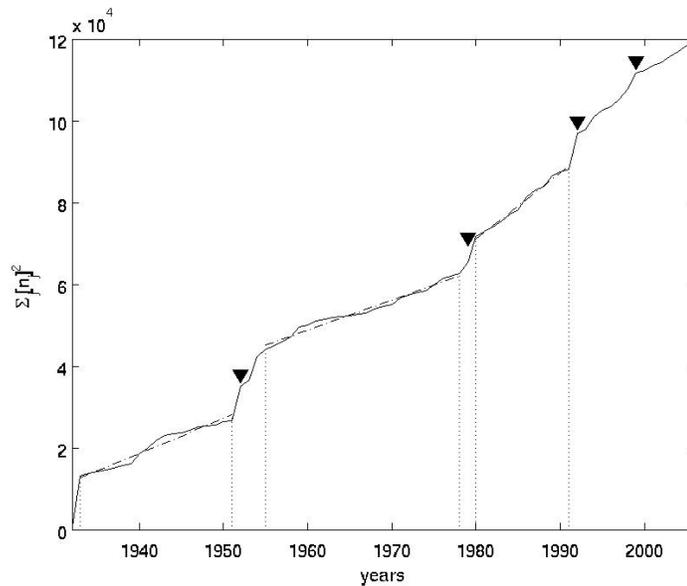


Figure 2.3: Plot of Eq. (2.4) for the southern California dataset from 1932 to 2006. Dash-dot lines are linear regressions for the three effective ergodic periods from 1933 to 1951, from 1955 to 1978, and from 1980 to 1991. Upside triangles locate the years of the occurrences of large earthquakes: the 1952 Kern County, the 1979 Imperial Valley, the 1992 Landers and the 1999 Hector Mine.

The same set of data is used with this method to better understand the previous

results in terms of seismicity clustering. A mesh of $0.1^\circ \times 0.1^\circ$ is considered for the region between latitudes 32°S and 40°S and longitudes -115° and -125° . Figure 2.3 shows the plot of Eq. (2.4) for the data considered. The dashed lines identify the large events that disrupt the effective ergodic periods: the 1952 Kern County, the 1979 Imperial Valley, the 1992 Landers, and the 1999 Hector Mine earthquakes. A vertical jump is observed for the years in which these events occurred. The exception is the 1979 Imperial Valley event, due to its late occurrence in the year (October), so that the bulk of the disruption occurs in the following year. It is important to stress that no premonitory pattern can be obtained from Fig. 2.3 due to the scale of the temporal discretization of the system: the effects of foreshocks/quiescence/aftershocks are all combined into one when considering a yearly time-discretization for the system.

Interesting features can be observed in Fig. 2.3. The first is the constant decrease in the amplitude of the vertical jumps as the years pass. This is attributed to the saturation of the cumulative number of events per box over time, which makes the effects of clustering less pronounced over the years due to the accrual of larger numbers events over the entire region and the asymptotic behavior of $\Omega_n \rightarrow \infty$ as $t \rightarrow \infty$. This illustrates importance of the choice of t_0 : the later t_0 is, the larger the response for the 1992 Landers and 1999 Hector mine earthquakes. This result also was noted by Tiampo et al. (2007). Here, this saturation results in the method not being able to detect the clustering due to the 1989 Loma Prieta earthquake.

The second important feature observed in Fig. 2.3 is the change in the slope of the linear regressions obtained with a Pearson's correlation coefficient greater than 0.97 between large events: from 1933 to 1951, from 1955 to 1978, and from 1980 to 1991.

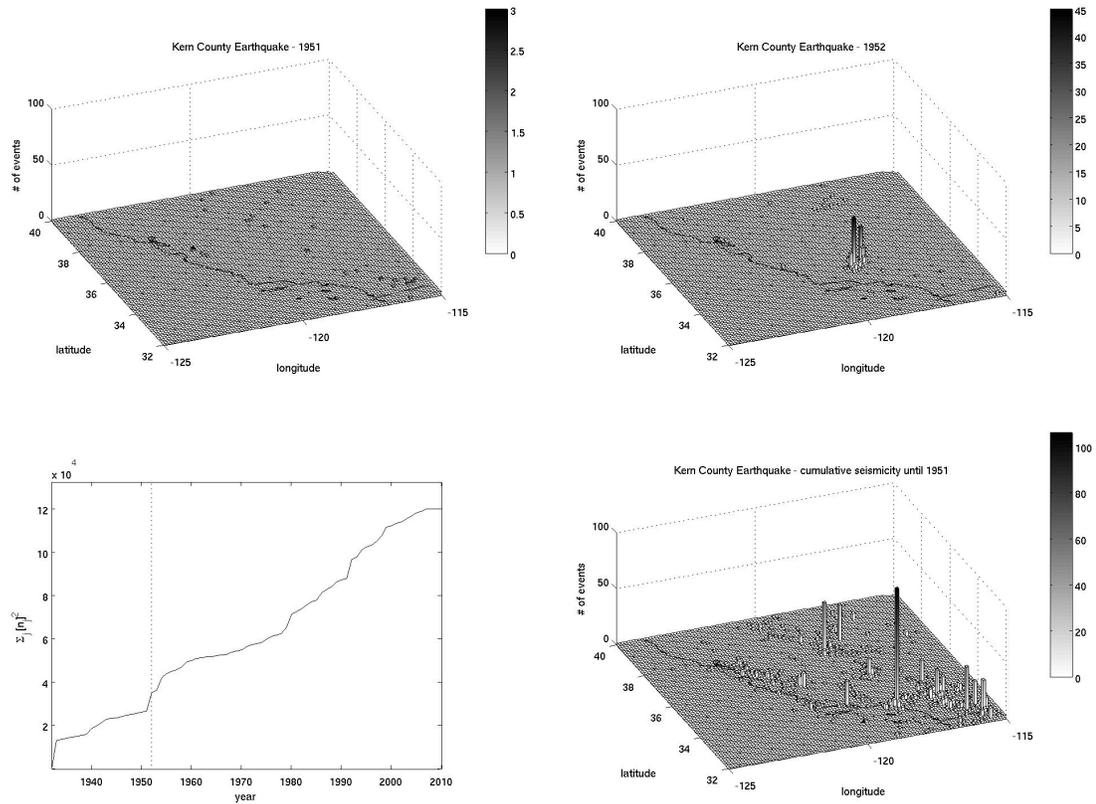


Figure 2.4: Distribution of events before the 1952 Kern County earthquake before the mainshock (top left), after the mainshock (top right) and up until the mainshock since 1932 (bottom right). The approximation in Eq. (2.4) with the time of occurrence of the 1952 Kern County event as the dashed line are also plotted (bottom left).

As mentioned previously, these periods are effective ergodic and they correspond to intervals of time in which the system displays a normal diffusive behaviour. The different slopes indicate that these diffusive processes sample the different subsets of the phase space at different rates. Additional work remains to determine whether this apparent rate change is due to a sampling effect from changes (primarily increases) in the seismic network, local and/or regional effects related to changes in the stress field from the large events themselves combined with tectonic and geologic heterogeneities,

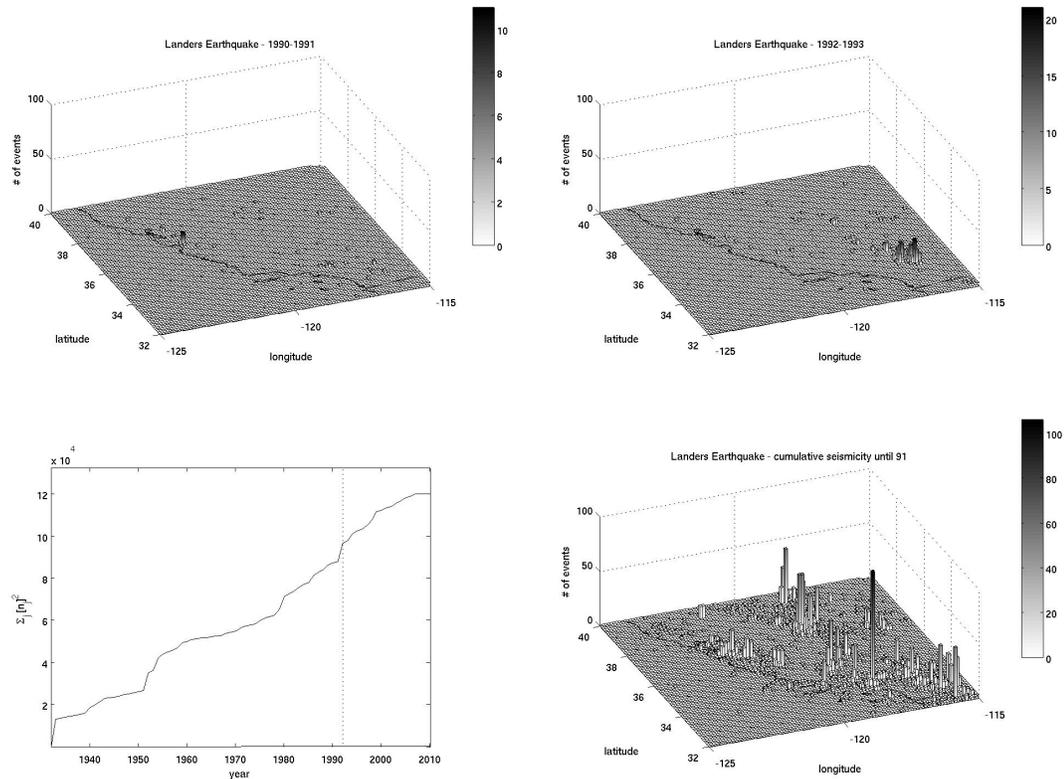


Figure 2.5: Distribution of events before the 1992 Landers earthquake before the mainshock (top left), after the mainshock (top right) and up until the mainshock since 1932 (bottom right). The approximation in Eq. (2.4) with the time of occurrence of the 1992 Landers event as the dashed line are also plotted (bottom left).

or some combination thereof.

Figures 2.4 to 2.6 display the distribution of the cumulative number of events prior to, along with the number of events before and after the 1952 Kern County, the 1999 Landers and the 1989 Loma Prieta Earthquakes. The cumulative activity prior to the 1952 Kern County event is not considerably larger than the activity following the mainshock. As a result, the computation of Eq. (2.4) for $t = 1952$ leads to the considerable vertical jump observed in Fig. 2.3. The effects of the increase in the

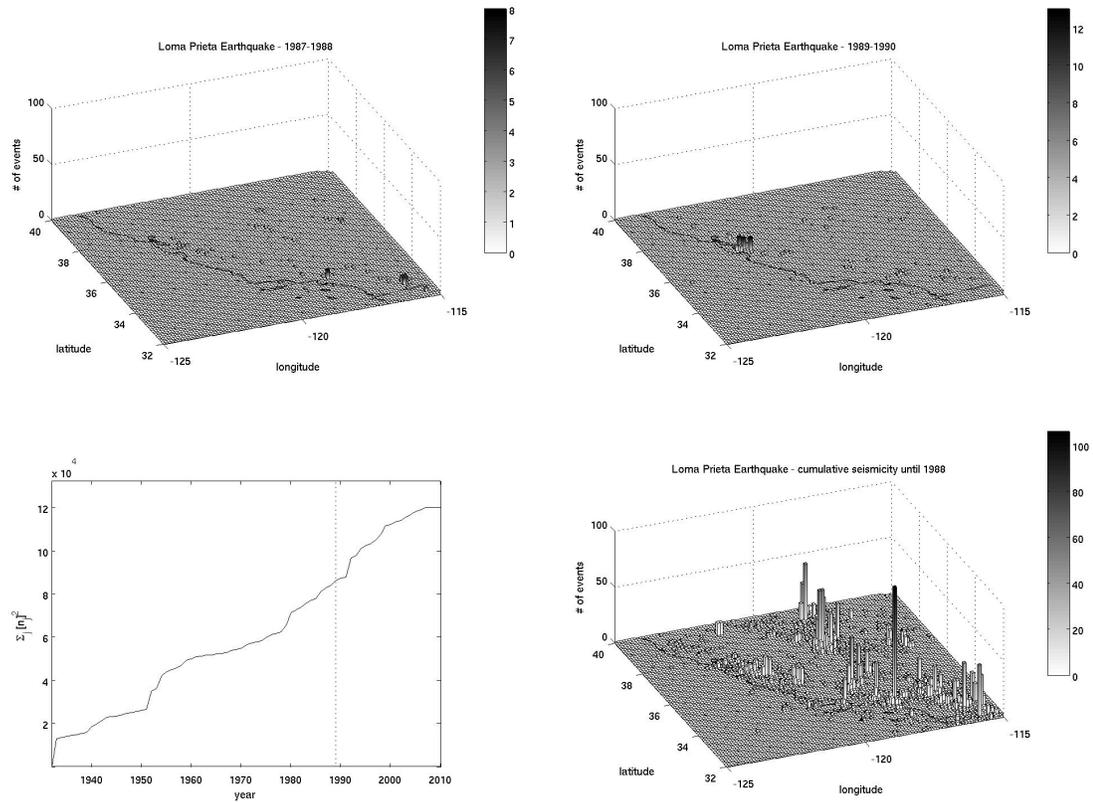


Figure 2.6: Distribution of events around the 1989 Loma Prieta earthquake: before the mainshock (top left), after the mainshock (top right) and up until the mainshock since 1932 (bottom right). The approximation in Eq. (2.4) with the time of occurrence of the 1989 Loma Prieta event as the dashed line are also plotted (bottom left).

cumulative number of events per box can be seen for the 1999 Landers earthquake illustrated in Fig. 2.5: the seismic activity after this event occurred in a region of considerable historic activity. The 1989 Loma Prieta earthquake, as displayed in Fig. 2.6, occurred in a region where the cumulative number of events prior to 1989 was considerably lower than the rest of the map. As a result, the clustering of seismic activity after this mainshock was not enough to promote a large disruption in the evaluation of Eq. (2.4). The latter can be attributed to the differences in seismic

activity between the northern and southern California.

2.3.3 Mining Seismicity

Mining induced seismicity (MIS) represents an interesting source of information due to the range of magnitudes that are involved: between lab controlled experiments and crustal seismicity. It can then offer important information on the scaling laws of seismicity and the nature of earthquake triggering. Economical factors also play an important role and, as a result, this topic has been extensively studied (Gibowicz and Kijko, 1994; Richardson and Jordan, 2002).

As a result, this method was also applied to MIS from two mines in Ontario, Canada. The dataset from Kidd Creek D Mine was obtained from August 2004 to May 2007 and it consists of 23000 event recordings. For Macassa Mine, over 10000 events were recorded from December 2004 to May 2007. A 3D version of the method was used and different space/time configurations were tested. The outcomes displayed similar behaviours and, in the present work, only the results from cubes with an edge length of 10m and a time binning of 7 consecutive days are shown for both mines.

Blasting activity is the main mechanism that drives the occurrence of small earthquakes in mines. As a result, MIS display a bimodal nature: small events highly clustered in time or space that are created by the blasts and larger, tectonic-like seismicity. Different bimodal distributions are constantly used to describe MIS (Gibowicz and Kijko, 1994). Richardson and Jordan (2002) used a set of simple criteria based on the space/time distance between events to identify the highly clustered blasting related events.

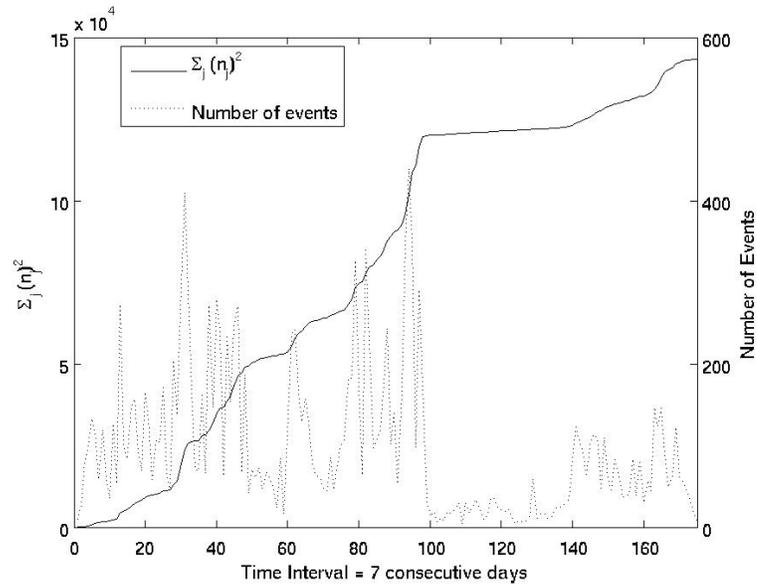


Figure 2.7: Plot of Eq. (2.4) for Kidd Creek D (solid line) and the number of events recorded per time bin t (dashed lines).

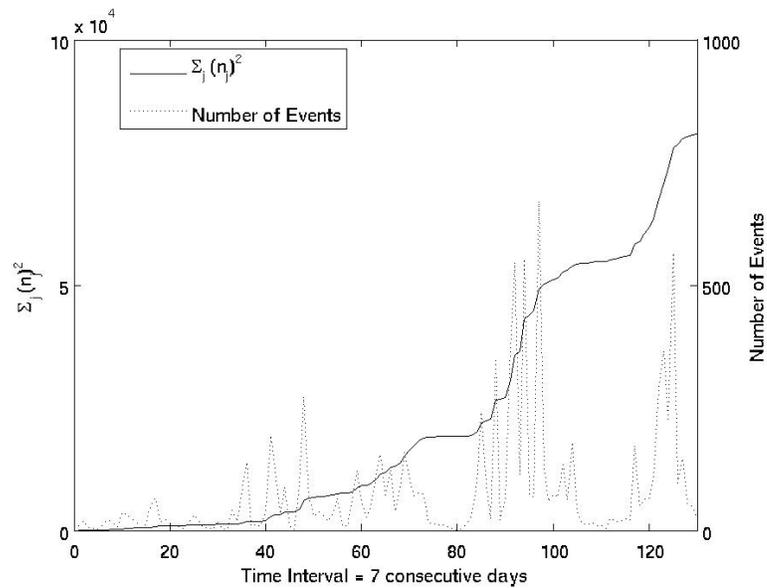


Figure 2.8: Plot of Eq. (2.4) for Macassa mine (solid line) and the number of events recorded per time bin t (dashed lines)

From the results obtained so far with this method, significant variability of Eq. (2.4) is expected. Its plot along with the seismic activity for Kidd Creek D is illustrated in Fig. 2.7. Frequent low-magnitude blasting activity generates most of the clustering observed in the various discontinuities in the plot. Note the horizontal feature in Fig. 2.7 around the period between the 100th and 140th weeks. The drop in the seismicity rate observed in Fig. 2.7 and the change of blasting sites are the reasons for the constant values of Eq. (2.4). The first automatically decreases the values obtained for Eq. (2.4) whereas the second has a more substantial role in the observed feature. As mentioned previously for the 1989 Loma Prieta earthquake, the method is unable to detect clustering in areas with a considerably lower seismic activity compared to the rest of the region. During the instances in which Σn_j^2 is constant, the blasting activity was shifted to regions with no or low previous seismic activity.

For Macassa Mine, the result of the method is plotted in Fig. 2.8. As indicated for the previous mine, two plateaus in the values of Eq. (2.4) are observed and they coincide with periods of low seismic activity. This mine displayed a completely different blasting regime compared to Kidd Creek D: rare high energy blasts. During the first period, around the 80th week, there was a shift in the blasting sites followed by a stoppage in the blasting activities. The second period around the 110th week is highlighted by a complete halt in the blasting process resulting in the sudden drop in seismic activity, leading to a second plateau. Once again, the historic seismicity played an important role as observed for the previous mine and southern California.

2.4 Conclusions

The TM metric is a simple metric that was first applied to study glass and liquid systems. Tiampo et al. (2007) showed that this metric can be used to identify periods of effective ergodicity, in which the system is considered to be in a metastable equilibrium state as a gas in thermal equilibrium, and that seismicity clustering seems to play an important role in this framework. These effective ergodic periods are interpreted as periods of time in which the evolution of cumulative number of events per box behave as a normal diffusive process. It is shown that the metric and the phenomena observed in the previous work can be attributed to seismic clustering and that, under the right assumptions, this metric can be simplified. Spatial clustering in seismicity allows for an approximation to the metric in which the effects of clustering in both time and space are simpler to account for.

While the determination of effective ergodic periods which ensure that spatial and temporal averages are stationary and confident is important for seismic hazard analysis, here we investigated the insight gained into seismicity clustering from the behaviour of the metric under various spatial and temporal end members clustering models. This interpretation was tested for three sets of data: synthetic, the southern California dataset and mining induced seismicity. From the synthetic data, it was observed that the effects of spatial and temporal clustering are of a different nature. The first can be seen as a change in the rate that the subset of the phase space is being browsed whereas the second is a change in the subset itself.

The implementation of this method to the southern California dataset showed

that the disruptions in the effective ergodic periods observed in Tiampo et al. (2007) were due to the aftershock sequences following large earthquakes. It was noticed that these disruptions were highly dependent on the distribution of past seismicity across the region. Large earthquakes such as the 1989 Loma Prieta did not disturb the metric due to the generally lower seismic activity in the region.

Mining seismicity from Kidd Creek D and Macassa mines in Canada were also tested under this interpretation of the TM metric. These mines displayed very different blasting patterns in the periods considered: the first with frequent low-magnitude blasts and the second with rarer larger ones. Regardless of these differences, the same dependence on past seismicity in the analysis was verified in the datasets from both mines. Changes in locations and rates of blasting activity induced changes in seismicity rates in locations with previously low activity, resulting in the constant metric during these periods. It is the same as that which appeared to occur for the 1989 Loma Prieta event and the low seismicity rate in the surrounding local region.

Past studies concerning the application of the TM metric to seismicity have shown that it can be used to highlight well-behaved statistical features of seismicity (Tiampo et al., 2007, 2010). Here another interesting feature of the TM metric is examined and it is shown that the metric provides a simple way to quantify seismicity clustering, and can differentiate between spatial and temporal clustering, given that the region space is chosen carefully. We also demonstrate that its simplicity comes at a cost: the non-uniqueness of the values of the metric, in which different distributions of the cumulative number of events in each box might yield the same score means that various spatial patterns can produce the same value over different time periods. Finally,

ongoing studies using simple models of earthquake fault systems suggest that there is a link between the TM results for seismicity that are both unique and more complicated than originally anticipated. Future work will attempt to link the clustering results for both historic data and models of earthquake processes.

2.5 Appendix

Consider a positive integer A written in terms of n numbers so that

$$A = \sum_i^n a_i \quad (2.6)$$

where a_i are non-negative real numbers.

From (2.6),

$$\begin{aligned} A^2 &= \left(\sum_i^n a_i \right)^2 = \sum_i^n a_i^2 + 2 \overbrace{\sum_{i \neq j}^n a_i \left(\sum_j^n a_j \right)}^{\geq 0} \\ \sum_i^n a_i^2 &= A^2 - 2 \sum_{i \neq j}^n a_i \left(\sum_j^n a_j \right) \end{aligned} \quad (2.7)$$

The extreme values of Eq. (2.7) depend on

$$f(a_j) = \sum_i^n \left(a_i \sum_{j \neq i}^n a_j \right) \quad (2.8)$$

The maximum value of Eq. (2.7) is obtained when Eq. (2.8) is null. Since a_j are non-negative integers and A is a positive integral, $f(a_j) = 0 \iff a_j = A\delta_{ij}$. This means that one value of $a_j = A$ when $j = i$ and the others are null, i.e. one box contains all events and the others have no events.

The maximum of (2.8) can be obtained by using the Lagrange multiplier method

using Eq. (2.6) as a constraint $g(a_i) = \sum_i^N a_i - A \equiv 0$

$$\begin{aligned} \nabla_{a_i} (f(a_i) - \lambda g(a_i)) &= 0 \\ \nabla_{a_i} \left[\sum_{j:j \neq i}^n a_j \left(\sum_i^n a_i \right) \right] - \lambda \nabla_{a_i} \left[\sum_i^N a_i - A \right] &= 0 \end{aligned} \quad (2.9)$$

where λ is a Lagrange multiplier.

The evaluation of Eq. (2.9) leads to a system of n -equations

$$\sum_{j \neq i} a_j = \lambda, \forall i \in [1, n] \quad (2.10)$$

which results in

$$a_j = \frac{A}{n}. \quad (2.11)$$

so that $\sum_j a_j = A$.

The result obtained in Eq. (2.11) means that Eq. (2.8) is maximized when the values of a_j are the same, i.e. for the case in which the events are evenly distributed in the non-empty boxes. The same procedure can be used to one of the a_i numbers. It is important to stress that different sets of a_i may yield the same $\sum_i a_i^2$.

Acknowledgments

The work presented here was supported by the WSIB Grant for Mining Re-Entry Protocols and the NSERC & Aon Benfield/ICLR Chair in Earthquake Hazard Assessment. The seismic catalog for southern California was obtained from ANSS (anss.org), while the mining data was obtained under the same WSIB grant.

References

- Baiesi, M. and Paczuski, M. (2004). Scale-free networks of earthquakes and aftershocks. *Physical Review E*, 69(066106).
- de Oliveira, C. R. and Werlang, T. (2007). Ergodic hypothesis in classical statistical mechanics. *Revista Brasileira de Ensino de Física*, 29(2):189–201.
- Dieterich, J. (1994). A constitutive law for rate of earthquake production and its application to earthquake clustering. *Journal of Geophysical Research*, 99(B2):2601–2618.
- Farquhar, I. E. (1964). *Ergodic theory in statistical mechanics, Monographs in Statistical Physics Volume 7*. Inerscience Publishers.
- Felzer, K. R., Becker, T. W., Abercrombie, R. E., Ekström, G., and Rice, J. J. (1994). Triggering of the 1999 m_w 7.1 Hector mine earthquake by aftershocks of the 1992 m_w 7.3 Landers earthquake. *Journal of Geophysical Research*, 107(B9):2190.
- Ferguson, C. D., Klein, W., and Rundle, J. B. (1999). Spinodals, scaling, and ergodicity in a threshold model with long-range stress transfer. *Physical Review E*, 60:1359–1373.
- Gibowicz, S. J. and Kijko, A. (1994). *An introduction to mining seismology*. Academic New York.
- Helmstetter, A. and Sornette, D. (2002). Subcritical and supercritical regimes

- in epidemic models of earthquake aftershocks. *Journal of Geophysical Research*, 107(B10):1–21.
- Holliday, J. R., Rundle, J. B., Tiampo, K. F., and Turcotte, D. L. (2006). Using earthquake intensities to forecast earthquake occurrence times. *Nonlinear Processes in Geophysics*, 13:585–593.
- Jimenez, A., Tiampo, K. F., Levin, S., and Posadas, A. M. (2006). Testing the persistence in earthquake catalogs: the iberian peninsula. *Europhysics Letters*, 73(2):171–177.
- Kagan, Y. Y. and Jackson, D. D. (1991). Long-term earthquake clustering. *Geophysical Journal International*, 104:117–133.
- Kanamori, H. (1981). *The nature of seismicity patterns before large earthquakes*, in *Earthquake Prediction: an International Review*. American Geophysical Union.
- Lees, J. M. (1998). Multiplet analysis at coso geothermal. *Bulletin of the Seismological Society of America*, 88(5):1127–1143.
- Marsan, D. and Lengliné, O. (2008). Extending earthquakes’ reach through cascading. *Science*, 319:1076–1079.
- Mendonza, C. and Hartzell, S. H. (1988). Aftershock patterns and mainshock faulting. *Bulletin of the Seismological Society of America*, 78(4):1438–1449.
- Mogi, K. (1985). *Earthquake Prediction*. Academic Press, NY.

- Mountain, R. D. and Thirumalai, D. (1989). Measures of effective ergodic convergence in liquids. *Journal of Physical Chemistry*, 93:6975–6979.
- Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point process. *Journal of the American Statistical Association*, 83(401):9–27.
- Reasenber, P. (1985). Second-order moment of central california seismicity, 1969-1982. *Journal of Geophysical Research*, 90(B7):5479–5495.
- Richardson, E. and Jordan, T. H. (2002). Seismicity in deep gold mines of south africa: implications for tectonic earthquakes. *Bulletin of the Seismological Society of America*, 92(5):1766–1782.
- Shcherbakov, R., Yakovlev, G., Turcotte, D. L., and Rundle, J. B. (2005). Model for the distribution of aftershock interoccurrence times. *Physical Review Letters*, 95(218501).
- Thirumalai, D. and Mountain, R. D. (1993). Activated dynamics, loss of ergodicity, and transport in supercooled liquids. *Physical Review E*, 47:479–489.
- Thirumalai, D., Mountain, R. D., and Kirkpatrick, T. R. (1989). Ergodic behavior in supercooled liquids and in glasses. *Physical Review A*, 39:3563–3574.
- Tiampo, K. F., Klein, W., Li, H. C., Mignan, A., Toya, Y., Kohen-Kadoshand, S. Z. L., Rundle, J. B., and Chen, C. C. (2010). Ergodicity and earthquake catalogs: Forecast testing and resulting implications. *Pure and Applied Geophysics*, 167(6–7):763–782.

- Tiampo, K. F., Rundle, J. B., Klein, W., Holliday, J. B., Martins, J. S. S., and Ferguson, C. D. (2007). Ergodicity in natural earthquake fault networks. *Physical Review E*, 75(066107):1–15.
- Tiampo, K. F., Rundle, J. B., Klein, W., Martins, J. S. S., and Ferguson, C. D. (2003). Ergodic dynamics in a natural threshold system. *Physical Review Letters*, 91(23):1–4.
- Tiampo, K. F., Rundle, J. B., McGinnis, S., and Klein, W. (2002). Pattern dynamics and forecast methods in seismically active regions. *Pure and Applied Geophysics*, 159(10):2429–2467.
- Utsu, T. (2002). *Statistical features of seismicity. International Handbook of Earthquake and Engineering Seismology, Part A*. Academic Press, Amsterdam.
- Zaliapin, I., Gabrielov, A., Keilis-Borok, V., and Wong, H. (2008). Clustering analysis of seismicity and aftershock identification. *Physical Review Letters*, 101(018501).

Chapter 3

Pattern Informatics in Mining Induced Seismicity and its Applications to Rockburst Hazard Assessment²

3.1 Introduction

Human activities that change the stress field of a region, such as the filling of reservoirs and mining activities, often lead to variations in the local or regional levels of seismicity (Simpson, 1986). Mining seismicity provides an interesting source for a better understanding of the nature of earthquakes. The magnitudes of these events fill the gap between controlled experiments and regional seismicity, turning these environments into excellent laboratories to understand different properties of seismicity. Richardson and Jordan (2002) analyzed source parameters from mining seismicity in deep gold mines in South Africa to better understand rock failure processes, estimating the critical slip distance of “friction-dominated” events in the order of 10^{-4} m.

One of the most important issues concerning mining seismicity is to ensure safe work conditions in mines while maintaining high productivity through the mitigation of rockburst hazard. Rockbursts are induced seismic events that result in damage to excavations and become more severe as the depth of mining increases. Vallejos and McKinnon (2010) have established that aftershock sequences in mining seismicity fol-

²resubmitted to Geophysical Journal International

low the modified Omori's law (Omori, 1894; Utsu, 2002) to improve the guidelines used to establish re-entry protocols in different mines in Canada. Other examples of the assessment of seismic hazard in mines are: the use of a model in which seismic events are Poissonian in time was proposed for coal mines in Poland (Lasocki, 1990); methods based on the relation between mining activity and induced seismicity (Glowacka et al., 1990); the development of a seismic hazard scale in Australian mines (Hudyma and Potvin, 2004) and the quantification of seismic hazard by using the cumulative Benioff strain during a time interval in Czech mines (Holub, 2007).

Most of the methods mentioned above take into account a particular region of the mine for the hazard assessment and/or they rely on analysis of a power law distribution of the induced seismicity. The use of power laws for mining events may be a troublesome assumption due to its seemingly bimodal nature (Gibowicz and Kijko, 1994; Holub, 1996; Richardson and Jordan, 2002), in which mining seismicity can be divided into events directly related to the mining activity and events related to local geological features. Gibowicz and Kijko (1994) classify events of the first type as low-to-medium magnitudes with their locations close to mining sites and local geological features, whereas the second type is comprised of events with larger magnitudes that are located farther from mining activities. A similar classification has been done for mining seismicity in South African gold mines by Richardson and Jordan (2002), where events are divided into Type A and B. The first is considered to be fracture-dominated and related to mining activity. The latter is hypothesized to be friction-related events that result from the long-term extraction of ore rather than mining activity.

A series of methods to forecast regional seismicity in the intermediate-to-long term are available and they rely on different precursory signals (Cicerone et al., 2009; Keilis-Borok, 2005), and seismicity pattern analysis is one of these potential indicators (Kanamori, 1981). The application of the latter to mining seismicity might then offer valuable information regarding the nucleation of large events in mines. Eneva (1998) applied pattern recognition methods to analyze the relationship between large and micro earthquakes as well as their temporal distribution in mines in Ontario, Canada, by using three parameters previously employed in the analysis of a different mine (Eneva and Young, 1993) and synthetic regional seismic data (Eneva and Ben-Zion, 1997).

Two techniques that are used to assess seismic hazard on a regional scale are the Pattern Informatics (Rundle et al., 2002; Tiampo et al., 2002) and Relative Intensity (Holliday et al., 2006, 2005). Both methods use past seismicity to quantify seismic hazard. The first quantifies temporal variations in seismicity patterns using phase dynamics, whereas the second measures historic seismicity rates. In this study, the application of the PI method of quantification of seismicity patterns is proposed to mining seismicity and its results are compared to those from the RI method. These methods are applied to seismicity from two mines in Canada: Kidd Creek D and Macassa.

Determining the parameters to be used in both methods, such as the temporal and spatial discretization for a mining induced seismicity scale, is a critical step. These parameters are estimated from the Thirumalai-Mountain (TM) metric (Thirumalai et al., 1989; Thirumalai and Mountain, 1993; Tiampo et al., 2003). With the pa-

rameters set, the retrospective forecasts for mining seismicity are compared using a contingency table originally used for tornado and weather forecasting and a simple skill score called the Heidke Skill Score (Heidke, 1926).

The structure of this paper is as follows: section 3.2 describes the data used in this study as well as both the PI and RI along with the Heidke skill score and the TM metric. The results of these methods applied to the seismicity of the mines studied are presented and discussed in section 3.3. Finally, the conclusions are presented in section 3.4.

3.2 Data and Methods

3.2.1 Mining Data

The dataset for Kidd Creek D mine consists of over 23000 recordings obtained from August 2004 to May 2007 within a volume of size 500m x 600m x 550m. These recordings were classified as blasts, noise, seismic events and reported events, where the latter includes events felt by mine workers, large mine tremors and rockbursts. The monitoring system in this mine consisted of 15 uniaxial and 4 triaxial accelerometers.

For Macassa mine, over 10000 events were recorded during the period from December 2004 to May 2007 within a volume of size 460m x 340m x 280m. The recordings for this mine were classified as events and blasts only during the time period considered. A dense microseismic monitoring array of 66 uniaxial accelerometers was used in this mine during the period considered.

