University of Granada

Doctorate Program in Mathematics and Statistics

Analysis of fractality characteristics in spatio-temporal processes. Applications to geophysical data

Francisco Javier Esquivel Sánchez

Thesis supervised by:
Prof. José Miguel Angulo Ibáñez

Granada 2013

Editor: Editorial de la Universidad de Granada
Autor: Francisco Javier Esquivel Sánchez
D.L.: GR 707-2014
ISBN: 978-84-9028-900-6

El doctorando Francisco Javier Esquivel Sánchez y el director de la tesis José Miguel Angulo Ibáñez garantizan, al firmar esta tesis doctoral, que el trabajo ha sido realizado por el doctorando bajo la dirección del director de la tesis y, hasta donde nuestro conocimiento alcanza, en la realización del trabajo se han respetado los derechos de otros autores a ser citados, cuando se han utilizado sus resultados o publicaciones.

Granada, 11 de noviembre de 2013

Director de la Tesis | Doctorando

Fdo.: José Miguel Angulo Ibáñez | Fdo.: Francisco Javier Esquivel Sánchez

Contents

SUMMARY **1**

RESUMEN **3**

Introduction **5**

Introducción **11**

I PRELIMINARIES **17**

1 Entropy and Complexity **19**
1.1 Uncertainty and Information Theory 20
1.1.1 Shannon entropy 20
1.1.2 Other measures of entropy 22
1.1.3 Divergences 23
1.1.4 Mutual information 25
1.1.5 A box-counting-based illustration 26
1.2 Complexity Measures 31
1.2.1 LMC measure of complexity 31
1.2.2 Other measures of complexity 32

2 Fractality, Multifractality, and Seismic Scaling Laws **35**
2.1 Fractality 36
2.1.1 Scale invariance and self-similarity 36
2.1.2 Fractal set 37
2.1.3 Fractal dimension 37
2.2 Multifractality 39
2.2.1 Introduction 39
2.2.2 The multifractal formalism 40

2.2.3 Illustrative example 43
2.3 Seismic Scaling Laws 47

II METHODOLOGY AND APPLICATIONS 51

3 Proposed Techniques and Methodological Aspects 53
3.1 Technical Extensions 54
3.1.1 A Tsallis entropy-based formulation of generalized dimensions 54
3.1.2 Some limiting behaviour aspects 56
3.1.3 Incorporation of effects through distributions 59
3.2 Further Methodological Aspects 60
3.2.1 Transformations on the temporal component 60
3.2.2 Sliding windows 61
3.2.3 Fitting of the frequency-magnitude distribution by means of weighted non-linear regression 64

4 Seismic Data Analysis 67
4.1 Agrón Data 68
4.1.1 Descriptive analysis 69
4.1.2 Study of dimensional interaction 71
4.1.3 Study of structural characteristics at different scales 78
4.2 El Hierro Data 92
4.2.1 Descriptive analysis 92
4.2.2 Study of dimensional interaction 95
4.2.3 Study of structural characteristics at different scales 97
4.2.4 Frequency-magnitude distribution analysis 98
4.3 Torreperrogil Data 107
4.3.1 Descriptive analysis 107
4.3.2 Study of dimensional interaction 109
4.3.3 Study of structural characteristics at different scales 114
4.3.4 Frequency-magnitude distribution analysis 124
4.4 Synthesis 129

III COMPLEMENTARY ASPECTS AND CONCLUSIONS 131

5 Synthesis, Complementary Aspects and Open Lines 133
5.1 Synthesis 133
5.2 Complementary Aspects 136

5.3 Open Lines 145

5 Síntesis, Aspectos Complementarios y Líneas Abiertas 151
5.1 Síntesis 151
5.2 Aspectos Complementarios 154
5.3 Líneas Abiertas 163

Bibliography 168

Appendix 179

List of Figures

1.1 Representation of p^q vs. p, depending of q. 24
1.2 Scatter plots for the scenarios A (left) and B (right) of the illustrative example. . 28
1.3 Representations of Shannon, Rényi and Tsallis entropies vs. q for scenarios A (left) and B (right), considering a box size = $1/10$. 29
1.4 Representations of Shannon entropy vs. box size for scenarios A (left) and B (right) using varying box sizes between 1/15 and 1/2. 30
1.5 Representations of Rényi (left) and Tsallis (right) entropies vs. q for scenarios A and B, considering a box size = $1/10$. 30

2.1 Fractal-like images in nature (reproduced from http://webecoist.momtastic.com/2008/09/07/17-amazing-examples-of-fractals-in-nature). 36
2.2 Cantor set (reproduced from http://classes.yale.edu/fractals/Labs/PaperFolding-Lab/MTSConstr.html). 38
2.3 Triadic Koch island (reproduced from http://what-when-how.com/computer-graphics-and-geometric-modeling/chaos-and-fractals-special-computer-graphics-part-1). 38
2.4 Representation of D_q vs. q for monofractal and multifractal data sets. 41
2.5 Representation of $f(\alpha)$ vs. α for monofractal and multifractal data sets. 42
2.6 Epicenter locations of the three periods. 44
2.7 Representation of $\tau(q)$ vs. q for three periods. 45
2.8 Representation of D_q vs. q for the three periods. 45
2.9 Representation of $f(\alpha)$ vs. α for the three periods. 46

3.1 3D scatter plots of components (X,Y,T), corresponding to TR, TN, TS and TU. 62
3.2 Illustration of the sliding windows technique with a window of fixed size according to a time interval. 63
3.3 Illustration of the sliding windows technique with a window of fixed size according to a number of events. 64

4.1 Agrón data. Epicenter locations on contoured Andalusian region. 68

4.2 Agrón data. Epicenter locations for the whole period by time, represented by colour bar, and magnitude (left) or depth (right), represented by size of circle mark. 69
4.3 Agrón data. Epicenter locations for the previous period by time, represented by colour bar, and magnitude (left) or depth (right), represented by size of circle mark. 70
4.4 Agrón data. Epicenter locations for the high seismicity activity period by time, represented by colour bar, and magnitude (left) or depth (right), represented by size of circle mark. 70
4.5 Agrón data. Epicenter locations for the subsequent period by time, represented by colour bar, and magnitude (left) or depth (right), represented by size of circle mark. 71
4.6 Agrón data. Daily frequencies for the whole period and the three subperiods. . 72
4.7 Agrón data. Magnitude distribution over time for the whole period and the three subperiods. 73
4.8 Agrón data. Frequency-magnitude histogram for the whole period and the three subperiods. 74
4.9 Agrón data. Magnitude reciprocal accumulated distribution for the whole period and the three subperiods. 75
4.10 Agrón data. Depth distribution over time for the whole period and for the three subperiods (from top to bottom). 76
4.11 Agrón data. Shannon entropy values $H(D)$, $H(X,Y)$, $H(X,Y,D)$, $H(X,Y)+H(D)$ and $H(X,Y)+H(D)-H(X,Y,D)$, for varying box sizes between 1/15 and 1/2 with D corresponding to time, depth and magnitude. 77
4.12 Agrón data. Values of the dependence coefficient for the interactions $(X,Y) \leftrightarrow T$, $(X,Y) \leftrightarrow Z$ and $(X,Y) \leftrightarrow M$, based on sliding windows (size 150 events). . . 78
4.13 Agrón data. 3D scatter plots of components (X,Y,T), corresponding to TR, TN, TS and TU. 79
4.14 Agrón data. Shannon entropy values $H(T)$, $H(X,Y)$, $H(X,Y,T)$, $H(X,Y)+H(T)$ and $H(X,Y)+H(T)-H(X,Y,T)$ based on sliding windows (size 150 events), considering a fixed box size $\varepsilon = 1/10$, corresponding to TR, TN, TS and TU. 80
4.15 Agrón data. Shannon entropy values $H(T)$, $H(X,Y)$, $H(X,Y,T)$, $H(X,Y)+H(T)$ and $H(X,Y)+H(T)-H(X,Y,T)$ for varying box sizes between 1/15 and 1/2, corresponding to TR, TN, TS and TU. 81
4.16 Agrón data. Rényi entropy values $H_q(T)$, $H_q(X,Y)$, $H_q(X,Y,T)$, $H_q(X,Y)+H_q(T)$ and $H(X,Y)+H(T)-H(X,Y,T)$ for fixed box size $\varepsilon = 1/10$ and varying deformation parameter q between -10 and 10, corresponding to TR, TN, TS and TU. 82

4.17 Agrón data. Tsallis entropy values $HT_q(T)$, $HT_q(X,Y)$, $HT_q(X,Y,T)$ and $HT_q(X,Y)+HT_q(T)-(q-1)HT_q(X,Y)HT_q(T)$ for fixed box size $\varepsilon = 1/10$ and varying deformation parameter q between 1 and 3, corresponding to TR, TN, TS and TU. 83
4.18 Agrón data. Variations in the 'capacity' dimension values over time based on sliding windows (size 150 events), with and without considering the effects of the magnitude. 84
4.19 Agrón data. Variations in the 'entropy' dimension values over time based on sliding windows (size 150 events), with and without considering the effects of the magnitude. 84
4.20 Agrón data. Variations in the 'correlation' dimension values over time based on sliding windows (size 150 events), with and without considering the effects of the magnitude. 85
4.21 Agrón data. Variations in the multifractal step values over time based on sliding windows (size 150 events), with and without considering the effects of the magnitude. 85
4.22 Agrón data. Variations in the values of the complexity measure $D_1 - D_2$ over time based on sliding windows (size 150 events), with and without considering the effects of the magnitude. 86
4.23 Agrón data. Variations in the values of the exponential version LMC over time based on sliding windows (size 150 events), with and without considering the effects of the magnitude. 87
4.24 Agrón data. Generalized Rényi dimension curves corresponding to TR, TN, TS and TU, for the three subperiods (from letf to right). 88
4.25 Agrón data. Multifractal spectra corresponding to TR, TN, TS and TU, for the three subperiods (from letf to right). 90
4.26 Agrón data. Values of generalized Rényi dimensions D_0, D_1, D_2, $D_1 - D_2$ and $D_{-\infty} - D_{\infty}$, corresponding to TR, TN, TS and TU, based on sliding windows (size 150 events). 91
4.27 El Hierro data. Epicenter locations on contoured island coastal line. 93
4.28 El Hierro data. Latitudes over time. 93
4.29 El Hierro data. Distribution of daily frequencies over time. 94
4.30 El Hierro data. Epicenter locations by time, represented by colour bar, and magnitude (left) or depth (right), represented by size of circle mark. 94
4.31 El Hierro data. Depth distribution over time. 95
4.32 El Hierro data. Magnitude distribution over time. 95
4.33 El Hierro data. Frequency-magnitude histogram (left) and reciprocal accumulated distribution (right). 96

4.34 El Hierro data. Shannon entropy values $H(D)$, $H(X,Y)$, $H(X,Y,D)$ and $H(X,Y)+H(D)$, for varying box sizes between 1/15 and 1/2, with D corresponding to time, depth and magnitude. 96
4.35 El Hierro data. Shannon entropy values $H(D)$, $H(X,Y)$, $H(X,Y,D)$, $H(X,Y)+H(D)$ and $H(X,Y)+H(D)-H(X,Y,D)$ based on sliding windows (size 150 events), considering a fixed box size $\varepsilon = 1/10$, with D corresponding to time, depth and magnitude. 97
4.36 El Hierro data. Values of the 'dependence coefficient' for components $(X,Y) \leftrightarrow T$, $(X,Y) \leftrightarrow Z$ and $(X,Y) \leftrightarrow M$, based on sliding windows (size 150 events). . . 98
4.37 El Hierro data. Variations in the 'capacity' dimension values over time based on sliding windows (size 150 events), with and without considering the effects of the magnitude. 99
4.38 El Hierro data. Variations in the 'entropy' dimension values over time based on sliding windows (size 150 events), with and without considering the effects of the magnitude. 99
4.39 El Hierro data. Variations in the 'correlation' dimension values over time based on sliding windows (size 150 events), with and without considering the effects of the magnitude. 100
4.40 El Hierro data. Variations in the step values over time based on sliding windows (size 150 events), with and without considering the effects of the magnitude. . . 100
4.41 El Hierro data. Variations in the complexity measure $D_1 - D_2$ values over time based on sliding windows (size 150 events), with and without considering the effects of the magnitude. 101
4.42 El Hierro data. Variations in the values of the LMC_{exp} over time based on sliding windows (size 150 events), with and without considering the effects of the magnitude. 101
4.43 El Hierro data. Representations of the fitted curves for the whole sequence and three phases. 103
4.44 El Hierro data. Magnitude distribution and estimated values of parameter q over time based on sliding windows (size 150 events), for the non-weighted (top) and weighted models (bottom). 104
4.45 El Hierro data. Magnitude distribution and estimated values of parameter a over time based on sliding windows (size 150 events), for the non-weighted (top) and weighted models (bottom). 105
4.46 El Hierro data. Misfit and R^2 values over time based on sliding windows (size 150 events), for the non-weighted (top) and weighted models (bottom). 106
4.47 Torreperogil data. Epicenter locations on contoured Andalusian region. 108
4.48 Torreperogil data. Daily frequencies for the whole period (top) and for a restricted interval corresponding to the high activity period (bottom). 108

4.49 Torreperogil data. Epicenter locations by time, represented by colour bar, and magnitude (left) or depth (right), represented by size of circle mark. 109
4.50 Torreperogil data. Magnitude distribution over time for the whole period (top) and for a restricted interval corresponding to the high activity period (bottom). . 110
4.51 Torreperogil data. Depth distribution over time for the whole period (top) and for a restricted interval corresponding to the high activity period (bottom). . . . 110
4.52 Torreperogil data. Shannon entropy values $H(D)$, $H(X,Y)$, $H(X,Y,D)$, $H(X,Y)+H(D)$ and $H(X,Y)+H(D)-H(X,Y,D)$, for varying box sizes between 1/15 and 1/2, with D corresponding to time, depth and magnitude 111
4.53 Torreperogil data. Rényi entropy values $H_q(D)$, $H_q(X,Y)$, $H_q(X,Y,D)$, $H_q(X,D)+H_q(D)$ and $H_q(X,Y)+H_q(D)-H_q(X,Y,D)$ for a specific box size $\varepsilon = 1/10$, with D corresponding to time, depth and magnitude. 112
4.54 Torreperogil data. Tsallis entropies $HT_q(D)$, $HT_q(X,Y)$, $HT_q(X,Y,D)$ and $HT_q(X,Y)+HT_q(D)-(q-1)HT_q(X,Y)HT_q(D)$ for a specific box size $\varepsilon = 1/10$, with D corresponding to time, depth and magnitude. 113
4.55 Torreperogil data. Visualization for the whole period of the Shannon entropy values $H(D)$, $H(X,Y)$, $H(X,Y,D)$, $H(X,Y)+H(D)$ and $H(X,Y)+H(D)-H(X,Y,D)$, based on sliding windows (size 150 events), considering a fixed box size $\varepsilon = 1/10$ with D corresponding to time, depth and magnitude. 114
4.56 Torreperogil data. Visualization for a restricted temporal interval of the Shannon entropy values $H(D)$, $H(X,Y)$, $H(X,Y,D)$, $H(X,Y)+H(D)$ and $H(X,Y)+H(D)-H(X,Y,D)$,based on sliding windows (size 150 events), considering a fixed box size $\varepsilon = 1/10$ with D corresponding to time, depth and magnitude. . . 115
4.57 Torreperogil data. Values of the dependence coefficient for the interactions $(X,Y) \leftrightarrow T$, $(X,Y) \leftrightarrow Z$ and $(X,Y) \leftrightarrow M$, based on sliding windows (size 150 events). 116
4.58 Torreperogil data. Rényi (left) and Tsallis (right) generalized dimensions curves for (X,Y,T), (X,Y,Z) and (X,Y,M). 116
4.59 Torreperogil data. Variations in the 'capacity' dimension values over time based on sliding windows (size 150 events), with and without considering the effects of the magnitude. 118
4.60 Torreperogil data. Variations in the 'entropy' dimension values over time based on sliding windows (size 150 events), with and without considering the effects of the magnitude. 119
4.61 Torreperogil data. Variations in the 'correlation' dimension values over time based on sliding windows (size 150 events), with and without considering the effects of the magnitude. 120

4.62 Torreperogil data. Variations in the multifractal step values over time based on sliding windows (size 150 events), with and without considering the effects of the magnitude. 121
4.63 Torreperogil data. Variations in the $D_1 - D_2$ values over time based on sliding windows (size 150 events), with and without considering the effects of the magnitude. 122
4.64 Torreperogil data. Variations in the LMC_{exp} values over time based on sliding windows (size 150 events), with and without considering the effects of the magnitude. 123
4.65 Torreperogil data. Representations of the fitted curves for the whole sequence and the three phases for both models. 124
4.66 Torreperogil data. From top to bottom, magnitude distribution, and estimated values of parameter q over time based on sliding windows (size 150 events), for the non-weighted and weighted models. 126
4.67 Torreperogil data. From top to bottom, magnitude distribution, and estimated values of parameter a over time based on sliding windows (size 150 events), for the non-weighted and weighted models. 127
4.68 Torreperogil data. From top to bottom, misfit and R^2 values over time based on sliding windows (size 150 events), for the non-weighted and weighted models. . 128

5.1 Simulated realizations using the Cauchy class with $\alpha = 0.5, 2$ (from left to right) and $\beta = 0.1, 0.9$ (from top to bottom). 137
5.2 Random field (top) and excursion sets for the thresholds corresponding to the quartiles (bottom from left to right), based on realization of Figure 5.1 for $\alpha = 0.5$ and $\beta = 0.1$ 138
5.3 Random field values (top) and excursion sets for the thresholds corresponding to the quartiles (bottom from left to right), based on realization of Figure 5.1 for $\alpha = 2$ and $\beta = 0.1$ 139
5.4 Random field (top) and excursion sets for the thresholds corresponding to the quartiles (bottom from left to right), based on realization of Figure 5.1 for $\alpha = 0.5$ and $\beta = 0.9$ 140
5.5 Random field (top) and excursion sets for the thresholds corresponding to the quartiles (bottom from left to right), based on realization of Figure 5.1 for $\alpha = 2$ and $\beta = 0.9$ 141
5.6 Excursion sets for the threshold corresponding to the 90th percentile, based on realizations of Figure 5.1. Black dots represent the centroids of connected components. 143
5.7 Excursion sets for the threshold corresponding to the 90th percentile, based on realizations of Figure 5.1. Black dots represent the A-exit points. 144

5.8 Dependence coefficient values for the spatial point patterns defined by centroids (left) and *A*-exit points (right) for the four models 145
5.9 Values of D_0, D_1, D_2, $D_{-\infty} - D_{\infty}$ and $D_1 - D_2$ for the spatial point patterns defined by centroids for the four models 146
5.10 Values of D_0, D_1, D_2, $D_{-\infty} - D_{\infty}$ and $D_1 - D_2$ for the spatial point patterns defined by *A*-exit points for the four models 147

5.1 Realizaciones simuladas mediante la clase de Cauchy con $\alpha = 0.5, 2$ (de izquierda a derecha) y $\beta = 0.1, 0.9$ (de arriba a abajo). 156
5.2 Campo aleatorio (arriba) y conjuntos de excursión considerando como umbrales los cuartiles (abajo de izquierda a derecha), basados en la realización de la Figura 5.1 para $\alpha = 0.5$ y $\beta = 0.1$ 157
5.3 Campo aleatorio (arriba) y conjuntos de excursión considerando como umbrales los cuartiles (abajo de izquierda a derecha), basados en la realización de la Figura 5.1 para $\alpha = 2$ y $\beta = 0.1$ 158
5.4 Campo aleatorio (arriba) y conjuntos de excursión considerando como umbrales los cuartiles (abajo de izquierda a derecha), basados en la realización de la Figura 5.1 para $\alpha = 0.5$ y $\beta = 0.9$ 159
5.5 Campo aleatorio (arriba) y conjuntos de excursión considerando como umbrales los cuartiles (abajo de izquierda a derecha), basados en la realización de la Figura 5.1 para $\alpha = 2$ y $\beta = 0.9$ 160
5.6 Conjuntos de excursión considerando como umbral el percentil 90th, basados en las realizaciones de la Figura 5.1. Los puntos negros representan los centroides de las componentes conexas. 161
5.7 Conjuntos de excursión considerando como umbral el percentil 90th, basados en las realizaciones de la Figura 5.1. Los puntos negros representan los *A-exit points*. 162
5.8 Valores del coeficiente de dependencia para los patrones puntuales espaciales definidos por los centroides (izquierda) y por los *A-exit points* (derecha) para los cuatro modelos 163
5.9 Valores de $D_0, D_1, D_2, D_{-\infty} - D_{\infty}$ y $D_1 - D_2$ para los patrones puntuales definidos por los centroides para los cuatro modelos 164
5.10 Valores de $D_0, D_1, D_2, D_{-\infty} - D_{\infty}$ y $D_1 - D_2$ para los patrones puntuales definidos por los *A-exit points* para los cuatro modelos 165

List of Tables

1.1 Shannon, Rényi and Tsallis entropy measures, and related measures of divergence and mutual information. 27

2.1 Values of D_0, D_1, D_2 and $D_{-\infty} - D_{\infty}$ for the three periods. 46

4.1 Agrón data. Values of generalized Rényi dimensions D_0, D_1, D_2, $D_1 - D_2$ and $D_{-\infty} - D_{\infty}$, corresponding to TR, TN, TS and TU, for the three subperiods. . . 89

4.2 Agrón data. Comparison of generalized Rényi and Tsallis dimension values, for $q = 0, 1, 2$. 89

4.3 El Hierro data. Parameters, misfit and R^2 of the fitted curves for the whole sequence and the three phases based on the non-weighted and weighted models. 102

4.4 Torreperogil data. Values of generalized Rényi dimensions D_0, D_1, D_2, $D_1 - D_2$, $D_{-\infty} - D_{\infty}$, and values of generalized Tsallis dimensions DT_0, DT_1, DT_2. . . 117

4.5 Torreperogil data. Values of parameters q and a, misfit and R^2 of the fitted curves based on the non-weighted and weighted models. 124

SUMMARY

Over the last decades, the consequences of earthquakes on the inhabitants and the environment where they occur have led to a considerable effort in their study. The increasing interest in understanding earthquakes, jointly with the close relationship of seismicity analysis with Statistics (see, as a reference, the works about Statistical Seismology by Vere-Jones 2000, 2001, 2010) have motivated us to center our attention in this research on this geophysical phenomenon.

In this thesis work, we focus on the statistical exploitation of information from regional space-time series of seismic events. We are interested in the intrinsic structural complexity of the spatio-temporal evolution patterns of seismic occurrences within a certain region and a certain period of time. Two main objectives are addressed: the study of the dimensional interaction and the analysis of the structural features at different scales. In particular, we focus on the following aspects: study of scaling behaviour, analysis of dimensional interaction, assessment of evolutionary changes, incorporation of effects, application to relevant real seismic data, and application to random fields excursion sets.

With this purpose, we propose various technical extensions in the context of information, complexity and multifractal analysis, combined with some methodological complementary aspects. Specifically, we formulate a new version of generalized dimensions based on Tsallis non-extensive statistics; we establish a limiting connection between some well-known complexity measures and generalized dimensions, leading to a concept of 'multifractal complexity'; we construct a dependence coefficient in the multifractal domain, useful for dimensional interaction assessment; and we propose a formal approach to incorporate local effects, such as those of the magnitude or potential information from covariates. Some transformations on the temporal component, combined with the former techniques, allow the evaluation of the effect of the occurrence times on the structural complexity displayed by the data. Further, the sliding windows technique provides a useful tool for assessing evolutionary changes (relevant, in particular, to describe temporal structural heterogeneities, as well as for detection of possible precursory elements). The fitting of the non-extensive frequency-magnitude distribution of earthquakes by means of a weighted non-linear regression, considering weights according to the released energy, provides improved results in terms of goodness of fit.

Three real seismic series are studied, namely corresponding to the areas of Agrón, El Hierro and Torreperogil, in Spain. These applications are developed not only for illustration of the

usefulness of the technical and methodological aspects introduced to describe seismicity, but also considering their intrinsic interest in relation to relevant aspects underlaying their different nature.

In particular, regarding Agrón series the study performed combining the techniques with temporal transformations has been found to be very practical for assessment of the spatio-temporal association. As for El Hierro series, the techniques applied have detected temporal changes in the structural characteristics which could be seen as a precursory indicator to the volcanic eruption. Finally, the study of the Torreperogil series has shown some results not present in the other data sets, which may suggest the existence of external factors. However, such an assertion must be further investigated based on comparative analysis with similar data sets.

As for the application to random field excursion sets defined by threshold exceedances, we perform a preliminary study using the above techniques on the variations of structural characteristics with respect to increasing threshold, and depending on the local variability and dependence range properties, based on the spatial point patterns defined by centroid of connected components and by A-exit points. The results show that multifractal analysis and complexity measures constitute a useful complementary approach for studying structural characteristics of spatial threshold exceedances.

Various open lines and further directions for continuing research are pointed out from this thesis work. These involve, in particular, further analytical aspects and generalization, incorporation of errors and data extensions, directionality assessment, and sensitivity analysis.

RESUMEN

En las últimas décadas, las consecuencias de los terremotos sobre los habitantes y el medio ambiente donde ocurren han conducido a un considerable esfuerzo en su estudio. El interés creciente en entender los terremotos, junto con la estrecha relación del análisis de la sismicidad con la Estadística (ver, como referencia, los trabajos acerca de Sismología Estadística de Vere-Jones 2000, 2001, 2010) nos ha motivado a centrar nuestra atención en esta investigación sobre este fenómeno geofísico.

En este trabajo de tesis, nos enfocamos sobre la explotación estadística de la información proveniente de series espacio-temporales regionales de eventos sísmicos. Estamos interesados en la complejidad estructural intrínseca de los patrones de evolución espacio-temporal de ocurrencias sísmicas dentro de una cierta región y un cierto periodo de tiempo. Se abordan dos objetivos principales: el estudio de la interacción dimensional y el análisis de las características estructurales en diferentes escalas. En particular, nos enfocamos sobre los siguientes aspectos: estudio del comportamiento escalar, análisis de interacción dimensional, evaluación de cambios evolucionarios, incorporación de efectos, aplicación a datos sísmicos reales relevantes y aplicación a conjuntos de excursión en campos aleatorios.

Con este propósito, proponemos varias extensiones técnicas en el contexto de información, complejidad y análisis multifractal, combinadas con algunos aspectos metodológicos complementarios. Específicamente, formulamos una nueva versión de dimensiones generalizadas basadas en la estadística no extensiva de Tsallis; establecemos una conexión límite entre algunas medidas de complejidad bien conocidas y dimensiones generalizadas, que conduce al concepto de “complejidad multifractal”; construimos un coeficiente de dependencia en el dominio multifractal, útil para la evaluación de interacción dimensional; y proponemos un enfoque formal para la incorporación de efectos locales, tales como los de la magnitud o información potencial de covariables. Algunas transformaciones sobre la componente temporal, combinadas con las técnicas anteriores, permiten la evaluación del efecto de los instantes de ocurrencia sobre la complejidad estructural exhibida por los datos. Adicionalmente la técnica de ventanas deslizantes proporciona una herramienta útil para la evaluación de cambios evolucionarios (relevante, en particular, para describir heterogeneidades estructurales temporales, así como para la detección de posibles elementos precursores). El ajuste de la distribución frecuencia-magnitud no extensiva de terremotos mediante una regresión no lineal ponderada, considerando pesos conforme a

la energía liberada, proporciona resultados mejorados en términos de bondad del ajuste.

Se estudian tres series sísmica, correspondientes a las áreas de Agrón, El Hierro y Torreperogil, en España. Estas aplicaciones se desarrollan no sólo para ilustrar la utilidad de los aspectos técnicos y metodológicos introducidos para describir sismicidad, sino también considerando su interés intrínseco en relación con aspectos relevantes subyacentes a su diferente naturaleza.

En particular, con respecto a la serie de Agrón el estudio realizado combinando las técnicas con transformaciones temporales ha resultado ser muy práctico para la evaluación de la asociación espacio-temporal. En cuanto a la serie de El Hierro, las técnicas aplicadas han detectado cambios temporales en las características estructurales que podrían verse como un indicador precursor a la erupción volcánica. Finalmente, el estudio de la serie de Torreperogil ha mostrado algunos resultados no presentes en los otros conjuntos de datos, que pueden sugerir la existencia de factores externos. No obstante, tal afirmación debe ser más investigada basándose en el análisis comparativo con conjuntos de datos similares.

En cuanto a la aplicación a conjuntos de excursión definidos por excedencias de umbrales en campos aleatorios, realizamos un estudio preliminar usando las técnicas anteriores sobre las variaciones de características estructurales con respecto al incremento del umbral, y dependiendo de las propiedades de variabilidad local y rango de dependencia, basado sobre los patrones puntuales espaciales definidos por centroides de las componentes conexas y por *A-exit points*. Los resultados muestran que el análisis multifractal y las medidas de complejidad constituyen un enfoque complementario útil para el estudio de características estructurales de excedencias de umbrales espaciales.

Se indican diversas líneas abiertas y direcciones adicionales para la continuación de la investigación a partir de este trabajo de tesis. Estás involucran, en particular, nuevos aspectos analíticos y generalizaciones, incorporación de errores y extensión en los datos, evaluación de direccionalidad y analysis de sensibilidad.

Introduction

This thesis work has been developed in the framework of granted projects P08-FQM-03834 of Consejería de Innovación, Ciencia y Empresa, Junta de Andalucía, and MTM2009-13250 and MTM2012-32666 of Ministerio de Economía y Competitividad, Spain. The title of the thesis, 'Analysis of fractality characteristics in spatio-temporal processes. Applications to geophysical data', makes reference to the study of geophysical processes with scaling properties in space and time. The focus of this work is the study of dimensional interaction and analysis of structural characteristics at different scales of spatio-temporal processes; in particular, we analyze some seismic data sets. The techniques used are based on entropy-complexity related measures. Besides, the study is complemented with the analysis of the frequency-magnitude distribution. Different methods and tools are developed to quantify and visualize specific features of seismic patterns, and some applications to relevant real seismic series are performed. In addition, multifractal tools are applied to the analysis of structural characteristics of excursion sets from simulated random fields.

Motivation and Background

Throughout history, the human being has always observed nature trying to understand it. From the first philosophers to the modern scientists, natural phenomena (rain, tsunamis, wind, earthquakes, geysers, volcanos, etc.) have always been widely studied. These phenomena can cause, among other consequences, a variety of immediate and long-term environmental impacts related to the geophysical medium transformations as well as effects on living individuals and systems. Over the last decades, the consequences of earthquakes on the inhabitants and the environment where they occur have led to a considerable effort in their study. Such increasing interest in understanding earthquakes jointly with their close relationship with statistics (see, as a reference, the works about statistical seismology by Vere-Jones 2000, 2001, 2010) and their scaling behaviour have motivated us to focus on this geophysical phenomenon.

An earthquake is the result of a movement in the earth's crust caused by a sudden release of the accumulated stress along fault planes. This stress release can be associated with different causes, such as the dynamics of tectonic plates (which is usually the cause of most earthquakes), volcanic eruptions or hydrologic dynamics. Different forms of measurement reveal the

high structural complexity of the phenomenon. Beyond the general geophysical laws, knowledge about the spatio-temporal dynamics characterizing regional seismicity regimes is gained through observations provided by continuous recordings based on a global network of stations, from which synthesized imputations of seismic events are determined in terms of space-time coordinates and magnitude with a certain precision depending on geophysical and technical factors. In some studies, information from covariates is considered with the aim of establishing possible interrelations; for instance, variations in certain variables (e.g. gas, electromagnetic wave energy or acoustic emissions, temperature changes, etc.) are considered as potentially precursory elements of large shocks (see, for example, Kalimeri et al. 2008).

In this thesis work, we concentrate on the statistical exploitation of information from regional space-time series of seismic events (i.e. the registered earthquakes with magnitude above a certain level), without consideration of additional geophysical information nor information from covariates. We are then interested in the intrinsic structural complexity of the spatio-temporal evolution patterns of seismic occurrences within a certain region and a certain period of time.

Among other characteristics, two aspects are of special significance: scaling properties and dimensional interaction. Specifically, seismic occurrences are characterized by multifractality and complex space-time structuring in seismic event patterns, also depending on levels of magnitude. Increasing knowledge on such aspects, including evolutionary changes, is interesting to understand the space-time dynamics of the phenomenon, and can be of potential importance for predictability and risk assessment.

Different methods and tools have been developed to quantify and visualize specific features from seismic patterns. Many of them are essentially based on information and entropy concepts and measures in relation to organization and uncertainty. In this context, some earthquake forecasting methods are based on the concept of the entropy score; a review can be found in Harte and Vere-Jones (2005). Furthermore, the usefulness of entropy to describe seismic phenomena can be seen in the relation between the *b* value of Gutenberg-Richter law and Shannon entropy (De Santis et al. 2011). Some applications of Shannon entropy to earthquakes can be found in Main and Burton (1984), Telesca et al. (2004) or De Santis (2010), among other references. Besides entropy measures, complexity measures have proved to be useful tools for studying the structural properties of seismic events. From a physical point of view, complexity measures are intended to quantify the ability of a system for organization; also to reflect richness of behaviour. In the literature there are diverse measures of 'complexity', which partially capture related features under different approaches. Interesting reviews of complexity measures can be found in Chelidze and Matcharashvili (2007) and Martín et al. (2006).

As mentioned, a variety of geophysical phenomena, among others, are characterized by certain forms of scaling behaviour intrinsic to spatiotemporal dynamics. We must note that the dynamics of earthquakes and many natural phenomena is governed by scaling powerlaws related to asymptotic self-similarity (fractality) of possibly different orders (see, for example,

Bak et al. 1987, Bak and Tang 1989). State distributions associated with such systems are formally described in terms of multifractal measures characterized by the presence of different local Hölder exponents. Numerous authors have shown the fractal structure of earthquake distribution in time, space and magnitude (Kagan and Knopoff 1980, Sadovskiy et al. 1984) and it has been acknowledged that seismicity may be described and interpreted in terms of fractal distributions with power law scaling (see, for instance, Hirata 1989, Caneva and Smirnov 2004, Öncel and Wilson 2007, Roy et al. 2011). Generalized dimensions and multifractal spectrum, based on the scaling powerlaw limiting behaviour of global and local Rényi entropies for the distributions associated with realizations at different resolution scales, constitute the main tools for multifractal analysis (see, for example, Harte 2001).

In this study, we also consider the first tools used to describe seismicity, namely the scaling empirical laws. Some of these laws appeared long time ago but they are still a reference due to their usefulness to study seismic phenomena, along with the fact that they have inspired several techniques. Among these laws, the most important ones are: Omory (1894), Gutenberg-Richter (1944) and Båth (1965) laws. The latter provides a frequency-magnitude distribution of earthquakes and it is the most widely used law. Nevertheless, the non-additive character of the process of stick-slip between tectonic plates (Sotolongo-Costa 2012) has led us to focus in this thesis on the non-extensive frequency-magnitude distribution of earthquakes formerly proposed in Sotolongo-Costa and Posadas (2004).

Objectives and Methodological Aspects

As mentioned above, in this research, the emphasis is mainly on two aspects: the study of dimensional interaction, and the analysis of structural characteristics at different scales. More specifically, the aims of this thesis work are concisely:

- Study of scaling behaviour.
- Analysis of dimensional interaction.
- Assessment on evolutionary changes.
- Incorporation of effects through distributions.
- Application to real seismic data of different nature, and discussion on differential aspects.
- Application to structural analysis of random field excursion sets.

We propose various techniques based on entropy combined with some methodological aspects. Firstly, to describe the interaction between the components of a point process we propose two approaches: the first one consist of studying the degree of dependence by means of entropy

measures based on the extensive or non-extensive expression of the measure in the case of independence; the second one is based on a dependence measure constructed following the former idea of assessment for independence based on entropy considering its limiting behaviour. In particular, a 'dependence coefficient' is defined based on the information dimension (Shannon case).

To study the structural characteristics of dynamical systems at different scales we introduce some technical extensions based on the limiting behaviour of information and complexity related quantities. Specifically, we propose a complexity measure based on the limiting behaviour of the LMC complexity measure, establishing its connection with the generalized Rényi dimensions. Besides, we define a version for generalized dimensions based on the non-extensive statistics of Tsallis.

In the study of seismic series it is common to focus on the spatio-temporal distribution of earthquakes. In some cases it is also of interest to consider the effects of other components, such as the magnitude. With this aim, we propose a practical way to incorporate these effects to the techniques considered, consisting of the use of a weighting function for obtaining the box-counting distributions.

Along with the proposed techniques, we have considered complementary methodological procedures. Firstly, for a better assessment of the spatio-temporal interaction we combine the exposed techniques considering different transformations on the temporal component. Secondly, temporal evolutionary changes are studied performing the different analyses in data segments by means of sliding windows. Finally, the fitting of the non-extensive frequency-magnitude distribution is carried out by weighted least-squares method, considering a weighting function based on the relation between magnitude and energy to improve the estimation of the parameters.

The former techniques and methodological aspects are applied to three real seismic series localized in the areas of Agrón, El Hierro and Torreperogil in Spain. The study of the series non only illustrates the above technical and methodological proposals, but they are especially interesting because of the relevant particular aspects underlying their nature.

In addition, we introduce the assessment on structural characteristics of excursion sets determined by threshold exceedances, by the study of the spatial point patterns defined by centroids and A-exit points, using complexity and multifractality tools, for increasing threshold and varying the properties of the underlying process.

Numerical implementation has been performed creating font code in **MATLAB** language, especially appropriate for matrix manipulation, function and data representation, as well as algorithm implementation. Random field realizations have been generated using **R** library *RandomFields* written by Martin Schlather. The representations of the data epicenters on contoured maps have been performed with **SURFER** software. The main font codes developed in this thesis are included in the Appendix.

Outline

This thesis work is structured in three parts:

Part I. Preliminaries

Part I presents a short review of the main approaches and tools used here to analyze the dynamics of earthquakes. This part is divided into two chapters. In Chapter 1 we introduce basic ideas about information theory and complexity measures. Here, we refer to the Shannon, Rényi and Tsallis entropies and their main characteristics, also including the concepts of divergence and mutual information related to the three entropy measures. Besides, we introduce the 'box-counting' methodology used for practical calculation of these measures, together with an example composed of two data sets which have been created to illustrate the usefulness of these techniques. The chapter concludes with the study of 'complexity' measures, focusing on the LMC complexity measure and some generalizations and alternatives. Chapter 2 consist of three sections. In the first section, we introduce the concepts of fractal dimension and scale invariance. The multifractal formalism in terms of the generalized Rényi dimensions and the multifractality spectrum is introduced in the second section following the 'lattice box-counting approach'. Finally, in section 3 we present the three most important scaling laws used to describe earthquakes, and a frequency-magnitude distribution (FMD) based on a non-extensive approach introduced Sotolongo-Costa and Posadas (2004) and the re-analized version by Silva et al. (2006).

Part II. Methodology and Applications

Part II is divided into two chapters. In Chapter 3, we introduce some technical extension with the aim of improving the study of seismicity as well as achieving the established objectives. To asses the interaction between components, we propose a study of the dimensional dependence in terms of entropy measures. This methodology takes into account whether the entropy measure possesses or not the extensivity property. In the context of the limiting behaviour (for box size decreasing to 0), we introduce a dependence measure based on generalized dimensions to analyze the spatio-temporal interaction. Moreover, we introduce a formulation of generalized dimensions based on the non-extensive statistics of Tsallis and a complexity measure is proposed in terms of the generalized dimensions. We introduce a weighting function in order to incorporate the effects of the magnitude through box-counting distributions. Further, we combine the former techniques with different transformations on the temporal component with the aim of evaluating its effect on the structural complexity displayed by the data. We describe the sliding windows technique to assess temporal evolutionary changes. Finally, we introduce a weighting function based on the relation between magnitude and energy to perform

the fitting of the non-extensive frequency-magnitude distribution.

In Chapter 4, three relevant seismic data sets of different nature are studied using the methodological approaches and technical proposals introduced. The first data set is associated with tectonic activity in the seismic zone of Agrón (Granada, Spain), and consist of a seismic series registered between October 1987 and September 1990. Here we are particulary interested in the assessment of evolutionary changes between regular and high activity periods. The second one is a well-known seismic sequence recently occurred in the volcanic island of El Hierro (Canary Islands, Spain). This sequence is especially interesting because of the large number of events occurred, more than 11.000 earthquakes in just 173 days, and also because during this series the famous submarine eruption of 2011 occurred. In this data set the main aim is to assess particular structural changes regarding the volcanic nature of the series. Finally, the last data set analyzed is the seismic series occurred in the area near the village of Torreperogil (Jaén, Spain) between March 2009 and March 2013. This series is also particulary interesting because of the implications in terms of environmental impact on the area. The uncertainty about its origin has produced controversy in the media and, as a consequence, several hypotheses have appeared and are under debate. Thus, in this case the aim is to try to detect whether there is a presence of external factors in the phenomenon.

Part III. Complementary Aspects and Conclusions

The final Part III is composed of only Chapter 5. In this chapter, we firstly introduce a summary, along with a synthetic discussion of the main aspects addressed in this thesis work. Besides, we expose the references of the articles and contributions to meetings derived from this research. Next, we present a brief exposition of the application of the techniques proposed to the study of structural characteristics of excursion sets in random fields. Finally, we point out several directions and open lines for continuing research.

Introducción

Este trabajo de tesis ha sido desarrollado en el marco de los proyectos P08-FQM-03834 de la Consejería de Innovación, Ciencia y Empresa, Junta de Andalucía, y MTM2009-13250 y MTM2012-32666 del Ministerio de Economía y Competitividad. El título de la Tesis “Análisis de características de fractalidad en procesos espacio-temporales. Aplicación a datos geofísicos”hace referencia al estudio de procesos geofísicos con propiedades de escalamiento en el espacio y en el tiempo. Este trabajo se enfoca al estudio de la interacción dimensional y el análisis de características estructurales a diferentes escalas de procesos espacio-temporales; en particular, se consideran algunos conjuntos de datos sísmicos. Las técnicas usadas están basadas en medidas de entropía y complejidad. Además, el estudio se complementa con el análisis de la distribución frecuencia-magnitud. Se desarrollan diferentes métodos y herramientas para cuantificar y visualizar rasgos específicos de patrones sísmicos y se elaboran algunas aplicaciones a series sísmicas reales relevantes. Además, se aplican herramientas multifractales al análisis de las características estructurales de conjuntos de excursión a partir de campos aleatorios simulados.

Motivación y Contexto

A lo largo de la historia, el ser humano ha observado siempre la naturaleza con el fin de entenderla. Desde los primeros filósofos hasta los científicos modernos, los fenómenos naturales (la lluvia, los tsunamis, el viento, los terremotos, los volcanes, etc.) han sido siempre ampliamente estudiados. Estos fenómenos pueden causar, entre otras consecuencias, una variedad de impactos medioambientales a largo plazo relacionados con transformaciones en el medio geofísico además de efectos en los seres vivos y los sistemas. En las últimas décadas, las consecuencias de los terremotos en los habitantes y el entorno en el que se producen han llevado a un aumento considerable en su estudio. Este interés creciente por entender los terremotos junto con su cercana relación con la estadística (ver como referencia los trabajos sobre sismología estadística de Vere-Jones 2000, 2001, 2010) y su estructura de escalamiento nos ha motivado a enfocar el estudio en este fenómeno geofísico.

Un terremoto es el resultado de un movimiento en la corteza terrestre causado por una liberación repentina de la energía acumulada en los planos de fallas. Esta liberación de energía puede

asociarse a diferentes causas tales como la dinámica de las placas tectónicas (que es generalmente la causa de la mayoría de los terremotos), erupciones volcánicas o dinámicas hidrológicas. Diferentes formas de medición revelan la gran complejidad estructural del fenómeno. Más allá de las leyes geofísicas generales, el conocimiento sobre las dinámicas espacio-temporales que caracterizan a la sismicidad regional se obtenie mediante observaciones proporcionadas por el registro continuo a partir de una red global de estaciones sísmicas. En algunos estudios, se considera la información de covariables para establecer posibles interrelaciones; por ejemplo, variaciones en ciertas variables (e.g. emisiones de gas, de energía de ondas electromagnéticas o acústicas, cambios de temperatura, etc.) se consideran como elementos particularmente precursores de grandes terremotos (ver, por ejemplo, Kalimeri et al. 2008).

En este trabajo de tesis nos centramos en la explotación estadística de la información proveniente de series sísmicas espacio-temporales (es decir, los terremotos registrados con magnitud por encima de un cierto nivel), sin considerar información geofísica adicional o información obtenida de covariables. Estamos, pues, interesados en la complejidad estructural inherente a los patrones de evolución temporal de eventos sísmicos dentro de una región y un período de tiempo dados.

Entre otras características, dos aspectos son de especial relevancia: propiedades de escalamiento e interacción dimensional. Específicamente, las ocurrencias sísmicas se caracterizan por tener multifractalidad y una estructura espacio-temporal compleja de los patrones de eventos sísmicos, también, dependencia de los niveles de magnitud. Un mayor conocimiento de dichos aspectos, incluyendo cambios evolucionarios, es interesante para entender la dinámica espacio-temporal del fenómeno y puede se de potencial importancia para la previsibilidad y la evaluación riesgos.

Diferentes métodos y herramientas han sido desarrollados para cuantificar y visualizar los rasgos específicos de los patrones sísmicos. Muchos de ellos están basados esencialmente en conceptos y medidas de información y entropía en relación con la organización y la incertidumbre. En este contexto, algunos métodos de predicción de terremotos están basados en el concepto de marcador de entropía (Harte y Vere-Jones (2005)). Además, la utilidad de la entropía para describir fenómenos sísmicos puede verse en la relación entre el valor b de la ley de Gutenberg-Richter y la entropía de Shannon (De Santis et al. 2011). Algunas aplicaciones de la entropía de Shannon a los terremotos puede encontrarse en Main y Burton (1984), Telesca et al. (2004) o De Santis (2010), entre otras referencias. Además de las medidas de entropía, las medidas de complejidad han demostrado ser herramientas útiles para estudiar las propiedades estructurales de eventos sísmicos. Desde un punto de vista físico, las medidas de complejidad tienen como objetivo cuantificar la capacidad de organización de un sistema; también, reflejar la riqueza de comportamiento. En la literatura, existen diversas medidas de “complejidad”, que capturan parcialmente rasgos relacionados bajo diferentes enfoques. Revisiones interesantes sobre medidas de complejidad puede encontrarse en Chelidze y Matcharashvili (2007) y en Martín et al. (2006).

Como mencionamos anteriormente, gran variedad de fenómenos geofísicos están caracterizados por ciertas formas de comportamiento de escalamiento intrínseco a la dinámica espacio-temporal. Debemos señalar que la dinámica de los terremotos y de muchos fenómenos naturales está gobernada por leyes potenciales de escalamiento relacionadas con la auto-similitud (fractalidad) de diferentes órdenes (ver, por ejemplo, Bak et al. 1987, Bak y Tang 1989). Las distribuciones de los estados asociadas a dichos sistemas se describen formalmente mediante medidas multifractales caracterizadas por la presencia de diferentes exponentes de Hölder locales. Numerosos autores han mostrado la estructura fractal de la distribución de los terremotos en el tiempo, espacio y magnitud (Kagan y Knopoff 1980, Sadovskiy et al. 1984) y se ha reconocido que la sismicidad puede ser descrita e interpretada mediante distribuciones fractales con leyes potenciales de escalamiento (ver, por ejemplo, Hirata 1989, Caneva y Smirnov 2004, Öncel y Wilson 2007, Roy et al. 2011). Las dimensiones generalizadas y el espectro multifractal, basados en el comportamiento límite de la ley potencial de escalamiento de las entropías locales y globales de Rényi para las distribuciones asociadas con realizaciones a diferentes escalas de resolución, constituyen las herramientas principales para el análisis multifractal (ver, por ejemplo, Harte 2001).

También consideramos en este estudio las primeras herramientas usadas para describir sismicidad, a saber, las leyes empíricas de escalamiento. Algunas de estas leyes aparecieron hace mucho tiempo, pero sirven todavía de referencia debido a su utilidad para estudiar los fenómenos sísmicos, junto con el hecho de que han inspirado diversas técnicas. Entre estas leyes, las más importantes son las de Omory (1894), Gutenberg-Richter (1944) y Båth (1965). Esta última describe la distribución frecuencia-magnitud de los terremotos y es la ley empírica más usada. No obstante, el carácter no aditivo del proceso de empuje-deslizamiento entre placas tectónicas (Sotolongo-Costa 2012) nos lleva a considerar en esta tesis la distribución no extensiva frecuencia-magnitud de terremotos anteriormente propuesta en Sotolongo-Costa y Posadas (2004).

Objetivos y Aspectos Metodológicos

Como mencionados más arriba, este trabajo se centra principalmente en dos aspectos: el estudio de la interacción dimensional y el análisis de las características estructurales a diferentes escalas. En particular, los objetivos específicos de este trabajo de tesis son, de forma concisa:

- Estudio del comportamiento de escalamiento.
- Análisis de la interacción dimensional.
- Evaluación de cambios evolucionarios.
- Incorporación de efectos a través de distribuciones.

- Aplicación a datos sísmicos reales de diferente naturaleza, y discusión sobre aspectos diferenciales.
- Aplicación al análisis estructural de conjuntos de excursión en campos aleatorios.

Proponemos diversas técnicas basadas en la entropía combinadas con algunos aspectos metodológicos. En primer lugar, para describir la interacción entre los componentes de un proceso puntual proponemos dos enfoques: el primero consiste en estudiar el grado de dependencia mediante medidas de entropía basadas en la expresión extensiva o no extensiva de la medida en el caso de independencia; el segundo enfoque está basado en una medida de dependencia construida siguiendo la idea anterior considerando su comportamiento límite. En particular, se define un "coeficiente de dependencia"basado en la dimensión de información (caso de Shannon).

Para estudiar las características estructurales de sistemas dinámicos a diferentes escalas, introducimos algunas extensiones técnicas basadas en el comportamiento límite de medidas relativas a información y complejidad. Específicamente, proponemos una medida de complejidad basada en el comportamiento límite de la medida de complejidad LMC, estableciendo su conexión con las dimensiones generalizadas de Rényi. Además, definimos una versión de las dimensiones generalizadas basada en la estadística no-extensiva de Tsallis.

En el estudio de series sísmicas es común centrarse en la distribución espacio-temporal de terremotos. En algunos casos es interesante considerar los efectos de otras componentes, como, por ejemplo, la magnitud. Con este objetivo, proponemos una vía práctica para incorporar dichos efectos a las técnicas consideradas consistente en el uso de una función de pesos para obtener las distribuciones del conteo por cajas.

Junto con las técnicas propuestas, hemos considerado algunos procedimientos metodológicos complementarios. En primer lugar, para una mejor evaluación de la interacción espacio-temporal combinamos las técnicas expuestas considerando diferentes transformaciones en la componente temporal. En segundo lugar, se estudian cambios evolucionarios temporales realizando los diferentes análisis a segmentos de datos obtenidos mediante ventanas deslizantes. Finalmente, el ajuste de la versión no extensiva de la distribución frecuencia-magnitud se lleva a cabo mediante el método de mínimos cuadrados ponderados, considerando una función de pesos basada en la relación entre magnitud y energía para mejorar las estimaciones de los parámetros.

Las técnicas y aspectos metodológicos anteriores se aplica a tres series sísmicas reales localizadas en las áreas de Agrón, El Hierro y Torreperogil en España. El estudio de estas series no sólo ilustra las propuestas técnicas y metodológicas anteriores, sino que son especialmente interesantes por los particulares aspectos relevantes subyacentes a su naturaleza.

Adicionalmente, introducimos la evaluación de características estructurales de conjuntos de excursión determinados por excedencia de umbrales, a través del estudio de los patrones puntuales espaciales definidos por centroides y *A-exit points*, usando herramientas de complejidad y multifractalidad, para umbrales crecientes y variando las propiedades del proceso subyacente.

La implementación numérica ha sido realizada creando código fuente en lenguaje **MATLAB**, especialmente apropiado para manipulación de matrices, representación de funciones y de datos, así como implementación de algoritmos. Realizaciones de campos aleatorios han sido generadas usando la librería *RandomField* escrita en **R** por Martin Schlather. Las representaciones de los epicentros en los mapas han sido realizadas mediante el software **SURFER**. Los principales códigos fuente desarrollados en esta tesis se incluyen en el Apéndice.

Líneas generales

Este trabajo de tesis está estructurado en tres partes.

Parte I. Preliminares

La Parte I presenta una revisión breve sobre los principales enfoques y herramientas usadas en este trabajo para analizar la dinámica de los terremotos. Esta parte está dividida en dos capítulos. En el Capítulo 1, introducimos ideas básicas sobre la teoría de la información y medidas de complejidad. Aquí, nos referimos a las entropías de Shannon, Rényi y Tsallis y sus principales características, incluyendo también los conceptos de divergencia e información mutua relativos a las tres entropías. Además, introducimos la metodología del conteo por cajas, usada para el cálculo práctico de estas medidas, junto con un ejemplo compuesto por dos conjuntos de datos creados para ilustrar la utilidad de estas técnicas. Este capítulo concluye con el estudio de las medidas de complejidad, centrándonos principalmente en la medida de complejidad LMC y algunas generalizaciones y alternativas. El Capítulo 2 consta de tres secciones. En la primera sección, introducimos los conceptos de dimensión fractal e invariancia de escala. El formalismo multrifractal basado en las dimensiones generalizadas de Rényi y el espectro multifractal se introduce en la segunda sección, siguiendo el enfoque de conteo por cajas sobre una malla. Finalmente, en la Sección 3 presentamos las leyes de escalamiento más importantes usadas para describir terremotos, y una distribución frecuencia-magnitud (FMD) basada en un enfoque no extensivo introducida por Sotolongo-Costa y Posadas (2004) y la versión revisada de Silva et al. (2006).

Parte II. Metodología y Aplicaciones

La Parte II está dividida en dos capítulos. En el Capítulo 3, introducimos algunas extensiones técnicas con el objetivo de mejorar el estudio de la sismicidad además de alcanzar los objetivos establecidos. Para evaluar la interacción entre componentes, proponemos un estudio de la dependencia dimensional en términos de medidas de entropía. Esta metodología tiene en cuenta si la medida de entropía posee o no la propiedad de extensividad. En el contexto del comportamiento límite (para tamaños de caja decreciendo a 0), introducimos una medida

de dependencia basada en las dimensiones generalizadas para analizar la interacción espacio-temporal. Además, introducimos una formulación de las dimensiones generalizadas basada en la estadística no extensiva de Tsallis y se proponemos una medida de complejidad basada en las dimensiones generalizadas. Introducimos una función de pesos con el fin de incorporar los efectos de la magnitud a través de las distribuciones de conteo por cajas. Adicionalmente, combinamos las técnicas anteriores con diferentes transformaciones en la componente temporal con el objetivo de evaluar su efecto sobre la complejidad estructural mostrada por los datos. Definimos la técnica de ventana deslizante para evaluar los cambios evolucionarios temporales. Finalmente, introducimos una función de pesos basada en la relación entre magnitud y energía para realizar el ajuste la la distribución no extensiva frecuencia-magnitud.

En el Capítulo 4, estudiamos tres series sísmicas de diferente naturaleza usando las propuestas técnicas y metodológicas introducidas. El primer conjunto de datos está asociado con actividad tectónica en la zona sísmica de Agrón (Granada, España), y consiste en una serie sísmica registrada entre octubre de 1987 y septiembre de 1990. En este caso estamos particularmente interesados en la evaluación de los cambios evolucionarios entre periodos de actividad sísmica regular y alta. El segundo conjunto está constituido por la conocida serie sísmica ocurrida recientemente en la isla volcánica de El Hierro (Islas Canarias, España). Esta secuencia es especialmente interesante debido al gran número de eventos ocurridos, más de 11.000 terremotos en sólo 173 días, y también debido a la erupción volcánica ocurrida en 2011 durante esta serie. En este conjunto de datos el principal objetivo es evaluar cambios estructurales particulares relacionados con la naturaleza volcánica de la serie. Finalmente, el último conjunto de datos analizado es la serie sísmica ocurrida en el área cercana a la localidad de Torreperogil (Jaén, España) entre marzo de 2009 y marzo de 2013. Esta serie es también interesante, debido a las implicaciones en lo que respecta al impacto medioambiental en la zona. La incertidumbre acerca de su origen ha producido controversia en los medios de comunicación y, como consecuencia, varias hipótesis han surgido y están en debate. En este caso, pues, el objetivo es tratar de detectar la posible presencia de factores externos en el fenómeno.

Parte III. Aspectos Complementarios y Conclusiones

La Parte final III consta sólo de del Capítulo 5. En este capítulo, primero introducimos un resumen, junto con una breve discusión sintética de los principales aspectos abordados en este trabajo de tesis. Además, exponemos las referencias de los artículos y contribuciones a congresos derivados de esta investigación. A continuación, presentamos una breve exposición de la aplicación aplicación de las técnicas propuestas al estudio de características estructurales de conjuntos de excursión en campos aleatorios. Finalmente, señalamos varias direcciones y líneas abiertas para continuación de la investigación.

Part I

PRELIMINARIES

Chapter 1

Entropy and Complexity

To understand the behaviour of dynamical systems we need to study their structural properties, being predictability an important one. In 1948, Shannon defined the concept of entropy as the average unpredictability in a random variable, which is equivalent to its information content. Then, information measures quantify the uncertainty existing in a probability distribution. In this sense, information theory provides a basis to measure the degree of chaos or disorder in the behaviour of a dynamical system. In the context of seismicity, some earthquake forecasting methods are based on the concept of the entropy score; a review is given in Harte and Vere-Jones (2005). The usefulness of entropy measures to describe seismic phenomena can be seen in the relation between the *b* value of the Gutenberg-Richter law (Chapter **??**) and Shannon entropy (De Santis et al., 2011). Some applications of Shannon entropy to earthquakes can be found, for example, in Main and Burton (1984), De Santis (2010) or Telesca et al. (2004). In addition, self-similarity observed in the dynamics of earthquakes has motivated the use of generalized entropies to describe the structural features at different scales. In particular, multifractal analysis tools (Chapter 2), based on the limiting behaviour of Rényi entropies, are powerful instruments to study theoretically complex systems.

A further step in recent research for characterization of the structural properties of a system has been focused on the development of complexity measures. 'Complexity' is partially and synthetically related to concepts such as uncertainty, randomness, disorder and delocalization. From a physical point of view, complexity measures are intended to quantify the ability of a system for organization; also, to reflect richness of behaviour. Generally, statistical complexity measures are jointly based on information entropy and divergence in some way.

The chapter is structured in two sections. In section 1.1, we present a short review of information theory concepts and elements. Here, we introduce Shannon, Rényi and Tsallis entropies, which are the most widely used entropies, and their main characteristics. Besides, the concepts of divergence and mutual information related to the three entropies are also exposed. We finish the section introducing the 'box-counting' technique used to carry out the calculus of

these measures on point patterns; we illustrate these tools by means of an example composed of two simple data sets created to study the usefulness of these techniques. Section 1.2 deals with 'complexity'. In this section, we expose one of the most used complexity measures, the LMC complexity, as well as some well-known generalizations and alternatives.

1.1 Uncertainty and Information Theory

One of the main aspects that have motivated a considerable interest in describing dynamical systems is predictability. Let us assume that we try to predict the result of a random experiment. This task may be difficult to achieve, but it does not mean that we are in a situation of complete ignorance. We can associate with one potential realization of the experiment a probability distribution which includes all the information that we have about it. In this sense, the aim is to quantify the degree of uncertainty that we have when trying to carry out predictions. Information theory provides measures which are intend to quantify the uncertainty in a dynamical system; these are known as entropy measures. The problem of constructing a measure which assesses the degree of uncertainty a system was proposed and solved by Claude E. Shannon in 1948. In this work, Shannon introduced three simple axioms that a measure has to comply to define univocally the uncertainty and he proposed a measure which consistent with these axioms. Later, other measures have appeared as generalizations of this one and some of them have proved to be particulary useful, specially Rényi and Tsallis entropies. The first one keeps the additivity property, whereas the second one is pseudoadditive.

1.1.1 Shannon entropy

Shannon (1948) established three axioms which a measure of uncertainty or disorder must comply to be defined univocally. Let $\bar{p} = (p_1, ..., p_N)$ be a probability distribution and $H(p_1.., p_N)$ ($H : \mathbb{R}^N \to \mathbb{R}$) a quantitative measure. H means the uncertainty that exists in a probability distribution if:

- H is a continuous function of $\bar{p}$.
- $H(1/N, ..., 1/N) = H\left(\left[\frac{\overline{1}}{N}\right]\right)$ is a monotonically increasing function on N.
- H satisfies that

$$\begin{aligned} H(p_1, ..., p_N) &= H(p_1, ..., p_{N-2}, p_{N-1} + p_N) \\ &+ (p_{N-1} + p_N) H\left(\frac{p_{N-1}}{p_{N-1} + p_N}, \frac{p_N}{p_{N-1} + p_N}\right). \end{aligned}$$

The first axiom is to ensure a good mathematical behaviour. The second one establishes that the uncertainty in the equiprobable case increases when so does the number of events. Finally, the third axiom determines that the uncertainty associated with a random system is equal to the weighted sum of the uncertainties associated with the different subsystems which it can be divided into. In the same work, the author introduced a measure which satisfies these axioms, known as Shannon entropy.