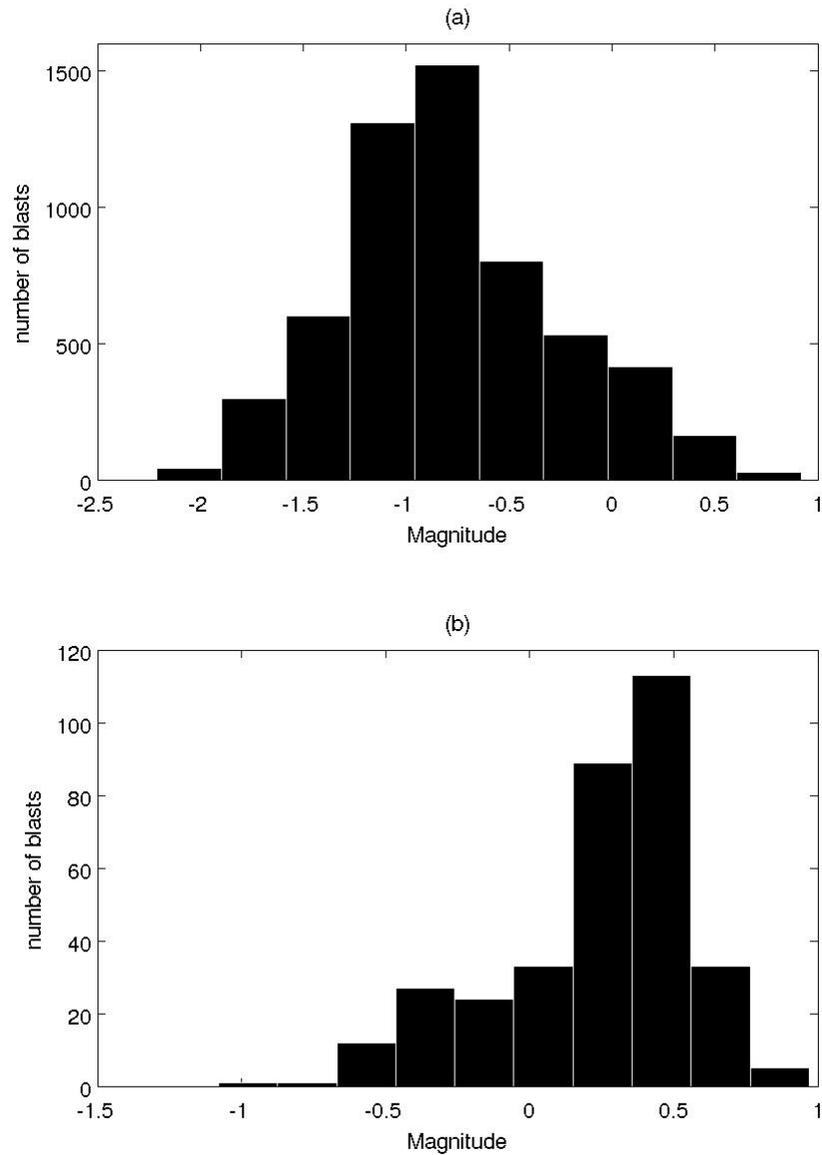


Figure 3.1: Histogram of blasts for Kidd Creek D (a) and Macassa (b) mines that highlight the difference in blasting activity between the mines. The first mine has frequent low-magnitude blasts whereas the second has rarer high-magnitude blasts.

The blasts during the time period considered for both mines were identified manually or automatically. In the first, the microseismic technologist at the mine knows

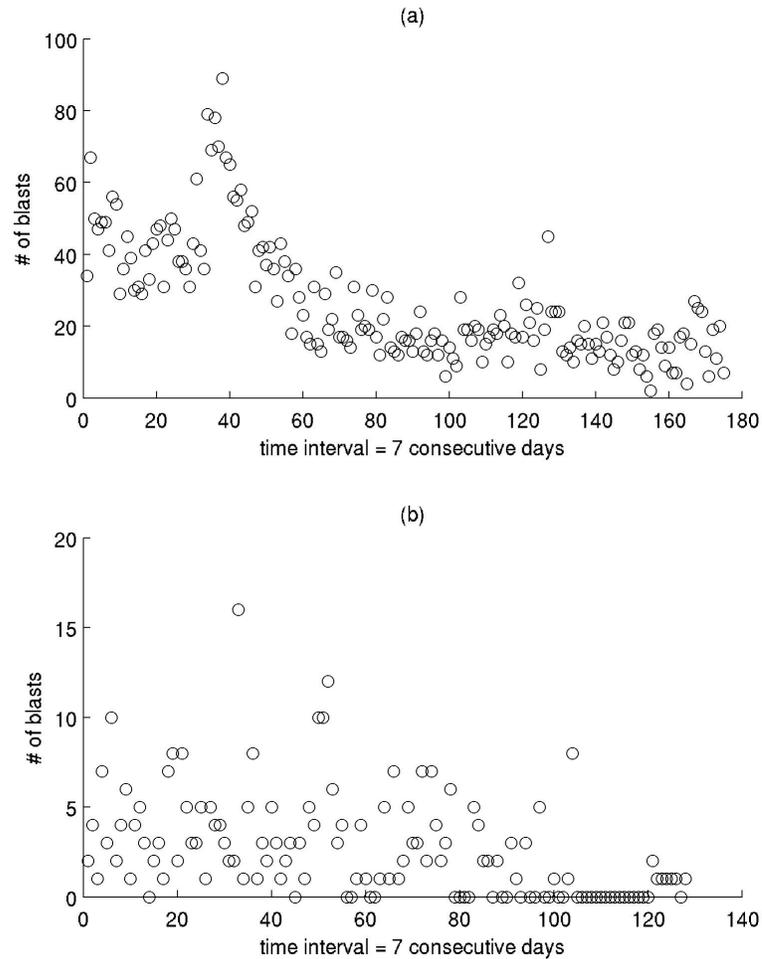


Figure 3.2: Time series of blasts for Kidd Creek D (a) and Macassa (b) mines per week. At Kidd Creek D mine, there is always blasting activity in the period considered whereas Macassa mine has periods in which blasting is halted.

approximately when the blasts were scheduled but exact times were not recorded in blast notices or daily blast logs. The blasts are then manually matched to the recorded seismicity by the examination of the waveforms. The automatic identification is performed by tagging events as blasts if a specific number of events occur within a certain time interval. Figs. 3.1 and 3.2 illustrate the distribution of magnitudes and the rate

of occurrence of blasts for both mines, respectively.

3.2.2 The Pattern Informatics and Relative Intensity

The PI and RI methods have been used for seismic hazard estimation on a regional scale in several different parts of the world such as central Japan (Nanjo et al., 2005), southern California (Holliday et al., 2005; Rundle et al., 2002; Tiampo et al., 2002) and Taiwan (Chen et al., 2005) for intermediate-to-long term hazard assessment. These methods rely on the premise that seismicity patterns can be used as a precursor to larger events to assess the seismic hazard in a given region (Kanamori, 1981). The first quantifies changes in seismicity rate while the second measures the long term seismicity in the target region. Tiampo et al. (2006) showed that the PI can be used as a proxy to measure stress in a given region of interest.

The starting point for both techniques is to have a complete catalog of seismic events from an initial time t_0 to a final time t_2 . A mesh is then created to divide the region of interest into N boxes. Time series $n_i(t)$ for the cumulative number of events in box i at a given time unit interval, or time bin t , are then generated for each box.

The RI hazard map is created by computing the number of earthquakes $n_i(t_0, t_2)$ in each box i from t_0 to t_2

$$n_i(t_0, t_2) = \sum_{t=t_0}^{t_2} n_i(t). \quad (3.1)$$

Previous work (Rundle et al., 2002; Holliday et al., 2005) regarded (3.1) as a non-normalized probability for the location of future events with a magnitude larger than a certain value M_f for times $t > t_2$. This probability can be normalized so that the

probability is unity over the entire region

$$P_i = \frac{n_i(t_0, t_2)}{\sum_{i=1}^N n_i(t_0, t_2)}. \quad (3.2)$$

The RI poses as a reasonable method for evaluation of seismic hazard based on long-term background rates of seismicity, regardless of the simplicity of (3.2). Past seismicity in each element of the mesh is the main signal used in this method: the more events in the past in a certain location, the more likely events are to occur there. However, past recordings have the same importance regardless of when they occurred. The result is a lack of sensitivity to foreshocks and quiescence periods prior to large magnitude events by the RI. These precursory signals are better identified by the PI method.

The computation of a PI hazard map requires a series of steps. Initially, a reference time $t_1 \in [t_0, t_2]$ is chosen and the time interval $\Delta t = t_2 - t_1$ is set as a standard. Seismicity patterns in a period from a time base t_b to t_1 are then compared to Δt . The seismic intensity in a box i from t_b to t is the average number of events per unit time and is written as

$$I_i(t_b, t) = \frac{n_i(t_b, t)}{t - t_b + 1}. \quad (3.3)$$

Seismic intensities are normalized for comparison purposes between different time intervals. This is done by subtracting the average of I_i , μ , over all boxes and dividing the result by the variance, σ

$$\hat{I}_i(t_b, t) = \frac{I_i(t_b, t) - \mu}{\sigma}. \quad (3.4)$$

Equation (3.4) is evaluated for t_b and the measurement of anomalous seismicity

in box i , $\Delta I_i(t_b, t_1, t_2)$, during $\Delta t = t_2 - t_1$ considering the period from t_b to t_2 is

$$\Delta I_i(t_b, t_1, t_2) = \hat{I}_i(t_b, t_2) - \hat{I}_i(t_b, t_1). \quad (3.5)$$

Summing (3.5) over all possible values of t_b reduces random variability of the measurement of anomalous seismicity. The square of this sum is proportional to the probability P_i of a future event with magnitude greater than M_f in box i during the period from t_2 to $t_2 + \Delta t$:

$$P_i \propto \left(\sum_{t_b} \Delta I_i(t_b, t_1, t_2) \right)^2. \quad (3.6)$$

Equation (3.6) offers a comparison between the seismic activity from t_1 until t_2 and all different time intervals between t_b and t_1 . It is important to emphasize that the anomalous seismicity quantified by equation (3.6) is related to activation and quiescence periods, which can be related to seismicity clustering. The first stands for an increase in seismic activity (clustering) whereas the second characterizes a decrease in seismic activity in the time interval considered.

Equations (3.2) and (3.6) provide a set of scores for each box i that can be normalized and, based on a decision threshold D , a binary forecast can be issued for the period from t_2 to $t_2 + \Delta t$ for both the PI and the RI. Locations with a score greater than D are considered to be hotspots in which it is hypothesized that events with magnitude greater than M_f are most likely to occur.

3.2.3 The Heidke Skill Score

The results from binary-type forecasts can be represented in contingency tables (Table 3.1), where a , b , c and d represent correct forecasts, false alarms, misses and correct

negatives respectively. The efficiency of forecasting methods can be measured by means of tests called Skill Scores (SS). These tests quantify the improvement of a given forecast over a standard reference one, and thus can evaluate their performance. They rely on the choice of the reference forecast and are intrinsically incomplete (Woodcock, 1976; Wilks, 1995).

The broadness of these tests can be seen in the literature. Shcherbakov et al. (2010) used the Pierce Skill Score (PSS) to quantify the efficiency of their method to obtain hotspots for large earthquakes worldwide. Comparisons showing that the PI outperforms the RI have been made using Relative Operating Characteristic diagrams (ROC) (Chen et al., 2005; Holliday et al., 2005). The comparison between the PI and RI using Molchan diagrams (Zechar and Jordan, 2008) establish that there is not a significant gain when using the PI for southern California.

The quantification of a SS can be made in terms of a ratio that uses a particular measure of accuracy of a forecast method, A , and the reference forecast, A_{ref} . It is given by

$$SS = \frac{A - A_{\text{ref}}}{A_{\text{perf}} - A_{\text{ref}}} \quad (3.7)$$

where A_{perf} is the value of accuracy for a perfect forecast.

One of the most used tests in atmospheric sciences is the Heidke Skill Score (HSS) (Heidke, 1926), which uses the hit rates, $P_h = a/n$ and $P_m = d/n$, of a forecast and a random one. It offers a straightforward comparison between the forecast method of interest and a random one. Considering table 3.1 and the total number of events

		Observed	
		yes	no
Forecasts	yes	a	b
	no	c	d

Table 3.1: Example of a contingency table: a , b , c and d represent correct forecasts, false alarms, misses and correct negatives respectively.

$n = a + b + c + d$, the marginal probability of obtaining a “yes” forecast and “yes” observation are $P_{yf} = (a + b)/n$ and $P_{yo} = (a + c)/n$. The probability of a correct forecast by chance is then $P_{yr} = P_{yf}P_{yo} = (a + b)(a + c)/n^2$. The same rationale can be made to conclude that the probability of a correct “no” forecast to be $P_{nr} = (b + d)(c + d)/n^2$. Thus $A = P_h + P_m = (a + d)/n$, $A_{\text{ref}} = P_{yr} + P_{nr}$, $A_{\text{perf}} = 1$, and applying these to equation (3.7), the HSS is given by

$$\begin{aligned}
 HSS &= \frac{(a + d)/n - P_{yr} - P_{nr}}{1 - P_{yr} - P_{nr}} \\
 &= \frac{2(ad - bc)}{(a + c)(c + d) + (a + b)(b + d)}.
 \end{aligned} \tag{3.8}$$

Note that a perfect and a random forecast yield $HSS = 1$ and $HSS = 0$, respectively.

A negative HSS means that the forecast performs worse than a random guess.

In this case, the HSS is chosen over the PSS due to the scarcity of large magnitude events in the mining catalog and the asymptotic behavior of the latter for a small number of forecasts. Using the notation from Table 3.1, the Pierce skill score (PSS) is written as

$$PSS = \frac{a}{a + c} - \frac{b}{b + d}. \tag{3.9}$$

In this case, the number of cubes in the 3D mesh is considerably larger than the number of large events. The number of correct negatives (d) becomes too large and equation (3.9) will tend towards $PSS = \frac{a}{a+c}$, the probability of detection of the method. To obtain a perfect PSS in this context, an alarm should be issued for all elements of the mesh. It is clear, therefore, that the PSS can be biased in this framework and the HSS is a more accurate measure of forecast skill here.

3.2.4 The Thirumalai-Mountain Metric

The ergodic hypothesis offers an interesting way for classical statistical mechanics to relate micro and macro states (Farquhar, 1964; de Oliveira and Werlang, 2007). The Thirumalai-Mountain (TM) metric was first developed to study the effective ergodicity in liquids and glasses (Thirumalai et al., 1989; Thirumalai and Mountain, 1993). In its original form, the TM metric is given as

$$\Omega_G(t) = \frac{1}{N} \sum_{j=1}^N [g_j(t) - \bar{g}]^2 \quad (3.10)$$

where g_j is the time average of an observable for box j until t and \bar{g} is the average of g_j over all boxes. Equation (3.10) can be interpreted as the spatial variance of the temporal mean of the observable $g_j(t)$.

Effective ergodicity refers to the relationship between both ensemble and particle time averages for long but finite time intervals. They occur in periods of time when the inverse TM metric is linear in time with a positive slope, $\frac{1}{\Omega_G(t)} = \frac{t}{D_e}$. In the latter relation, $D_e = \frac{1}{\Omega_G(t_0)}$ is a diffusion parameter (Tiampo et al., 2003, 2007) and t_0 is the initial time considered.

Tiampo et al. (2007) applied this metric to regional seismicity to identify periods of effective ergodicity by using the cumulative number of events in each box $n_j(t)$ as a proxy for the observable $g_j(t)$, the seismic energy released. Because stationarity is a necessary condition for ergodicity, during periods of effective ergodicity, the spatial average is constant and approaches the temporal average. The system is in a state of metastable equilibrium and because the spatial and temporal moments become constant in this case, and techniques that identify seismicity pattern changes such as the PI are more accurate and effective (Tiampo et al., 2010).

In this context, equation (3.10) can be re-written in terms of the variance of the cumulative number of events in each box $n_j(t)$. The TM metric can be written as

$$\begin{aligned}\Omega_G(t) &= \frac{1}{t^2} \frac{\sum_j^N (n_j(t))^2}{N} - \frac{1}{t^2} \left(\frac{\sum_j^N n_j(t)}{N} \right)^2 \\ &= \frac{1}{t^2} (\langle n_j(t)^2 \rangle - \langle n_j(t) \rangle^2).\end{aligned}\tag{3.11}$$

Equation (3.11) shows that the TM metric is proportional to the variance $n_j(t)$. Cho et al. (2010) considered the latter expression to interpret the time evolution of the TM metric as a normal diffusive process during effective ergodic periods. This led to the conclusion that the metric is sensitive to seismicity clustering in both space and time and that effective ergodic periods represent intervals of time in which there is no clustering.

3.3 The PI in mining seismicity

The application of the PI to mining seismicity should be performed with caution given the fact that the latter is not driven with a constant driving rate as regional

scale seismicity usually is. However, an examination of equations (3.5) and (3.6) indicate that the PI quantifies anomalous seismic activity in a given box i during a time interval Δt compared to the historical seismicity. Tiampo et al. (2006) also compared this measurement of anomalous seismicity to a proxy of the change of the underlying stress in different parts of California. In the current framework, the application of the PI in mining seismicity should then be considered strictly in terms of the quantification of seismicity anomalies initially and its role as precursory signal for larger seismic events in mines.

The implementation of the PI and RI methods to mining induced seismicity depends on a series of parameters that are known for regional seismicity. Moulik (2009) dealt with the estimation for the PI in a regional scale by searching for the set of parameters that yield the optimal value of the Pierce's Skill Score in retrospective forecasts. It is important to notice that the procedure followed in the latter study is not based on a deterministic approach to obtain the desired parameters.

Establishing these parameters for mining seismicity is achieved by a different means to ensure a physical explanation for particular choices, and the TM metric plays an important role here. The completeness of the catalog is analyzed based on the frequency-magnitude distribution of events in both mines. Spatial/temporal discretization and forecast magnitudes M_f are addressed with the TM metric. Finally, different time intervals Δt are tested using the Hedke Skill Score to evaluate retrospective forecasts for all active boxes in the system. In this framework, a correct forecasts is issued if a large event occurs in one of the locations identified as a hotspot or within its vicinity as performed in Holliday et al. (2005). The latter is the boxes

immediately around the hotspot.

3.3.1 The Magnitude of Completeness

A complete catalog is a requirement for a good performance of both the PI and RI methods (Rundle et al., 2002; Tiampo et al., 2002). The cumulative frequency-magnitude distribution of seismic events for both Kidd Creek D and Macassa mines are plotted in Fig. 3.3. The estimation of the b -values was made using the method described in Tinti and Mulargia (1985). Error estimations for the b -values reported here are to a 98% confidence level.

Two b -values were calculated in Fig. 3.3a for Kidd Creek D mine for different cutoff magnitudes. The lower cutoff at $-1.93\mathbf{M}$ yields a distribution with $b = 1.22 \pm 0.03$ that describes well the lower-magnitude range but not the high-magnitude end. At this end, another linear trend is dominant for a cutoff magnitude of $-0.73\mathbf{M}$ with $b = 1.91 \pm 0.22$. Similar trends have been reported for Pyhäsalmi ore mine in Finland (Oye et al., 2005) and gold mines in South Africa (Richardson and Jordan, 2002).

For Macassa mine, one cutoff magnitude at $-1.25\mathbf{M}$ was used and a distribution with a $b = 3.20 \pm 0.09$ was obtained and plotted in Fig. 3.3(b). The reason for the use of a single cutoff magnitude in this mine is the clipping of the seismic waveforms of large events that occurred far from the sensors in the mine, preventing an accurate determination of their magnitudes. The result is a superposition in the intermediate-to-high magnitude range of both accurately measured magnitudes and saturated ones. The difference in the b -values obtained indicate that the seismicity in both mines is different, as expected given the blasting regimes adopted (Figs. 3.1 and 3.2).

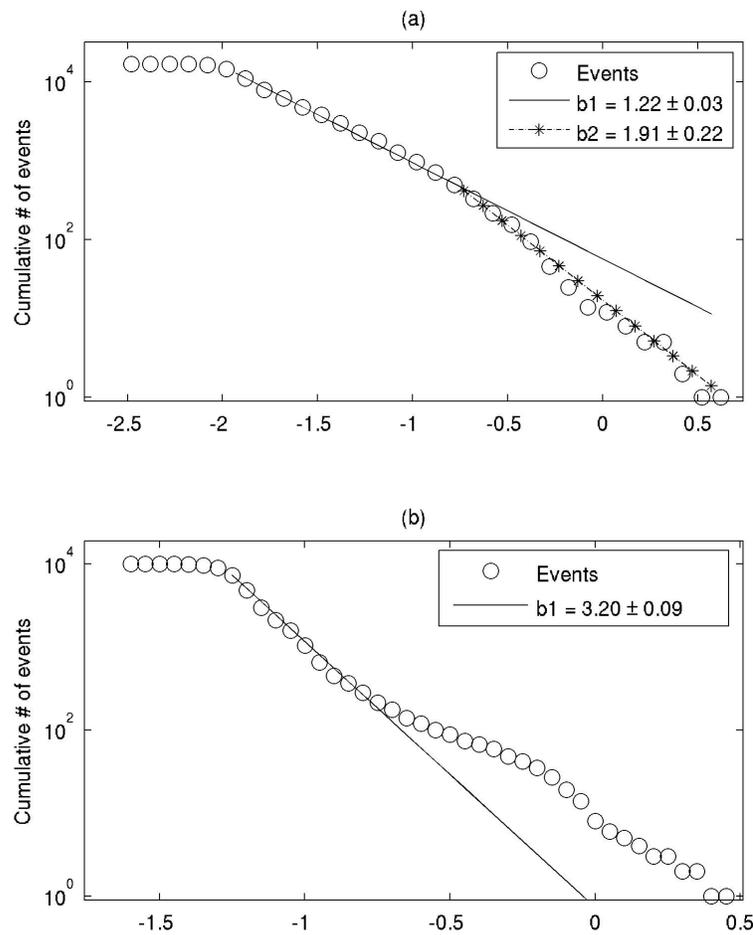


Figure 3.3: Cumulative frequency-magnitude distribution of events for Kidd Creek D (a) and Macassa (b) mines. Solid lines represent the power law.

The low b values calculated for Kidd Creek D mine suggests that the small but frequent blasts accelerate the rupture mechanism in the region, regardless of the wider range of blast magnitudes in Fig. 3.1(a). On the other hand, the larger but infrequent blasts at Macassa mine seems to have triggered a large number of events that disrupt the expected power-law behavior. Even though the recordings of these two mines obey the Modified Omori's Law (Vallejos and McKinnon, 2010), it still

does not guarantee that their magnitude distributions will follow a power law.

One of the factors that contributes to the difference in the calculated b -values for the considered mines is the distinct blast regimes adopted (Blake and Hedley, 2004). Kidd Creek D was in development and operated with small (Fig 3.1a) but frequent (Fig. 3.2a) blasts of an average magnitude of $-0.80\mathbf{M}$, whereas large (Fig. 3.1b) and isolated (Fig. 3.2b) blasts with an average magnitude of $0.23\mathbf{M}$ were used at Macassa mine.

The employment of tests used for regional seismicity such as the one proposed by Wiemer and Wyss (2000) can be biased in a mining seismicity scale to determine the minimum magnitude of completeness. The reason is the underlying assumptions in these methods, that often consider the frequency-magnitude distribution of events to have an exponential form. Aki (1987) used borehole data in California to show that this assumption may be troublesome even in a regional scale. Figs. 3.3(a) and 3.3(b) clearly do not display a linear trend such as the one observed for the regional seismicity.

The magnitude of completeness M_c in this case was then determined as the cutoff magnitudes in Vallejos and McKinnon (2011). Events below the magnitude bin with the highest number of events in a non-cumulative frequency-magnitude distribution of the mining seismicity were discarded. This corresponds to the magnitudes in Figs. 3.3(a) and 3.3(b) where the roll-off ends. These magnitudes are $M_c = -1.93\mathbf{M}$ and $M_c = -1.25\mathbf{M}$ for Kidd Creek D and Macassa Mine, respectively.

3.3.2 Spatial/Temporal Discretization

The TM metric can offer important constraints to both spatial and temporal discretization of the system as well as the forecast magnitude related to the application of the PI and RI methods to mining seismicity. Ergodicity can be crucial to classical statistical analysis of systems given the simplifications that result from it, where time averages of a single particle are the same as the ensemble averages. The TM metric is of great importance in the estimation of the parameters considered in this case. However, it is important to stress that ergodicity is a critical feature frequently needed to ensure the validation of a classical statistical study of a system, but it does not guarantee the success of the method.

In this procedure, seismic events excluding blasts were considered for the computation of the TM metric. Both regions of interest were divided into a 3D mesh of cubes with different edge lengths and various time increments - from 7 to 30 days - to construct the time series of the cumulative number of events for each box i . Fig. 3.4 displays the inverse TM metric for Kidd Creek D mine using different edge lengths and time units. The same is plotted in Fig. 3.5 for Macassa mine. All the linear trends to define effective ergodic periods were obtained for a Pearson's correlation coefficient greater than 0.99.

For Kidd Creek D, long periods of effective ergodicity are observed from the 49th to the 91th week and from 98th to the 175th week (Fig. 3.4a). The same time periods are also effective ergodic on a monthly base, from the 12nd to 22nd month and from the 23rd to 41st month (Fig. 3.4b). A disruption in the effective ergodic periods is

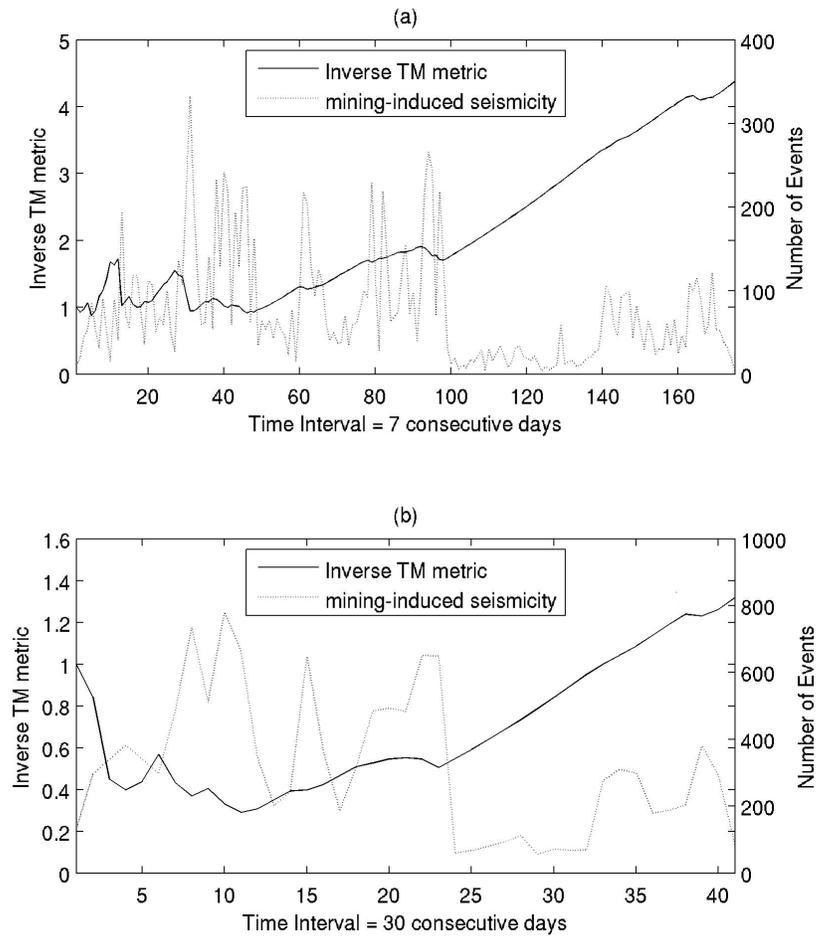


Figure 3.4: Inverse TM metric and seismicity rates for a 3D mesh at Kidd Creek D mine considering: 7-day time series with cube edge of 10m (a) and 30-day time series with cube edge of 20m (b). Most features of the inverse TM metric are roughly the same for the different space/time configurations.

observed in both cases is due to the change in seismicity rate around the 95th week and 41st month. This change is attributed to the drop in blasting activity and their low magnitude, as illustrated in Figs. 3.1(a) and 3.2(a) respectively.

Unlike the previous case, no long linear trend is established in Fig. 3.5 for the

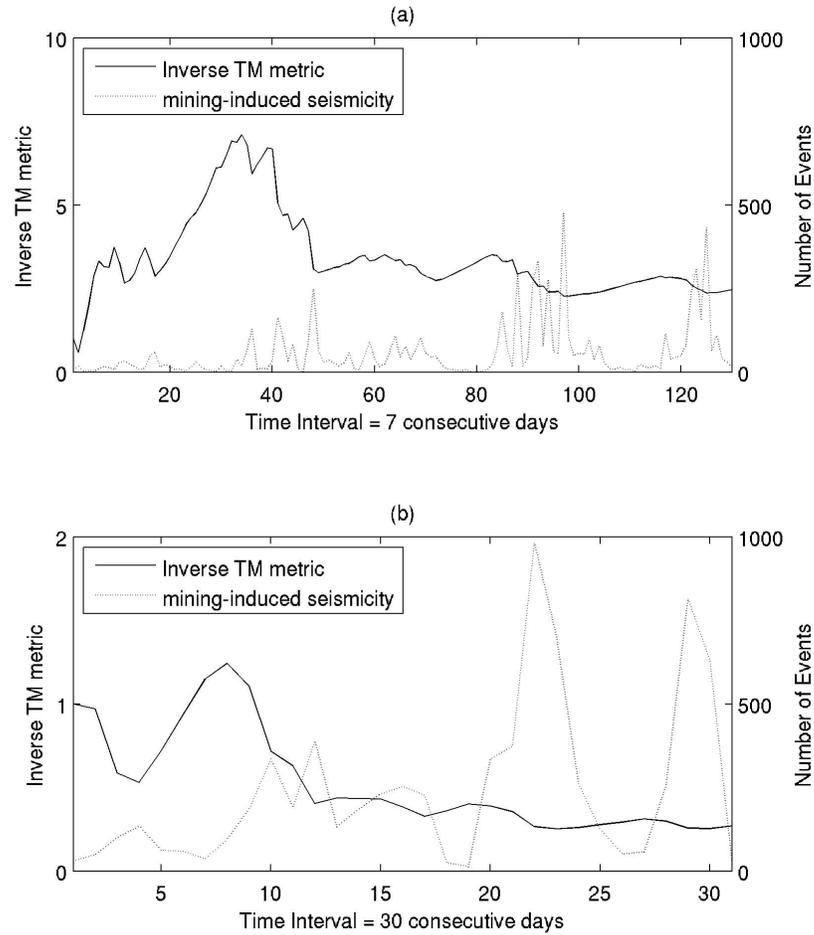


Figure 3.5: Inverse TM metric and seismicity rates for a 3D mesh at Macassa mine considering: 7-day time series with cube edge of 10m (a) and 30-day time series with cube edge of 20m (b). Most features of the inverse TM metric are roughly the same for the different space/time configurations.

inverse TM metric at Macassa mine regardless of the space/time configuration used. The sudden changes of the seismicity rate at this location in both cases prevent the system from reaching long effective ergodic periods. As observed for Kidd Creek D, the blasting activity in this mine is a key factor for such behavior. The large-

magnitude blasting regime implemented in this mine (Fig. 3.2b) yields the highly clustered seismicity observed in Fig. 3.4(b). These sequences of aftershocks in mining seismicity are clustered in time and they obey the modified Omori's law (Vallejos and McKinnon, 2010).

The geology of the mines may also contribute to the results obtained by the TM metric for the mines considered. A number of sub-vertical faults cutting through the orebody that are close enough to be influenced by the mining-induced stress changes are located at Kidd Creek D mine. This configuration is more susceptible to smoother stress adjustments that would lead to states of metastable equilibrium. At Macassa mine, faults are not as pervasive as the latter and may be less likely to respond to mining-induced stress changes. The result is a system most likely dominated by a coarser stress adjustment and, thus less likely to establish effective ergodic periods.

The variation of the space/time configurations does not change considerably the trends in the inverse TM metric in either mines. Based on the interpretation of the TM metric as a quantification of seismicity clustering, this indicates that the seismicity clustering observed in the different spatial/temporal discretizations seems invariant. The blasting regimes adopted in each mine seem to deeply affect the establishment of effective ergodic periods at the mines as observed in their frequency-magnitude distributions. Given the observed invariance of the TM metric with respect to different gridding, the smallest time window of a week and spatial configuration of 10m edge cubes can be chosen for both mines for the application of the PI and RI to mining seismicity.

3.3.3 Forecast Magnitude

Most of the parameters needed for the implementation of the PI and RI for mining seismicity have been determined. The magnitude of the events to be identified in the retrospective forecasts can also be estimated by the TM metric instead of being hypothesized as done for regional seismicity (Holliday et al., 2005). The key point to achieve this goal is the interpretation of the TM metric as a measurement of seismicity clustering.

Kagan and Jackson (1991) presented the idea of mainshocks to be Poisson processes in time with a fixed rate. It is then reasonable to assume that the inverse TM metric for the mainshocks would yield long periods of effective ergodicity for mainshocks. With this in mind, the forecast magnitude M_f can then be estimated by calculating the inverse TM metric for different sets of seismicity that are bounded from a varying lower limit M^* to the largest magnitude recorded M_{\max} . The scan is performed by constantly increasing M^* from the smallest magnitude recorded until a value M^{**} that yields a long effective ergodic period. This ensures that the considered set of events is declustered in space and time while displaying larger magnitudes, and thus M^{**} is set to be M_f .

Fig. 3.6 displays the inverse TM metric for both mines using 10m edge cubes and 7 day time bins. The minimum of the reported events of -1.5M was used for Kidd Creek D mine and the inverse TM metric is plotted in Fig. 3.6(a). The period of effective ergodicity in this case is the same as the one obtained in Fig. 3.4(a) for the complete catalog. Once again, the blasting regime during this period is the main

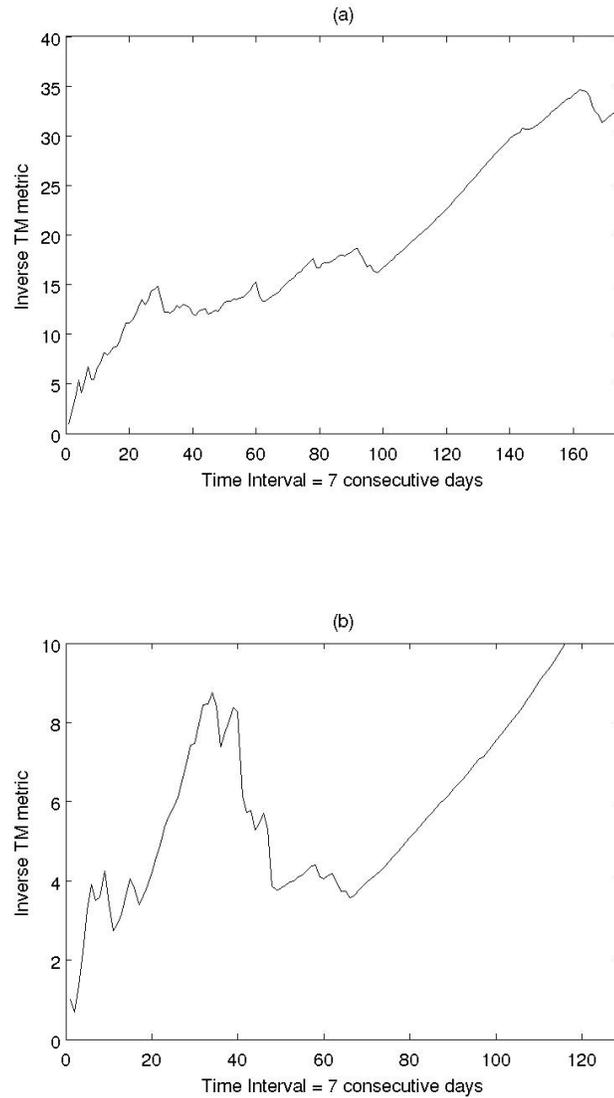


Figure 3.6: Inverse TM metric for Kidd Creed D (a) and Macassa (b) mines considering events with magnitude greater than $-1.5M$ and $-1.07M$ respectively. The same effective ergodic period highlighted in Figure 3.4 is observed for Kidd Creek D mine and a new effective ergodic period is observed for Macassa mine from the 66th onwards.

cause for such stability. It is important to stress that events that yield linear inverse TM metrics are not necessarily mainshocks, but they comprise a set of events of larger magnitude that yield a system in which the cumulative number of events in each box evolves as a normal diffusive process.

For Macassa mine, there was no reported event classification and M_f was determined by employing the method described here. Fig. 3.6(b) displays the inverse TM metric for a cutoff magnitude of -1.07M which was chosen as M_f . A period of effective ergodicity, not present when the complete catalog was considered, is now observed from the 63rd week onwards. The blasting activity in this mine that constantly disrupted the inverse TM metric for the entire catalog does not disrupt the effective ergodic periods observed for events larger than -1.07M.

3.3.4 PI and RI in Mining Induced Seismicity

With the majority of the parameters for the PI and RI defined, a series of retrospective forecasts with different final times t_2 were computed to analyze the use of seismicity patterns as a precursory signal for large events determined in the previous section in the mines considered. Two time intervals Δt , one of two and another of four weeks, were considered in these retrospective forecasts for different values of t_2 for both mines. The results of the PI were then compared to the results obtained by the RI.

The decision threshold, D , for each retrospective forecast was chosen to maximize the Heidke score for both PI and RI forecasts. The choice of the HSS over ROC diagrams to quantify the efficiency of the forecasts was to allow for a series of forecasts

at different instances of time. The result is a simple and systematical comparison between the results from the PI and RI for a set of retrospective forecasts. Figs. 3.7 and 3.8 display the maximum Heidke score for both methods considering a given t_2 and each Δt for Kidd Creek D and Macassa mines respectively. The overall trend is a better performance of the PI for Kidd Creek D mine regardless of the time interval Δt . The performances of the PI and RI are comparable for Macassa mine with some instances of time where the PI performs better, in particular for $\Delta t = 4$.

It is important to stress that the RI highlights locations based on the historical seismic activity and it does not take into account the effects of blasts immediately before t_2 . On the other hand, the PI considers such activity when choosing the Δt time interval prior to t_2 as a training period. The plots of the maxima for the HSS at Kidd Creek D mine in Fig. 3.7 for both values of Δt show that the PI outperforms the RI consistently due to the constant blasting rate in the period considered at Kidd Creek D mine (Fig. 3.2a). The time period considered fall within the long effective ergodic periods in Fig. 3.4(a), during which the system is in a metastable equilibrium state as observed for regional seismicity (Tiampo et al., 2007). The requirements of the PI of a slowly driven system are met for mining seismicity in this case.

The cumulative number of events in each box evolves as a random walk during the effective ergodic periods (Cho et al., 2010). As a result, the occurrences of quiescence/activation periods are more easily identified, and the PI method is sensitive to these particular precursory signals. The result is the overall better performance of the PI observed in Fig. 3.7 for both values of Δt . This is in agreement with recent studies on the performance of the PI on a regional scale (Tiampo et al., 2010) and

it offers a physical explanation for such behavior. Instances of poor performance of the PI during this period, around the 105th and 126th weeks, occur when there was a change in the blasting sites. This leads to a sudden change in seismicity patterns that the PI cannot take into account instantaneously. Some time after the relocation of the blasting sites, the performance of the PI increases again as expected.

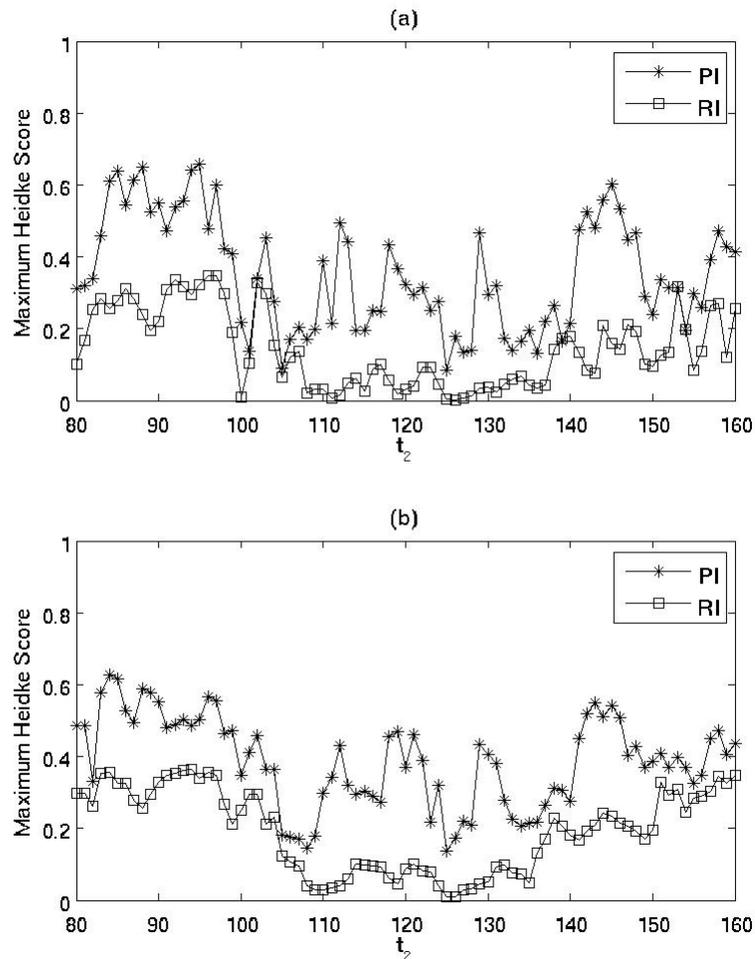


Figure 3.7: Plot of the maximum values of the Heidke score for a 7-day time series with cube edge of 10m for both PI and RI at Kidd Creek D mine for $\Delta t = 2$ (a) and 4 weeks (b). A magnitude forecast of $-1.25M$ was considered in this case.

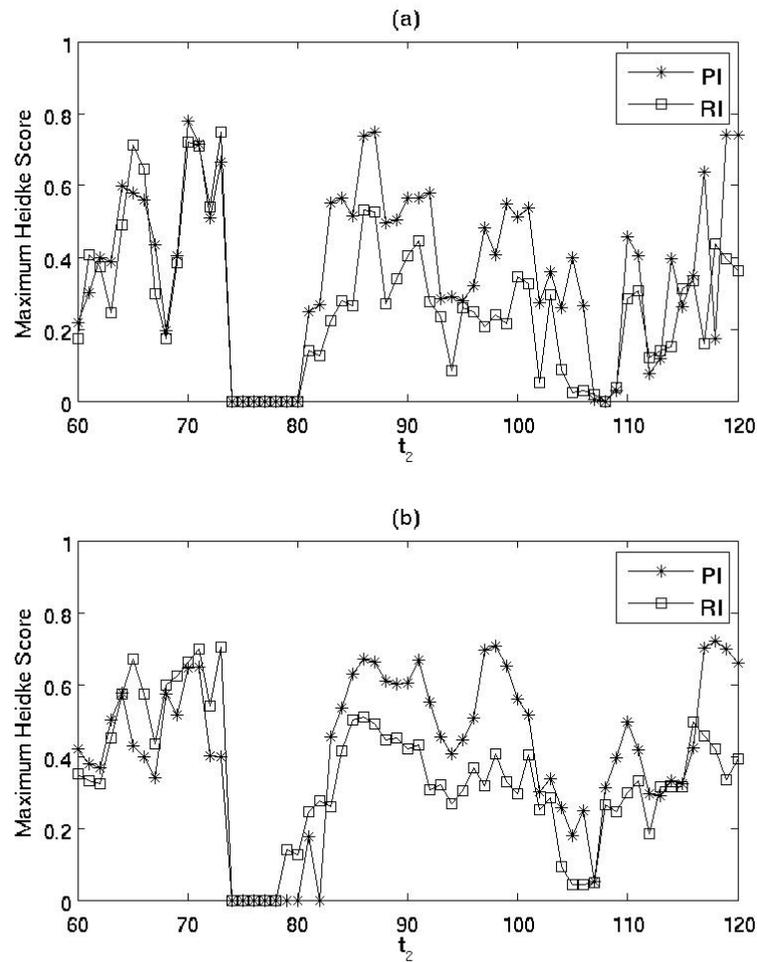


Figure 3.8: Plot of the maximum values of the Heidke score for a 7-day time series with cube edge of 10m for both PI and RI at Macassa mine for $\Delta t = 2$ (a) and 4 weeks (b). A magnitude forecast of $-1.07M$ was considered in this case.

For Macassa mine, the highly irregular blasting regime adopted generated strong clustering throughout the period considered and an effective ergodic period was not obtained (Fig. 3.5a). This uneven clustering weakens the efficiency of seismicity patterns as a precursory signal for large events in the mine, resulting in similar performances of the PI and RI for $\Delta t = 2$ in Fig. 3.8(a). From the 74th to the 80th

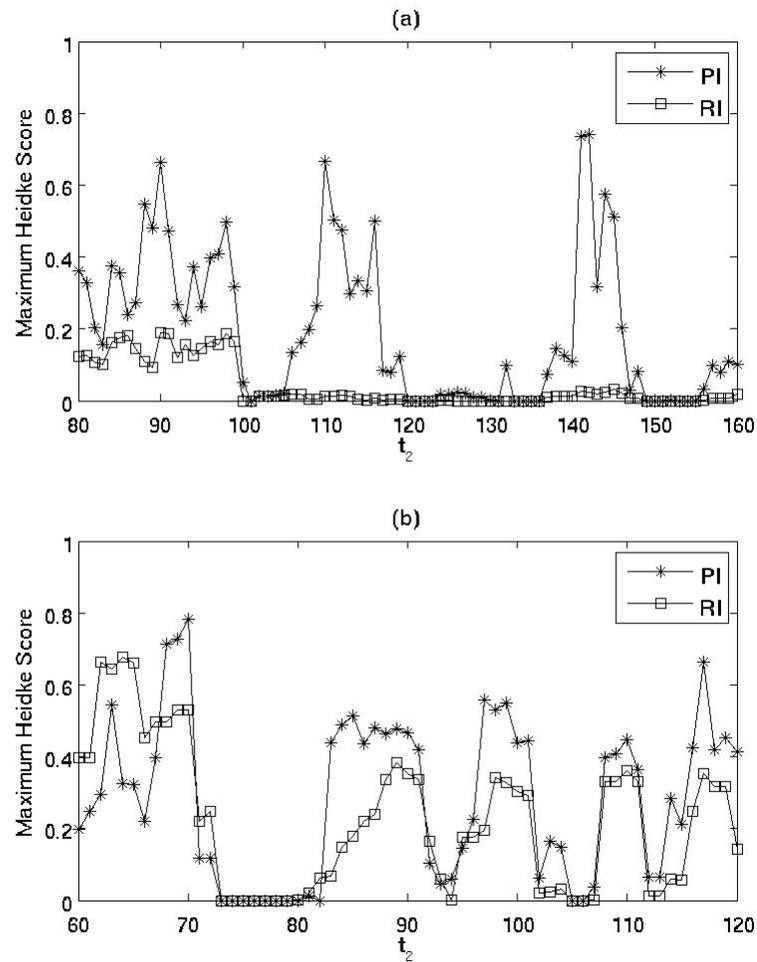


Figure 3.9: Plot of the maximum values for the Heidke score for a 7-day time series with cube edge of 10m for both PI and RI at Kidd Creek D (a) and at Macassa (b) mine for $\Delta t = 4$. Forecast magnitudes of $-0.73M$ and $-0.7M$ was considered for Kidd Creek D and Macassa mines, respectively.

week, there was a change in the location of the majority of the blasts followed by a period without blasting at the mine considered. This resulted in the occurrence of seismicity in locations with no previous history of activity and a sudden decrease in the occurrence of events altogether that resulted in the poor performance of both PI

and RI displayed in Fig. 3.8. A change in blasting sites occurred around the 93th week, resulting in the decrease in performance for both Δt . For $\Delta t = 4$, better performance of the PI is observed for t_2 values from 82 to 100 (Fig. 3.8b). A comparison between Figs. 3.2(b) and 3.8 indicates that these isolated instances of poor performance of both the PI and RI around the 110th week can be attributed to the halt in blasting during this period. Problems in the estimation of intermediate-to-high magnitudes due to the clipping of seismic waveforms may have influenced estimation of the forecast magnitude for Macassa mine, and thus the retrospective forecasts themselves.

The differences between the blasting regimes adopted in the mines are most likely the cause of these divergent responses. The frequent but small blasting activity at Kidd Creek D mine displayed in Figs. 3.1(a) and 3.2(a) seems to accelerate without changing drastically the dynamics of the system, given how constant they were during the period considered. This is analogous to the idea that small earthquakes are an integral part of the dynamics that includes the triggering of earthquakes for regional seismicity in southern California (Helmstetter, 2003). The use of large blasts (Fig. 3.1b), regardless of their lower frequency (Fig. 3.2b), seems to offset the effect of the abundance of small events. The main issue is the high variability in blasting practice, which does not allow for a pattern to be established. This poses as an important issue for methods that rely on pattern recognition such as the PI, since the triggering mechanism is affected considerably by external factors such as blasting. The result is that the PI does not offer a considerable gain over the RI.

A comparison between the performances of the PI and the RI for both mines suggest that seismicity patterns can be useful for seismic hazard in mines. In particular, the PI performs better at Kidd Creek D mine and for an instance of time at Macassa mine for $\Delta t = 4$. The good performance of the PI for $\Delta t = 4$ is in agreement with previous studies on the hazard assessment in deep gold mines in South Africa (Spottiswoode, 2010), where monthly assessment based on overall seismicity rate was suggested.

The same scenarios were considered with higher forecast magnitudes in Fig. 3.9 for both mines and $\Delta t = 4$. This choice was based on the changes in the linear trends of the frequency-magnitude plots in Fig. 3.3 that occur at $-0.73M$ and $-0.7M$ for Kidd Creek D and Macassa mines, respectively. The performance of the RI in Fig. 3.9(a) drops considerably compared to the performance shown in Fig.3.7(a) with the increase of the forecast magnitude for Kidd Creek D mine, especially after the 100th week. The PI performance displays a similar trend to the one in Fig. 3.7(b), with the exception of a sudden drop from the 120th to the 135th weeks and from the 150th week onwards due to a decrease in the occurrence of events with magnitude $-0.73M$ or greater during these periods. For Macassa mine, the performances of both PI and RI displayed in Fig. 3.9(b) are similar to the ones in Fig. 3.8(b). The drops of efficiency in both methods around the 93th and 112th weeks were enhanced when the forecast magnitude increases to $-0.7M$. These results corroborate the idea that the PI may offer useful information to seismic hazard assessment in mines.

3.4 Conclusion

The employment of the PI and RI, two intermediate-to-long term forecast methods well established for regional seismicity, was attempted for mining induced seismicity to analyze the applicability of seismicity patterns as precursory signal to large events in two distinct mining environments. The catalogs of induced seismicity from the mines considered are completely different with respect to their blasting regimes. The main difficulty for this implementation was to determine the most appropriate parameters to be used in the low magnitude range that encompasses seismic activity in mines. The TM metric was of extreme importance in this task to determine the spatial/temporal discretization of the system as well as the magnitudes of large events to be considered in the PI and RI. As a result, the choice of parameters was made on a physical basis rather than on an *ad-hoc* one.

The interpretation of the TM metric as a measurement of seismicity clustering was crucial in this framework. During effective ergodic periods in Kidd Creek D mine, the PI performed considerably better than the RI given the declustered nature of seismicity. This allowed for a better identification of instances of anomalous seismic activity at Kidd Creek D mine. The lack of an effective ergodic period at Macassa mine showed that precursory signal is damped by the clustering generated by the large blasts, which occurred at irregular time intervals. The PI did not perform consistently and considerably better than the RI in this case.

The blasting regime of the mines played a crucial role in the different aspects of mining induced seismicity that were covered. Less frequent high-magnitude blasts

and changes in blasting sites had considerable influence on the effective ergodicity of mining induced seismicity. These factors suppressed the establishment of a seismicity trend long enough to improve the efficiency of the PI. Taking them into consideration when performing a PI analysis is an important issue to be considered.

Acknowledgments

This research was supported by a grant from the Workplace Safety and Insurance Board (WSIB) of Ontario to develop re-entry protocols in seismically active mines and the NSERC & Aon Benfield/ICLR Industrial Research Chair in Earthquake Hazard assessment. The authors wish to acknowledge the permission of the mines studied to publish this work and thank J. B. Setyobudhi for his valuable comments.

References

- Aki, K. (1987). Magnitude-frequency relation for small earthquakes: a clue to the origin of f_{\max} of large earthquakes. *Journal of Geophysical Research*, 92(B2):1349–1355.
- Blake, W. and Hedley, D. G. F. (2004). *Case studies from north American hard-rock mines*. Society for Mining Metallurgy & Exploration.
- Chen, C. C., Rundle, J. B., Holliday, J. R., Nanjo, K. Z., Turcotte, D. L., Li, S. C., and Tiampo, K. F. (2005). The 1999 chi-chi, taiwan, earthquake as a typical example of seismic activation and quiescence. *Geophysical Research Letters*, 32(L22315):1–4.
- Cho, N. F., Tiampo, K. F., McKinnon, S. D., Vallejos, J. A., Klein, W., and Dominguez, R. (2010). A simple metric to quantify seismicity clustering. *Non-linear Processes in Geophysics*, 17:293–302.
- Cicerone, R. D., Ebel, J. E., and Britton, J. (2009). A systematic compilation of earthquake precursors. *Tectonophysics*, 476:371–396.
- de Oliveira, C. R. and Werlang, T. (2007). Ergodic hypothesis in classical statistical mechanics. *Revista Brasileira de Ensino de Física*, 29(2):189–201.
- Eneva, M. (1998). In search for a relationship between induced microseismicity and larger events in mines. *Tectonophysics*, 289:91–104.
- Eneva, M. and Ben-Zion, Y. (1997). Application of pattern recognition techniques

- to earthquake catalogs generated by model of segmented fault systems in three-dimensional elastic solids. *Journal of Geophysical Research*, 102(B11):24513–24528.
- Eneva, M. and Young, P. (1993). Evaluation of spatial patterns in the distribution of seismic activity in mines: A case study at creighton mine, northern ontario (canada). *Rockbursts and Seismicity in Mines*, pages 175–180.
- Farquhar, I. E. (1964). *Ergodic theory in statistical mechanics, Monographs in Statistical Physics Volume 7*. Inerscience Publishers.
- Gibowicz, S. J. and Kijko, A. (1994). *An introduction to mining seismology*. Academic New York.
- Glowacka, E., Stankiewicz, T., and Holub, K. (1990). Seismic hazard estimate based on the extracted deposit volume and bimodal character of seismic activity. *Gerlands Beitrage zur Geophysik*, 99:35–43.
- Heidke, P. (1926). Berechnung des erfolges und der gute der windstarkevorhersagen im sturmwarnungsdienst. *Geografiska Annaler*, 8:301–349.
- Helmstetter, A. (2003). Is earthquake triggering driven by small earthquakes? *Physical Review Letters*, 91(5):1–4.
- Holliday, J. R., Nanjo, K. Z., Tiampo, K. F., Rundle, J. B., and Turcotte, D. L. (2005). Earthquake forecasting and its verification. *Nonlinear Processes in Geophysics*, 12:965–977.
- Holliday, J. R., Rundle, J. B., Tiampo, K. F., and Turcotte, D. L. (2006). Using

- earthquake intensities to forecast earthquake occurrence times. *Nonlinear Processes in Geophysics*, 13:585–593.
- Holub, K. (1996). Space-time variations of the frequency-energy relation for mining induced seismicity in the ostrava-karviná mining district. *Pure and Applied Geophysics*, 146(2):265–280.
- Holub, K. (2007). A study of mining-induced seismicity in czech mines with longwall coal exploitation. *Journal of Mining Science*, 43(1):32–39.
- Hudyma, M. R. and Potvin, Y. (2004). Seismic hazard in western australian mines. *Journal of the South African Institute of Mining and Metallurgy*, 104(5):265–276.
- Kagan, Y. Y. and Jackson, D. D. (1991). Long-term earthquake clustering. *Geophysical Journal International*, 104:117–133.
- Kanamori, H. (1981). *The nature of seismicity patterns before large earthquakes, in Earthquake Prediction: an International Review*. American Geophysical Union.
- Keilis-Borok, V. (2005). Earthquake prediction: state-of-the-art and end emerging possibilities. *Annual Review of Earth and Planetary Sciences*, 30:1–33.
- Lasocki, S. (1990). Prediction of strong mining tremors. *Geofizyka Stosowana*, 7:1–110.
- Moulik, P. (2009). Computational studies in earthquake physics and fault-system science: implications on fault network reconstruction and earthquake characterization. Technical report.

- Nanjo, K. Z., Holliday, J. R., Chen, C. C., Rundle, J. B., and Turcotte, D. L. (2005). Application of a modified pattern informatics method to forecasting the locations of future large earthquakes in the central japan. *Tectonophysics*, 424:351–366.
- Omori, F. (1894). On the aftershocks of earthquakes. *The Journal of the College of Science, Imperial University of Tokyo*, 7:111–200.
- Oye, V., Bungum, H., and Roth, M. (2005). Source parameters and scaling relations for mining-related seismicity within the pyhasalmi ore mine, finland. *Bulletin of the Seismological Society of America*, 95(3):1011–1026.
- Richardson, E. and Jordan, T. H. (2002). Seismicity in deep gold mines of south africa: implications for tectonic earthquakes. *Bulletin of the Seismological Society of America*, 92(5):1766–1782.
- Rundle, J. B., Tiampo, K. F., Klein, W., and Martins, J. S. S. (2002). Self-organization in leaky threshold systems: The influence of near-mean field dynamics and its implications for earthquakes, neurobiology, and forecasting. *Proceedings of the National Academy of Sciences of the United States of America*, 99(Supplement 1):2514–2521.
- Shcherbakov, R., Turcotte, D. L., Rundle, J. B., Tiampo, K. F., and Holliday, J. R. (2010). Forecasting the locations of future large earthquakes: an analysis and verification. *Pure and Applied Geophysics*, 167:743–749.
- Simpson, D. W. (1986). Triggered earthquakes. *Annual Review of Earth and Planetary Sciences*, 14:21–42.

- Spottiswoode, S. M. (2010). Mine seismicity: prediction or forecasting. *Journal of the South African Institute of Mining and Metallurgy*, 110:11–20.
- Thirumalai, D. and Mountain, R. D. (1993). Activated dynamics, loss of ergodicity, and transport in supercooled liquids. *Physical Review E*, 47:479–489.
- Thirumalai, D., Mountain, R. D., and Kirkpatrick, T. R. (1989). Ergodic behavior in supercooled liquids and in glasses. *Physical Review A*, 39:3563–3574.
- Tiampo, K. F., Klein, W., Li, H. C., Mignan, A., Toya, Y., Kohen-Kadoshand, S. Z. L., Rundle, J. B., and Chen, C. C. (2010). Ergodicity and earthquake catalogs: Forecast testing and resulting implications. *Pure and Applied Geophysics*, 167(6–7):763–782.
- Tiampo, K. F., Rundle, J. B., and Klein, W. (2006). Stress shadows determined from a phase dynamical measure of historic seismicity. *Pure and Applied Geophysics*, 263:2407–2416.
- Tiampo, K. F., Rundle, J. B., Klein, W., Holliday, J. B., Martins, J. S. S., and Ferguson, C. D. (2007). Ergodicity in natural earthquake fault networks. *Physical Review E*, 75(066107):1–15.
- Tiampo, K. F., Rundle, J. B., Klein, W., Martins, J. S. S., and Ferguson, C. D. (2003). Ergodic dynamics in a natural threshold system. *Physical Review Letters*, 91(23):1–4.
- Tiampo, K. F., Rundle, J. B., McGinnis, S., and Klein, W. (2002). Pattern dynamics

- and forecast methods in seismically active regions. *Pure and Applied Geophysics*, 159(10):2429–2467.
- Tinti, S. and Mulargia, F. (1985). Effect of magnitude uncertainties on estimating the parameters in the gutenbergrichter frequency-magnitude law. *Bulletin of the Seismological Society of America*, 75(6):1681–1697.
- Utsu, T. (2002). *Statistical features of seismicity. International Handbook of Earthquake and Engineering Seismology, Part A*. Academic Press, Amsterdam.
- Vallejos, J. A. and McKinnon, S. D. (2010). Omoris law applied to mining-induced seismicity and re-entry protocol development. *Pure and Applied Geophysics*, 167(1–2):91–106.
- Vallejos, J. A. and McKinnon, S. D. (2011). Correlations between mining and seismicity for re-entry protocol development. *International Journal of Rock Mechanics and Mining Sciences*, 48(4):616–625.
- Wiemer, S. and Wyss, M. (2000). Minimum magnitude of completeness in earthquake catalogs: examples from alaska, the western united states, and japan. *Bulletin of the Seismological Society of America*, 90(4):859–869.
- Wilks, D. S. (1995). *Statistical methods in the atmospheric sciences*. Academic Press.
- Woodcock, F. (1976). The evaluation of yes/no forecasts for scientific and administrative purposes. *Geophysical Research Letters*, 104(10):1209–1214.

Zechar, J. D. and Jordan, T. H. (2008). Testing alarm-based earthquake predictions.
Geophysical Journal International, 172:715–724.

Chapter 4

Declustering seismicity using the TM metric³

4.1 Introduction

Clustering is an important branch in data analysis and several methods to study agglomerations are available in the literature (Jain and Dubes, 1988; Hartigan, 1975). A simple examination of the spatial and temporal distributions of regional earthquakes show that they primarily occur along tectonic features such as faults (spatial clustering) and close in time to large events (temporal clustering). For example, the Omori Law (Omori, 1894) is an empirical relation that describes the clustering of earthquake aftershocks in time. Dobrovolsky et al. (1979) calculated a strain radius that delineates the area where events can be related to a precursor and Utsu (2002) offers a broad review on the spatial distribution of seismicity. Stein (1999) also incorporated the principle of spatial clustering when associating the occurrence of triggered events or aftershocks to regions of stress increase.

The inherent clustered nature of earthquakes has motivated several studies to understand and model seismicity. The Epidemic Type Aftershock Sequence (ETAS) (Ogata, 1988; Helmstetter and Sornette, 2002) and the Branching Aftershock Sequence (BASS) (Turcotte and Holliday, 2007) are examples of techniques that model seismic nucleation in terms of larger events that generate smaller ones. Other exam-

³submitted to Bulletin of the Seismological Society of America

ples are the analysis of the interoccurrence times between events as a non-homogeneous Poisson process by Shcherbakov et al. (2005) and the study of the correlation between the locations of aftershocks and regions of coseismic slip in southern California by Mendonza and Hartzell (1988).

Declustering data has been the subject of intensive study over the years, where the main goal is to separate a given data set into subsets in which their elements share similar characteristics based on a particular set of criteria. Well-known methods for declustering are the k-means (MacQueen, 1967; Huang, 1998) and techniques based on fuzzy c-means (Dunn, 1973; Bezdek, 1981). In the k-means technique, centroids represent clusters that are a function of the average of a given characteristic over all elements in a small subset. The likelihood that an element belongs to more than one cluster is considered in the fuzzy c-means approach.

For seismicity, several attempts have been made in the past to decluster catalogs. Gardner and Knopoff (1974) developed a declustering method for southern California using time and space windows for different magnitude cutoffs to identify clustered events. Events within these windows are considered clustered and removed from the catalog. A more refined procedure was introduced by Reasenber (1985) in which events are analyzed in pairs and aftershock sequences are modeled as time-dependent Poisson processes. Based on the space/time distance of the events, the pair is classified as clustered or not. Both these methods were developed mostly for California and are heavily dependent on the parameters used in the analysis. Zaliapin et al. (2008) upgraded the work presented in Baiesi and Paczuski (2004), in which time-space-magnitude distances between earthquakes are analyzed, to identify clustered and

declustered events in a framework that is not parameter-dependent.

A method to decluster seismicity based on determining the probability of direct and indirect aftershock triggering was proposed by Marsan and Lengliné (2008). No a priori model is needed to perform this declustering and it was found that a cascading effect for small earthquakes, in which events trigger other events indirectly, plays an important role in the dynamics of the system. Hidden Markov models also have been used to decluster seismicity (Wu, 2010), where the seismic activity before and after an event is considered in order to determine if this event belongs to the background seismicity. In this study, a different method to decluster seismicity is proposed based on the statistical properties of the data over long time intervals. This technique employs a concept from statistical mechanics, the Thirumalai-Mountain (TM) metric (Thirumalai et al., 1989; Thirumalai and Mountain, 1993), as a basis for measuring seismic clustering as deviations from the long-term ergodic averages of the fault system (Cho et al., 2010).

In section 4.2, the TM metric and its relationship to seismicity clustering is introduced. The proposed declustering method is presented in section 4.3 and applied to a synthetic catalog, the southern California catalog, the Taiwanese catalog and a region comprised of southern Spain and northern Africa in section 4.4. Finally, the conclusion and final remarks are presented in section 4.5.

4.2 The TM Metric and Clustering

The ergodic hypothesis is an important assumption in different formulations of classical statistical mechanics because it provides a simple link to relate micro and macro states of a system (Farquhar, 1964). For a thermodynamic system, a simple definition of the ergodic hypothesis states that

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T f(t) dt = \langle f \rangle \quad (4.1)$$

where $f(t)$ is a phase-space function and $\langle f \rangle$ is its ensemble average. This definition means that the temporal average of $f(t)$ of a single particle of the system is equal to the ensemble average of $f(t)$ for a given time t .

The original application of the TM metric was to study effective ergodicity in liquids and glasses (Thirumalai et al., 1989). In this context, the term effective meant that long but finite time intervals are considered in the analysis in order to assure that the phase space is sampled evenly (Thirumalai and Mountain, 1993). For the N -body system, the metric is written as

$$\Omega_G(t) = \frac{1}{N} \sum_{i=1}^N [g_i(t) - \langle g \rangle]^2 \quad (4.2)$$

where $g_i = \frac{1}{t} \int_0^t G_i(t) dt$ is the time average of $G_i(t)$ until t , $G_i(t)$ is an observable for particle i , N is the number of particles in the system, and $\langle g \rangle = \frac{1}{N} \sum_i g_i(t)$ is the average of g_i over all particles i . From these definitions, equation 4.2 can be written

as:

$$\begin{aligned}
\Omega_G(t) &= \frac{1}{N} \sum_{j=1}^N [g_j(t)^2 - 2g_j(t) \langle g \rangle + \langle g \rangle^2] \\
&= \left[\frac{\sum_j g_j(t)^2}{N} \right] - 2 \left[\frac{\sum_j g_j(t)}{N} \right] \langle g \rangle + \frac{\sum_j \langle g \rangle^2}{N} \\
&= \langle g(t)^2 \rangle - \langle g(t) \rangle^2 \\
&= \frac{1}{t^2} \left[\left\langle \int_0^t G_j(t') dt' \int_0^t G_j(t'') dt'' \right\rangle - \left\langle \int_0^t G_j(t) dt \right\rangle^2 \right]. \quad (4.3)
\end{aligned}$$

Equation (4.3) is simply the variance of $\int_0^t G_j(t) dt$ and effective ergodic periods occur when its inverse is linear in time with a positive slope.

Tiampo et al. (2007) employed this metric to regional seismicity by gridding a region of interest and using the cumulative number of events $n_j(t)$ in each box j until time step t as a proxy for the released seismic energy. In this framework, $\int_0^t G_j(t') dt' = \int_0^t G_j(t'') dt'' = n_j(t)$ and equation (4.3) becomes

$$\Omega_G(t) = \frac{1}{t^2} [\langle n_j(t)^2 \rangle - \langle n_j(t) \rangle^2] = \frac{1}{t^2} \sigma^2. \quad (4.4)$$

where $\sigma^2 = \langle n_j(t)^2 \rangle - \langle n_j(t) \rangle^2$ is the variance of $n_j(t)$.

It was shown that effective ergodic periods occur and that during these periods the system is stationary and in a state of meta-equilibrium. These states are disrupted by large events and it was later shown that these disruptions in the TM metric during otherwise ergodic periods were caused by seismicity clustering (Cho et al., 2010).

The latter study offers an in-depth analysis of the TM metric when applied to earthquake systems. It employed the description of the TM metric in terms of the variance of $n_j(t)$ to establish that $\langle n_j(t)^2 \rangle$ is dependent on both spatial and temporal clustering whereas $\langle n_j(t) \rangle^2$ is a measurement of temporal clustering only for a fixed

number of events in a time step t . The scaling of the inverse of the metric during effective ergodic periods plays a crucial role in understanding the effects of spatiotemporal seismicity clustering in the TM metric. During these instances of time,

$$\langle n_j(t)^2 \rangle - n_j(t)^2 \propto t. \quad (4.5)$$

Equation (4.5) describes the variance of a normal diffusion process, such as a Brownian motion. As a result, seismicity clustering can be quantified by examining the features of σ^2 : if it is linear in time, seismicity is considered to be random in both space and time; any deviations from the linear trend are the result of clustering in space and/or time.

4.3 Declustering Method

In this framework, effective ergodic periods are considered those in which the occurrence of seismicity is random in both space and time, ie. it is not clustered. The declustering of seismic catalogs using the TM metric then becomes a simple problem once the metric is applied to a particular region: given a complete dataset, how many events must be extracted from each box so that the TM metric displays a long effective ergodic period over a given time period? The method is straightforward and it consists of the following steps:

1. an initial interval from a time t_0 to t^* is chosen in which relation (4.5) holds.

This trend represents an upper-limit for the seismic activity required for the establishment of a single long effective ergodic period in the declustered catalog,

and it is extrapolated to be used as a reference line to obtain the random portion of seismicity.

2. events that occur in the timestep $t^* + 1$ are removed so that $\langle n_j(t^* + 1)^2 \rangle - \langle n_j(t^* + 1) \rangle^2$ is the same as the value given by the reference trend obtained at step 1.
3. with the new set of $n_j(t^* + 1)$, step 2 is repeated by excluding events that occur in the next time step $t^* + 2$ and so on.

The reference line can be seen as the targeted σ^2 for the declustered portion of the catalog. Determining the number of events that must be removed from boxes in a given time step so that σ^2 for the remaining events matches the value extrapolated in the reference line is the goal. This is an optimization problem that focus on minimizing the difference between the the variances yielded by the reference line and the remaining events. This type of problem is not trivial and has been the subject of extensive studies to develop effective methods to solve them. Computational efficiency and the convergence to local extremes rather than the global extreme are the most significant challenges of such problems.

Several methods have been used over the years to address optimization problems: neural networks (Crick, 1989; Bishop, 1995), genetic algorithms (GAs) (Holland, 1992), particle swarm optimization (PSO) (Kennedy and Eberhart, 1995; Pedersen, 2010), dynamic programming (Bellman, 1972), among others. In this particular study, GAs and the PSO were tested and the PSO was chosen given its superior performance.

The PSO technique considers a group of particles located in the targeted parameter space as possible solutions to the problem in hand. The location of each particle represents the set of parameters that constitutes one single solution. The dynamics of each particle is dictated by the best position of the particle under consideration and the best position recorded from all the particles, resembling a gradient-based search. The optimized version of the PSO described in Pedersen (2010) was used in this study. In addition, boxes with three earthquakes or less were automatically considered declustered for the synthetic seismicity in order to improve the efficiency of the technique. A similar limit of boxes with two earthquakes was applied to the regional seismicity for the same reason.