Definition 1.1.1. *Shannon entropy* of a discrete finite probability distribution, with mass probability function $\bar{p} = (p_1, p_2, \ldots, p_N)$, is defined as

$$H(\bar{p}) := -K \sum_{i=1}^{N} p_i \log(p_i) = -KE[\log(\bar{p})]. \tag{1.1}$$

where K is a positive constant.

The constant K fixes a scale and, because of that, we can consider $K = 1$ without loss of generality. The base of the logarithm is arbitrary but the most extended ones are 2 and e. The first one is usual when we work with a binary probability distribution while the second one simplifies the mathematical operations. In this work, we operate with the second choice. Then, we denote

$$H(\bar{p}) = -\sum_{i=1}^{N} p_i \ln(p_i) = -E[\ln(\bar{p})]. \tag{1.2}$$

From the definition, $H(\bar{p}) \geq 0$, $H(\bar{p})$ has concave shape on $\bar{p}$, and it satisfies the extensivity property, or additivity law for independent subsystems:

$$H(\bar{p_1}, \bar{p_2}) = H(\bar{p_1}) + H(\bar{p_2}).$$

In order to determine the maximum possible entropy value we just consider the constraint $\sum_{i=1}^{N} p_i = 1$. The probability function that maximizes the entropy is the equiprobable probability distribution ($p_i = \frac{1}{N},\ i = 1, ..., N$), thus:

$$H^{max}(\bar{p}) = H\left(\left[\frac{1}{N}\right]\right) = -\sum_{i=1}^{N} \frac{1}{N} \ln\left(\frac{1}{N}\right) = -\ln\left(\frac{1}{N}\right) = \ln(N).$$

Shannon entropy has the problem of its lack of delimitation. For this reason, some authors use the configuration entropy, which is a normalization for $[0, 1]$ defined as

$$H^*(\bar{p}) = \frac{H(\bar{p})}{H^{max}(\bar{p})} = \frac{-\sum_{i=1}^{N} p_i \ln(p_i)}{\ln(N)}.$$

Shannon configuration entropy is interpreted as the degree of 'disorder' and it is used in subsection 1.1 in the definition of 'SDL complexity'.

1.1.2 Other measures of entropy

Different authors have tried to generalize Shannon entropy, relaxing its axiomatic. These measures do not possess some the of good mathematical properties characterizing Shannon entropy but they have a significant interest and usefulness. The most extended generalizations are Rényi (Rényi, 1961) and Tsallis (Tsallis, 1988) entropies.

Definition 1.1.2. *Rényi entropy of order q,* $H_q(\overline{p})$, of a discrete finite probability distribution with a probability mass function $\overline{p} = (p_1, p_2, \ldots, p_N)$, is defined as

$$H_q(\overline{p}) := \frac{1}{1-q}\ln\left(\sum_{i=1}^{N} p_i^q\right) = \frac{1}{1-q}\ln\left(E[\overline{p}^{q-1}]\right) \qquad (q \neq 1) \tag{1.3}$$

$$H_1(\overline{p}) := -\sum_{i=1}^{N} p_i \ln(p_i) = -E[\ln(\overline{p})] = \lim_{q\to 1} H_q(\overline{p}). \tag{1.4}$$

Then, Rényi entropy contains as a particular case Shannon entropy, when $q = 1$. The main properties of this entropy are that it preserves the extensivity property, or additivity law for independent subsystems, and it is a non-increasing function of q. For $q = 2$, the negative exponential of H_q is known as the 'energy', $E = \sum_{i=1}^{N} p_i^2$, also named Onicescu information measure (Onicescu 1966). The maximum Rényi entropy is obtained for an equiprobable distribution and coincides the maximum Shannon entropy for all values of q:

$$H_q^{max} = H_q\left(\left[\overline{\frac{1}{N}}\right]\right) = \frac{1}{1-q}\ln\left(\sum_{i=1}^{N}\left(\frac{1}{N}\right)^q\right) = \frac{1}{1-q}\ln\left(N\left(\frac{1}{N^q}\right)\right) = \frac{1}{1-q}\ln\left(N^{1-q}\right) = \ln(N).$$

The Rényi configuration entropy is defined as

$$H_q^*(\overline{p}) = \frac{H_q(\overline{p})}{H_q^{max}(\overline{p})} = \frac{\frac{1}{1-q}\ln\left(\sum_{i=1}^{N} p_i^q\right)}{\ln(N)}.$$

Definition 1.1.3. *Tsallis entropy of order q,* $T_q(\overline{p})$, of a discrete finite probability distribution with a probability mass function $\overline{p} = (p_1, p_2, \ldots, p_N)$, is defined as

$$T_q(\bar{p}) := \frac{1}{q-1}\left(1-\sum_{i=1}^{N} p_i^q\right) = \frac{1}{q-1}\left(1-E[\bar{p}^{q-1}]\right) \qquad (q \neq 1) \tag{1.5}$$

$$T_1(\bar{p}) := -\sum_{i=1}^{N} p_i \ln(p_i) = -E[\ln(\bar{p})] = \lim_{q\to 1} T_q(\bar{p}). \tag{1.6}$$

This measure is a non-extensive generalization of Shannon entropy, first defined in 1967 by Havrda and Charvát, but it was somehow forgotten until Tsallis rediscovered it in 1988. Hence, this entropy measure is known in the literature both as Tsallis entropy and Havrda-Charvát entropy. For $q = 1$, Tsallis entropy is also defined as Shannon entropy, its limiting case for $q \to 1$; for $q = 2$, $T_2(\bar{p}) = 1 - e^{-H_2(\bar{p})}$. Formally, $T_q(\bar{p})$ can be seen as a linearization of $H_q(\bar{p})$. Conceptually, however, it has been justified, in particular, in relation to the characterization of complex systems in non-extensive statistical mechanics (Tsallis 2009). Rényi and Tsallis entropies are useful measures for capturing multifractal characteristics (see, for instance, Jizba and Arimitsu 2004, Tsallis 2009). (A comparative discussion of Shannon, Rényi and Tsallis entropies in terms of relevant analytical properties can be found, for example, in Tsallis 2004.) It is important to note that the non-extensive statistical mechanics is based on this generalized entropy measure (see Gell-Mann and Tsallis 2004; Furuichi 2006). The maximum Tsallis entropy is also obtained for an equiprobable distribution, and its value depends on q:

$$T_q^{max} = T_q\left(\left[\frac{1}{N}\right]\right) = \frac{1-\sum_{i=1}^{N}\left(\frac{1}{N}\right)^q}{q-1} = \frac{1-N^{1-q}}{q-1}.$$

Tsallis configuration entropy is then defined as

$$T_q^*(\bar{p}) = \frac{T_q(\bar{p})}{T_q^{max}(\bar{p})} = \frac{\frac{1}{q-1}\left(1-\sum_{i=1}^{N} p_i^q\right)}{\frac{1-N^{1-q}}{q-1}} = \frac{1-\sum_{i=1}^{N} p_i^q}{1-N^{1-q}}.$$

In both Rényi and Tsallis entropies q plays the role of a 'deformation parameter', which enhances in a certain way and extent higher and lower values of the distribution $\bar{p}$, according to the sign of $q-1$ (see Figure 1.1).

1.1.3 Divergences

The concept of divergence was originally introduced to quantify dissimilarity or departure of a probability distribution, $\bar{p}_2$, with respect to another given probability distribution, $\bar{p}_1$. Among

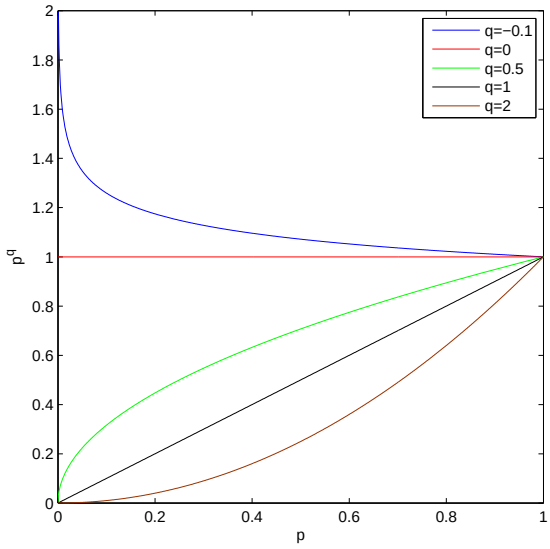

Figure 1.1: Representation of p^q vs. p, depending of q.

the most used divergences are the ones related to the above entropy definitions.

Specifically, the divergence measure related to Shannon entropy is known as Kullback–Leibler divergence, information divergence or relative entropy (Kullback and Leibler 1951). Formally, given two discrete finite probability distributions with probability mass functions $\bar{p_1} = (p_{11}, p_{12}, ..., p_{1N})$ and $\bar{p_2} = (p_{21}, p_{22}, ..., p_{2N})$, respectively, the Kullback-Leibler divergence, denoted $D(\bar{p_1} \| \bar{p_2})$, is defined as follows:

$$D(\bar{p_1} \| \bar{p_2}) := \sum_{i=1}^{N} p_{1i} \ln \frac{p_{1i}}{p_{2i}} = E_{\bar{p_1}} \left[\ln \frac{\bar{p_1}}{\bar{p_2}} \right]. \tag{1.7}$$

The most relevant properties of Kullback-Leibler divergence are:

- Kullback-Leibler divergence is not symmetric. That is, $D(\bar{p_1} \| \bar{p_2}) = D(\bar{p_2} \| \bar{p_1})$ is not necessarily true.
- Kullback-Leibler divergence is always non-negative.
- $D(\bar{p_1} \| \bar{p_2}) = 0 \iff \bar{p_1} \equiv \bar{p_2}$.

Rényi divergence was introduced by Rényi (1961) and it is related to Rényi entropy much like Kullback-Leibler divergence is related to Shannon entropy. In particular, Rényi divergence of order 1 equals Kullback-Leibler divergence. Rényi divergence is defined, for $q \neq 1$, as

$$D_q^R(\bar{p}_1 \| \bar{p}_2) := \frac{1}{q-1} \ln \left(\sum_{i=1}^{N} \frac{p_{1i}^q}{p_{2i}^{q-1}} \right) = \frac{1}{q-1} \ln \left(E_{\bar{p}_1} \left[\left(\frac{\bar{p}_1}{\bar{p}_2} \right)^{q-1} \right] \right). \tag{1.8}$$

The most relevant properties of Rényi divergence are:

- Rényi divergence is not symmetric.
- Rényi divergence is always non-negative.
- $D_q^R(\bar{p}_1 \| \bar{p}_2) = 0 \Longleftrightarrow \bar{p}_1 = \bar{p}_2$.
- For $q \to 1$, $D_q^R(\bar{p}_1 \| \bar{p}_2) \to D(\bar{p}_1 \| \bar{p}_2)$, and hence $D_1^R(\bar{p}_1 \| \bar{p}_2) := D(\bar{p}_1 \| \bar{p}_2)$.

Similarly, Tsallis divergence is defined, for $q \neq 1$, as

$$D_q^T(\bar{p}_1 \| \bar{p}_2) := \frac{1}{1-q} \left(1 - \sum_{i=1}^{N} \frac{p_{1i}^q}{p_{2i}^{q-1}} \right) = \frac{1}{1-q} \left(1 - E_{\bar{p}_1} \left[\left(\frac{\bar{p}_1}{\bar{p}_2} \right)^{q-1} \right] \right), \tag{1.9}$$

and its most relevant properties are:

- Tsallis divergence is not symmetric.
- Tsallis divergence is always non-negative.
- $D_q^T(\bar{p}_1 \| \bar{p}_2) = 0 \Longleftrightarrow \bar{p}_1 = \bar{p}_2$.
- For $q \to 1$, $D_q^T(\bar{p}_1 \| \bar{p}_2) \to D(\bar{p}_1 \| \bar{p}_2)$, and hence, $D_1^T(\bar{p}_1 \| \bar{p}_2) := D(\bar{p}_1 \| \bar{p}_2)$.

1.1.4 Mutual information

Mutual information measures are intended to quantify the degree of dependence existing between two random variables or vectors X and Y, or, in other words, the amount of information they share; the more independent the variables the lesser their mutual information. A natural definition of mutual information between X and Y is based on the divergence of the product of their marginal distributions (that is, their joint distribution under the assumption of mutual independence), with respect to their actual joint distribution. We introduce the definitions of mutual information based on Shannon, Rényi and Tsallis entropies. For $X \sim \bar{p}_X$, $Y \sim \bar{p}_Y$, and $(X,Y) \sim \bar{p}_{XY}$, Shannon mutual information is given by

$$
\begin{aligned}
I(X,Y) &:= D(\bar{p}_{XY} \| \bar{p}_X \bar{p}_Y) \\
&= \sum_{x\in X}\sum_{y\in Y} \bar{p}_{XY} \ln\left(\frac{\bar{p}_{XY}}{\bar{p}_X \bar{p}_Y}\right) \\
&= \sum_{x\in X}\sum_{y\in Y} \bar{p}_{XY} \ln(\bar{p}_{XY}) - \sum_{x\in X}\sum_{y\in Y} \bar{p}_{XY} \ln(\bar{p}_X) - \sum_{x\in X}\sum_{y\in Y} \bar{p}_{XY} \ln(\bar{p}_Y) \\
&= \sum_{x\in X}\sum_{y\in Y} \bar{p}_{XY} \ln(\bar{p}_{XY}) - \sum_{x\in X} \bar{p}_X \ln(\bar{p}_X) - \sum_{y\in Y} \bar{p}_Y \ln(\bar{p}_Y) \\
&= H(X) + H(Y) - H(X,Y).
\end{aligned}
$$

The last equality has a complementary interest for computation and for predictive interpretations. However, such an equivalence is not kept for the cases of Rényi and Tsallis mutual information, and then two alternative formulations can be given in each case. For Rényi mutual information, we denote

$$
\begin{aligned}
I_q^R(X,Y) &:= D_q^R(\bar{p}_{XY} \| \bar{p}_X \bar{p}_Y) && (1.10)\\
\tilde{I}_q^R(X,Y) &:= H_q(X) + H_q(Y) - H_q(X,Y). && (1.11)
\end{aligned}
$$

Similarly, for Tsallis mutual information,

$$
\begin{aligned}
I_q^T(X,Y) &:= D_q^T(\bar{p}_{XY} \| \bar{p}_X \bar{p}_Y) && (1.12)\\
\tilde{I}_q^T(X,Y) &:= T_q(X) + T_q(Y) - T_q(X,Y). && (1.13)
\end{aligned}
$$

For a synthetic view, the entropy, divergence and mutual information measures introduced are summarized in Table 1.1.

1.1.5 A box-counting-based illustration

The primary problem of the implementation of these techniques with experimental data is the determination of the probability distribution $\bar{p}$ of interest associated with the dynamical system under study. In the case of realizations represented by point patterns, box counting is a method widely used to collect measurements by analyzing the local density of a pattern when breaking the original data set into smaller and smaller boxes. The methodology usually consists of choosing a box size (side) ε (assuming a hypercubic box design) and to construct a regular grid on the data set. The probability distribution is then determined based on the points falling within each box.

Table 1.1: Shannon, Rényi and Tsallis entropy measures, and related measures of divergence and mutual information.

Entropy	Shannon	$H(\bar{p}) = -E[\ln(\bar{p})]$
	Rényi	$H_q(\bar{p}) = \frac{1}{1-q}\ln\left(E[\bar{p}^{q-1}]\right) \qquad (q \neq 1)$ $H_1(\bar{p}) = H(\bar{p})$
	Tsallis	$T_q(\bar{p}) = \frac{1}{q-1}\left(1 - E[\bar{p}^{q-1}]\right) \qquad (q \neq 1)$ $T_1(\bar{p}) = H(\bar{p})$
Divergence	Kullback-Leibler	$D(\bar{p}_1 \| \bar{p}_2) = E_{\bar{p}_1}\left[\ln \frac{\bar{p}_1}{\bar{p}_2}\right]$
	Rényi	$D_q^R(\bar{p}_1 \| \bar{p}_2) = \frac{1}{q-1}\ln E_{\bar{p}_1}\left[\left(\frac{\bar{p}_1}{\bar{p}_2}\right)^{q-1}\right] \qquad (q \neq 1)$ $D_1^R(\bar{p}_1 \| \bar{p}_2) = D(\bar{p}_1 \| \bar{p}_2)$
	Tsallis	$D_q^T(\bar{p}_1 \| \bar{p}_2) = \frac{1}{1-q}\left(1 - E_{\bar{p}_1}\left[\left(\frac{\bar{p}_1}{\bar{p}_2}\right)^{q-1}\right]\right) \qquad (q \neq 1)$ $D_1^T(\bar{p}_1 \| \bar{p}_2) = D(\bar{p}_1 \| \bar{p}_2)$
Mutual Information	Shannon	$I(X,Y) = D(\bar{p}_{XY} \| \bar{p}_X \bar{p}_Y) = H(X) + H(Y) - H(X,Y)$
	Rényi	$I_q^R(X,Y) = D_q^R(\bar{p}_{XY} \| \bar{p}_X \bar{p}_Y)$ $\tilde{I}_q^R(X,Y) = H_q(X) + H_q(Y) - H_q(X,Y)$
	Tsallis	$I_q^T(X,Y) = D_q^T(\bar{p}_{XY} \| \bar{p}_X \bar{p}_Y)$ $\tilde{I}_q^T(X,Y) = T_q(X) + T_q(Y) - T_q(X,Y)$

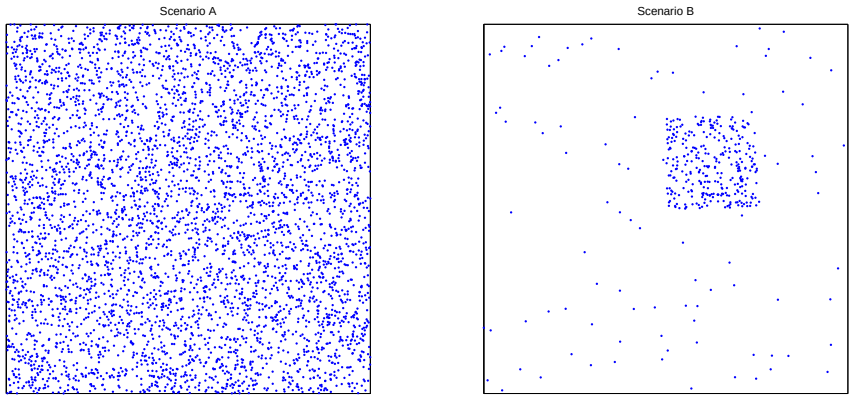

Figure 1.2: Scatter plots for the scenarios A (left) and B (right) of the illustrative example.

In this work, we use this technique to obtain the probability distributions considering different box sizes for griding the study area. For computational convenience and illustrative purpose, in all cases studied we normalize the data set to the unit hypercube. Thus, the study domain has unit sides and the box sizes are usually taken with side $\varepsilon_i = 1/i,\ i = 2,3,4,\ldots$ for the different grids. In each box we measure some type of magnitude. Here, we count the number of events which falls into each box. We construct the basic probability distributions for the considered box sizes by means of the above counts.

To illustrate the above tools we apply these techniques to two artificially constructed cases, see Figure 1.2. To construct the scenarios we divide the area into 16 boxes and in each box we generate points by means of the uniform distribution. Scenario A is constructed with the same characteristics in each box; and this data set is near to the equiprobable distribution. As to scenario B, we generate points for the complete area with a low point density and only one box having high point density.

Figure 1.3 contains the representations of Shannon entropy, and Rényi and Tsallis entropies vs. q, for scenarios A and B considering a box size = $1/10$. (We plot Shannon entropy independent from the parameter q, only as a reference for level comparisons.) In these representations, we can observe the shape of the curves for the different scenarios and how these curves coincide for $q = 1$.

To study better the data sets we represent both scenarios for each entropy measure. Figure 1.4 displays the Shannon entropy values calculated using varying box sizes between $1/2$ and $1/10$. The Shannon entropy values in scenario A are larger than in scenario B, which indicates that the uncertainty in this case is higher for all the box sizes considered. Besides, in this plot

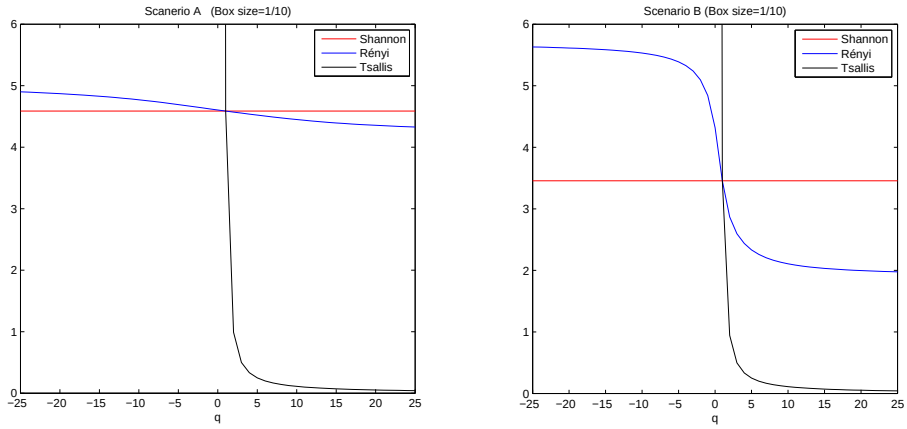

Figure 1.3: Representations of Shannon, Rényi and Tsallis entropies vs. q for scenarios A (left) and B (right), considering a box size = $1/10$.

we can observe how the behaviour of the Shannon entropies with respect to the box sizes is more unstable in scenario B than in scenario A. The generalized measures, Rényi and Tsallis entropies, are respectively represented in the left and right plots of Figure 1.5, considering a box size equal to 1/10, and varying q between -25 and 25. It is important to note, as we mentioned previously, that the values of $q < 1$ emphasize the zones with little representation of points whereas the values of $q > 1$ enhance the zones with larger representation of points. Hence, here we can see larger differences between both scenarios.

In scenario A, the values of Rényi entropy show little variation with respect to q. In contrast, we appreciate significant differences between the Rényi entropy values in scenario B. This shows the sensitivity of Rényi entropies to the variation of the deformation parameter q, depending on the degree of structuring in the pattern.

In scenario A, the values of Rényi entropy for varying q are similar due to the fact that all areas have approximately the same representations, which indicate that the level of disorder it not sensitive to variations in the deformation parameter. In contrast, we appreciate significant differences between the Rényi entropy values in scenario B. In this case, the level of disorder increases for $q < 1$, that is, when the areas with few representation are emphasized, whereas it decreases when the areas with high concentration of points are accentuated, for $q > 1$. It is important to note that these results are obtained for a specific box size. The analysis at different scales, regarding the limiting behaviour, is performed by means of multifractal tools which are based on Rényi entropy, as it will be shown in the next chapter.

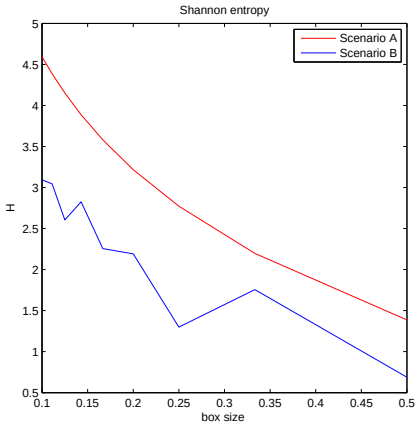

Figure 1.4: Representations of Shannon entropy vs. box size for scenarios A (left) and B (right) using varying box sizes between 1/15 and 1/2.

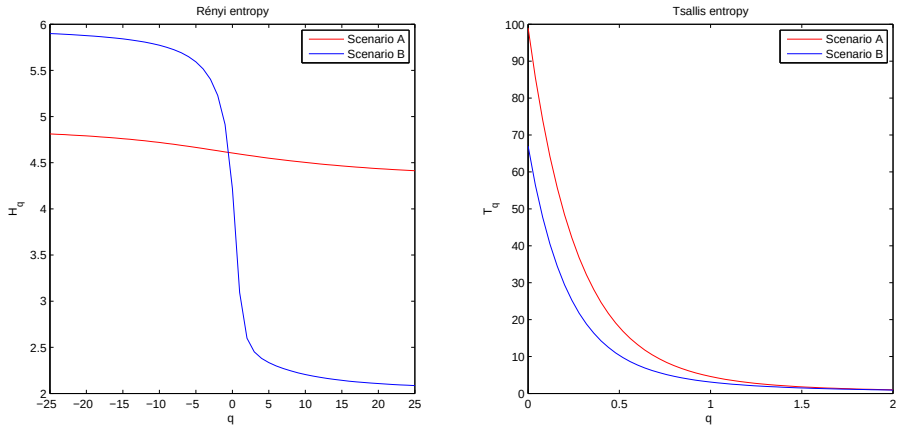

Figure 1.5: Representations of Rényi (left) and Tsallis (right) entropies vs. q for scenarios A and B, considering a box size = $1/10$.

1.2 Complexity Measures

A further step in recent research for the characterization of the structural properties of a system has been focused on the development of complexity measures. In the past two decades, the interest in the study of the complexity of dynamical systems has led to some physical modeling approaches for the analysis of complexity (see, for example, Korvin 1992, Sornette 2000, in the earth sciences; and Goltz 1998, Kanamori and Brodsky 2004, Main 1996, Sadovski and Pisarenko 1991, particularly in seismicity). The concept of 'complexity' is partially and synthetically related to concepts such as uncertainty, randomness, disorder and delocalization. From a physical point of view, complexity measures are intended to quantify the ability of a system for organization; also, to reflect richness of behaviour. In the literature there are several measures of complexity, with different analytical properties and interpretations. In this work, we expose one of the best known, namely the LMC measure of complexity, as well as other related proposals.

1.2.1 LMC measure of complexity

López-Ruiz et al. (1995) proposed a statistical complexity measure which is jointly based on information entropy and divergence in some way. The authors developed this measure considering the perfect crystal (related with order) and the isolated ideal gas (related with equilibrium) as physical examples of systems with complexity zero. In particular, they defined the 'disequilibrium' as some type of distance from the equiprobable distribution. In this context, in the case of the perfect crystal the 'disequilibrium' would be the maximum, and in the case of the isolated ideal gas the 'disequilibrium' would be zero. In this sense, the authors defined a type of complexity measure given by the product of 'information' and 'disequilibrium' factors:

$$C = H \cdot D. \tag{1.14}$$

These authors also proposed, specifically, the Shannon entropy as measure of 'information' and the quadratic distances to equiprobability as mesure of 'disequilibrium',

$$D = \sum_{i=1}^{N} \left(p_i - \frac{1}{N} \right)^2 = \sum_{i=1}^{N} p_i^2 - \frac{1}{N}. \tag{1.15}$$

Then, the LMC (López-Ruiz, Mancini and Calbet) complexity measure was defined as follows:

$$C_{LMC} = H \cdot D = \left(-\sum_{i=1}^{N} p_i \ln(p_i) \right) \left(\sum_{i=1}^{N} \left(p_i - \frac{1}{N} \right)^2 \right). \tag{1.16}$$

This measure satisfies the condition that complexity vanishes for distributions corresponding to perfect order and maximum disorder. In the case of perfect order H is zero and, then, C_{LMC} is zero too. Maximum disorder occurs when $p_i = 1/N,\ i = 1,...,N$ and then $D = 0$ and C_{LMC} is also zero. Besides, Anteneodo and Plastino (1996) showed that the probability distribution that maximizes the LMC complexity measure for $N \mapsto \infty$ is the distribution which has one state with probability $2/3$ and the other ones with equal probabilities.

There are some complexity measures closely related to the LMC complexity measure. A modification proposed by López-Ruiz et al. (1995) is the normalization of this measure replacing the Shannon entropy for its normalized version (Shannon configuration entropy)

$$\bar{C}_{LMC} = \bar{H} \cdot D = \left(\frac{\sum_{i=1}^{N} p_i \ln(p_i)}{\ln(N)} \right) \left(\sum_{i=1}^{N} \left(p_i - \frac{1}{N} \right)^2 \right). \tag{1.17}$$

Anteneodo and Plastino (1996) showed some problems of the LMC complexity measure, specially in the continuous case, such as drastic changes under rescaling transformations or under replications of the system. To solve these problems, Catalan et al. (2002) proposed another extension of the LMC complexity measure, known as the exponential form of the LMC complexity measure, defined replacing the Shannon entropy by its exponential:

$$C_{LMC}^{exp} = e^{H} \cdot D = \left(\exp\left\{ -\sum_{i=1}^{N} p_i \ln(p_i) \right\} \right) \left(\sum_{i=1}^{N} \left(p_i - \frac{1}{N} \right)^2 \right). \tag{1.18}$$

1.2.2 Other measures of complexity

In this subsection we refer to other complexity measures, which are, to some extent, related to the LMC complexity measure. The first one is a generalized complexity measure which was proposed by López-Ruiz et al. (2009), derived from the relation between the Rényi entropies and the LMC complexity measure exposed in López-Ruiz (2005). This measure, $\tilde{C}^{\alpha,\beta}$, depends on two parameters α and β and it is defined by

$$\tilde{C}^{\alpha,\beta} = e^{H_\alpha - H_\beta}, \qquad 0 < \alpha, \beta < \infty. \tag{1.19}$$

Then, this family of complexity measures contains as particular case, in some way, the exponential form of LMC measure when $\alpha = 1$ and $\beta = 2$. The main property of these generalized measures is that they are invariant under translation, rescaling transformations and replication. Moreover, in the special case of β going to infinity the exposed complexity measure is the product between a global account and a local information of the distribution.

Apart from this measure, a variety of related definitions have been proposed in the literature under the same scheme ($C = H \cdot D$) than the LMC complexity measure, with recent ones in terms of generalized entropy/divergence measures, as well as conceptual extensions such as 'relative complexity' for comparison purposes. A generalized synthetic formulation is given by Martín et al. (2006), involving different forms of entropy for 'disorder', and distance/divergence for 'disequilibrium'. In this sense, an interesting version is the 'Jensen-Shannon complexity measure' which replaces the term $D(\bar{p})$ with Jensen-Shannon divergence of $\bar{p}$ w.r.t. equiprobability, the latter being calculated as

$$D_{JS}(\bar{p}) = H\left(\frac{1}{2}\bar{p} + \frac{1}{2}\left[\overline{\frac{1}{N}}\right]\right) - \frac{1}{2}H(\bar{p}) - \frac{1}{2}\ln N. \tag{1.20}$$

Finally, we refer to the parametric family known as 'SDL complexity' introduced by Shiner, Davison and Landsberg in 1999. This family is also based on a product of 'order' and 'disorder' factors, according to the following formulation:

$$SDL(\bar{p}) = \Gamma_{\alpha,\beta} = \Delta^{\alpha}\Omega^{\beta} = \Delta^{\alpha}(1-\Delta)^{\beta} \qquad (\alpha,\beta \geq 0), \tag{1.21}$$

with

$$\Delta = \frac{H(\bar{p})}{H_{max}} \qquad \text{('configuration entropy')} \tag{1.22}$$

$$\Omega = 1 - \Delta, \tag{1.23}$$

where α is the strength of disorder, β is the strength of order, and, as before, $H_{max} = -\ln(N)$ is the entropy for an equiprobable distribution. These measures are characterized by a high degree of flexibility by means of the combination of the parameters α and β. When β is equal to 0 and $\alpha > 0$ the SDL complexity is an increasing function with respect to 'disorder' (Δ). When α is equal to 0 and $\beta > 0$ the SDL complexity is a increasing function regarding 'order' (Ω). When both are positive SDL complexity is a concave function regarding 'order' or 'disorder'. The maximum complexity occurs for the distribution which satisfies

$$\Delta = \frac{-\sum_{i=1}^{N} p_i \ln(p_i)}{\ln(N)} = \frac{1}{2}. \tag{1.24}$$

Shiner et al. (1999) showed that the 'normalized LMC complexity measure' introduced by López-Ruiz et al. (1995) is an approximation of the 'SDL complexity measure' for $\alpha = 1$ and $\beta = 1$.

$$\bar{C} = \Delta D \approxeq \frac{2\ln(N)}{N}\Omega = \frac{2\ln(N)}{N}\Gamma_{1,1}. \tag{1.25}$$

The situations when both complexity measures take the maximum value are not similar. As mentioned before, the normalized LMC complexity measure takes its maximum value when the distribution has one state with probability $2/3$ and the other ones equal probabilities. However, SDL complexity measure takes its maximum value for the distributions which satisfy the constraint 1.24.

Chapter 2

Fractality, Multifractality, and Seismic Scaling Laws

The study of irregular objects has always had great interest in the scientific community. This interest increased rapidly in 1967 when Mandelbrot introduced the concepts of statistical self-similarity and fractal dimension to describe the length of the British coast. In his subsequent books (1977, 1982), he established the bases of fractal theory and showed the fractal geometry of nature. The concept of multifractality is a generalization of the concept of fractals, which was developed to describe irregular objects which need more than one dimension to be explained. Such objects need a spectrum of dimensions at different scales to be characterized. The main tools of the multifractal analysis are the generalized Rényi dimensions and the multifractal spectrum. It is important to note that before the concept of fractals appeared the applicability of power-laws in nature was widely used. This is the case of the Gutenberg-Richter law (1945) which is introduced in Section 2.3.

This chapter is structured as follows: in Section 2.1, we refer to the concepts of scale invariance and self-similarity. Besides, we introduce the definition of fractal dimension and we carry out its calculus for two well-known sets (the Cantor set and triadic Koch island). In section 2.2, we present the multifractal tools, namely the generalized Rényi dimensions and the multifractal spectrum, along with their calculus based on the lattice formalism. To understand better the interpretations of the spectra and the measures we apply these techniques to an illustrative example. Finally, in Section 2.3 we introduce the most used scaling laws in seismicity analysis (Omory, Båth, Gutenberg-Richter), as well as a non-extensive version of the frequency-magnitude distribution.

Figure 2.1: Fractal-like images in nature (reproduced from http://webecoist.momtastic.com/2008/ 09/07/17-amazing-examples-of-fractals-in-nature).

2.1 Fractality

2.1.1 Scale invariance and self-similarity

An object possesses the self-similarity property when it is exactly or approximately similar to some part of itself. The scale invariance is the case of exact self-similarity. Nature is full of self-similar objects, such as mountain ranges, forests, rocky coastline, snowflake, etc (see Figure 2.1). Many geophysical phenomena also possess the self-similarity property, for instance, earthquakes. In this context, Mandelbrot (1967) introduced the concept of fractals which is developed in Subsection 2.1.3. From this work, many authors have considered fractal applications in geophysics.

2.1.2 Fractal set

From the books by Mandelbrot (1977, 1983), the concept of fractals has been applied in different studies to describe objects with irregular form. A possible definition of fractal set could be a geometric shape that can be subdivided into parts, each of which is exactly or approximately similar to a smaller copy of the whole object. Hence, fractal objects possess the self-similarity property. Although many definitions of fractals have appeared none of them has been univocally accepted. In this sense, Falcone (1999) introduced the idea of understanding the concept of fractals by means of a collection of properties, being the main characteristics of a fractal object the self-similarity and the irregular shape.

2.1.3 Fractal dimension

The Euclidean dimension is well-known and its value for a point is 0, for a segment is 1, for a square is 2, etc. The problem to describe an object appears when the its shape is irregular. In this case, the fractal dimension is a useful measure to describe the object. In 1967, Mandelbrot introduced the concept of fractals and calculated the fractal dimension of the West coastline of Great Britain. With the aim of exposing this development we follow the quantitative discussion of Turcotte (1989). A fractal set is defined according to the relation:

$$N_i = \frac{C}{r_i^D}, \tag{2.1}$$

where i is the order of scale, N_i is the numbers of objects with a characteristic linear dimension r_i, C is a constant of proportionality, and D is the fractal dimension, which can be calculated as

$$D = \frac{\log(N_{i+1}/N_i)}{\log(r_i/r_i+1)} = \frac{\ln(N_{i+1}/N_i)}{\ln(r_i/r_i+1)}. \tag{2.2}$$

To illustrate this measure we use two of the most common examples in the literature, the Cantor set and the triadic Koch island; see Figures 2.2 and 2.3 respectively. The Cantor set is obtained by iteration from the segment with length 1. In the first stage the segment is divided into three equal sections and the central section is deleted. In the next stage, we repeat the procedure is repeated in the sections that were not previously deleted, etc. Thus, the fractal dimension is obtained as

$$D = \frac{\ln(N_1/N_0)}{\ln(r_0/r_1)} = \frac{\ln(2/1)}{\ln(3)} = \frac{\ln(2)}{\ln(3)} = 0.6309. \tag{2.3}$$

As for the the triadic Koch island the fractal dimension is

$$D = \frac{\ln(N_1/N_0)}{\ln(r_0/r_1)} = \frac{\ln(12/3)}{\ln(3)} = \frac{\ln(4)}{\ln(3)} = 1.2619. \tag{2.4}$$

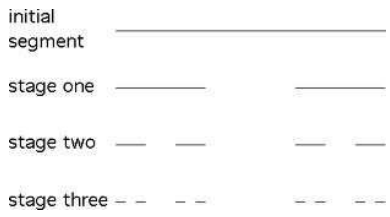

Figure 2.2: Cantor set (reproduced from http://classes.yale.edu/fractals/Labs/PaperFolding-Lab/MTSConstr.html).

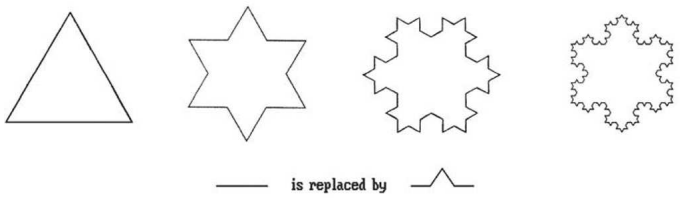

Figure 2.3: Triadic Koch island (reproduced from http://what-when-how.com/computer-graphics-and-geometric-modeling/chaos-and-fractals-special-computer-graphics-part-1).

We can consider the triadic Koch island as the deterministic case of the problem of determining the coastline length of an island. The length of the perimeter, P_i, of a fractal island is calculated as

$$P_i = r_i N_i, \tag{2.5}$$

and substituting in (2.1),

$$P_i = \frac{C}{r_i^{D-1}}. \tag{2.6}$$

Based of this expression, Mandelbrot (1967) obtained the length of the coastline of the Great Britain using different dividers of length r_i and considering $D = 1.25$. The author showed that the data points of the log-log representation of P_i vs. r_i define a straight line, which indicates that the coastline is a fractal.

2.2 Multifractality

2.2.1 Introduction

As already stated, the concept of multifractality is a generalization of the concept of fractals and it was developed to describe irregular objects which need more than one dimension to be explained. These objects need a spectrum of dimensions at different scales to be characterized. Thus, a fractal object is homogeneous or monofractal when it has the same properties at all scales and it is heterogeneous or multifractal when it needs more than one scaling exponent to be described. Multifractal techniques have been applied to a variety of practical situations such as modeling stock market time series, interpreting medical images, characterization of telecommunications traffic, etc. Multifractal systems are very common in nature, especially in geophysics and, in particular, numerous authors have shown the fractal structure in time, space and magnitude of the earthquake distribution (see Kagan and Knopoff 1980, Sadovskiy et al. 1984). They refer to a certain form of heterogeneous dynamics observed in a wide range of natural phenomena characterized by different orders of self-similarity, or 'fractality orders' (see, for instance, Mandelbrot 1982, Stanley and Meakin 1988). Distributions related to such systems are formally described in terms of the generalized Rényi dimensions and the singularity spectrum, known as multifractal spectrum. In addition, multifractality is widely used as a measure of the complexity of dynamical systems; see, for example, Sen (2007, 2011).

Multifractal tools are useful to describe the global and local structural characteristics of a point pattern at different scales. Global behaviour at different scales is depicted by the generalized Rényi dimensions whereas the multifractal spectrum shows local behaviour. Two main approaches have been adopted in this context (see Harte 2001, for an accessible introduction and critical discussion). The first one, closer to the intuitive entropy-related interpretation, is based on a box-counting procedure from regular lattice partitions, subsequently refined, of the domain. The second one, which may be computationally advantageous in some cases, is based on evaluation of the interpoint distances between events. Here, according to our objectives, we adopt the former approach, the essential elements of which are summarized in the next lines.

As mentioned, multifractal analysis is performed in practice focusing on two related sets of quantities: the multifractal spectrum, and the generalized (Rényi) dimensions. Various well-known methods are available to obtain such measures: multifractal detrended fluctuation analysis which is dedicated to analyze the scaling properties of time series (Kantelhardt et al. 2002, Mohaved et al. 2006); the Wavelet transform modulus maxima method (see Silchenko and Hu 2001, Bacry et al. 1993); and the partition function approach (Hentschel and Procaccia 1983, Halsey et al. 1986, Zhou et al. 2013). In this work, we adopt the former approach, the essential elements of whose are introduced in the next subsection.

2.2.2 The multifractal formalism

In this thesis, we perform the calculations of generalized Rényi dimensions and multifractal spectrum using the partition function formalism by means of the box-counting algorithm (see Hentschel and Procaccia 1983, Halsey et al. 1986). We assume that μ is a multifractal probability measure on a certain bounded domain D (in $\mathbb{R}^d$), and $\mathscr{X}$ is a regular lattice covering (box width ε) of D. We denote the k-th box in K_ε by $B_\varepsilon(k)$, where K_ε represents the set of boxes with positive measure. Then, the Rényi entropy of order q based on this partition is given by

$$H_q(P_\varepsilon) = \frac{1}{1-q} \ln \sum_{k \in K_\varepsilon} \mu^q(B_\varepsilon(k)). \tag{2.7}$$

In this context, the $\ln \sum_{k \in K_\varepsilon} \mu^q(B_\varepsilon(k))$ is known as 'partition function' and the scaling exponent, $\tau(q)$, is defined as

$$\tau(q) = \lim_{\varepsilon \to 0} \frac{\ln \sum_{k \in K_\varepsilon} \mu^q[B_\varepsilon(k)]}{\ln \varepsilon} \qquad -\infty < q < \infty. \tag{2.8}$$

Generalized dimensions are defined by the rate of divergence of the Rényi entropies from the partition probability distribution P_ε of μ w.r.t. decreasing box width ε, in the following sense:

$$D_1 = \lim_{\varepsilon \to 0} \frac{\sum_{k \in K_\varepsilon} \mu[B_\varepsilon(k)] \ln \mu[B_\varepsilon(k)]}{\ln \varepsilon} \tag{2.9}$$

$$D_q = \lim_{\varepsilon \to 0} \frac{\tau(q)}{q-1} = \lim_{\varepsilon \to 0} \frac{1}{q-1} \frac{\ln \sum_{k \in K_\varepsilon} \mu^q[B_\varepsilon(k)]}{\ln \varepsilon} \qquad (q \neq 1). \tag{2.10}$$

Thus, D_q defines the scaling powerlaw of the Rényi entropy of order q of P_ε w.r.t. ε,

$$e^{-H_q(P_\varepsilon)} \sim \varepsilon^{D_q} \qquad (\forall q). \tag{2.11}$$

It can be easily proved that D_q is a monotonically decreasing function of q, constant for simple scaling ('monofractal' behaviour); see Figure 2.4. Then, we will say that a data set has multifractal behaviour when $\cdots > D_{q_1} > D_{q_2} > D_{q_3} > \cdots$ for $\cdots < q_1 < q_2 < q_3 < \cdots$. Varying the 'deformation parameter' q, one can extract different features of the distribution of the measure on its support. In particular, based on their meaning and interpretation, there are four measures of special interest:

- 'Capacity' dimension, D_0. It shows how the points of pattern fill the domain under study. The larger the value of this dimension the better the space is covered.

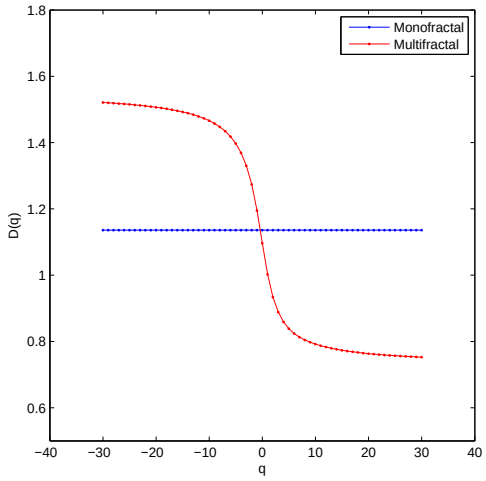

Figure 2.4: Representation of D_q vs. q for monofractal and multifractal data sets.

- 'Entropy' dimension, D_1. It is a measure of order-disorder of the points in the domain under study. Larger values indicate higher disorder.
- 'Correlation' dimension, D_2. It quantifies the degree of clustering-inhibition. Lower values correspond to a higher level of clustering.
- 'Multifractal step', $D_{-\infty} - D_{\infty}$. It quantifies the degree of multifractality. Larger values indicate a stronger multifractal behavior.

The generalized dimensions reflect global scaling behaviour based on average quantities. Alternatively, as mentioned before, multifractality can be analyzed through local scaling behaviour in terms of the multifractal spectrum. For each $\alpha > 0$, in the range of Hölder exponents present in μ, $f(\alpha)$ is defined as the 'fractal' dimension of the support of μ for that particular exponent. Formally, in the lattice approximation, with

$$K_\varepsilon(\alpha,\delta) = \left\{ k \in K_\varepsilon : \alpha - \delta < \frac{\ln \mu[B_\varepsilon(k)]}{\ln \varepsilon} \leq \alpha + \delta \right\}, \tag{2.12}$$

$f(\alpha)$ is given by

$$f(\alpha) = \lim_{\delta \to 0} \lim_{\varepsilon \to 0} \frac{\ln \# K_\varepsilon(\alpha,\delta)}{-\ln \varepsilon}, \tag{2.13}$$

that is,

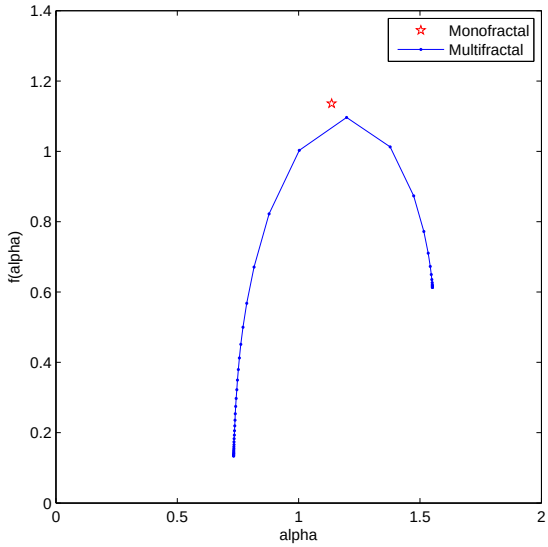

Figure 2.5: Representation of $f(\alpha)$ vs. α for monofractal and multifractal data sets.

$$\#K_\varepsilon(\alpha,\delta) \sim \varepsilon^{-f(\alpha)}. \tag{2.14}$$

In other words, $f(\alpha)$ can be interpreted as the 'density' of points where the local order of singularity of μ is equal to α. The representation $f(\alpha)$ vs α, multifractal spectrum, shows local scaling behaviour and it typically has a shape of hump in the case of multifractality while it is reduced to a point in the case of monofractality (see Figure 2.5).

The multifractal spectrum shows different aspects of interest. Its range indicates the level of multifractal behavior. In the case of monofractal structures we find scale-invariance and the spectrum concentrates on a unique point. In contrast, a wide range indicates high multifractal behavior. The peak of the curve coincides with the 'capacity' dimension. The left-hand branch is related to the degree of spatial clustering, and the curves whose left-hand branch has a slow decay correspond to stronger structured spatial patterns. Small values of α correspond to parts of the set where the measure μ is very high. However, high values of α are associated with areas with low representation, i.e, where the measure μ is very low.

Under suitable conditions, the functions D_q and $f(\alpha)$ are related in terms of a Legendre transform (see, for example, Harte 2001, Kravchenko et al. 1999), hence providing equivalent information but with related complementary interpretations. Formally:

$$\alpha(q) = \frac{d\tau(q)}{dq}, \tag{2.15}$$

$$f(\alpha) = q\alpha(q) - \tau(q). \tag{2.16}$$

In practice, the computation of the multifractal spectrum based on this relation may lead to some errors (Veneciano et al. 1995). A direct estimation method for computing the multifractal spectrum without applying the Legendre transform was proposed by Chhabra and Jensen (1989), and it has been used in a wide range of applications by several authors (Chhabra et al. 1989, Cuevas 2003, Perrier et al. 1989, Posadas et al. 2005, Wang et al. 2005, Zhou 2013). Following the development of Chhabra and Jensen (1989), the values of $\alpha(q)$ and $f(\alpha)$ can be calculated as

$$\alpha(q) = \lim_{\varepsilon \to 0} \frac{\sum_{k \in K_\varepsilon} \tilde{\mu}[B_\varepsilon(k)] \ln \mu[B_\varepsilon(k)]}{\ln \varepsilon}. \tag{2.17}$$

$$f(q) = \lim_{\varepsilon \to 0} \frac{\sum_{k \in K_\varepsilon} \tilde{\mu}[B_\varepsilon(k)] \ln \tilde{\mu}[B_\varepsilon(k)]}{\ln \varepsilon}. \tag{2.18}$$

where

$$\tilde{\mu}[B_\varepsilon(k)] = \frac{\mu[B_\varepsilon(k)]^q}{\sum_{k \in K_\varepsilon} \mu[B_\varepsilon(k)]^q}. \tag{2.19}$$

2.2.3 Illustrative example

For illustration, we apply the former techniques to the seismic series of Agrón studied in Chapter 4. We analyze the three data sets corresponding to the three periods of equal size in number of events (551), with the central one corresponding to a high concentration of seismic activity. Figure 2.6 displays the representations of the epicenter locations for the three periods, being clearly visible that the central period possesses higher level of spatial clustering.

The representations $\tau(q)$ vs. q for the three periods (see Figure 2.7) confirm that the high activity period has a higher degree of spatial clustering, as shown by the deviation of its right branch ($q > 1$) with respect to the other ones.

As for the generalized Rényi dimensions curve, depicted in Figure 2.8, we can observe how the multifractal behaviour ($D_{-\infty} - D_{\infty}$) is markedly larger in the high activity period. The main differences between the curves can be found in the values of D_q for $q > 1$, where the measures corresponding to the areas with higher density of points are emphasized.

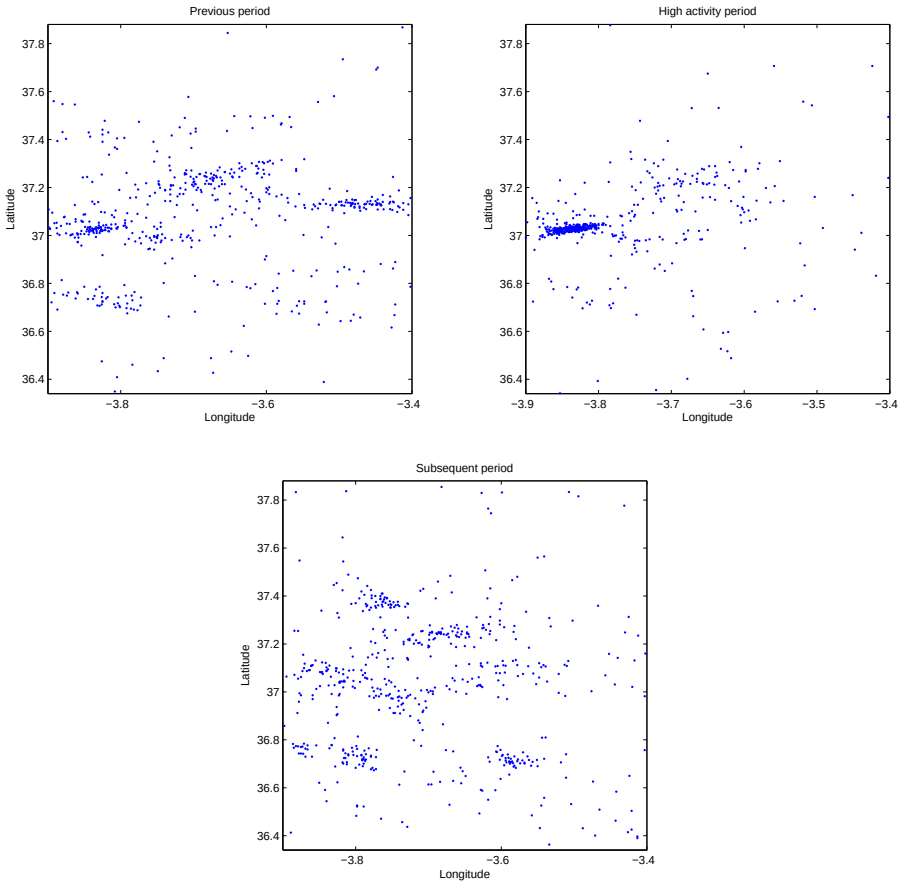

Figure 2.6: Epicenter locations of the three periods.

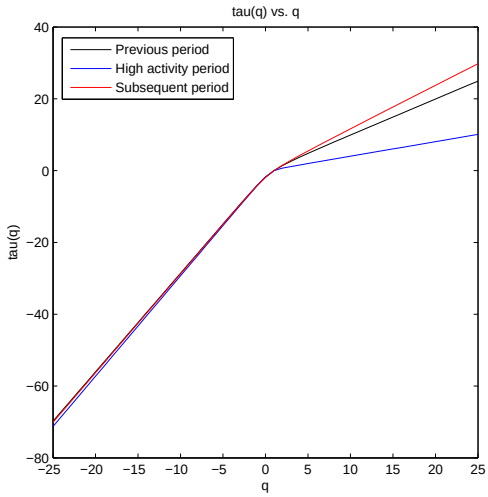

Figure 2.7: Representation of $\tau(q)$ vs. q for three periods.

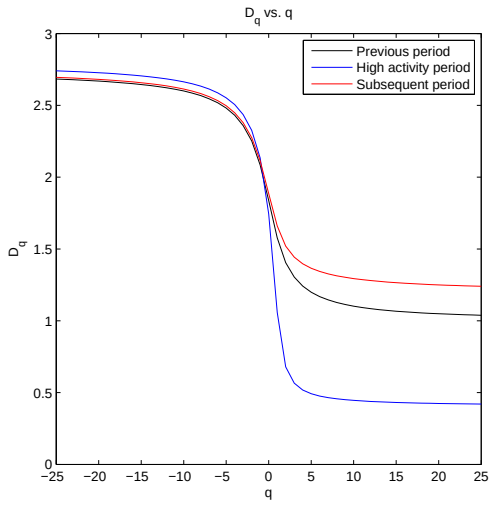

Figure 2.8: Representation of D_q vs. q for the three periods.

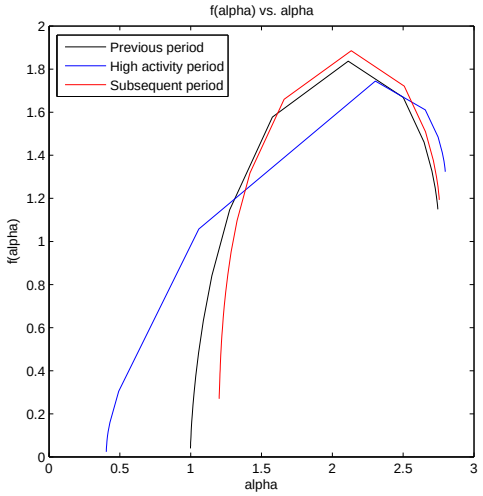

Figure 2.9: Representation of $f(\alpha)$ vs. α for the three periods.

Values of the coefficients D_0, D_1, D_2 and increments $D_{-\infty} - D_{\infty}$ for the three periods are collected in Table 2.1. These coefficients confirm that the high activity period has larger multifractal behaviour along with a higher level of clustering and a lower degree of disorder.

Table 2.1: Values of D_0, D_1, D_2 and $D_{-\infty} - D_{\infty}$ for the three periods.

	D_0	D_1	D_2	$D_{-\infty} - D_{\infty}$
Previous period	1.9055	1.6053	1.4349	2.0252
High activity period	1.8540	1.0521	0.6871	2.8381
Subsequent period	1.9562	1.6895	1.5480	1.8561

Finally, analogous interpr3tations can be derived from the multifractal spectra (see Figure 2.9). The higher multifractal behaviour and degree of clustering of the high activity period is shown by the larger range of its spectrum and the slope of its left-hand branch. The peaks of the spectra indicate that in the subsequent period the filling of the study area is better than in the other ones.

2.3 Seismic Scaling Laws

The first attempts to describe the statistical features of seismicity were made by means of empirical laws, the most used laws to describe the statistical features of earthquakes are: the frequency-magnitude distribution of events (Gutenberg-Richter's law), the decay over time in the number of events after a large mainshock (Omori's law), and the constant relationship between the magnitudes of a mainshock and its largest aftershock (Båth's law). Although some of these laws appeared long time ago, they are still used nowadays due to their usefulness for studying seismic phenomena along with the fact that they continue inspiring different techniques. The first empirical law to appear was Omory's law (1894), which was introduced to describe how the aftershock frequency decreases by roughly the reciprocal of time after the main shock. This law's expression is:

$$n(t) = \frac{K}{c+t}, \tag{2.20}$$

where A and K are positive constants. The modified version of Omori's law, now commonly used, was proposed by Utsu (1961), its expression being

$$n(t) = \frac{K}{(c+t)^p}, \tag{2.21}$$

where p varies usually between 0.7 and 1.5. In this context, Kagan and Knopoff (1978, 1981) and Ogata (1988) showed the self-similar structure in the temporal dimension of Omori's power law.

In 1944 Gutenberg and Richter proposed its empirical law which describes the distribution of the cumulative number of earthquakes with magnitude greater than a magnitude threshold m. The frequency-magnitude distribution of earthquakes is scale-invariant and follows the power law

$$\log_{10} N(\geq m) = a - bm. \tag{2.22}$$

The slope b informs about the proportion between large and small earthquakes and it is usually a value between 0.8 and 1.2. This parameter has been associated with stress conditions in the crust and material heterogeneity (Mogi 1962; Scholz 1968; Wyss 1973). In this sense, a value less than 1 is related to high tectonic stress and homogeneous fragments. In contrast, an increase in the b value indicates a highly fractured rock and heterogeneous material. The constant a has less scientific interest and it indicates the seismicity level of the region.

Finally, Båth's law (1965) states that the difference of magnitude between a mainshock and its largest aftershock,

$$\Delta m = m_{ms} - m_{as}^{max}, \tag{2.23}$$

where m_{ms} is the magnitude of the mainshock and m_{as}^{max} is the magnitude of the largest aftershock, is approximately constant, independent of the mainshock magnitude, and $\Delta m \simeq 1.2$.

The Omori and the Gutenberg-Richter empirical laws are the basis for the well-known epidemic-type aftershock sequence (ETAS) model (Ogata 1988). The ETAS model has been widely studied for over the last two decades to describe real seismicity catalogues and to obtain short-term probabilistic forecasts.

Despite the fact that the Gutenberg-Richter law is the most widely accepted law to describe seismic phenomena, it presents difficulties for events having small and high magnitudes (see, for example, Ibáñez et al. 2012). The misfit for the small values of magnitude has been usually attributed to the lack of sensibility of the measurement devices. In practice, a correct estimate of a and b is crucial to better understand the seismicity sequence.

On the one hand, with the aim of improving the estimation it is usual to consider a minimum threshold of magnitude. This threshold is known as the magnitude of completeness, M_c, and it is defined as the magnitude at which the lower end of the frequency-magnitude distribution (FMD) departs from the exponential decay (a interesting review can be found in Mignan and Woessner 2012). There exist some techniques to estimate M_c but the most used ones are: the maximum curvature method, MCM (Wiemer and Wyss 2000); the Goodness-of-Fit Test technique, GFT (Wyss et al. 1999, Wiemer and Wyss 2002); and the Bayesian Magnitude of Completeness method, BMC (Mignan et al. 2011).

On the other hand, Sotolongo-Costas and Posadas (2004) argued that before arriving to the threshold value of instrument sensitivity the behaviour between frequency and magnitude can be explained by the existence of material between the planes of fault. In this context, the size distribution of faults and the material between the planes of fault are important factors to study the seismicity of an area (see, for example, Saleur et al. 1996, Sornette 1999). Sotolongo-Costa and Posadas (2004) proposed a model for earthquake dynamics, known as fragment-asperity model, based on the idea of considering the irregularity of the tectonic plates surface and the fragments between the planes of fault. This model is obtained using the Tsallis entropy formalism and it provides a new frequency-magnitude distribution in the non-extensive domain:

$$\ln\left(\frac{N(m>m_{th})}{N}\right) = \left(\frac{2-q}{1-q}\right)\ln\left\{1+k(q-1)(2-q)^{\frac{1-q}{q-2}}10^{2m_{th}}\right\}, \tag{2.24}$$

where N is the total number of earthquakes, $N(m>m_{th})$ represents the number of earthquakes with magnitude m greater than the threshold m_{th}, and q is the non-extensivity parameter.