Once the number of events to be removed from each box is determined, a choice must be made regarding which events to exclude from the boxes. For the TM metric, events in the same box and time step are indistinguishable. This task can be accomplished by introducing constraints on the declustering process. In this case, the frequency-magnitude distribution of the entire region is used as the appropriate constraint, given both its simplicity of implementation and the fact that it is an opportunity to include magnitude information to the method, which otherwise is not included in this particular TM formulation. As a result, events are drawn from each box in each time step according to an exponential probability distribution that yields the Gutenberg-Richter distribution for the entire region, expressed as

$$\log N(m) = a - bm, \quad (4.6)$$

where a and b are parameters and $N(m)$ is the number of events with magnitude greater than m .

Tinti and Mulargia (1985) define the probability density function (pdf) $p(m)$ that generates (4.6) as:

$$p(m) = \begin{cases} \beta\lambda e^{-\beta m}, & \text{if } m_1 \leq m \leq m_2 \\ 0, & \text{otherwise} \end{cases}, \quad (4.7)$$

where $\beta = b/\log e$ and $\lambda = (e^{-\beta m_1} - e^{-\beta m_2})^{-1}$. For each box with events to be extracted in a timestep t , a magnitude is randomly drawn from (4.7) with m_1 as the minimum and m_2 as the maximum magnitudes in the box. The event with the closest magnitude to the one drawn is removed and the process is repeated until the required number of events is extracted. Declustering is considered complete when σ^2 is linear, where the criteria for linearity is when the Pearson's correlation coefficient on its linear regression is greater than 0.99.

4.4 Results

In this section, results of the proposed declustering method are presented for a synthetic catalog based on the ETAS model and natural seismicity from southern California, Taiwan, Switzerland and a region comprised of southern Spain and northern Africa. These regions offer a good challenge for the proposed method given their variability in tectonic and seismic characteristics. Kreemer et al. (2003) indicates that the first two regions undergo high shear strain rates, whereas the last two regions display low shear strain rates.

The proposed method was applied in the same manner to all the different catalogs considered once the spatial/temporal discretization and a starting period was identified for a stationary seismicity rate in order to obtain the reference line. The results were compared to two classical declustering techniques, as appropriate: Reasenberg's method (Reasenberg, 1985) using parameters optimized for California and a modified Gardner-Knopoff (GK) method (Gardner and Knopoff, 1974; Wiemer et al., 2009) for central Europe.

For the synthetic catalog, the b -value is estimated using Aki's formulation (Aki, 1965) for continuous magnitude distributions. The estimations for southern California, Taiwan and Switzerland were made using the estimation in Tinti and Mulargia (1985) to account for the binning of magnitudes. Magnitude bins of 0.01 were considered for Taiwan and southern California from 1972 onwards, whereas 0.1 bins were considered for Switzerland and southern California prior to 1971 for the latter. Errors in the b -value were estimated for these regions to a 98% confidence level. The b -values for the region comprised of southern Spain and northern Africa were calculated from least-square regressions as done by previous studies in the region (Jimenez et al., 2006; Stich et al., 2007) and their 95% confidence level intervals were used.

To quantify whether the declustered seismicity is Poissonian in time, the coefficient of variation (C_v) of earthquake interoccurrence time δt was applied. It is defined as the ratio of the standard deviation of δt and its average. This approach was used in Kagan and Jackson (1991), where a Poissonian system in time displays $C_v = 1$, a quasiperiodic system yields a C_v in the range $(0, 1)$, and a system with temporal clustering results in $C_v > 1$. In this context, the computation of C_v should be

seen as a statistical testing scheme and not a direct implication that a given dataset is Poissonian in time or not. For the historic catalogs, the interoccurrence times between earthquakes were calculated in days. Finally, the MATLAB toolbox ZMAP (Wiemer, 2001) was used to implement the two classical declustering techniques, as noted above, for every seismic datasets.

4.4.1 Synthetic Catalog

Synthetic catalogs offer controlled test subjects for any statistical seismological method. Here the ETAS model was used to generate 27,132 events distributed in a square area with an side length of 128 arbitrary units. A total of 17,800 were randomly displaced in space and time, constituting the declustered portion of the catalog. No correlated seismicity was generated in the initial portion of the catalog, in order to enable a good estimation of the linear trend of σ^2 needed in the declustering process.

A spatial mesh of 8×8 arbitrary units boxes and a total of 50 timesteps were considered in order to generate the $n_j(t)$ time series. These parameters were chosen to loosely emulate southern California regarding the number of events per box per time unit. Figure 4.1a displays the obtained σ^2 for the original synthetic catalog along with the reference line calculated from the extrapolation of the initial trend of the latter. The deviations of σ^2 for the original catalog from a linear trend indicate the presence of clustering in space and time as expected. Increases in seismic activity in the entire catalog, as illustrated in Figure 4.1c, correlate well with the discontinuities in σ^2 for the entire catalog in Figure 4.1a, in particular the large discontinuity seen at approximately $t = 3800$.

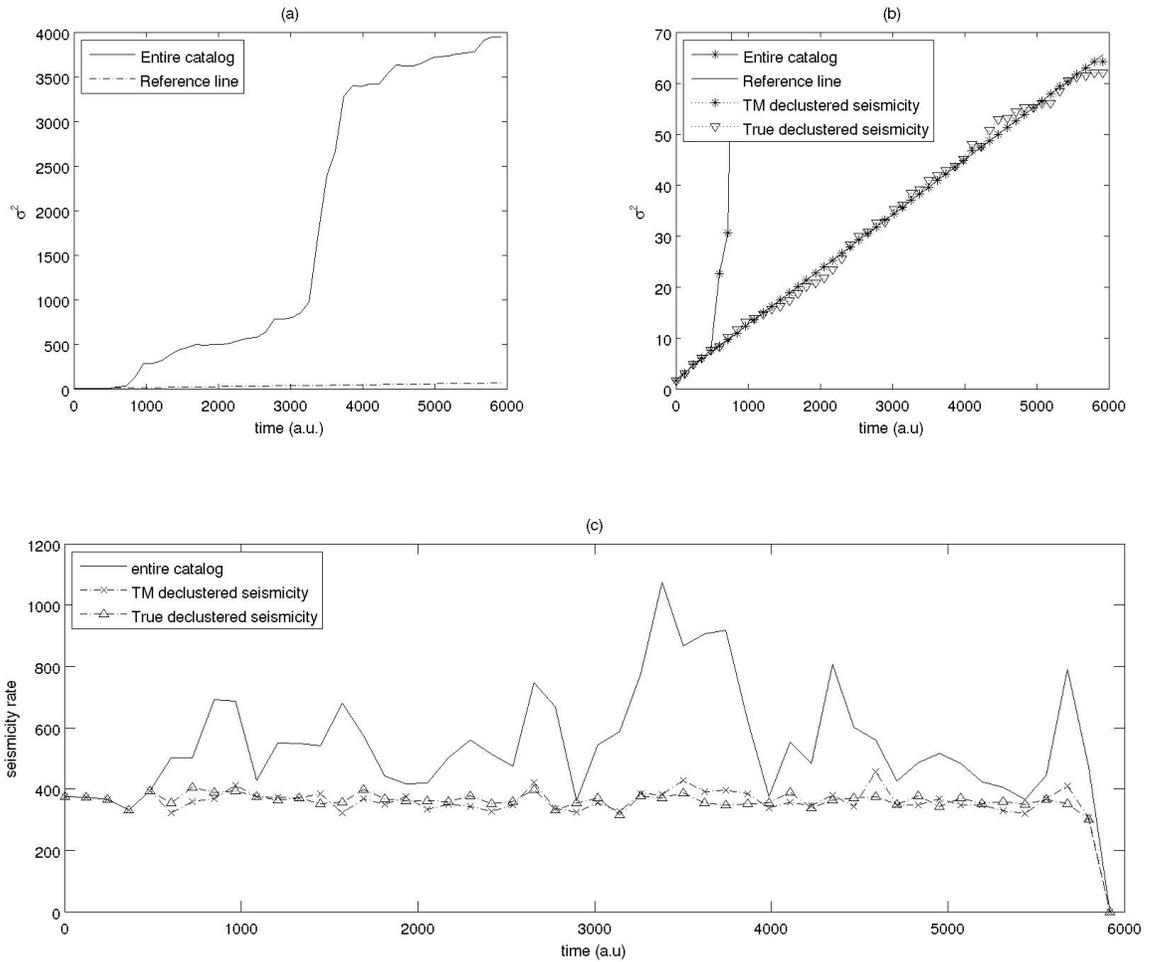


Figure 4.1: Variance of $n_j(t)$ for (a) the entire catalog along with the reference line. (b) Enlargement of the reference line and the variance of both the true declustered portion of the catalog and the events identified as declustered by the TM method. (c) Plot of the seismicity rates of the entire catalog, the original declustered portion and the identified declustered seismicity.

The declustering method based on the TM metric was applied to this dataset and the PSO determined the number of events to be removed from each box so that a linear σ^2 would be obtained. Figure 4.1b displays a closer look at the reference line along with the variances from the 17,800 events randomly generated by the ETAS, the

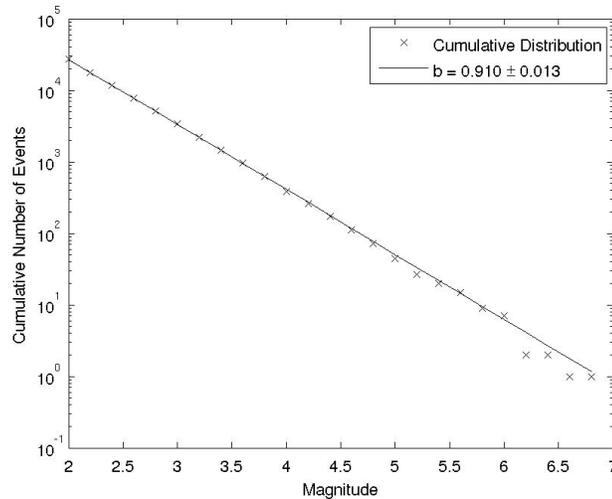


Figure 4.2: Frequency-magnitude distribution of the original synthetic dataset.

true background seismicity, and the events not removed by the PSO. The reference line fits the σ^2 of both the true background and the events not removed by the PSO as expected. A comparison between the seismicity rates from the true background portion of the catalog and the events that were not removed from the catalog by the PSO is plotted in Figure 4.1c. Both portions are constant over the period considered, as expected, and they agree reasonably well with each other.

The PSO identified a total of 9,386 clustered events from the catalog to obtain the long effective ergodic period outlined in Figure 4.1b, meaning that 17,746 events remained as declustered seismicity. Equation (4.7) was used to determine which events were removed from every box and its parameters were obtained from the Gutenberg-Richter distribution of the entire catalog illustrated in Figure 4.2. The result is a set of two catalogs identified by the method: one of clustered and another of declustered events. The declustered catalog yielded a $C_v = 1.03$, indicating that this portion of

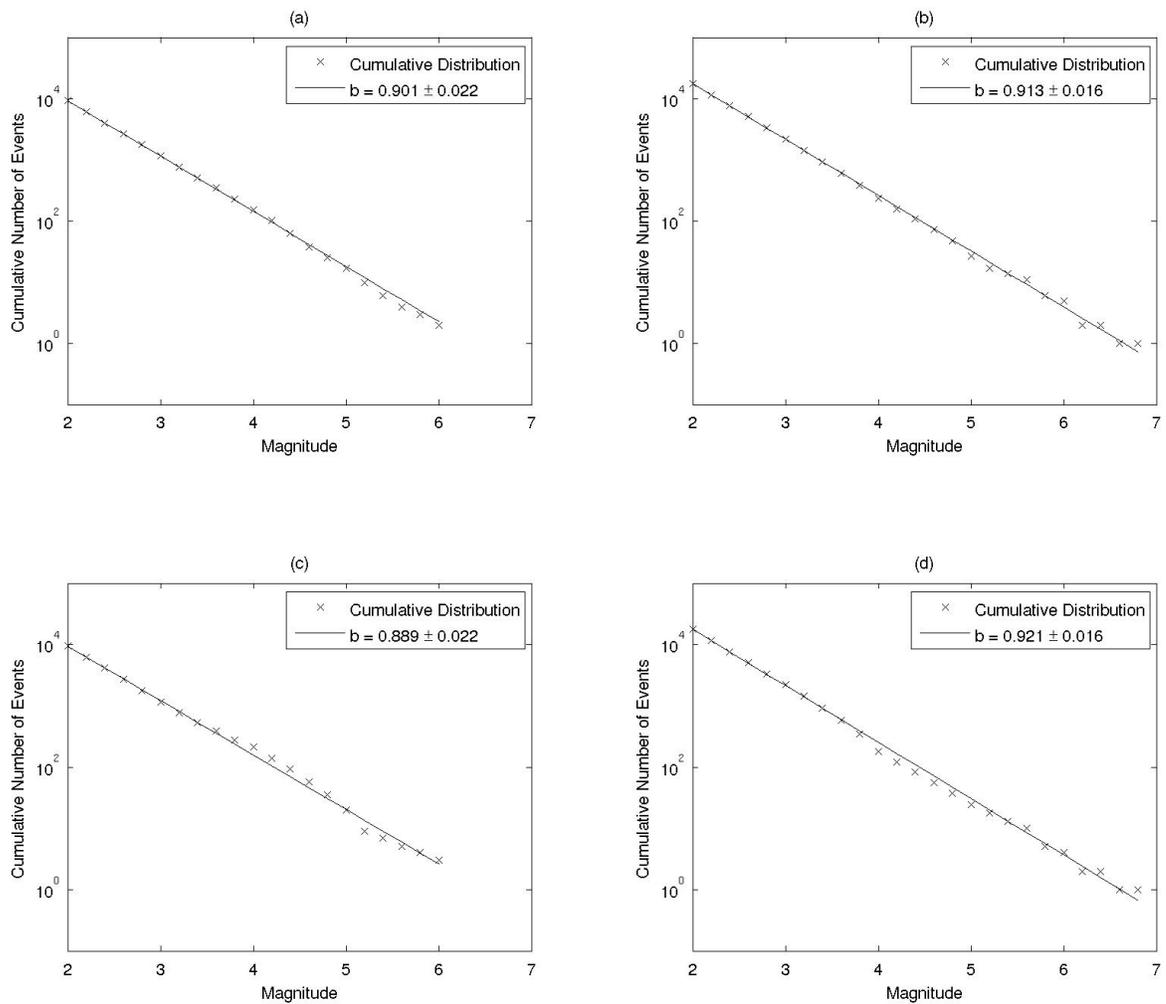


Figure 4.3: Frequency-magnitude distribution for the synthetic (a) clustered and (b) declustered events, along with the distributions for the (c) clustered and (d) declustered events identified by the proposed method.

the seismicity is Poissonian.

Figure 4.3 displays the Gutenberg-Richter distributions of the four catalogs considered: the true random seismicity, as defined by the original ETAS algorithm, and the events identified as declustered by the method, along with their respective clustered counterparts. The b -value obtained for the clustered seismicity of 0.889 ± 0.022

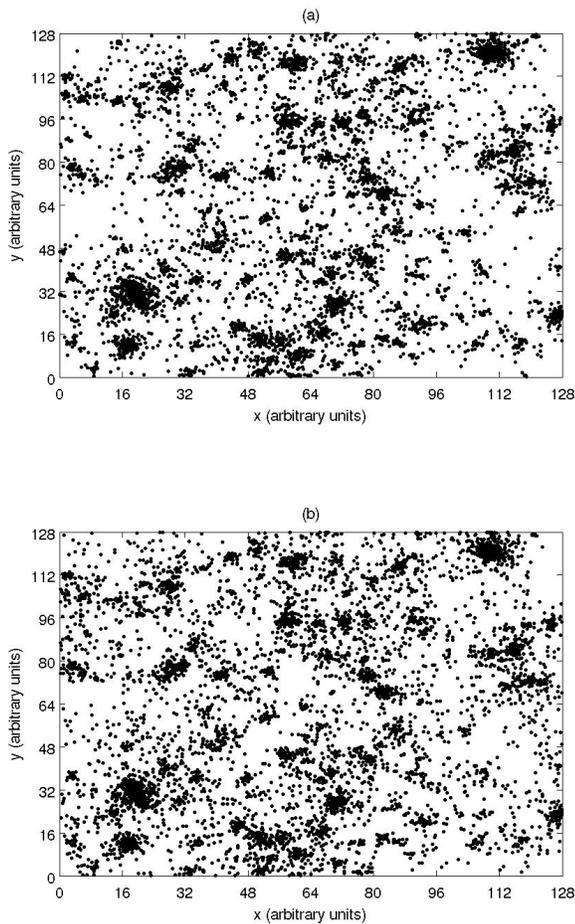


Figure 4.4: Spatial distribution of the (a) original clustered synthetic seismicity along with the (b) identified clustered seismicity by the TM metric.

is close to the original regional value for b of 0.901 ± 0.013 . This is expected given the use of (4.7) as a criteria to extract events from each box, which introduces a magnitude dependence to the method as a constraint that the clustered seismicity is imposed to. The lack of such dependence might lead to unrealistic results. For example, Wu (2010) presented a declustering method based on a hidden Markov model that does not take into account the magnitudes of events, and the clustered seismicity

identified in Japan yielded a low b -value around 0.68.

The main difference in the distributions plotted in Figures 4.3a and 4.3c occurs for events with magnitudes between 4 and 5. More events in this range are observed for the clustered seismicity identified by the TM metric, leading to a higher mean magnitude in this case. Aki's estimation of the b -value is inversely proportional to the mean magnitude of the events, and thus results in a lower estimation for the clustered seismicity as identified by the TM metric. The converse explains the higher estimation of the b -value for the declustered seismicity compared to the true declustered seismicity: a lower mean magnitude for the declustered seismicity identified by the TM metric yields a higher estimation of the b -value. In part this also is a result of the fact that the TM method is identifying moderate-sized events as clustered because they generate aftershocks, while the original ETAS algorithm designates them as part of the random process of initial background earthquake generation. In actuality, these events are related to the clustered nature of seismicity and the algorithm is designating them correctly in this context.

The spatial distribution of the clustered catalog as identified by this technique is compared to the original clustered portions of the synthetic catalog in Figures 4.4a and 4.4b, respectively. They display the same features for the true and derived clustered portions of the catalog and their differences arise from the inability to distinguish between events in the same box and the simplicity of the criteria used to extract events that only takes into account their magnitudes. The method presented here consists of two decision-making processes: establishing how many events must be removed from each box so that σ^2 becomes linear in time using a minimization

method, the PSO in this case, and then choosing which events from each box to be removed using a given criteria, equation (4.7) in this case. The PSO allows for up to 7,990 of the 9,332 true clustered events to be identified, meaning that 1,396 of the events were wrongly assigned for removal by the PSO, but the use of the probabilistic criteria (4.7) yielded the correct identification of 7,272 clustered events. The difference was the higher number of events with magnitudes between 4M and 5M and the lower number of events 5M or greater identified as clustered when comparing Figures 4.3a and 4.3c.

The proposed declustering method was applied to 100 different synthetic catalogs comprised of 20,000 events each in order to better understand its efficiency. The PSO allowed for up to 91% of the true clustered events to be correctly identified on average, meaning that 9% of the events were poorly assigned for removal by the PSO. However, the employment of the probabilistic criteria (4.7) yielded a lower number of correct clustered events identified, or 84% on average. The efficiency of the proposed method depends on the criteria in (4.7), but the trade-off between its simplicity and the efficiency loss was considered advantageous in this case.

4.4.2 Southern California

Southern California is an ideal candidate for the declustering method presented given the quality of the available data and the past declustering studies done in the region (Gardner and Knopoff, 1974; Reasenber, 1985; Zaliapin et al., 2008; Wang et al., 2010). A similar dataset to the one used in Tiampo et al. (2007) was subjected to the TM declustering technique and the results were qualitatively compared to the ones

obtained by Reasenbergs declustering method. The parameters used in Reasenbergs method were the same listed in Helmstetter et al. (2006) to a similar region in southern California. They were obtained so that the declustered portion of the seismicity was characterized as a Poisson process. The dataset considered was compiled from both the Advanced National Seismic System (ANSS) and Northern California Seismic Network (NCSN) and it consists of events $3M$ or greater recorded between 1933 and 2006, distributed between latitudes 32° and 40° north and longitudes 115° and 125° west.

Figure 4.5a illustrates the spatial distribution of events in the region of interest with magnitude $3M$ or greater along with their rate of occurrence on an yearly base. This is the same region of interest of previous studies on seismicity patterns (Rundle et al., 2002; Tiampo et al., 2002) that required a complete catalog. It is clear from Figure 4.5b that the number of events recorded since the 1970's is considerably higher than previously. Thus, the dataset was divided into events recorded before 1971 and events recorded from 1972 onwards and the declustering using the TM metric was performed separately. This division was chosen to avoid the clustered seismicity from the 1971 San Fernando earthquake in the estimation of the reference line for the second time period considered.

In both cases, the region was divided into a mesh of $0.1^\circ \times 0.1^\circ$ boxes and yearly time steps were considered to construct the time series $n_j(t)$, as in Tiampo et al. (2007). The different variances of $n_j(t)$ are plotted in Figure 4.6a and 4.6b for recordings before and after 1971, respectively. The discontinuities in the variances of $n_j(t)$ for the entire catalog in 1952, 1971, 1979, 1992 and 1999 agree with the occurrence

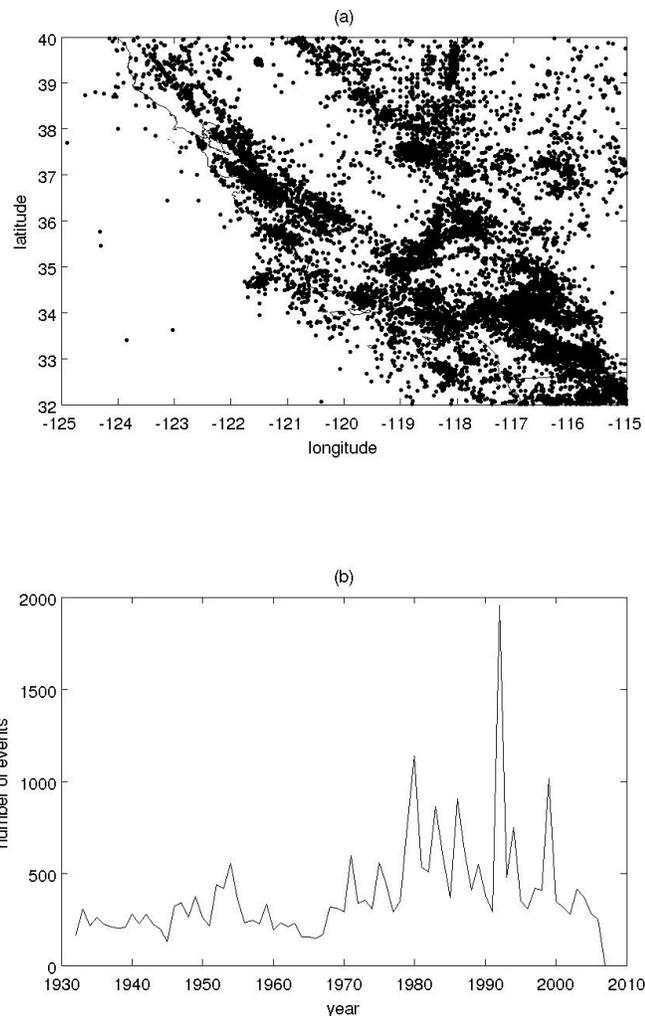


Figure 4.5: (a) Spatial distribution and (b) rate of occurrence of events recorded with magnitudes $3M$ or greater from 1933 to 2006 in southern California.

of large earthquakes in the region as expected: the 1952 Kern County, the 1971 San Fernando, the 1979 Imperial Valley, the 1992 Landers and the 1999 Hector Mine earthquakes. Other deviations from the linear trend are observed in the 1940's are likely due to the initiation of recordings in the northwest portion of the region.

The proposed method was applied to these two time periods and the variances of $n_j(t)$ for the declustered seismicity did collapse to the reference lines in both cases, as

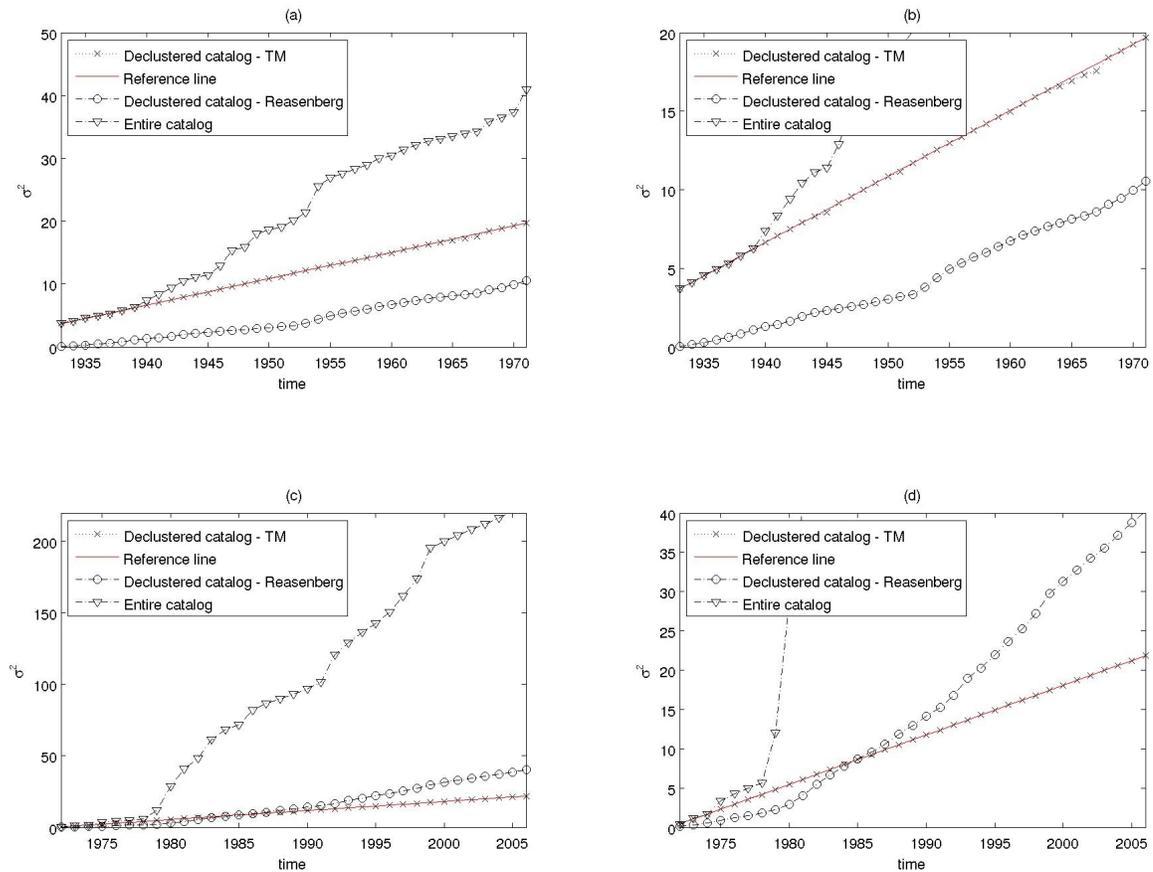


Figure 4.6: Plot of the reference line to perform the declustering and the variances of $n_j(t)$ for the original catalog and the declustered seismicity obtained from the TM and Reasenberg's method from southern California for (a) events recorded from 1933 to 1971 and (b) an enlargement of the variances for the declustered catalogs. The same plots for events recorded in the same region from 1972 onwards are shown in (c) and (d), respectively.

indicated in Figures 4.6a for 1933 to 1971 and 4.6c for 1972 to 2006. A closer look at the variances of the declustered events obtained by Reasenberg's method in Figures 4.6b and 4.6d show that some fluctuations are still present in both time periods that make σ^2 not quite linear, indicating that some clustering is still present. Jones and Hauksson (1997) employed Reasenberg's declustering method to a smaller region in

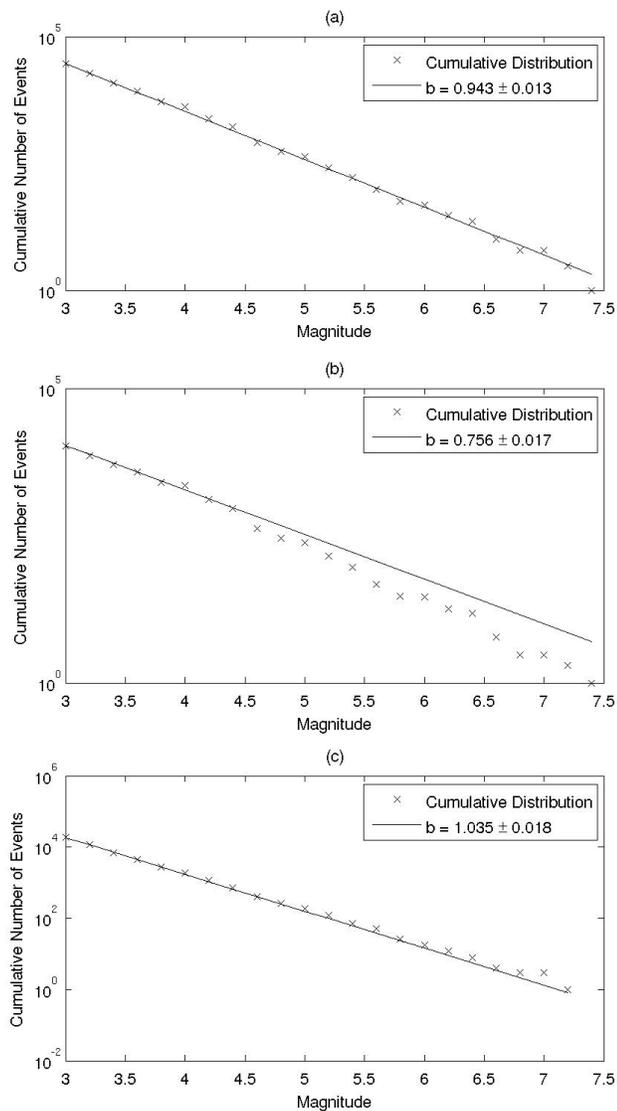


Figure 4.7: Frequency-magnitude distribution for the entire catalog of southern California for (a) the entire period, (b) 1933-1971 and (c) 1972-2006.

southern California and temporal clusters in the declustered catalog were observed. It was argued that this temporal clustering was due to the existence of seismic cycles.

A vertical offset is observed between σ^2 of the declustered catalogs obtained by the proposed and Reasenbergs methods in Figures 4.6b and 4.6d. The reason is the need to use the initial portion of recorded seismicity to determine the reference line in the

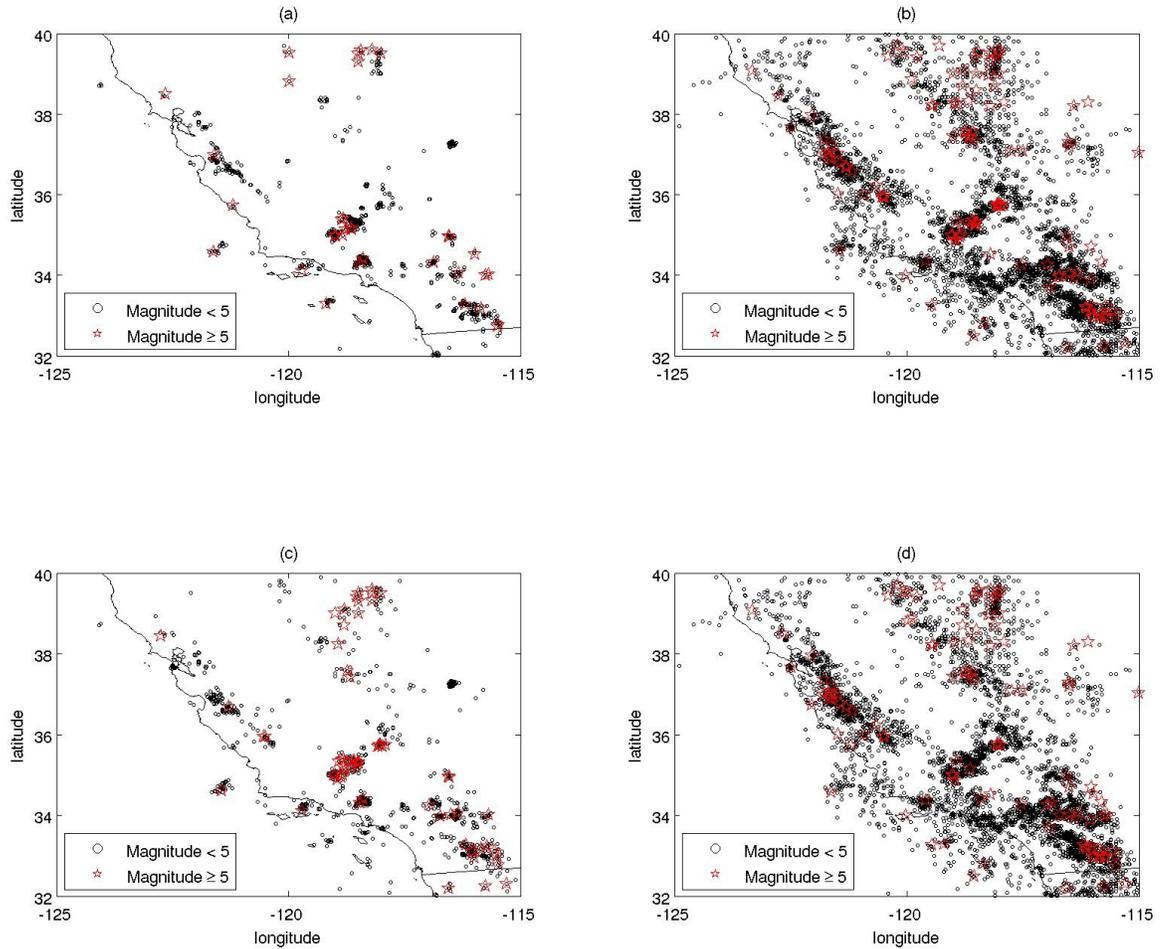


Figure 4.8: Spatial distribution of (a) clustered and (b) declustered events identified by the TM metric along with (c) clustered and (d) declustered events identified by Reasenber's method for events recorded in the 1940-1971 period.

method presented here. As a result, the declustering takes place during the period in which the linear trend of σ^2 is extrapolated, in this case 1940-1971 and 1975-2006. In Reasenber's method, no training period is needed and hence the observed vertical shift. In addition, the primary departure from a linear trend observed in Figure 4.6d in the late 1970's may be attributed to the changes in the network coverage, indicating that more events are being classified as declustered since the 1980's.

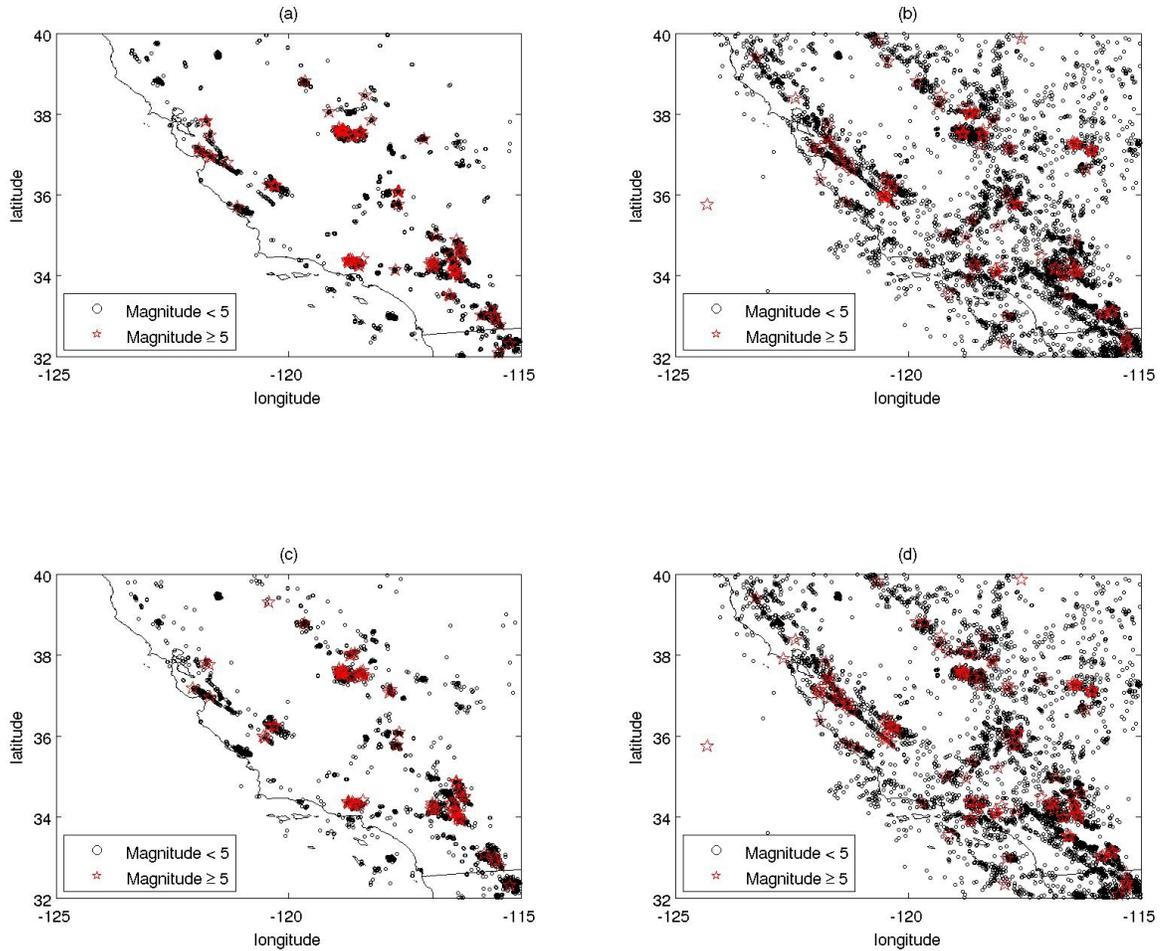


Figure 4.9: Spatial distribution of (a) clustered and (b) declustered events identified by the TM metric along with (c) clustered and (d) declustered events identified by Reasenber's method for events recorded in the 1975-2006 period.

Given the division of the original catalog motivated by the change in the seismic network coverage of the region, the removal of events was performed accordingly. Figure 4.7a displays the frequency-magnitude distribution for the entire catalog and the b -values for the probability distribution in equation (4.7) were calculated for the distributions in Figures 4.7b and 4.7c for events recorded before and after 1971. The two sets of results were then concatenated into a single representation so that

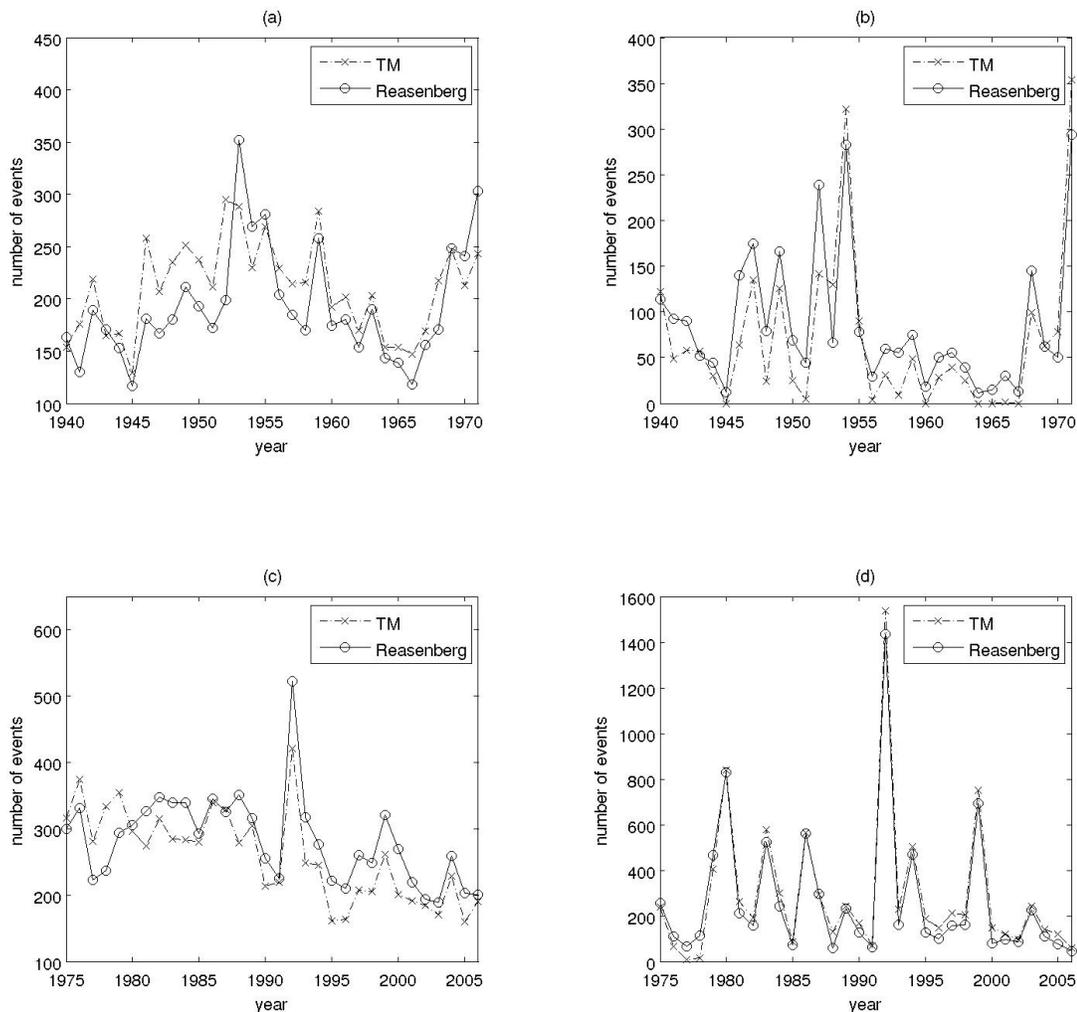


Figure 4.10: Plot of seismicity rates for the (a) declustered and (b) clustered seismicity identified by the method based on the TM metric and Reasenberg's algorithm for the period 1940-1971. The same plots for the identified (c) declustered and (d) clustered events for the 1975-2006 period from both methods.

qualitative comparisons with the results obtained from Reasenberg's method could be made. In order for these comparisons to be unbiased, events in the periods 1933-1939 and 1972-1975 that were used to extrapolate σ^2 were excluded from all subsets of events. The performance of both methods was very similar overall, especially for the

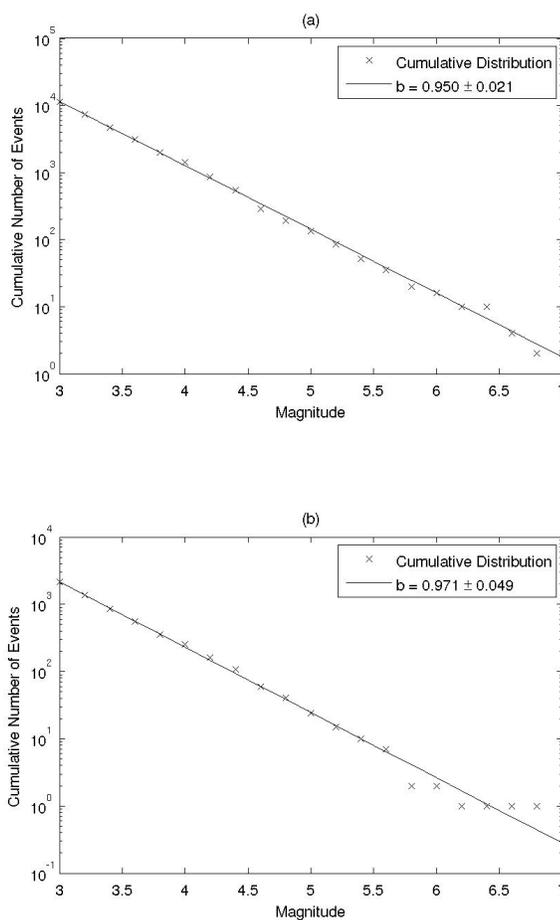


Figure 4.11: Frequency-magnitude distribution for (a) all clustered events identified by the TM method from 1940 until 1971 and 1975 until 2006 and (b) aftershock sequences extracted from the latter for the 1992 Landers earthquake. A square area of $1.25^\circ \times 1.25^\circ$ and events recorded in the year of the mainshock and the following year were considered.

period from 1975 until 2005 when the coverage of the seismic network was improved.

The spatial distribution of epicenters of the clustered and declustered seismicity identified by both methods from 1940 to 1971 and 1975 onwards are plotted in Figures 4.8 and 4.9, respectively. A total number of clustered events of 11,318 and 11,154 were

identified by the TM and Reasenbergs' methods for 1940-1971 and 1975 onwards, respectively. An overall good agreement between the distribution of the clustered events is observed in Figures 4.8a and 4.8c for the first considered period. The main difference is in the distribution of clustered events with magnitude 5M or greater in the northern and southern parts of the map, which might be attributed to the different seismicity rate in northern California and the poorer network coverage in the period. The TM method is dependent on a homogeneous catalog and a uniform declustered seismicity rate. In Figures 4.8 and 4.9 the TM declustering method appears to isolate clusters more effectively in space, with fewer random background events included in the clustered portion. With the improvement of coverage in the 1970s, the distribution of clustered seismicity identified by both methods becomes very similar, even for events with magnitudes greater than or equal to 5M with some isolated exceptions around the considered region, as illustrated in Figures 4.9a and 4.9c for the 1975-2006 period. These differences in the distribution of large events may be attributed to the probabilistic criteria (4.7) employed in the TM declustering method.

The calculated C_v for the declustered portions illustrated in Figures 4.8 and 4.9 were 1.30 and 1.16 for the TM and Reasenbergs' methods for the 1940-1971 period, respectively. For the 1975-2006 period, the declustered catalogs yielded C_v values of 1.27 for the TM method and 1.10 for Reasenbergs'. These differences were expected given the design of both methods: the parameters used in Reasenbergs' method were optimized to yield a Poissonian declustered catalog, whereas no consideration regarding the interoccurrence times between declustered events is made in the proposed

method.

Figure 4.10 displays the seismicity rates of the clustered and declustered portions identified by both methods for the periods 1940-1971 and 1975-2006. Throughout 1940-1971 and the initial years of 1975-2006, the TM method consistently identifies more declustered events, as illustrated in Figures 4.10a and 4.10c. This trend changes from the early 1980s onwards and coincides with the increase in σ^2 for Reasenbergs' declustered catalog, as expected and observed in Figure 4.6d. Figure 4.10b shows that a low number of clustered events is identified by both methods from the mid 1950s to the mid 1960s, suggesting the existence of a period of low seismic activity in the region. This drop in seismic activity was observed by Jones and Hauksson (1997) from August, 1952 until July, 1969 for a smaller region in southern California. Only 33 large events with magnitude 5M or greater were recorded during the periods 1955-1964. This is less than the average number of large events of 68 per 10 years recorded from 1935 to 1954 and from 1975 to 1994.

It is important to stress that any conclusions regarding the temporal clustering during this period should be evaluated with caution. Network coverage plays an important role in the number of aftershocks recorded in the 1950s and 1960s. The periods 1965-1974 and 1995-2005 can be used as example to illustrate this concern. In the first, 29 events with magnitude 5M or greater but the TM declustering method identified more than 500 in its first-half alone as illustrated in Figure 4.10b. From 1995 until 2005, a similar small number of large earthquakes was recorded in the region: 32 events. However, far more aftershocks were recorded during this 10-year period. The year 1999 alone displayed around 800 clustered events identified by both

methods due to the 1999 Hector Mine earthquake (Figure 4.10d).

It is possible to extend the investigation of individual aftershock sequences identified by the TM metric. Figure 4.11 displays the frequency-magnitude distribution of the clustered seismicity and the aftershock sequences of the 1992 Landers earthquake identified by the proposed TM method. For this sequence, seismicity recorded within a square region of $1.25^\circ \times 1.25^\circ$ around the mainshock was used as in Shcherbakov et al. (2005) for events recorded in 1992 and 1993. The frequency-magnitude distribution shown in Figure 4.11 for the 1992 Landers aftershock sequence appears to display a characteristic distribution (Schwartz and Coppersmith, 1984), because all events in the boxes where the mainshock occurred were classified as a clustered and were removed from the declustered dataset. The b -value obtained for this sequence is 0.97 ± 0.05 , which is higher than the value of 0.950 ± 0.021 for the clustered events, indicating that the low magnitude seismic activity in this sequence was considerably high.

4.4.3 Taiwan

The declustering method based on the TM metric is a statistical approach that does not depend on the geology or faulting characteristics of the system. The Taiwanese catalog poses as an interesting challenge for the proposed technique since it comprises events recorded in a subduction zone. Chen et al. (2006) analyzed seismicity patterns in Taiwan for a dataset of events recorded from 1987 to 1999 with magnitude 3.4M or greater between latitudes 19.4° and 26° north and longitudes 117.9° and 126.5° east. In the latter study, the region of interest was divided into a mesh of $0.1^\circ \times 0.1^\circ$ boxes

and the time series $n_j(t)$ were constructed on an yearly base. Figures 4.12 displays the spatial distribution of the events in the entire catalog obtained from the Central Weather Bureau of Taiwan.

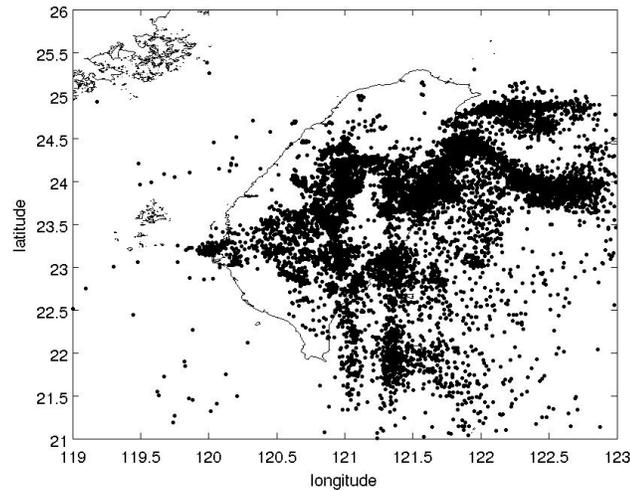


Figure 4.12: Spatial distribution of the entire Taiwanese catalog for events with magnitude $3.4M$ or greater recorded from 1987 to 2008.

The same discretization was employed here to perform the TM declustering technique to an extended dataset of the Taiwanese catalog, with recordings from 1987 to 2008. This catalog was also declustered using Reasenbergs method with the default parameters for California as performed by Chen and Wu (2006) and the results were compared. The variances of $n_j(t)$ for the entire catalog and the declustered portions along with the reference line are plotted in Figure 4.13a. The discontinuities in Figure 4.13a of σ^2 for the entire catalog observed in 1990, 1994, 1999 and 2002 are attributed to the clustering of events. The PSO established the long effective ergodic period for the Taiwanese catalog in Figure 4.13b as expected. A closer examination of σ^2 for the declustered catalogs obtained by both methods in Figure 4.13b show that the

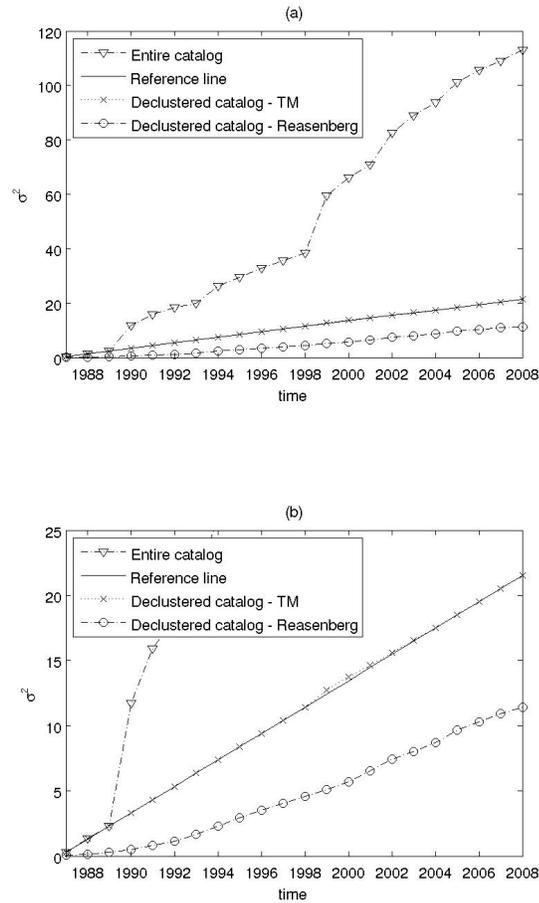


Figure 4.13: (a) Plot of the declustering reference line and the variances of $n_j(t)$ for the original catalog and the declustered seismicity obtained from the TM and Reasenberg's method for Taiwan, from 1987 to 2008. (b) Enlargement of the variances of $n_j(t)$ for the Taiwanese catalog along with its declustered portion and the reference line for the TM declustering method.

variance for the Reasenberg's declustered catalog can also be considered linear.

Figure 4.14 displays the seismicity rates of the entire catalog along with the declustered portions obtained from the TM and Reasenberg's method. The years of high seismic activity (1990, 1994, 1999 and 2002) correlate well with the discontinuities of

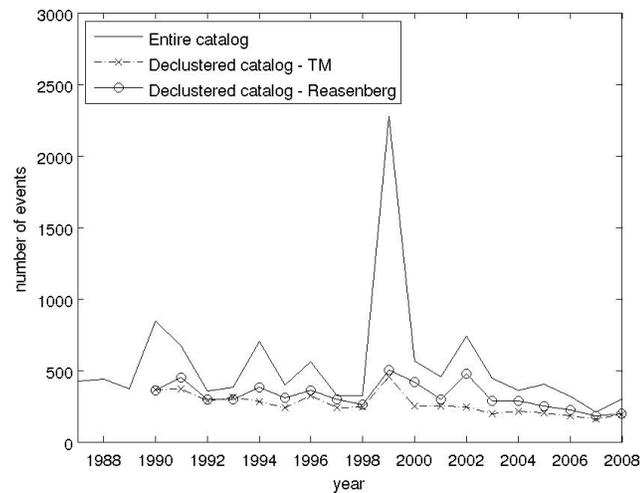


Figure 4.14: Yearly seismicity rate for the entire Taiwanese catalog from 1987 to 2008 along with declustered portions obtained from the TM and Reasenberg's method from 1990 to 2008.

σ^2 of the entire catalog in Figure 4.13b as expected. The activity in the year 1999 was considerably higher due to the Chi-Chi earthquake. The declustered portions in Figure 4.14 are more homogeneous throughout the considered time period for both methods, with similar seismicity rates up until 1999 and from 2003 onwards. From 2000 to 2002, the TM method identified more clustered events in 2000 and 2002 due to the 1999 Chi-Chi earthquake and an 6.8M event occurred in March, 31 of 2002, respectively.

The choice of the events to be removed from each box in each timestep was made using the criteria outlined in equation (4.7) and the b -value calculated for the frequency-magnitude distribution for the entire region is plotted in Figure 4.15a. The clustered portion of the seismicity plotted in Figure 4.15b displays a similar b -value to the one obtained for the entire catalog, whereas the declustered portion in Figure 4.15c displays a slightly lower b -value. For the 1999 Chi-Chi earthquake,

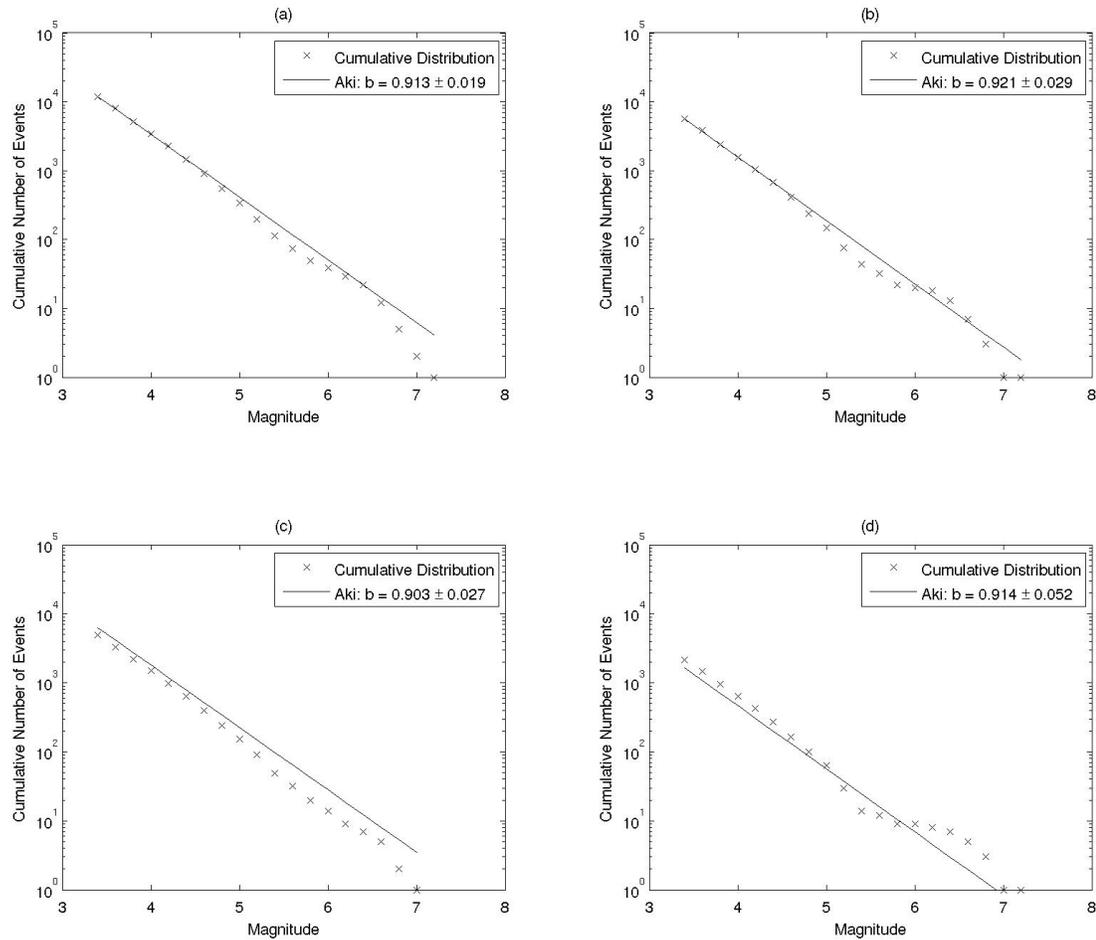


Figure 4.15: Frequency-magnitude distribution for (a) the original Taiwanese catalog from 1987 until 2008, (b) its clustered portion as identified by the TM metric from 1990 until 2008, (c) the declustered portion for the same period and (d) the seismicity associated with the 1999 Chi-chi earthquake. For the latter, the seismicity identified as clustered recorded in 1999 and 2000 within a $1^\circ \times 1^\circ$ square centered at the mainshock was considered in the aftershock sequence.

seismicity located in a $1^\circ \times 1^\circ$ area around the mainshock for the years 1999 and 2000 was considered. As observed for the aftershock sequences in southern California, the mainshock of the 1999 Chi-Chi earthquake was tagged as clustered due to the decision of the PSO to remove all events in the box where the mainshock was located. The

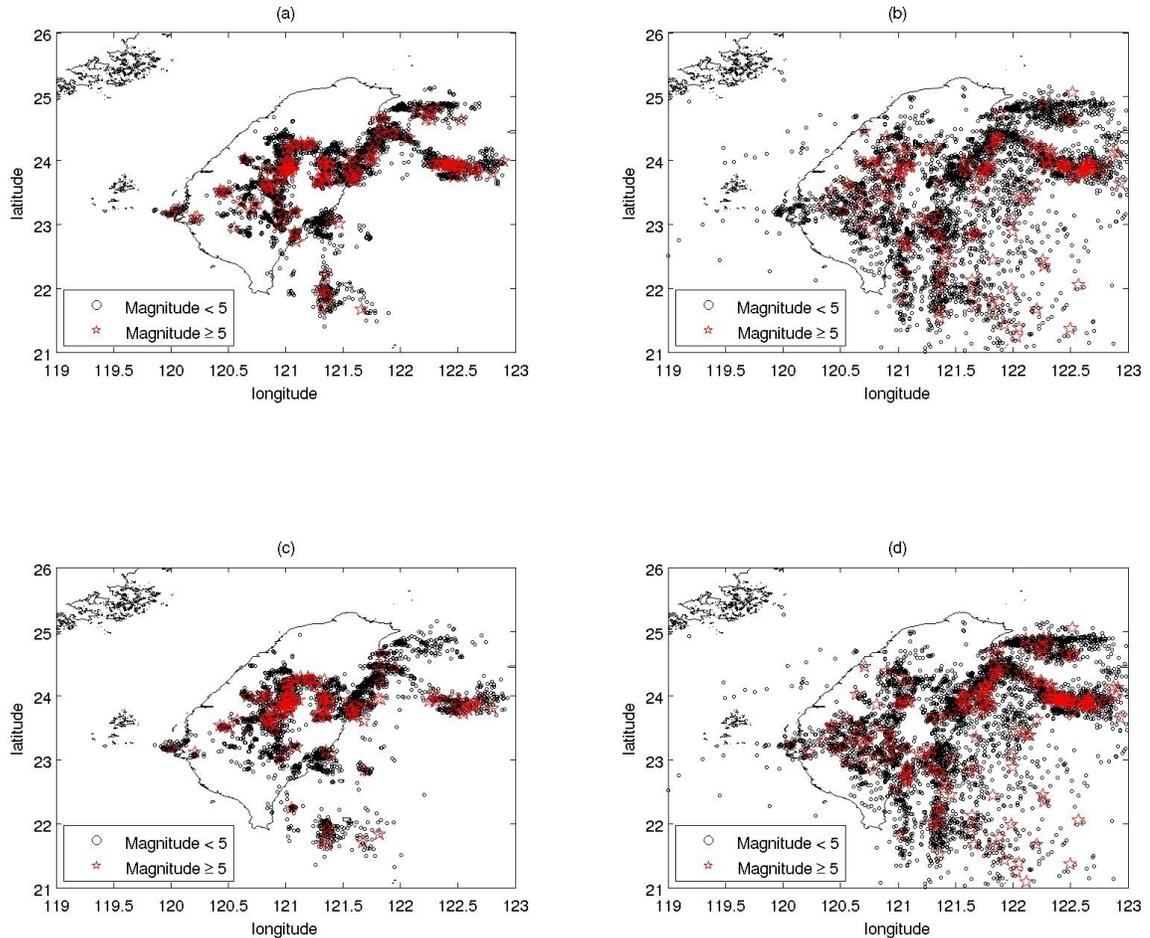


Figure 4.16: Spatial distribution of the (a) clustered and (b) declustered seismicity identified by the TM metric and (c) clustered and (d) declustered seismicity identified by Reasenber's method in Taiwan from 1990 to 2008.

result is the frequency-magnitude relationship plotted in Figure 4.15d, which follows the Gutenberg-Richter law.

Figure 4.16 displays the spatial distribution of the clustered and declustered portions of the catalog from 1990 to 2008 identified by both declustering techniques. Overall, the same features in both clustered distributions (Figures 4.16a and 4.16b) for the TM and Reasenber's methods are observed. In addition to the better spatial

isolation of clustered events in Figure 4.16a, the main difference is the lack of large events with magnitude 5M or greater identified as clustered by Reasenbergs method in the northeast coast in Figure 4.16b. Despite the many similarities in the spatial distribution of clustered events obtained by both methods, the total number of events identified as clustered was quite different. For the TM method, a total of 5,633 events were determined clustered whereas Reasenbergs method identified only 4,498 events as clustered. The latter constantly identified more declustered events over the years, especially during the 2000-2003 period. This difference likely is due to the choice of the standard parameters for California to perform this declustering, but a study of the optimal parameters for the Reasenbergs method in Taiwan falls outside the scope of this work.

The values for C_v obtained for the declustered seismicity obtained by the proposed TM method and Reasenbergs method were 1.36 and 1.19, respectively. Reasenbergs method in this case displayed an analogous performance to that of southern California. This may be due to fact that there is a concentration of high shear strain rates in both regions that results in larger mainshocks and more concentrated aftershock sequences.

4.4.4 Switzerland

Another interesting region with good quality data to employ the proposed declustering method is Switzerland. Nanjo et al. (2010) used a probabilistic method that estimated the magnitude of completeness of the Swiss catalog between 1983 and 2008 to be 2.5M in the far southwest of the country and lower than 1.6M in regions of high seismic activity. The moderate seismic activity in the region poses a challenge to the method

given the general requirement for a substantial number of recordings to apply the method and, as such, is a good test of the method.

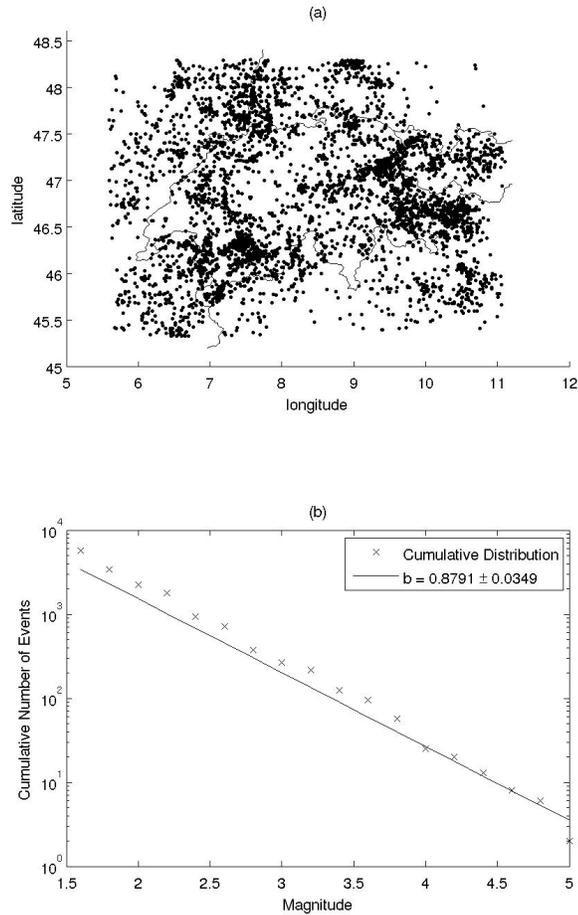


Figure 4.17: (a) Spatial distribution of the entire Swiss catalog for events with magnitude $1.6M$ or greater from 1983 until 2008. (b) The frequency-magnitude distribution for the entire Swiss catalog for the same time period.

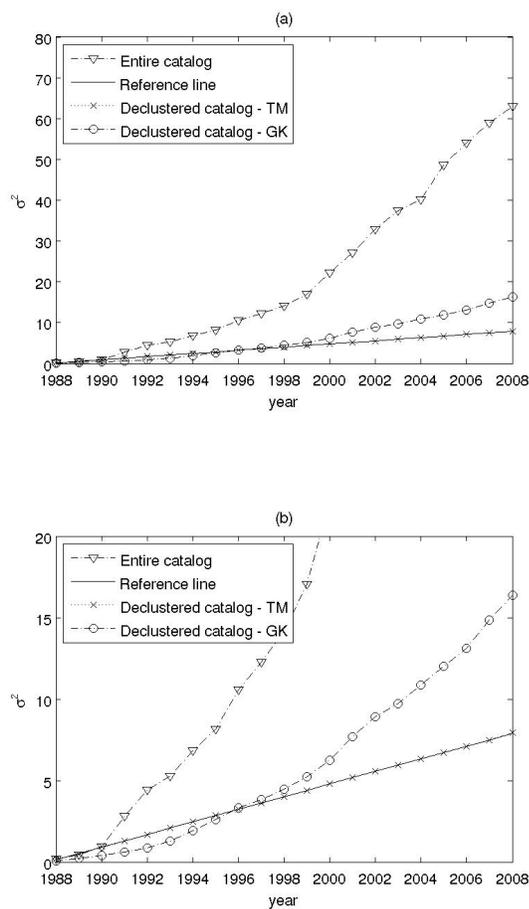


Figure 4.18: (a) Plot of the reference line to perform the declustering and the variances of $n_j(t)$ for the original catalog and the declustered seismicity obtained from the TM and the GK method from Switzerland for events recorded from 1988 to 2008 and (b) enlargement of the variances for the declustered catalogs.

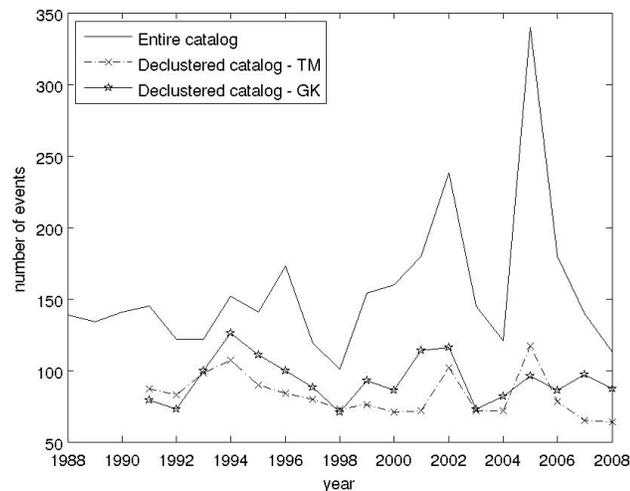


Figure 4.19: Seismicity rates for the entire Swiss catalog from 1988 until 2008 and the declustered seismicity identified by the TM metric and the GK method from 1991 until 2008.

A cutoff magnitude of $1.6M$ was used to seismicity recorded from 1988 to 2008 between latitudes 45° and 48.5° north and longitudes 5.2° and 11.5° east. This dataset was obtained from the Swiss Seismological Service and it is considered complete by the method described in Wiemer and Wyss (2000) to a 95% level of confidence. Figures 4.17a and 4.17b display the spatial distribution of epicenters and the frequency-magnitude distribution of the entire catalog, respectively. For comparison, the same dataset was declustered with the method described in Gardner and Knopoff (1974), optimized for central Europe as described in Wiemer et al. (2009). Their choice of the GK method was based on the high variability of the results obtained for Reasenbergs' method for different parameters.