Silva et al. (2006) presented a re-analized version of this fragment-asperity interaction model, and derived the following frequency-magnitude distribution:

$$\ln\left(\frac{N(m>m_{th})}{N}\right) = \left(\frac{2-q}{1-q}\right)\ln\left\{1-\left(\frac{1-q}{2-q}\right)\left(\frac{10^{2m_{th}}}{a^{2/3}}\right)\right\}. \tag{2.25}$$

In addition, Telesca (2011, 2012), using different relation between the magnitude and released energy obtained the next expression:

$$\ln\left(\frac{N(m>m_{th})}{N}\right)=\left(\frac{2-q}{1-q}\right)\ln\left\{1-\left(\frac{1-q}{2-q}\right)\left(\frac{10^{m_{th}}}{a^{2/3}}\right)\right\}. \tag{2.26}$$

These approaches have been used to study seismic catalogues obtaining excellent estimates in recent studies (see, for example, Celikoglu 2012; Telesca 2010, 2011, 2012; Vallianatos 2009, 2013). The constant a expresses the proportionality between the released energy and the fragment size, and, as before, q is the non-extensivity parameter. In the context of the fragment-asperity model, q is related to the scale of the spatial interaction. The physical state is at equilibrium when q is equal to 1. When the value of q is increased the physical state is far from the equilibrium. In this sense, a significant increase in the non-extensivity parameter suggests that the system enters in an instability state where important amounts of energy can be released (see Kalimeri et al. 2008, Telesca 2010a).

Part II

METHODOLOGY AND APPLICATIONS

Chapter 3

Proposed Techniques and Methodological Aspects

In the former chapters we have exposed some of the most used statistical techniques in the study of seismic phenomena. The common use of such tools is mainly motivated by the nature of earthquakes, which, among other aspects, involve a relation to complex dynamics governed by scaling powerlaws related to self-similarity (fractality) of possibly different orders (see, for instance, Chen et al. 2007; Godano et al. 1999; Marsan et al. 1999) and the nonadditive character of the fragmentation process in the tectonic plates (see Silva et al. 2006; Sotolongo-Costas and Posadas 2004). These aspects jointly with the aims on which this thesis work is focused, such as the dimensional interaction, scaling behaviour, evolutionary changes or the incorporation of the magnitude effects to the analysis, have taken us to propose some extensions. Thus, in the first section of this chapter we formulate a new generalized dimension based on Tsallis non-extensive statistics, we establish a limiting connection between some complexity measures and generalized dimensions, we construct a dependence coefficient in the multifractal domain, and we propose a formal approach to incorporate effects such as those of the magnitude to the analysis. In Section 2 we deal with some methodological aspects used to apply the exposed and proposed techniques. These procedures are relevant to study some aspects of interest in the analysis of seismic series, and they are also necessary to understand the analysis performed in the following section and their corresponding interpretations. Specifically, we introduce: some transformations on the temporal component which combined with the former techniques allow us to evaluate the effect of the occurrence time on the structural complexity displayed by the data; the sliding windows technique, useful for assessing evolutionary changes (very relevant, in particular, to describe temporal structure heterogeneities and for detection of possible precursory elements); and the fitting of the frequency-magnitude distribution of earthquakes by means of a non-linear regression considering weights according to the released energy.

3.1 Technical Extensions

As we mentioned before, this section comprises new contributions to the study of seismicity proposed in this thesis. Firstly, we propose a Tsallis entropy-based version of generalized dimensions, motivated by the nonadditive character of earthquakes along with the fact that they are governed by scaling powerlaws of possibly different orders. Next Subsection, we study some limiting aspects based on scaling behaviour. Specifically, we establish limiting connections between well-known complexity measures and generalized dimensions, showing the potential use of related functionals for quantitative assessment of 'multifractal complexity'. Besides, with the aim of assessing the degree of association between components, we introduce a dependence coefficient based on the limiting behaviour of mutual information to quantify the dimensional interaction, in particular the spatio-temporal interaction. Finally, we deal with the problem of incorporating the local effects of some marker to the analysis, formulating a weighted box-counting procedure to construct the reference probability distributions.

3.1.1 A Tsallis entropy-based formulation of generalized dimensions

Non-extensive statistics, based on Tsallis entropy, has become a very useful tool to describe the complexity of dynamical systems. The fact that Tsallis non-extensive formulation allows interactions at different orders and that the process of shock fragmentation is non-additive (long-range interactions occur in the whole object) has led to the broad application of non-extensive statistics to model and describe seismic phenomena in the last years. In this context, some contributions have studied seismic catalogues from a non-extensive frequency-magnitude distribution of earthquakes obtaining successful fittings (see Celikoglu 2012; Telesca 2010, 2011, 2012; Vallianatos 2009; Vallianatos et al. 2013). Besides, Darooneh and Dadashinia (2008) performed an analysis of the spatial and temporal distributions between successive earthquakes from the point of view of the non-extensive statistical mechanics. This fact along with the multifractal and self-similar structure exhibited by seismic phenomena has led us to propose a new formulation of generalized dimensions based on the non-extensive formalism. We have constructed this new formulation looking at the structure of the generalized Rényi dimensions equations. As introduced in Chapter 2, the usual concept of generalized dimensions is directly related to the limiting behaviour of the Rényi entropies based on the box-counting distribution given by

$$D_q = \lim_{\varepsilon\to 0}\frac{\tau(q)}{q-1} = \lim_{\varepsilon\to 0}\frac{1}{q-1}\frac{\ln\sum_{k\in K_\varepsilon}\mu^q[B_\varepsilon(k)]}{\ln\varepsilon} \qquad (q\neq 1) \tag{3.1}$$

$$D_1 = \lim_{\varepsilon\to 0}\frac{\sum_{k\in K_\varepsilon}\mu[B_\varepsilon(k)]\ln\mu[B_\varepsilon(k)]}{\ln\varepsilon}. \tag{3.2}$$

The denominator, thus, reflects the pure effect of information gain associated with the lattice subdivision, without consideration of the distribution of the events within. These equations can be rewritten for $\bar{p} = (p_1, p_2, ..., p_N)$ p.d.f. of the r.v. X in terms of Rényi entropy and replacing ε by $1/N$ (since the component X is normalized to fit within the unit hypercube) as

$$D_q = \lim_{N\to\infty} \frac{1}{q-1} \frac{\sum_{i=1}^{N} p_i(\varepsilon)^q}{\ln\left(\frac{1}{N}\right)} = \lim_{N\to\infty} \frac{H_q(\bar{p}(\varepsilon))}{\ln N} \qquad (q \neq 1) \tag{3.3}$$

$$D_1 = \lim_{N\to\infty} \frac{\sum_{i=1}^{N} p_i(\varepsilon) \ln p_i(\varepsilon)}{\ln\left(\frac{1}{N}\right)} = \lim_{N\to\infty} \frac{H_1(\bar{p}(\varepsilon))}{\ln N}. \tag{3.4}$$

In these expressions we can identify that the term in the denominator, $\ln N$, is directly related to Rényi entropy for a system with equiprobable states (maximum entropy):

$$\begin{aligned} H_q\left(\left[\frac{1}{N}\right]\right) &= \frac{1}{1-q} \ln\left(\sum_{i=1}^{N} \left(\frac{1}{N}\right)^q\right) = \frac{1}{1-q}\left(-\ln(N^{q-1})\right) = \ln N \qquad (q \neq 1) \\ H_1\left(\left[\frac{1}{N}\right]\right) &= -\sum_{i=1}^{N} \frac{1}{N} \ln\left(\frac{1}{N}\right) = \ln N. \end{aligned} \tag{3.5}$$

Then, we use here this interpretation to propose a suitable definition of generalized dimensions related to the limiting behaviour of Tsallis entropy (or 'generalized Tsallis dimensions'). Now, the Tsallis entropy for a system with equiprobable states is given by

$$\begin{aligned} T_q\left(\left[\frac{1}{N}\right]\right) &= \frac{1 - \sum_{i=1}^{N} \left(\frac{1}{N}\right)^q}{q-1} = \frac{1-N^{1-q}}{q-1} \qquad (q \neq 1) \\ T_1\left(\left[\frac{1}{N}\right]\right) &= H_1\left(\left[\frac{1}{N}\right]\right) = \ln N. \end{aligned} \tag{3.6}$$

Thus, we can formulate the 'generalized Tsallis dimensions' as follows: For $q \neq 1$,

$$DT_q = \lim_{N\to\infty} \frac{\frac{1}{q-1}\left(1 - \sum_{i=1}^{N} p_i(\varepsilon)^q\right)}{\frac{1-N^{1-q}}{q-1}} = \lim_{N\to\infty} \frac{1 - \sum_{i=1}^{N} p_i(\varepsilon)^q}{1-N^{1-q}}, \tag{3.7}$$

and replacing N with $1/\varepsilon$ we have the expression

$$DT_q = \lim_{\varepsilon \to 0} \frac{1 - \sum_{k \in K_\varepsilon} \mu^q[B_\varepsilon(k)]}{1 - \varepsilon^{q-1}}. \tag{3.8}$$

Obviously, given the coincidence of Rényi and Tsallis entropies with Shannon entropy for $q = 1$, we define by continuity $DT_1 := D_1$. In particular, for $q > 1$, we can just write

$$DT_q = (q-1) \lim_{\varepsilon \to 0} T_q(P_\varepsilon). \tag{3.9}$$

In this case, we directly have, for $q \neq 1$, the approximation

$$T_q(P_\varepsilon) \sim DT_q \frac{1 - \varepsilon^{q-1}}{q-1}. \tag{3.10}$$

3.1.2 Some limiting behaviour aspects

This subsection is focused on analyzing two aspects of special interest, both regarding the limiting behaviour of information and complexity related quantities. First, we establish a limiting connection between widely used forms of complexity and generalized dimensions, justifying the potential use of functionals based on generalized dimensions for complexity assessment in this context. Second, we introduce a mutual information-based dependence coefficient in the generalized dimension domain to quantify, in particular, the space-time interaction.

Complexity measures in terms of generalized dimensions The complexity measures introduced in Chapter 1 have proved to be useful to quantify the ability of a system for organization but they present the inconvenience that they are only calculated for a single scale. As we mentioned previously, the dynamics of many natural phenomena is commonly related to complex behaviour characterized by various forms of heterogeneity of possibly different orders. Then, in this context, and in particular in seismicity, a measure which quantifies the degree of complexity of the system independently of the scale seems appropriate. This leads us to propose complexity measures from their limiting behaviour, establishing a connection with generalized dimensions. To begin with, we assume, for convenience, the exponential version of LMC complexity, given by

$$\begin{aligned} LMC_{\exp} &= H \cdot D = e^{H(\bar{p})} D(\bar{p}) = e^{H(\bar{p})} \sum_{i=1}^{N} \left(p_i - \frac{1}{N} \right)^2 \\ &= e^{H(\bar{p})} \left[E - \frac{1}{N} \right] = e^{H(\bar{p})} \left[e^{-H_2(\bar{p})} - \frac{1}{N} \right]. \end{aligned} \tag{3.11}$$

We are interested in studying the limiting behaviour of complexity with respect to decreasing box size in the lattice approximation. The choice of the exponential form for the information

factor has here a certain justification in relation to the nondiscrete form of the limit distribution, but it is mainly adopted for technical convenience. Then, considering the scaling powerlaw of Rényi entropy of order q,

$$e^{-H_q(P_\varepsilon)} \sim \varepsilon^{D_q} \qquad (\forall q) \tag{3.12}$$

we can write the approximation

$$\begin{aligned} e^{H(P_\varepsilon)}\left[e^{-H_2(P_\varepsilon)} - \frac{1}{\#K_\varepsilon}\right] &= e^{H(P_\varepsilon)}\left[e^{-H_2(P_\varepsilon)} - e^{-\ln \#K_\varepsilon}\right] \\ &= e^{H(P_\varepsilon)}\left[e^{-H_2(P_\varepsilon)} - e^{-H_0(P_\varepsilon)}\right] \\ &\sim \varepsilon^{-D_1}\left(\varepsilon^{D_2} - \varepsilon^{D_0}\right). \end{aligned} \tag{3.13}$$

Now, since $D_2 \leq D_0$, and excluding the trivial case of equality, which would correspond to a monofractal behaviour, we subsequently follow the approximation with

$$\cdots \sim \varepsilon^{-D_1}\varepsilon^{D_2} = \varepsilon^{D_2 - D_1}. \tag{3.14}$$

Such a relationship means that the difference $D_1 - D_2$ can in fact be interpreted as a simple case of a complexity measure in the generalized dimension domain (or 'multifractal complexity'), directly related to the conceptual approach given by LMC complexity.

Similarly, we can construct a generalized complexity mesure family in the multifractal domain considering the extension on the limit behaviour of the family of generalized complexity measures proposed by López-Ruiz et al. (2009), $\tilde{C}^{\alpha,\beta}$. We start from its expression

$$\tilde{C}^{\alpha,\beta}(\bar{p}) = e^{H_\alpha(\bar{p}) - H_\beta(\bar{p})}, \tag{3.15}$$

and considering the scaling powerlaw of Rényi entropy of order q, we can write

$$\tilde{C}^{\alpha,\beta}(P_\varepsilon) = e^{H_\alpha(P_\varepsilon) - H_\beta(P_\varepsilon)} \sim \varepsilon^{-D_\alpha}\varepsilon^{D_\beta} = \varepsilon^{D_\beta - D_\alpha}. \tag{3.16}$$

Hence, $D_\alpha - D_\beta$ can be interpreted as a generalized complexity measure in the generalized dimension domain.

Following this approach, similar results can be established for other versions of complexity measures; for instance, in relation to some forms included in the general formulation proposed by Martín et al. (2006). In some cases, formalization through the limiting approximation $e^{-H_q(P_\varepsilon)} \sim \varepsilon^{D_q}$ $(\forall q)$ is not analytically as immediate as to provide a simple closed form in terms of generalized dimensions. However, the above results constitute significant cases of the potential usefulness of suitably selected generalized dimension functionals for quantitative description of complexity characteristics in the multifractal domain.

Analyzing dimensional interaction: a dependence coefficient Over the last decades, spatio-temporal marked point processes have been used in a wide variety of applications, such as earthquakes, wildfires or epidemics (see Ripley 1977; Diggle 1983; Guttorp 1995; Schoenberg et al. 2002a). In these applications the models are commonly separable, which has led to an interest in the study of the independence or separability of some components of a point process. Several separability tests have appeared in the last decades, most of them based on nonparametric conditional intensity, $\lambda(X,Y,T)$ (see, for example, Assunção and Maia 2007; Díaz-Avalos et al. 2013; Schoenberg 2004). In this context, a spatio-temporal process (X,Y,T) is separable with respect to the temporal component T if

$$\lambda(X,Y,T) = f(T)\lambda_1(X,Y). \tag{3.17}$$

where f is a fixed nonnegative function and λ_1 is a predictable nonnegative process (Schoenberg 2004).

Another relevant precedent is the work of Furuichi (2006) in which two possible correlation coefficients are constructed based on the Tsallis 'mutual entropy' given by

$$\tilde{I}_q^T(X,Y) := T_q(X) + T_q(Y) - T_q(X,Y), \tag{3.18}$$

with $q > 1$.

The expressions of the correlation coefficients proposed by Furuichi (2006) are

$$\rho_q(X,Y) = \frac{\tilde{I}_q^T(X,Y)}{T_q(X,Y)} \quad \text{and} \quad \hat{\rho}_q(X,Y) = \frac{\tilde{I}_q^T(X,Y)}{\max\{T_q(X), T_q(Y)\}}, \tag{3.19}$$

for $T_q(X) > 0$, $T_q(Y) > 0$, $q > 1$.

The multifractal characteristics of seismicity in space and time lead us to propose a coefficient to quantify the degree of association between space and time independently of the scale. Specifically, we formulate a dependence coefficient which quantifies departure from independence in the generalized dimensions domain. According to the scaling power law of Rényi entropy of order q of P_ε w.r.t. ε,

$$e^{-H_q(P_\varepsilon)} \sim \varepsilon^{D_q} \qquad (\forall q). \tag{3.20}$$

For simplicity, we introduce here the case $q = 1$. Applying (3.20) separately to (X,Y) and T, we have

$$e^{-H\left(P_\varepsilon^{(X,Y)}\right)} \sim \varepsilon^{D_1^{(X,Y)}} \tag{3.21}$$

$$e^{-H\left(P_\varepsilon^{T}\right)} \sim \varepsilon^{D_1^{T}}, \tag{3.22}$$

and hence,

$$e^{-H\left(P_\varepsilon^{(X,Y)}\right)-H\left(P_\varepsilon^{T}\right)} \sim \varepsilon^{D_1^{(X,Y)}+D_1^{T}}. \tag{3.23}$$

Note that the exponent in the left-hand side of the latter expression would correspond, except for the minus sign, to evaluation of the joint entropy under the condition of independence between the spatial and temporal components. For the actual joint vector (X,Y,T), we would have

$$e^{-H\left(P_\varepsilon^{(X,Y,T)}\right)} \sim \varepsilon^{D_1^{(X,Y,T)}}. \tag{3.24}$$

Hence, we can assert that under independence of (X,Y) and T we would have the equality $D_1^{(X,Y,T)} = D_1^{(X,Y)} + D_1^{T}$. In fact, the difference $D_1^{(X,Y)} + D_1^{T} - D_1^{(X,Y,T)}$ represents the scaling powerlaw behaviour of mutual information (in Shannon sense) between (X,Y) and T, and quantifies departure from independence in the generalized dimension domain, for $q = 1$ (i.e. in terms of 'entropy' dimensions). Normalization provides with respect to $D_1^{(X,Y)} + D_1^{T}$ a dependence coefficient (or 'dimensional interaction coefficient') in the form

$$DC = \frac{D_1^{(X,Y)} + D_1^{T} - D_1^{(X,Y,T)}}{D_1^{(X,Y)} + D_1^{T}}. \tag{3.25}$$

This coefficient is especially useful when we focus on assessment of evolutionary structural heterogeneities in spatio-temporal interaction. For this purpose, we can calculate the dependence coefficient over time by means of sliding windows (see Subsection 3.2.2). Other alternative versions, with certain differential properties and interpretations, can also be formulated by replacing the denominator with $D_1^{(X,Y,T)}$ or $\max\{D_1(X), D_1(Y)\}$. Moreover, some extensions of this proposal under similar arguments can be obtained for $q \neq 1$.

3.1.3 Incorporation of effects through distributions

In the study of seismic series it is usual to analyze point processes composed of space-time coordinates identifying epicenters or hypocenters and occurrence times. However, in some cases one might want to incorporate, in some appropriate way, extra information discriminating events; for instance, in terms of their respective magnitude (related to the energy release in a logarithmic scale). Thereby, we can assess evolutionary changes in the association between multifractal characteristics inherent to the spatio-temporal interaction and the effects of the magnitude. From the point of view of point processes, this (and other types of potentially complementary information) can be considered as a mark associated with each individual event. A simple way to formally incorporate such effects, that we explore here, consists of modifying the derivation of the box-counting probability distributions by weighting each event in terms of

a suitably selected function of its magnitude. In this respect, we propose to consider the next expression to construct the probability distribution in the box-counting procedure: For $k \in K_\varepsilon$.

$$p_k^{M,\phi}(\varepsilon) = \frac{\sum\limits_{i \in B_\varepsilon(K)} \phi(m_i)}{\sum\limits_{i \in D} \phi(m_i)} = \frac{\sum\limits_{i \in D} \mathbb{1}_{B_\varepsilon(k)}(i)\phi(m_i)}{\sum\limits_{i \in D} \phi(m_i)}, \tag{3.26}$$

where D is the complete domain, m_i is the magnitude (mark) of point event i, and $\mathbb{1}_{B_\varepsilon(k)}$ is the indicator function for box $B_\varepsilon(k)$,

$$\mathbb{1}_{B_\varepsilon(k)}(i) = \begin{cases} 1 & \text{if } i \in B_\varepsilon(k) \\ 0 & \text{otherwise.} \end{cases}$$

In the case of considering $\phi(m_i) \equiv 1$ we are in the standard situation where the measure value $p_k^{M,\phi}(\varepsilon)$ is defined as the proportion of points falling within the box $B_\varepsilon(k)$. If $\phi(m_j)$ is a suitable function of the magnitude certain aspects about the intrinsic association between the magnitude and the spatio-temporal dynamics of the phenomenon can be assessed. Here, we propose to consider the released energy to construct ϕ. In this sense, the relation between magnitude and released energy (E) has the logarithmic form $m = \lambda \ln E$, where λ usually is 1 or $3/2$ (see, for instance, Silva et al. 2006; Telesca 2010, 2011). In the next chapter we illustrate this approach in the applications to real seismic series.

3.2 Further Methodological Aspects

This section is devoted to the description of some complementary methodological aspects used to apply the exposed and proposed techniques to the analysis of the seismic series performed in the following chapter. A proper understanding of these methodologies is important to identify the aims which are pursued in each analysis and for the interpretations of the results obtained. The first aspect addresses the evaluation of the significance of time in the spatio-temporal structuring of the phenomenon, and consists of the application of different transformations on the temporal component which partially break the spatio-temporal association. The aim of the second aspect, referred to as 'sliding windows', is the assessment of evolutionary changes in the seismic series based on sliding windows. Finally, the last aspect consists of considering a weighted function in terms of the released energy associated with each threshold to perform the fitting of the earthquakes frequency-magnitude distribution proposed by Silva (2006).

3.2.1 Transformations on the temporal component

As we have pointed out, one of the main aims of this thesis work is the assessment of the spatio-temporal association displayed by the data. In particular, we apply the exposed and pro-

posed techniques to different spatio-temporal point processes obtained modifying the temporal component in a certain way. The purpose of these transformations is to partially break the space-time association. This methodology is applied in the study of a seismic series in the next chapter and here we use the corresponding patterns to illustration.

The original data based on the real temporal component is denoted as TR, and its 3D scatter plot representing longitude, latitude and time can be visualized in the top left plot of Figure 3.1. In this representation we can observe the existence of strong spatio-temporal clusters. The first transformation on the temporal component is performed to keep the order of the events, but discarding the interevent time information by assigning equally spaced time intervals. This first transformation is known as natural time, TN (see Tanaka et al., 2004; Varotsos, 2011). In the spatio-temporal scatter plot, shown in the top right plot of Figure 3.1, the spatio-temporal clusters are vertically deformed, and the marginal purely temporal clustering disappears. In the next transformation, denoted shuffled time, TS, we randomly permute the temporal component, which may be seen as a random reallocation of the spatial components (X,Y) to each occurrence time. In this case, the spatio-temporal association is broken but the original marginal temporal and spatial clusters are separately kept, as we can observe in the bottom left plot of Figure 3.1. Finally, we fully break the spatio-temporal association generating new independent time values from a continuous uniform distribution, fixing its parameters according to the real time ($U(min(T),max(T))$), this case denoted TU. The 3D scatter representation of the resulting patterns is shown in the bottom right plot of Figure 3.1, in which the lack of spatio-temporal association and temporal clusters is visible and only the marginal spatial structure is kept. Another related transformation, also used in the literature, is given by 'shuffled interevent times' (TSI), keeping the temporal order of events but randomly permuting –i.e. shuffling– the series of interevent times. The effects of the transformation are similar in a certain sense to those of natural time. Some related results (not collected here) are referred in Angulo and Esquivel (2013b).

A comparative analysis of results for these different spatio-temporal point patterns provides important information about the spatio-temporal association in relation to the study of scaling behaviour, structural complexity or evolutionary changes.

3.2.2 Sliding windows

In order to assess structural evolutionary changes over time we analyze the variations in the results obtained by applying the different techniques to data segments based on sliding windows. The basic form of this approach consists of fixing a size of time interval, or window size, analyzing the spatio-temporal pattern of events within this time interval, moving this window over time with a certain level of overlapping, and, finally, comparing the results obtained for the different subsets over time. Figure 3.2 shows an illustrative diagram, where we can observe how the different subsets (S_i) have the same duration although the number of events within each

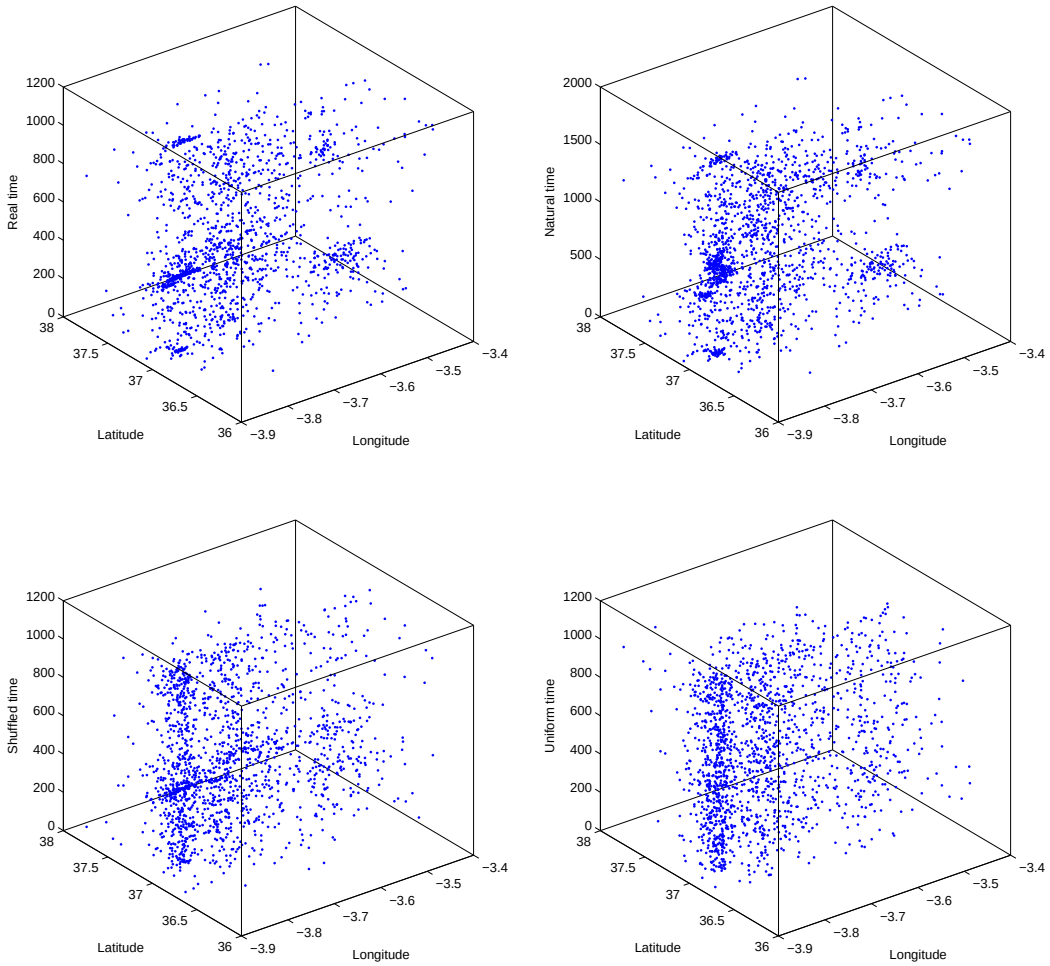

Figure 3.1: 3D scatter plots of components (X, Y, T), corresponding to TR, TN, TS and TU.

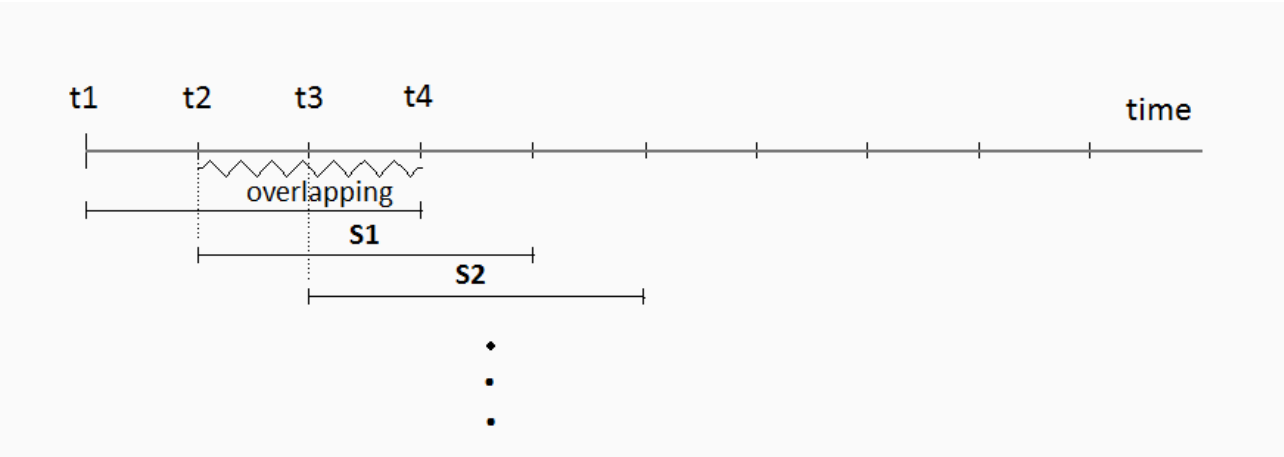

Figure 3.2: Illustration of the sliding windows technique with a window of fixed size according to a time interval.

subset can vary.

When analyzing seismicity we can find the problem of the existence of windows with low representation or even with no events at all. In fact, it is known that the estimates of the scaling properties of seismic systems may be significantly affected by the size of data sets (Eneva, 1996) and, in addition, the fitting of the Gutenberg-Richter law approach is affected by intervals with low levels of seismicity. Because of this, in the context of seismicity, it is often convenient to consider windows with a fixed number of events rather than with equally spaced temporal intervals. This approach, which can be viewed as to take a window with fixed time length in the natural time (TN), is illustrated in Figure 3.3 in which both, each data segment (S_i) and the overlapping are chosen to have the same number of events, respectively.

Moreover, there are some characteristics of this methodology we have to take into account to perform the appropriate analysis and understand the results conveniently. The first one is the size of the overlapping, which must be high in order to obtain more accurate results on transitions. In general, this overlapping is expressed in terms of a percentage of the window size. The second one is the choice of the window size, which depends on the type of evolutionary changes we are interested in for our study. Thus, if the aim is to study evolutionary changes at a local level the window size has to be small in proportion with the total number of events. In contrast, a window size with a higher number of events allows to detect global evolutionary changes. Finally, regarding the graphical representation of the estimated values for each window it is convenient to associate each estimate with the occurrence time corresponding to the last event included in the window, to avoid misleading artificially anticipatory effects.

In our analysis of seismic data, shown in the following chapter, we have widely applied

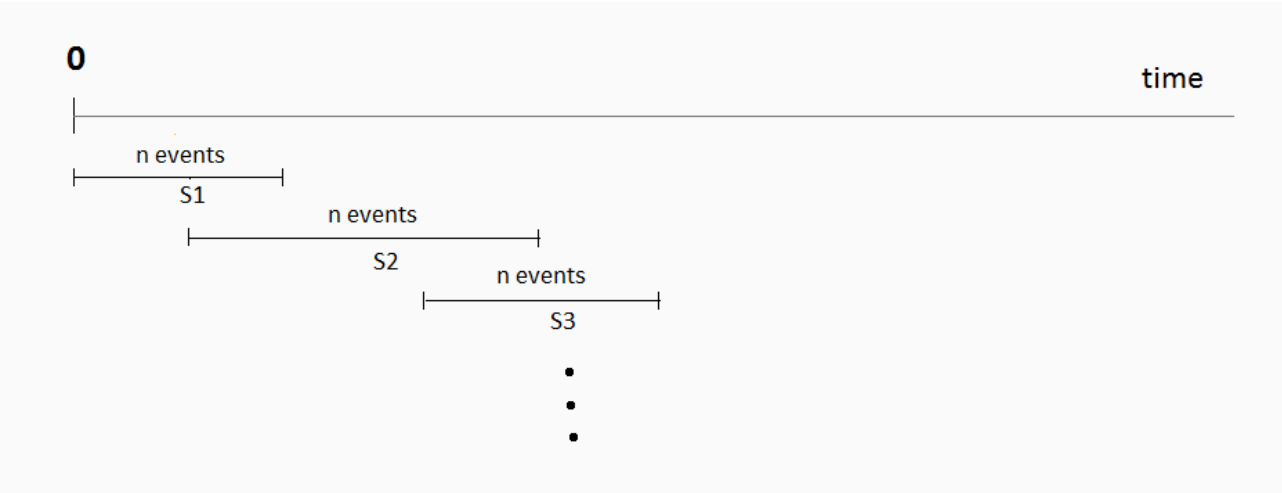

Figure 3.3: Illustration of the sliding windows technique with a window of fixed size according to a number of events.

this technique, always using a window size of 150 events and an overlapping of 90 per cent for illustration.

3.2.3 Fitting of the frequency-magnitude distribution by means of weighted non-linear regression

In the last decade, the use of the non-extensive statistical mechanics has increased considerably in the study of geophysical dynamics, especially in seismicity. In this context, the modification of the Gutenberg-Richter law based on nonextensive statistics (proposed firstly by Sotolongo-Costa and Posadas 2004, and re-analized by Silva et al. 2006) has been applied to different seismic catalogues with slight adaptations obtaining excellent fitted approximations over the whole range of magnitudes (see Celikoglu 2012; Darooneh and Mehri 2010; Telesca 2010, 2011, 2012; Vallianatos 2013). In addition, the estimated parameters of the expression provide relevant information to describe the physical characteristics of seismic series. Thus, the fact that the goodness of fit of the expression is good is fundamental for the interpretations of the parameters to be correct.

As a consequence, the choice of the estimation method used to perform the fitting has significant importance to carry out a correct analysis of the seismic series. To fit this model, several methods have been used such as the Levenberg-Marquardt non-linear least squares method (Levenberg 1944, Marquardt 1963), or the estimation of the parameters by means of the max-

imum likelihood method (Telesca 2012). These methods consider all magnitude thresholds with equal weight for the estimation and the highest misfits are obtained for the largest thresholds. With the aim of correcting such misfits, we propose to improve the fitting increasing the weight for the thresholds according to magnitude. Hence, we carry out the fitting by means of the weighted non-linear regression performed by using the least squared method considering a weighting function based on the following relation between magnitude and energy:

$$\phi_\lambda(m) = e^{\lambda m} = (e^m)^\lambda . \tag{3.27}$$

In the above expression λ is usually taken equal to 1 or $3/2$ depending on the relation between magnitude and energy used to obtain the expression of the fragment size distribution function (see, for instance, Silva et al. 2006; Telesca 2010, 2011; Vallianatos et al. 2013). Besides, it is convenient to normalize the weighting values to sum up to 1. Thus, we propose to perform the fitting of the cumulative number of earthquakes with magnitude larger than m with respect to the threshold m using the following normalized weighting function:

$$w(m_i) = \frac{\phi_\lambda(m_i)}{\sum_{j=1}^{n} \phi_\lambda(m_j)} \tag{3.28}$$

where n is the total number of thresholds considered.

In the next chapter, we perform the fitting to two seismic series by means of the nonweighted and weighted nonlinear regression for comparison. As it will be shown, the results indicate that the goodness of fit improves in all cases when we use weights to perform the estimations.

Chapter 4

Seismic Data Analysis

In this chapter, three relevant seismic data sets are studied by means of the proposed techniques introduced in this thesis work. These data sets have been chosen due to their different nature and the special interest in their study. The first studied data set is associated with tectonic activity in the seismic zone of Agrón (Granada, Spain). This series presents the behaviour of an active seismic area and it has been chosen for two main reasons: firstly, to assess the structural changes of the spatio-temporal interaction in the evolution of the seismic phenomenon before, during and after a high activity period; and secondly, to serve as a reference for comparison to the other data sets, which present special particularities. The second one is a well-known seismic series recently occurred in the volcanic island of El Hierro (Canary Islands, Spain). The study of this seismic series is interesting mainly due to the fact that a volcanic eruption is involved. Besides, it is very interesting because of the great number of events occurred, more than 11.000 earthquakes, in just 173 days. In this data set the main aim is to assess structural changes regarding the volcanic nature of the series. Finally, the last seismic series analyzed occurred also recently in the area near the village of Torreperogil (Jaén, Spain). Among the particularities of this last series we can find the fact that this area had never had a relevant seismic activity, when suddenly a high number of earthquakes began to occur concentrated on a small zone. The extreme and unexpected changes in the behaviour of the seismic activity in this area have caused a big impact on the zone and its inhabitants. The uncertainty about the origin has produced controversy and, consequently, several theories have appeared. Thus, in this case the aim is to try to detect whether there is a possible presence of external factors in the phenomenon.

It is important to note that we have carried out a descriptive analysis, a study of the dimensional interactions and an assessment of the structural characteristics at different scales common for the three data sets. However, these analyses have been completed in different ways according to the established objectives for each series. Thus, in the case of Agrón we have performed a deeper study of the spatio-temporal interaction combining the techniques with temporal trans-

formations. Such a study has not been carried out for the other two series to avoid repetition. On the other hand, in the cases of El Hierro and Torreperogil the interesting characteristics involved have motivated us to complete the analysis by means of the non-extensive frequency-magnitude distribution.

4.1 Agrón Data

The first data series we analyze corresponds to a seismic series occurred before, during and after a high activity period in an seismic area where activity is associated with tectonic dynamics. This seismic series is composed of 1.653 seismic events registered in an area near the village of Agrón (Granada, Spain) during the period between 8 October 1987, and 10 September 1990. This selection contains three subperiods of equal size in number of events (551), with the central one corresponding to a high concentration of seismic activity. Here, we pay special attention to assessing structural changes in the evolution of the phenomenon between (‘pre’, ‘central’ and ‘post’) subperiods, based on the data. Figure 4.1 shows the swarm of events depicted according to their 2D geographical coordinates (epicenters), with delimitation of the rectangular domain used for study on the contoured region of Andalusia. With the aim of assessing the effects of the occurrence time on the structural complexity displayed by the data we combine the exposed techniques with different transformations on the temporal component (or, more precisely, on the space-time association).

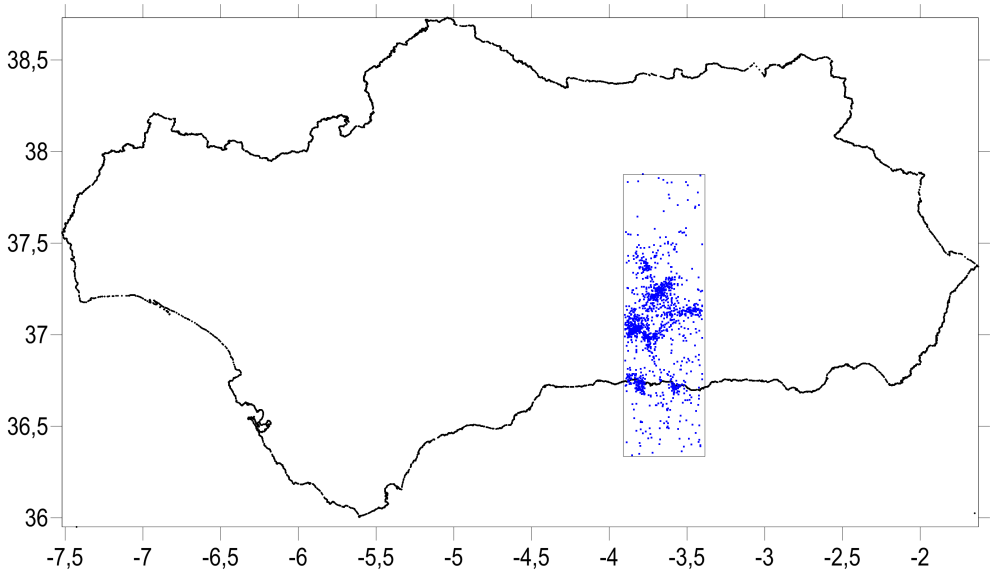

Figure 4.1: Agrón data. Epicenter locations on contoured Andalusian region.

The study is structured in three parts. Firstly, we carry out a descriptive analysis of the epicenter location as well as the time, depth, and magnitude components of the seismic series. Secondly, we perform a general analysis of the dimensional interaction between the epicenter location and the occurrence time, depth or magnitude components. Subsequently, we analyze the structural characteristics of the interactions $(X,Y) \leftrightarrow T$ and $(X,Y) \leftrightarrow Z$ with and without considering the effects of the magnitude at different scales by means of generalized dimensions over time. Moreover, we carry out a deeper study of the spatio-temporal interaction by means of the combinations of the former techniques with different transformations on the temporal component.

4.1.1 Descriptive analysis

Before applying the exposed techniques to the data we carry out a first exploration by means of graphical descriptive analyses. Firstly, differences in the space-time structure of points, and particularly in clustering patterns for the whole period and within subperiods are visualized in the plots in Figures 4.2, 4.3, 4.4 and 4.5 (in each case, time is distinguished by colour transitions, and the size of each point represents individual magnitude or depth, from left to right). These plots show that in the whole period and in the three subperiods there is a strong clustering, especially during the central period where there is a marked cluster. In the plots representing the whole period we can also observe spatio-temporal associations. However, regarding the relations between magnitude/depth and the locations and time we can not detect clear patterns.

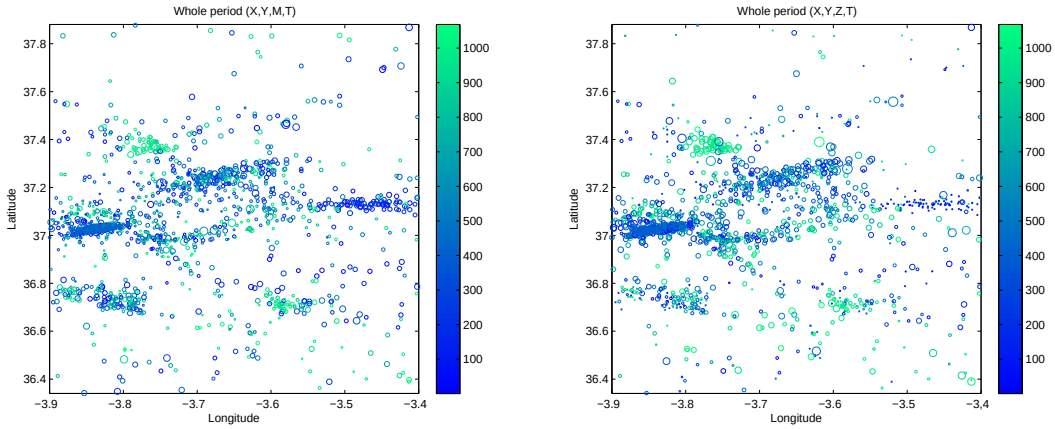

Figure 4.2: Agrón data. Epicenter locations for the whole period by time, represented by colour bar, and magnitude (left) or depth (right), represented by size of circle mark.

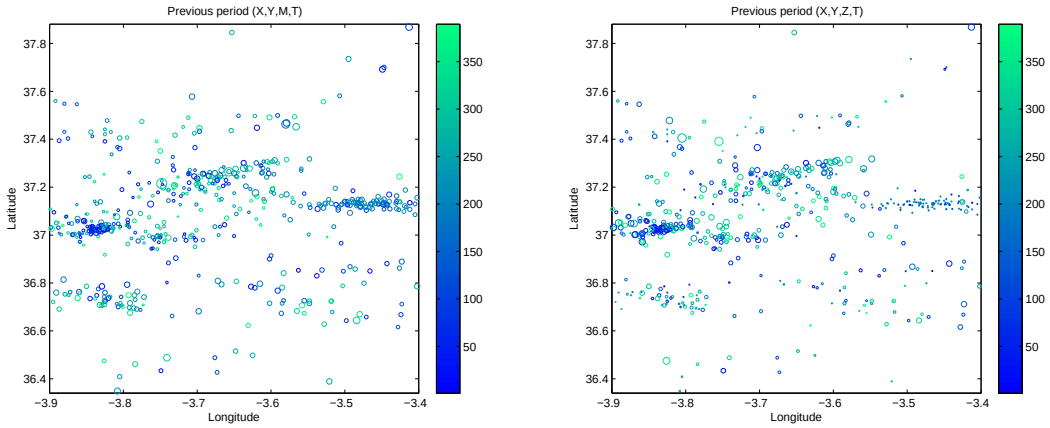

Figure 4.3: Agrón data. Epicenter locations for the previous period by time, represented by colour bar, and magnitude (left) or depth (right), represented by size of circle mark.

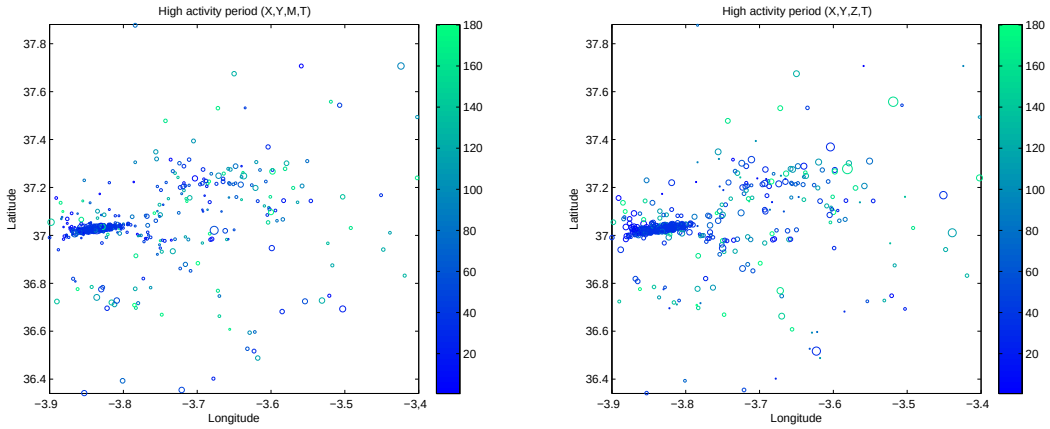

Figure 4.4: Agrón data. Epicenter locations for the high seismicity activity period by time, represented by colour bar, and magnitude (left) or depth (right), represented by size of circle mark.

We perform a second descriptive analysis in order to study the daily frequency of the events during the whole period and within subperiods, see Figure 4.6. In the top plot we can detect a segment of three days with high seismic activity during December 1998. Moreover, in all the plots the daily frequencies show typical variations of a multifractal behaviour, properly quantified later.

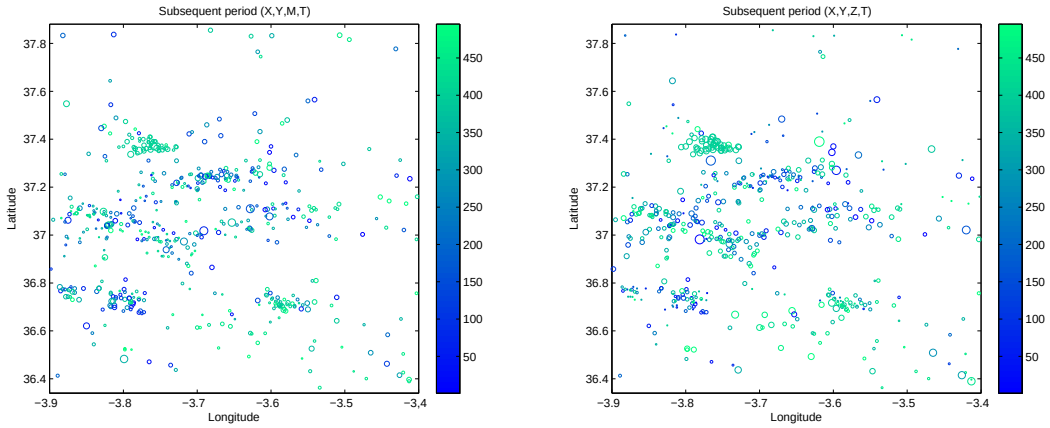

Figure 4.5: Agrón data. Epicenter locations for the subsequent period by time, represented by colour bar, and magnitude (left) or depth (right), represented by size of circle mark.

The third analysis deals with the study of the magnitude jointly with its association with time. As mentioned before, the magnitude component has a special relevance in the study of seismicity and, therefore, its study here is more complete than the study of depth. The distribution of the magnitude over time is plotted in Figure 4.7, by means of which we can assess the degree of association between both components. In these plots we can observe how during the first subperiod there is a lack of low magnitude events but, overall, we can not detect clear trends over time. Finally, in Figures 4.8 and 4.9 we represent the histogram and the reciprocal accumulated distribution of the event magnitudes for the whole period and the three subperiods.

To finish the descriptive analysis, we complete the assessment on depth analyzing its degree of association with time; see Figure 4.10. It must be stated that, as in the study of the magnitude study, we cannot detect clear temporal evolutionary patterns.

4.1.2 Study of dimensional interaction

In this subsection, we study dimensional interaction between the epicenter location (X,Y) and the time T, depth Z and magnitude M components. We perform a brief analysis of the three cases and subsequently we carry out a broader study of the spatio-temporal case. The techniques used are based on entropy measures and the dependence coefficient introduced in Chapter 3.

Firstly, we perform a short analysis of the dimensional interaction between the epicenter location and the time, depth and magnitude components by means of Shannon entropy and the 'dependence coefficient'. The Shannon entropy values $H(D)$, $H(X,Y)$, $H(X,Y,D)$, $H(X,Y)+$

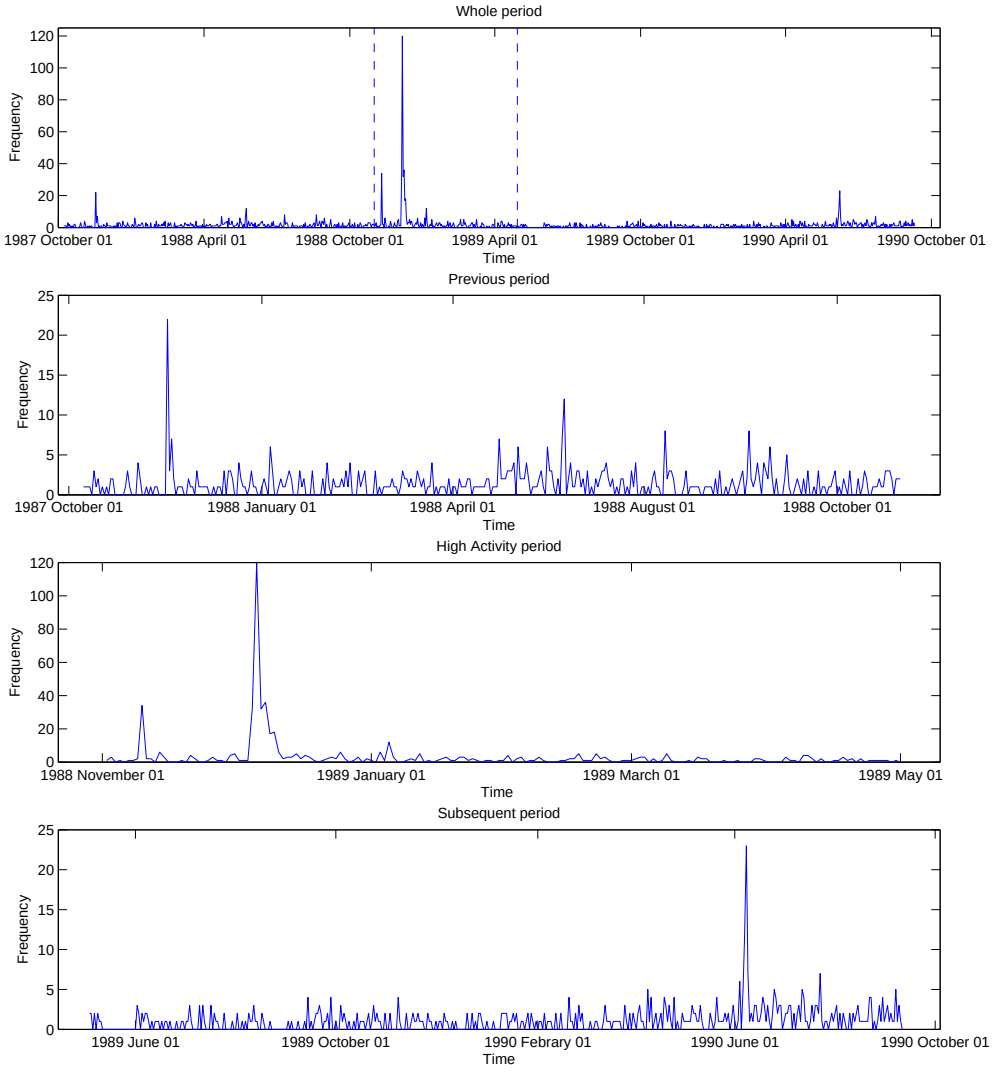

Figure 4.6: Agrón data. Daily frequencies for the whole period and the three subperiods.

$H(D)$ and $H(X,Y)+H(D)-H(X,Y,D)$ are calculated using varying box sizes between $1/15$ and $1/2$ for the whole period and the three subperiods, considering as D component T, Z or M; see Figure 4.11. In these plots we can observe how the departures (brown line) for the three cases show that the degree of dependence between (X,Y) and Z is lower than for T and M.

The next step is the assessment of the degree of dependence over time by means of the values of the dependence coefficient calculated for data segments based on sliding windows, considering window sizes of 150 events with an overlapping of 90 per cent; see Figure 4.12.

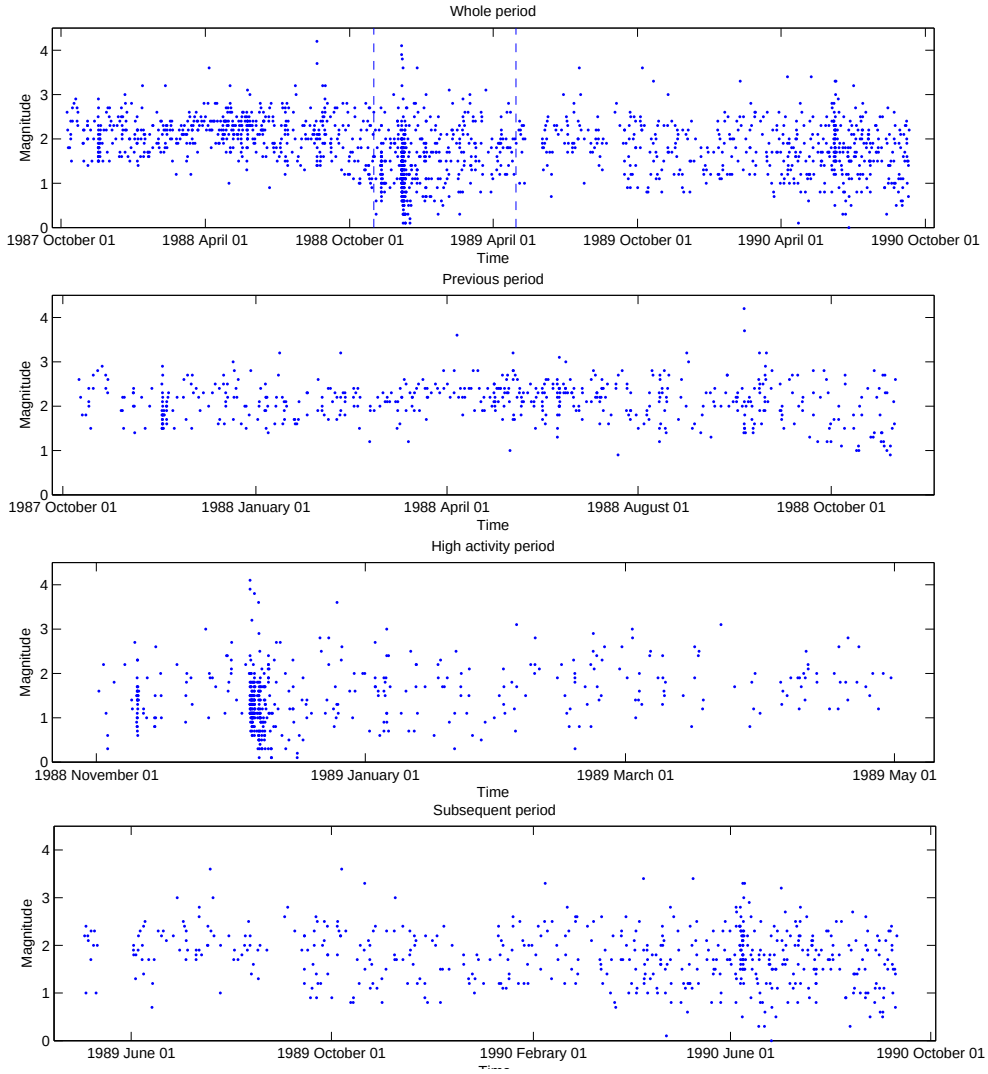

Figure 4.7: Agrón data. Magnitude distribution over time for the whole period and the three subperiods.

The curves in these plots show that the degree of dependence between the spatial location and the occurrence time is markedly higher than in the case of the other components in the regular activity periods. However, during the high activity period the degree of dependence decreases strongly in the three cases, in particular regarding the association $(X,Y) \leftrightarrow T$ and $(X,Y) \leftrightarrow M$.

These results show that the degree of association between the spatial component (X,Y) and occurrence time T is significantly higher than for magnitude M and depth Z, and that the tem-

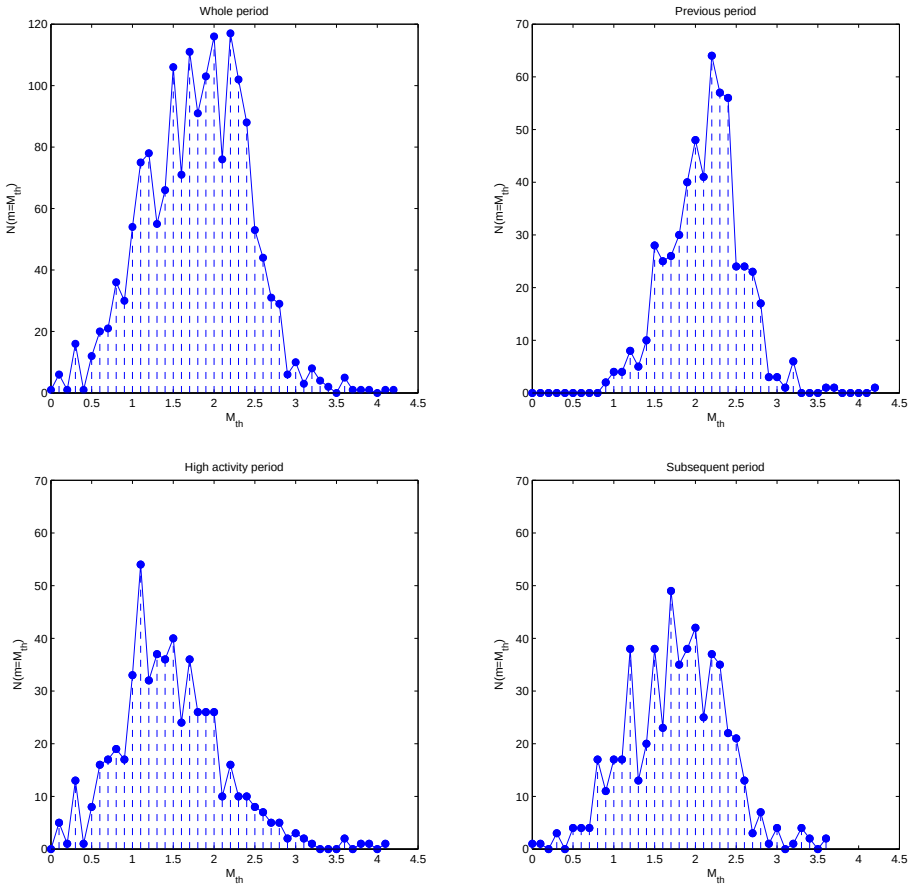

Figure 4.8: Agrón data. Frequency-magnitude histogram for the whole period and the three subperiods.

poral changes in the degree of dependence between (X,Y) and T are stronger than for the other cases. Consequently, we perform a deeper study of spatio-temporal interaction. For a better assessment of the effects of the occurrence time on the structural complexity displayed by the data, we combine the techniques with different transformations on the temporal component or, more precisely, on the space-time association. Specifically, as mentioned in Chapter 3, apart from the registered 'real time' (TR), we consider the 'natural time' (TN), 'shuffled time' (TS) and 'random uniform time' (TU). A first exploration of the space-time interaction can be obtained by looking at the different 3D scatter plots based on the components (X,Y,T) representing longitude, latitude, and time in correspondence to the former time-related transformations; see Figure 4.13. In the top left plot we can observe the existence of strong spatio-temporal clusters.

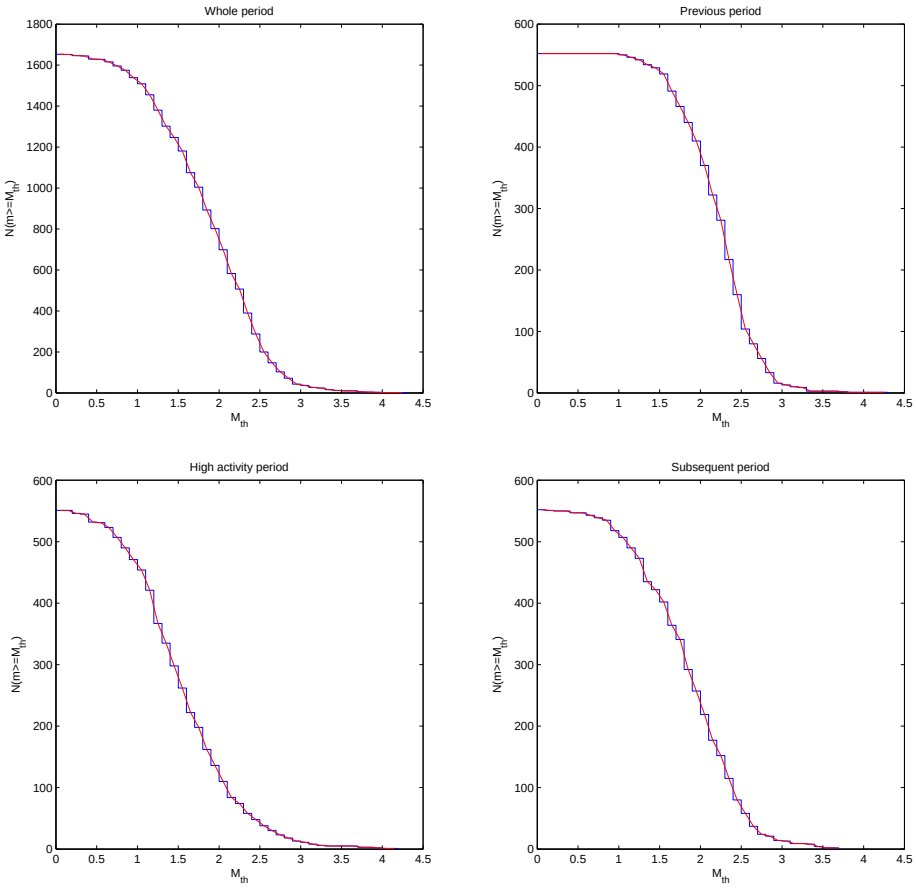

Figure 4.9: Agrón data. Magnitude reciprocal accumulated distribution for the whole period and the three subperiods.