A mesh of $0.1^\circ \times 0.1^\circ$ boxes and yearly timesteps were used to construct the time series $n_j(t)$ and the variances of $n_j(t)$ for the entire catalog and the two declustered

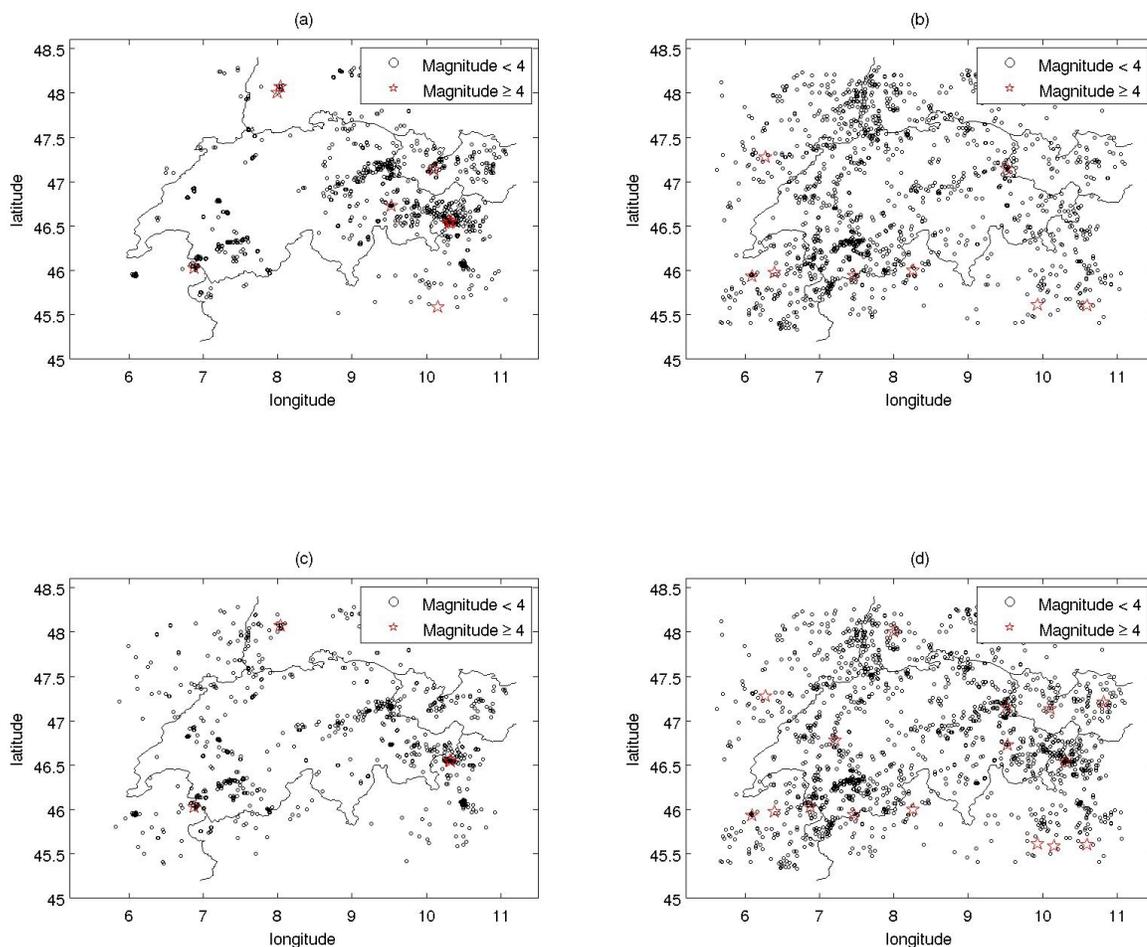


Figure 4.20: Spatial distribution of (a) clustered and (b) declustered events identified by the TM metric along with (c) clustered and (d) declustered events identified by the Gardner-Knopoff method for the Swiss catalog from 1991 until 2008.

portions are plotted in Figure 4.18a, with a close-up of the variances from the declustered portions in Figure 4.18b. There are no abrupt changes in σ^2 of the entire catalog in Figure 4.18a, but considerable changes in the slope are observed in the periods of 1999-2003 and 2003-2005. These changes correlate well with periods of high seismic activity in Figure 4.19, indicating that seismicity clustering is disrupting σ^2 of the entire catalog as expected. As for the declustered seismicity identified by the TM

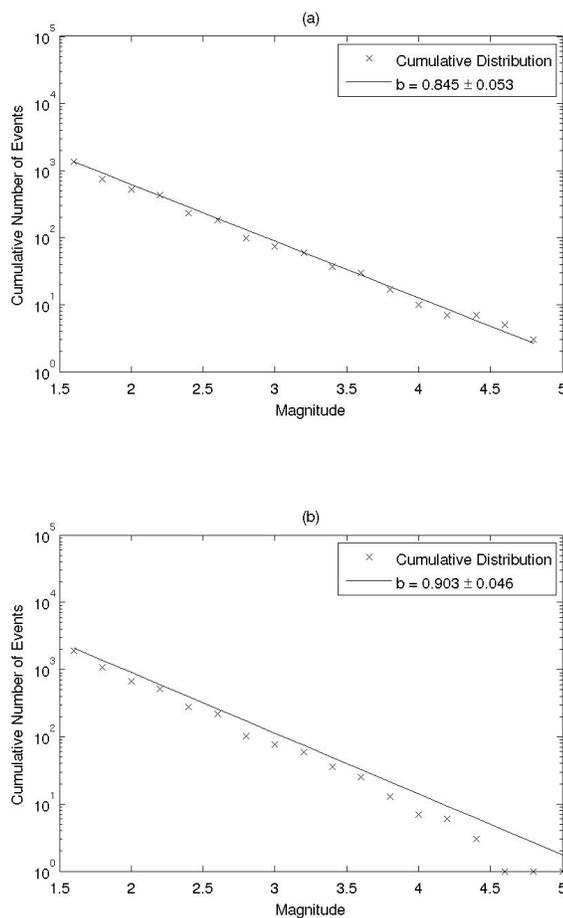


Figure 4.21: Frequency-magnitude distribution for the (a) clustered and (b) declustered portions of the Swiss catalog identified by the TM metric from 1991 until 2008.

metric, the PSO performed well in determining the number of events to be removed from the catalog to obtain a long linear trend for σ^2 as seen in Figure 4.18b. No linear trend is obtained for the declustered seismicity from the GK method, and an apparent scaling $\langle n_j(t)^2 \rangle \propto n_j(t)^2 \propto t^\alpha$ with $\alpha > 1$ is observed in this case. Previous results on the TM metric (Cho et al., 2010) showed that this is due to the presence of spatial clustering.

The seismicity rate of the Swiss catalog along with both declustered portions identified by the TM metric are plotted in Figure 4.19. It is clear that the total number of events considered in this case is less than the previous cases (Figures 4.1c, 4.10, and 4.14). The determination of the reference line could be biased given the high seismic activity during the 1983-1987 period, therefore the declustering method was applied from 1988 onwards. The rates for the declustered portions differ from each other, with an overall higher seismicity rate for the declustered portion obtained by the GK method. A total of 1,678 declustered events were obtained by the GK method and 1,491 by the TM method.

The frequency-magnitude distribution of the region plotted in Figure 4.17b was used to choose the events to be removed from each box. Figure 4.20 displays the spatial distributions of the clustered and declustered events identified by both methods from 1990 onwards. In this case, events with magnitude $4M$ or larger are considered large, given the moderate seismic activity in the region. There are only a few large events in the catalog and the TM metric identifies more of these as clustered. A simple visual inspection shows that the clustered seismicity obtained by the TM metric in Figure 4.20c again is more localized than the clustered seismicity in Figure 4.20a. Also, the declustered sets in 4.20b and 4.20d indicate that the GK method identifies more than double the number of declustered events (264 versus 115 events) in the region delimited by longitudes 9.5° and 11° and latitudes 46° and 47° . This contributes to the spatial clustering remaining in the declustered portion of the catalog identified by the GK method in Figure 4.18b and is likely due to the heterogeneity in the coverage of the Swiss seismic network (Nanjo et al., 2010).

The values for C_v obtained for the declustered seismicity identified by the TM and the GK methods were of 1.25 and 1.05, respectively. This region differs considerably from both Taiwan and southern California regarding the distribution of shear strain rates (Kreemer et al., 2003), displaying a more homogeneous rate and less intense shear strain concentration. This illustrates the difference between the performances of the TM method and the classical declustering methods (Reasenber and the GK). Both were originally applied to California and their parameters are not fitted for this region. The first displays high variability with respect to the chosen parameters while, for the second it is necessary to use parameters optimized specifically for central Europe (Wiemer et al., 2009).

Both the distributions for the clustered and declustered seismicity identified by the TM metric in Figures 4.21a and 4.21b follow the Gutenberg-Richter law as expected. The low number of larger magnitude clustered events in Figures 4.21a is due to the removal of all seismicity in boxes with large events and as seen for the individual aftershock sequences in southern California.

4.4.5 Southern Spain and northern Africa

The region displays complicated and diffuse seismic activity, comprised of moderate seismicity driven by the convergence of the Eurasian and African plates. This region displays a similar shear strain rate distribution regime as seen for Switzerland, with low and diffused shear strain rates (Kreemer et al., 2003). A dataset obtained from the Spanish National Geographic Institute (IGN) from 1990 to 2010 was considered here. The completeness of this catalog was addressed in Gómez (2009) and it was

determined that from 1990 onwards, the magnitude of completeness for southern Spain is approximately $3M_b$. Seismicity recorded between latitudes 40° and 32° north and longitudes 20° west and 6° east and depths less than 60km was declustered using the proposed declustering method and Reasenberg's method. For the latter, the standard parameters for California were employed for declustering.

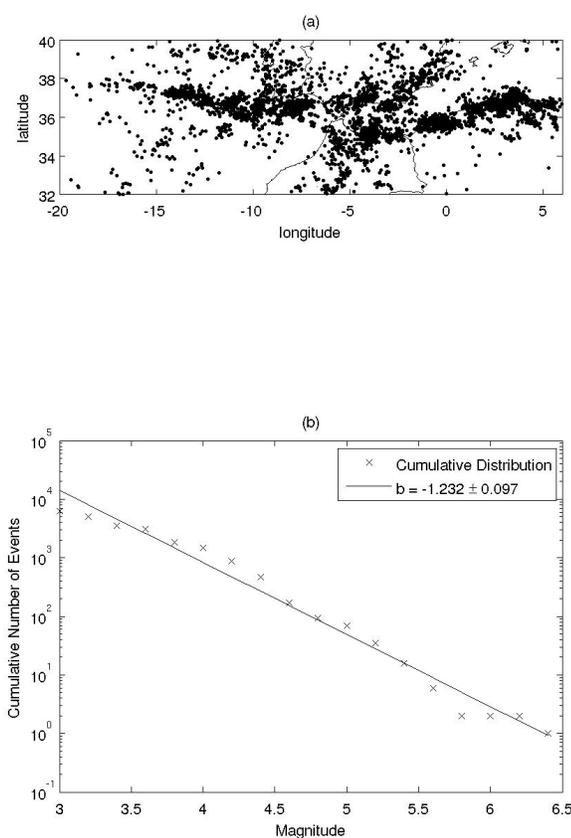


Figure 4.22: (a) Spatial distribution of the entire Spanish catalog for events with magnitude $3M_b$ or greater from 1990 until 2010. (b) The frequency-magnitude distribution for the entire Spanish catalog for the same time period.

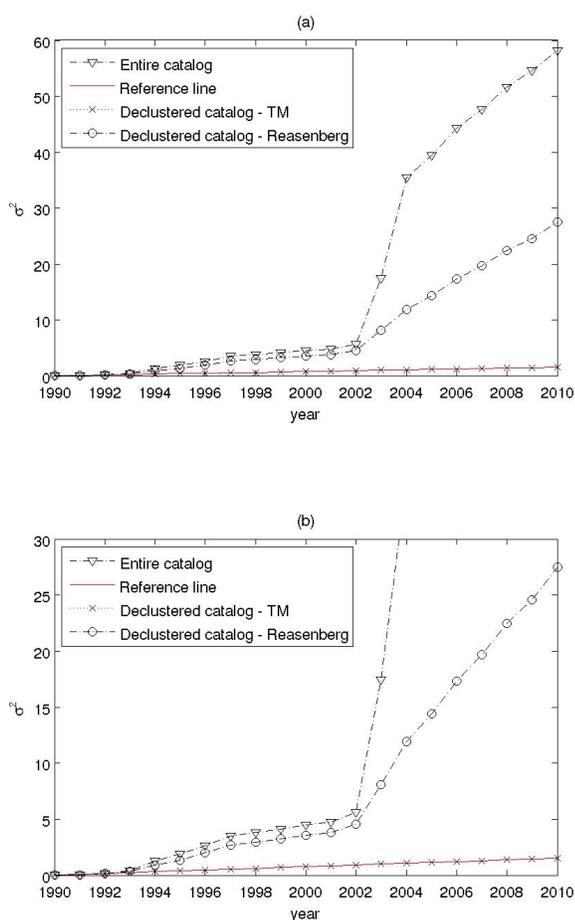


Figure 4.23: (a) Plot of the reference line to perform the declustering and the variances of $n_j(t)$ for the original catalog and the declustered seismicity obtained from the TM and Reasenberg's method from Spain for events recorded from 1990 to 2010 and (b) enlargement of the variances for the declustered catalogs.

Figure 4.22a displays the spatial distribution of the seismicity recorded during the time period considered. The region also encompasses the northern parts of Morocco and Algeria. Most seismicity is distributed in the northern tip of Africa and in the coast south and southwest from Portugal, with a smaller fraction in the southeast of

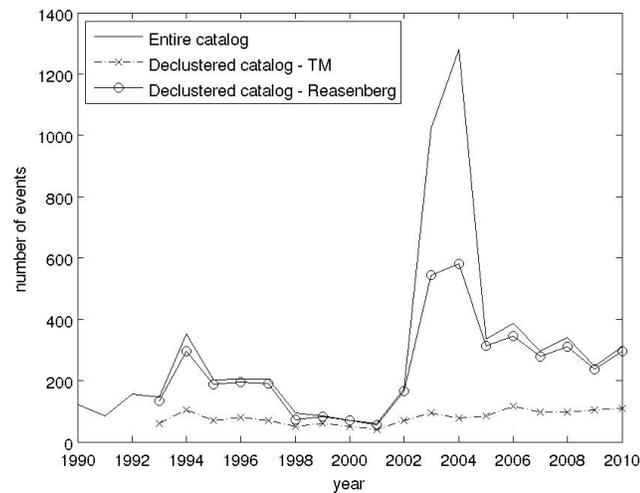


Figure 4.24: Seismicity rates for the entire Spanish catalog from 1990 until 2010 and the declustered seismicity identified by the TM metric and Reasenberg's methods from 1993 until 2010.

Spain. The frequency-magnitude distribution for the region considered is plotted in Figure 4.22b, and the b -value of 1.232 ± 0.097 was obtained from a least squares linear fit. The least square estimation of the b -value was employed to the dataset provided by IGN for a different time period (Jimenez et al., 2006) and a different seismic catalog maintained by the Andaluz Institute of Geophysics (IAG) (Stich et al., 2007) and a b -values of 1.24 was obtained by the latter in the period 1986-2006 for events $3.6M_b$ or greater.

To perform the proposed declustering technique, the discretization of Tiampo et al. (2007) for the entire Iberian Peninsula, comprised of yearly timesteps and a mesh of $0.1^\circ \times 0.1^\circ$ boxes to construct the time series $n_j(t)$ was employed. This spatial grid is the same as that of Gómez (2009) for the study of the completeness of the catalog. The variances of $n_j(t)$ for the entire catalog and the declustered portions of the seismicity

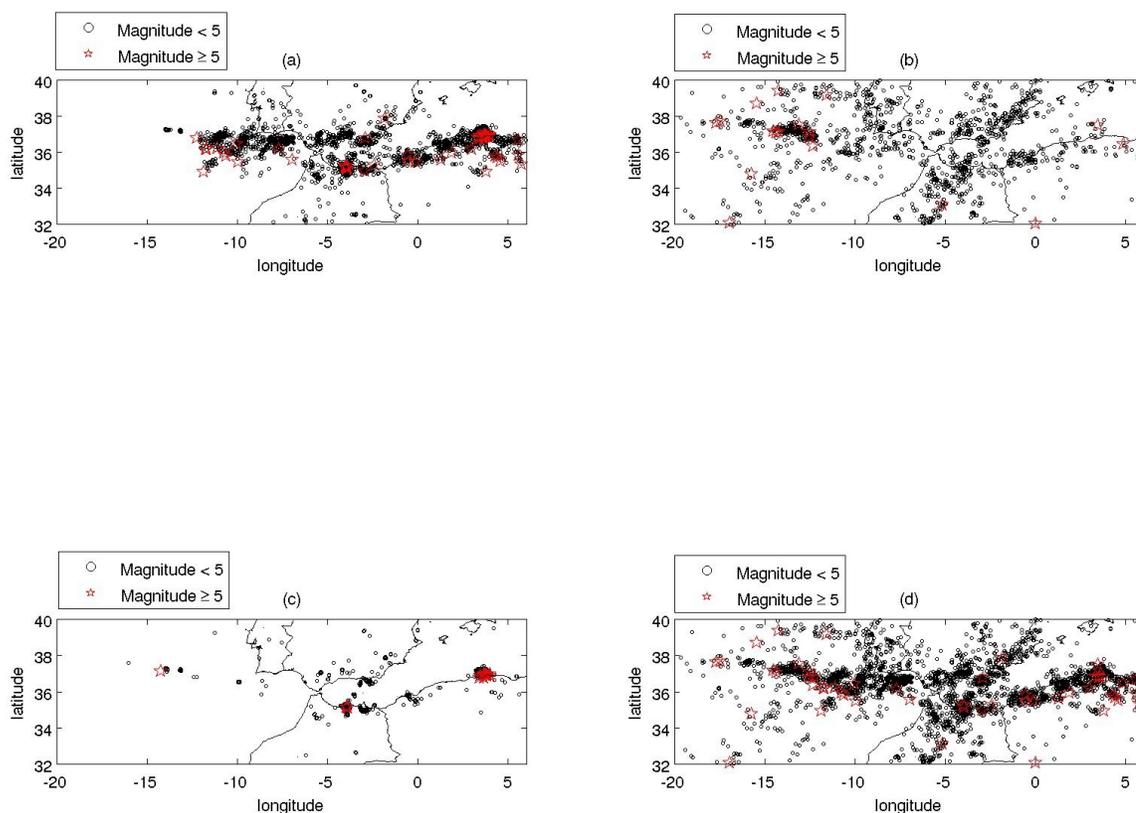


Figure 4.25: Spatial distribution of (a) clustered and (b) declustered events identified by the TM metric along with (c) clustered and (d) declustered events identified by Resenberg's method from 1993 until 2010.

identified by both methods are plotted in Figure 4.23a, with an enlargement of the variances of the declustered portions in Figure 4.23b. A long effective ergodic period for the declustered seismicity was obtained by the TM method in Figure 4.23, as expected. The same was not observed for the declustered seismicity identified by Resenberg's method, and the overall trend of σ^2 in this case is the same as σ^2 for the entire catalog until 2003. This is an indication that Resenberg's method did not perform well in southern Spain with the standard parameters for California. The

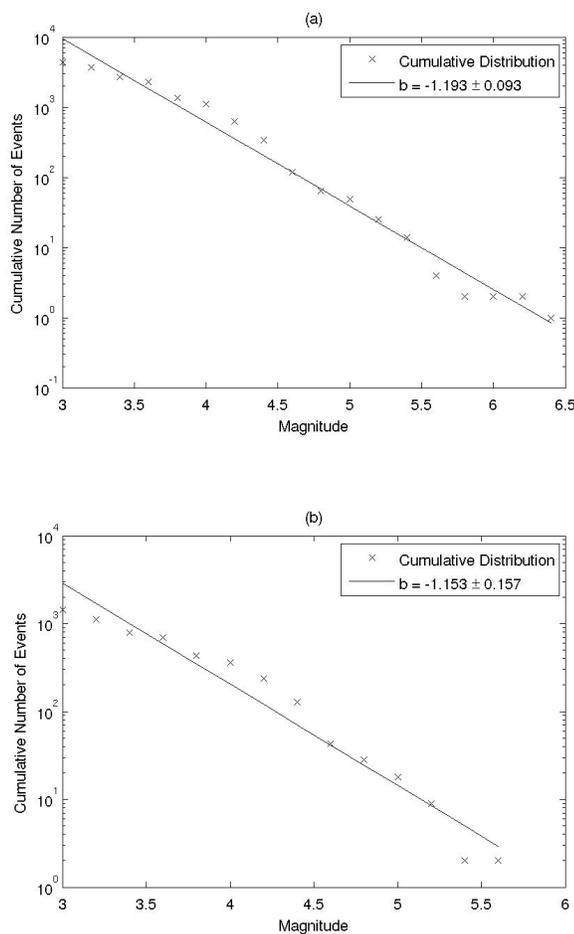


Figure 4.26: Frequency-magnitude distribution for the (a) clustered and (b) declustered portions of the Spanish catalog identified by the TM metric from 1993 until 2010.

first departure of σ^2 of the entire catalog from the linear trend extrapolated from the initial stationary period occurred in 1994, followed by a second in 2003/2004. These discontinuities were due to large events in the region: two events $5.7M_b$ in the north of Morocco and Algeria in 1994, the $6.6M_b$ 2003 Boumerdès earthquake, and the $6.2M_b$ 2004 Al-hoceima earthquake.

The seismicity rate for the entire catalog plotted in Figure 4.24 corroborates the

observations above. The two mainshocks in 2004 and 2003 generated a large peak in seismic activity in the region, and a small activity peak is observed in 1994. A quiescent period with a few clustered events is observed prior to the large events in northern Africa from 1998 until 2001, and an increase in the average number of events per year is observed since 2000. The latter is due to the installation of wide band seismometers in the region (Gómez, 2009). The seismicity rate of the declustered seismicity identified by Reasenberg's method shows that most of the events were considered declustered, whereas the declustered seismicity obtained by the TM metric is stationary over the entire period considered.

From the regional b -value obtained in Figure 4.22b, the events to be removed from each box were chosen based on criteria (4.7). A total of 4,380 out of the original 6,180 events were classified as clustered by the TM method, compared to the 1,461 clustered events identified by Reasenberg's method. This striking difference may be attributed to the differences in shear strain rates between southern Spain and California (Kreemer et al., 2003): the parameters used in California take into account clustered seismicity that is much more concentrated than that of regions of lower shear strain rates, and only a few events are classified as clustered in this case. This would also explain the difficulties mentioned by Wiemer et al. (2009) regarding the use of this classical declustering techniques for Switzerland.

Figure 4.25 illustrates the spatial distribution of clustered and declustered seismicity obtained by both methods. Most of the clustered seismicity identified by the TM metric in Figure 4.25a is located along northern Africa, where the large-magnitude seismicity occurs more frequently, in the ocean south of Portugal where a set of faults

are presented, and in the southeast of Spain. Its counterpart, the declustered seismicity, in Figure 4.25b is more sparse, as expected, with the exception of a cluster southwest of Portugal that was generated from the constant occurrence of isolated events around Gloria fault. An interesting feature in Figure 4.25b is the transverse trend across the map, from the southeast of Spain through Morocco.

The combination of the distribution of epicenters for the declustered seismicity identified by Reasenbergs method in Figure 4.25d and its seismicity rate in Figure 4.24 confirms that the method does not perform well for this region, and optimum parameters must be found to ensure a better performance of this technique. Most of the clustered seismicity identified this classical declustering method in Figure 4.25c comes from the 2003 Boumerdès and the 2004 Al-hoceima earthquakes.

The frequency-magnitude distributions for the clustered and declustered portions of the catalog identified by the TM metric are plotted in Figures 4.26a and 4.26b, respectively. The b -value obtained for the clustered seismicity was of 1.193 ± 0.093 , which is very close to the regional value as expected. The lower b -value obtained for the declustered portion of the catalog indicates that a lower proportion of small events is found in this case.

The values for C_v obtained for the declustered seismicity obtained by the proposed TM method and Reasenbergs were 1.24 and 1.58, respectively. In this case, Reasenbergs method appears to perform less effectively than in southern California, incorrectly identifying clustered events as declustered, likely due to the differences in the seismic regime in this region. As mentioned previously, an optimization of the parameters has to be employed to ensure a more accurate identification of clustering

in this region with Reasenbergs method.

4.5 Conclusion

A new and simple declustering technique is proposed based on the TM metric, and its premise is to remove clustered events from a catalog so that the remaining seismicity displays a long effective ergodic period. No information regarding the geology of the region or a large number of parameters is needed for the performance of this method, and its performance is shown to be robust for different and varied datasets. A system of particles is created from the spatial and temporal discretization of the region of interest and the cumulative number of events in each box $n_j(t)$ is used as a proxy for the seismic-released energy. The number of events to be removed from each box is determined by an optimization method, the PSO in this case, and the regional frequency-magnitude distribution then is used to determine which events are to be excluded from each box. In this framework, the evolution of $n_j(t)$ for the declustered catalog is considered as a normal diffusive process.

The method was initially used to decluster a synthetic catalog generated from the ETAS model and then four locations that display different tectonic settings: southern California, Taiwan, Switzerland and a region comprised of southern Spain and northern Africa. The good performance achieved for the declustering of the synthetic catalog was a positive indicator of the efficiency of the method and long periods of effective ergodicity were found for all the cases of actual, historic seismicity. The proposed method was applied consistently, with no change in its application once the

spatial discretization and the cutoff magnitude of the catalog were determined. The regions considered display different distributions of shear strain rates, and these characteristics appear to influence the performance of the classical declustering methods considered for comparison purposes. The performance of the method was consistent regardless of the region under consideration for a homogeneous dataset, while the performance of the classical methods depended on the parameters used for the different regions.

Applying the probability density function that yielded the regional Gutenberg-Richter law as a constraint for the clustered seismicity proved to be simple criteria that enabled the incorporation of the magnitude of events into the proposed declustering method. The frequency-magnitude distributions for the clustered seismicity identified by the proposed method in all the cases followed the Gutenberg-Richter law with b -values close to unity. The individual aftershock sequences displayed the same agreement with the Gutenberg-Richter law, and the discrepancies observed were caused by the inclusion of large-magnitude events located in boxes where all events were identified as clustered by the PSO. The results obtained by the proposed method agreed reasonably well with other established declustering methods for southern California and Taiwan, but differed for Switzerland, most likely due to the heterogeneity of the catalog, and Spain, due to the application of Reasenbergs method with parameters used for California. These last two analyses illustrate the dependence of these classical declustering methods on their parameters.

The declustered seismicity identified by the proposed method displayed constant values of C_v that deviate from the unity value expected for Poisson processes. This

topic has to be approached with caution given the fact that the proposed method does not take into account the interoccurrence times between neither the clustered or declustered events, unlike the classical methods used here. An attempt to rescue a part of the physics of the region under consideration was made with the introduction of an exponential probability distribution to the magnitudes of the clustered events. A logical second step to this method would be the incorporation of aspects of the physics of interoccurrence times of the clustered events.

Despite its simplicity, a series of factors had to be taken into account to ensure a proper employment of the proposed declustering method. The extrapolation of the initial linear trend of σ^2 was essential, and a period of stationary seismic activity must be considered in order to obtain an accurate reference line. For regional seismicity, homogeneous coverage in the region of interest is needed to perform the declustering and account for changes in seismic network coverage. In addition the completeness of the catalog also is an important issue because a sufficiently large record is required to validate the statistical approach. Further improvements to the proposed method include the improvement of the efficiency of the optimization method and the employment of a more elaborated criteria to identify events for extraction.

Acknowledgements

This research was supported by the NSERC & AON Benfield/ICLR Chair in Earthquake Hazard assessment. The authors would also like to thank R. Naryshkin and P. Gonzalez for their valuable comments, as well as the Central Weather Bureau of

Taiwan for the Taiwanese earthquake catalog.

References

- Aki, K. (1965). Maximum likelihood estimate of the b in the formula $\log n = a - bm$ and its confidence limits. *Bulletin of the Seismological Society of America*, 43:237–239.
- Baiesi, M. and Paczuski, M. (2004). Scale-free networks of earthquakes and aftershocks. *Physical Review E*, 69(066106).
- Bellman, R. (1972). *Dynamic programming*. Princeton University Press.
- Bezdek, J. C. (1981). *Pattern recognition with Fuzzy objective function algorithms*. Plenum Press, New York.
- Bishop, C. M. (1995). *Neural Networks for pattern recognition*. Clarendon, Oxford, UK.
- Chen, C. C., Rundle, J. B., Holliday, J. R., Nanjo, K. Z., Turcotte, D. L., Li, S. C., and Tiampo, K. F. (2006). From tornadoes to earthquakes: forecast verification for binary events applied to the 1999 chi-chi, taiwan, earthquake. *Terrestrial, Atmospheric and Oceanic Sciences*, 17(3):503–516.
- Chen, C. C. and Wu, Y. X. (2006). An improved region-time-length algorithm applied to the 1999 chi-chi, taiwan earthquake. *Geophysical Journal International*, 166:1144–1147.
- Cho, N. F., Tiampo, K. F., McKinnon, S. D., Vallejos, J. A., Klein, W., and

- Dominguez, R. (2010). A simple metric to quantify seismicity clustering. *Non-linear Processes in Geophysics*, 17:293–302.
- Crick, F. (1989). The recent excitement about neural networks. *Nature*, 337:129–132.
- Dobrovolsky, I. P., Zubkov, S. I., and Miachkin, V. I. (1979). Estimation of the size of earthquake preparation zones. *Pure and Applied Geophysics*, 117:1025–1044.
- Dunn, J. C. (1973). A fuzzy relative of the isodata process and its use in detecting compact well-separated clusters. *Journal of Cybernetics*, 3:32–57.
- Gardner, J. K. and Knopoff, L. (1974). Is the sequence of earthquakes in southern california, with aftershocks removed, poissonian? *Bulletin of the Seismological Society of America*, 64(5):1363–1367.
- Gómez, A. G. (2009). Mapa de localizaciones probables de futuros terremotos en la península ibérica, baleares y canarias. *Seguridad y Medio Ambiente*, 114:44–54.
- Hartigan, J. A. (1975). *Clustering algorithms*. John Wiley and Sons, New York.
- Helmstetter, A., Kagan, Y. Y., and Jackson, D. D. (2006). Comparison of short-term and time independent earthquakes forecast models for southern california. *Bulletin of the Seismological Society of America*, 96(1):90–106.
- Helmstetter, A. and Sornette, D. (2002). Subcritical and supercritical regimes in epidemic models of earthquake aftershocks. *Journal of Geophysical Research*, 107(B10):1–21.

- Holland, J. (1992). *Adaptation in natural and artificial systems : an introductory analysis with applications to biology, control, and artificial intelligence*. MIT Press, Cambridge, Mass.
- Huang, Z. (1998). Extensions to the k-means algorithm for clustering large data sets with categorical values. *Data Mining and Knowledge Discovery*, 2(2):283–304.
- Jain, A. K. and Dubes, R. C. (1988). *Algorithms for clustering data*. Prentice Hall, Wood Cliffs, NJ.
- Jimenez, A., Tiampo, K. F., Levin, S., and Posadas, A. M. (2006). Testing the persistence in earthquake catalogs: the iberian peninsula. *Europhysics Letters*, 73(2):171–177.
- Jones, L. M. and Hauksson, E. (1997). The seismic cycle in southern california: precursor or response? *Geophysical Research Letters*, 24(4):469–472.
- Kagan, Y. Y. and Jackson, D. D. (1991). Long-term earthquake clustering. *Geophysical Journal International*, 104:117–133.
- Kennedy, J. and Eberhart, R. (1995). Particle swarm optimization. *Proceedings of IEEE International Conference on Neural Networks*, IV:1942–1948.
- Kreemer, C., Holt, W. E., and Haines, A. J. (2003). An integrated global model of present-day plate motions and plate boundary deformation. *Geophysical Journal International*, 154:8–35.
- MacQueen, J. B. (1967). Some methods for classification and analysis of multivariate.

- Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability*, pages 281–297.
- Marsan, D. and Lengliné, O. (2008). Extending earthquakes’ reach through cascading. *Science*, 319:1076–1079.
- Mendonza, C. and Hartzell, S. H. (1988). Aftershock patterns and mainshock faulting. *Bulletin of the Seismological Society of America*, 78(4):1438–1449.
- Nanjo, K. Z., Schorlemmer, D., Woessner, J., Wiemer, S., and Giardini, D. (2010). Earthquake detection capability of the swiss seismic network. *Geophysical Journal International*, 181(3):1713–1724.
- Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point process. *Journal of the American Statistical Association*, 83(401):9–27.
- Omori, F. (1894). On the aftershocks of earthquakes. *The Journal of the College of Science, Imperial University of Tokyo*, 7:111–200.
- Pedersen, M. E. H. (2010). Good parameters for particle swarm optimization. *Technical Report HL1001 (Hvass Laboratories)*.
- Reasenber, P. (1985). Second-order moment of central california seismicity, 1969–1982. *Journal of Geophysical Research*, 90(B7):5479–5495.
- Rundle, J. B., Tiampo, K. F., Klein, W., and Martins, J. S. S. (2002). Self-organization in leaky threshold systems: The influence of near-mean field dynamics and its implications for earthquakes, neurobiology, and forecasting. *Proceedings of*

- the National Academy of Sciences of the United States of America*, 99(Supplement 1):2514–2521.
- Schwartz, D. P. and Coppersmith, K. J. (1984). Fault behavior and characteristic earthquakes: examples from the wasatch and san andreas fault zones. *Journal of Geophysical Research*, 89(B7):5681–5698.
- Shcherbakov, R., Yakovlev, G., Turcotte, D. L., and Rundle, J. B. (2005). Model for the distribution of aftershock interoccurrence times. *Physical Review Letters*, 95(218501).
- Stein, R. S. (1999). The role of stress transfer in earthquake occurrence. *Nature*, 402:605–609.
- Stich, D., Martín, J. B., and Morales, J. (2007). Deformación sísmica y asísmica en la zona béticas-rifalborán. *Revista de la Sociedad Geológica de España*, 20(3-4):311–319.
- Thirumalai, D. and Mountain, R. D. (1993). Activated dynamics, loss of ergodicity, and transport in supercooled liquids. *Physical Review E*, 47:479–489.
- Thirumalai, D., Mountain, R. D., and Kirkpatrick, T. R. (1989). Ergodic behavior in supercooled liquids and in glasses. *Physical Review A*, 39:3563–3574.
- Tiampo, K. F., Rundle, J. B., Klein, W., Holliday, J. B., Martins, J. S. S., and Ferguson, C. D. (2007). Ergodicity in natural earthquake fault networks. *Physical Review E*, 75(066107):1–15.

- Tiampo, K. F., Rundle, J. B., McGinnis, S., and Klein, W. (2002). Pattern dynamics and forecast methods in seismically active regions. *Pure and Applied Geophysics*, 159(10):2429–2467.
- Tinti, S. and Mulargia, F. (1985). Effect of magnitude uncertainties on estimating the parameters in the gutenbergrichter frequency-magnitude law. *Bulletin of the Seismological Society of America*, 75(6):1681–1697.
- Turcotte, D. L. and Holliday, J. R. (2007). Bass, an alternative to etas. *Geophysical Research Letters*, 34(L12303).
- Utsu, T. (2002). *Statistical features of seismicity. International Handbook of Earthquake and Engineering Seismology, Part A*. Academic Press, Amsterdam.
- Wang, Q., Jackson, D. D., and Zhuang, J. (2010). Are spontaneous earthquakes stationary in california? *Journal of Geophysical Research*, 115(B08310):1–12.
- Wiemer, S. (2001). A software package to analyze seismicity: Zmap. *Seismological Research Letters*, 72(3):373–382.
- Wiemer, S., Giardini, D., Fäh, D., Deichmann, N., and Sellami, S. (2009). Probabilistic seismic hazard assessment of switzerland: best estimates and uncertainties. *Journal of Seismology*, 13(4):449–478.
- Wiemer, S. and Wyss, M. (2000). Minimum magnitude of completeness in earthquake catalogs: examples from alaska, the western united states, and japan. *Bulletin of the Seismological Society of America*, 90(4):859–869.

Wu, Z. (2010). A hidden markov model for earthquake declustering. *Journal of Geophysical Research*, 115(B03306):24513–24528.

Zaliapin, I., Gabrielov, A., Keilis-Borok, V., and Wong, H. (2008). Clustering analysis of seismicity and aftershock identification. *Physical Review Letters*, 101(018501).

Chapter 5

Effects of location errors in the Pattern

Informatics⁴

5.1 Introduction

Earthquakes are a significant natural hazard in many inhabited areas, and the determination of both the spatial and temporal locations prone to the occurrence of large events has become a topic of great interest over the years (Keilis-Borok, 2005; Kanamori, 2003). The seismic hazard map for the United States (Petersen et al., 2008) and the working group on Regional Earthquake Likelihood Models (RELM) are two recent examples of these efforts (Field, 2007). A series of different phenomena, including but not limited to electromagnetic activity, gas emissions, and/or surface deformation have been evaluated over the years as precursory signals to accomplish this goal, with limited success (Cicerone et al., 2009).

Among the potential precursory signals are seismicity patterns that might occur prior to a strong earthquake. Kanamori (1981) identified different precursory seismicity patterns based on different large events distributed worldwide, but many of these focused on patterns in moderate-to-large events. However, methods such as these have been limited statistically by their relatively small numbers. The Pattern Informatics (PI) is a technique used in intermediate-to-long regional seismic hazard assessment that rely on seismic patterns in small-to-moderate magnitude events as a

⁴submitted to Pure and Applied Geophysics

precursory signal for larger earthquakes (Rundle et al., 2002; Tiampo et al., 2002). This method therefore utilizes a large number of seismic events in a phase dynamics framework that uses rotations of a phase function to quantify anomalies in seismic activity. From a statistical standpoint, the PI offers a solid comparison between long-term background seismicity rates and local variations in order to identify periods of activation and/or quiescence. This technique has been used worldwide (Chen et al., 2005; Nanjo et al., 2005) and it is also a part of the RELM working group.

When applying statistical methods to study earthquakes as a point-source phenomena, events are characterized by their occurrence time, location and magnitude. These estimations are subjected to errors that may affect the outcomes of different statistical analyzes. An attempt to address this issue in the PI was performed in Nanjo et al. (2005) by smoothing seismicity rates of a particular region with the activity in its vicinity. van Stiphout et al. (2011) analyzed the effects of different sources of error in the estimation of background seismicity levels, commonly used in hazard assessment (eg. Petersen et al. (2008) and Wiemer et al. (2009)). It was found that hypocentral errors influenced these estimations, but the choice of the declustering method had the greatest impact on the estimation of the background seismicity rate. A different approach to address the same issue was employed by Werner and Sornette (2008) to determined the deviation of estimations of seismicity rates due to magnitude uncertainties in a simple clustering model.

The use of large datasets in the PI is an important feature with respect to fluctuations introduced by epicenter locations uncertainties. If errors are considered independent and identically distributed, the large number of events would tend to

decrease the effects of location errors in the PI. This is because of the Central Limit Theorem, that reduces these variabilities. In this paper, the effects of locations errors in the performance of the PI are analyzed. This is accomplished by generating different scenarios from the introduction of normally distributed noise with different standard deviations randomly distributed over the epicenters of events recorded in California. The PI method is applied to the various datasets and its performance for retrospective forecasts is quantified by means of Skill Scores (SS) commonly used in atmospheric sciences (Wilks, 1995; Woodcock, 1976). This simple analysis is extended to investigation of the role of the number of events in the catalog by randomly removing events from the original dataset before the introduction of the noise as described previously.

The next section introduces the PI along with the two SS employed to quantify the performance of the latter and the TM metric, used to quantify seismicity clustering. The process of generating the perturbed catalogs is also described in Section 5.2. The effects of the introduction of noise in the locations being discussed in Section 5.3. Finally, the implications of the process are discussed in section 5.4.

5.2 Methodology

5.2.1 The Pattern Informatics

The PI (Rundle et al., 2002; Tiampo et al., 2002) is a method used to offer intermediate-to-long term seismic hazard assessment by quantifying local changes in seismicity patterns in order to identify instances in space and time characterized by seismic ac-

tivation or quiescence. These variations are considered a proxy for local stress changes which precede large events (Tiampo et al., 2006). The PI method has been used in different regions in the globe such as central Japan (Nanjo et al., 2005), southern California (Holliday et al., 2005) and Taiwan (Chen et al., 2005).

The computation of a PI hazard map is performed by initially obtaining a complete seismic catalog from a region of interest with recordings from an initial time t_0 until t_2 . A mesh is then created to divide the location considered into a set of N_b boxes and time series $n_j(t)$ of the cumulative number of events in each box i up until a certain time t are computed. The seismic intensity in a box i between t and a time base $t_b \in [t_0, t_2]$ can be calculated as

$$I_i(t_b, t) = \frac{n_i(t_b, t)}{t - t_b + 1}. \quad (5.1)$$

Seismic intensities given by equation (5.1) are then normalized for comparison purposes by subtracting the averages μ over all boxes of I_i and dividing the result by the variance

$$\hat{I}_i(t_b, t) = \frac{I_i(t_b, t) - \mu}{\sigma}. \quad (5.2)$$

A base time $t_1 \in [t_0, t_2]$ is chosen so that a training period Δt from t_1 until t_2 is considered. Equation (5.1) is evaluated for t_1 and t_2 to determine the anomalous seismicity $\Delta I_i(t_b, t_1, t_2)$ in box i during Δt

$$\Delta I_i(t_b, t_1, t_2) = \hat{I}_i(t_b, t_2) - \hat{I}_i(t_b, t_1). \quad (5.3)$$

The sum of equation (5.3) over different values t_b is performed to reduce the random variability in the measurement of anomalous seismic activity. The probability

of a future event of magnitude M_f or greater to occur in box i during the period from t_2 to $t_2 + \Delta t$ is then given as

$$P_i \propto \left(\sum_{t_b} \Delta I_i(t_b, t_1, t_2) \right)^2. \quad (5.4)$$

For a given threshold D , a binary forecast can be issued for the forecast period from t_2 until $t_2 + \Delta t$. Events with magnitude M_f or greater are considered most likely to occur on or adjacent to these hotspots: locations where P_i greater than D . Holliday et al. (2005) hypothesized that $M_f = M_c + 2$ for southern California, where M_c is the magnitude of completeness of the catalog.

5.2.2 The Heidke and Pierce Skill Scores

The results of binary-type forecasts can be represented in contingency tables, as illustrated in table 5.1. In this case, a , b , c and d represent correct forecasts, false alarms, misses and correct negatives respectively, and $a+b+c+d = N_b$. A series of skill scores have been developed over the years to quantify the efficiency of binary forecasts based on contingency tables. These methods, commonly used in atmospheric sciences, rely on the comparison between forecasts and thus are intrinsically incomplete (Wilks, 1995; Woodcock, 1976).

One of these quantitative measures is the Heidke skill score (HSS) (Heidke, 1926). It compares the considered forecast to a random one and it is calculated as

$$HSS = \frac{2(ad - bc)}{(a + c)(c + d) + (a + b)(b + d)}, \quad (5.5)$$

where $HSS \in [-1, 1]$. A perfect and a random forecast yield $HSS = 1$ and $HSS = 0$ respectively. Negative scores mean that the forecast considered perform worse than

		Observed	
		yes	no
Forecasts	yes	a	b
	no	c	d

Table 5.1: Example of a contingency table: a , b , c and d represent correct forecasts, false alarms, misses and correct negatives respectively

random guessing.

A second score is the Pierce skill score (PSS), which uses the hit rate $H = \frac{a}{a+c}$ and the false alarm rate $F = \frac{b}{b+d}$ of the considered forecast. The first is the fraction of events successfully forecasted, whereas the second is the fraction of false alarms given that an event did not occur. The PSS is given by:

$$PSS = H - F = \frac{ad - bc}{(a + c)(b + d)}, \quad (5.6)$$

where $PSS \in [-1, 1]$ and the same interpretation for the HSS can be made for the PSS.

5.2.3 The TM Metric and Clustering

In classical statistical physics, the ergodic hypothesis offers a simple link to relate micro and macro states of a system. Thirumalai et al. (1989) studied the effective ergodicity for liquid and glasses using a simple metric defined as

$$\Omega_G(t) = \frac{1}{N_b} \sum_{i=1}^{N_b} [g_i(t) - \langle g \rangle]^2 \quad (5.7)$$

where N_b is the number of particles in the system considered, $g_i(t) = \frac{1}{t} \int_0^t G_i(t') dt'$ is the time-average of a phase-space function $G_i(t)$ and $\langle g \rangle = \frac{1}{N_b} \sum_i g_i(t)$ is its ensemble

average. The term effective addresses the fact that long but finite time intervals are considered in which the phase space is sampled evenly (Thirumalai and Mountain, 1993).

Tiampo et al. (2007) applied the TM metric to earthquake fault systems by dividing a region of interest and using the number of events above a certain magnitude as a proxy for $G_i(t)$: the seismic released energy. In this framework, $\int_0^t G_i(t') dt' \equiv n_i(t)$, where $n_i(t)$ is the cumulative number of events in each box, and equation (5.7) can then be written as

$$\Omega_G(t) = \frac{1}{t^2} [\langle n_j(t)^2 \rangle - n_j(t)^2], \quad (5.8)$$

where the angular brackets stand for ensemble averages.

Effective ergodic periods are identified as instances of time when the inverse TM metric is linear in time, $\frac{D_e}{\Omega_G(t)} = t$, with a positive slope. The parameter $D_e = \Omega_G(t_0)$ is a diffusion coefficient that is related to the rate at which the phase space is sampled. During these periods, the system is in a state of metastable equilibrium that is disrupted by large earthquakes (Tiampo et al., 2007). Cho et al. (2010) established that the evolution of $n_j(t)$ can be seen as a diffusive process during these periods, when the occurrence of events is considered declustered, and deviations of the inverse TM metric from a linear trend occur as a result of the clustering of events in space and/or time.

The first term in the right-hand side of equation (5.8) is sensitive to both spatial and temporal seismicity clustering whereas the second term is sensitive to temporal clustering only. For a fixed number of events, equation (5.8) displays its extreme

value when all events are concentrated in a single box (maximum) and all events are evenly distributed over the boxes (minimum). In this case, D_e can be also interpreted as a coefficient that scales the TM metric with respect to the initial conditions of the system regarding its clustering.

5.2.4 Generation of Perturbed Catalogs

An original dataset comprised of events recorded by the Advanced National Seismic System (ANSS) and Northern California Seismic Network (NCSN) of events recorded between latitudes 32° and 40° north and longitudes -115° and -125° west from an initial time $t_0 = 1932$ until 2006 was considered. A magnitude of completeness $M_c = 3$ was employed, which means that the forecast magnitude in this case is $M_f = 5$. A mesh of boxes with edge length $\Delta x = 0.1^\circ$ and yearly timesteps is used to generate the time series $n_j(t)$. This is the same space/time discretization used to perform the PI in California and several other regions (Chen et al., 2005; Holliday et al., 2005).

Errors in the estimation of epicenter locations in southern California vary over the years due to changes in the network coverage and the use of digitalized waveforms since the 1970s (Hutton and Yu, 2008). The average accuracy in the horizontal component from 1932 until 2006 is $\approx 2.5\text{km}$ and it drops to $\approx 0.7\text{km}$ from 1970 onwards (Perlock, 2009). The perturbed datasets were obtained by adding noise to epicenter locations from the original southern Californian catalog. This noise was considered uniformly distributed around the epicenters and normally distributed with zero mean and different standard deviations σ_n in the radial direction. The lowest $\sigma_n = 0.007^\circ$ was chosen to be the average error of $\approx 0.7\text{km}$ for events recorded from

1970 onwards. The remaining $\sigma_n = 0.05^\circ$, 0.1° and 0.2° were chosen to take into account higher levels of noise in the analysis and they were chosen based on Δx .

The effects of a lower number of events was also considered. Initially, N_r events were randomly removed from the original southern California dataset so that the remaining events would constitute a spatially unperturbed, but smaller, catalog. The same procedure of adding noise to epicenter locations, as detailed above, was applied to the events in this cropped catalog to generate the perturbed scenarios. Different $N_r = 17500$, 20000 and 25000 were considered, and a total of $N_r = 22500$ events were removed from the original recordings from 1932 until 2006 in this particular case. Furthermore, noise levels $\sigma_n = 0.007^\circ$, 0.1° and 0.2° were considered in this analysis.

5.3 Results

The PI was performed for two different training periods: one from $t_1 = 1981$ until $t_2 = 1991$ (Δt_1) and another from $t_1 = 1989$ until $t_2 = 1999$ (Δt_2). These Δt periods yielded forecasts from 1992 until 2002 and from 2000 until 2010, respectively. A total of 55 large events (set E_1) were recorded in the first period. For the second forecast period, the number of large events dropped to 29 (set E_2). Figure 5.1 illustrates the resulting PI hazard map obtained from the original and different perturbed catalog for Δt_1 along with the set of events to be forecasted E_1 . Both the HSS and PSS were used to quantify the effectiveness of the retrospective forecasts for Δt_1 and Δt_2 . The computation of the scores was performed for all active boxes in the time period

considered from t_0 until $t_2 + \Delta t$. The same criteria used in Holliday et al. (2005) to issue a successful forecast was used: if one of the targeted events is located in a hotspot or in one of its immediate neighbors, the forecast is considered correct.

The PI scores obtained from equation (5.4) are a measurement of seismic activation and quiescence for a given box i , and this quantity is related to the seismicity clustering in the system. Figure 5.2 is a plot of the inverse TM metric of the original catalog as well as the perturbed ones for different values of σ_n . In order to compare the degree of clustering of the different catalogs using the TM metric, D_e was set to unity. The level of seismicity clustering decreases with the increase of the noise levels added to the locations of the epicenters, represented by the direct correlation between the inverse TM metric and σ_n in Figure 5.2, as expected. Despite this constant decrease of seismicity clustering with the increase of noise levels, the main features of the inverse TM plot remain the same. The latter points towards persistence of clustering patterns despite the addition of noise to event locations, although these cluster patterns become more diffuse, as seen in Figure 5.1.

This invariance of the seismicity clustering patterns indicate that the effectiveness of the PI should not be drastically affected by the addition of perturbation to the epicenter locations, due to its intrinsic dependence on seismic patterns for forecasting. The Central Limit Theorem also corroborates this hypothesis: the large number of small events in the catalog reduces the effective variance in isolating the cluster locations, reducing the overall effect of the addition of noise in epicenter locations. Figure 5.3 displays the HSS and PSS for Δt_1 and the different catalogs considered. The maximum values for both skill scores are similar, regardless of the noise level of

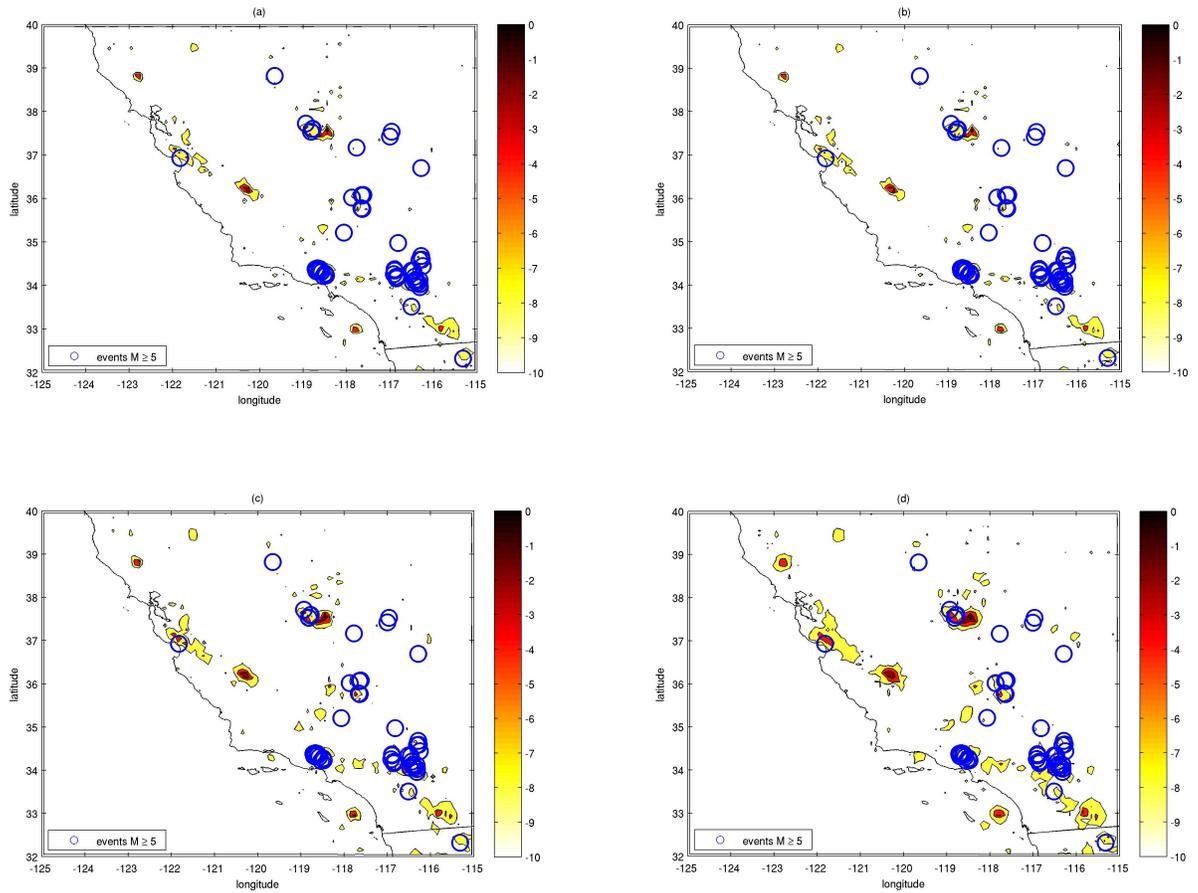


Figure 5.1: Spatial distribution of results from equation (5.4) for the training period Δt_1 considering (a) no noise, (b) $\sigma_n = 0.007^\circ$, (c) $\sigma_n = 0.05^\circ$ and (d) $\sigma_n = 0.1^\circ$. Events with magnitude M_f or greater that occurred during the forecast period are represented by open blue circles. Color scale represents $\log_{10}(PI/PI_{max})$.

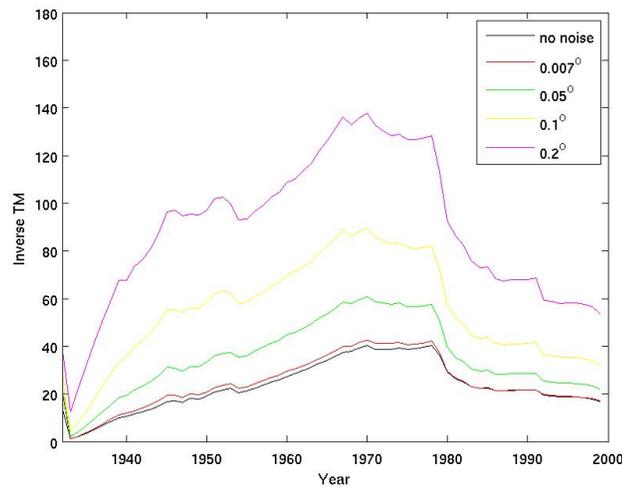


Figure 5.2: Inverse TM metric of seismicity recorded from 1932 until 2000 along with plots for the datasets obtained from different noise levels σ_n . For these plots, $D_e = 1$ for a better comparison of clustering intensity of the different catalogs.

the catalog. The difference between the curves arises from the decrease in seismicity clustering with the addition of noise to the epicenter locations, as verified in Figure 5.2, resulting in a decrease in the variability of the values calculated by equation (5.4). However, the location of hotspots should not change drastically with the addition of noise to the catalogs as a result of the persistence of the clustering patterns. This leads to the horizontal shift in the scores plotted in Figure 5.3 as the noise level increases.

To verify the hypothesis that the efficiency of the PI does not drastically change with location errors, a series of 300 PI retrospective forecasts for Δt_1 and for each noise level was performed. Figure 5.4 displays the distribution of maximum scores for both HSS and PSS for the forecast of the events E_1 and the different noise levels. These scores are estimated to follow normal distributions with standard deviations

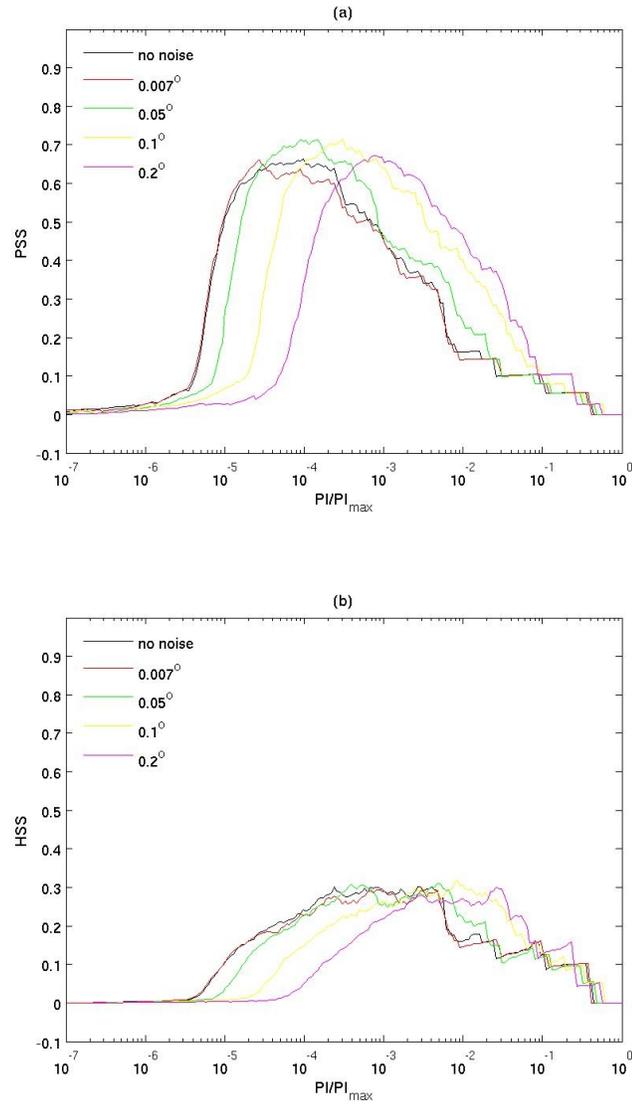


Figure 5.3: Plot of the (a) HSS and (b) PSS for Δt_1 and different threshold levels.

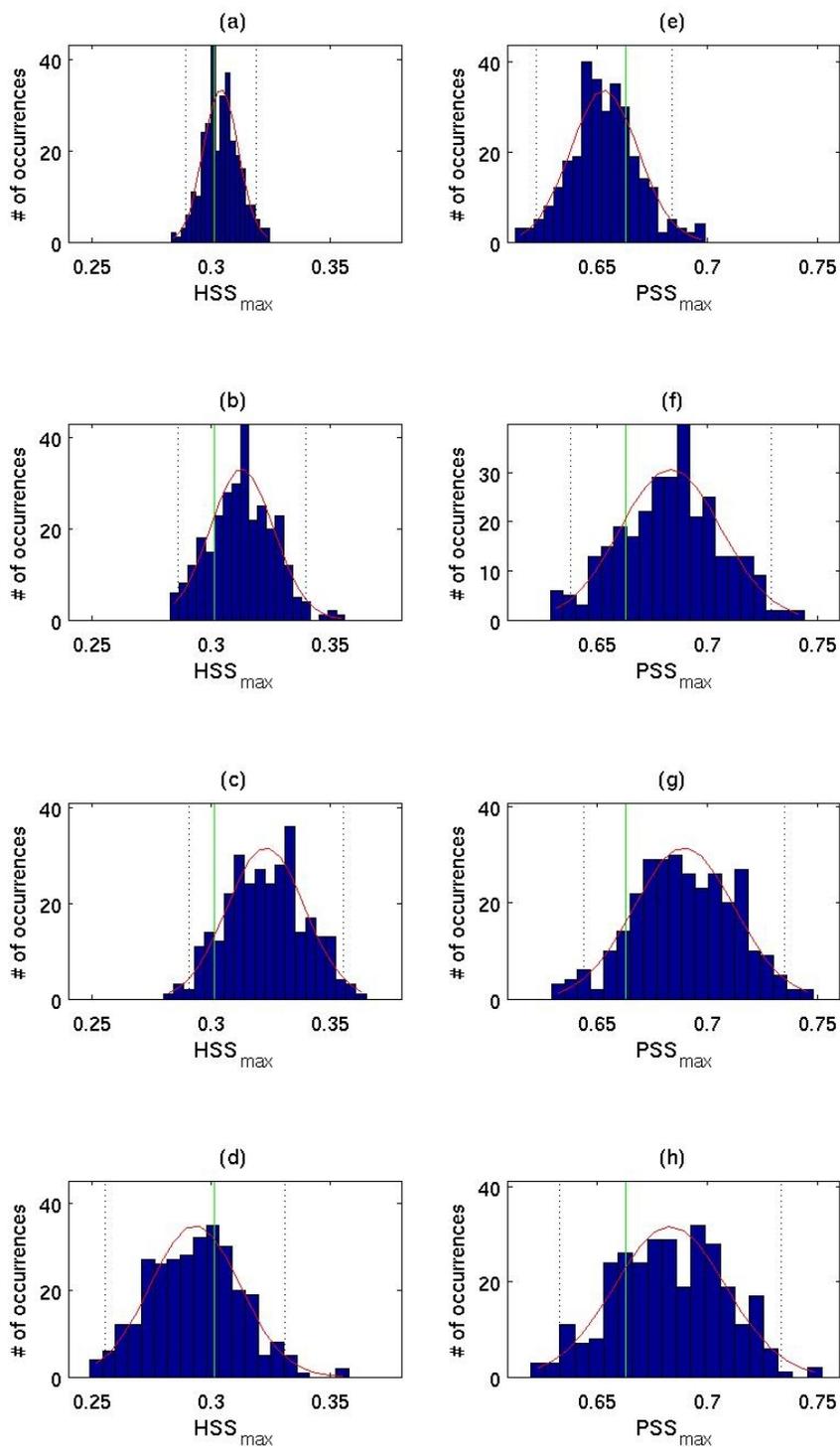


Figure 5.4: Distribution of maximum HSS from (a) to (d) and PSS from (e) to (h) for Δt_1 and $\sigma_n = 0.007^\circ$, 0.05° , 0.1° and 0.2° , respectively. Red lines represent normal distributions generated from the means and standard deviations σ_{SS} of the score values in each case, while the dot-lines represent the $\pm 2\sigma_{SS}$. Green lines represent the maximum score obtained from the unperturbed catalog.

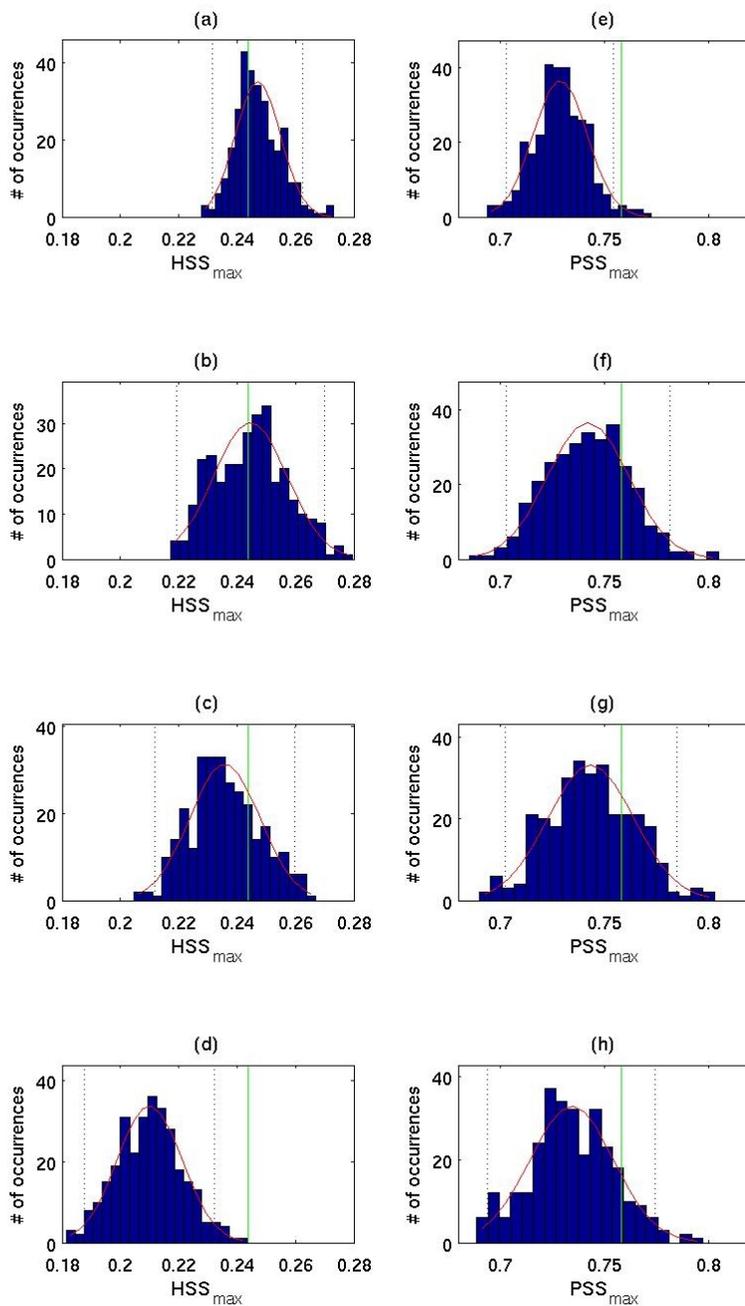


Figure 5.5: Distribution of maximum HSS from (a) to (d) and PSS from (e) to (h) and Δt_2 for $\sigma_n = 0.007^\circ$, 0.05° , 0.1° and 0.2° , respectively. Red lines represent normal distributions generated from the means and standard deviations σ_{SS} of the score values in each case, while the dot-lines represent the $\pm 2\sigma_{SS}$. Green lines represent the maximum score obtained from the unperturbed catalog.

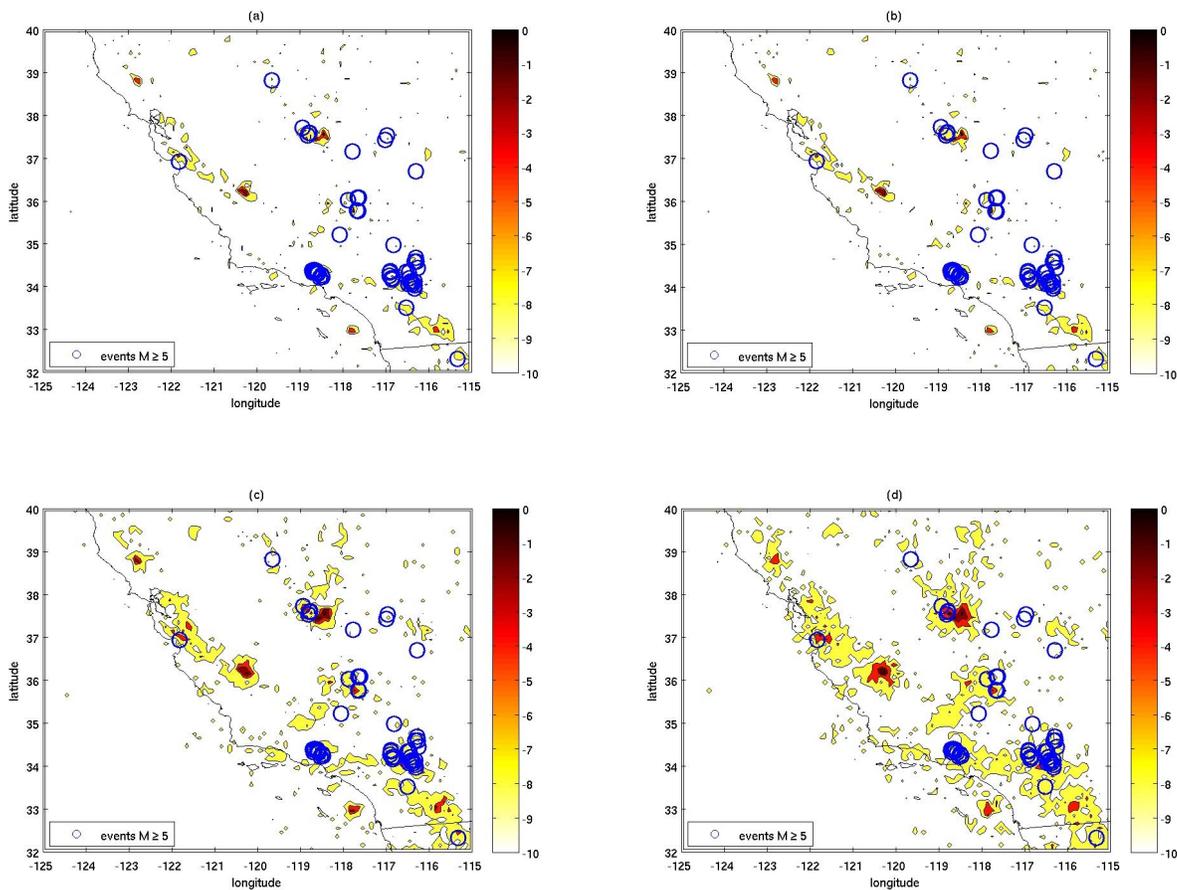


Figure 5.6: Spatial distribution of results from equation (5.4) considering the cropped catalog, where $N_r = 22500$ events are removed, for the training period Δt_1 considering (a) no noise, (b) $\sigma_n = 0.007^\circ$, (c) $\sigma_n = 0.1^\circ$ and (d) $\sigma_n = 0.2^\circ$. Events with magnitude M_f or greater that occurred during the forecast period are represented by open blue circles. Color scale represents $\log_{10}(PI/PI_{max})$.

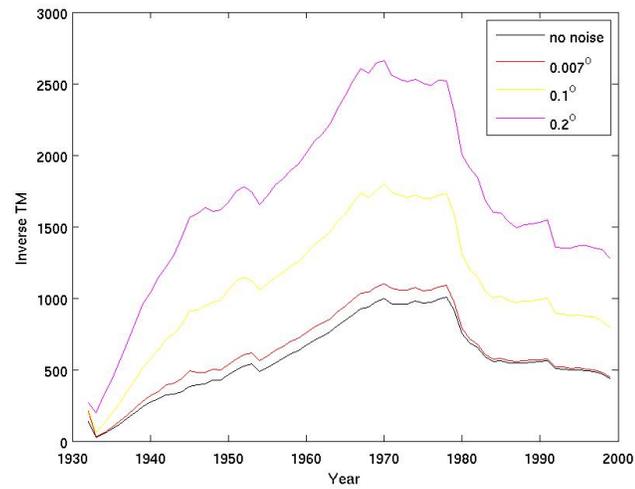


Figure 5.7: Inverse TM metric of seismicity of part of the seismicity recorded from 1932 until 2000. In this case, more than 20000 events were randomly removed from the original catalog. The inverse TM metric is also plotted for the datasets obtained from the addition of different noise levels σ_n to this cropped dataset. For these plots, $D_e = 1$ for a better comparison of clustering intensity of the different catalogs.

σ_{SS} , which became wider as σ_n increased due to the direct correlation between the latter and the number of possible configurations of n_j . In all cases, the maximum score of the unperturbed catalog lies within the $\pm 2\sigma_{SS}$ interval of the distribution of maximum scores of the perturbed catalogs. This indicates that the average result obtained by the different perturbed catalogs are not considerably different from what was obtained by the unperturbed catalog.

Figure 5.5 displays the same distribution of maximum scores for PI retrospective forecasts for the second training period Δt_2 and the different catalogs considered. The same direct correlation between σ_{SS} and the noise level σ_n is also observed. In this case, the maximum score of the unperturbed catalog was outside the $\pm 2\sigma_{SS}$ interval in two distinct scenarios: for the PSS at $\sigma_n = 0.007^\circ$ and for the HSS at $\sigma_n = 0.2^\circ$. The first performance decline is due to an average drop of H generated by the increase of misses c in the configuration that yields the maximum PSS while the false alarm rate remained constant (refer to the definition of the PSS in (5.6)). The fewer events in E_2 enhances the effects of this increase in c in the computation of H and thus the maximum PSS.