Likewise, when we consider the natural time these clusters are clearly visible although they exhibit a vertical deformation due to the uniformization of interevent times. In the bottom left plot we can observe how the temporal ordering is maintained whereas the space-time association is eliminated. Finally, in the bottom right plot, unlike the other cases, the spatio-temporal association jointly with the temporal structure is removed and only the purely spatial clustering is kept.

To assess the temporal changes in the partial elimination of the spatio-temporal association produced by the different temporal transformations, as observed in the previous exploration, we perform an analysis based on Shannon entropy by means of sliding windows. Figure 4.14, displays Shannon entropy values $H(T)$, $H(X,Y)$, $H(X,Y,T)$, $H(X,Y)+H(T)$ and $H(X,Y)+$

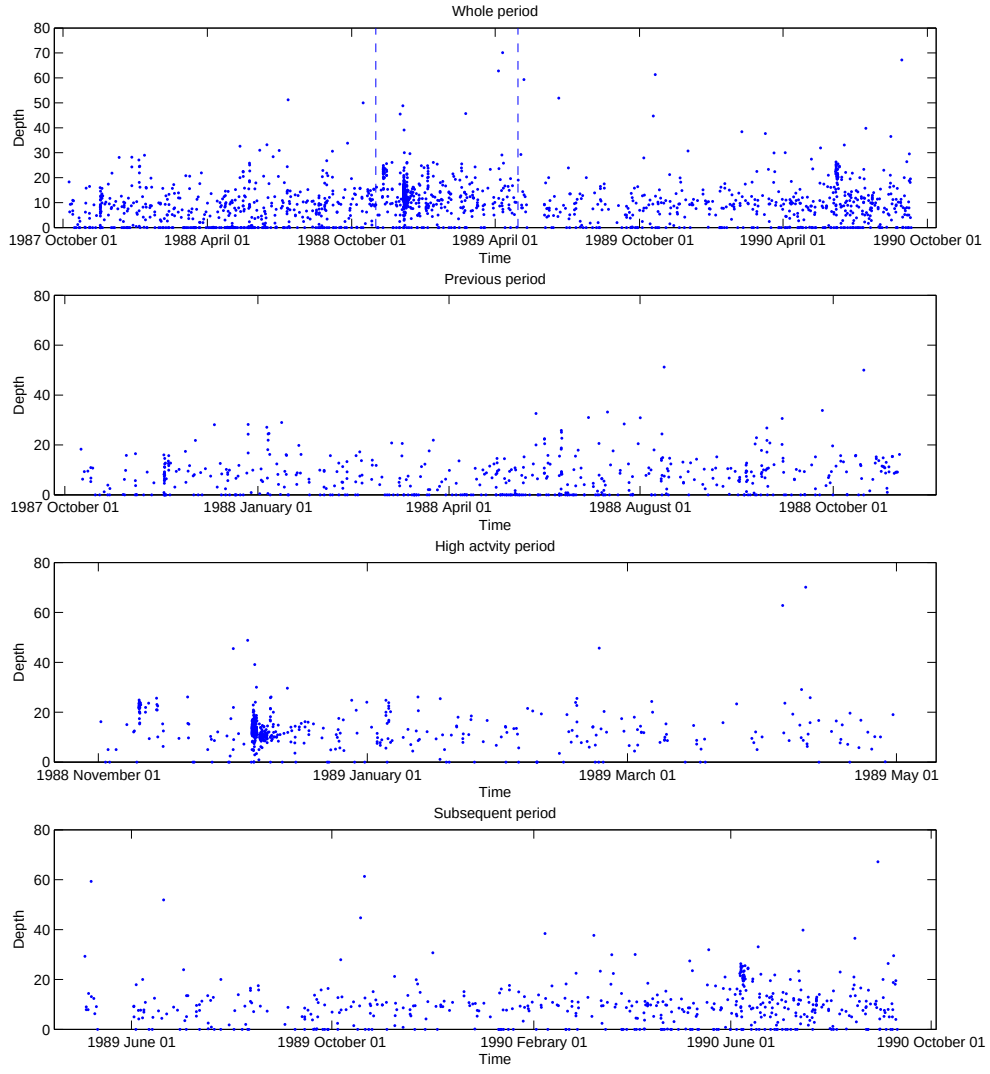

Figure 4.10: Agrón data. Depth distribution over time for the whole period and for the three subperiods (from top to bottom).

$H(T)-H(X,Y,T)$ in relation to the four different scenarios for TR, TN, TS, and TU considering a specific box size $\varepsilon = 1/10$. The top plot shows that during the days of higher seismic activity the strength of structuring in (X,Y,T) increases whereas the level of association between the occurrence time and the epicenter location decreases. When the interevent time information is discarded by assigning equally spaced time intervals the spatio-temporal association is only slightly affected. The third plot shows that during the high activity days the original temporal

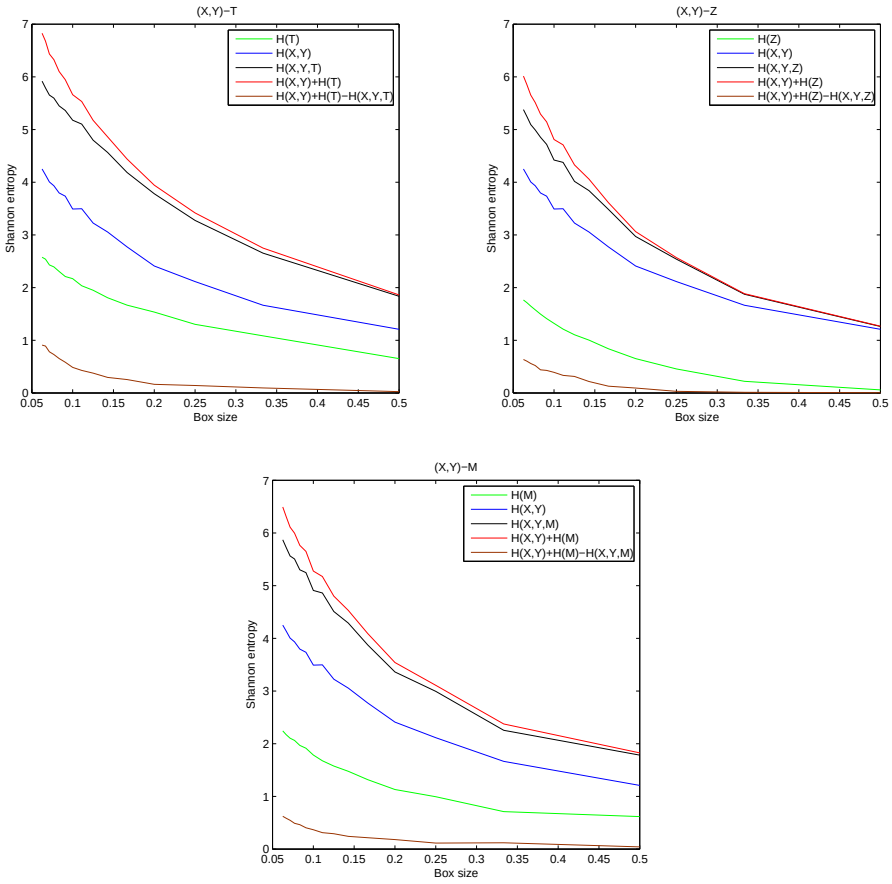

Figure 4.11: Agrón data. Shannon entropy values $H(D)$, $H(X,Y)$, $H(X,Y,D)$, $H(X,Y)+H(D)$ and $H(X,Y)+H(D)-H(X,Y,D)$, for varying box sizes between 1/15 and 1/2 with D corresponding to time, depth and magnitude.

ordering of the epicenter location is eliminated and the decrease of the level of disorder in the spatio-temporal interaction is only due to the temporal component. The bottom plot corresponds to the elimination of the original temporal structure and the space-time association, as it can be noticed in the corresponding curves.

Once we have shown that the central period is the one exhibiting a higher spatio-temporal association, we focus on the analysis of this subperiod to obtain clearer results. With this purpose, the analysis is now performed combining the mentioned temporal transformations with the techniques only for this subperiod. Plots of Figure 4.15, in correspondence to the four different scenarios for TR, TN, TS, and TU, display Shannon entropies $H(T)$, $H(X,Y)$,

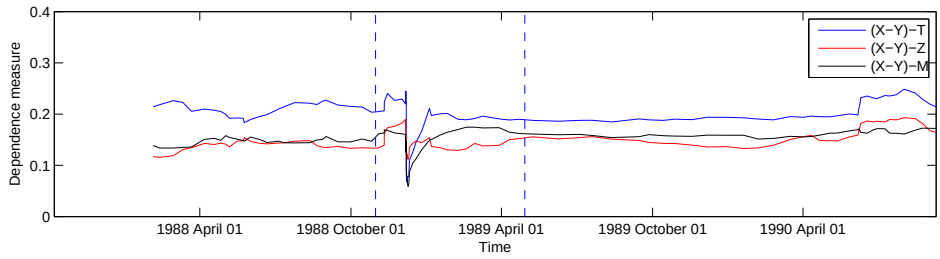

Figure 4.12: Agrón data. Values of the dependence coefficient for the interactions $(X,Y) \leftrightarrow T$, $(X,Y) \leftrightarrow Z$ and $(X,Y) \leftrightarrow M$, based on sliding windows (size 150 events).

$H(X,Y,T)$ and $H(X,Y)+H(T)$, calculated using varying box sizes between 1/2 and 1/15. The presence of a certain degree of spatio-temporal interaction, in terms of departure of $H(X,Y,T)$ from $H(X,Y)+H(T)$, is reflected, in particular, in the transition from the top plots to the bottom ones, due to destruction of the real time ordering associated with the data spatial coordinates. Moreover, purely temporal structuring of the data can be assessed by the entropy $H(T)$ increase induced in the right plots compared to the original in the left plots.

As for Rényi entropy values, namely $H_q(T)$, $H_q(X,Y)$, $H_q(X,Y,T)$ and $H_q(X,Y)+H_q(T)$, here we select for display, as before, a specific box size $\varepsilon = 1/10$, which corresponds to a middle range where a certain computational scaling stability is achieved. Pairwise comparisons between the four plots in Figure 4.16 provide interesting interpretations, in the same lines, regarding dimensional interaction, specially looking at varying non-negative values of the deformation parameter q.

To finish, in Figure 4.17 the Tsallis entropy values are displayed considering a specific box size $\varepsilon = 1/10$ and a variation of q between 1 and 3. In this case, the degree of spatio-temporal interaction is reflected in terms of departure of $T_q(X,Y,T)$ from $T_q(X,Y)+T_q(T)-(q-1)T_q(X,Y)T_q(T)$. We must highlight that the results obtained have a similar interpretation than the former analysis.

4.1.3 Study of structural characteristics at different scales

The next step is to study the structural characteristics of the seismic series in relation to the scaling properties inherent to the phenomenon by means of multifractal tools. Firstly, we perform an analysis of the structural changes for (X,Y,T) and (X,Y,Z) over time considering the effects of the magnitude. Secondly, since the case (X,Y,T) shows a richter structure, we complete the analysis combining the techniques with the different time transformations.

The analysis of the structural characteristics of (X,Y,T) and (X,Y,Z) over time is performed by means of the generalized dimensions D_0, D_1, D_2, multifractal step $D_{-\infty} - D_{\infty}$, and the com-

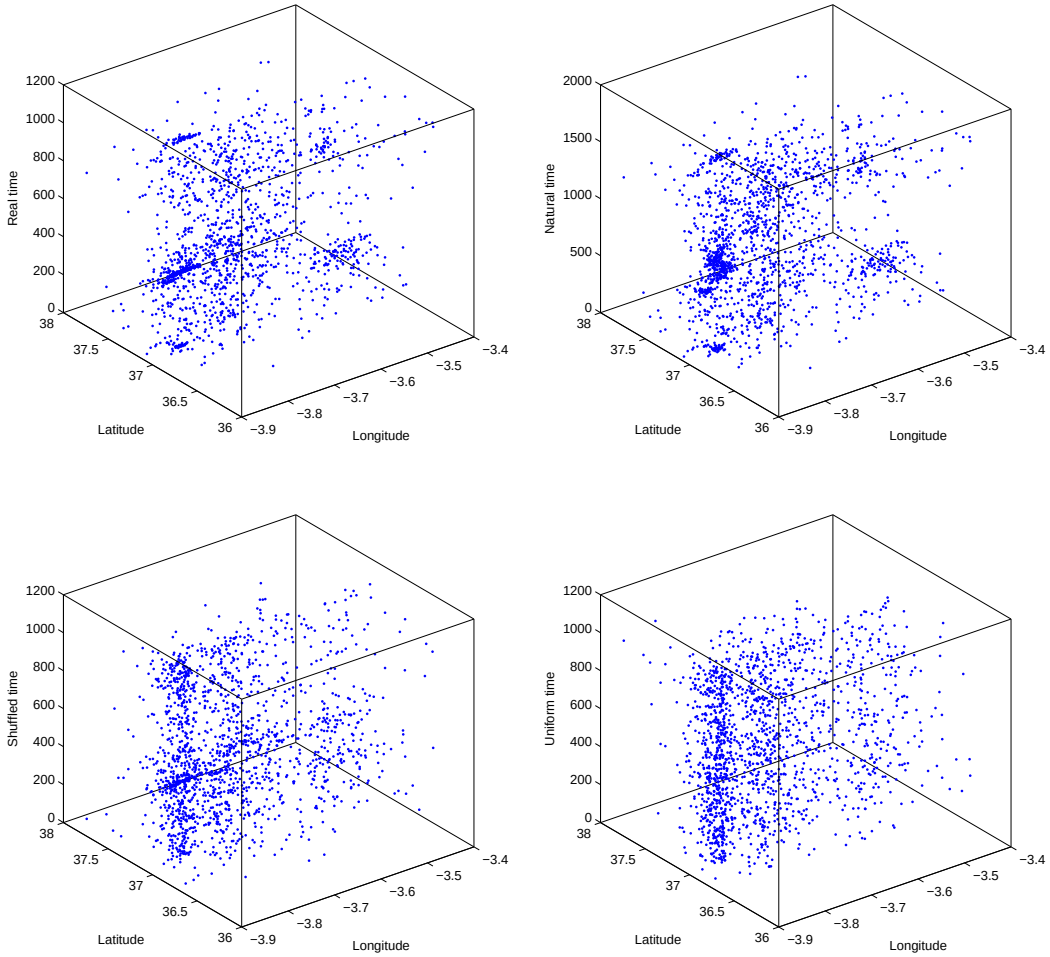

Figure 4.13: Agrón data. 3D scatter plots of components (X,Y,T), corresponding to TR, TN, TS and TU.

plexity measure $D_1 - D_2$ calculated on sliding windows, with and without considering magnitude effects. In addition, the results for the LMC_{exp} are calculated for a box size $\varepsilon = 1/10$, with the aim of comparing to the limiting values given by $D_1 - D_2$. The window sizes are taken of 150 events with an overlapping of 90 per cent and the effects of the magnitude are incorporated considering the weighting function $\phi(m) = \exp(3m/2)$.

In Figures 4.18, 4.19, and 4.20, which display the D_0, D_1, and D_2 values over time, we can observe how during the days with greater concentration of events the covering of the study area worsens whereas the degrees of ordering and clustering increase markedly for both cases (we

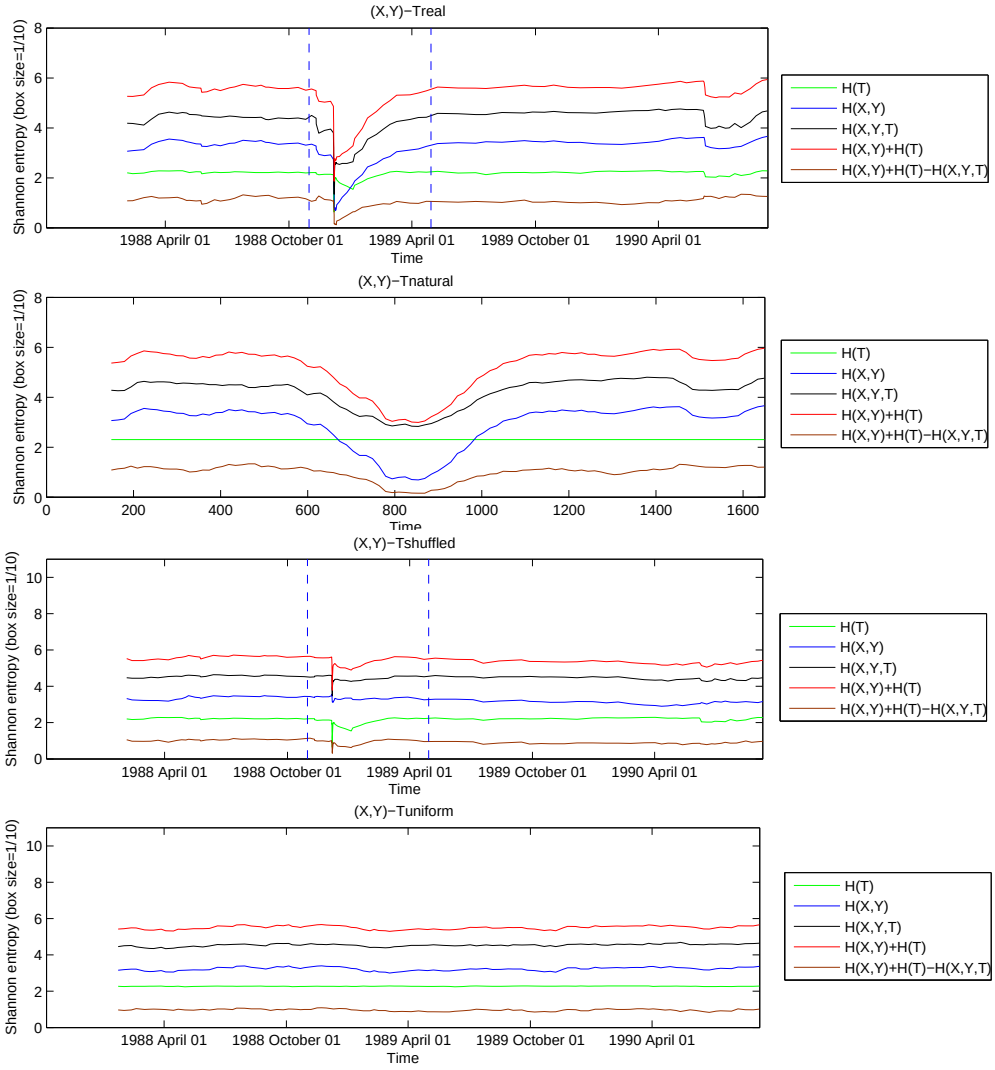

Figure 4.14: Agrón data. Shannon entropy values $H(T)$, $H(X,Y)$, $H(X,Y,T)$, $H(X,Y)+H(T)$ and $H(X,Y)+H(T)-H(X,Y,T)$ based on sliding windows (size 150 events), considering a fixed box size $\varepsilon = 1/10$, corresponding to TR, TN, TS and TU.

must note that D_0 is not affected by function $\phi(m)$). However, it is interesting to note that these features are more visible for (X,Y,T). In agreement with these results, the multifractal step increases during the same days; see Figure 4.21. An analogous interpretation can be obtained when analyzing complexity by means of the $D_1 - D_2$ and LMC_{exp} measures; see Figures 4.22 and 4.23. As expected, the corresponding results display similar patterns.

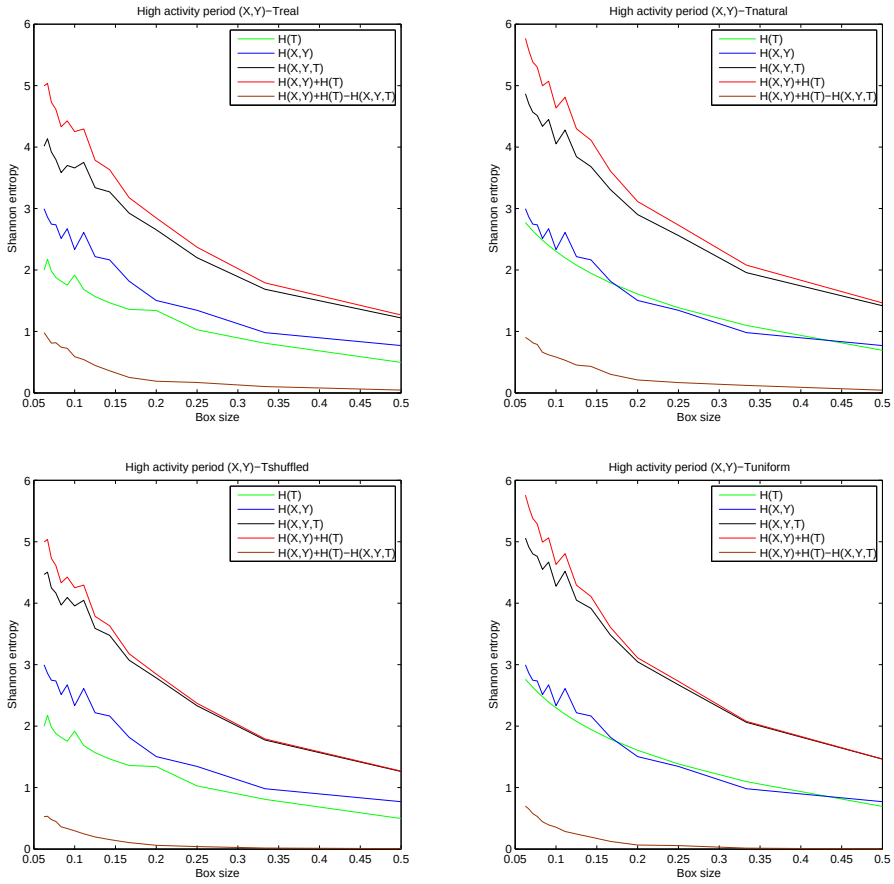

Figure 4.15: Agrón data. Shannon entropy values $H(T)$, $H(X,Y)$, $H(X,Y,T)$, $H(X,Y)+H(T)$ and $H(X,Y)+H(T)-H(X,Y,T)$ for varying box sizes between 1/15 and 1/2, corresponding to TR, TN, TS and TU.

Finally, when we compare the results with those obtained incorporating the magnitude effects we can observe how the two corresponding curves in each case follow a similar bahaviour, suggesting that the distribution of magnitude in general has an intrinsic structural association with that of space-time coordinates. Interestingly, however, deviations are more significant in association with high activity episodes whose nature could be subject to discrimination depending on the sign of deviates (see, for instance, a clear reversion of the curves of $D_1 - D_2$ obtained for (X,Y,T), comparing the central subperiod and the final part of the subsequent subperiod). In all cases, a drastic increase in the degree of multifractality is observed in association with the concentration of seismic activity in the central subperiod.

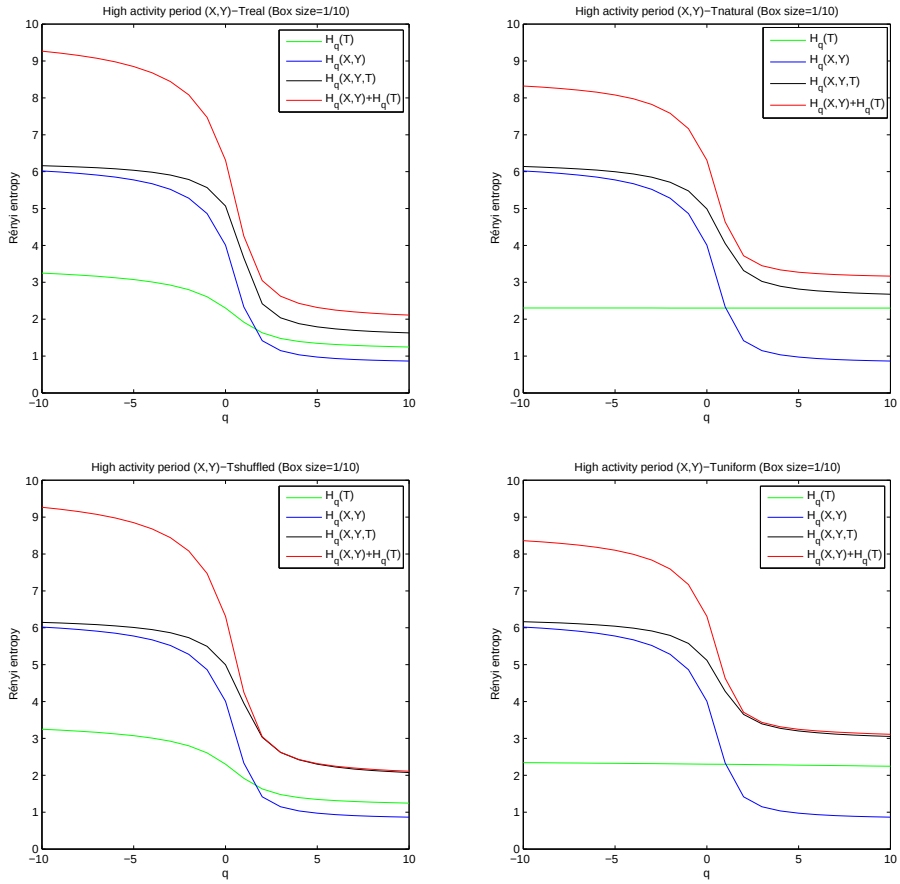

Figure 4.16: Agrón data. Rényi entropy values $H_q(T)$, $H_q(X,Y)$, $H_q(X,Y,T)$, $H_q(X,Y)+H_q(T)$ and $H(X,Y)+H(T)-H(X,Y,T)$ for fixed box size $\varepsilon = 1/10$ and varying deformation parameter q between -10 and 10, corresponding to TR, TN, TS and TU.

To complete the study of the structural characteristics of (X,Y,T) we combine the techniques for scaling analysis with temporal transformations. Firstly, we calculate the generalized dimensions and the multifractal spectrum for the three subperiods of the series. Figure 4.24 shows the generalized dimension curves within each subperiod for the different time-related transformations. These plots show clearly different features between the subperiods. The range in the curves indicate a higher degree of multifractality in the central seismic subseries. Note also that the transformation TS has stronger effects in the 'pre' and 'post' subperiods, suggesting that the association between space and time coordinates has a relatively larger contribution to the spatio-temporal interaction in such subperiods. Furthermore, although generalized dimen-

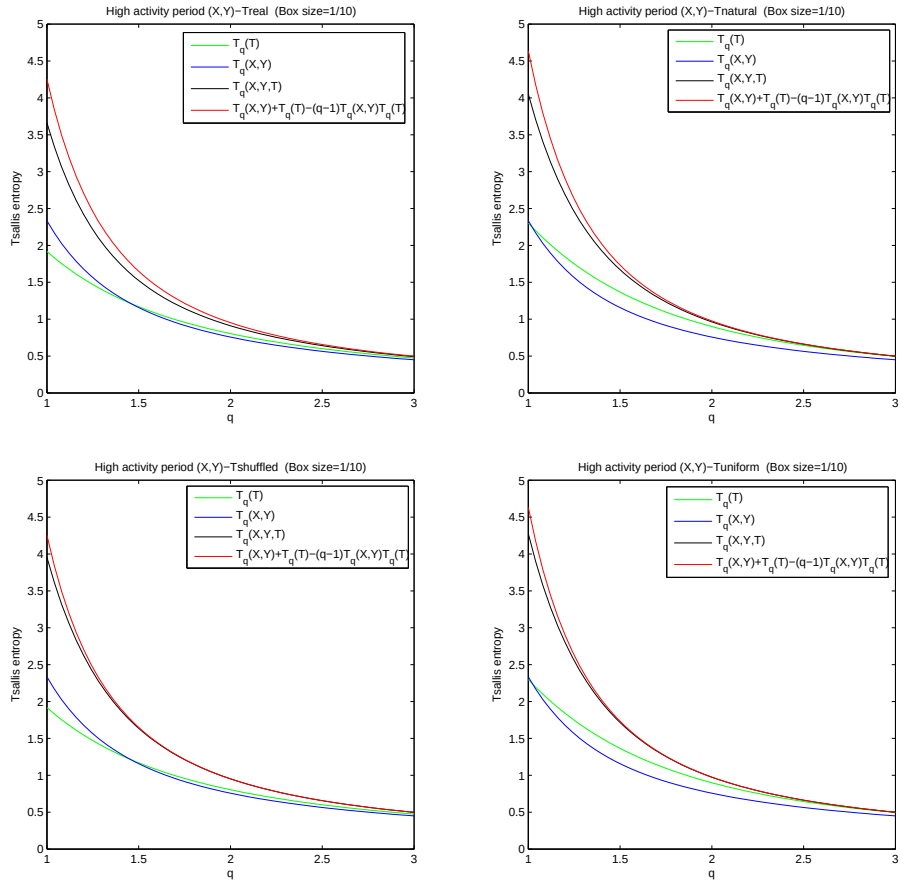

Figure 4.17: Agrón data. Tsallis entropy values $HT_q(T)$, $HT_q(X,Y)$, $HT_q(X,Y,T)$ and $HT_q(X,Y)+HT_q(T)-(q-1)HT_q(X,Y)HT_q(T)$ for fixed box size $\varepsilon = 1/10$ and varying deformation parameter q between 1 and 3, corresponding to TR, TN, TS and TU.

sions are similar between the 'pre' and 'post' subperiods, spatio-temporal interaction shows a different structure, as provided by the curves corresponding to transformed time coordinates, specially under TN and TU; that is, in relation to the interevent times, which have a higher significance in structuring in the 'post' subperiod. In the three curves, the high values of D_q for $q < 1$ show that in all cases there are areas poorly represented while the low D_q values for $q > 1$ in the period of high activity reflect of the existence of well represented areas.

These interpretations are confirmed from the values of the coefficients D_0, D_1, D_2, $D_1 - D_2$ and $D_{-\infty} - D_{\infty}$ shown in Table 4.4. The values of the 'capacity' dimension are smaller during the central subperiod, which involves that the covering of the study area is worse whereas the

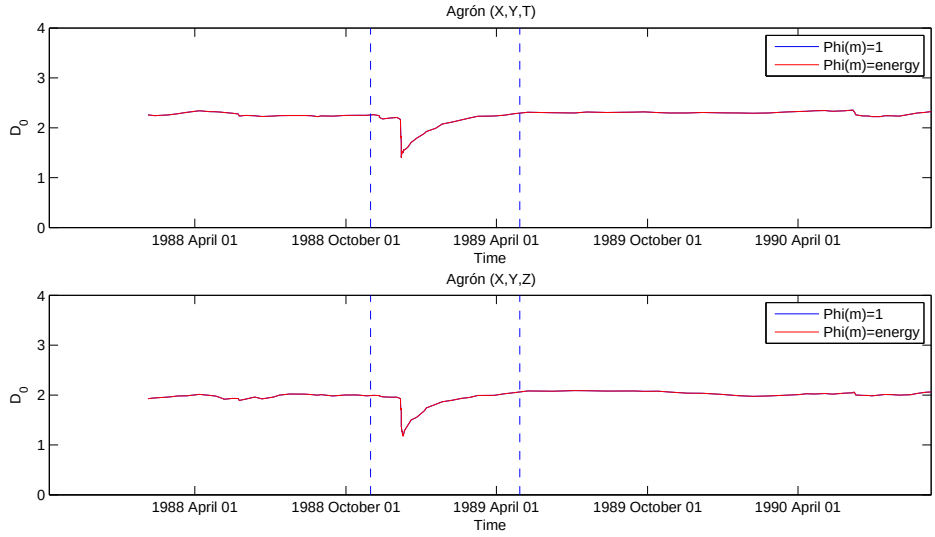

Figure 4.18: Agrón data. Variations in the 'capacity' dimension values over time based on sliding windows (size 150 events), with and without considering the effects of the magnitude.

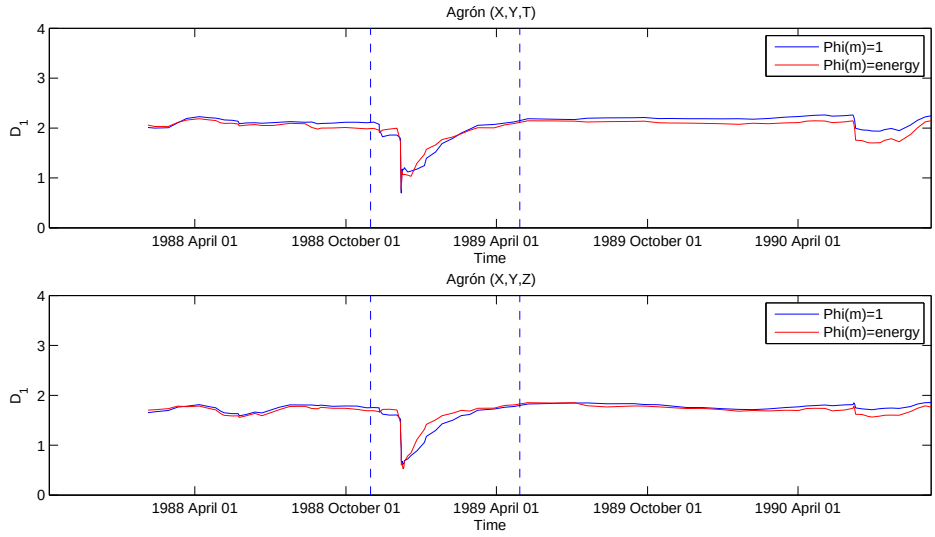

Figure 4.19: Agrón data. Variations in the 'entropy' dimension values over time based on sliding windows (size 150 events), with and without considering the effects of the magnitude.

values of the 'information' and 'correlation' dimensions are lower, showing the existence of a higher degree of ordering and spatial clustering during this period. Moreover, high values of

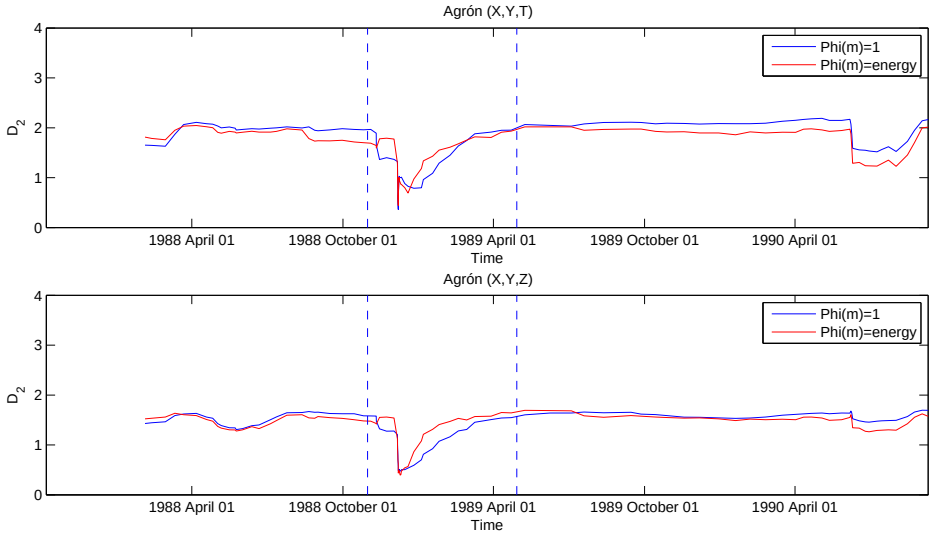

Figure 4.20: Agrón data. Variations in the 'correlation' dimension values over time based on sliding windows (size 150 events), with and without considering the effects of the magnitude.

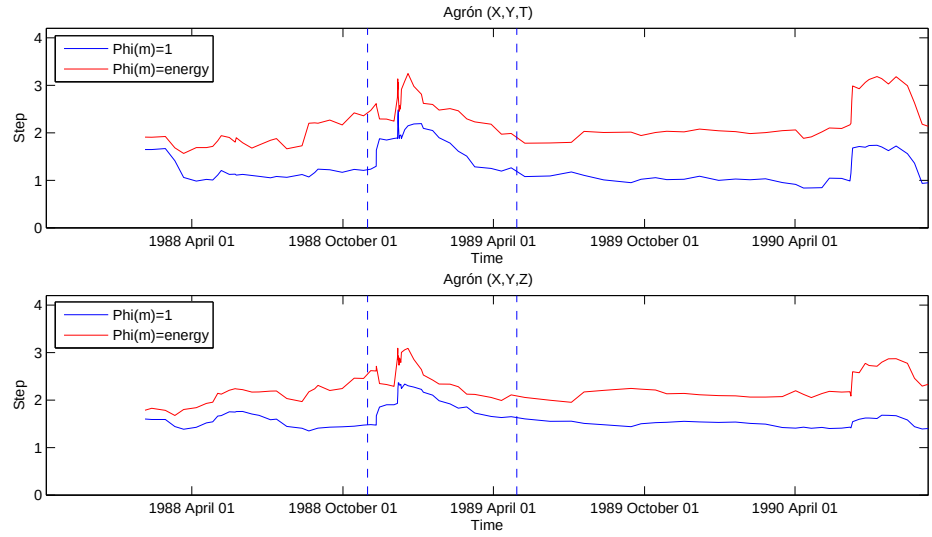

Figure 4.21: Agrón data. Variations in the multifractal step values over time based on sliding windows (size 150 events), with and without considering the effects of the magnitude.

the $D_{-\infty} - D_{\infty}$ and $D_1 - D_2$ reflect that the degrees of multifractality and complexity are higher during the central subperiod.

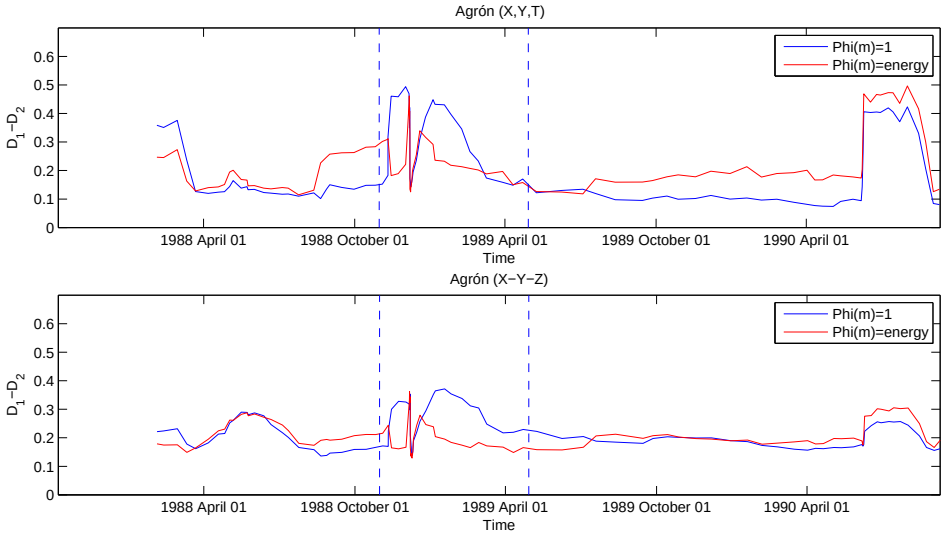

Figure 4.22: Agrón data. Variations in the values of the complexity measure $D_1 - D_2$ over time based on sliding windows (size 150 events), with and without considering the effects of the magnitude.

A similar interpretation can be obtained comparing the three related plots of multifractal spectra displayed in Figure 4.25. Here, we can see that the range of local fractality coefficients (Hölder exponents) is much richer in the central period, which partially points out a higher degree of complexity in relation to the contribution of the information factor. In these plots, the peaks ('capacity' dimension) indicate that in the 'pre' and 'post' subperiods the filling of the study area is higher. The slopes of the left branches reflect strong spatial clusters in the high activity period and those of the right branches reflect the existence of poorly represented areas in the three subperiods.

These results are compared to those obtained in terms of generalized dimensions based on the non-extensive statistics of Tsallis, proposed in Chapter 3. The Tsallis generalized dimensions for $q = 0, 1$ and 2 are calculated for the whole period and the three subperiods, and collected together with the corresponding generalized Rényi dimensions in Table 4.2. We must state that the interpretations derived in both cases are similar.

Finally, for a better assessment of the structural evolutionary changes, the study is complemented with the examination of the generalized dimensions for data segments based on sliding windows, again taken with a window size of 150 and an overlapping of 90 per cent; see Figure 4.26.

The values of $D_{-\infty} - D_{\infty}$ for (X, Y, TR) show that the degree of multifractality is higher during the periods of high seismic activity. This fact occurs in association with a strong decrease

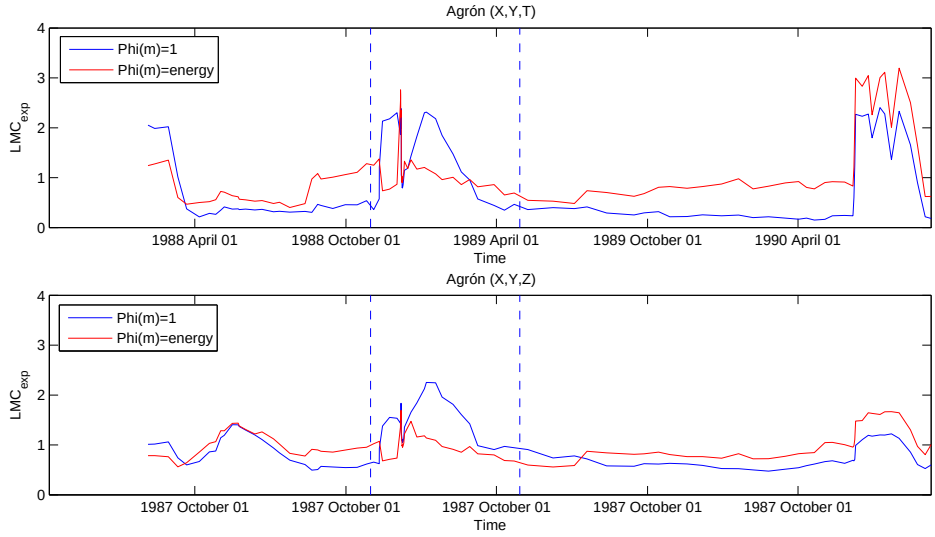

Figure 4.23: Agrón data. Variations in the values of the exponential version LMC over time based on sliding windows (size 150 events), with and without considering the effects of the magnitude.

in both D_1 and D_2, which indicates an increase in the degrees of ordering and clustering.

When we consider TN as temporal component, the temporal ordering is preserved but the interevent time information is discarded. This effect is visible in the second plot where we can also see the increase in the degrees of multifractality, ordering and clustering observed in the central period.

As for TS, represented in the third plot, we keep the temporal component but the spatial locations are randomly changed. Here, the increase in the degrees of ordering, clustering and multifractalilty in the period with high seismic activity is essentially due to the temporal structure.

Finally, the bottom plot, representing TU, is the extreme scenario where we lose all information about the real ordering and interevent times. In this case, there are not evolutionary changes, as expected.

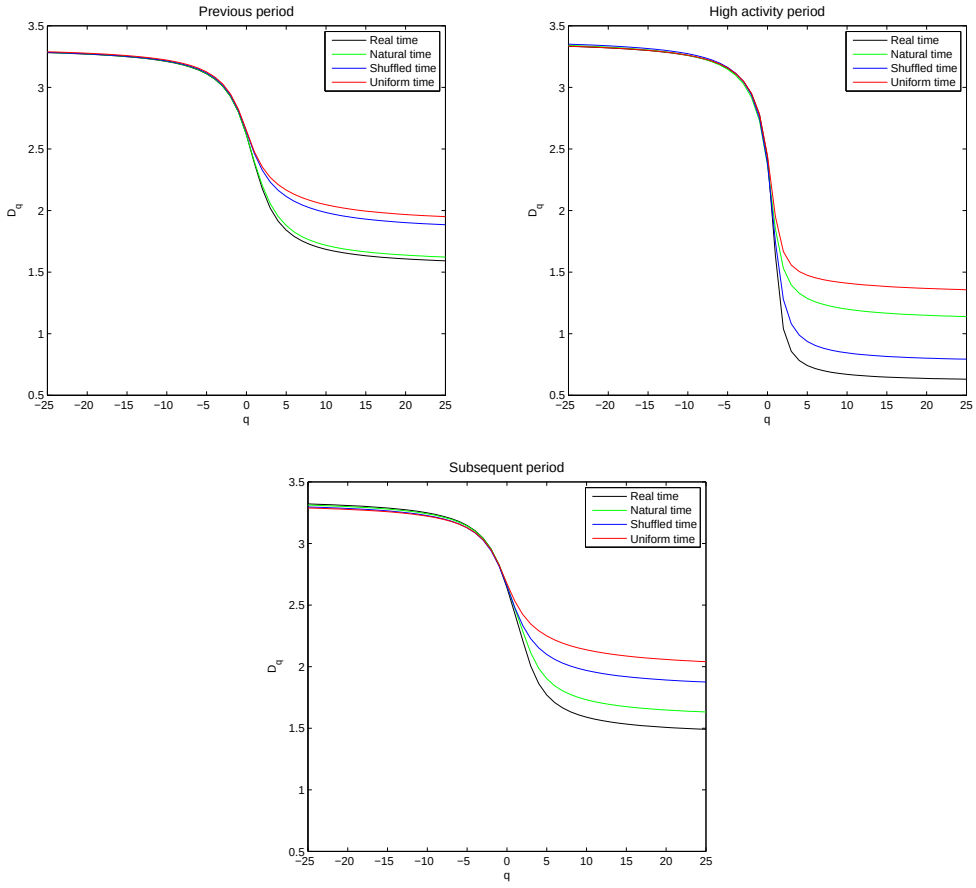

Figure 4.24: Agrón data. Generalized Rényi dimension curves corresponding to TR, TN, TS and TU, for the three subperiods (from letf to right).

Table 4.1: Agrón data. Values of generalized Rényi dimensions D_0, D_1, D_2, $D_1 - D_2$ and $D_{-\infty} - D_{\infty}$, corresponding to TR, TN, TS and TU, for the three subperiods.

Previous period	D_0	D_1	D_2	$D_1 - D_2$	$D_{-\infty} - D_{\infty}$
Real time	2.6111	2.3895	2.1771	0.2124	1.6893
Natural time	2.6115	2.3945	2.1995	0.1950	1.6739
Shuffled time	2.6439	2.4662	2.3323	0.1339	1.4385
Uniform time	2.6438	2.4728	2.3423	0.1305	1.3583
High activity period	D_0	D_1	D_2	$D_1 - D_2$	$D_{-\infty} - D_{\infty}$
Real time	2.4300	1.6363	1.0389	0.5974	2.7028
Natural time	2.3696	1.8503	1.5304	0.3199	2.2015
Shuffled time	2.3789	1.7457	1.2769	0.4688	2.5576
Uniform time	2.4454	1.9507	1.6666	0.2841	1.9771
Subsequent period	D_0	D_1	D_2	$D_1 - D_2$	$D_{-\infty} - D_{\infty}$
Real time	2.6399	2.4294	2.2091	0.2203	1.8299
Natural time	2.6595	2.4724	2.2845	0.1879	1.6798
Shuffled time	2.6525	2.4778	2.3321	0.1457	1.4216
Uniform time	2.6735	2.5304	2.4227	0.1077	1.2478

Table 4.2: Agrón data. Comparison of generalized Rényi and Tsallis dimension values, for $q = 0, 1, 2$.

	D_0	DT_0	D_1	DT_1	D_2	DT_2
Agrón – subperiod 1	2.6111	27.2807	2.3895	2.3895	2.1771	1.2007
Agrón – subperiod 2	2.4300	18.0316	1.6363	1.6363	1.0389	1.0036
Agrón – subperiod 3	2.6399	29.0456	2.4294	2.4294	2.2091	1.2018
Agrón – whole period	2.7820	41.5404	2.3256	2.3256	1.8041	1.1737

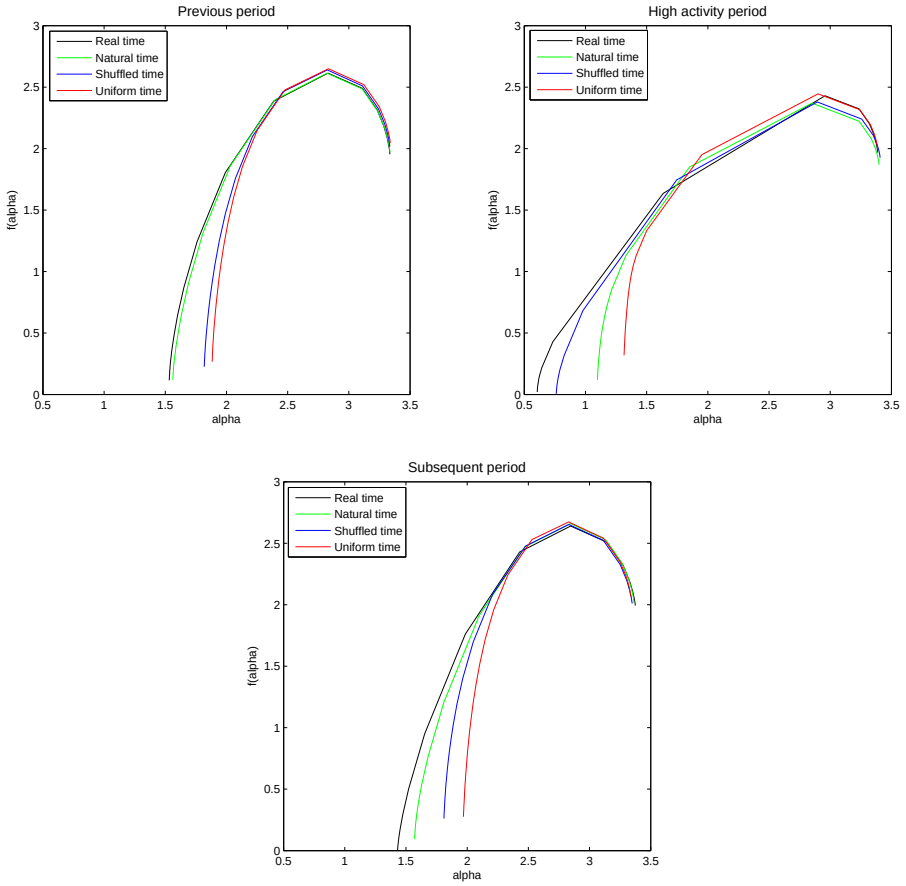

Figure 4.25: Agrón data. Multifractal spectra corresponding to TR, TN, TS and TU, for the three subperiods (from letf to right).

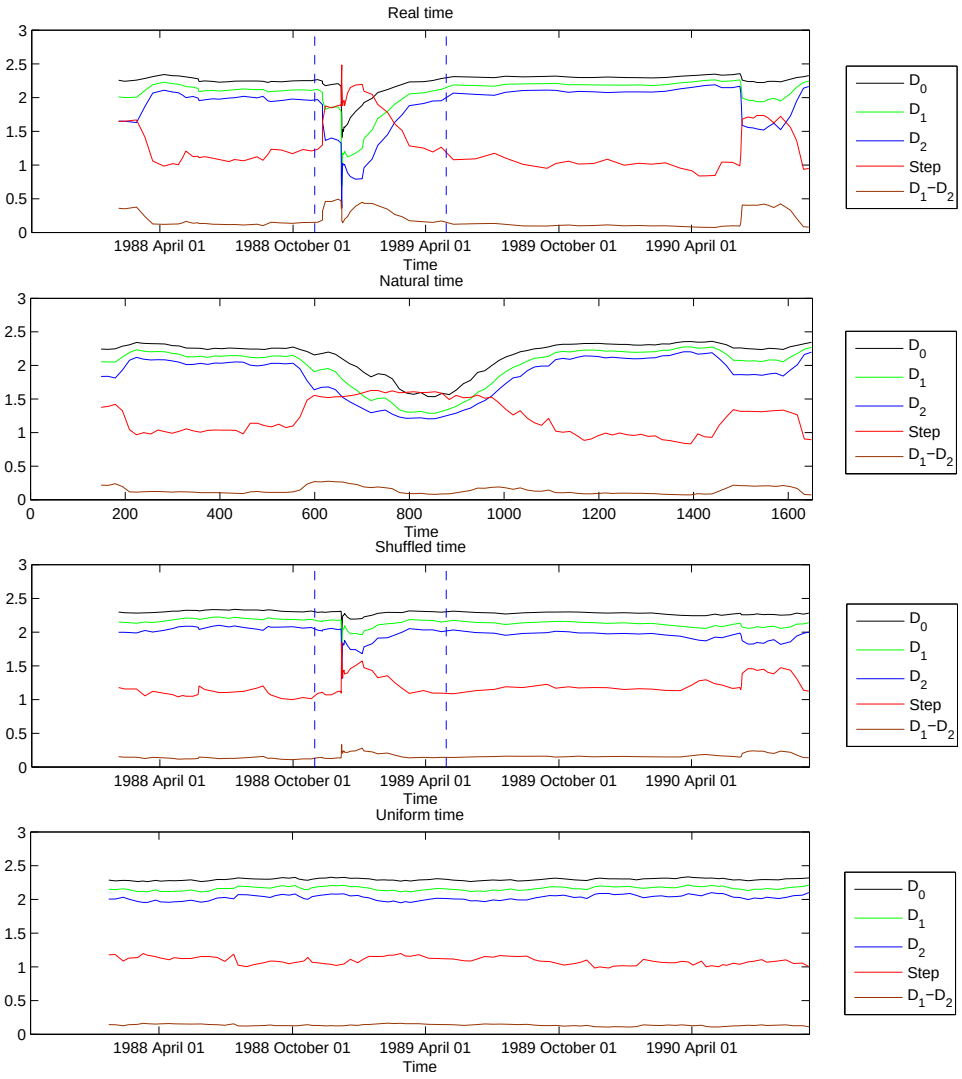

Figure 4.26: Agrón data. Values of generalized Rényi dimensions D_0, D_1, D_2, $D_1 - D_2$ and $D_{-\infty} - D_{\infty}$, corresponding to TR, TN, TS and TU, based on sliding windows (size 150 events).

4.2 El Hierro Data

The volcanic island of El Hierro is the youngest and smallest of the Canary Islands archipelago, in the Atlantic Ocean near the coast of Africa. Seismic activity is mainly related to its volcanic nature. For a longtime, the seismicity of this island has been characterized by a low intensity. Suddenly in July 2011, a large number of low magnitude earthquakes were located in the middle of the island. In the second half of September their locations moved towards the South, and on October 10, 2011, a submarine eruption occurred in the South of the island. The volcanic eruption was confirmed two days later by the appearance of a large lightgreen coloured area in the sea produced by magmatic gases, and five days later by floating volcanic fragments. Since then, the earthquakes magnitude increased and at the end of October the events migrated to the North. The largest earthquake of the series was recorded on November 8 and from this date the seismic activity began to decrease. It is important to point out the high quality of the catalogue due to the fact that at the beginning of the sequence the Spanish IGN (National Geographic Institute) decided to increase its monitoring network by deploying seven temporary seismic stations.

It must be noted that this data set is a well-known seismic sequence due to the impact on the environment and its inhabitants of the volcanic eruption occurred during the series. The pictures of the submarine volcanic eruption have had a great coverage in the media. Besides these aspects, this data set is very interesting due to the large number of events occurred, more than 11.000 in just 173 days, and also because of the submarine volcanic eruption involved. The seismic activity has continued on the island and, nowadays, its physical state is widely followed by experts all over the world. Then, in this data set the main aim is to assess particular structural changes regarding the volcanic nature of the series

4.2.1 Descriptive analysis

In this study, we use the IGN earthquake catalogue between 19 July 2011 and 7 January 2012 taken from the IGN web site, *http://www.01.ign.es/ign/layoutIn/sismoFormularioCatalogo.do*, considering a study area located within longitude $[-18.2, -17.8]$ and latitude $[27.5, 27.9]$. We have discarded from this catalogue all earthquakes with depth error greater than 8 km, with the data set finally being composed of 11.142 events.

As mentioned above, this seismic series can be divided into three phases corresponding to well separated spatio-temporal clusters; see Ibáñez et al. (2012). A first exploration of the spatio-temporal interaction can be obtained looking at Figures 4.27 and 4.28. In both figures we can observe how at the beginning of the series the events are located in the middle of the study area (Phase A); subsequently, they move towards the South (Phase B); and finally, the locations migrate to the North (Phase C). These migrations in latitude are specially clear in the second figure.

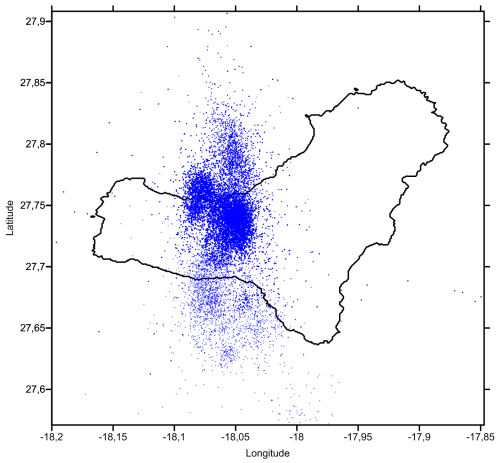

Figure 4.27: El Hierro data. Epicenter locations on contoured island coastal line.

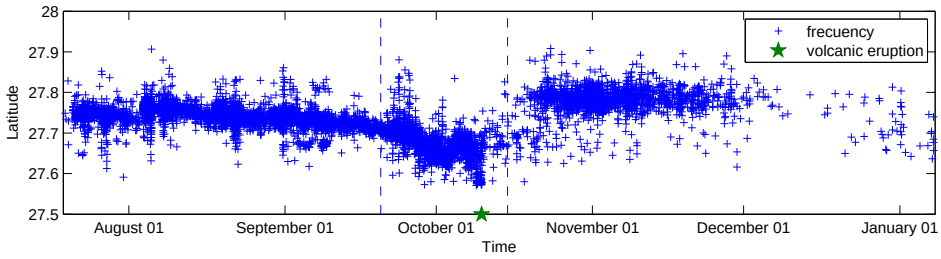

Figure 4.28: El Hierro data. Latitudes over time.

Daily frequencies, displayed in Figure 4.29, show typical variations inherent to multifractal behaviour. Here, we can see how in Phases A and B there is a greater seismic activity than in Phase C. Overall, the seismic activity started to decrease from the beginning of the volcanic eruption.

This first exploration of the spatio-temporal interaction is completed with the incorporation of the magnitude and depth, plotted in Figure 4.30. This figure shows the epicenter locations, with the colour representing the occurrence time, and the circle size representing either magnitude (left plot) or depth (right plot). The epicenter locations are very concentrated over time, which suggests a strong spatio-temporal interaction $(X,Y) \leftrightarrow T$. In the left plot, we observe that the larger earthquakes occurr in the South and North of the study area, corresponding to the middle and the end of the seismic series. In contrast, in the right plot it is difficult to see the degree of relation between depth and location, or between depth and time.

The depth distribution, Figure 4.31, shows clear trends over time. Depths are very clustered

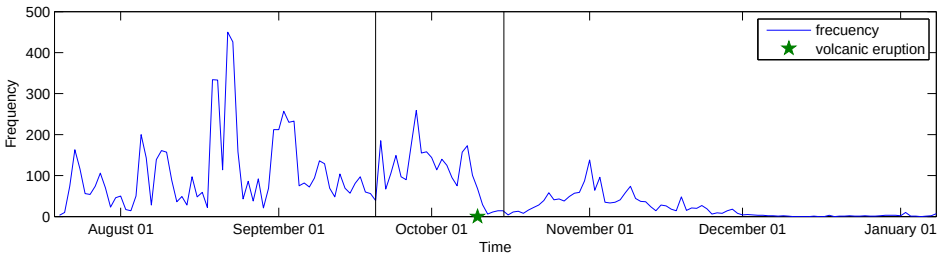

Figure 4.29: El Hierro data. Distribution of daily frequencies over time.

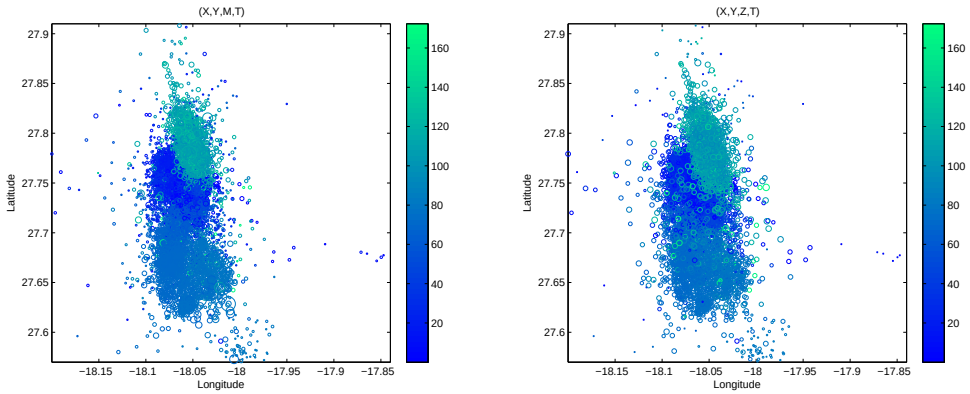

Figure 4.30: El Hierro data. Epicenter locations by time, represented by colour bar, and magnitude (left) or depth (right), represented by size of circle mark.

near 11 km during the whole Phase A. In Phase B, this clustering continues but the values increase slightly until the days previous to the volcanic eruption. During the days around the beginning of the volcanic eruption some earthquakes are located between 0 and 5 km, and after those days the depths keep increasing.

Finally, the magnitude distribution is displayed in Figures 4.32 and 4.33. In Figure 4.32, we can observe how the magnitudes of the events have a particular distribution in each phase. Thus, in Phase A, the magnitudes are low and very concentrated; Phase B, however, is composed of earthquakes with variable magnitude; and in Phase C, low magnitude events cannot be found. In Figure 4.33 the magnitude distribution in the series is shown by means of the frequency-magnitude histogram and the reciprocal empirical distribution, that is, the number of events with magnitude above a given level. Both plots markedly reflect the Gutenberg-Richter law,

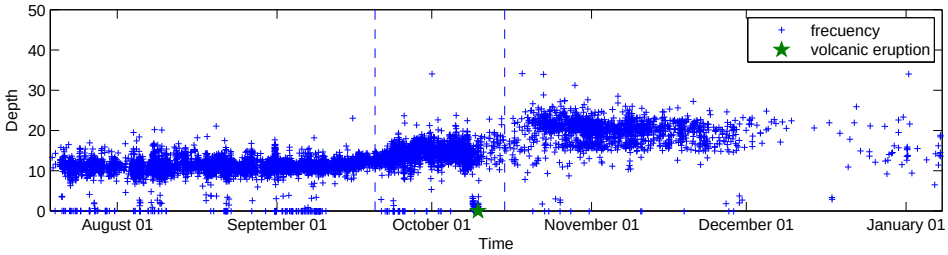

Figure 4.31: El Hierro data. Depth distribution over time.

which relates the frequencies of the events to their magnitude. Besides, the volcanic eruption plays an important role in the behaviour of the series.

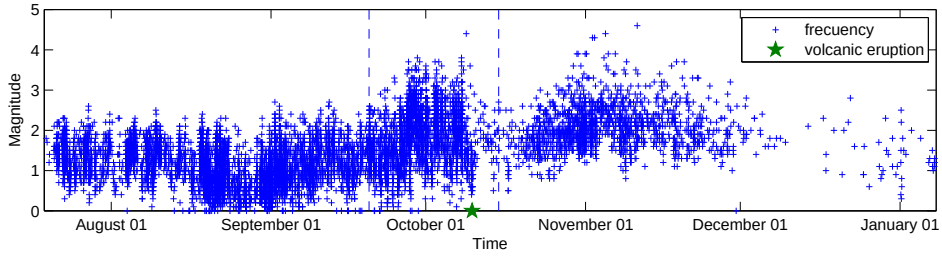

Figure 4.32: El Hierro data. Magnitude distribution over time.

4.2.2 Study of dimensional interaction

The descriptive study has shown that the temporal, depth and magnitude components have an important effect on the epicenter location. For a quantitative assessment, we study the dimensional interaction between these components and the epicenter location by means of entropy measures and the dependence coefficient. Firstly, we calculate the Shannon entropy values $H(D)$, $H(X,Y)$, $H(X,Y,D)$, $H(X,Y)+H(D)$ and $H(X,Y)+H(D)-H(X,Y,D)$, using varying box sizes between 1/15 and 1/2, and considering either time T, depth Z or magnitude M as the D component. Results are displayed in Figure 4.52. The deviations (brown line) show that the three components have significant effects on the epicenter location, with the effects of the temporal component being greater and those of the magnitude being smaller.

For a better assessment of the dimensional interaction, we perform a study of the temporal evolution of the T, D and M effects on the epicenter location. With this purpose, we calculate the values of the Shannon entropy taking a fixed box size $\varepsilon = 1/10$, for data segments of 150

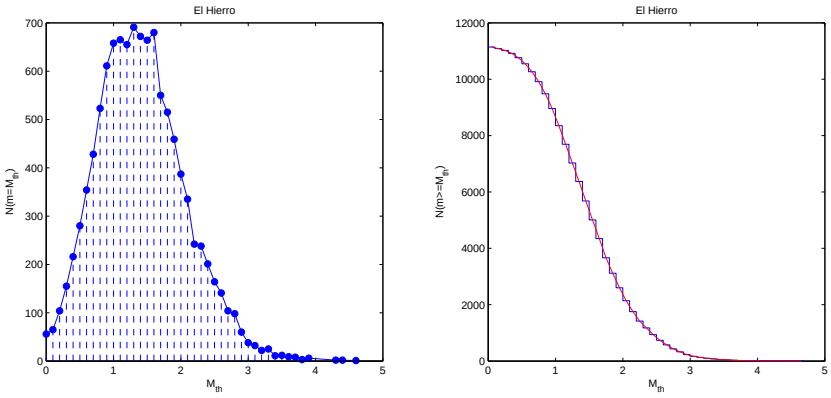

Figure 4.33: El Hierro data. Frequency-magnitude histogram (left) and reciprocal accumulated distribution (right).

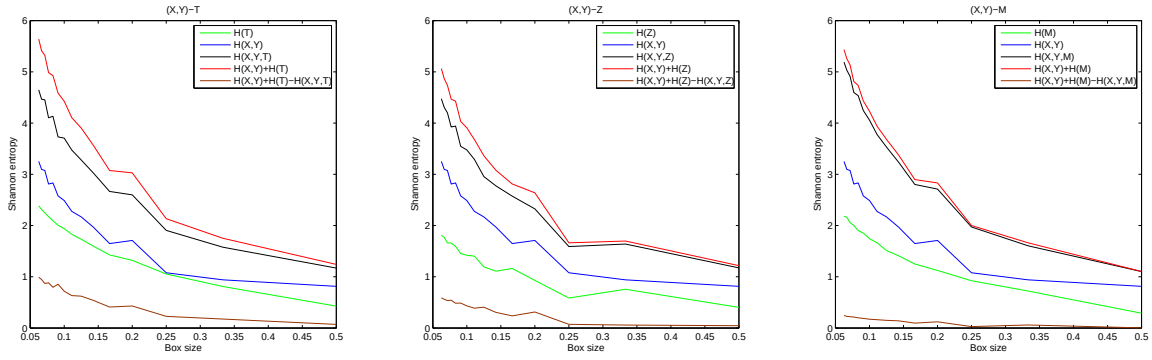

Figure 4.34: El Hierro data. Shannon entropy values $H(D)$, $H(X,Y)$, $H(X,Y,D)$ and $H(X,Y)+H(D)$, for varying box sizes between 1/15 and 1/2, with D corresponding to time, depth and magnitude.

events with an overlapping of 90 per cent; see Figure 4.35. These plots show that the degree of dependence between the dimension D (time, depth or magnitude) and the epicenter location increases during the days previous to the volcanic eruption, particulary in the case of time and depth. On those days, the degree of spatial disorder or uncertainty increases.

Finally, we study the dimensional interaction by calculating the dependence coefficient based on the 'entropy' dimension for the interactions $(X,Y) \leftrightarrow T$, $(X,Y) \leftrightarrow Z$ and $(X,Y) \leftrightarrow M$. The results, displayed in Figure 4.36, show that the degree of dependence between the epi-

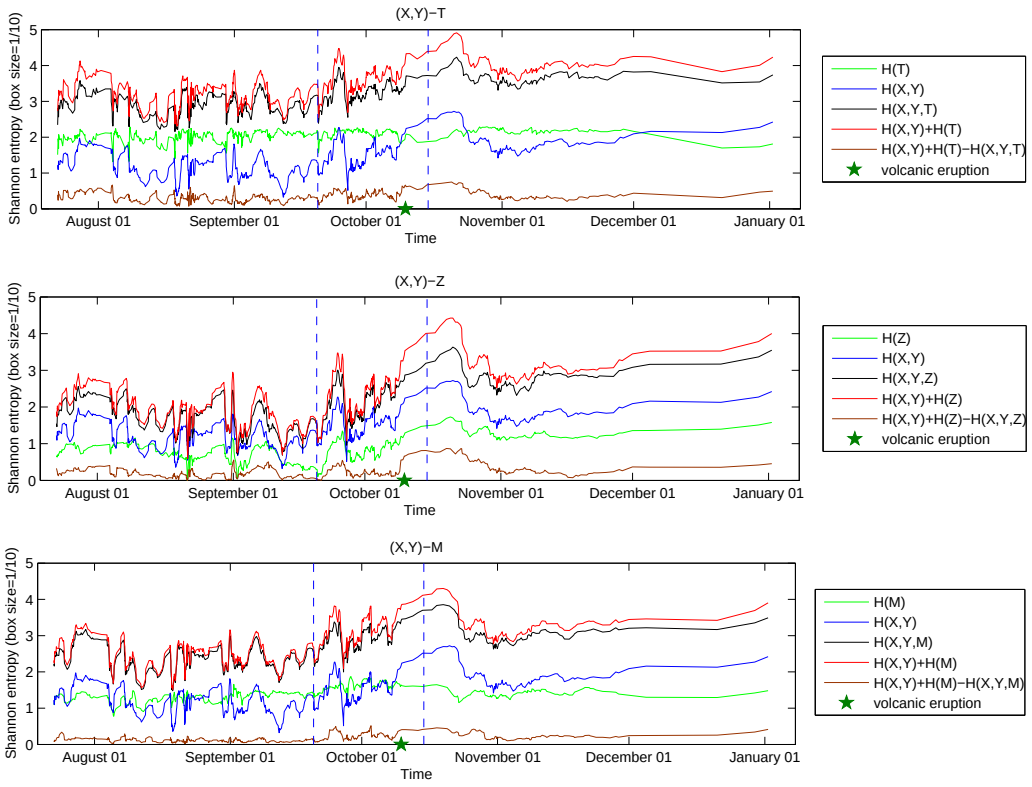

Figure 4.35: El Hierro data. Shannon entropy values $H(D)$, $H(X,Y)$, $H(X,Y,D)$, $H(X,Y)+H(D)$ and $H(X,Y)+H(D)-H(X,Y,D)$ based on sliding windows (size 150 events), considering a fixed box size $\varepsilon = 1/10$, with D corresponding to time, depth and magnitude.

center location and the magnitude is lower than in the other cases. Moreover, in all cases the dependence markedly increases during the volcanic eruption.

These results mainly indicate two aspects: firstly, that occurrence time and depth have a higher degree of association with the epicenter location; and secondly, that this association increases strongly since the volcanic eruption occurred.

4.2.3 Study of structural characteristics at different scales

In previous analyses, we have concluded that the temporal and depth components have significant effects on the epicenter location and so does the magnitude but to a smaller extent. Thus, to perform a complete study of the structural characteristics of the seismic series we analyze the

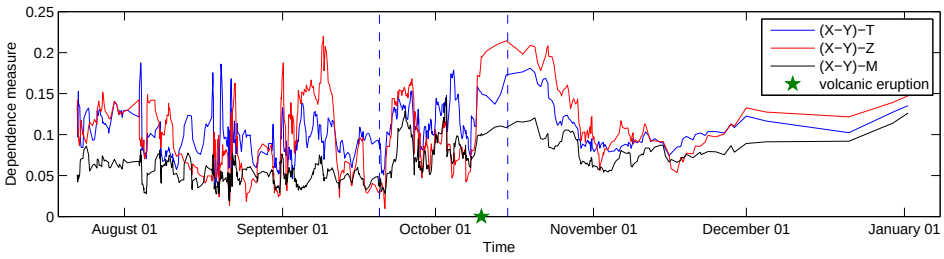

Figure 4.36: El Hierro data. Values of the 'dependence coefficient' for components $(X,Y) \leftrightarrow T$, $(X,Y) \leftrightarrow Z$ and $(X,Y) \leftrightarrow M$, based on sliding windows (size 150 events).

(X,Y,T) and (X,Y,Z) patterns over time, with and without considering the effects of the magnitude, in association with the scaling properties inherent to the phenomenon. The procedure used consists of evaluating the temporal changes in the generalized dimensions, D_0, D_1, and D_2, multifractal step $D_{-\infty} - D_{\infty}$, and complexity measure $D_1 - D_2$, besides the LMC_{exp} calculated for a box size $\varepsilon = 1/10$. The analysis is performed based on sliding windows of 150 events with 90 per cent of overlapping, considering two weighting functions in the box-counting: $\phi(m) = 1$ and $\phi(m) = \exp(3m/2)$.

The temporal variations in the values of the former coefficients are displayed in Figures 4.37 to 4.42. Here, we can see that the curves with and without considering the effects of the magnitude show similar patterns in all cases, although we can detect strong deviations during the volcanic eruption and the periods with larger earthquakes. This suggests, as in the Agrón data set, that the distribution of magnitude in general has an intrinsic structural association with (X,Y,T) and (X,Y,Z). In these plots we can also observe that the significant departures between both curves in all cases start the days previous to the volcanic eruption, which might be interpreted as a precursory indicator. In conclusion, the results obtained in this series suggest the advisability of incorporating the effects of the magnitude in this study. Besides, we must note that, as in the Agrón series analysis, the results of both complexity measures $D_1 - D_2$ and LMC_{exp} show similar patterns.

4.2.4 Frequency-magnitude distribution analysis

In this subsection, we analyze the seismic series by means of the non-extensive frequency-magnitude distribution (FMD) exposed in Chapter **??**. We perform the fitting by means of the weighted non-linear regression introduced in Chapter 3 considering the following weighting functions: $\phi(m) = 1$ and $\phi(m) = \exp(3m/2)$. We refer to these fittings as non-weighted and weighted models.

Firstly, we consider the three phases corresponding to the different spatio-temporal clusters

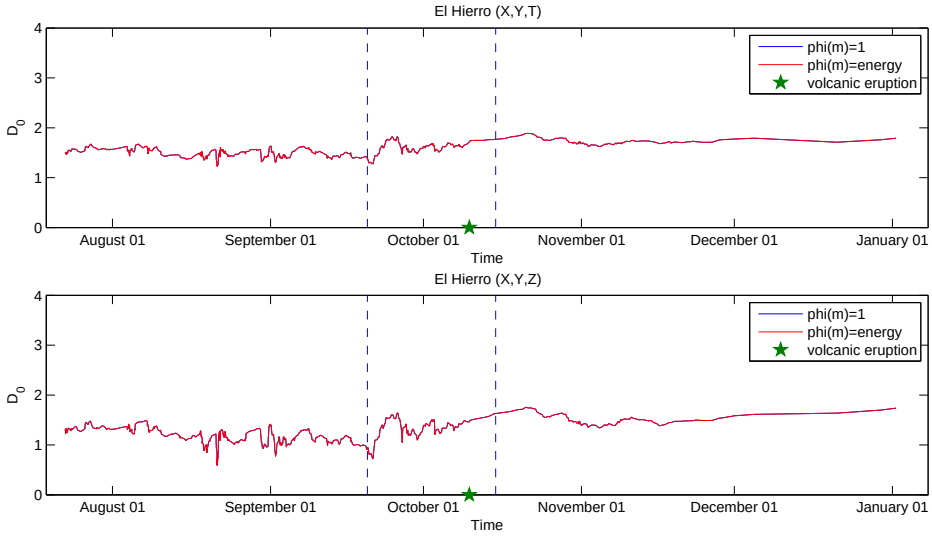

Figure 4.37: El Hierro data. Variations in the 'capacity' dimension values over time based on sliding windows (size 150 events), with and without considering the effects of the magnitude.

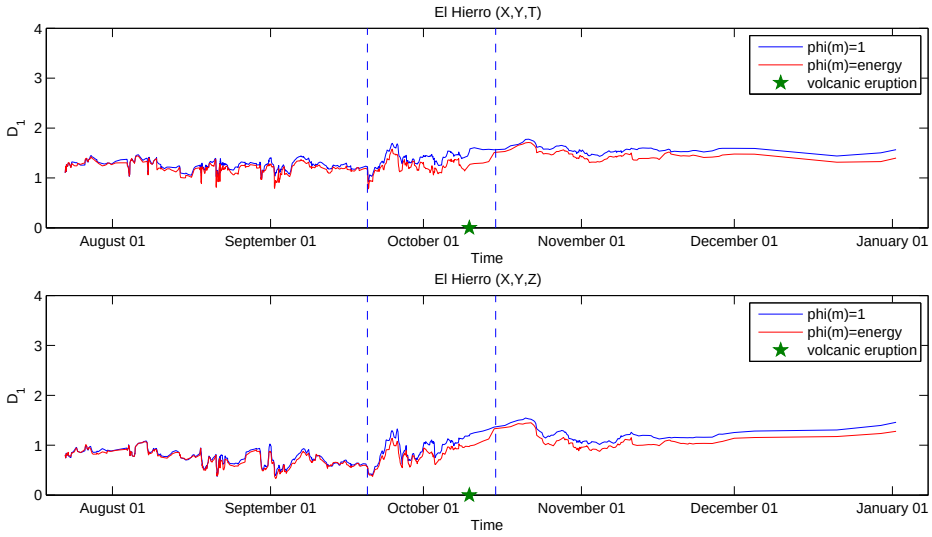

Figure 4.38: El Hierro data. Variations in the 'entropy' dimension values over time based on sliding windows (size 150 events), with and without considering the effects of the magnitude.

observed in the data; subsequently, we carry out a study over time for the whole data set. Figure 4.43 shows the representations of the fitted curves and Table 4.3 collects the values of

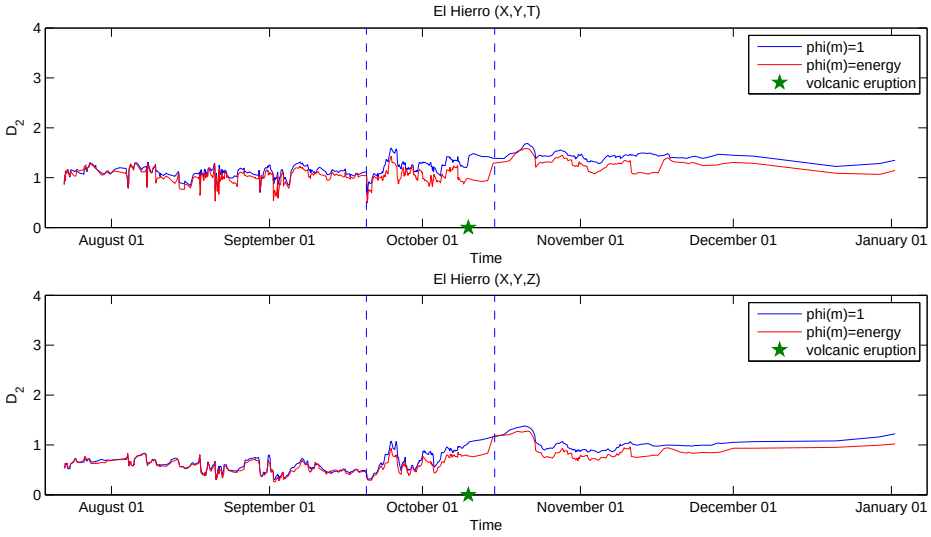

Figure 4.39: El Hierro data. Variations in the 'correlation' dimension values over time based on sliding windows (size 150 events), with and without considering the effects of the magnitude.

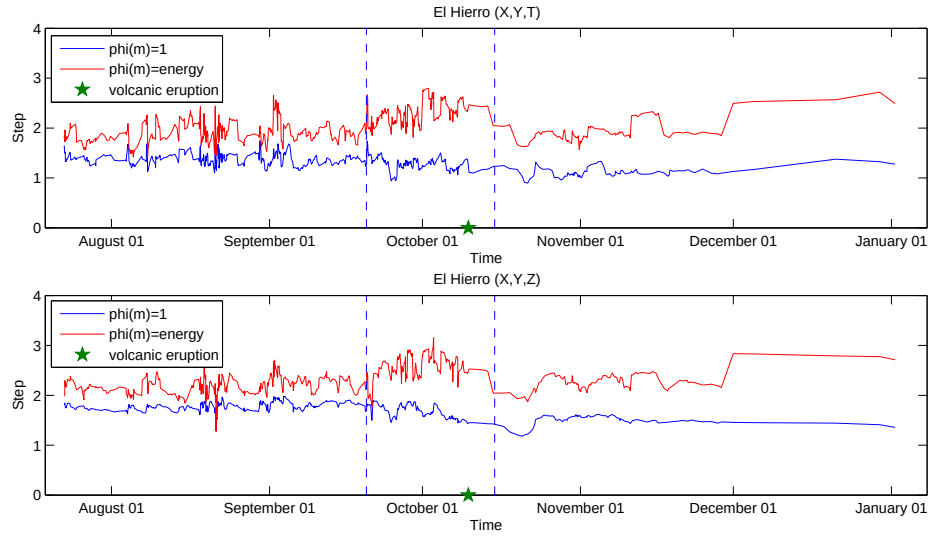

Figure 4.40: El Hierro data. Variations in the step values over time based on sliding windows (size 150 events), with and without considering the effects of the magnitude.

the parameter estimates, misfit and R^2 for the two models, considering the whole series and the three phases. We must note that the misfit is calculated here as the average of the absolute

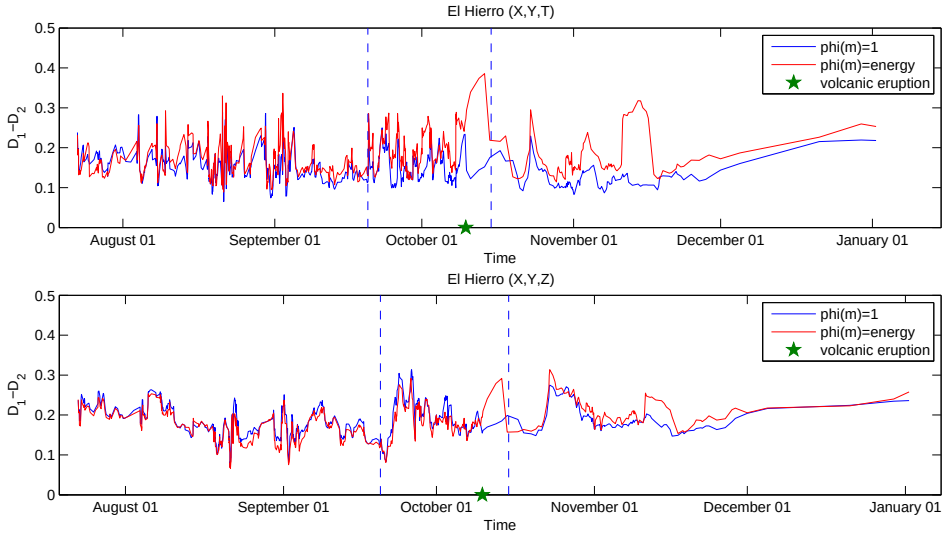

Figure 4.41: El Hierro data. Variations in the complexity measure $D_1 - D_2$ values over time based on sliding windows (size 150 events), with and without considering the effects of the magnitude.

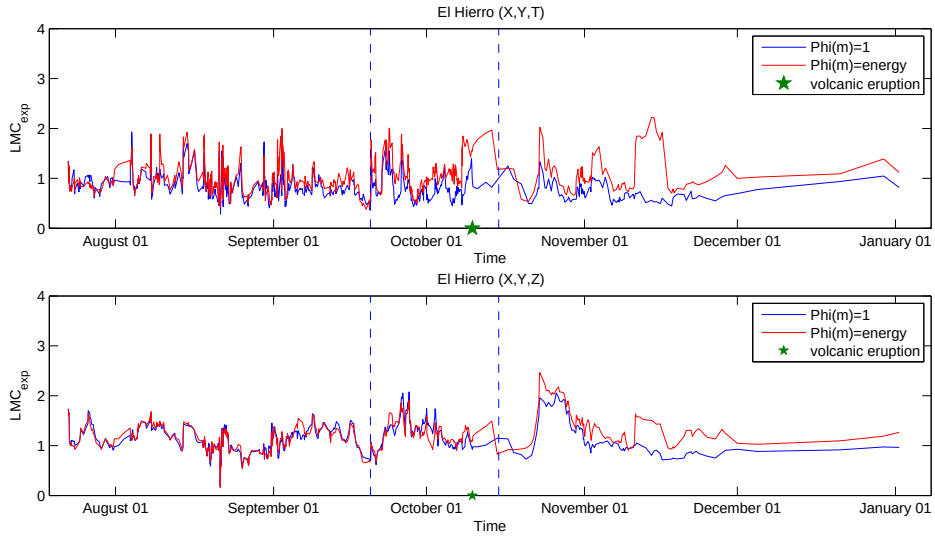

Figure 4.42: El Hierro data. Variations in the values of the LMC_{exp} over time based on sliding windows (size 150 events), with and without considering the effects of the magnitude.

values of residuals. In both cases, the misfit is lower in Phase A, which can be due to the fact that in this phase events of high magnitude do not occur. In contrast, Phase B has only one large event and it might be the reason wherefore the misfit is higher. It is important to highlight that the weighted model clearly improves the fitting in all cases, especially when it is worse (Phase B).

The estimated value of the non-extensivity parameter q is higher in Phase B and Phase C than in Phase A for both models, which indicates that the physical state in these two phases is farther from equilibrium than in Phase A. In the context of seismicity, this implies that larger earthquakes can be expected. In addition, the higher values of the parameter a in Phase B and Phase C show that during these periods a large amount of energy in proportion to the size of the fragments is released.

Table 4.3: El Hierro data. Parameters, misfit and R^2 of the fitted curves for the whole sequence and the three phases based on the non-weighted and weighted models.

	q	a	misfit	R^2
Whole sequence				
$\phi(m) = 1$	1.4260	143.4045	0.0542	0.9972
$\phi(m) =$energy	1.4152	185.8459	0.0084	0.9979
Phase A				
$\phi(m) = 1$	1.1995	63.1251	0.0221	0.9993
$\phi(m) =$energy	1.1960	65.0538	0.0047	0.9995
Phase B				
$\phi(m) = 1$	1.3669	644.7999	0.0734	0.9887
$\phi(m) =$energy	1.3840	468.3539	0.0173	0.9913
Phase C				
$\phi(m) = 1$	1.4436	570.3165	0.0540	0.9955
$\phi(m) =$energy	1.4472	533.3538	0.0075	0.9978

To complete the study, we analyze the temporal evolution of the series through the variations in the parameters (q and a) and the goodness of fit (misfit and R^2) over time. The estimates are calculated using both models applied on sliding windows.

The results show that the variations in the estimated values of parameter q are irregular and smaller in Phase A, whereas in Phase B and Phase 0C the values are far from equilibrium ($q = 1$); see the bottom plot of Figure 4.44. Besides, we can observe that the higher values of q occur during the volcanic eruption and the larger earthquakes, which indicates that in these

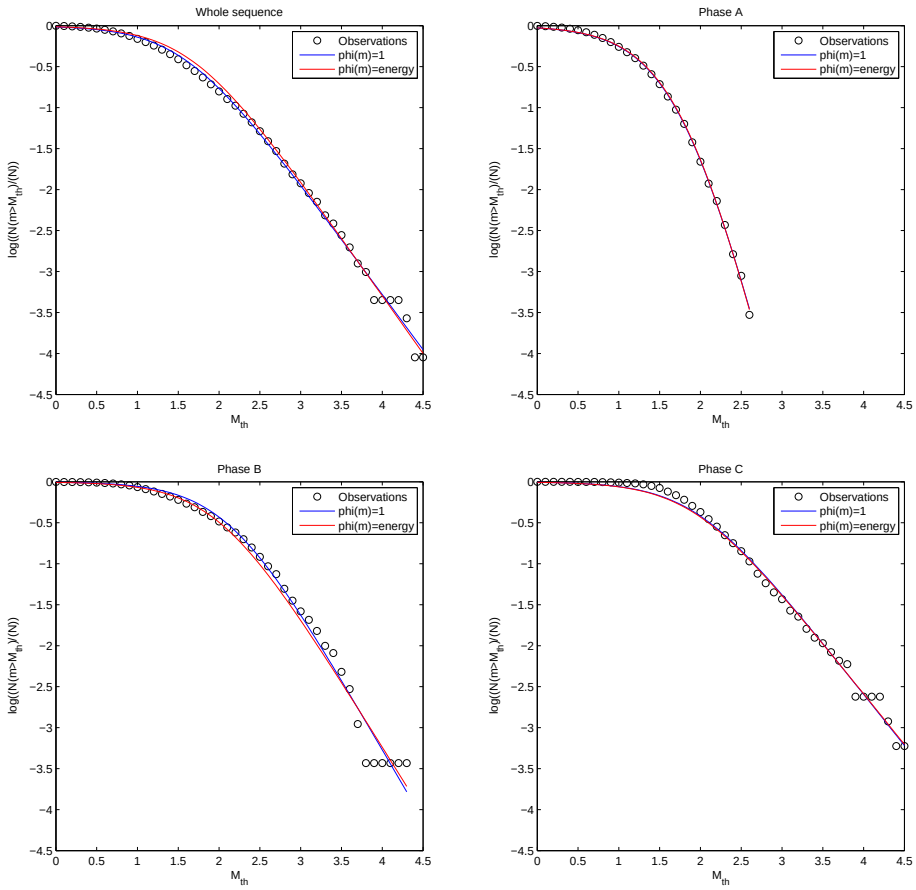

Figure 4.43: El Hierro data. Representations of the fitted curves for the whole sequence and three phases.

periods the physical state is farther from equilibrium. Regarding the differences in the results obtained for both estimation models we can observe how the estimated values for q are slightly higher during the volcanic eruption and the days with larger earthquakes for the weighted model.

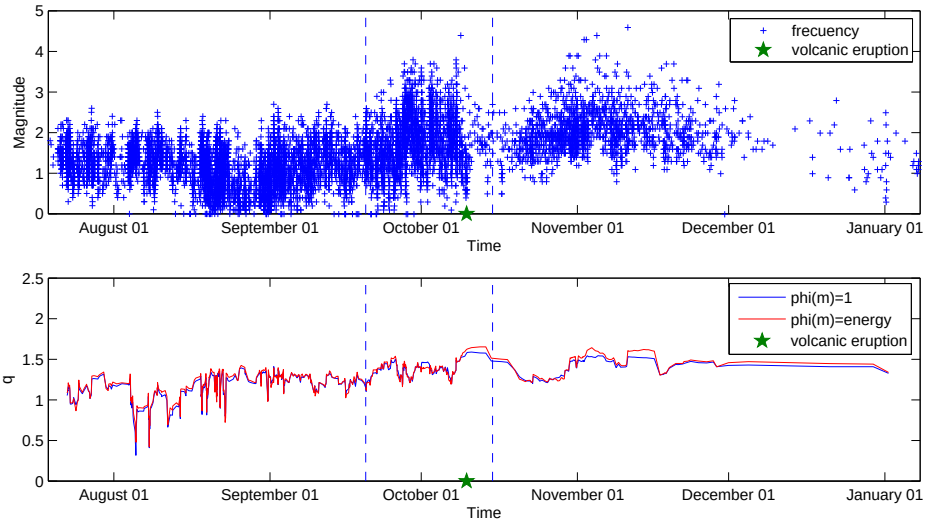

Figure 4.44: El Hierro data. Magnitude distribution and estimated values of parameter q over time based on sliding windows (size 150 events), for the non-weighted (top) and weighted models (bottom).

As for the estimated values of the parameter a over time, displayed for both models in the bottom plot of Figure 4.45, we can observe how according with the variations of a the series can be divided in the same three phases that when we consider the spatio-temporal clustering. In Phase A, the parameter a takes small values showing that the energy released in proportion to the fragment size is not significant. To interpret this, we must note that the magnitudes (see the top plot of the former figure) are small during this period, but it can also be due to the fact that the fragment sizes are large. Nevertheless, in Phase B, the estimated values of parameter a show a great amount of energy in proportion to the fragments size during the period previous to the volcanic eruption. Despite the fact that during these days large earthquakes occurred, the great increase in the estimated values makes us think that the fragment size is not large. This can be seen as a precursory behaviour to the volcanic eruption. Since its occurrence, the amount of energy released in proportion to the fragment size is very small. Finally, in Phase C, the estimated values of a are smaller than in Phase B although in this period the largest earthquakes occurred, which could indicate that the fragment sizes are larger in Phase C than in Phase B.

To finish, we study the goodness of fit for both models. The plots of Figure 4.46 display the variations in the misfit and R^2 over time (from top to bottom). These plots clearly show that

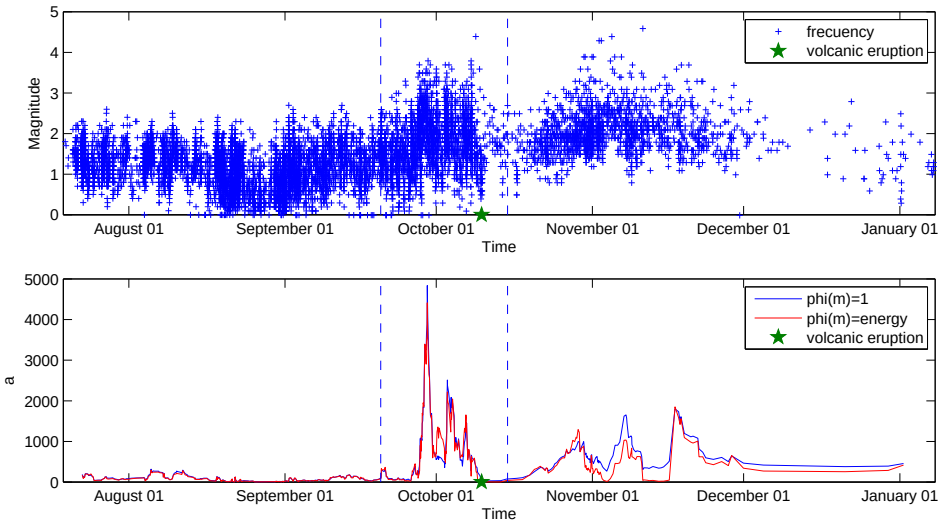

Figure 4.45: El Hierro data. Magnitude distribution and estimated values of parameter *a* over time based on sliding windows (size 150 events), for the non-weighted (top) and weighted models (bottom).

when we use the weighted non-linear model the goodness of fit improves, especially when the misfit is worse.

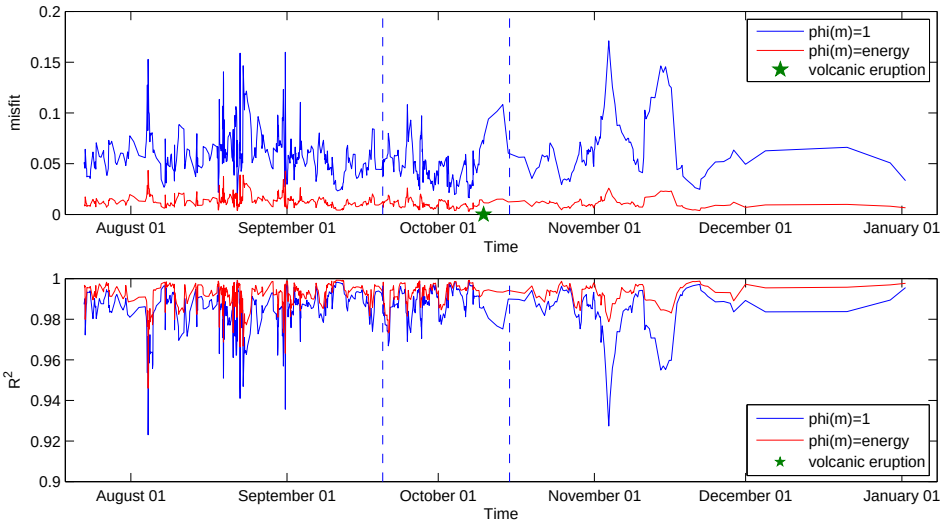

Figure 4.46: El Hierro data. Misfit and R^2 values over time based on sliding windows (size 150 events), for the non-weighted (top) and weighted models (bottom).

4.3 Torreperrogil Data

The last data set we analyze corresponds to the seismic series occurred in the area near the village of Torreperogil (Jaén, Spain) from 1 March 2009 to 1 March 2.013, with 2132 events registered. The classical seismic approach, based on dynamics of the tectonic plates, does not explain well this unexpected seismic phenomenon and, as a result, different theories have emerged to justify its presence. Among such theories the most relevant ones are 'hydroseismicity' and 'fracking'. In the context of 'hydroseismicity', Doblas et al. (2013) defend that the causes which have produced this phenomenon are the pronounced drought of the year 2012 together with numerous extractions of water and the heavy rainfall occurred from September 2012. The other theory assures that the causes of the sudden change in the seismic activity are not natural but caused by 'fracking' in this area. The Andalusian Government granted permission on 10 April 2012 to the company Oil&Gas to carry out a research on hydrocarbons in La Loma district. This company states that they have just carried out some research on prospecting but not drilling works and this method is incapable of producing earthquakes. On the other hand, the citizen platform against fracking has taken a complaint to Court to investigate the possible works of hydraulic fracturing of the rocks to extract gas. This theory has precedents in other regions of the world. A very relevant case occurred in Groningen district, Holland, which is the major deposit of natural gas in Europe. The gas extracted is consumed by 97 per cent of the population and produces enormous incomes for the state. In this case, the government has assumed publically the seismic risk in exchange for extracting natural gas. Thus, the main aim in the analysis of this data set is to try to detect whether there is a presence of external factors in the phenomenon.

We perform the study of this seismic series following the same approaches used for the analysis of EL Hierro data set: firstly, we carry out a descriptive analysis of the data; the dimensional interactions are then analyzed; the next step is to assess the structural characteristics at different scales; finally, we complete the study with the analysis of the series by means of the non-extensive frequency-magnitude distribution.

4.3.1 Descriptive analysis

As in the case of El Hierro, the data set has been taken from the Spanish IGN (National Geographic Institute). The study area is located in the Northeast of Andalusia within longitude $[-3.6, -3]$ and latitude $[37.7, 38.2]$. The swarm of epicenters occurred between 1 March 2009 and 1 March 2013, in relation to the region border line, is displayed in Figure 4.47. This area has always been characterized by a low seismic activity until December 2012, when an extreme increase in the seismic activity was observed (daily frequencies are shown in Figure 4.48). The sequence took place during 48 months, but approximately 89 per cent of the total number of events occurred during the last 5 months.

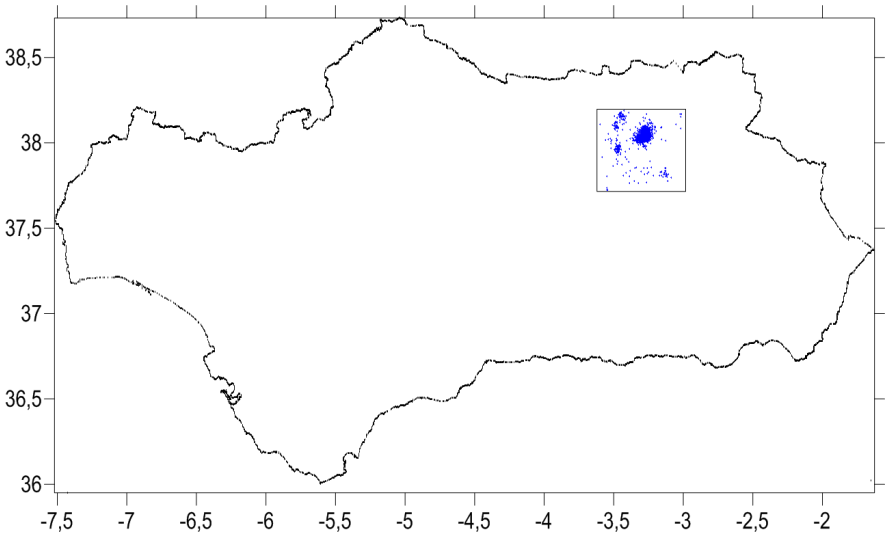

Figure 4.47: Torreperogil data. Epicenter locations on contoured Andalusian region.

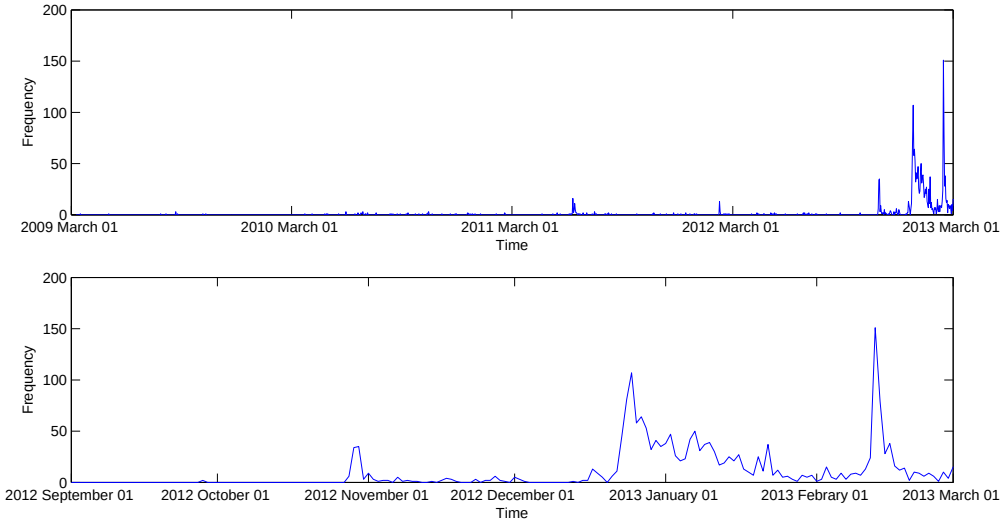

Figure 4.48: Torreperogil data. Daily frequencies for the whole period (top) and for a restricted interval corresponding to the high activity period (bottom).

As was done in the other data series, we perform a first exploration of the spatio-temporal interaction jointly with the incorporation of the magnitude and depth; see Figure 4.49. In this figure we can see the swarm of epicenters, with the colour representing the occurrence time and

the circle size representing either magnitude (left plot) or depth (right plot). Both plots show that most events occurred at the end of the series, and they are located around only one spatial cluster, which indicates a strong spatio-temporal interaction. In the left plot, it is difficult to detect whether there is an association between epicenter location and magnitude. In the right plot, however, we can observe significant associations between depth and epicenter location, and between depth and time.

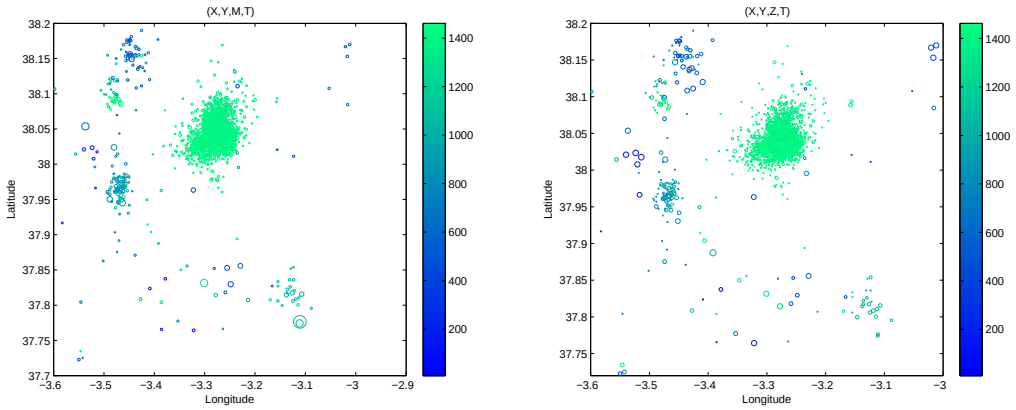

Figure 4.49: Torreperogil data. Epicenter locations by time, represented by colour bar, and magnitude (left) or depth (right), represented by size of circle mark.

For a better assessment of the relation between time and magnitude, and between time and depth, we plot their distributions in Figures 4.50 and 4.51, respectively. In both figures the top plots display the whole series whereas the bottom plots allow an easier visualization restricted to the last five months. Regarding the magnitude distribution, we can observe how during the higher seismic frequency periods the magnitude values are more disperse. As for the depth distribution, the plots confirm that the depths of the earthquakes seem scattered until October 2012, when the observed depths are very concentrated between 0 and 10 km.

4.3.2 Study of dimensional interaction

As was done before for Agrón and El Hierro, we now use the tools based on entropy and the dependence coefficient to study the degree of dimensional interaction of $(X,Y) \leftrightarrow T$, $(X,Y) \leftrightarrow Z$ and $(X,Y) \leftrightarrow M$. Figure 4.52 displays the Shannon entropy values $H(D)$, $H(X,Y)$, $H(X,Y,D)$, $H(X,Y)+H(D)$ and $H(X,Y)+H(D)-H(X,Y,D)$ calculated using varying box sizes between $1/15$ and $1/2$, considering as D component T, Z or M. In these plots we can observe that the Shannon entropy values for $(X,Y) \leftrightarrow M$ (bottom plot) are significantly higher than for the

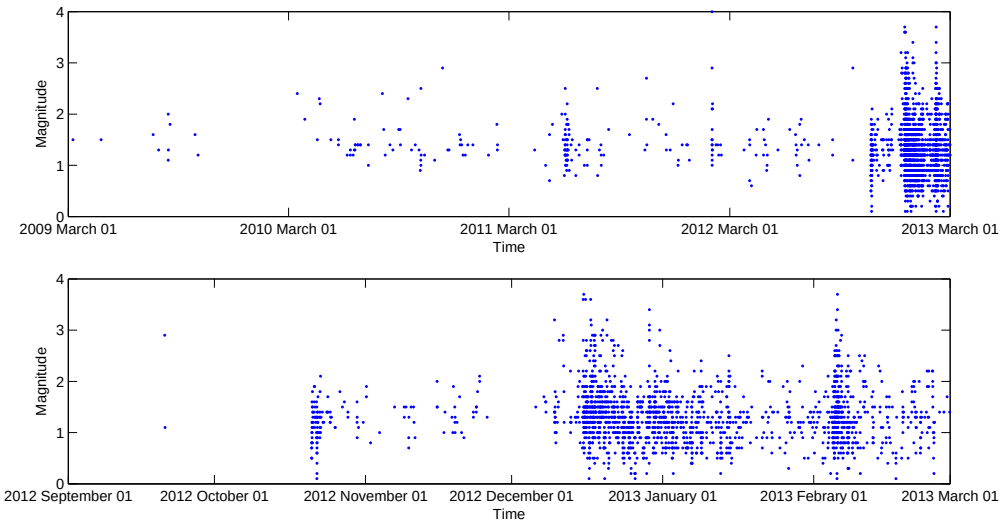

Figure 4.50: Torreperogil data. Magnitude distribution over time for the whole period (top) and for a restricted interval corresponding to the high activity period (bottom).

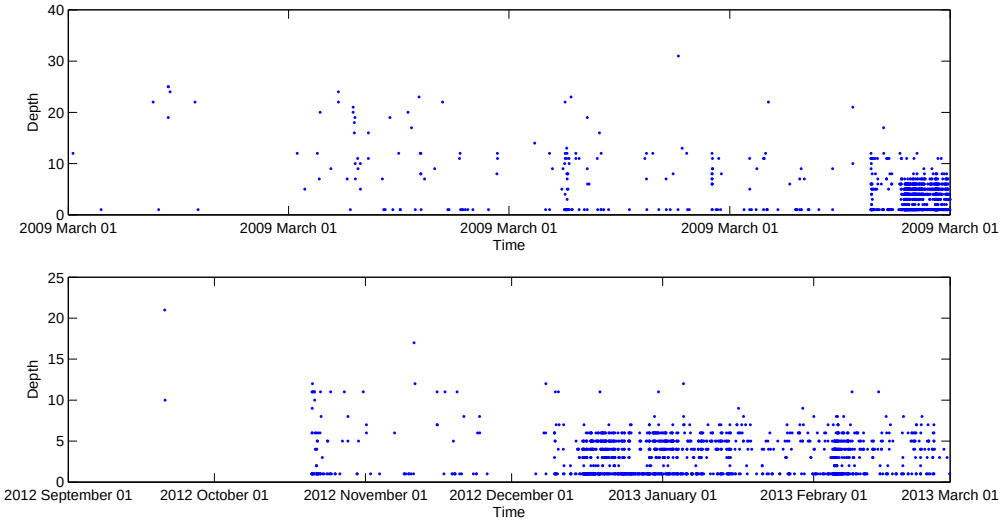

Figure 4.51: Torreperogil data. Depth distribution over time for the whole period (top) and for a restricted interval corresponding to the high activity period (bottom).

other interactions (left and right top plots), which reflects that this interaction contains high uncertainty or disorder. The significant departure of $H(X,Y,T)$ from $H(X,Y)+H(T)$ indicates

that the dimensional interaction between the earthquake epicenter and the occurrence time is very strong; see left top plot. The right top plot of the same figure shows the presence of a certain degree of association in the interaction $(X,Y) \leftrightarrow Z$. The barely departure between $H(X,Y,Z)$ and $H(X,Y)+H(Z)$ in the bottom plot of the figure shows a weak dimensional interaction between (X,Y) and Z.

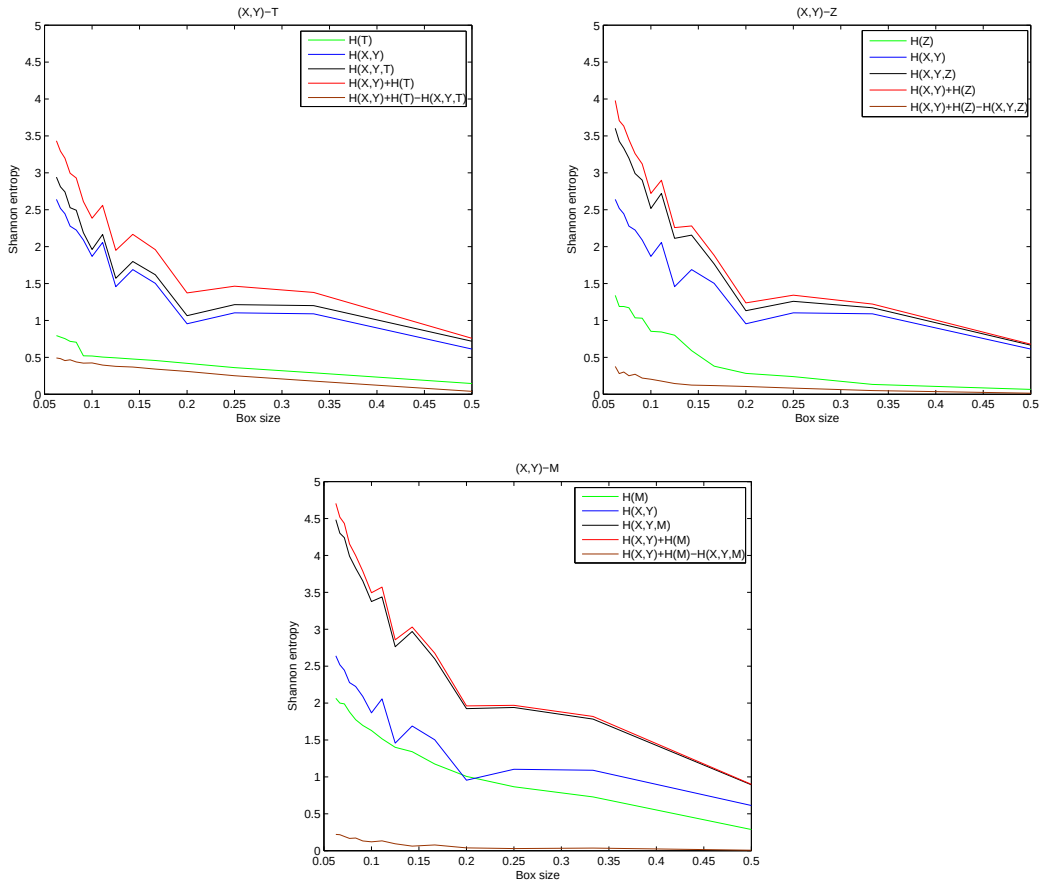

Figure 4.52: Torreperogil data. Shannon entropy values $H(D)$, $H(X,Y)$, $H(X,Y,D)$, $H(X,Y)+H(D)$ and $H(X,Y)+H(D)-H(X,Y,D)$, for varying box sizes between 1/15 and 1/2, with D corresponding to time, depth and magnitude

Now, we study the dimensional interaction between the components by means of the Rényi entropy. We calculate the Rényi entropy values $H_q(D)$, $H_q(X,Y)$, $H_q(X,Y,D)$, and $H_q(X,Y)+H(D)$ taking the component D as T, M or Z. With this purpose, we select a specific box size $\varepsilon = 1/10$ with q varying from -8 to 8. The plots of Figure 4.53 reflect that the dimensional

interaction $(X,Y) \leftrightarrow T$ is greater than in the other cases, whereas for $(X,Y) \leftrightarrow M$ it is very low, especially considering non-negative values of q.

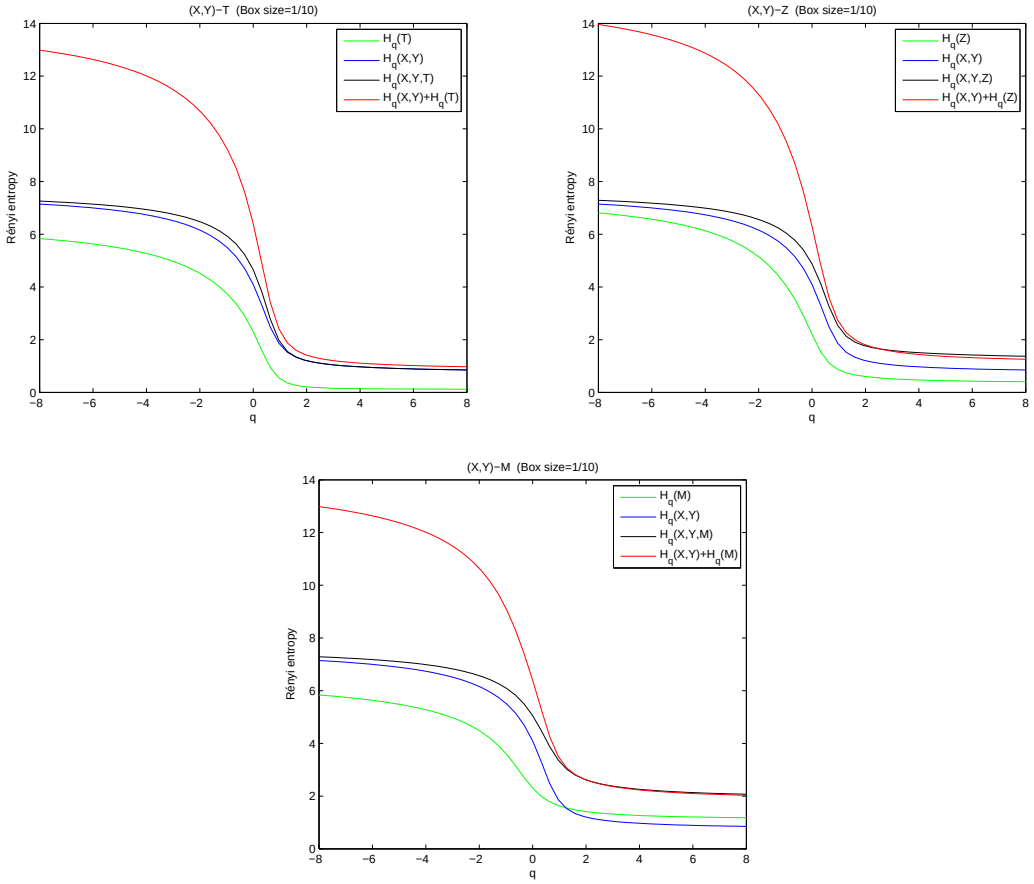

Figure 4.53: Torreperogil data. Rényi entropy values $H_q(D)$, $H_q(X,Y)$, $H_q(X,Y,D)$, $H_q(X,D)+H_q(D)$ and $H_q(X,Y)+H_q(D)-H_q(X,Y,D)$ for a specific box size $\varepsilon = 1/10$, with D corresponding to time, depth and magnitude.

As for Tsallis entropy, we calculate $T_q(D)$, $T_q(X,Y)$, $T_q(X,Y,D)$ and $T_q(X,Y)+T_q(D)-(q-1)T_q(X,Y)T_q(D)$, considering as D the T, M or Z components. As was done above with the Rényi entropy, we select a specific box size $\varepsilon = 1/10$, with the q parameter now varying from 1 to 3. The deviation of $T_q(X,Y,T)$ from $T_q(X,Y)+T_q(T)-(q-1)T_q(X,Y)T_q(T)$ indicates that the degree of dependence of the spatio-temporal interaction is clearly larger than for the other interactions; see Figure 4.54.

For a better assessment of the dimensional interaction, the study is complemented with the

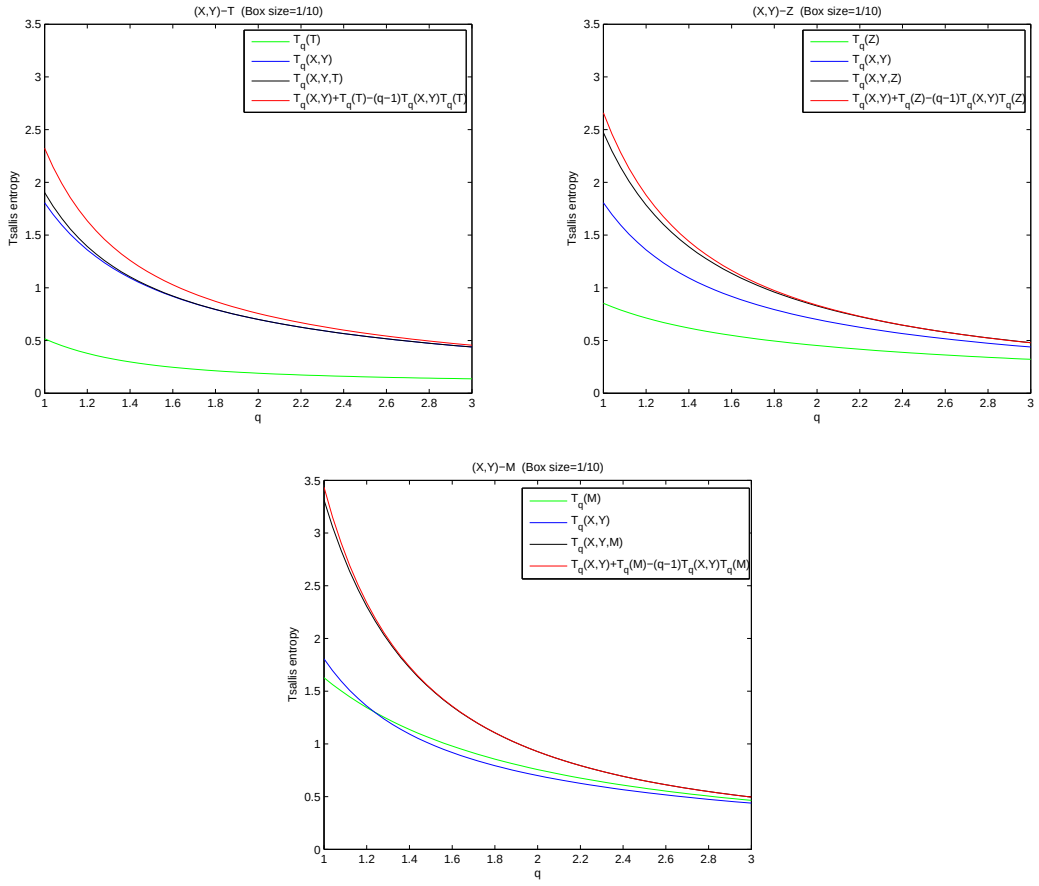

Figure 4.54: Torreperogil data. Tsallis entropies $HT_q(D)$, $HT_q(X,Y)$, $HT_q(X,Y,D)$ and $HT_q(X,Y)+HT_q(D)-(q-1)HT_q(X,Y)HT_q(D)$ for a specific box size $\varepsilon = 1/10$, with D corresponding to time, depth and magnitude.

values of the Shannon entropy calculated on sliding windows of 150 events considering an overlapping of 90 per cent for the thee interactions. The analysis is performed by taking a box size of $1/10$ and the results are displayed in Figure 4.55. Besides, Figure 4.56 allows easier visualization restricted to the high activity period. In the three plots we can observe how the degree of disorder markedly decreases for (X,Y) during the high activity periods. As for the dimensional interaction the association decreases in these periods, particulary for $(X-Y) \leftrightarrow T$.