The decrease in the variability of the PI scores introduced by the high $\sigma_n = 0.2^\circ$ is the main reason for the second observed degradation, which can be seen as an increase in the random nature of this particular forecast. This results in a worse performance by the HSS standard, since it is a comparison of a given forecast to a random one. This effect is also enhanced by the fewer number of events in E_2 , as in the previous case. While this scenario holds no major concern given that the noise level in this case is of $\approx 22\text{km}$, which does not reflect the true epicenter location errors in southern

California, it might prove important in other regions where errors are significantly higher, particularly those associated with depth.

It is important to stress that both the PSS and the HSS use different criteria to quantify the effectiveness of a binary forecast. For example, the increase in misses at $\sigma_n = 0.007^\circ$ was fundamental for the degradation of the PI performance within the scope of the PSS. A similar degradation was not obtained when employing the HSS at the same noise σ_n . A systematic decrease in the PI performance would be a good indicator of a significant effect errors in epicenter locations. However, the localized degradations due to the nature of the skill scores do not offer significant evidence of an overall impact in the PI due to location errors.

The effects of reducing the total number of events in the catalog also was considered. Figure 5.7 is a plot of the inverse TM metric for the unperturbed cropped catalog and its perturbed versions. As before, the values of the inverse TM metric are directly correlated to σ_n due to the decrease of clustering as the noise levels increase. The invariance of the clustering trends in this plot indicates that the PI performance should not be drastically affected in this case. This is supported by Figure 5.6, which displays the PI hazard maps obtained from the cropped catalog along with its perturbed versions for Δt_1 along with the set of events to be forecasted E_1 . The same overall persistence of the locations of hotspots is observed in these maps, where the patterns become more diffuse as the noise level increases once again. This increase is larger compared to what is observed in Figure 5.1 and it is a direct consequence of the lower number of events in this case.

The persistence of the clustering patterns identified by the TM in this case implies

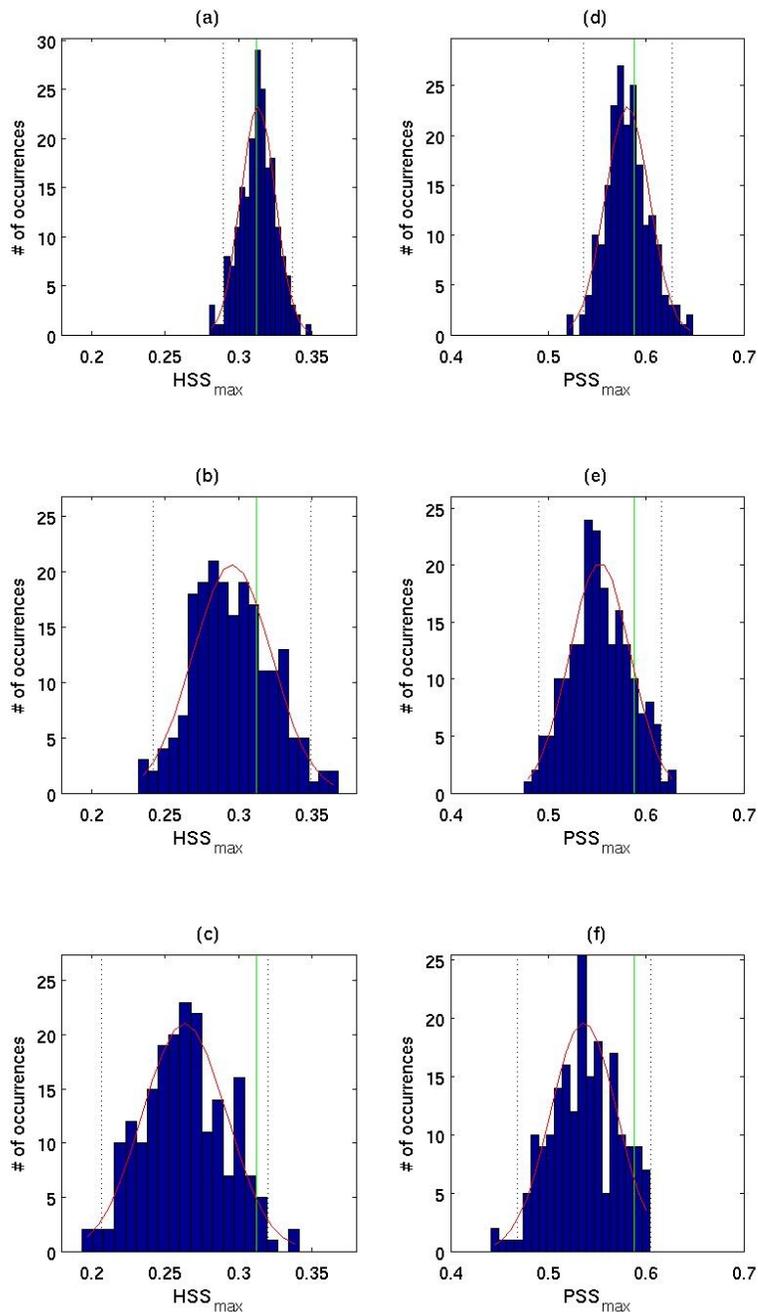


Figure 5.8: Distribution of maximum HSS from (a) to (c) and PSS from (d) to (f) for Δt_1 and $\sigma_n = 0.007^\circ$, 0.1° and 0.2° , respectively. Events were randomly removed in this case to total 5359 events from 1932 until 2002. Red lines represent normal distributions generated from the means and standard deviations σ_{SS} of the score values in each case, while the dot-lines represent the $\pm\sigma_{SS}$. Green lines represent the maximum score obtained from the unperturbed catalog.

again that the PI performances should not vary greatly for these catalogs, as well. Figure 5.8 displays the distributions of the maximum HSS and PSS for the 200 retrospective forecasts for $\Delta 1$ for each σ_n . The PI performances in the perturbed cases do not display a systematic degradation compared to the performance obtained from the unperturbed catalog, as expected.

5.4 Conclusion

The effects of location errors in seismic catalogs to the performance of the PI was addressed here by artificially introducing different levels of noise to the southern Californian dataset. This allowed for a simple analysis of the PI performance with respect to error in epicenter locations. The TM metric results indicates that the seismicity clustering patterns remain relatively constant despite the addition of noise to the original catalog. This invariance leads to a similar performance of the PI for the different noise levels considered when considering the Heidke and Pierce skill scores to quantify the effectiveness of the method.

Significant deviations between the performance of the unperturbed and perturbed catalog were observed for individual scenarios. These differences occurred due to the sensitivity of the skill scores to a particular feature of the retrospective forecast. An overall systematic decrease of the maximum performance of the PI with respect to both skill scores was not observed for any of the noise levels used. This suggests that the PI performance is not sensitive to the locations errors for catalogs with large numbers of events as a result of its dependence on seismic clustering for its forecasting

skill.

The random removal of events to analyze the combined effect of the number of events and location errors also was considered. A significant number of events was removed from the original dataset, but no significant change was observed in the clustering of the system. The enhanced diffusion of clustering patterns due to higher noise levels is as a direct consequence of the fewer number of events. No significant difference between the performances of the PI for the perturbed and unperturbed catalogs was also observed. The latter once again indicates the importance of the clustering trends in the performance of the PI.

This work also suggests that the efficacy of seismicity-based forecasting techniques that rely on large numbers of events over small-to-medium magnitudes will display similar behaviour. Future work will extend this analysis to techniques such as the Relative Intensity method (Nanjo et al., 2005). In addition, additional studies should include the removal of more events from the original catalog in order to determine the minimum number necessary to offset the error in location, and an analysis based on a non-random criteria in order to assess different factors such as changes in recording capabilities. The effects of the choice of the training period and grid sizes associated with location errors will also be considered.

Acknowledgments

This research was supported by the NSERC & Aon Benfield/ICLR Industrial Research Chair in Earthquake Hazard assessment. The authors would like to thank H.

C. Li for his contribution to this work.

References

- Chen, C. C., Rundle, J. B., Holliday, J. R., Nanjo, K. Z., Turcotte, D. L., Li, S. C., and Tiampo, K. F. (2005). The 1999 chi-chi, taiwan, earthquake as a typical example of seismic activation and quiescence. *Geophysical Research Letters*, 32(L22315):1–4.
- Cho, N. F., Tiampo, K. F., McKinnon, S. D., Vallejos, J. A., Klein, W., and Dominguez, R. (2010). A simple metric to quantify seismicity clustering. *Nonlinear Processes in Geophysics*, 17:293–302.
- Cicerone, R. D., Ebel, J. E., and Britton, J. (2009). A systematic compilation of earthquake precursors. *Tectonophysics*, 476:371–396.
- Field, E. H. (2007). Overview of the working group for the development of regional earthquake likelihood models (relm). *Seismological Research Letters*, 78:7–16.
- Heidke, P. (1926). Berechnung des erfolges und der gute der windstarkevorhersagen im sturmwarnungsdienst. *Geografiska Annaler*, 8:301–349.
- Holliday, J. R., Nanjo, K. Z., Tiampo, K. F., Rundle, J. B., and Turcotte, D. L. (2005). Earthquake forecasting and its verification. *Nonlinear Processes in Geophysics*, 12:965–977.
- Hutton and Yu (2008). News flash! scsn earthquake catalog completed. Technical report.
- Kanamori, H. (1981). *The nature of seismicity patterns before large earthquakes, in Earthquake Prediction: an International Review*. American Geophysical Union.

- Kanamori, H. (2003). *Earthquake Prediction: An Overview, in the International Handbook of Earthquake and Engineering Seismology*. Academic Press, Amsterdam.
- Keilis-Borok, V. (2005). Earthquake prediction: state-of-the-art and end emerging possibilities. *Annual Review of Earth and Planetary Sciences*, 30:1–33.
- Nanjo, K. Z., Holliday, J. R., Chen, C. C., Rundle, J. B., and Turcotte, D. L. (2005). Application of a modified pattern informatics method to forecasting the locations of future large earthquakes in the central japan. *Tectonophysics*, 424:351–366.
- Perlock, P. (2009). Imaging faults in 3-d using the pattern informatics technique - m. sc. thesis. Technical report.
- Petersen, M. D., Frankel, A. D., Harmsen, S. C., Mueller, C. S., Haller, K. M., Wheeler, R. L., Wesson, R. L., Zeng, Y., Boyd, O. S., Perkins, D. M., Luco, N., Field, E. H., Wills, C. J., and Rukstales, K. S. (2008). Documentation for the 2008 update of the united states national seismic hazard maps. Technical report, USGS.
- Rundle, J. B., Tiampo, K. F., Klein, W., and Martins, J. S. S. (2002). Self-organization in leaky threshold systems: The influence of near-mean field dynamics and its implications for earthquakes, neurobiology, and forecasting. *Proceedings of the National Academy of Sciences of the United States of America*, 99(Supplement 1):2514–2521.
- Thirumalai, D. and Mountain, R. D. (1993). Activated dynamics, loss of ergodicity, and transport in supercooled liquids. *Physical Review E*, 47:479–489.

- Thirumalai, D., Mountain, R. D., and Kirkpatrick, T. R. (1989). Ergodic behavior in supercooled liquids and in glasses. *Physical Review A*, 39:3563–3574.
- Tiampo, K. F., Rundle, J. B., and Klein, W. (2006). Stress shadows determined from a phase dynamical measure of historic seismicity. *Pure and Applied Geophysics*, 263:2407–2416.
- Tiampo, K. F., Rundle, J. B., Klein, W., Holliday, J. B., Martins, J. S. S., and Ferguson, C. D. (2007). Ergodicity in natural earthquake fault networks. *Physical Review E*, 75(066107):1–15.
- Tiampo, K. F., Rundle, J. B., McGinnis, S., and Klein, W. (2002). Pattern dynamics and forecast methods in seismically active regions. *Pure and Applied Geophysics*, 159(10):2429–2467.
- van Stiphout, T., Schorlemmer, D., and Wiemer, S. (2011). The effect of uncertainties on estimates of background seismicity rate. *Bulletin of the Seismological Society of America*, 101(2):482–494.
- Werner, M. J. and Sornette, D. (2008). Magnitude uncertainties impact seismic rate estimates, forecasts, and prediction experiment. *Journal of Geophysical Research*, 113(B08302):1–20.
- Wiemer, S., Giardini, D., Fäh, D., Deichmann, N., and Sellami, S. (2009). Probabilistic seismic hazard assessment of switzerland: best estimates and uncertainties. *Journal of Seismology*, 13(4):449–478.
- Wilks, D. S. (1995). *Statistical methods in the atmospheric sciences*. Academic Press.

Woodcock, F. (1976). The evaluation of yes/no forecasts for scientific and administrative purposes. *Geophysical Research Letters*, 104(10):1209–1214.

Chapter 6

Conclusion

Statistical seismology offers many different tools to better understand earthquakes. This work presents an in-depth analysis of the study of ergodicity in an earthquake system using the TM metric. Initially, it was explicitly determined that this metric can be considered a simple measurement of seismicity clustering (Cho et al., 2010). Periods of effective ergodicity were interpreted to be instances of time in which seismicity is declustered, during which the evolution of the cumulative number of events can be considered a normal diffusive process much like a random walk. The latter is an interesting interpretation for declustered seismicity because it relates declustered seismicity to a system that evolves randomly. The TM metric was applied successfully as a clustering measurement to both regional and mining seismicity, but some precautions need to be taken. Regions with different seismicity rates should be addressed individually when using the TM metric.

The concept of effective ergodicity and its relation to seismicity clustering was key in the adaptation of methods for regional seismic hazard assessment to a mining-induced seismicity scale. The choices of parameters to be used in the PI and RI for mining seismicity were made based on the inspection of clustering patterns of seismicity identified by the TM metric, rather than in an *ad-hoc* fashion. The PI outperformed the RI in most cases, especially during effective ergodic periods. The link between effective ergodicity and the occurrence of declustered seismicity allowed for a clear understanding of this improvement in the PI performance: during these

instances of time, quiescence is better identified due to the lack of clusters that damp its signal. This is analogous to the use of declustered catalogs in long-term regional seismic hazard assessment (Petersen et al., 2008; Wiemer et al., 2009) and it complements previous observations of high efficiency of the PI during periods of effective ergodicity for regional seismicity (Tiampo et al., 2010).

Another straightforward application of the TM metric was the formulation of a declustering technique for seismicity, an important step in the assessment of long-term seismic hazard (Wiemer et al., 2009; Petersen et al., 2008) or even the study of aftershock sequences. The proposed method is generic while maintaining a simple formulation by approaching the removal of clustered events as an optimization problem. It consists in removing events from a given dataset so that the remaining seismicity displays a long effective ergodic period. The latter allows for a physical interpretation of declustered seismicity as a system in equilibrium such as the cellular automaton in Rundle et al. (1995). The choice of the events to be considered clustered is constrained by the frequency-magnitude distribution of events of the considered catalog. This is a broad approach and it was successfully tested in a synthetic catalog and performed well for regions with various tectonic settings, as long as a homogeneous coverage and good data is available. The Gardner-Knopoff method was modified for central Europe (Wiemer et al., 2009) and Reasenbergs methods displayed a poor performance for the region comprised of southern Spain and northern Africa. On the other hand, the proposed declustering technique displayed the significant advantage of requiring no adjustment in its application to the different regions once the spatial/temporal discretization was defined. The determination of the initial trend to be extrapolated

in order to obtain the long effective ergodic period is of extreme importance to the method and should be addressed with caution. Other declustering methods based on different approaches that have strived towards improving declustering seismicity are also available (e.g. Marsan and Lengliné (2008); Zaliapin et al. (2008)).

Lastly, the effects of uncertainties in epicenter locations in the performance of the PI was examined. Different studies have shown that magnitude uncertainties (Werner and Sornette, 2008) and hypocenter errors (van Stiphout et al., 2011) may affect the estimation of seismic rates, a key step in seismic hazard estimation. The southern Californian dataset was used to generate various perturbed catalogs by adding normally distributed noise at different levels to event locations. Here the PI method was used to evaluate the changes in the resulting changes in the catalog. Retrospective PI forecasts using the unperturbed catalog and various perturbed catalogs were performed. No systematic degradation in the PI performance was observed when considering the different perturbed catalogs. The effect of reducing the number of catalog events also was analyzed by randomly removing events from the unperturbed catalog. Perturbed datasets then were generated from the cropped catalog with different noise levels and the same analysis was performed. No systematic discrepancies between the PI performances for the unperturbed and perturbed catalogs were observed for this case as well. The TM metric suggested that clustering patterns remained relatively invariant for all the different perturbed datasets when compared to the original catalog. The dependence of the PI on clustering patterns suggests that the performance of the method should not change considerably. The latter is corroborated by the Central Limit Theorem given the large number of events considered, even with the removal

of seismicity, that would offset the overall effect of the perturbations in epicenter locations.

6.1 Future Work

The study of earthquake systems in the framework of effective ergodicity has provided important insights into the long-term behavior of earthquake fault systems, as well as a practical application for the declustering of seismic catalogs. Interpreting the TM metric as a measurement of clustering allowed for different applications and the better comprehension of previous results. More can be done in this line of research, including:

1. the implementation of the declustering method in 3-D for regions with good coverage and a long recording history;
2. the application of constrains other than the frequency-magnitude distribution of events to the declustering method based on the TM metric;
3. a further examination of better ways to estimate the initial trend used in the declustering method proposed;
4. the improvement of the optimization methods used in the proposed declustering technique to reduce computation times;
5. the performance of the PI in declustered catalogs obtained from the TM metric as performed in other long-term seismic hazard assessment studies (Wiemer et al., 2009; Petersen et al., 2008)

6. a deep analysis of the clustered seismicity identified by the TM metric.

References

- Cho, N. F., Tiampo, K. F., McKinnon, S. D., Vallejos, J. A., Klein, W., and Dominguez, R. (2010). A simple metric to quantify seismicity clustering. *Non-linear Processes in Geophysics*, 17:293–302.
- Marsan, D. and Lengliné, O. (2008). Extending earthquakes’ reach through cascading. *Science*, 319:1076–1079.
- Petersen, M. D., Frankel, A. D., Harmsen, S. C., Mueller, C. S., Haller, K. M., Wheeler, R. L., Wesson, R. L., Zeng, Y., Boyd, O. S., Perkins, D. M., Luco, N., Field, E. H., Wills, C. J., and Rukstales, K. S. (2008). Documentation for the 2008 update of the united states national seismic hazard maps. Technical report, USGS.
- Rundle, J. B., Klein, W., Gross, S., and Turcotte, D. L. (1995). Boltzmann fluctuations in numerical simulations of nonequilibrium lattice threshold systems. *Physical Review Letters*, 75:1658–1661.
- Tiampo, K. F., Klein, W., Li, H. C., Mignan, A., Toya, Y., Kohen-Kadoshand, S. Z. L., Rundle, J. B., and Chen, C. C. (2010). Ergodicity and earthquake catalogs: forecast testing and resulting implications. *Pure and Applied Geophysics*, 167(6–7):763–782.
- van Stiphout, T., Schorlemmer, D., and Wiemer, S. (2011). The effect of uncertainties on estimates of background seismicity rate. *Bulletin of the Seismological Society of America*, 101(2):482–494.

Werner, M. J. and Sornette, D. (2008). Magnitude uncertainties impact seismic rate estimates, forecasts, and prediction experiment. *Journal of Geophysical Research*, 113(B08302):1–20.

Wiemer, S., Giardini, D., Fäh, D., Deichmann, N., and Sellami, S. (2009). Probabilistic seismic hazard assessment of switzerland: best estimates and uncertainties. *Journal of Seismology*, 13(4):449–478.

Zaliapin, I., Gabrielov, A., Keilis-Borok, V., and Wong, H. (2008). Clustering analysis of seismicity and aftershock identification. *Physical Review Letters*, 101(018501).

A Computer code

A.1 decluster.m

```

function decluster(catalog,t_start,p)

[var,time,num_boxes] = ergodicity(catalog);

%% Declustering

load count

cum_events1 = cum_events;
count_declust = count;

%%%% This part gets the linear trend of the approx of the TM
metric %%%
t = t_start;
fit = polyfit(time(1:t_start),var(1:t_start),1);
y = time.*fit(1) + fit(2);

%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% get the boxes that have at least one event in the 11th-time

t = t + 1; % time that we are taking the clustering into
consideration

for i = t:length(time),
    i
    index = find(count(:,t) > 3);
    if (isempty(index) == 1),
        index = find(count(:,t) > 0);
    end
    count_t = count(:,t);
    param = count(index,t)';
    length(param)
    save('par','index','cum_events1','t','count_t','num_boxes')

    low_bound = zeros(1,length(param));
    up_bound = param;

    [gbest,fit_func] = PSO(up_bound,low_bound,y(i));
    %[gbest,fit_func] = MOL(up_bound,low_bound,y(i));

```

```

%[gbest, fit_func] = PSO-AIW(up_bound, low_bound, y(i));

x = gbest(1:length(param));
N = size(cum_events1,1);
cum_test = count_t;
cum_test(index) = x;
cum_events1(:,t) = cum_test + cum_events1(:,(t-1));

count_declust(:,t) = cum_test;
t = t + 1;
end

cum_events_declust = cum_events1;
cum_events1 = cum_events;

%% Selection of events, where p(1) = b-value and p(2) = a-value
beta = p(1)/log10(exp(1));
m1 = min(catalog(:,4));
m2 = max(catalog(:,4));
lambda = 1/(exp(-beta*m1) - exp(-beta*m2));

p_dist = @(x) beta*lambda*exp(-beta*x);

aft = count - count_declust;

catalog_index = [catalog(:,1:5) ind];
random_catalog = [];
k = 1;

for i = 1:length(time),
    index = find(catalog_index(:,6) == i); % get events that
        happened in time-step "i"
    chunk_time = catalog_index(index,:); % crop of the catalog
        for time-step "i"

    index_remove = find(aft(:,i) > 0); % where there is at least
        1 aftershock in time-step "i"
    box_random = find(aft(:,i) == 0); % finding boxes with no
        correlated events

    for j = 1:length(box_random),
        ind = find(chunk_time(:,7) == box_random(j));
        random_catalog = vertcat(random_catalog, chunk_time(ind,:))
            ); % grouping events from random boxes to the
            declustered catalog
        chunk_time(ind,:) = []; % remove random events from
            chunk_time
    end
end

```

```

end

for j = 1:length(index_remove), % analyzing boxes
    individually to extract aftershocks
    index_box = find(chunk_time(:,7) == index_remove(j));
    chunk = chunk_time(index_box,:); % events in box "j" and
        year "i"

    for l = 1:aft(index_remove(j),i), % choosing events one
        by one

        m_min = min(chunk(:,4));
        m_max = max(chunk(:,4));

        argu = rand()*(exp(-beta*m_min) - exp(-beta*m_max)) +
            exp(-beta*m_max); % so that the sorted magnitude
            varies from m_min to m_max

        probs = -log(argu)/beta; % chosen magnitude
        comp = abs(chunk(:,4) - probs); % to find the event
            with the closest magnitude

        ind_search = find(comp == min(comp)); % vector that
            shows which magnitudes are the closest from the
            chosen magnitude "probs"

        diff(k) = min(comp);
        diff_mag(k) = (m_max - m_min);
        if diff_mag(k) == 0,
            diff_rel(k) = 0;
        else
            diff_rel(k) = min(comp)/(m_max - m_min);
        end

        aft_catalog(k,:) = chunk(ind_search(1),:); % choosing
            the event as an aftershock
        chunk(ind_search(1),:) = []; % deleting the chosen
            event
        k = k + 1;
    end

    random_catalog = vertcat(random_catalog,chunk); % adding
        the left overs to the random catalog
    clear chunk
end
end
end

```



```

    box_index = (index_dep-1)*(x_box*y_box) + (index_lat - 1)*
        x_box + index_lon;
    count(box_index, index_time) = single(count(box_index,
        index_time) + 1);

    ind(i,:) = [index_time box_index];
end

num_boxes = x_box*y_box; % total number of boxes

% Calculating the TM metric

integral = zeros(num_boxes,1);
epsilon = zeros(num_boxes, length(time));

for i = 1:length(time),
    integral = count(:,i) + integral;
    epsilon(:,i) = (1/i).*integral; % eq (2) from "ergodicity in
        natural earthquake fault networks"
    mean_epsilon = (1/num_boxes)*sum(epsilon(:,i)); % eq 3
    var = (epsilon(:,i) - mean_epsilon).^2;
    term(i) = sum(var);
    TM(i) = (1/(num_boxes))*term(i); % eq 1
end

D = 1/TM(1); % The diffusion coefficient;
TM = TM.*D;

cum_events(:,1) = count(:,1);
for i = 2:length(time);
    cum_events(:,i) = cum_events(:,(i-1)) + count(:,i);
end

save('count.mat', 'count', 'ind', 'cum_events')

av_square = sum(cum_events.^2)./num_boxes;
square_av = (sum(cum_events).^2)./num_boxes.^2;

var = av_square - square_av;

```

A.3 PSO.m

```

function [gbest, fit_func] = PSO(up_bound, low_bound, y);

% INPTUS
% up_bound: row vector with upper limits of number of events in
    each box with at least one event (from clustered data)

```

```

% low_bound: row vector with lower limits of number of events in
    each box with at least one event
% y: y-coordinates of the reference point

% OUTPUTS
% gbest: best solution of the PSO
% fit_func: scores based on the distance between the model and
    the possible
% solutions
% residual: vector with the difference between global minimum and
    j-th
% minimum found

load par

% Options for the PSO
N_ind = 400; % number of individuals in the population
gen = 20000; % number of generations
param = length(up_bound); % number of parameters

conv = 1;
count = 1;

%c0 = linspace(0.5,1.4,gen);
%c1 = 2; % learning factor
%c2 = 2; % learning factor

coef = [20 -0.4438 -0.2699 3.395; 30 -0.6031 -0.6485 2.6475; 50
    -0.2256 -0.1564 3.8876; 100 -0.2089 -0.0787 3.7637];

compa = abs(coef(:,1) - param);
ind = find(compa == min(compa));

c0 = coef(ind(1),2).*ones(gen,1);
c1 = coef(ind(1),3);
c2 = coef(ind(1),4);

amp = up_bound - low_bound;
for i = 1:N_ind,
    in_sol(i,:) = round(amp.*rand(1,param) + low_bound); %
        generating initial solutions randomly from "low_bound" to
        "up_bound"
    v_initial(i,:) = round(0.5.*(amp.*rand(1,param) + low_bound))
        ; % generating initial velocities
end
sol = in_sol;

```

```

while (conv >= 0.0001 || count < 2*gen/100),

    for i = 1:N_ind,
        cum_test = count_t;
        cum_test(index) = sol(i,:)';
        cum_test = cum_test + cum_events1(:,(t-1));
        y_mod = sum(cum_test.^2)./num_boxes - sum(cum_test./
            num_boxes).^2;

        fit_func(i,:) = [sol(i,:) sum((y - y_mod).^2)]; %
            matrix where each row is a particle and the last
            element of the row is the fitness score
    end

    if (count == 1), % it is the first interaction of them
        ALL
        pbest = fit_func;
        scores = pbest(:,param+1);
        ind = find(scores == min(scores));
        gbest = pbest(ind(1),:); % one single particle and
            the fitness score in the last element of the
            vector
    else
        scores_old = scores; % fitness scores of previous
            generation
        scores = fit_func(:,param+1); % fitness scores of
            current generation
        comp = scores < scores_old;

        ind = find(comp == 1); % particles that have improved
            their fitness
        pbest(ind,:) = fit_func(ind,:); % rewriting the
            particles that have improved in the "pbest" matrix
        scores = pbest(:,param+1);
        ind = find(scores == min(scores)); % locating the
            minimum fitness score in generation "j"
        if (scores(ind(1)) < gbest(param+1)), % comparing
            scores
            gbest = pbest(ind(1),:);
        end
    end

    end

    for k = 1:N_ind, % this loop updates each solution
        if (count == 1), % first interaction ever
            v = round(c0(count).*v_initial(k,:) + c1.*rand(1,
                param).*(gbest(1:param) - sol(k,:)) + c2.*rand
                (1,param).*(pbest(k,1:param) - sol(k,:)));

```

```

        sol(k,:) = sol(k,:) + v;
        v_last(k,:) = v;
    else
        v = round(c0(count).*v_last(k,:) + c1.*rand(1,
            param).*(gbest(1:param) - sol(k,:)) + c2.*rand
            (1,param).*(pbest(k,1:param) - sol(k,:)));
        sol(k,:) = sol(k,:) + v;
        v_last(k,:) = v;
    end

    comp = find(sol(k,:) > up_bound); % checking if
        current solution exceeds upper bound
    if (isempty(comp) == 0),
        sol(k,comp) = up_bound(comp);
    end

    comp = find( sol(k,:) < low_bound); % checking if
        current solution is lower than lower boundary
    if (isempty(comp) == 0),
        sol(k,comp) = low_bound(comp);
    end

    end
    conv = gbest(length(gbest));
    if count == gen,
        break
    end
    count = count + 1;
end
count
total = sum(up_bound)
remaining = sum(gbest(1:length(gbest)-1))
conv = gbest(length(gbest))

```

A.4 locations.m

```

function loc = locations(boxes);

boxsize = 0.1;
d_size = 30; % for 2D case, use d_size = dep_max-dep_min
lat_max = 40;
lat_min = 32;
lon_min = -125;
lon_max = -115;
dep_min = 0;
dep_max = 30;
mag_min = 3;
mag_max = 100;

```



```

% t1 in the PI paper
% t2 in the PI paper
% pi_score: the score for each box
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% getting the vector with the time and the location of t1 in this
  vector.
n = (t1-t0)/(time(2)-time(1)) + 1; %location of t1
limit = length(time);
num_boxes = size(series ,1);
factor_std = sqrt(1/(num_boxes-1));

% running the calculations for the pairs tb to t2 for all boxes.
  There will be
% length(time) pairs.
scores_t2 = zeros(num_boxes, limit);
for i=1:limit ,
    den = limit-i+1;
    test = (1/den).*sum(series(:,i:limit),2); % equation (1)
    mean_boxes = (1/num_boxes)*sum(test);
    var = (test-mean_boxes).^2;
    std = sqrt(sum(var));
    scores_t2(:,i) = (test - mean_boxes)./std; % eq (2)
end

% calculations of the pairs tb to t1. They are divided into 2 for
  tb less
% and greater than t1 respectively.
scores_t1 = zeros(num_boxes, limit);
for i = 1:n,
    den = n-i+1;
    test = (1/den).*sum(series(:,i:n),2); % equation (1)
    mean_boxes = (1/num_boxes)*sum(test);
    var = (test-mean_boxes).^2;
    std = sqrt(sum(var));
    scores_t1(:,i) = (test - mean_boxes)./std; % eq (2)
end

for i=(n+1):limit ,
    den = i-n+1;
    test = -(1/den).*sum(series(:,n:i),2); % equation (1)
    mean_boxes = (1/num_boxes)*sum(test);
    var = (test-mean_boxes).^2;
    std = sqrt(sum(var));
    scores_t1(:,i) = (test - mean_boxes)./std; % eq (2)
end

```

```

eq3 = scores_t2-scores_t1; % eq (3)

%eq4 = (1/limit).*sum(eq3,2); % eq (4)
eq4 = sum(eq3,2); % eq (4) de acordo com o programa em PERL
pi_score = eq4.^2;
norm_space = sum(pi_score);
pi_score = pi_score./norm_space; % eq (5)

%hotspots = pi_score - (1/num_boxes).*sum(pi_score); % eq (6)
hotspots = pi_score;
hotspots_norm = hotspots./max(hotspots);

```

A.6 moore_box_update.m

```

function [result , test] = moore_box_update(threshold , bigevents ,
      locations , hotspots , count , fore_count)

```

```

% This function calculates the number of hits , correct negatives ,
  false

```

```

% alarms and misses when using the PI for a given threshold.

```

```

%%% Inputs and Outputs %%%

```

```

% threshold: the limit in the PI score to consider a box to be a
  positive forecast

```

```

% bigevents: catalog with the big events in the forecast period

```

```

% result: vector with the Pierce's Skill Score (in this order)

```

```

% test: vector with yy, yn, ny, nn (in this order)

```

```

boxsize = 0.1;
d_size = 30; % for 2D case, use d_size = dep_max-dep_min
lat_max = 40;
lat_min = 32;
lon_min = -125;
lon_max = -115;
dep_min = 0;
dep_max = 30;
mag_min = 3;
mag_max = 100;
mag_big = 5;

```

```

t0 = 1932;

```

```

t1 = 1989;

```

```

t2 = 1999;

```

```

x_box = ceil((lon_max - lon_min)/boxsize); % number of boxes in
  the x-axis

```



```

        end
    end

    if (proxy1 == 1);
        ny = ny + 1;
        chkp = 3;
    else
        nn = nn + 1;
        chkp = 4;
    end

else % there IS A FORECAST
    for j = 1:size(bigevents,1), % there is a big event
        in the vicinity of box "i"
            checkpoint1 = abs(bigevents(j,2) - locations(i,1)
                ) - 1.5*boxsize;
            checkpoint2 = abs(bigevents(j,3) - locations(i,2)
                ) - 1.5*boxsize;
            checkpoint3 = abs(bigevents(j,4) - locations(i,3)
                ) - 1.5*d_size;

            if ((checkpoint1 <= 0) && (checkpoint2 <= 0) && (
                checkpoint3 <= 0)),
                proxy2 = 1;
                break
            end
        end
    end

    if (proxy2 == 1);
        yy = yy + 1;
        chkp = 1;
    else
        yn = yn + 1;
        chkp = 2;
    end

end

end
box_tag(i) = chkp;

end
end

p_score = ((yy*nn - yn*ny)/((yy + ny)*(yn + nn))); % Pierce's
skill score
H = yy/(yy + ny);
FA = yn/(yn + nn);
num = 2*(yy*nn - ny*yn);

```

```
den = (yy + ny)*(ny + nn) + (yy + yn)*(yn + nn);  
heidke = num/den;  
  
result = [p_score threshold H FA heidke];  
  
test = [yy yn ny nn];  
  
end
```

VITA

Nelson Cho

Post Secondary Education and Degrees:

The University of Western Ontario
2007-2011 Ph.D. in Geophysics
Universidade de Brasília, Brasília, Brasil
2000-2005 B.Sc. and M.Sc. in Physics

Honours and Awards:

WGRS 2007-2010
CAPES Graduate Scholarship - Master Degree 2003-2005

Related work experience:

2007-2010, Teaching Assistant, The University of Western Ontario, London, Canada
2005, Substitute Professor, Universidade de Brasília, Brasília, Brasil
2003-2004, Substitute Teacher, School of the Nations, Brasília, Brasil

Professional memberships:

AGU

Presentations:

Geocomplexity Workshop, Toronto-ON, 2009

Publications:

1. Cho N. F., Tiampo K. F., Mckinnon S. D., Vallejos J. A., Klein W., and Dominguez R., A Simple Metric to Quantify Seismicity Clustering, *Nonlin. Processes Geophys.*, 17, 293-302, 2010.
2. Cho N. F., Tiampo K. F., Vallejos J. A., and Mckinnon S. D., Pattern Informatics in Mining Induced Seismicity and its Application to Seismic Hazard Assessment in Mines, submitted to peer-reviewed journal.
3. Cho N. F., Tiampo K. F., Bhattacharya P., Shcherbakov R., Klein W. and Mansinha L., Declustering seismicity using the TM metric, submitted to peer-reviewed journal.
4. Cho N. F., Tiampo K. F., and LI H. C., Effects of location errors in the Pattern Informatics, submitted to peer-reviewed journal.