Finally, we study the dimensional interaction over time by means of the dependence coefficient for the cases based on three interactions by sliding windows; see in Figure 4.57. The plots show, again, that the levels of dependence between the epicenter location and the time, depth

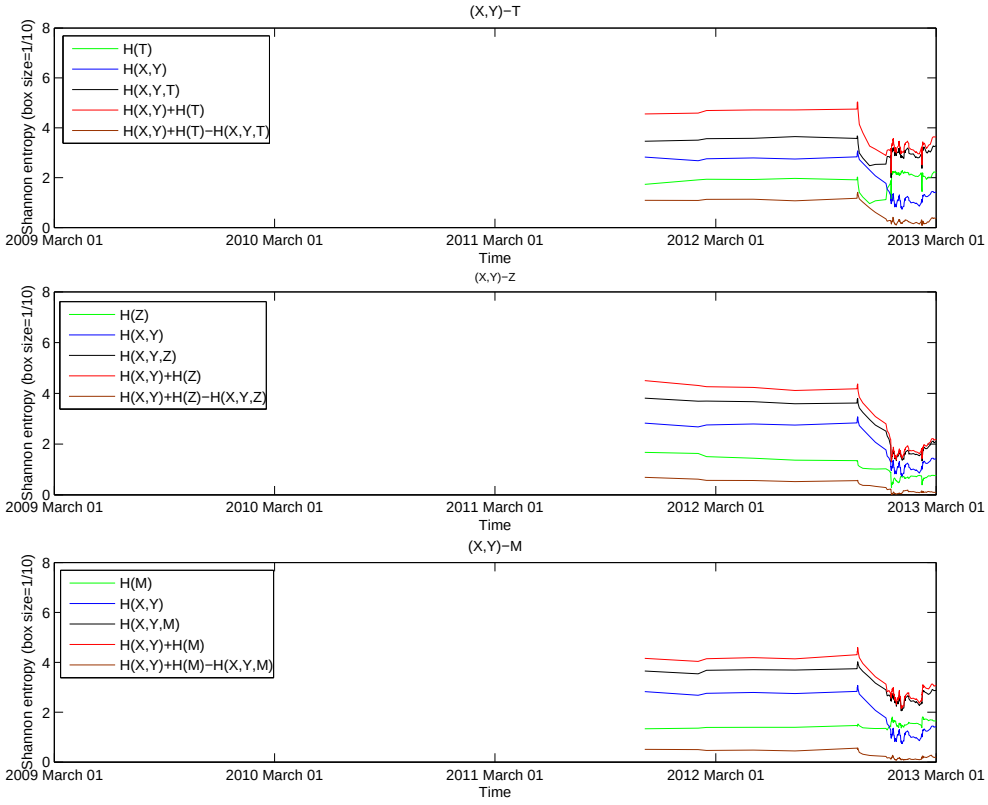

Figure 4.55: Torreperogil data. Visualization for the whole period of the Shannon entropy values $H(D)$, $H(X,Y)$, $H(X,Y,D)$, $H(X,Y)+H(D)$ and $H(X,Y)+H(D)-H(X,Y,D)$, based on sliding windows (size 150 events), considering a fixed box size $\varepsilon = 1/10$ with D corresponding to time, depth and magnitude.

and magnitude components are very low during the high activity period.

4.3.3 Study of structural characteristics at different scales

The above results show some evidence of dynamical structural changes. Now, we study the structural characteristics in association with the scaling properties inherent to the phenomenon. Firstly, we carry out a multifractal analysis for (X,Y,T), (X,Y,Z) and (X,Y,M) by means of the Rényi and Tsallis generalized dimensions. Secondly, we perform the analysis of the structural characteristics at different scales over time, looking at the variations the coefficients D_0, D_1, D_2, $D_1 - D_2$ and $D_{-\infty} - D_{\infty}$ calculated based on sliding windows.

The Rényi and Tsallis generalized dimension spectra for (X,Y,T), (X,Y,Z) and (X,Y,M)

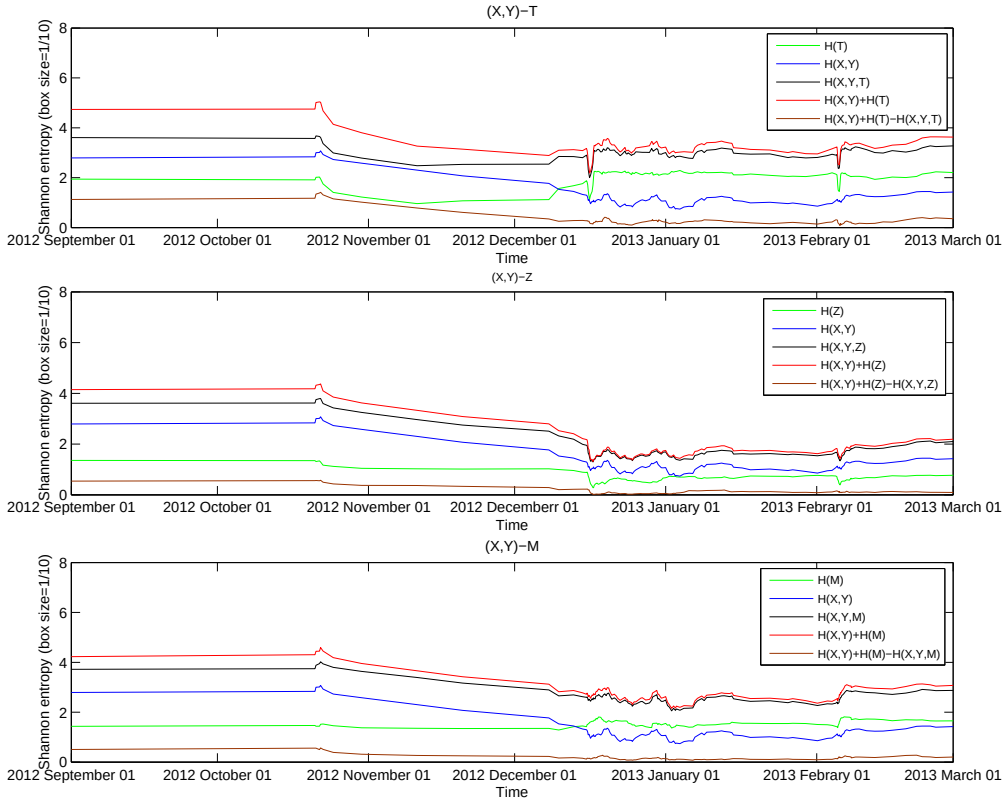

Figure 4.56: Torreperogil data. Visualization for a restricted temporal interval of the Shannon entropy values $H(D)$, $H(X,Y)$, $H(X,Y,D)$, $H(X,Y)+H(D)$ and $H(X,Y)+H(D)-H(X,Y,D)$,based on sliding windows (size 150 events), considering a fixed box size $\varepsilon = 1/10$ with D corresponding to time, depth and magnitude.

are plotted in Figure 4.58. We must note that the calculations for both generalized dimensions are performed considering varying box sizes between 1/15 and 1/2. In the Rényi case the parameter q varies between -25 and 25, whereas for the Tsallis case it varies between 0 and 2. These spectra reflect that (X,Y,T) has a greater degree of multifractal behaviour than the other cases.

These interpretations are confirmed by the values of the coefficients D_0, D_1, D_2, $D_1 - D_2$ and $D_{-\infty} - D_{\infty}$ and those of the generalized Tsallis dimensions DT_0, DT_1, DT_2, shown in Table 4.4.

To complete the study of structural characteristics at different scales, as was done in the other data sets, we analyze the variations in the structural characteristics for (X,Y,T) and (X,Y,Z) by means of the coefficients D_0, D_1, D_2, $D_1 - D_2$ and $D_{-\infty} - D_{\infty}$ values calculated for data

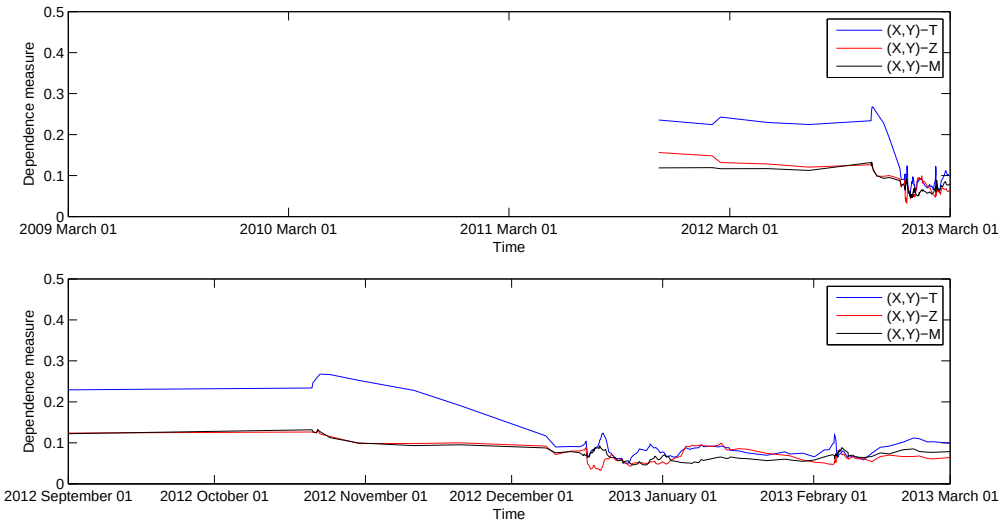

Figure 4.57: Torreperogil data. Values of the dependence coefficient for the interactions $(X,Y) \leftrightarrow T$, $(X,Y) \leftrightarrow Z$ and $(X,Y) \leftrightarrow M$, based on sliding windows (size 150 events).

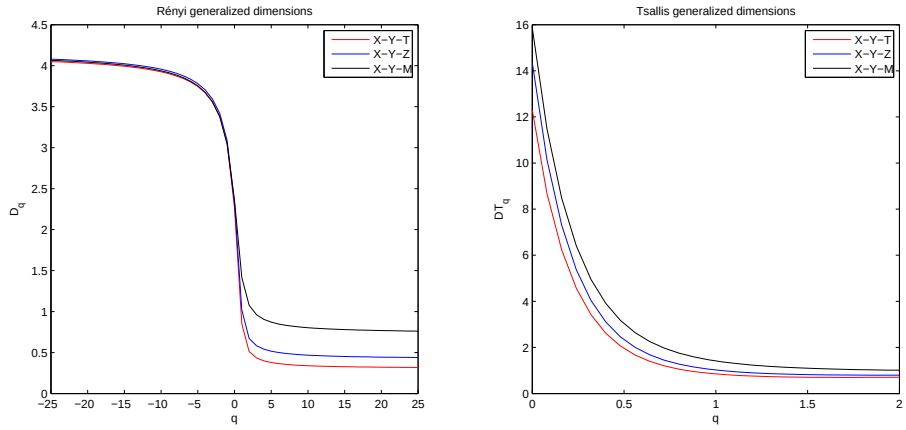

Figure 4.58: Torreperogil data. Rényi (left) and Tsallis (right) generalized dimensions curves for (X,Y,T), (X,Y,Z) and (X,Y,M).

segments using sliding windows, with and without considering the effects of the magnitude.

The D_0, D_1, and D_2 values displayed in Figures 4.59, 4.60, and 4.61, respectively, show similar patterns for both interactions except for the period with high seismic activity, during which significant differences can be observed. These plots reflect that the degrees of ordering and

Table 4.4: Torreperogil data. Values of generalized Rényi dimensions D_0, D_1, D_2, $D_1 - D_2$, $D_{-\infty} - D_{\infty}$, and values of generalized Tsallis dimensions DT_0, DT_1, DT_2.

	D_0	DT_0	D_1	DT_1	D_2	DT_2	$D_1 - D_2$	$D_{-\infty} - D_{\infty}$
$X-Y-T$	2.2643	12.2877	0.8497	0.8497	0.5131	0.1427	0.3366	3.7322
$X-Y-Z$	2.3200	14.3404	1.0205	1.0205	0.6721	0.7948	0.3834	3.6401
$X-Y-M$	2.3540	15.8491	1.4142	1.4142	1.0767	1.0098	0.4044	3.3041

clustering decrease during the days with higher concentration of events for (X,Y,T), whereas they keep high for (X,Y,Z). The departures between the curves with and without considering the magnitude effect show similar patterns in all cases.

In agreement with these results, there are not significant differences between (X,Y,T) and (X,Y,Z) with regard to the degree of multifractalilty, as displayed in Figure 4.62. Interestingly, during the high activity days it does not increase markedly unlike in the data sets of Agrón and El Hierro. In addition, when we compare the curves for $\phi(m) = 1$ and $\phi(m) = \exp(3m/2)$, we detect differences in the patterns for both cases. On the contrary, in the Agrón and El Hierro series these differences are not observed. This lack of association in terms of the multifractal characteristics between (X,Y,T) and magnitude indicates anomalies in the structural changes of this series.

As for complexity (see Figures 4.63 and 4.3.3), the two measures show similar patterns, as expected. However, the results reflect a low level of complexity during the periods with higher concentration of events, again, unlike in the data sets of Agrón and El Hierro.

In conclusion, the results of this study suggest the possible presence of external factors.

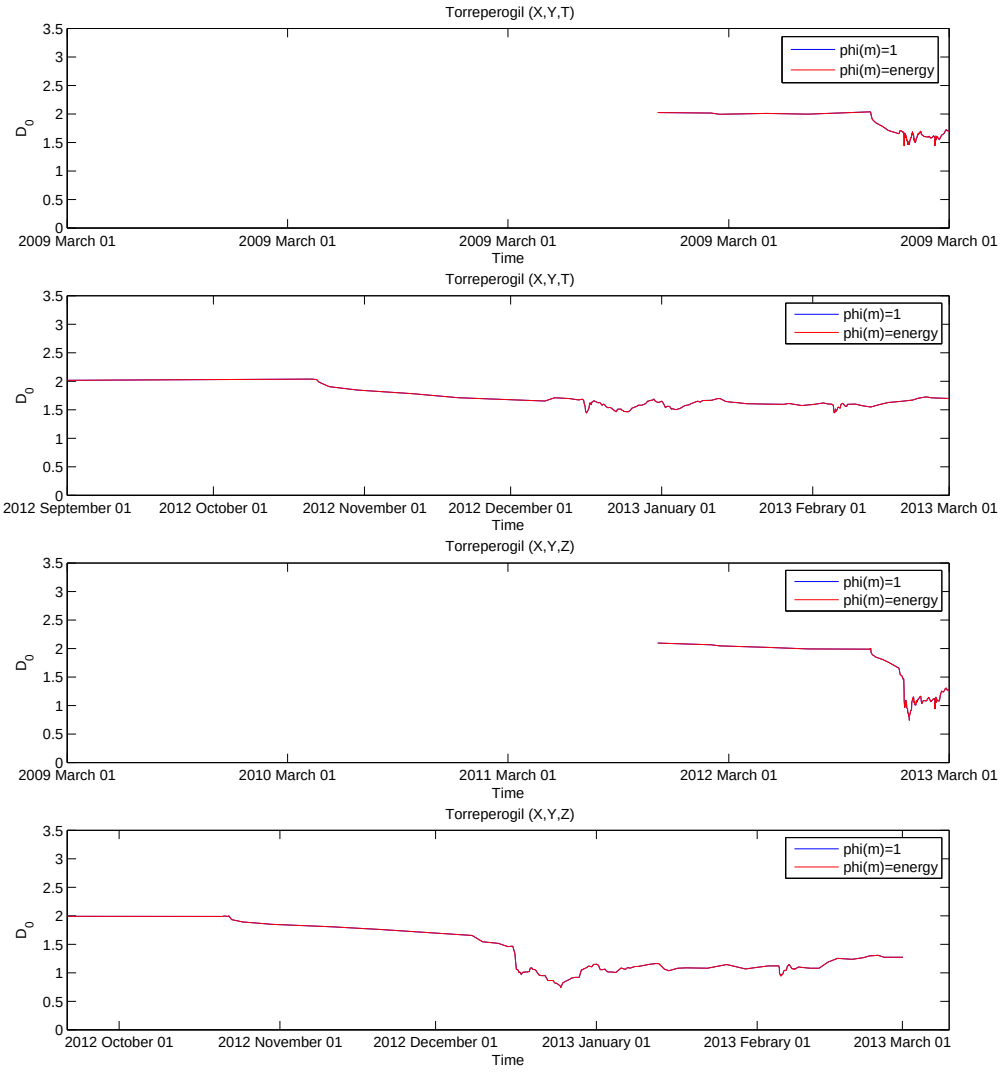

Figure 4.59: Torreperogil data. Variations in the ‘capacity’ dimension values over time based on sliding windows (size 150 events), with and without considering the effects of the magnitude.

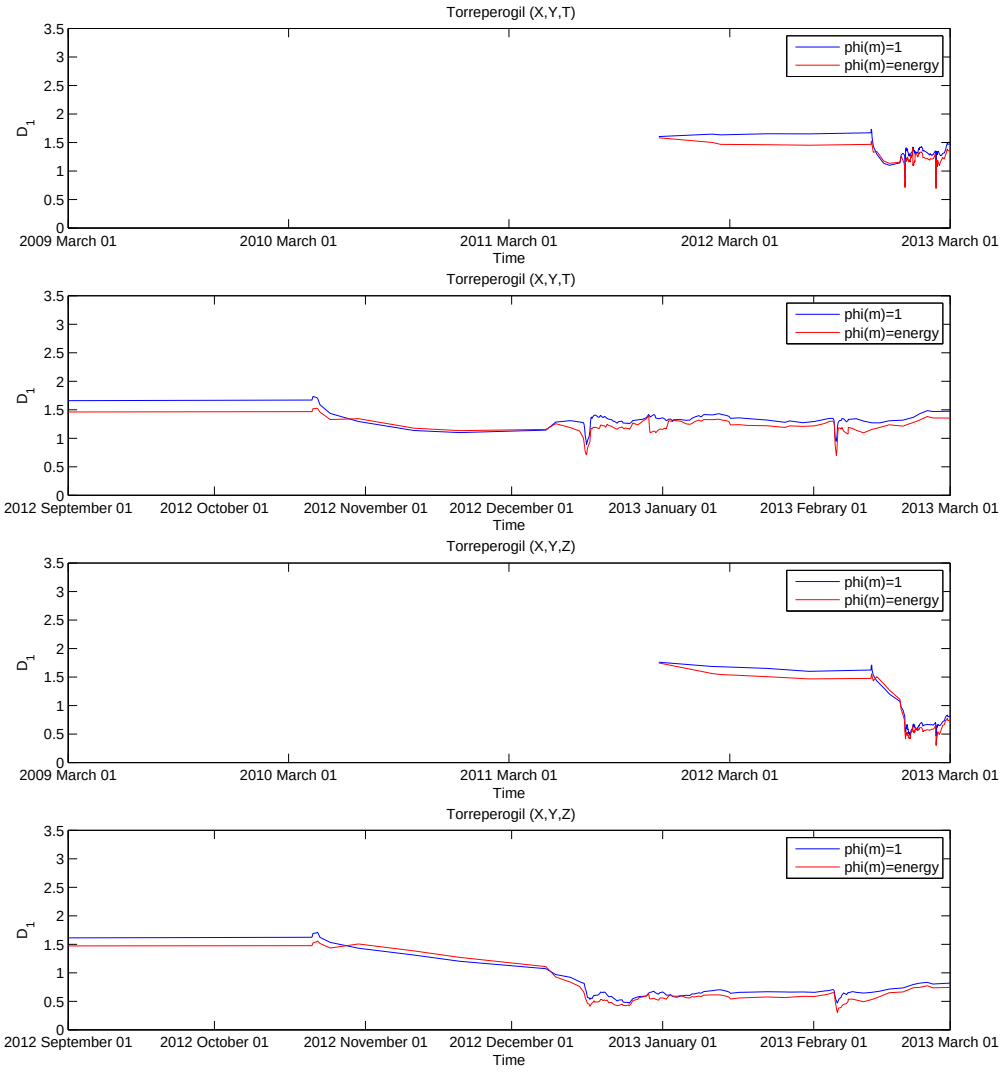

Figure 4.60: Torreperogil data. Variations in the ‘entropy’ dimension values over time based on sliding windows (size 150 events), with and without considering the effects of the magnitude.

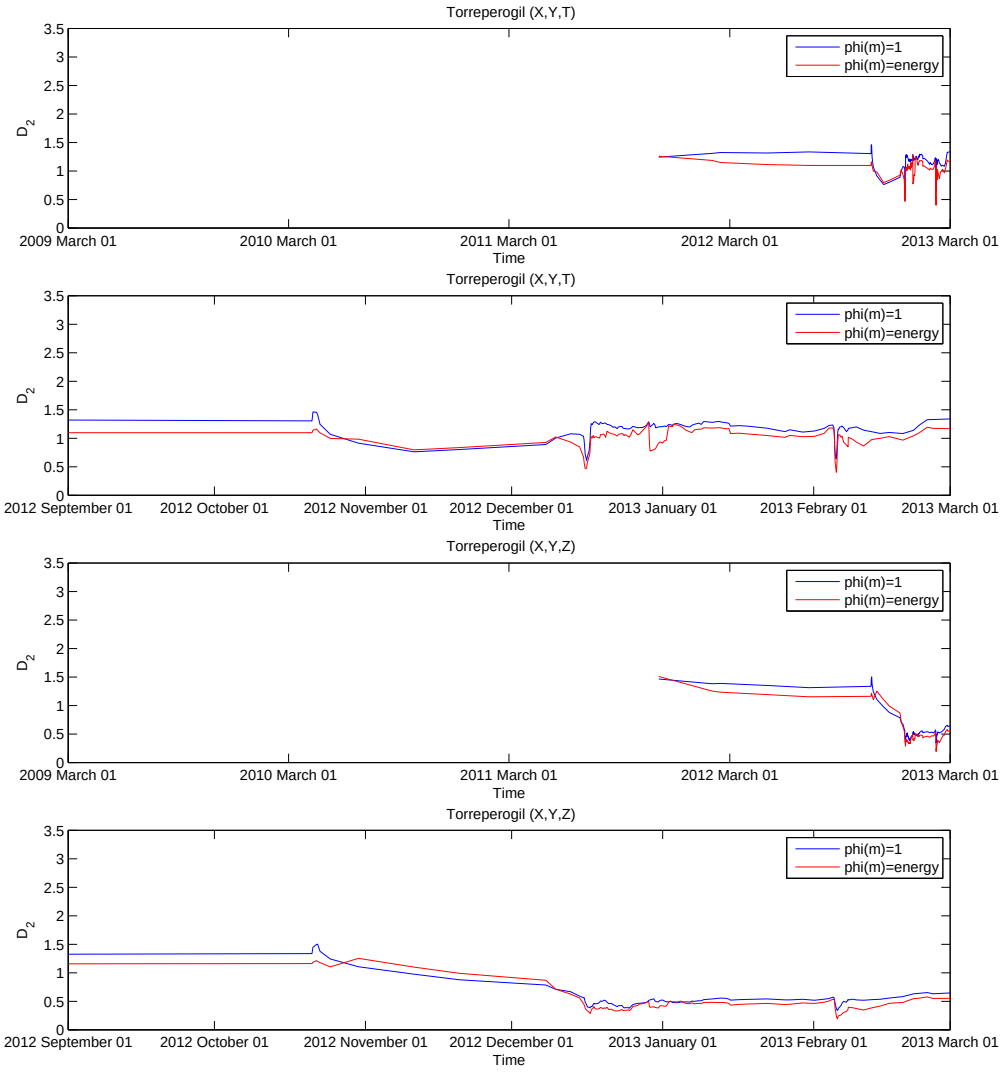

Figure 4.61: Torreperogil data. Variations in the 'correlation' dimension values over time based on sliding windows (size 150 events), with and without considering the effects of the magnitude.

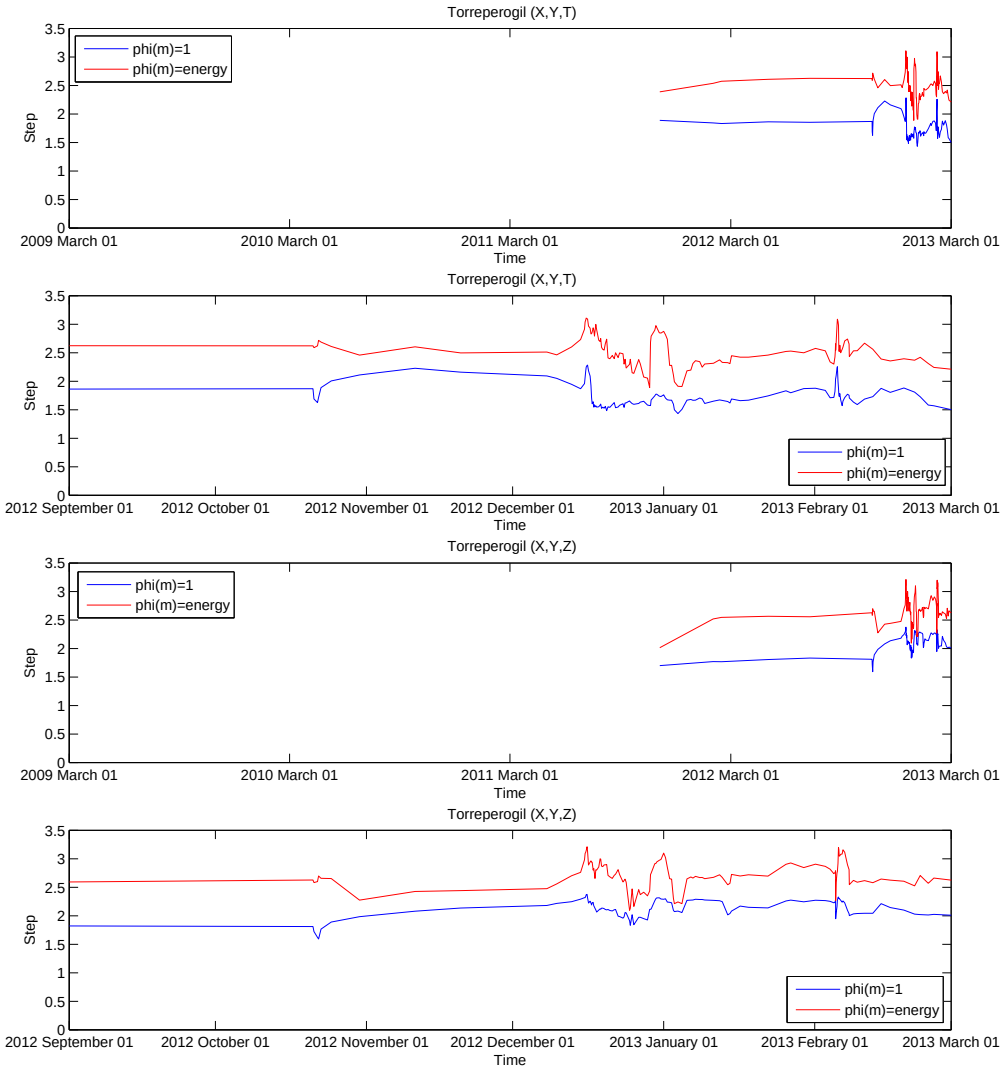

Figure 4.62: Torreperogil data. Variations in the multifractal step values over time based on sliding windows (size 150 events), with and without considering the effects of the magnitude.

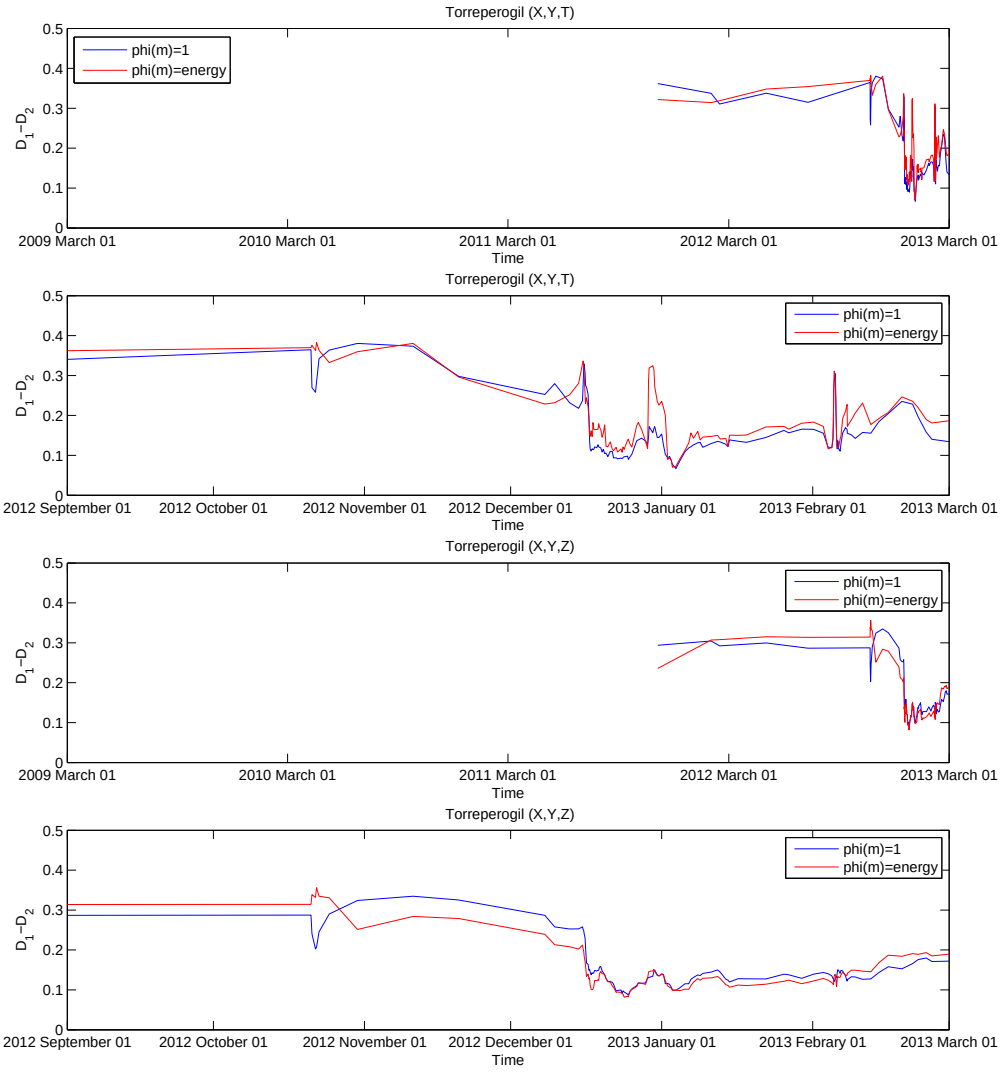

Figure 4.63: Torreperogil data. Variations in the $D_1 - D_2$ values over time based on sliding windows (size 150 events), with and without considering the effects of the magnitude.

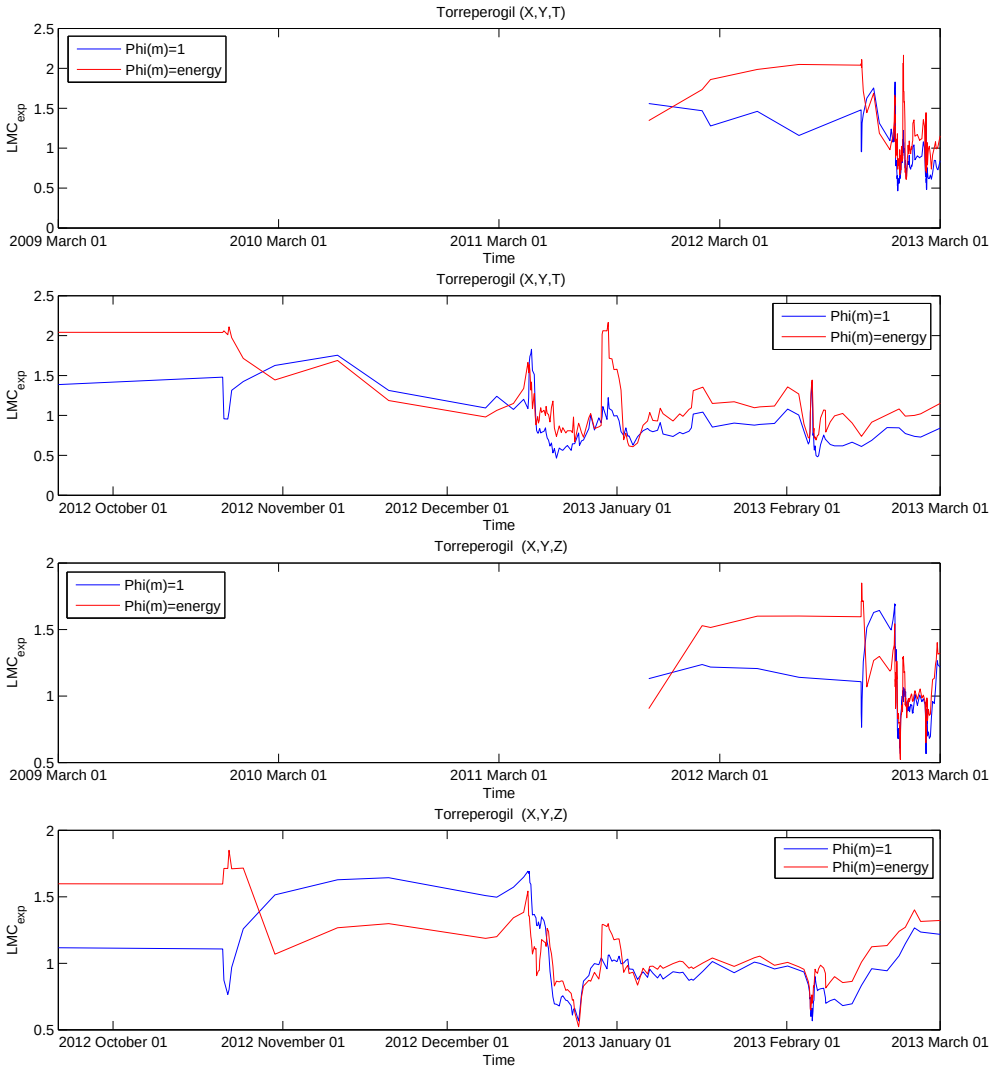

Figure 4.64: Torreperogil data. Variations in the LMC_{exp} values over time based on sliding windows (size 150 events), with and without considering the effects of the magnitude.

4.3.4 Frequency-magnitude distribution analysis

Now, we study the Torreperogil series by means of the frequency-magnitude distribution (FMD). As was done in the case of El Hierro series, we carry out an analysis of the whole catalogue, and also a study of the temporal evolution of the estimated parameters, using both models, non-weighted ($\phi(m) = 1$) and weighted ($\phi(m) = \exp(3m/2)$). Figure 4.65 shows the fitted curves on the observations, and Table 4.5 is composed of the estimated parameter values, misfit and R^2 for the whole data set. The parameter estimates for both models do not present significant differences. The values of the parameter q (1.4423 and 1.4253, respectively) indicate that the physical state is away from equilibrium. This means that larger earthquakes can be expected. The released energy in proportion to the fragment size is not high. The values of misfit and R^2 show that the fitting performed using the weighted nonlinear least-squares model improves the estimation.

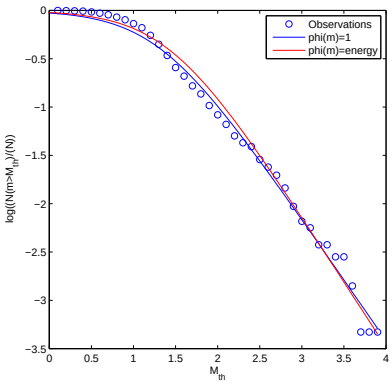

Figure 4.65: Torreperogil data. Representations of the fitted curves for the whole sequence and the three phases for both models.

Table 4.5: Torreperogil data. Values of parameters q and a, misfit and R^2 of the fitted curves based on the non-weighted and weighted models.

	q	a	misfit	R^2
$\phi(m) = 1$	1.4423	60.8195	0.0741	0.9924
$\phi(m) = \exp(3m/2)$	1.4253	86.5341	0.0117	0.9961

To complete the study we analyze the temporal evolution of the series through the variations

in the parameters over time. The parameters are calculated based on sliding windows of 150 events with 90 per cent of overlapping. Figure 4.66 displays the magnitude distribution over time and estimates of parameter q for both models, for the whole temporal range and for a restricted range corresponding to the high activity period. In the third plot, we can observe how the estimated values of q decrease in the period of high seismic activity, which indicates that during this period the physical state is close to equilibrium. This result might be interpreted as if larger earthquakes could not be expected, unlike the results obtained for the El Hierro series. In these plots, we can observe how the estimates of q, when we consider the energy effects, are generally higher, but particulary the deviations are higher during the regular activity periods.

The variations over time of the parameter a values obtained for both models are displayed in the bottom plots of Figure 4.67. In the top plots, we observe two periods in which a large amount of energy is released in proportion to the fragment size. During these days larger earthquakes occur but these strong increments can also be due to the small size of the fragments. The estimated values for the weighted model (red line) are higher in the periods with larger earthquakes and smaller in the other cases.

Finally, we study the goodness of fit of both models to the data. The plots of Figure 4.68 display the variation of the misfit and R^2 over time (from top to bottom). These results show that when we use the weighted model the fitting improves, especially when the non-weighted model presents a higher misfit.

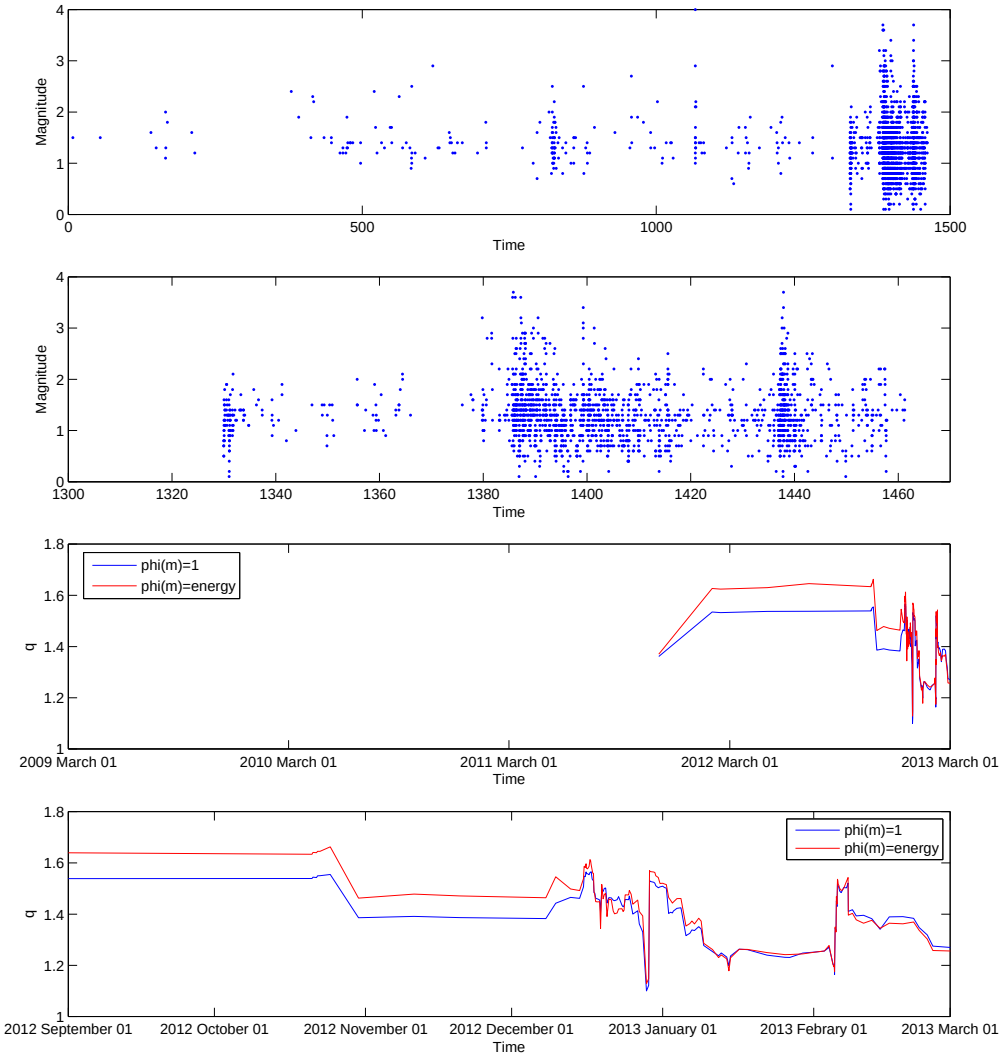

Figure 4.66: Torreperogil data. From top to bottom, magnitude distribution, and estimated values of parameter q over time based on sliding windows (size 150 events), for the non-weighted and weighted models.

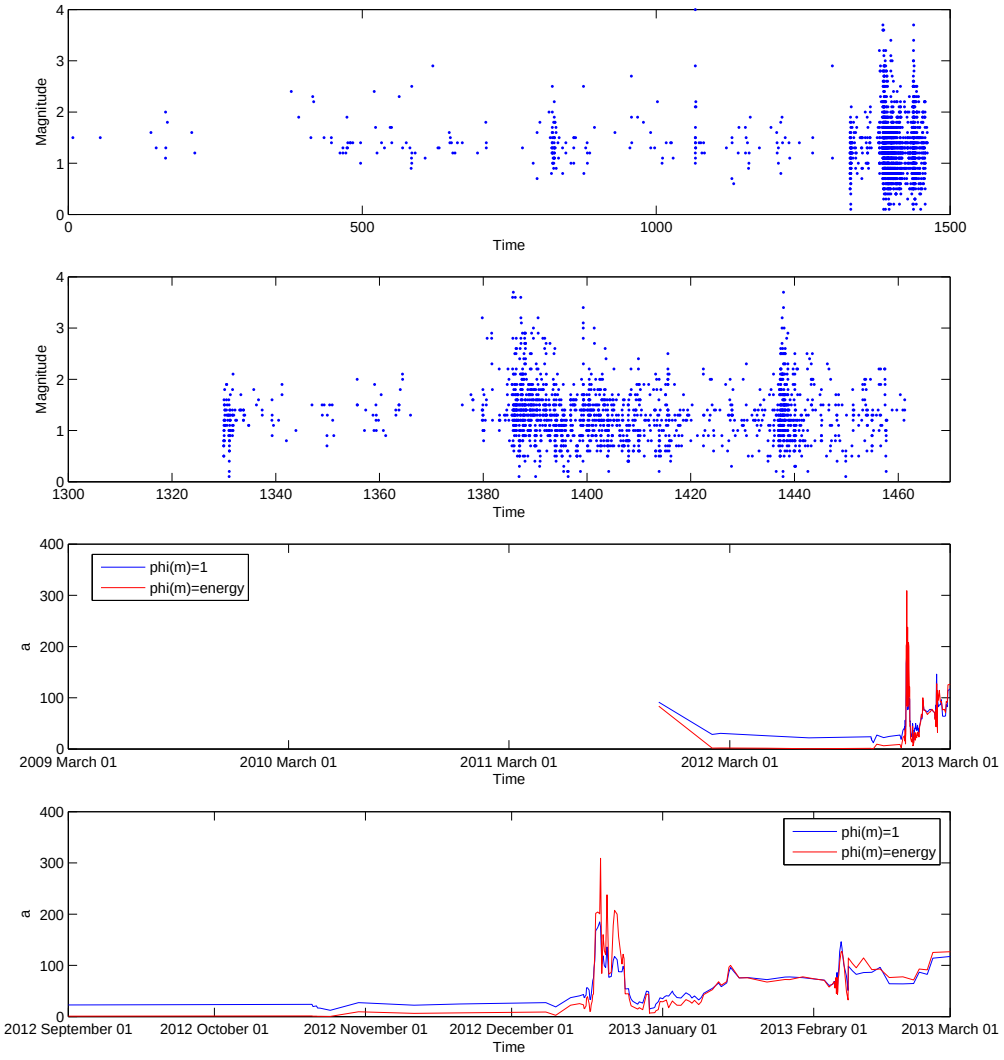

Figure 4.67: Torreperogil data. From top to bottom, magnitude distribution, and estimated values of parameter *a* over time based on sliding windows (size 150 events), for the non-weighted and weighted models.

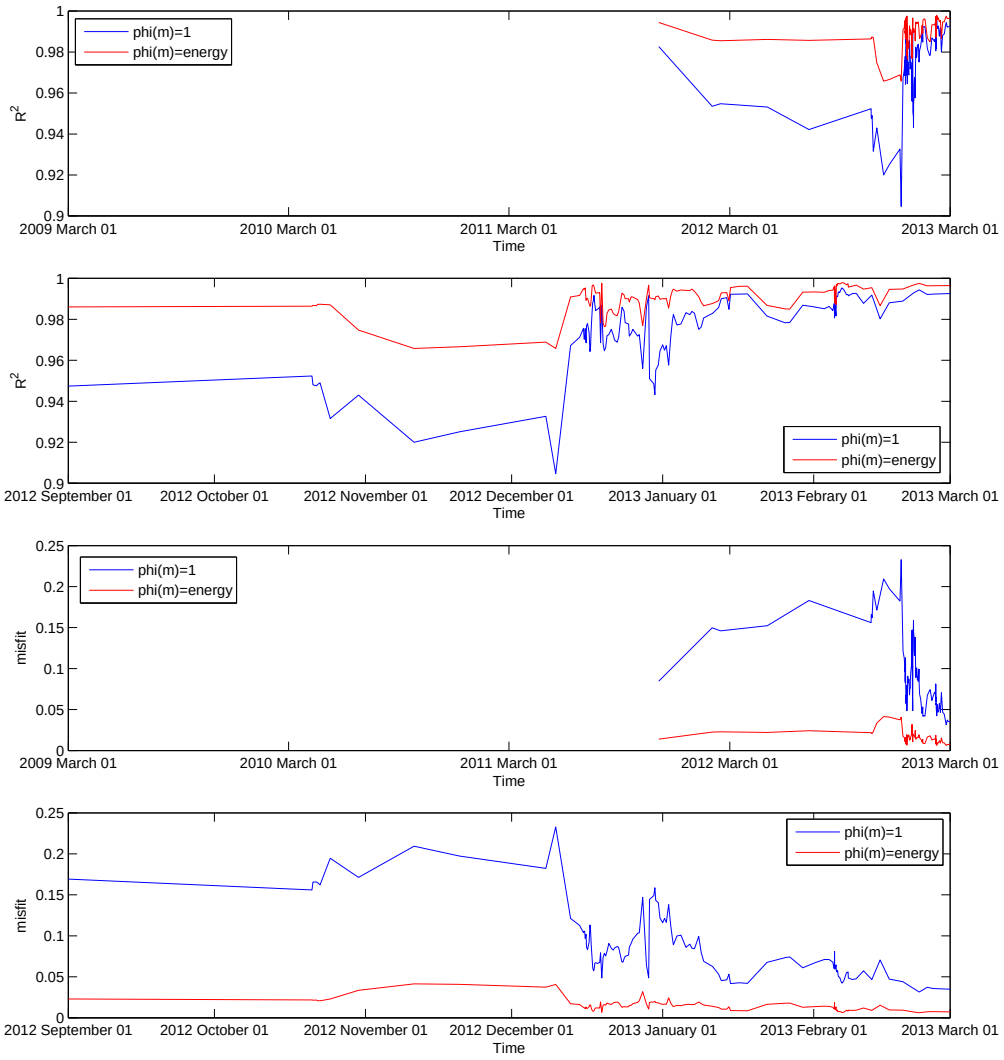

Figure 4.68: Torreperogil data. From top to bottom, misfit and R^2 values over time based on sliding windows (size 150 events), for the non-weighted and weighted models.

4.4 Synthesis

The study of the three data sets by means of the proposed techniques has mainly shown that these are very useful tools to describe seismicity. The results obtained provide relevant information relating to the established objectives for each series. In particular, regarding Agrón series the study performed combining the techniques with temporal transformations has been found to be very practical for assessment of the spatio-temporal association. As for El Hierro series, the techniques applied have detected temporal changes in the structural characteristics which could be seen as a precursory indicator to the volcanic eruption. Finally, the study of Torreperogil series has shown some results not present in the other data sets, which may suggest the existence of external factors. However, such an assertion must still be investigated based on comparative analysis with similar data sets.

Part III

COMPLEMENTARY ASPECTS AND CONCLUSIONS

Chapter 5

Synthesis, Complementary Aspects and Open Lines

5.1 Synthesis

In this thesis work, we study various aspects related to seismicity through the analysis of the point processes associated with this geophysical phenomenon. Two main aspects of interest in this context concern scaling properties and dimensional interaction. Specifically, we focus on the statistical use of information-theoretic concepts and mesures in the analysis of the structural complexity of seismic distributional patterns.

Specifically, in Part I we present a preliminary view on some of the most relevant and successful techniques in this context. Firstly (Chapter 1), we refer to usual definitions of entropy and complexity for discrete probability distributions and to some of their related properties. Secondly (Chapter 2), we introduce the main tools of multifractal analysis, namely the generalized Rényi dimensions and multifractal spectrum. Moreover, we introduce the main empirical laws used to describe seismicity, as well as the non-extensive versions for the frequency-magnitude distribution of earthquakes.

In Part II, we propose some technical and methodological extensions based on the former elements (Chapter 3), and we apply them to the study of three seismic series (Chapter 4). Specifically, by means of the limiting behaviour of information and complexity related quantities we establish a complexity measure connected to the LMC complexity measure, showing the potential use of related functionals for quantitative assessment of 'multifractal complexity', and we also define a 'dependence coefficient' based on the generalized Rényi dimensions. In addition, given the increasing interest in the non-extenive statistics of Tsallis to describe seismicity, we formulate a Tsallis entropy-based version of generalized dimensions. Further, we propose the incorporation of the magnitude effects to the different analyses in terms of weighted box-counting distributions. These techniques are applied in combination with several useful com-

plementary approaches, namely different transformations on the temporal component, sliding windows, and weighted non-linear regression. In Chapter 4, we show the implementation and application of these technical and methodological aspects through the study of three seismic series of an underlying different nature. In the case of Agrón series, we perform an extensive analysis of the spatio-temporal interaction, discriminating the behaviour in different activity subperiods. On the other hand, for the El Hierro series, which involves a volcanic phenomenon, the results obtained show structural changes in the temporal evolution, which might be interpreted as precursory behaviour to the volcanic eruption. Finally, the unexpected structural characteristics seen in the Torreperogil series might indicate the potencial existence of external factors.

In conclusion, the results obtained in the applications to seismic series show the usefulness of the proposed techniques to describe seismicity. In particular, the complexity measure proposed provides an appropriate tool to quantify complexity in the multifractal domain, in connection with the LMC complexity measure. As for the dimensional interaction, the analysis performed indicates that the dependence coefficient, formulated here in terms of the 'entropy' dimension, is a useful tool to study the degree of association between components in point processes with multifractal characteristics. Regarding the new formulation of generalized dimensions based on Tsallis entropy, we must point out that it has offered satisfactory results, although further study on analytical aspects is required. Besides, the combination of these techniques with temporal transformations provides extra information in the study of the spatio-temporal interaction. In addition, the incorporation of the magnitude effects in the different analyses has suggested an intrinsic structural association between the magnitudes and the space-time coordinates. Finally, in the study of seismicity by means of the non-extensive frequency-magnitude distribution, the results obtained show that estimates are improved in all cases when we consider a weighted non-linear regression to perform the fitting.

Some of the developments and applications derived in this thesis work are collected in the following publications:

- Angulo, J. M., Esquivel, F. J., Structural complexity in space-time seismic event data. *Stochastic Environmental Research and Risk Assessment*, DOI 10.1007/s00477-013-0807-x, 2013.

- Angulo, J. M., Esquivel, F. J., Statistical complexity analysis of spatiotemporal dynamics. *Mathematics of Planet Earth. Proceedings of the 15th Annual Conference of the International Association for Mathematical Geosciences (IAMG 2013)*. Lecture Notes in Earth System Sciences, ISBN 978-3-642-32407-9, Pardo-Igúzquiza, E., Guardiola-Albert, C., Heredia, J., Moreno-Merino, L., Durán, J. J., Vargas-Guzmán, J. A. (Eds.), Springer, 185-188, 2013.

Further aspects are incorporated as part of other articles in preparation, involving:

- The analysis of the seismic series of El Hierro by means of complexity and multifractal techniques and the weighted non-extensive frequency-magnitude distribution.

- Discrimination of effects in relation to the unexpected behaviour of seismicity in the region of Torreperogil.

- Application of the techniques proposed in this thesis to the study of structural properties of excursion sets in random fields. (A brief exposition of this idea and a preliminary study performed are descibed in the next section.)

In addition, the contents of this research have been presented in different meetings. As invited talks:

- Angulo J. M., Esquivel, F. J., Structural complexity in space-time seismic event data. *VI International Workshop on Spatio-Temporal Modelling (METMAVI)*, Guimãraes, Portugal, September 12-14, 2012.

- Angulo J. M., Esquivel, F. J., Information-related approaches for statistical complexity assessment. *The 25th Annual International Conference on Statistics and Modelling in Human and Social Sciences*, Cairo, Egypt, March 25-28, 2013.

- Angulo, J. M., Esquivel, F. J., Statistical complexity analysis of spatiotemporal dynamics. *The 15th Annual Conference of the International Association for Mathematical Geosciences (IAMG 2013)*, Madrid, Spain, September 2-6, 2013.

As contributed talks:

- Esquivel, F. J., Angulo J. M., Estudio del efecto de la dinámica temporal en la estructura de la situación geográfica de los epicentros de una secuencia sísmica. *XXXIII Congreso Nacional de Estadística e Investigación Operativa (SEIO 2012)*, Madrid, Spain, April 17-20, 2012.

- Esquivel, F. J., Madrid, A. E., Angulo, J. M., Multifractal characteristics of spatial threshold exceedances. *5th International Conference of the ERCIM Working Group on Computing & Statistics (ERCIM 2012)*, Oviedo, Spain, December 1-3, 2012.

- Esquivel, F. J., Angulo J. M., Some aspects of structural complexity dynamics in seismic data of Torreperogil. *XXXIV Congreso Nacional de Estadística e Investigación Operativa (SEIO 2013)*, Castellón, Spain, September 11-13, 2013.

In the modality of poster:

- Angulo, J. M., Esquivel, F. J., Un estudio sobre dimensiones fractales generalizadas en procesos puntuales espaciales. *XXXII Congreso Nacional de Estadística e Investigación Operativa (SEIO 2010)*, A Coruña, Spain, September 14-17, 2010.
- Esquivel, F. J., Angulo J. M., Information loss under projections from seismic data. *58th World Statistics Congress of the International Statistical Institute (ISI 2011)*, Dublin, Ireland, August 21-26, 2011.

5.2 Complementary Aspects

The possible relation between stochastic processes and fractal theory can be a useful tool to describe some properties of a process by means of fractal measures. In this context, Orey (1970) proved that the Hausdorff dimension can detect important properties such as stationarity and certain properties inherent to a Gaussian process with stationary increments, in the temporal domain. In addition, Madrid et al. (2012) showed that the analysis of spatial point patterns defined from excursion sets in terms of centroids of the connected component and A-exit points provides information on relevant characteristics of the underlying random field. Following this idea, we have considered in the context of this thesis work the study of structural characteristics of excursion sets by the application of the complexity and multifractality tools to such point patterns.

For this purpose, we have considered realizations of Gaussian random fields generated using the Cauchy class (Gneiting and Schlater, 2004), defined by the homogeneous and isotropic covariance function

$$C(h) = \sigma^2(1+|h|\alpha)^{\beta/\alpha}, \tag{5.1}$$

where σ^2 stands for the variance of the corresponding random field, and $\alpha \in (0,2]$ and $\beta \geq 0$. This class contains as particular case the 'Cauchy model', for $\alpha = 2$. The parameter α is associated with local variability and it determines the fractal dimension of realizations, $D = n + 1 - \alpha/2$. Independently, the parameter β is related to the dependence range, and for $\beta \in (0,1)$ the process has long memory with Hurst coefficient $H = 1 - \beta/2$. Figure 5.1 displays four simulated realizations on the square $[1,200]^2$ (using the same pseudorandom numbers) based on crossed combinations of parameter values $\alpha = 0.5, 2$ and $\beta = 0.1, 0.9$, with common $\sigma = 0.1$. The images show clear differences regarding local variation and larger scale structuring.

In this study, we consider the threshold exceedance sets from 50 realizations of each model, taking thresholds corresponding to the range of percentiles between 10 and 98. Figures 5.2, 5.3, 5.4 and 5.5 display the excursion sets for each realization of Figure 5.1, corresponding to the empirical quartiles as thresholds.

The spatial point patterns analyzed are determined by the centroids of the connected components and by A-exit points for each threshold exceedance set:

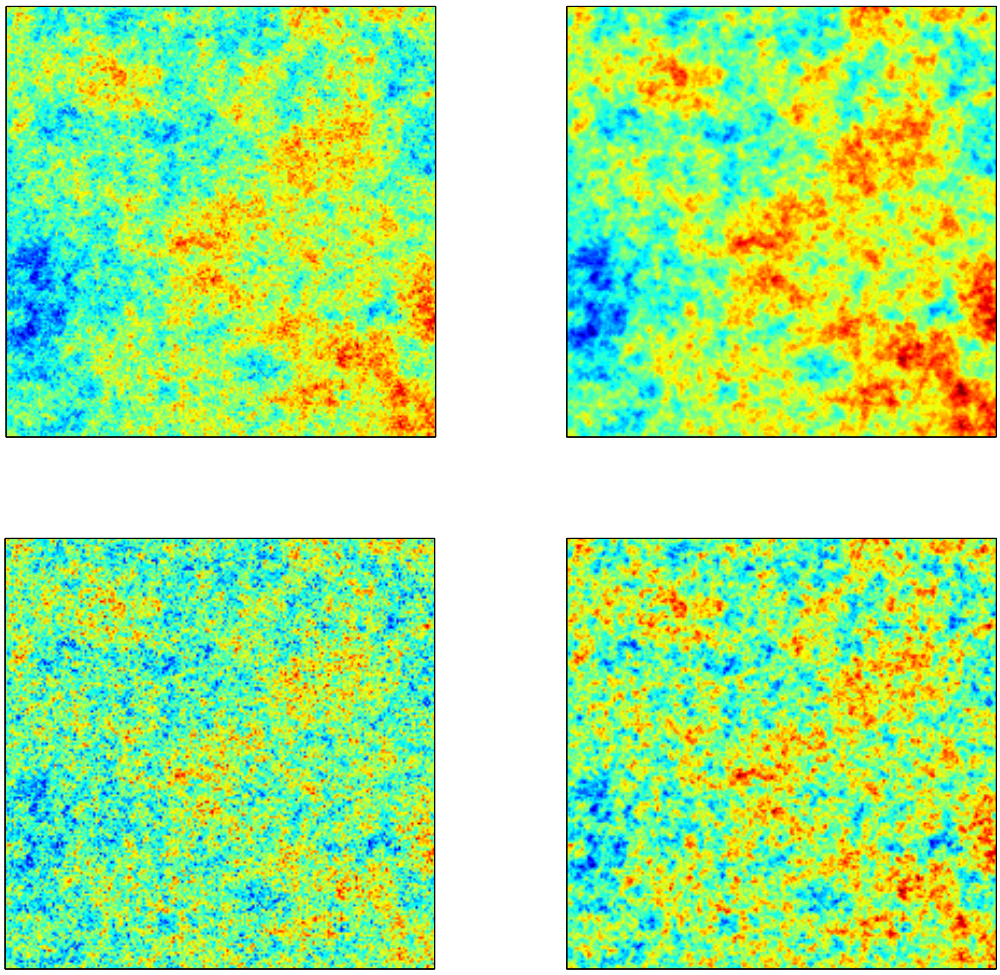

Figure 5.1: Simulated realizations using the Cauchy class with $\alpha = 0.5, 2$ (from left to right) and $\beta = 0.1, 0.9$ (from top to bottom).

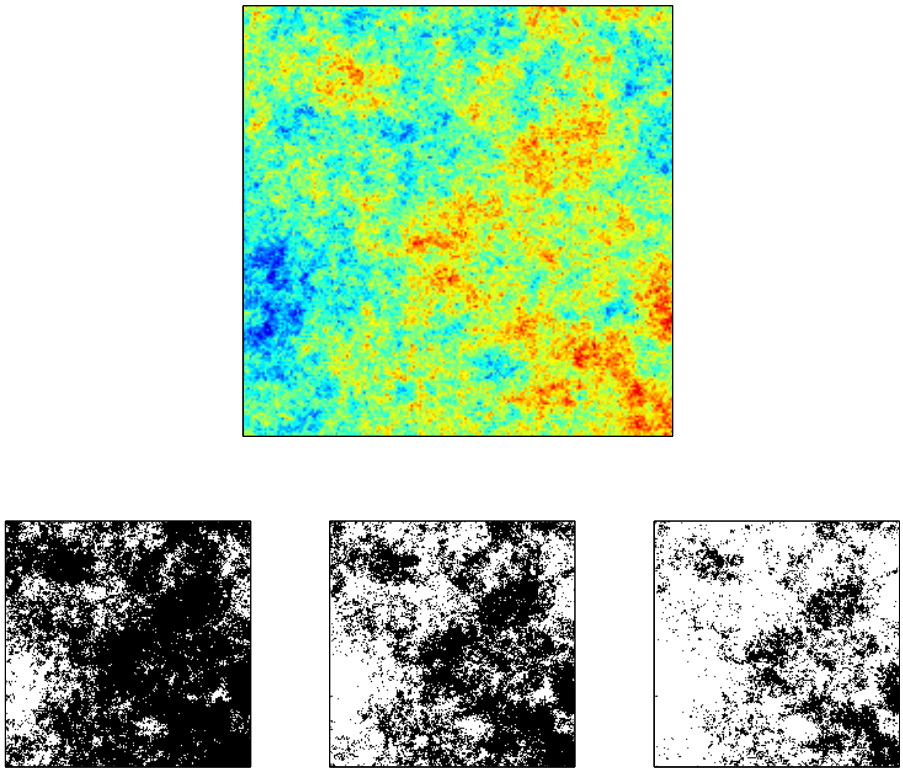

Figure 5.2: Random field (top) and excursion sets for the thresholds corresponding to the quartiles (bottom from left to right), based on realization of Figure 5.1 for $\alpha = 0.5$ and $\beta = 0.1$

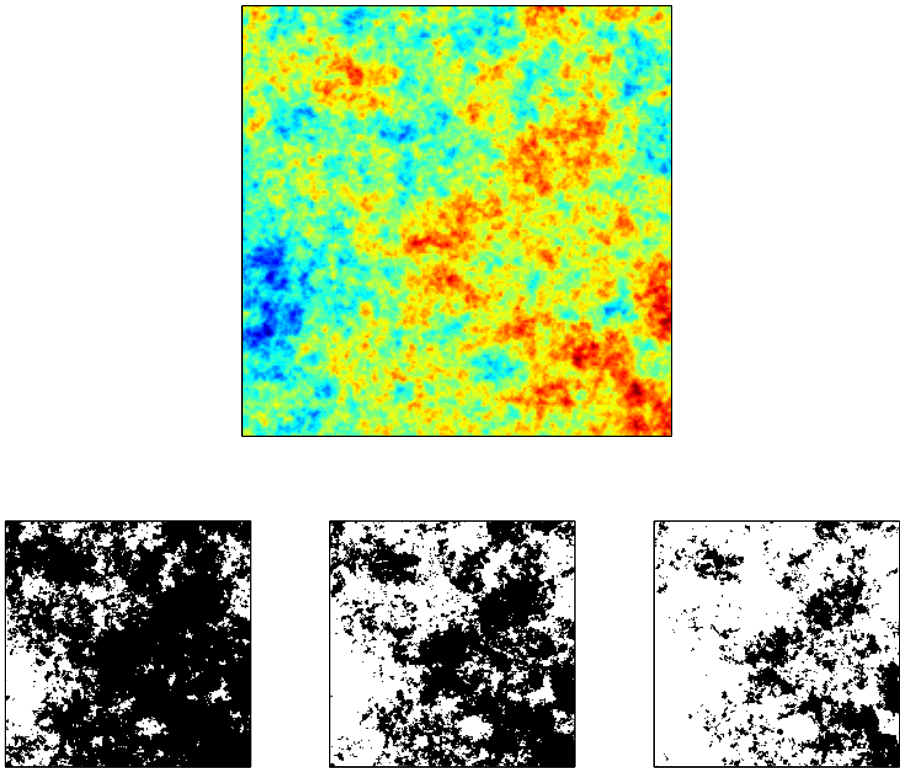

Figure 5.3: Random field values (top) and excursion sets for the thresholds corresponding to the quartiles (bottom from left to right), based on realization of Figure 5.1 for $\alpha = 2$ and $\beta = 0.1$

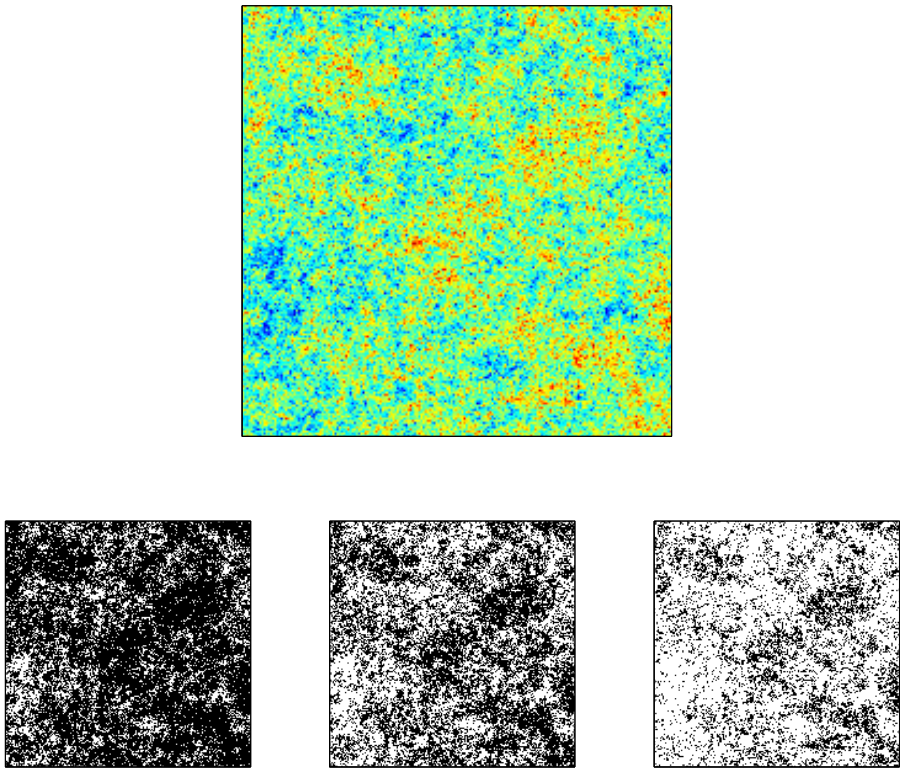

Figure 5.4: Random field (top) and excursion sets for the thresholds corresponding to the quartiles (bottom from left to right), based on realization of Figure 5.1 for $\alpha = 0.5$ and $\beta = 0.9$

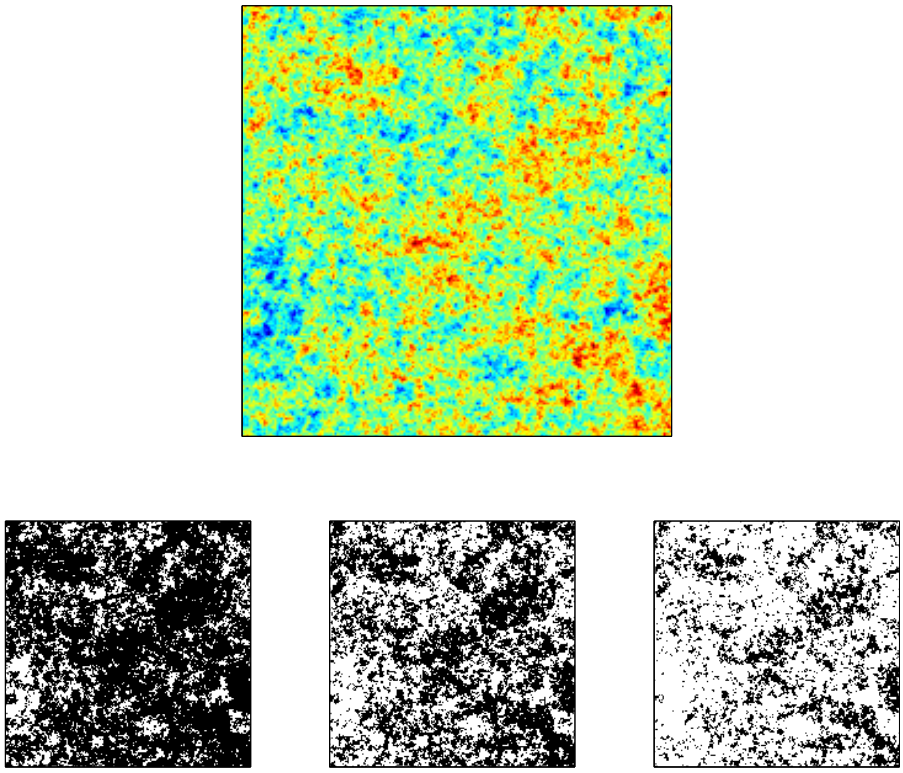

Figure 5.5: Random field (top) and excursion sets for the thresholds corresponding to the quartiles (bottom from left to right), based on realization of Figure 5.1 for $\alpha = 2$ and $\beta = 0.9$

- 'Centroids' case: We take each connected component as an isolated event, represented by its centroid coordinates (see, for example, Madrid et al., 2012).

- '*A*-exit points' case: *A*-exit points are defined, in a certain formal sense (see Piterbarg, 1996), as salient point in the threshold exceedance set towards a given reference focal point at infinity, then reflecting its boundary roughness among other characteristics (see, for instance, Piterbarg, 1996).

Figures 5.6 and 5.7 show the excursions sets for the threshold corresponding to the 90th percentile on the basis of the realizations displayed in Figure 5.1, with the points indicating the centroids and *A*-exit points, respectively (here a single left focal point is considered). We can observe how by columns, from left to right, the number of connected components decreases whereas their sizes increase. This is due to the fact that when the value of the parameter α increases the fractality degree decreases leading to less fragmentation and roughness. On the other hand, when the value of β increases, plots from top to bottom, the dependence ranges are reduced and, therefore, the level of disaggregation increases, with lesser heterogeneity in the sizes of connected components.

Based on the point patterns derived from the 50 realizations of each model and considering the thresholds between the percentiles 10 and 98, we study some structural characteristics. Firstly, we assess the spatial association $X \leftrightarrow Y$ by means of the dependence coefficient introduced in Chapter 3. Figure 5.8 displays the average of the dependence coefficients values corresponding to the 50 realizations varying the threshold considered, for the point patterns defined by centroids (left plot) and *A*-exit points (right plot). The representations show that the degree of spatial association is sensitive to the variation in the threshold considered, as well as to the parameter specifications in the underlying model, with clear differences between the centroids and *A*-exit points cases.

To perform the study of the structural characteristics at different scales, we calculate the coefficients D_0, D_1, D_2, the multifractal step $D_{-\infty} - D_{\infty}$, and the complexity measure $D_1 - D_2$. Figure 5.9 displays the results obtained considering the point patterns defined by centroids for the four models. In this figure, we can see, among other aspects, that the changes in the multifractal behaviour are more sensitive to the increase of the threshold for lower α and higher β. The corresponding results obtained when we consider the spatial point patterns defined by *A*-exit points are displayed in Figure 5.10. Similar interpretations can be derived, although there are visible differences in the location of the trend changes.

From the preliminary application performed, we can conclude that the complexity and multifractality tools considered can be useful for quantitative assessment on structural characteristics associated with spatial threshold exceedance sets (Esquivel and Angulo, 2012); in particular, for detection of differential features depending on the underlying random field properties and the threshold.

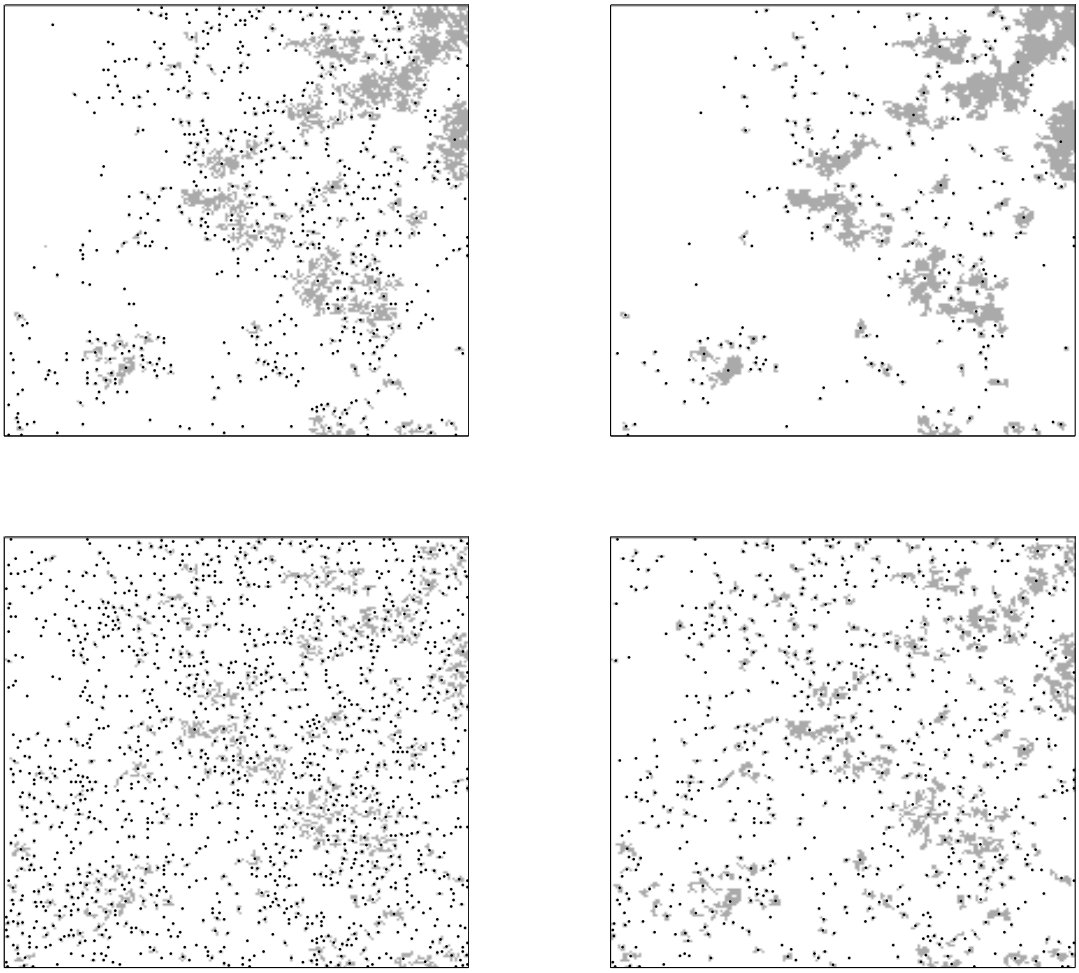

Figure 5.6: Excursion sets for the threshold corresponding to the 90th percentile, based on realizations of Figure 5.1. Black dots represent the centroids of connected components.

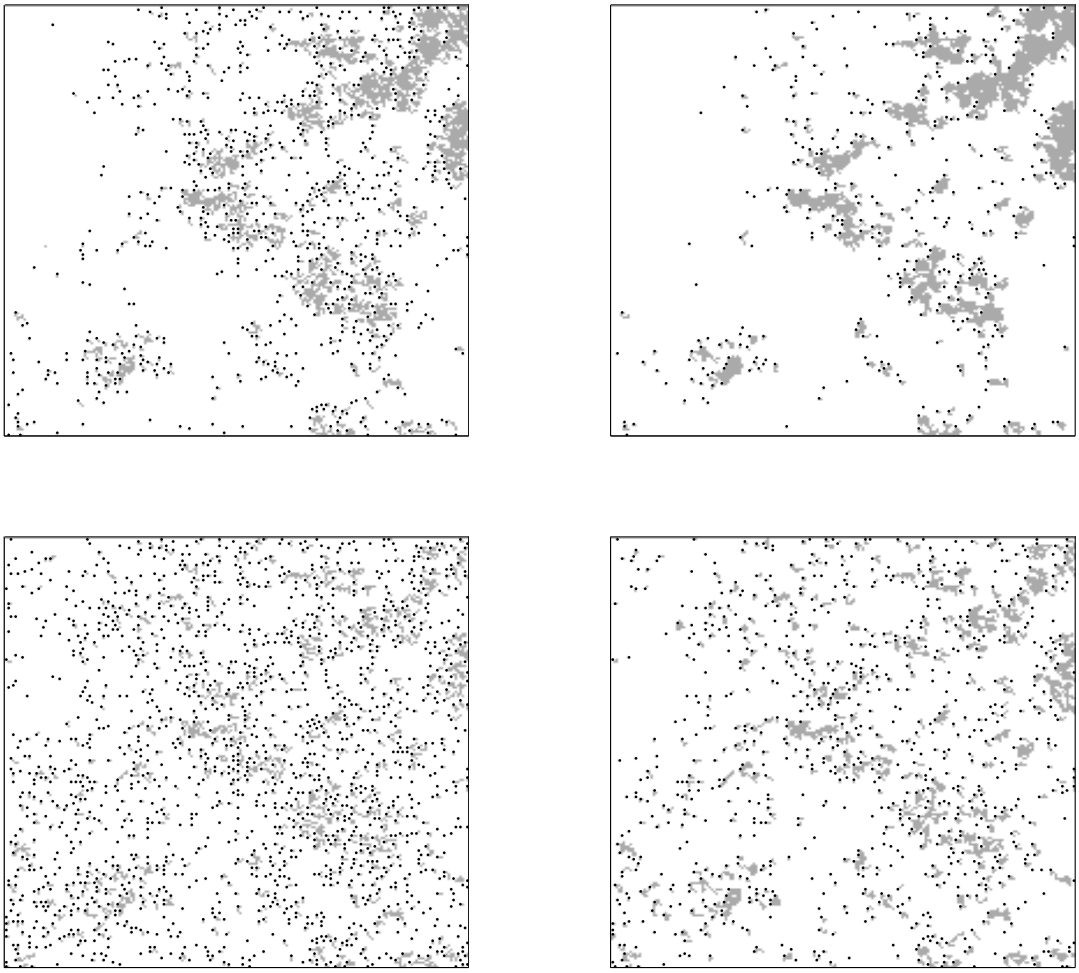

Figure 5.7: Excursion sets for the threshold corresponding to the 90th percentile, based on realizations of Figure 5.1. Black dots represent the *A*-exit points.

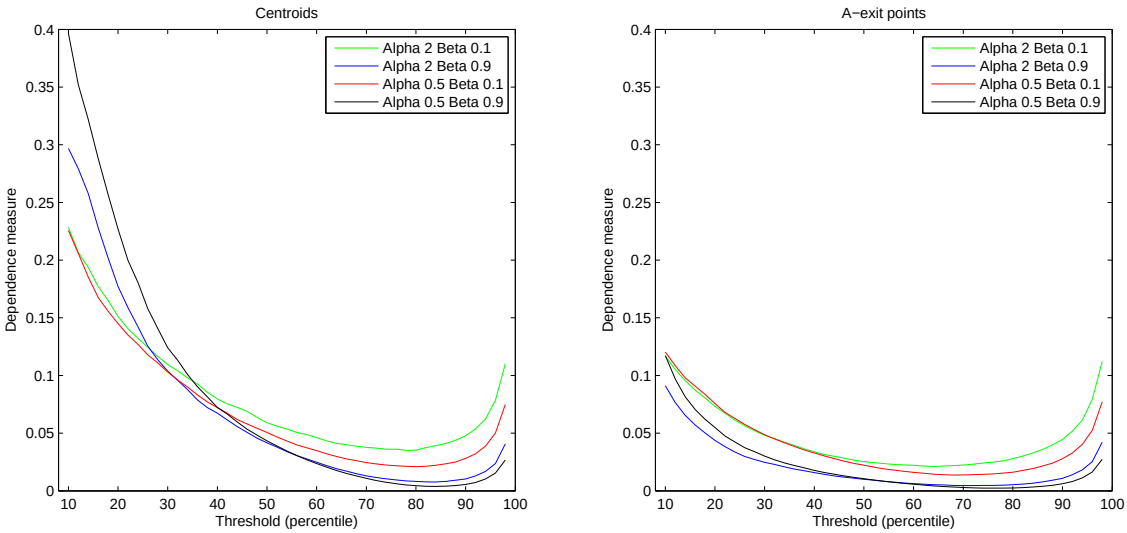

Figure 5.8: Dependence coefficient values for the spatial point patterns defined by centroids (left) and *A*-exit points (right) for the four models

5.3 Open Lines

The development of this thesis work has opened various lines and further directions for continuing related research, focused on different aspects including:

- Analytical study of the generalized dimensions based on Tsallis entropy

  In seismicity, the facts that the process of stick-slip between tectonic plates is non-additive (Sotolongo-Costa, 2012) and that the non-extensive statistics allows interactions at different orders, led us to introduction of a new formulation of generalized dimensions based on Tsallis entropy (Chapter 3). In this context, we are interested in performing an analytical study of related properties, and deriving a formulation of a proper related definition of an associated multifractal spectrum.

- Formulation of new complexity measures in the multifractal domain

  In Chapter 3 we propose a 'generalized complexity' measure in the multifractal domain based on the limiting behaviour of the Rényi entropy measure, and in connection with the generalized complexity measure $\tilde{C}^{\alpha,\beta}$ (López-Ruiz et al., 2009). In this context, with the aim of exploiting the variational information of the generalized dimension curve for discriminatory assessment of multifractal complexity characteristics, we consider the study

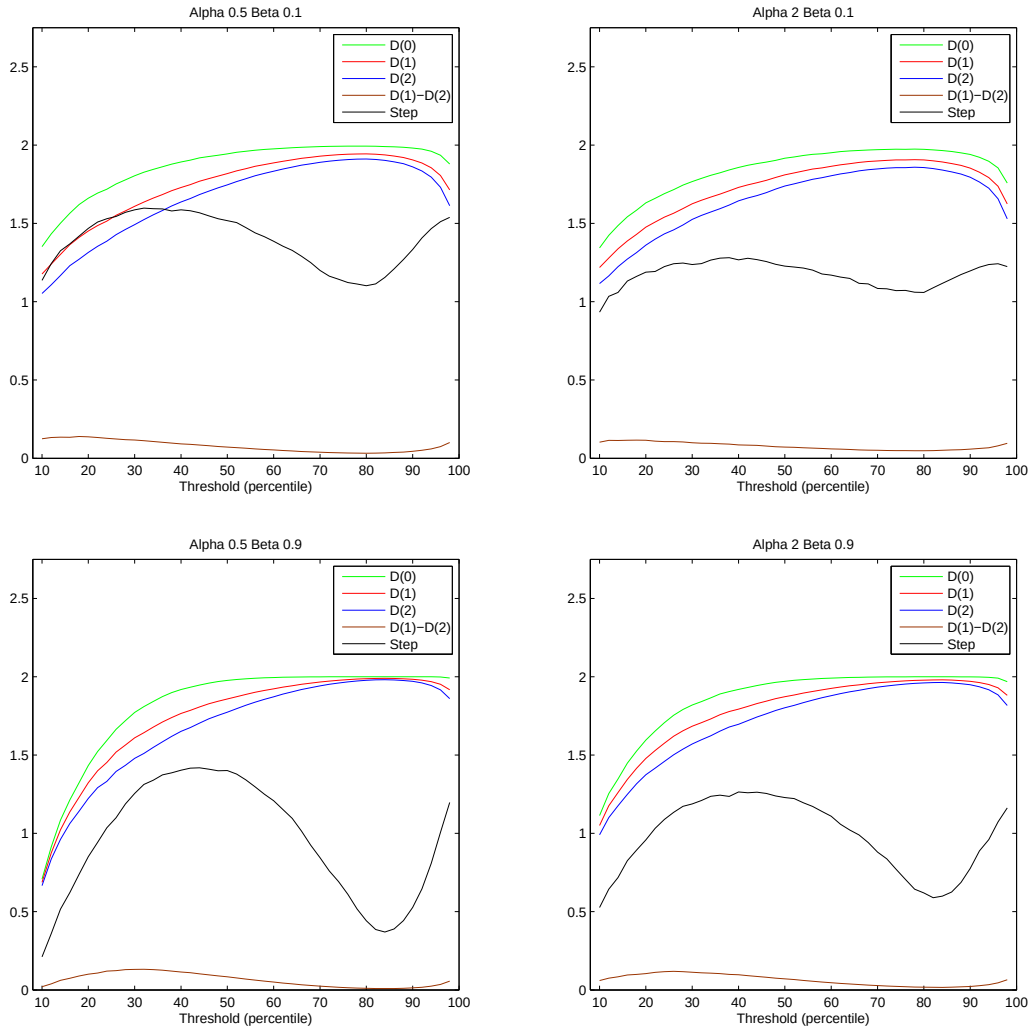

Figure 5.9: Values of D_0, D_1, D_2, $D_{-\infty} - D_{\infty}$ and $D_1 - D_2$ for the spatial point patterns defined by centroids for the four models

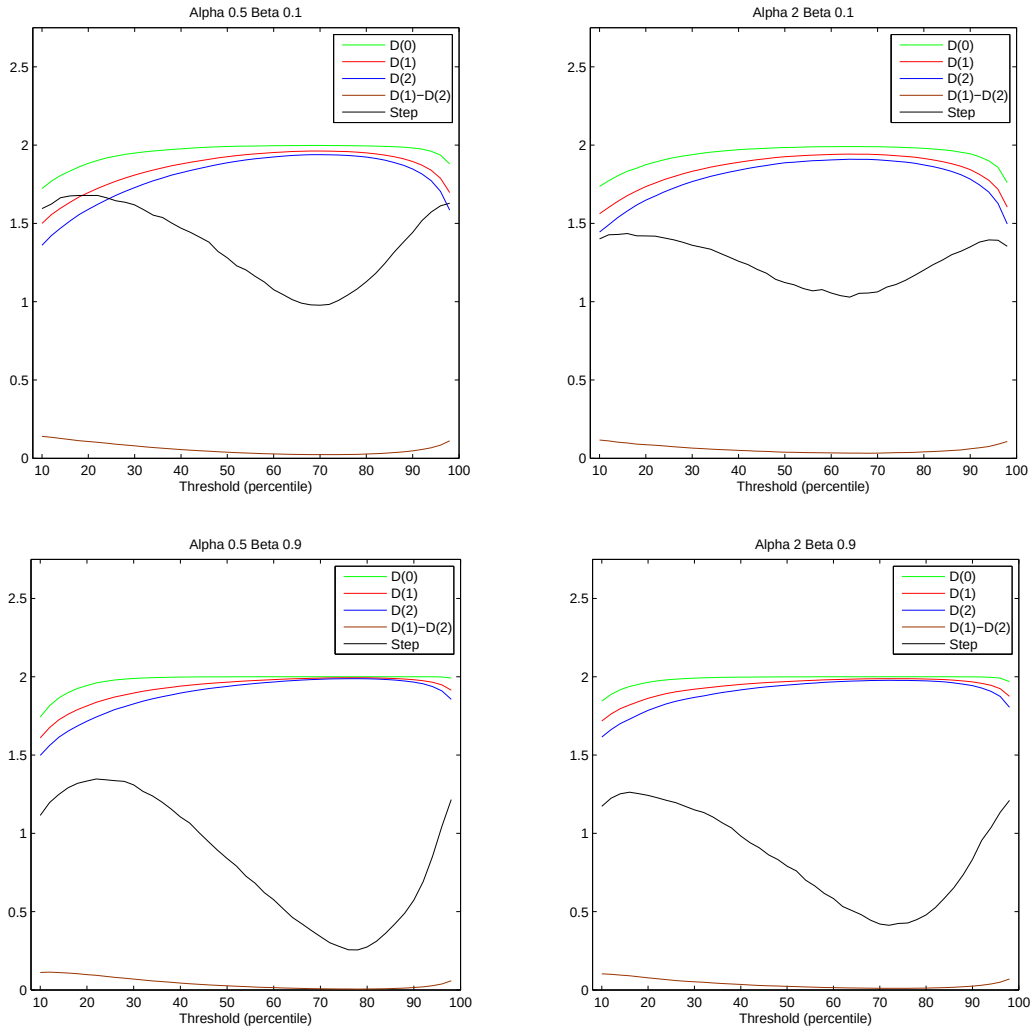

Figure 5.10: Values of D_0, D_1, D_2, $D_{-\infty} - D_{\infty}$ and $D_1 - D_2$ for the spatial point patterns defined by A-exit points for the four models

of different linear functionals in terms of the generalized dimensions such as derivatives. In particular, the first derivative of the generalized dimension curve can provide direct significant information about the system scaling complexity.

- Extensions of the dependence coefficient

  In Chapter 3, we formulate a dependence coefficient in the multifractal domain based on the deviation from the case of independence in the limiting behaviour of the Shannon mutual information (that is, in terms of 'entropy' dimension). In this sense, interesting aspects for continuing research are the formulation of related extended versions for dependence coefficients in the generalized dimension domain in relation to limiting behaviour of generalized forms of mutual information, as well as the analytical study of related properties. Thereby, some extensions of this proposal under similar arguments can be obtained for different values of the deformation parameter q. In this sense, in the case of the spatio-temporal interaction, considering different values of q the associated coefficient quantifies the degree of association between space and time independently of the scale emphasizing areas with low or high density of points.

  Furthermore, alternative versions can also be formulated by replacing the denominator with $D_q^{(X,Y,T)}$ or $\max\left\{D_q(X), D_q(Y)\right\}$ (see, for example, Furuichi 1967), or considering the limiting behaviour of Tsallis entropy instead of Rényi entropy.

- Extended analysis on the incorporation of effects

  The incorporation of the magnitude in the study of the three seismic series in Chapter 4 has provided relevant information regarding its effects. In this respect, an interesting extension consists of performing a sensitivity analysis on the choice of weighting function ϕ, regarding its influence on the different entropy and complexity related measures considered. Besides the magnitude, other characteristics might be of interest, such as potential information available from covariates.

- Incorporation of measurement errors and data extensions

  In practice, seismic registration involves various sources of errors. For instance, the determination of space-time coordinate is usually affected by the monitoring network configuration, the depth at which a specific event occurs, etc. Incorporation of this type of information, in particular through information measures, as well as the study of the seismicity of entropy measures and techniques derived to the presence of errors, are interesting with the aim of improving the accuracy in the assessment.

  Furthermore, such registered coordinates represent a space-time point of an event which in fact has a broader extent in space and time. Consideration of functional data or fuzzy data constitutes a another relevant direction for continuing research in this context.

- Extension of these techniques to study directionality aspects in space-time behaviour.

  The spatio-temporal structure of seismic event patterns usually shows directionality features related to the regional geodynamics underlying the phenomenon. The study of such characteristics at different scales and in space and time constitutes an interesting aspect for extended research, taking as a reference, in particular, existing related approaches in the analysis of point patterns.

- Assessment of the sensitivity of entropy measures by means of cluster analysis.

  Another interesting direction for research in the context of the application to seismic series is perform cluster analysis before studying the structural complexity with the aim of assessing the sensibility of entropy measures.

- Formal study of the application to excursion sets

  In the former section we have shown that the complexity and multifractalilty measures considered are useful tools to describe local heterogeneities in the structure of the point patterns associated to the excursions sets in random fields. This fact suggests the continuation of this research line in relation to analytical aspects, as well as by means of simulation studies.

Capítulo 5

Síntesis, Aspectos Complementarios y Líneas Abiertas

5.1. Síntesis

En este trabajo de tesis, abordamos varios aspectos relacionados con la sismicidad a través del estudio de los procesos puntuales asociados con este fenómeno geofísico. Dos aspectos principales de interés en este contexto son las propiedades de escalamiento y la interacción dimensional. Específicamente, este trabajo se centra en el uso estadístico de conceptos y medidas basados en la teoría de la información para analizar la complejidad estructural de patrones de distribución sísmicos.

Específicamente, en la Parte I, presentamos una visión preliminar de algunas de las técnicas más relevantes y que han tenido éxito en este contexto. En primer lugar (Capítulo 1), nos referimos a definiciones usuales de entropía y complejidad para distribuciones de probabilidad discretas así como a algunas de sus propiedades. En segundo lugar (Capítulo 2), introducimos las principales herramientas del análisis multifractal, esto es, las dimensiones generalizadas de Rényi y el espectro multifractal. Además, introducimos las principales leyes empíricas usadas para describir la sismicidad, así como las versiones no extensivas de la distribución frecuencia-magnitud de terremotos.

En la Parte II, proponemos algunas extensiones técnicas y metodológicas basadas en los anteriores elementos (Capítulo 3), y las aplicamos al estudio de tres series sísmicas (Capítulo 4). Específicamente, mediante el comportamiento límite de medidas de información y complejidad establecemos una medida de complejidad conectada con la medida de complejidad LMC, mostrando el uso potencial de funcionales relacionados para la evaluación cuantitativa de "complejidad multifractal", y también definimos un "coeficiente de dependencia"basado en las dimensiones generalizadas de Rényi. Adicionalmente, dado el creciente interés en la estadística no extensiva de Tsallis para describir la sismicidad, formulamos una versión de las dimensio-

nes generalizadas basada en la entropía de Tsallis. Además, proponemos la incorporación de los efectos de la magnitud a los diferentes análisis en términos de distribuciones ponderadas de conteo por cajas. Estas técnicas son aplicadas en combinación con varios enfoques complementarios tiles, a saber, diferentes transformaciones en la componente temporal, ventanas deslizantes y regresión no lineal ponderada. En el Capítulo 4, mostramos la implementación y aplicación de estos aspectos técnicos y metodológicos a través del estudio de tres series sísmicas de diferente naturaleza subyaccente. En el caso de la serie de Agrón, realizamos un análisis extendido de la interacción espacio-temporal, discriminando el comportamiento en diferentes subperiodos de actividad. Por otra parte, en el caso de la serie de El Hierro, que involucra un fenómeno volcánico, los resultados obtenidos muestran cambios estructurales en la evolución temporal, los cuales pueden ser interpretados como un comportamiento precursor a la erupción volcánica. En cuanto a la serie de Torreperogil, las inesperadas características estructurales observadas podrían indicar la posible existencia de factores externos.

En conclusión, los resultados obtenidos en las aplicaciones a las series sísmicas muestran la utilidad de las técnicas propuestas para describir la sismicidad. En particular, la medida de complejidad propuesta proporciona una herramienta apropiada para cuantificar la complejidad en el dominio multifractal, en conexión con la medida de complejidad LMC. En cuanto a la interacción dimensional, el análisis realizado indica que el coeficiente de dependencia, formulado en este trabajo en términos de la dimensión de “entropía, es una herramienta útil para estudiar el grado de asociación entre componentes en procesos puntuales con características multifractales. Con respecto a la nueva formulación de dimensiones generalizadas basada en la entropía de Tsallis, debemos señalar que está ha ofrecido resultados satisfactorios, si bien se requiere un estudio posterior sobre sus aspectos analíticos. Además, la combinación de estas técnicas con transformaciones temporales proporciona información extra en el estudio de la interacción espacio-temporal. Adicionalmente, la incorporación de los efectos de la magnitud en los diferentes análisis, ha sugerido una asociación estructural intrínseca entre las magnitudes y las coordenadas espacio-temporales. Por último, los resultados obtenidos en el estudio de la sismicidad mediante la distribución no extensiva frecuencia-magnitud muestran que las estimaciones son mejoradas en todos los casos cuando se considera una regresión no linear ponderada para realizar el ajuste.

Algunos de los desarrollos y aplicaciones derivados de este trabajo de tesis están recogidos en la siguientes publicaciones:

- Angulo, J. M., Esquivel, F. J., Structural complexity in space-time seismic event data. *Stochastic Environmental Research and Risk Assessment*, DOI 10.1007/s00477-013-0807-x, 2013.

- Angulo, J. M., Esquivel, F. J., Statistical complexity analysis of spatiotemporal dynamics. *Mathematics of Planet Earth. Proceedings of the 15th Annual Conference of the International Association for Mathematical Geosciences (IAMG 2013)*. Lecture Notes in Earth

System Sciences, ISBN 978-3-642-32407-9, Pardo-Igúzquiza, E., Guardiola-Albert, C., Heredia, J., Moreno-Merino, L., Durán, J. J., Vargas-Guzmán, J. A. (Eds.), Springer, 185-188, 2013.

Otros aspectos son incorporados como parte de artículos en fase de preparación, acerca de:

- El análisis de la serie sísmica de El Hierro mediante técnicas de complejidad y multifractalidad y la distribución ponderada no extensiva frecuencia-magnitud.
- Discriminación de efectos en relación al comportamiento sísmico inesperado en la región de Torreperogil.
- Aplicación de las técnicas propuestas en esta tesis al estudio de propiedades estructurales de los conjuntos de excursión en campos aleatorios. (En la siguiente sección se describe una breve exposición de esta idea y un estudio preliminar realizado.)

Adicionalmente, los contenidos de esta investigación han sido presentados en diferentes reuniones científicas. Como ponencias invitadas:

- Angulo J. M., Esquivel, F. J., Structural complexity in space-time seismic event data. *VI International Workshop on Spatio-Temporal Modelling (METMAVI)*, Guimãraes, Portugal, 12-14 de septiembre de 2012.
- Angulo J. M., Esquivel, F. J., Information-related approaches for statistical complexity assessment. *The 25th Annual International Conference on Statistics and Modelling in Human and Social Sciences*, El Cairo, Egipto, 25-28 de marzo de 2013.
- Angulo, J. M., Esquivel, F. J., Statistical complexity analysis of spatiotemporal dynamics. *The 15th Annual Conference of the International Association for Mathematical Geosciences (IAMG 2013)*, Madrid, España, 2-6 de septiembre de 2013.

Como comunicaciones:

- Esquivel, F. J., Angulo J. M., Estudio del efecto de la dinámica temporal en la estructura de la situación geográfica de los epicentros de una secuencia sísmica. *XXXIII Congreso Nacional de Estadística e Investigación Operativa (SEIO 2012)*, Madrid, España, 17-20 de abril de 2012.
- Esquivel, F. J., Madrid, A. E., Angulo, J. M., Multifractal characteristics of spatial threshold exceedances. *5th International Conference of the ERCIM Working Group on Computing & Statistics (ERCIM 2012)*, Oviedo, España, 1-3 de diciembre de 2012.

- Esquivel, F. J., Angulo J. M., Some aspects of structural complexity dynamics in seismic data of Torreperogil. *XXXIV Congreso Nacional de Estadística e Investigación Operativa (SEIO 2013)*, Castellón, España, 11-13 de septiembre de 2013.

En la modalidad de póster:

- Angulo, J. M., Esquivel, F. J., Un estudio sobre dimensiones fractales generalizadas en procesos puntuales espaciales. *XXXII Congreso Nacional de Estadística e Investigación Operativa (SEIO 2010)*, A Coruña, España, 14-17 de septiembre de 2010.
- Esquivel, F. J., Angulo J. M., Information loss under projections from seismic data. *58th World Statistics Congress of the International Statistical Institute (ISI 2011)*, Dublín, Irlanda, 21-26 de agosto de 2011.

5.2. Aspectos Complementarios

La posible relación entre procesos estocásticos y la teoría fractal puede ser una herramienta útil para describir algunas propiedades de un proceso mediante medidas fractales. En este contexto, Orey (1970) probó que la dimensión de Hausdorff puede detectar importantes propiedades tales como estacionariedad y ciertas propiedades inherentes a un proceso Gaussiano con incrementos estacionarios, en el dominio temporal. Adicionalmente, Madrid et al. (2012) mostraron que el análisis de patrones puntuales espaciales definidos a partir de conjuntos de excursión en términos de los centroides de las componentes conexas y de los *A-exit points* proporciona información sobre caractaerísticas relevantes del campo aleatorio subyacente. Siguiendo esta idea, hemos considerado en el contexto de este trabajo de tesis el estudio de características estructurales de los conjuntos de excursión mediante la aplicación de herramientas de complejidad y de multifractalidad a dichos patrones puntuales.

Con este propósito, hemos considerado realizaciones de campos aleatorios Gaussianos generadas usando la clase de Cauchy (Gneiting y Schlater (2004)), definida por la función de covarianza homogénea e isotrópica

$$C(h) = \sigma^2(1+|h|\alpha)^{\beta/\alpha}, \tag{5.1}$$

donde σ^2 representa la varianza del correspondiente campo aleatorio, y $\alpha \in (0,2]$ y $\beta \geq 0$. Esta clase contiene como caso particular el "modelo de Cauchy", para $\alpha = 2$. El parámetro α está asociado con la variabilidad local y determina la dimension fractal de las realizaciones, $D = n+1-\alpha/2$. Independientemente, el parámetro β está relacionado con el rango de dependencia, y para $\beta \in (0,1)$ el proceso tiene memoria larga con coeficiente de Hurst $H = 1-\beta/2$. La Figura 5.1 muestra cuatro realizaciones simuladas sobre el cuadro $[1,200]^2$ (usando los mismos números pseudoaleatorios) basadas en las combinaciones cruzadas de los valores de los

parámetros $\alpha = 0{,}5, 2$ y $\beta = 0{,}1, 0{,}9$, con $\sigma = 0{,}1$ común. Las imágenes muestran claras diferencias en relación con la variación local y la estructuración a gran escala.

En este estudio, consideramos los conjuntos de excedencias de umbrales a partir de de 50 realizaciones para cada modelo, tomando umbrales correspondientes al rango de percentiles entre 10 y 98. Las Figuras 5.2, 5.3, 5.4 y 5.5 muestran los conjuntos de excursión para cada una de las realizaciones de la figura 5.1 considerando los cuartiles empíricos como umbrales.

Los patrones puntuales espaciales analizados están determinados por centroides de las componentes conexas y por los *A-exit points* para cada conjunto de excedencia de umbral:

- Caso de "Centroides": Tomamos cada componente conexa como un evento aislado, representado por las coordenadas de su centroide (ver, como ejemplo, Madrid et al., 2012).
- Caso de "*A-exit points*": Los *A-exit points* se definen, en un sentido formal (ver Piterbarg, 1996), como puntos salientes en el conjunto de excedencia de umbral respecto a un punto focal de referencia en el infinito, reflejando así la rugosidad de su frontera entre otras características (ver, como ejemplo, Piterbarg, 1996).

Las Figuras 5.6 y 5.7 muestran los conjuntos de excursión para el umbral correspondiente al percentil 90 sobre la base de las realizaciones mostradas en la Figura 5.1, indicando los puntos los centroides y los *A-exit points*, respectivamente (aquí, un único punto local a la izquierda). Podemos observar cómo por columnas, de izquierda a derecha, el número de componentes conexas disminuye mientras que su tamaño aumenta. Esto es debido al hecho de que cuando el valor del parámetro α aumenta el grado de fractalidad disminuye dando lugar a una menor fragmentación y rugosidad. Por otro lado, cuando el valor de β aumenta, gráficas de arriba a abajo, los rangos de dependencia se reducen y, por tanto, el nivel de desagregación aumenta, con una menor heterogeneidad en los tamaños de las componentes conexas.

En base a los patrones puntuales derivados de las 50 realizaciones de cada modelo, y considerando los umbrales correspondientes a los percentiles entre 10 y 98, estudiamos algunas características estructurales. En primer lugar, evaluamos la asociación espacial $X \leftrightarrow Y$ mediante el coeficiente de dependencia introducido en el Capítulo 3. La Figura 5.8 muestra los valores promedio de los coeficientes de dependencia correspondientes a las 50 realizaciones variando el umbral considerado, para los patrones puntuales definidos por los centroides (gráfico de la izquierda) y por los *A-exit points* (gráfico de la derecha). Las representaciones muestran que el grado de asociación espacial es sensible a la variación del umbral considerado, así como a las especificaciones de los parámetros en el modelo subyacente, con claras diferencias entre los casos de los centroides y *A-exit points*.

Para realizar el estudio de las características estructurales a diferentes escalas, calculamos los coeficientes D_0, D_1, D_2, el salto multifractal $D_{-\infty} - D_{\infty}$ y la medida de complejidad $D_1 - D_2$. La Figura 5.9 muestra los resultados obtenidos considerando los patrones puntuales definidos por centroides para los cuatro modelos. En esta figura, podemos ver, entre otros aspectos, que

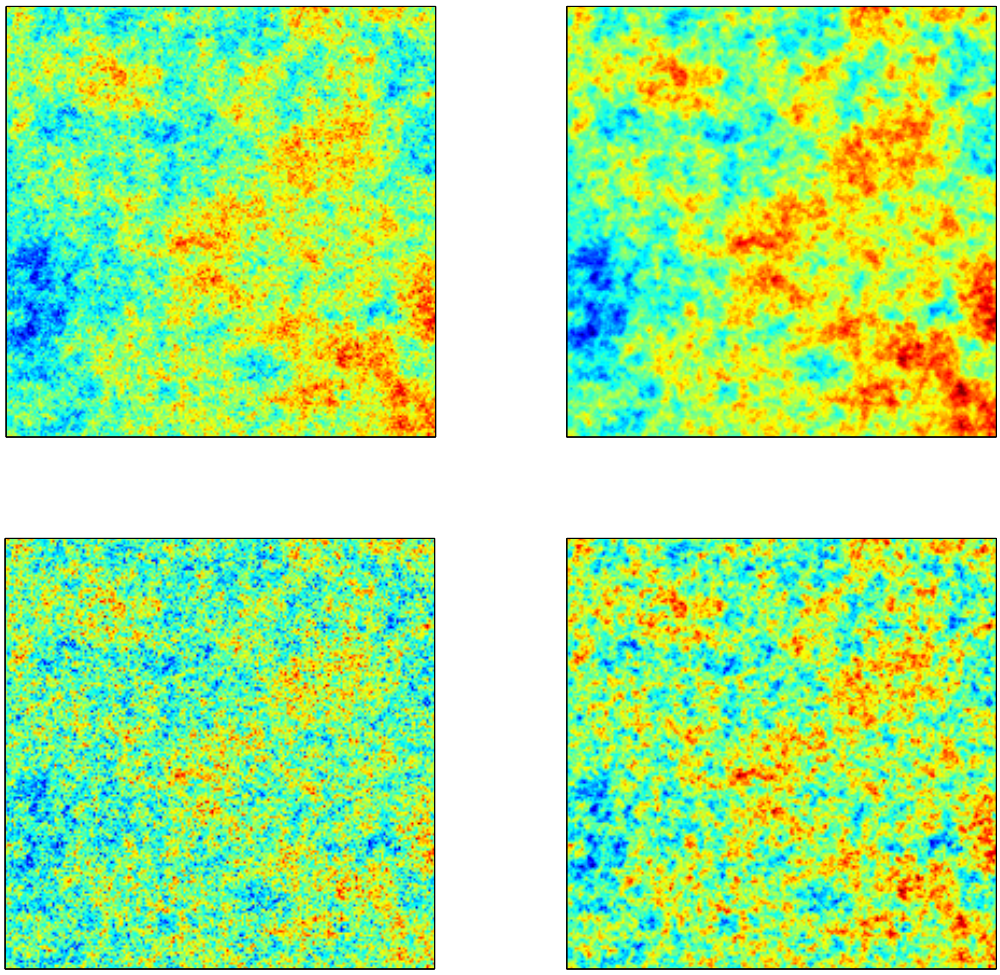

Figura 5.1: Realizaciones simuladas mediante la clase de Cauchy con $\alpha = 0{,}5, 2$ (de izquierda a derecha) y $\beta = 0{,}1, 0{,}9$ (de arriba a abajo).

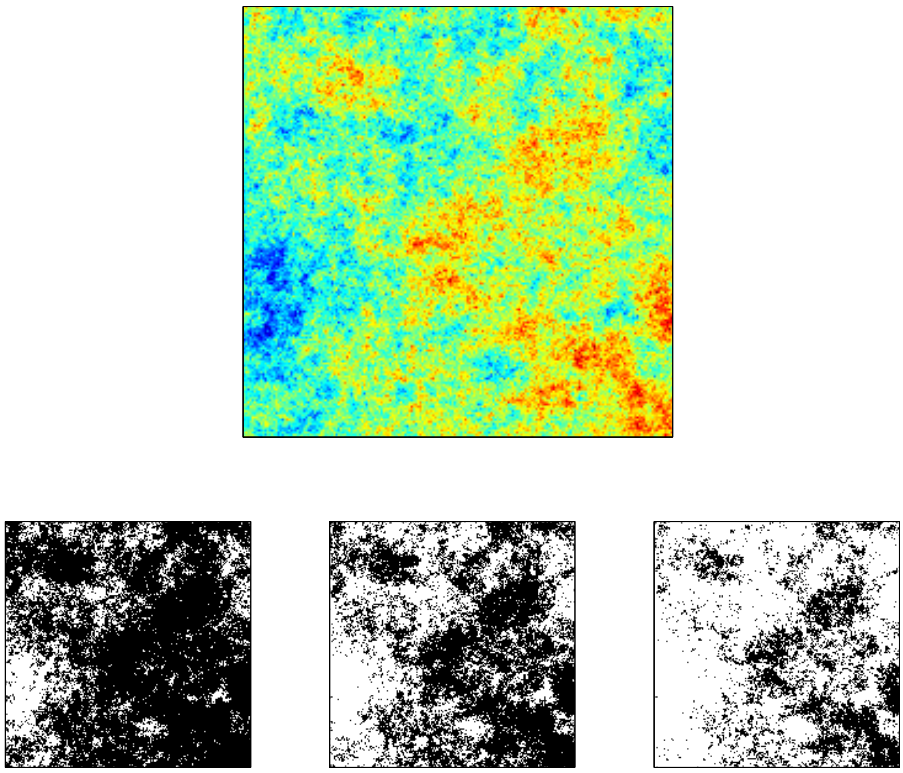

Figura 5.2: Campo aleatorio (arriba) y conjuntos de excursión considerando como umbrales los cuartiles (abajo de izquierda a derecha), basados en la realización de la Figura 5.1 para $\alpha = 0{,}5$ y $\beta = 0{,}1$

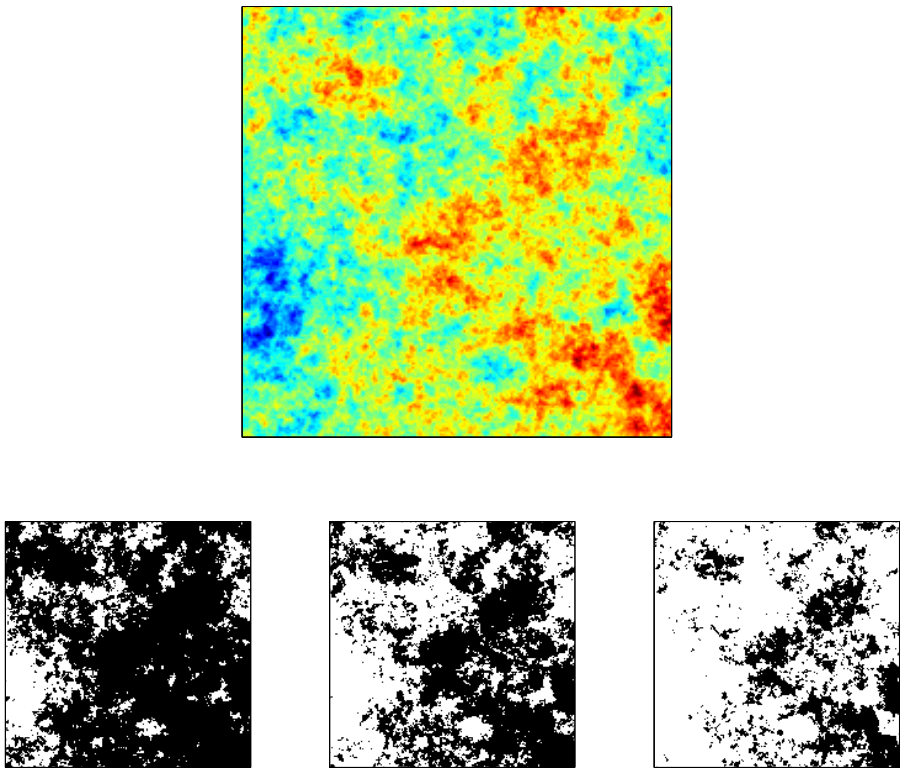

Figura 5.3: Campo aleatorio (arriba) y conjuntos de excursión considerando como umbrales los cuartiles (abajo de izquierda a derecha), basados en la realización de la Figura 5.1 para $\alpha = 2$ y $\beta = 0{,}1$

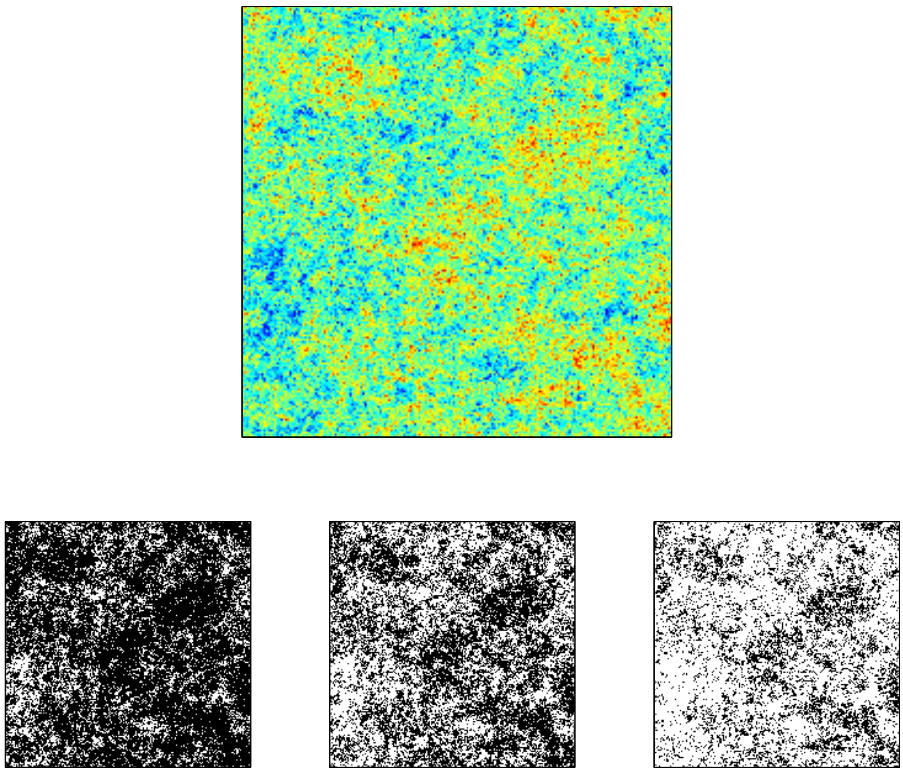

Figura 5.4: Campo aleatorio (arriba) y conjuntos de excursión considerando como umbrales los cuartiles (abajo de izquierda a derecha), basados en la realización de la Figura 5.1 para $\alpha = 0{,}5$ y $\beta = 0{,}9$

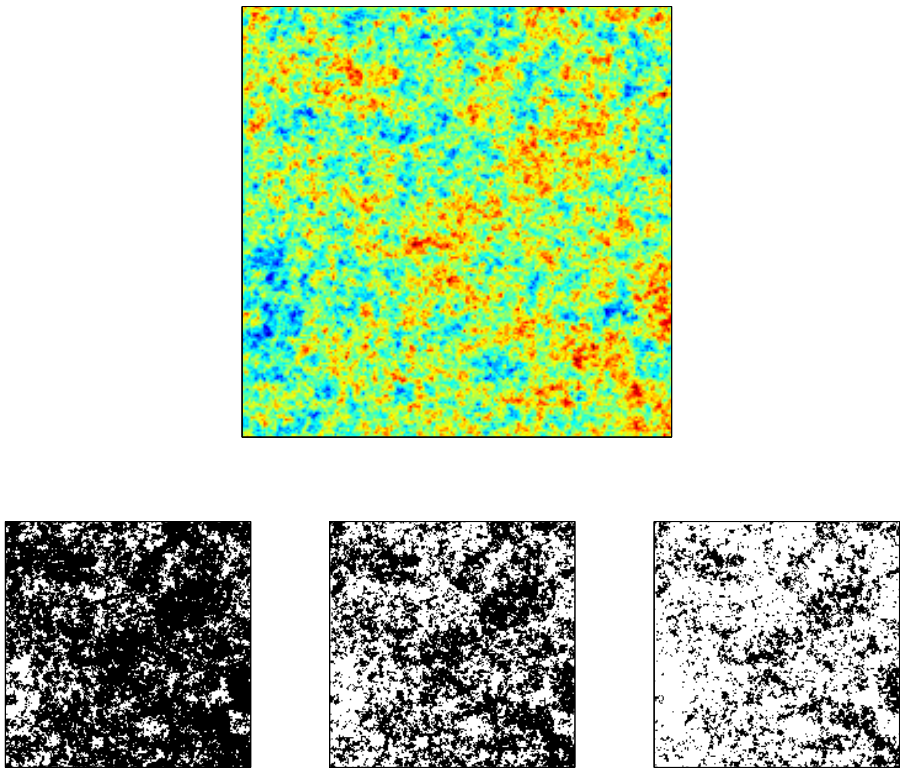

Figura 5.5: Campo aleatorio (arriba) y conjuntos de excursión considerando como umbrales los cuartiles (abajo de izquierda a derecha), basados en la realización de la Figura 5.1 para $\alpha = 2$ y $\beta = 0{,}9$

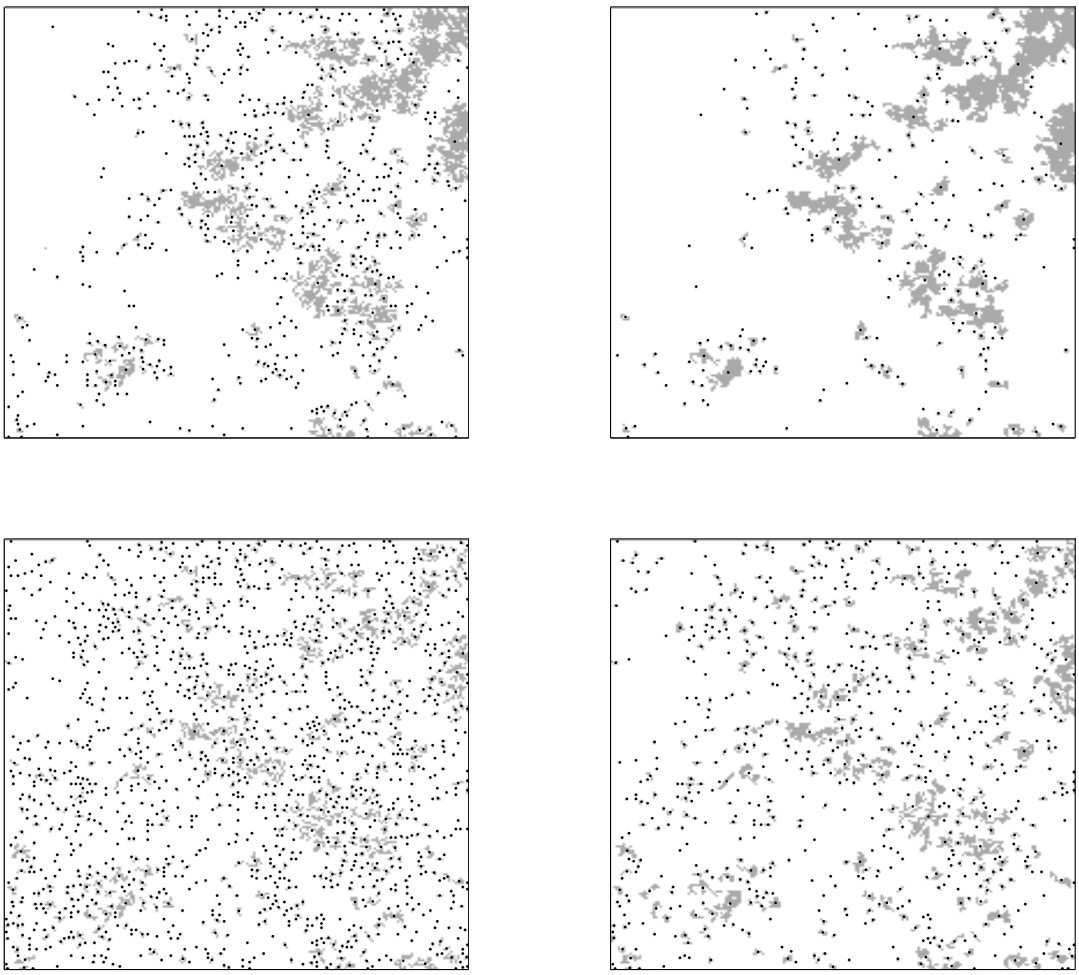

Figura 5.6: Conjuntos de excursión considerando como umbral el percentil 90th, basados en las realizaciones de la Figura 5.1. Los puntos negros representan los centroides de las componentes conexas.

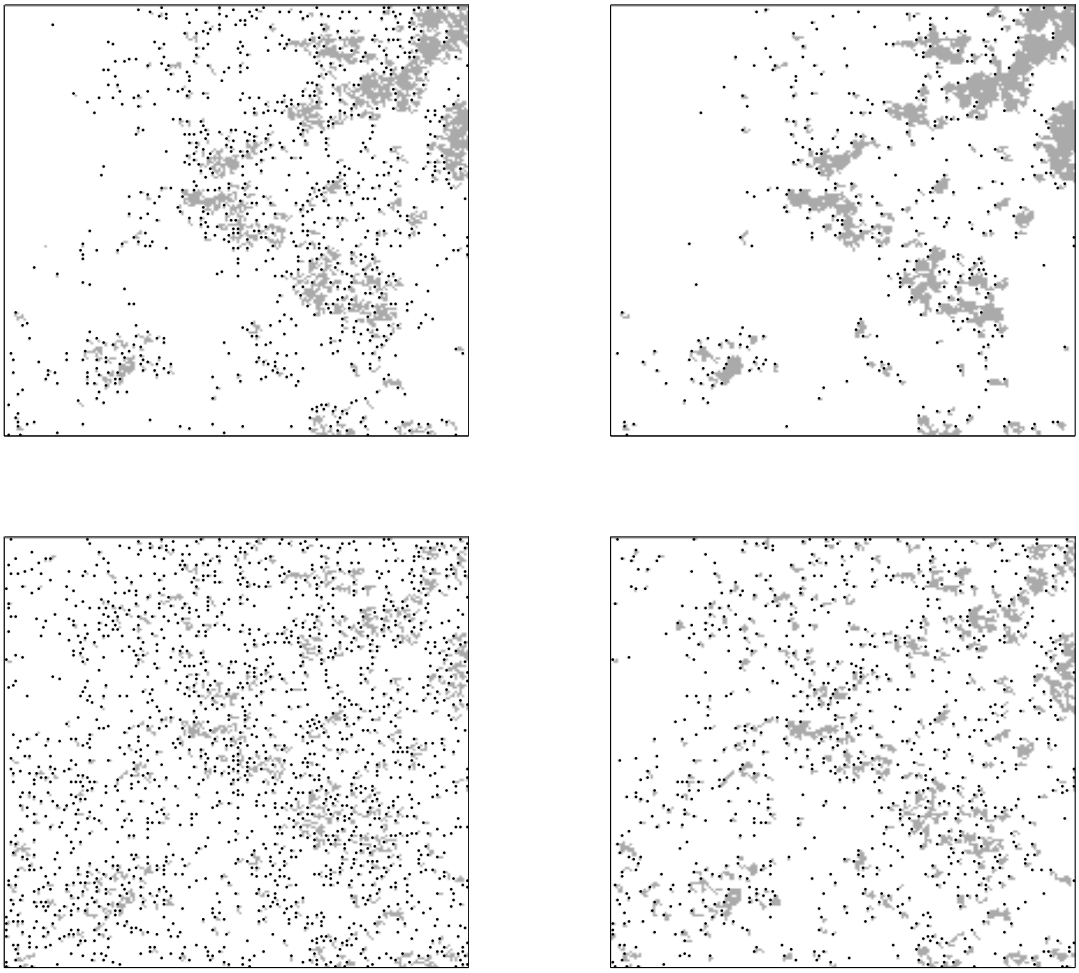

Figura 5.7: Conjuntos de excursión considerando como umbral el percentil 90th, basados en las realizaciones de la Figura 5.1. Los puntos negros representan los *A-exit points*.

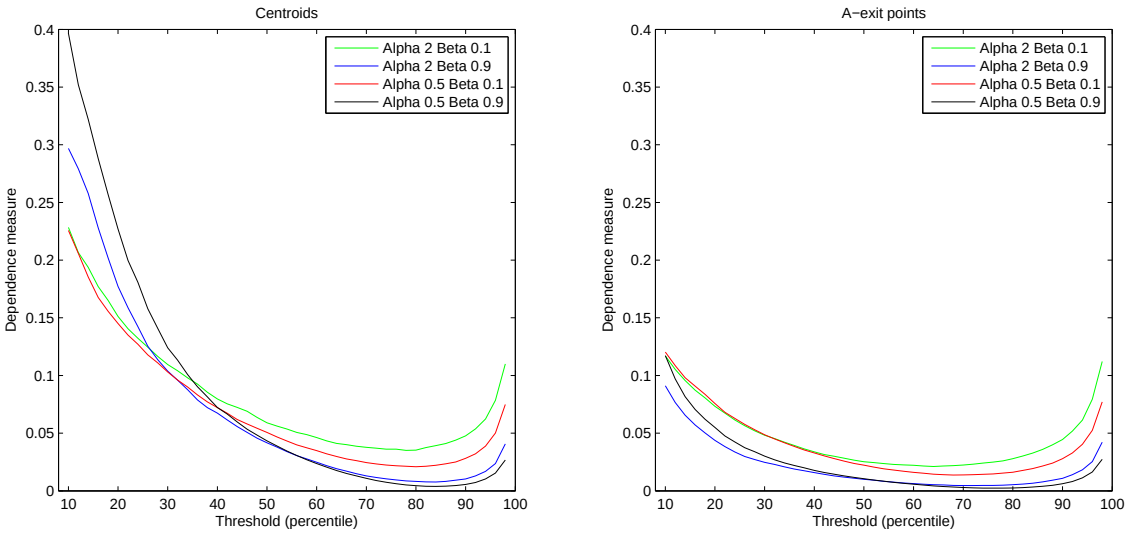

Figura 5.8: Valores del coeficiente de dependencia para los patrones puntuales espaciales definidos por los centroides (izquierda) y por los *A-exit points* (derecha) para los cuatro modelos

los cambios en el comportamiento multifractal son más sensibles al incremento en el valor del umbral para un α bajo y un β alto. Los correspondientes resultados obtenidos cuando consideramos los patrones puntuales espaciales definidos por los *A-exit points* se muestran en la Figura 5.10. Interpretaciones similares pueden derivarse, si bien obstante hay diferencias visibles en la localización de los cambios de tendencia.

De la aplicación preliminar realizada, podemos concluir que las herramientas de multifractalidad y complejidad consideradas pueden se útiles para la evaluación cuantitativa de características estructurales asociadas a los conjuntos de excedencias de umbrales espaciales (Esquivel y Angulo, 2012); en particular, para detección de rasgos diferenciales dependiendo de las propiedades de campo aleatorio subyacente y del umbral.

5.3. Líneas Abiertas

El desarrollo de este trabajo de tesis ha abierto varias líneas y direcciones adicionales para la continuación de investigaciones relacionadas, enfocadas sobre diferentes aspectos, incluyendo:

- Estudio analítico de las dimensiones generalizadas basadas en la entropía de Tsallis

  En sismicidad, los hechos de que el proceso de empuje-deslizamiento entre placas tectónicas es no aditivo (Sotolongo-Costa, 2012) y que la estadística no extensiva permite inter-

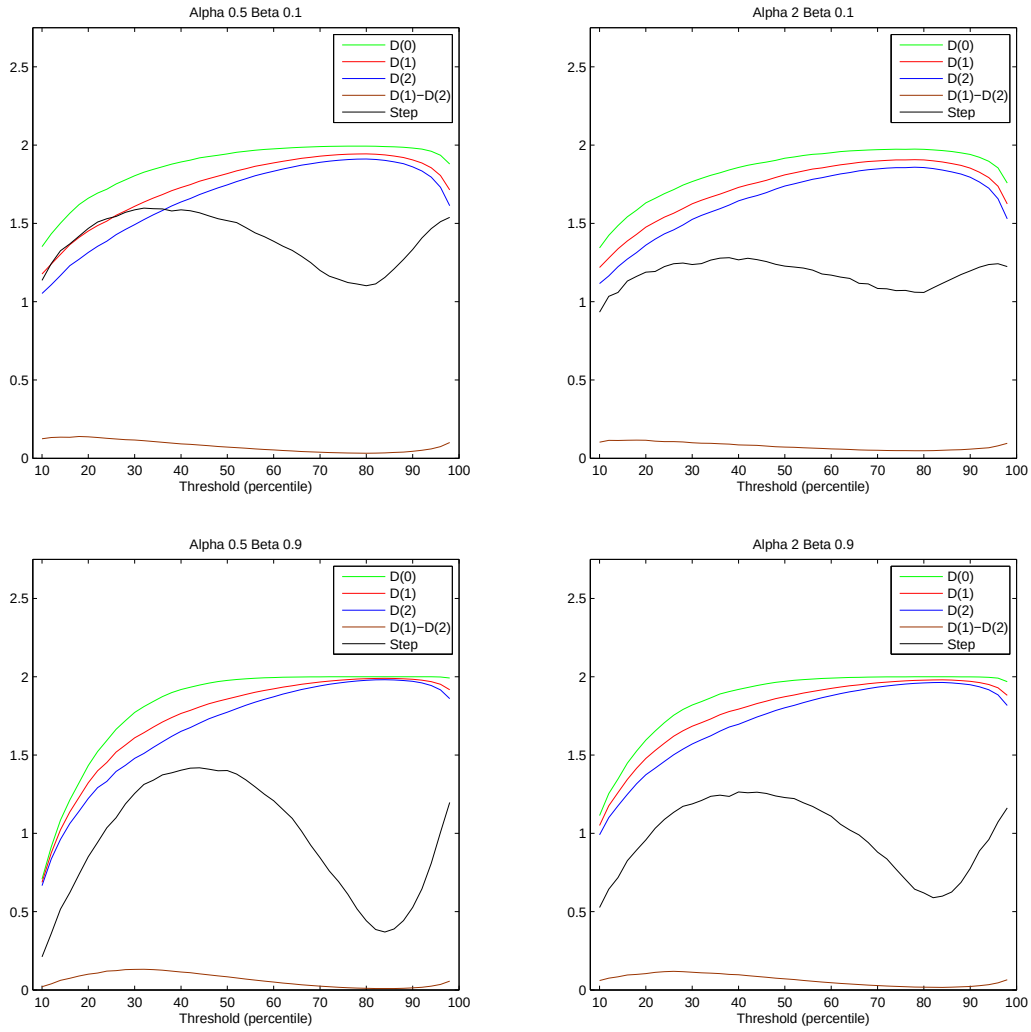

Figura 5.9: Valores de D_0, D_1, D_2, $D_{-\infty} - D_{\infty}$ y $D_1 - D_2$ para los patrones puntuales definidos por los centroides para los cuatro modelos

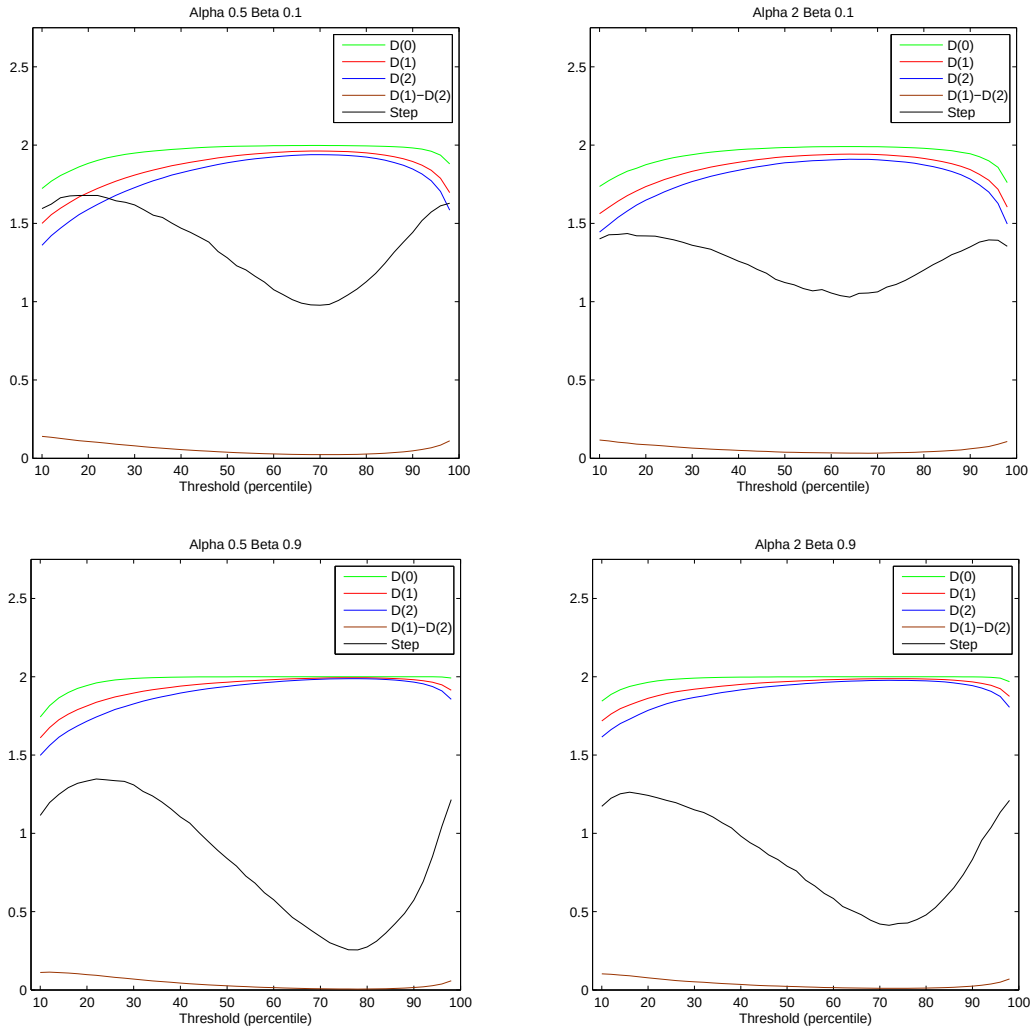

Figura 5.10: Valores de D_0, D_1, D_2, $D_{-\infty} - D_{\infty}$ y $D_1 - D_2$ para los patrones puntuales definidos por los *A-exit points* para los cuatro modelos

acciones en diferentes órdenes nos han conducido a introducir una nueva formulación de dimensiones generalizadas basadas en la entropía de Tsallis (Capítulo 3). En este contexto, estamos interesados en realizar un estudio analítico de las propiedades relacionadas, y en derivar la formulación de una definición apropiada de un espectro multifractal asociado.

- Formulación de nuevas medidas de complejidad en el dominio multifractal

  En el Capítulo 3 proponemos una medida generalizada de complejidad en el dominio multifractal basada en el comportamiento límite de la entropía de Rényi, y en conexión con la medida generalizada de complejidad $\tilde{C}^{\alpha,\beta}$ (López-Ruiz et al., 2009). En este contexto, con el objetivo de explorar la información variacional de la curva de dimensiones generalizadas para la evaluación discriminatoria de características de complejidad multifractal, consideramos el estudio de diferentes funcionales lineales en términos de las dimensiones generalizadas tales como las derivadas. En particular, la derivada primera de la curva de dimensiones generalizadas puede proporcionar información directa significativa acerca de la complejidad en el escalamiento del sistema de escalamiento.

- Extensiones del coeficiente de dependencia

  En el Capítulo 3 formulamos un coeficiente de dependencia en el dominio multifractal basado en la desviación con respecto al caso de independencia en el comportamiento límite de la información mutua de Shannon (esto es, en términos de la dimensión de "entropía"). En este sentido, la formulación de versiones extendidas relacionadas para coeficientes de dependencia en el dominio de las dimensiones generalizadas en relación con el comportamiento límite de formas generalizadas de información mutua, así como el estudio analítico de propiedades relacionadas, son aspectos interesantes para la continuación de la investigación. De este modo, algunas extensiones de esta propuesta bajo argumentos similares pueden obtenerse para valores diferentes del parámetro de deformación q. En este sentido, en el caso de la interacción espacio-temporal, considerando diferentes valores de q el coeficiente asociado cuantifica el grado de asociación entre espacio y tiempo independientemente de la escala enfatizando areas con baja o alta densidad de puntos.

  Adicionalmente, versiones alternativas pueden formularse también reemplazando el denominador por $D_q^{(X,Y,T)}$ o máx $\left\{D_q(X), D_q(Y)\right\}$ (ver, por ejemplo, Furuichi 1967), o considerando el comportamiento límite de la entropía de Tsallis en lugar de la entropía de Rényi.

- Análisis extendido sobre la incorporación de efectos

  La incorporación de de la magnitud en el estudio de las tres serie sísmicas realizado en el Capítulo 4 ha proporcionado información relevante en relación con sus efectos. En este sentido, una extensión interesante consiste en realizar un análisis de sensibilidad sobre

la elección de la función de ponderación ϕ, referente a su influencia sobre las diferentes medidas de entropía y medidas relacionadas de complejidad consideradas. Además de la magnitud, otras características podrían ser de interés, tales como la información potencial disponible de covariables.

- Incorporación de errores de medición y extensiones obre los datos

  En la práctica los registros sísmicos involucran diversas fuentes de errores. Por ejemplo, la determinación de las coordenadas espacio-temporales usualmente está afectado por la configuración de la red d monitorización, la profundidad a la que ocurre un evento específico, etc. La incorporación de este tipo de información, er particular a través de medidas de información, así como el estudio de la sensibilidad de medidas de entropía y técnicas derivadas con respecto a la presencia de errores, son interesantes con el fin de mejorar la precisión en la evaluación.

  Adicionalmente, tales coordenadas registradas representan un punto espacio-temporal de un suceso que en realidad tiene una extensión más amplia en el espacio y en el tiempo. La consideración de datos funcionales o datos difusos constituye otra dirección relevante para la continuación de la investigación en este contexto.

- Extensión de estas técnicas para estudiar aspectos de direccionalidad en el comportamiento espacio-temporal.

  La estructura espacio-temporal de patrones de eventos sísmicos usualmente muestra rasgos de direccionalidad relacionados con la geodinámica regional subyacente al fenómeno. El estudio de tales características en diferentes escalas y en el espacio y tiempo constituye un aspecto interesante para la extensión de la investigación, tomando como referencia, en particular, enfoques relacionados existentes en el análisis de patrones puntuales.

- Estudio formal de la aplicación a conjuntos de excursión.

  En la sección anterior hemos mostrado que las medidas de complejidad y multifractalidad consideradas son herramientas útiles para describir heterogeneidades locales en la estructura de los patrones puntuales asociados a los conjuntos de excursión en campos aleatorios. Este hecho sugiere la continuación de esta línea de investigación en relación con aspectos analíticos, así como mediante estudios de simulación.

Bibliography

[1] Angulo J. M., Esquivel, F. J., Un estudio sobre dimensiones fractales generalizadas en procesos puntuales espaciales. *Proceedings of the XXXII Congreso Nacional de Estadística e Investigación Operativa y VI Jornadas de Estadística Pública (SEIO 2010)*, A Coruña, Spain, 14-17 September. Costa, J., Fernández, E., Presedo, M. A., Vilar, J. N. (Eds.), ISBN: 978-84-693-6152-8

[2] Angulo, J. M., Esquivel, F. J., Structural complexity in space-time seismic event data. *Stochastic Environmental Research and Risk Assessment*, DOI 10.1007/s00477-013-0807-x (2013)

[3] Angulo, J. M., Esquivel, F. J., Statistical complexity analysis of spatiotemporal dynamics. *Mathematics of Planet Earth. Proceedings of the 15th Annual Conference of the International Association for Mathematical Geosciences (IAMG 2013)*. Lecture Notes in Earth System Sciences, ISBN 978-3-642-32407-9. Pardo-Igúzquiza, E., Guardiola-Albert, C., Heredia, J., Moreno-Merino, L., Durán, J. J., Vargas-Guzmán, J. A. (Eds.), Springer, 185-188 (2013a)

[4] Anteneodo, C., Plastino, A. R., Some features of the López-Ruiz-Mancini-Calbet (LMC) statistical measure of complexity. *Physics Letters A*, **223**, 348-354 (1996)

[5] Assunção, R., Maia, A., A note on testing separability in spatial-temporal marked point processes. *Biometrics*, **63**, 290-295 (2007)

[6] Bacry, E., Muzy, J. F., Arnéodo, A., Singularity spectrum of fractal signals from wavelet analysis: exact results. *Journal of Statistical Physics*, **70**, 635-674 (1993)

[7] Bak, P., Tang, C., Weisenfeld, K., Selforganized criticality: an explanation of 1/f noise. *Physical Review Letters*, **59**, 381-384 (1987)

[8] Bak, P., Tang, C., Earthquakes as selforganized critical phenomena. *Journal of Geophysical Research*, **949**, 15635-15637 (1989)

[9] Båth, M., Bagci, G. B., Lateral inhomogeneities of the upper mantle. *Tectonophysics*. **2**, 483-514 (1965)

[10] Bouvrie, P. A., Angulo, J. C., Antolín, J., A generalized relative complexity: application to atomic one-particle densities. *Chemical Physics Letters*, **539**, 191-196 (2012)

[11] Caneva, A., Smirnov, V., Using the fractal dimension of earthquake distributions and slope of the reccurence curve to forecast earthquakes in Colombia. *Earth Sciences Research Journal*, **8**, 3-9, (2004)

[12] Cátalan, R. G., Garay, J., López-Ruiz, R., Features of the extension of a statistical measure of complexity for continuous systems. *Physical Review E*, **66**, 11102 (2002)

[13] Celikoglu, A., Tirnakli, U., Earthquakes, model systems and connections to *q*-statistics. *Acta Geophysica*, **60**, 535-546 (2012)

[14] Chen, S., Sun, L., Ma, L., Multifractal characteristics of spatiotemporal distribution of earthquakes in intraplate and interplate regions. *Acta Seismologica Sinica*, **20**, 40-50 (2007)

[15] Chelidze, T., Matcharashvili, T., Complexity of seismic process; measuring and applications– a review. *Tectonophysics*, **431**, 49-60 (2007)

[16] Chhabra, A. B., Jensen, R. V., Direct determination of the $f(\alpha)$ singularity spectrum. *Physical Review Letters*, **62**, 1327-1330 (1989)

[17] Chhabra, A. B., Meneveau, C., Jensen, R. V., Sreenivasan, K. R., Direct determination of the $f(\alpha)$ singularity spectrum and its application to fully developed turbulence. *Physical Review A*, **40**, 5284–5294 (1989)

[18] Cuevas, E., $f(\alpha)$ multifractal spectrum at strong and weak disorder. *Physical Review B*, **68**, 24206–24212 (2003)

[19] Darooneh, A. H., Dadashinia, C., Qamili, E., Frepoli, A., Analysis of the spatial and temporal distributions between successive earthquakes: nonextensive statistical mechanics viewpoint. *Physica A*, **387**, 3647-3654 (2008)

[20] Darooneh, A. H., Mehri, A., A nonextensive modification of the Gutenberg-Richter law: q-stretched exponential form. *Physica A*,**389**, 509-514 (2010)

[21] De Santis, A., Cianchini, G., Qamili, E., Frepoli, A., The 2009 L'Aquila (Central Italy) seismic sequence as a chaotic process. *Tectonophysics*, **496**, 44-52 (2010)

[22] De Santis, A., Cianchini, G., Favali, P., Beranzoli, L., Boschi, E., The Gutenberg–Richter law and entropy of earthquakes: two case studies in central Italy. *Bulletin of the Seismological Society of America*, **101**, 1386-1395 (2011)

[23] Díaz-Avalos, C., Juan, P., Mateu, J., Similarity measures of conditional intensity functions to test separability in multidimensional point processes. *Stochastic Environmental Research and Risk Assessment*, **27**, 1193-1205 (2013)

[24] Diggle, P., *Statistical Analysis of Spatial Point Patterns*. Academic Press, London (1983)

[25] Doblas, M., Aretxabala, A., Torres, Y., Benito, B., The 2012/2013 seismic crisis of Jaén, Spain: a working hypothesis involving hydroseismicity. *Proceedings of the Meeting of the Americas*, Cancún, Mexico, 14-17 May 2013

[26] Eneva, M., Effect of limited data sets in evaluating the scaling properties of spatially distributed data: an example from mining-induced seismic activity. *Geophys. J. Int*, **124**, 773-786 (1996)

[27] Esquivel, F. J., Angulo, J. M., Information loss under projections from seismic data. *Proceedings of the 2011 World Statistics Congress (ISI 2011)*, Dublin, Ireland, 12-20 August 2011

[28] Esquivel, F. J., Angulo, J. M., Estudio del efecto de la dinámica temporal en la estructura de la situación geográfica de los epicentros de una secuencia sísmica. *Proceedings of the XXXIII Congreso Nacional de Estadística e Investigación Operativa e VII Jornadas de Estadística Pública (SEIO 2012)*, Madrid, Spain, 21-26 April 2012

[29] Esquivel, F. J., Madrid, A. E., Angulo, J. M., Multifractal characteristics of spatial threshold exeedances. *Proceedings of the 5th International Conference of the ERCIM Working Group on Computing & Statistics (ERCIM 2012)*, Oviedo, Spain, 1-3 December 2012

[30] Falconer, K. J., *Fractal Geometry: Mathematical Foundations and Applycations.* John Wiley and Sons, England (1990)

[31] Furuichi, S., Information theoretical properties of Tsallis entropies. *Journal of Mathematical Physics*, **47**, 023302 (2006)

[32] Gell-Man, M., Tsallis, C., *Nonextensive Entropy - Interdisciplinary Applications*. Oxford University Press, New York (2004)

[33] Gneiting, T., Schlather, M., Stochastic models that separate fractal dimension and the Hurst effect. *SIAM Review*, **46**, 269–282 (2004)

[34] Godano, C., Tosi, P., Derubeis, V., Augliera, P., Scaling properties of the spatio–temporal distribution of earthquakes: a multifractal approach applied to a Californian catalogue. *Geophysical Journal International*, **136**, 99-108 (1999)

[35] Goltz, M., *Fractal and Chaotic Properties of Earthquakes*. Springer, Berlin (1998)

[36] Gutenberg, B., Richter, C. F., Bulletin of the Seismological Society of America. *Bull. Seismol. Soc. Am.*, **34**, 185-188 (1944)

[37] Guttorp, P., *Stochastic Modeling of Scientific Data.* Chapman and Hall, London (1995)

[38] Halsey T. C., Jensen M.H., Kadanoff L. P., Procaccia I., Shraiman B. I., Fractal measures and their singularities: the characterization of strange sets. *Nuclear Physics B*, **33**, 1141–1151 (1986)

[39] Harte, D., *Multifractals: Theory and Applications.* Chapman and Hall/CRC, Boca Raton (2001)

[40] Harte, D., Vere-Jones, D., The entropy score and its uses in earthquake forecasting. *Pure and Applied Geophysics*, **162** ,1229–1253 (2005)

[41] Havrda, J., Charvát, F., Quantification method of classification processes: concept of structural *a*-entropy. *Kybernetika*, **3**, 30-35 (1967)

[42] Hentschel, H., Procaccia, I., The infinite number of generalized dimensions of fractals and strange attractors , *Physica D*, **8**, 435-444 (1983)

[43] Hirata, T., A correlation between the *b*-value and the fractal dimension of earthquakes. *Geophysical Journal International*, **94**, 7507–7514 (1989)

[44] Ibáñez, J. M., De Angleis, S., Díaz-Moreno, A., Hernández, P., Alguacil, G., Posadas, A., Pérez, N., Insights into the 2011-2012 submarine eruption off the coast of El Hierro (Canary Islands, Spain) from statistical analyses of earthquake activity. *Geophysical Journal International*, **191**, 659-670 (2012)

[45] Kagan, Y. Y., Knopoff, L., Statistical study of the occurrence of shallow earthquakes. *Geophysic Journal International*, **55**, 67-86 (1978)

[46] Kagan, Y. Y., Knopoff, L., Spatial distribution of earthquakes: the two-point correlation function. *Geophysical Journal of the Royal Astronomical Society*, **62**, 303–320 (1980)

[47] Kagan, Y. Y., Knopoff, L., Stochastic synthesis of earthquake catalogs. *Journal of Geophysical Research*, **86**, 2853-2862 (1981)

[48] Kalimeri, M., Papadimitriou, C., Balasis, G., Eftaxias, K., Dynamical complexity detection in pre-seismic emissions using nonadditive Tsallis entropy. *Physica A*, **387**, 1161-1172 (2008)

[49] Kanamori, K., Brodsky, E., The physics of earthquakes. *Reports on Progress in Physics*, **67**, 1429–1496 (2004)

[50] Kantelhardt, J. W., Zschiegner, S. A., Koscielny-Bunde, E., Havlin, S., Bunde, A., Stanley, H. E., Multifractal detrended fluctuation analysis of nonstationary time series. *Physica A*, **316**, 87-114 (2002)

[51] Korvin. G., *Fractal Models in the Earth Sciences*. Elsevier, Amsterdam (1992)

[52] Kravchenko, A. N., Boast, C. W., Bullock, D. G., Multifractal analysis of soil spatial variability. *Agronomy Journal*, **91**, 1033-1041 (1999)

[53] Kullback, S., Leibler, R. A., On information and sufficiency. *Annals of Mathematical Statistics*, **22**, 79–86 (1951)

[54] Levenberg, K., A method for the solution of certain non-linear problems in least squares. *Quarterly of Applied Mathematics*, **2**, 164–168 (1944)

[55] López-Ruiz, R., Mancini, H.L., Calbet, X., A statistical measure of complexity. *Physics Letters A*, **209**, 321-326 (1995)

[56] López-Ruiz, R., Shannon information, LMC complexity and Rényi entropies: a straightforward approach. *Biophysical Chemistry*, **115**, 215– 218 (2005)

[57] López-Ruiz, R., Nagy, Á., Romera, E., Sañudo, J., A generalized statistical complexity measure: applications to quantum systems. *Journal of Mathematical Physics*, **50**, 123528 (2009)

[58] Madrid, A. E., Angulo J. M., Mateu J., Spatial threshold exceedance analysis through marked point processes. *Environmetrics*, **23**, 108-118 (2012)

[59] Main, I., Burton, P., W., Information theory and the earthquake frequency-magnitude distribution. *Bull. Seism. Soc. Am.*, **74**, 1409-1426 (1984)

[60] Main, I., Statistical physics, seismogenesis, and seismic hazard. *Rev. Geophys.*, **34**, 433-462 (1996)

[61] Mandelbrot, B., How long is the coast of Britain? Statistical self-similarity and fractional dimension. *Science, New Series*, **156**, 636-638 (1967)

[62] Mandelbrot, B., *Fractals: Form, Chance and Dimension*. Freeman, San Francisco (1977)

[63] Mandelbrot, B., *The Fractal Geometry of Nature*. Freeman, New York (1982)

[64] Marquardt, D., An algorithm for least-squares estimation of nonlinear parameters. *SIAM Journal On Applied Mathematics*, **11**, 431-441 (1963)

[65] Marsan, D., Bean, C. J., Steacy, S., Spatio-temporal analysis of stress diffusion in a mining-induced seismicity system. *Geophysical Research Letters.* **26**, 3697-3700, (1999)

[66] Martín, M.T., Plastino, A., Rosso, O.A., Generalized statistical complexity measures: Geometrical and analytical properties. *Physica A*, **369**, 439-462 (2006)

[67] Mignan, A., Retrospective on the accelerating seismic release (ASR) hypothesis: controversy and new horizons. *Tectonophysics*, **505**, 1-16 (2011)

[68] Mignan, A., Woessner, J., Estimating the magnitude of completeness for earthquake catalogs, *Community Online Resource for Statistical Seismicity Analysis*, doi:10.5078/corssa-00180805. Available at http://www.corssa.org. (2012)

[69] Mogi, K., Magnitude–frequency relation for elastic shocks accompanying fractures of various materials and some related problems in earthquakes. *Journal of College of Science of the Imperial University of Tokyo*, **40**, 831-853 (1962)

[70] Movahed, M., Hermanis, S. E., Fractal analysis of river flow fluctuations. *Physica A*, **387**, 915–932 (2008)

[71] Ogata,Y., Statistical models for earthquake occurrence and residual analysis for point processes. *Journal of the American Statistical Association*, **83**, 9-27 (1988)

[72] Omori, F., On the aftershocks of earthquakes. *Journal of College of Science of the Imperial University of Tokyo*, **7**, 111-200 (1894)

[73] Öncel A.O., Wilson T. H., Anomalous seismicity preceding the 1999 Izmit event, NW Turkey. *Geophysical Journal International*, **169**, 259-270 (2007)

[74] Onicescu, O., Energie Informationelle.*Comptes Rendus de la Academie des Sciences Paris*, **A**, 841–842 (1966)

[75] Orey, S., Gaussian sample functions and the Hausdorff dimension of level crossings. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, **15**, 249-256 (1970)

[76] Perrier, E., Tarquis, A., Dathe, A., A program for fractal and multifractal analysis of 2D binary images: Computer algorithms versus mathematical theory. *Geoderma*, **134**, 284-294 (2006)

[77] Piterbarg V. I., *Asymptotic Methods in the Theory of Gaussian Processes and Fields*, American Mathematical Society, Providence (1996)

[78] Posadas, A. N. D., Quiroz, R., Zorogastua, P. E., Leon-Verlade, C., Multifractal characterization of the spatial distribution of ulexite in a Bolivian salt flat. *International Journal of Remote Sensing*, **26**, 615-627 (2005)

[79] Rényi, A., On measures of entropy and information. *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, **1**, 547-561, (1961)

[80] Rényi, A., On some basic problems of statistics from the point of view of information theory, *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, **1**, 531–543 (1967)

[81] Rényi, A., *Probability Theory*, North-Holland, Amsterdam (1970)

[82] Ripley, B., Modelling spatial patterns. *Journal of the Royal Statistical Society*, **39**, 172-192 (1977)

[83] Roy S., Ghosh U., Hazra S., Kayal J. R., Fractal dimension and *b*-value mapping in the Andaman-Sumatra subduction zone. *Natural Hazards*, **57**, 27–37 (2011)

[84] Sadovskiy, M. A., Golubeva, T. V., Pisarenko, V. F. Shnirman, M., G., Characteristic dimensions of rock and hierarchical properties of seismicity, *Izvestiya*, **20**, 87-96 (1984)

[85] Sadovskiy, M. A., Pisarenko, V. F., Seismic process in block environment. *Moscow: Nauka*, **96**, (1991)

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

[87] Sen, K. D., Multifractality as a measure of complexity in solar flare activity. *Solar Physics*, **241**, 67–76 (2007)

[88] Sen, K. D., *Statistical Complexity: Applications In Electronic Structure*. Springer (2011)

[89] Schoenberg, F. P., Brillinger, D. R., and Guttorp, P. M., Point processes, spatial-temporal. *In Encyclopedia of Environmetrics*, **3**, 1573–1577 (2002)

[90] Schoenberg, F. P., Testing separability in spatial-temporal marked point processes, *Biometrics*, **60**, 471-481 (2004)

[91] Scholz, C. H., The frequency–magnitude relation of microfracturing in rock and its relation to earthquakes. *Bulletin of the Seismological Society of America*, **58**, 399–415 (1968)

[92] Shannon, C. E., A mathematical theory of communication. *Bell System Technical Journal*, **27**, 379-423 (1948)

[93] Shiner, J. S., Davison, M., and Landsberg, P. T., Simple measure for complexity. *Physical Review E.*, **59**, 14591464 (1999)

[94] Silchenko, A., Hu, C., Multifractal characterization of stochastic resonance. *Physical Review E*, **63**, 041105 (2001)

[95] Silva, R., França, G. S., Vilar, C. S., Alcaniz J. S., Nonextensive models for earthquakes. *Physical Review E*, **73**, 026102 (2006)

[96] Sornette, D., Earthquakes: from chemical alteration to mechanical rupture, *Physics Reports*, **313**, 328-292 (1999)

[97] Sornette, D., *Critical Phenomena in Natural Sciences*, Springer, Berlín (2000)

[98] Sotolongo-Costa, O., Rodríguez, A. H., Rodgers, G. J., Tsallis entropy and the transition to scaling in fragmentation. *Entropy*, **2**, 172-177 (2000)

[99] Sotolongo-Costa, O. Posadas, A., Fragment-asperity interaction model for rarthquakes. *Phys. Rev. Lett.*, **92**, 048501 (2004)

[100] Sotolongo-Costa, A., Non-extensive framework for earthquakes: the role of fragments. *Pacta Geophysica*, **60**, 526-534 (2012)

[101] Stanley, H. E., Meakin, P., Multifractal phenomena in physics and chemistry, *Nature*, **335**, 405-409 (1988)

[102] Tanaka, P., Varotsos, P. A., Sarlis, N. V., Skordas, E. S., A plausible universal behavior of earthquakes in the natural time-domain. *Proc. Japan Acad. Ser. B*, **80**, 283-289 (2004)

[103] Telesca, L., Lapenna, V., Alexis, N., Multiresolution wavelet analysis of earthquakes. *Chaos, Solitons and Fractals*, **22**, 741-748 (2004)

[104] Telesca, L., A non-extensive approach in investigating the seismicity of L'Aquila area (central Italy). *Terra Nova*, **22**, 87-93 (2010a)

[105] Telesca, L., Analysis of Italian seismicity by usingh a nonextensive approach. *Technophysics*, **494**, 155-162 (2010b)

[106] Telesca, L., Tsallis-based nonextensive analysis of the Southern California seismicity. *Entropy*, **13**, 1267-1280 (2011)

[107] Telesca, L., Maximum likelihood estimation of the nonextensive parameters of the earthquake cumulative magnitude distribution. *Bulletin of the Seismological Society of America*, **102**, 886-891 (2012)

[108] Tsallis, C., Possible generalization of Boltzmann-Gibbs statistics. *J. Stat. Phys.*, **52**, 479-487 (1988)

[109] Tsallis, C., *Introduction to Nonextensive Statistical Mechanics: Approaching a Complex World*, Springer, New York (2009)

[110] Turcotte, D. L., Fractals in Geology and Geophysics. *Pure and Applied Geophysics*, **131**, 171-196 (1989)

[111] Utsu, T., A statistical study of the occurrence of aftershocks, *Geophysical Magazine*, **30**, 521–60 (1961)

[112] Vallianatos, F., A non-extensive approach to risk assessment. *Natural Hazards and Earth System Sciences*, **9**, 211-216 (2009)

[113] Vallianatos, F., Michas, G., Papadakis, G., Tzanis, A., Evidence of non-extensivity in the seismicity observed during the 2011-2012 unrest at the Santorini volcanic complex, Greece. *Nat. Hazards Earth Syst. Sci.*, **13**, 177-185 (2013)

[114] Varotsos, P., Sarlis, N. V., Skordas, E. S., Uyeda, S., Kamogawa, M., Natural time analysis of critical phenomena. *Proceedings of the National Academy of Sciences of the United States of America.*, **28**, 11361-11364 (2011)

[115] Veneziano, D., Moglen, G., Bras, R., Multifractal analysis: pitfalls of standard procedures and alternatives. *Physical Review E.*, **52**, 1387-1398 (1995)

[116] Vere-Jones, D., On the fractal dimensions of point patterns. *Journal of Applied Probability*, **31(3)**, 643–663 (1999)

[117] Vere-Jones, D., Seismology – a statistical vignette. *Journal of the American Statistical Association*, **95**, 975-978, (2000)

[118] Vere-Jones, D., The marriage of statistics and seismology. *Journal of Applied Probability*, **38**, 1-5 (2001)

[119] Vere-Jones, D., Foundations of statistical seismology. *Pure and Applied Geophysics*, **167**, 645-653 (2010)

[120] Wang, J., Ning, X., Ma, Q., Bian, C., Xu, Y., Chen, Y., Multiscale multifractality analysis of a 12-lead electrocardiogram. *Physical Review E.,* **71**, 062902 (2005)

[121] Wiemer, S., Wyss, M., Minimum magnitude of compelte reporting in earthquake catalogs: examples from Alaska, the western United States, and Japan. *Bulletin of the Seismological Society of America*, **90**, 859-869 (1973)

[122] Wiemer, S., Wyss, M., Mapping spatial variability of the frequency–magnitude distribution of earthquakes. *Advances Geophysics*, **45**, 259-302 (2002)

[123] Wyss, M., Towards a physical understanding of the frequency–magnitude distribution. *Geophysical Journal of the Royal Astronomical Society* , **31**, 341-359 (1973)

[124] Wyss, M., Hasegawa, A., Wiemer, S., Umino, N., Quantitative mapping of precursory seismic quiescence before the 1989, M7.1 off-Sanriku earthquake, Japan, *Annal Geofysics*, **42**, 851-869 (1999)

[125] Zhou Y., Zhang Q., Singh V., P., Fractal-based evaluation of the effect of water reservoirs on hydrological processes: the dams in the Yangtze River as a case study. *Stochastic Environmental Research and Risk Assessment*, doi 10.1007/s00477-013-0747-5 (2013)

Appendix

Appendix

This appendix collects a selection of **MATLAB** codes developed for the elaboration of the thesis.

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%% Shannon entropy values for %%%%
%%%%% the interaction (X,Y)-Z  %%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

clear all
clc

% Change the current folder

cd C:\DATA

% Load the data

load THESIS\Agron

% Definition of the study area

min_x=-3.9;max_x=-3.4;
min_y=36.34;max_y=37.88;
min_z=0;max_z=70.1;

% Number of events

n_data=length(x);
```

```
% Data normalization

x=x-min_x;
y=y-min_y;
z=z-min_z;
max_x=max_x-min_x;
max_y=max_y-min_y;
max_z=max_z-min_z;
x=x/max_x;
y=y/max_y;
z=z/max_z;

% Number of box sizes

f=15;

% Box sizes

for v=1:f
     E(v)=1/(f-v+2);
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Shannon entropy values of (Z) %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

for v=1:f

% Box size

e=E(v);

% Partitions in the axe

nz=f-v+2;

% Counting of the points fell in each box

zn=linspace(0,0,n_data);
```

```
for i=1:n_data
    for j=1:nz
        if 0+(j-1)*e<z(i) & z(i)<=0+(j)*e
            zn(i)=j;
        end
    end
end

for i=1:n_data
   if zn(i)==0
      zn(i)=1;
   end
end

% Frequencies and probabilities

FREC1=linspace(0,0,n_data);

a=0;
for i=1:n_data
        a=zn(i);
        FREC1(a)=FREC1(a)+1;
end

p1=linspace(0,0,nz);

for j=1:nz
    p1(j)=FREC1(j)/n_data;
end

% Shannon entropy values

sum1=0;
for i=1:nz
    if p1(i)~=0
       sum1=sum1+(p1(i)*log(p1(i)));
    end
end
```

```
S1(v)=-sum1;

end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Shannon entropy values of (X,Y) %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

for v=1:f

% Box size

e=E(v);

% Partitions in the axes

nx=f-v+2;
ny=f-v+2;

% Counter of the points fell in each box

xn=linspace(0,0,n_data);
yn=linspace(0,0,n_data);

for i=1:n_data
    for j=1:nx
       if 0+(j-1)*e<x(i) & x(i)<=0+(j)*e
            xn(i)=j;
       end
    end
    for j=1:ny
        if 0+(j-1)*e<y(i) & y(i)<=0+(j)*e
            yn(i)=j;
        end
    end
end

for i=1:n_data
```

```
   if xn(i)==0
      xn(i)=1;
   end
end

for i=1:n_data
   if yn(i)==0
      yn(i)=1;
   end
end

% Frequencies and probabilities

FREC2=zeros(nx,ny);

a=0;b=0;
for i=1:n_data
    a=xn(i);
    b=yn(i);
    FREC2(a,b)=FREC2(a,b)+1;
end

p2=zeros(nx,ny);

for i=1:nx
    for j=1:ny
         p2(i,j)=FREC2(i,j)/n_data;
    end
end

% Shannon entropy values

sum2=0;
for i=1:nx
   for j=1:ny
        if p2(i,j)~=0
           sum2=sum2+(p2(i,j)*log(p2(i,j)));
        end
    end
```

```
end

S2(v)=-sum2;

end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Shannon entropy values of (X,Y,Z) %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

for v=1:f

% Box size

e=E(v);

% Partitions in the axes

nx=f-v+2;
ny=f-v+2;
nz=f-v+2;

% Counter of the points fell in each box

xn=linspace(0,0,n_data);
yn=linspace(0,0,n_data);
zn=linspace(0,0,n_data);

for i=1:n_data
     for j=1:nx
        if 0+(j-1)*e<x(i) & x(i)<=0+(j)*e
            xn(i)=j;
        end
     end
     for j=1:ny
           if 0+(j-1)*e<y(i) & y(i)<=0+(j)*e
            yn(i)=j;
            end
     end
```

```
     for j=1:nz
           if 0+(j-1)*e<z(i) & z(i)<=0+(j)*e
            zn(i)=j;
           end
     end
end


for i=1:n_data
   if yn(i)==0
      yn(i)=1;
   end
end

for i=1:n_data
   if xn(i)==0
      xn(i)=1;
   end
end

for i=1:n_data
   if zn(i)==0
      zn(i)=1;
   end
end

% Frequencies and probabilities

FREC3=zeros(nx,ny,nz);

a=0;b=0;c=0;
for i=1:n_data
        a=xn(i);
        b=yn(i);
        c=zn(i);
        FREC3(a,b,c)=FREC3(a,b,c)+1;
end

p3=zeros(nx,ny,nz);
```

```
for i=1:nx
    for j=1:ny
        for l=1:nz
         p3(i,j,l)=FREC3(i,j,l)/n_data;
        end
    end
end

% Shannon entropy values

sum3=0;
for i=1:nx
   for j=1:ny
       for l=1:nz
           if p3(i,j,l)~=0
           sum3=sum3+(p3(i,j,l)*log(p3(i,j,l)));
           end
        end
    end
end

S3(v)=-sum3;

end

% Extensivity expression

for v=1:f
Extensivity(v)=S1(v)+S2(v);
end

% Departure between the interaction and the extensivity expression

for n=1:f
Departure(n)=Extensivity(n)-S3(n);
end

% Plotting the results
```

```
hold on
plot(E,S1,'g'),plot(E,S2,'b'),plot(E,S3,'k'),plot(E,Extensivity,'r'),
plot(E,Departure,'y');
title('Agrón (X,Y)-Z')
xlabel('Box size')
ylabel('Shannon entropy')
legend('H(T)','H(X,Y)','H(X,Y,T)','H(X,Y)+H(T)','H(X,Y)+H(T)-H(X,Y,T)')
hold off
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%% Dependence coefficient calculated by means of %%%%
%%%%% sliding windows for the intercation (X,Y)-Z %%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

clear all
clc

% Change the current folder.

cd C:\DATA

% Load the data

load THESIS\Agron

% Definition of the study area

min_x=-3.9;max_x=-3.4;
min_y=36.34;max_y=37.88;
min_z=0;max_z=70.1;

% Number of events

num_data=length(x);

% Data normalization

x=x-min_x;
y=y-min_y;
z=z-min_z;
max_x=max_x-min_x;
max_y=max_y-min_y;
max_z=max_z-min_z;
x=x/max_x;
```

```
y=y/max_y;
z=z/max_z;

% Duplicate the components

tt=t;
xx=x;
yy=y;
zz=z;
mm=m;


% window size

size=150;

% Overlapping

increment=(100-90)*size/100;

% Number of windows

num=floor((((num_data)/size)*((size/increment)))-((size/increment)-1));

for i=1:num
    e_inf(i)=increment*(i-1);
    e_sup(i)=e_inf(i)+size;
    centre(i)=(e_inf(i)+e_sup(i))/2;
end

e_inf=e_inf+1;
e_sup=e_sup;

if e_sup(num)-num_data>=0.5*increment
    num=num-1;
end

q=1;
```

```
% Start the sliding windows

for n=1:num

    clear m x y t z
    if e_sup(n)<=length(mm)
        x=xx(e_inf(n):e_sup(n));
        y=yy(e_inf(n):e_sup(n));
        z=zz(e_inf(n):e_sup(n));
        m=mm(e_inf(n):e_sup(n));
        t=tt(e_inf(n):e_sup(n));
        tmax=tt(e_sup(n));
        tmin=tt(e_inf(n));
    else
        m=mm(e_inf(n):length(mm));
        x=xx(e_inf(n):length(mm));
        y=yy(e_inf(n):length(mm));
        z=zz(e_inf(n):length(mm));
        t=tt(e_inf(n):length(mm));
        tmax=tt(length(mm));
        tmin=tt(e_inf(n));
    end


n_data=length(x);
time(n)=tmax;

% Number of box sizes

f=15;

% Box sizes

for v=1:f
    E(v)=1/(f-v+2);
end

% Scale logarithms
```

```
for v=1:f
    LogE(v)=log(E(v));
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% Information dimension values for (Z) %%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

for v=1:f

% Box size

e=E(v);

% % Partitions in the axe

nz=(f-v+2);

% Counter of the points fell in each box

zn=linspace(0,0,n_data);

for i=1:length(x)

     for j=1:nz
        if 0+(j-1)*e<z(i) & z(i)<=0+(j)*e
            zn(i)=j;
        end
     end
end

for i=1:n_data
   if zn(i)==0
      zn(i)=1;
   end
end

% Frequencies and probabilities with and without
% considering the effects of the magnitude
```

```
FREC1=linspace(0,0,nz);
Smag1=linspace(0,0,nz);

a=0;
for i=1:n_data
    a=zn(i);
    FREC1(a)=FREC1(a)+1;
    Smag1(a)=Smag1(a)+exp(m(i));
end

STmag1=0;
for i=1:n_data
    STmag1=STmag1+exp(m(i));
end


p1=linspace(0,0,nz);
p1_effects=linspace(0,0,nz);

for l=1:nz
    p1(l)=FREC1(l)/n_data;
    p1_effects(l)=Smag1(l)/STmag1;
end

% Calculations of the measures

sum1=0; sum2=0;
sum1_effects=0; sum2_effects=0;

      for l=1:nz
         if p1(l)~=0
           sum1=sum1+p1(l)^(q);
           sum2=sum2+p1(l)*log(p1(l));
          end
      end


      for l=1:nz
         if p1_effects(l)~=0
```

```
            sum1_effects=sum1_effects+p1_effects(l)^(q);
            sum2_effects=sum2_effects+p1_effects(l)*log(p1_effects(l));
           end
      end


measure1(v)=sum1;
logmeasure1(v)=log(measure1(v));
measure2(v)=sum2;

measure1_effects(v)=sum1_effects;
logmeasure1_effects(v)=log(measure1_effects(v));
measure2_effects(v)=sum2_effects;

end

% Fitting log-log

L=linspace(0,0,f);
W=[L;LogE]';
P=regress(measure2',W);
P_effects=regress(measure2_effects',W);
D_Z(n)=P(2);
D_Z_effects(n)=P_effects(2);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% Information dimension values for (X,Y) %%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

for v=1:f

% Box size

e=E(v); % Tamaño de caja

% Partitions of the axes

nx=(f-v+2);
ny=(f-v+2);
```

```
% Counter of the points fell in each box

xn=linspace(0,0,n_data);
yn=linspace(0,0,n_data);

for i=1:length(x)
     for j=1:nx
        if 0+(j-1)*e<x(i) & x(i)<=0+(j)*e
            xn(i)=j;
        end
     end
     for j=1:ny
        if 0+(j-1)*e<y(i) & y(i)<=0+(j)*e
            yn(i)=j;
        end
     end
end

for i=1:n_data
   if xn(i)==0
      xn(i)=1;
   end
end

for i=1:n_data
   if yn(i)==0
      yn(i)=1;
   end
end

% Frequencies and probabilities with and without
% considering the effects of the magnitude
FREC2=zeros(nx,ny);
Smag2=zeros(nx,ny);

a=0;b=0;

for i=1:n_data
```

```
        a=xn(i);
        b=yn(i);
        FREC2(a,b)=FREC2(a,b)+1;
        Smag2(a,b)=Smag2(a,b)+exp(m(i));
end

STmag2=0;

for i=1:n_data
    STmag2=STmag2+exp(m(i));
end

p2=zeros(nx,ny);
p2_effects=zeros(nx,ny);

for i=1:nx
    for j=1:ny
         p2(i,j)=FREC2(i,j)/n_data;
         p2_effects(i,j)=Smag2(i,j)/STmag2;
    end
end

% Calculations of the measures

sum1=0; sum2=0;
sum1_effects=0; sum2_effects=0;

for i=1:nx
   for j=1:ny
       if p2(i,j)~=0
           sum1=sum1+p2(i,j)^(q);
           sum2=sum2+p2(i,j)*log(p2(i,j));
       end
   end
end

for i=1:nx
   for j=1:ny
       if p2_effects(i,j)~=0
```

```
            sum1_effects=sum1_effects+p2_effects(i,j)^(q);
            sum2_effects=sum2_effects+p2_effects(i,j)*log(p2_effects(i,j));
        end
    end
 end

measure1(v)=sum1;
logmeasure1(v)=log(measure1(v));
measure2(v)=sum2;

measure1_effects(v)=sum1_effects;
logmeasure1_effects(v)=log(measure1_effects(v));
measure2_effects(v)=sum2_effects;

end

% Fitting log-log

L=linspace(0,0,f);
W=[L;LogE]';
P2=regress(measure2',W);
P2_effects=regress(measure2_effects',W);
D_XY(n)=P2(2);
D_XY_effects(n)=P2_effects(2);



%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% Information dimension values for (X,Y,Z) %%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

for v=1:f

% Box size

e=E(v);

% Partitions of the axes
```

```
nx=(f-v+2);
ny=(f-v+2);
nz=(f-v+2);

% Counter of the points fell in each box

xn=linspace(0,0,n_data);
yn=linspace(0,0,n_data);
zn=linspace(0,0,n_data);

for i=1:length(x)
     for j=1:nx
        if 0+(j-1)*e<x(i) & x(i)<=0+(j)*e
            xn(i)=j;
        end
     end
     for j=1:ny
        if 0+(j-1)*e<y(i) & y(i)<=0+(j)*e
            yn(i)=j;
        end
     end
     for j=1:nz
        if 0+(j-1)*e<z(i) & z(i)<=0+(j)*e
            zn(i)=j;
        end
     end
end

for i=1:n_data
   if xn(i)==0
      xn(i)=1;
   end
end

for i=1:n_data
   if yn(i)==0
      yn(i)=1;
   end
end
```

```
for i=1:n_data
   if zn(i)==0
      zn(i)=1;
   end
end

% Frequencies and probabilities with and without
% considering the effects of the magnitude
FREC3=zeros(nx,ny,nz);
Smag3=zeros(nx,ny,nz);

a=0;b=0;c=0;
for i=1:n_data
        a=xn(i);
        b=yn(i);
        c=zn(i);
        FREC3(a,b,c)=FREC3(a,b,c)+1;
        Smag3(a,b,c)=Smag3(a,b,c)+exp(m(i));
end

STmag3=0;
for i=1:n_data
    STmag3=STmag3+exp(m(i));
end

p3=zeros(nx,ny,nz);
p3_effects=zeros(nx,ny,nz);

for i=1:nx
    for j=1:ny
        for l=1:nz
         p3(i,j,l)=FREC3(i,j,l)/n_data;
         p3_effects(i,j,l)=Smag3(i,j,l)/STmag3;
        end
    end
end

% Calculations of the measures
```

```
sum1=0; sum2=0;
sum1_effects=0; sum2_effects=0;

for i=1:nx
   for j=1:ny
      for l=1:nz
         if p3(i,j,l)~=0
           sum1=sum1+p3(i,j,l)^(q);
           sum2=sum2+p3(i,j,l)*log(p3(i,j,l));
          end
       end
    end
end

for i=1:nx
   for j=1:ny
      for l=1:nz
         if p3_effects(i,j,l)~=0
           sum1_effects=sum1_effects+p3_effects(i,j,l)^(q);
           sum2_effects=sum2_effects+p3_effects(i,j,l)*
           log(p3_effects(i,j,l));
         end
       end
    end
 end

measure1(v)=sum1;
logmeasure1(v)=log(measure1(v));
measure2(v)=sum2;

measure1_effects(v)=sum1_effects;
logmeasure1_effects(v)=log(measure1_effects(v));
measure2_effects(v)=sum2_effects;

end

% Log-Log fitting
```

```
L=linspace(0,0,f);
W=[L;LogE]';
P2=regress(measure2',W);
P2_effects=regress(measure2_effects',W);
D_XYZ(n)=P2(2);
D_XYZ_effects(n)=P2_effects(2);

end

% Dependence coefficient

for i=1:num
    CD(i)=(D_XY(i)+D_Z(i)-D_XYZ(i))/(D_XY(i)+D_Z(i));
    CD_effects(i)=(D_XY_effects(i)+D_Z_effects(i)-D_XYZ_effects(i))/
    (D_XY_effects(i)+D_Z_effects(i));
end

% Plotting the results

hold on
xlabel('Time')
ylabel('Dependence measure')
title('Agrón (X,Y)-Z')
plot(time,CD,'b');
plot(time,CD_effects,'r');
legend('phi(m)=1','phi(m)=energy')
hold off
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%% Generalized dimensions, complexity measure %%%%%%
%%%%% and step calculated by means of sliding %%%%%%%%%
%%%%% windows for the interaction (X,Y,Z) with %%%%%%%%
%%%%% and without considering the effect of %%%%%%%%%%%
%%%%%%%%%%%%%%%%%% the magnitude %%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

clear all
clc

% Change the current folder

cd C:\DATA

% Load the data

load THESIS\Agron

% Definition of the study area

min_x=-3.9;max_x=-3.4;
min_y=36.34;max_y=37.88;
min_z=0;max_z=70.1;

% Number of events

num_data=length(x);

% Data normalization

x=x-min_x;
y=y-min_y;
z=z-min_z;
max_x=max_x-min_x;
max_y=max_y-min_y;
max_z=max_z-min_z;
x=x/max_x;
```

```
y=y/max_y;
z=z/max_z;

% Duplication of the components

tt=t;
xx=x;
yy=y;
zz=z;
mm=m;

% window size

size=150;

% Overlapping

increment=(100-90)*size/100;

% Number of windows

num=floor((((num_data)/size)*((size/increment)))-((size/increment)-1));

for i=1:num
    e_inf(i)=increment*(i-1);
    e_sup(i)=e_inf(i)+size;
    centre(i)=(e_inf(i)+e_sup(i))/2;
end

e_inf=e_inf+1;
e_sup=e_sup;

if e_sup(num)-num_data>=0.5*increment
    num=num-1;
end

% Start the sliding windows

for n=1:num
```

```
    clear m x y t z
    if e_sup(n)<=length(mm)
        x=xx(e_inf(n):e_sup(n));
        y=yy(e_inf(n):e_sup(n));
        z=zz(e_inf(n):e_sup(n));
        m=mm(e_inf(n):e_sup(n));
        t=tt(e_inf(n):e_sup(n));
        tmax=tt(e_sup(n));
        tmin=tt(e_inf(n));
    else
        m=mm(e_inf(n):length(mm));
        x=xx(e_inf(n):length(mm));
        y=yy(e_inf(n):length(mm));
        z=zz(e_inf(n):length(mm));
        t=tt(e_inf(n):length(mm));
        tmax=tt(length(mm));
        tmin=tt(e_inf(n));
    end

n_data=length(x);
time(n)=tmax;

% Number of box sizes

f=15;

% Box sizes

for v=1:f
    E(v)=1/(f-v+2);
end

% Scale logarithms

for v=1:f
    LogE(v)=log(E(v));
end
```

```
% Number of values of q

num_q=5;
QQ=linspace(-floor(num_q/2),floor(num_q/2),num_q);

% Start the calculations for each value of q

for w=1:num_q
q=QQ(w);

for v=1:f

% Box size

e=E(v);

% Partitions in the axes

nx=(f-v+2);
ny=(f-v+2);
nz=(f-v+2);

% Counter of the points fell in each box

xn=linspace(0,0,n_data);
yn=linspace(0,0,n_data);
zn=linspace(0,0,n_data);

for i=1:length(x)
     for j=1:nx
        if 0+(j-1)*e<x(i) & x(i)<=0+(j)*e
            xn(i)=j;
        end
     end
     for j=1:ny
        if 0+(j-1)*e<y(i) & y(i)<=0+(j)*e
            yn(i)=j;
        end
     end
```

```
     for j=1:nz
        if 0+(j-1)*e<z(i) & z(i)<=0+(j)*e
            zn(i)=j;
        end
     end
end

for i=1:n_data
   if xn(i)==0
      xn(i)=1;
   end
end

for i=1:n_data
   if yn(i)==0
      yn(i)=1;
   end
end

for i=1:n_data
   if zn(i)==0
      zn(i)=1;
   end
end


% Frequencies and probabilities with and without
% considering the effects of the magnitude
FREC3=zeros(nx,ny,nz);
Smag3=zeros(nx,ny,nz);

a=0;b=0;c=0;
for i=1:n_data
        a=xn(i);
        b=yn(i);
        c=zn(i);
        FREC3(a,b,c)=FREC3(a,b,c)+1;
        Smag3(a,b,c)=Smag3(a,b,c)+exp((3/2)*m(i));
end
```

```
STmag3=0;
for i=1:n_data
    STmag3=STmag3+exp((3/2)*m(i));
end

p3=zeros(nx,ny,nz);
p3_effects=zeros(nx,ny,nz);

for i=1:nx
    for j=1:ny
        for l=1:nz
         p3(i,j,l)=FREC3(i,j,l)/n_data;
         p3_effects(i,j,l)=Smag3(i,j,l)/STmag3;
        end
    end
end

% Calculations of the measures

sum1=0; sum2=0;
sum1_effects=0; sum2_effects=0;

for i=1:nx
   for j=1:ny
      for l=1:nz
         if p3(i,j,l)~=0
           sum1=sum1+p3(i,j,l)^(q);
           sum2=sum2+p3(i,j,l)*log(p3(i,j,l));
          end
       end
    end
end

for i=1:nx
   for j=1:ny
      for l=1:nz
         if p3_effects(i,j,l)~=0
           sum1_effects=sum1_effects+p3_effects(i,j,l)^(q);
```

```
            sum2_effects=sum2_effects+p3_effects(i,j,l)*
            log(p3_effects(i,j,l));
          end
        end
    end
 end

measure1(v)=sum1;
logmeasure1(v)=log(measure1(v));
measure2(v)=sum2;

measure1_effects(v)=sum1_effects;
logmeasure1_effects(v)=log(measure1_effects(v));
measure2_effects(v)=sum2_effects;

end

% Log-Log fitting

if q==1

L=linspace(0,0,f);
W=[L;LogE]';
P2=regress(measure2',W);
P2_effects=regress(measure2_effects',W);
D_XYZ(w,n)=P2(2);
D_XYZ_effects(w,n)=P2_effects(2);

else % q<>1

L=linspace(0,0,f);
W=[L;LogE]';
P=regress(logmeasure1',W);
P_effects=regress(logmeasure1_effects',W);
tau(w)=P(2);
tau_effects(w)=P_effects(2);
D_XYZ(w,n)=(1/(q-1))*(P(2));
D_XYZ_effects(w,n)=(1/(q-1))*(P_effects(2));
```

```
end

end

% Generalized dimensions, step and complexity measure

D_0(n)=D_XYZ(floor(num_q/2)+1,n);
D_1(n)=D_XYZ(floor(num_q/2)+2,n);
D_2(n)=D_XYZ(floor(num_q/2)+3,n);
Step(n)=D_XYZ(1,n)-D_XYZ(num_q,n);
Diff(n)=D_1(n)-D_2(n);

D_0_effects(n)=D_XYZ_effects(floor(num_q/2)+1,n);
D_1_effects(n)=D_XYZ_effects(floor(num_q/2)+2,n);
D_2_effects(n)=D_XYZ_effects(floor(num_q/2)+3,n);
Step_effects(n)=D_XYZ_effects(1,n)-D_XYZ_effects(num_q,n);
Diff_effects(n)=D_1_effects(n)-D_2_effects(n);

end

% Boundaries

bx1=linspace(388.7,388.7,10);
by1=linspace(0,5,10);
bx2=linspace(569.3,569.3,10);
by2=linspace(0,5,10);

% Plotting the results

hold on
xlabel('Time')
ylabel('Step')
title('Agrón (X,Y,Z)')
plot(time,Step,'b');
plot(time,Step_effects,'r');
plot(bx1,by2,'--b');
plot(bx2,by2,'--b');
legend('Phi(m)=1','Phi(m)=energy')
hold off
```