Multifractal Analysis of Geographical Structures and Processes: Concepts and Applications

ZHOU, Yu

A Thesis Submitted in Partial Fulfilment

of the Requirements for the Degree of

Doctor of Philosophy

in

Geography and Resource Management

The Chinese University of Hong Kong

September 2011

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

Thesis/Assessment Committee

Professor HUANG, Bo (Chair)

Professor LEUNG, Yee (Thesis Supervisor)

Professor LIN, Hui (Committee Member)

Professor LAM, Nina Siu-Ngan (External Examiner)

Professor WANG, Dong-Gen (External Examiner)

Abstract

of thesis entitled:
Multifractal Analysis of Geographical Structures and Processes: Concepts and Applications
Submitted by ZHOU, Yu
for the degree of Doctor of Philosophy
at The Chinese University of Hong Kong
in September 2011

Following its development by Benoit Mandelbrot, fractal has been used in a large number of studies of a wide variety of geographical phenomena exhibiting complexity. Fractal makes the study of highly irregular and complex structures and processes that defy traditional mathematical analysis feasible. One of the most popular uses of the fractal is to draw attention to the variance of structures and processes across multiple scales (e.g. scaling behavior). The concept of a power-law, which is expressed as a straight line in a double-logarithmic plot, is the most characteristic scaling behavior of fractal to measure self-similarity invariant across multiple scales. However, the single fractal dimension falls short in capturing scale-invariant geographical phenomena for the characterization of their non-linear variation across a wide range of scales. The concept of multifractal was, therefore, introduced to give a more complete description. The fractal dimension was extended to the generalized fractal dimensions. The main objective of this thesis is to improve the theoretical formulation of Multifractal Analysis (MFA) along a number of directions, and to make applications to investigate the self-similarity, long-range correlation or multifractality of some real-life geographical phenomena for substantiation.

Detrended Fluctuation Analysis (DFA) and Multifractal Detrended Fluctuation Analysis (MF-DFA) have become the most popularly used because of their

effectiveness and easy implementation. Actually, MF-DFA is based on DFA, which is designed to calculate the Hurst exponent, H, through the power-law between the square of fluctuations and the corresponding scales. H aims at quantifying the long-range correlation of a process and can be related to the fractal dimension. The generalized Hurst exponents $h(q)$ can be obtained by studying the scaling behavior of the qth moment of the fluctuations. On the conceptual level, MF-DFA and DFA (the base of MF-DFA) are explored in this thesis.

Two problems of DFA and MF-DFA in this study are: firstly, oscillations in the fluctuation function and significant errors in the crossover locations; and secondly, the negative influence of periodic trend on the scaling behavior in DFA. The Multifractal Moving-Window Detrended Fluctuation Analysis (MF-MWDFA) and the more general Multifractal Temporally Weighted Detrended Fluctuation Analysis (MF-TWDFA) are formulated as a solution of the first problem. The second problem is solved by a pre-detrending method on the basis of Empirical Mode Decomposition (EMD) for the elimination of the effect of the periodic trend in DFA. Furthermore, some classical relationships of the exponents in MF-DFA are revisited. This study will rectify the incorrectness of existing results found under some situations, and propose modified relationships to obtain the appropriate characterizations.

In terms of applications, the efficacy of the improved DFA is shown by two real-life examples, namely: temperature variations and sunspot activities. A substantial systematic analysis of the temporal and spatial patterns of the earthquake process is studied at length. As a complement to the inter-event spatial and temporal distance, the epicenter motion direction is investigated by DFA. The scaling behaviors under different conditions (e.g. the threshold magnitudes, boundary effect, random removal of some events, and different seismic zones) are also investigated. At the small scale, there is a general scaling behavior indicating the random process and independence of the different sensitive testing conditions. In the large scaling range, the long-range correlation appears. Furthermore, the behavior on the dependence of different conditions is uncovered.

This thesis, therefore, gives a rigorous and systematic study of geographical phenomena in multiple temporal and spatial scales.

摘要

自Benoit Mandelbrot提出概念，分形已经被广泛的应用到了对地理现象的复杂性研究中。分形概念使得对不能被传统数学分析方法所研究的极不规则、复杂的结构和过程的研究变得可行。多尺度下结构和过程变化的规律，即尺度律，是分形概念中最关心的问题之一。作为分形概念中最重要的一种尺度律，幂定律在双对数图里被表示为一条直线并能估计分形维数来量化对象的自相似性。但是，单一的分形维数常常无法有效而完备的提取出大范围尺度下非线性变异的全部特性。因此，我们引入多重分形的概念对研究对象进行一个更全面的描述，并推广得到多重分形维数。本文意在多重分形的若干个方面开展理论上探讨，从而进一步完善多重分形分析的理论体系，并对实际问题中的自相似性、长程相关性及多重分形特性进行分析和研究。

去趋势涨落分析(Detrended Fluctuation Analysis (DFA))和多重分形趋势涨落分析(Multifractal Detrended Fluctuation Analysis (MF-DFA))凭借其高效和易操作性成为了最为常用的分形和多重分形方法。通过涨落函数的平方和相应尺度之间的幂定律，Hurst指数，H，可被DFA计算来量化长程相关性并建立其与分形维数的关系。同时，DFA和Hurst指数还可以推广到更一般的MF-DFA和$h(q)$。本文将在理论层面上研究和讨论MF-DFA及DFA。

本文关注的DFA和MF-DFA的两个问题为：涨落函数的摆动和由此带来的在拐点定位时的误差；以及周期趋势对DFA尺度律的负面影响。对这两个问题，本文提出了多重分形滑动窗口去趋势涨落分析(Multifractal Moving-Window Detrended Fluctuation Analysis (MF-MWDFA))和更一般的多重分析时间加权去趋势涨落分析(Multifractal Temporally Weighted Detrended Fluctuation Analysis (MF-TWDFA))；以及基于经验模式分解(Empirical Mode Decomposition (EMD)) 的预处理。同时，我们在理论层面上研究了MF-DFA中

的几个经典关系的潜在问题，并由此提出对了这些关系的修正。

在应用部分，改进的DFA的有效性很好的体现在了对气温变化和太阳黑子活动的研究中。此外，在地震过程的时间和空间分布模式的系统分析中根据现有研究主要集中在相邻地震事件之间的时间和空间距离上的情况，提出了对地震迁移方向进行研究来补充现有结果。DFA得到的尺度律以及在几个不同方面，即震级阈值、边界条件、完整性和不同地震区域，的敏感性分析表明：在小尺度上，地震迁移方向具有随机性且不依赖于上述设定条件；而大尺度上，则表现出了长程相关性和对震级阈值的依赖性。

综上所述，本文的主要目的是对地理问题进行精细而系统的时间和空间上的多尺度研究。

Acknowledgements

First and foremost, I would like to thank Prof. Yee Leung, my supervisor, for the opportunity he granted me to study at The Chinese University of Hong Kong. I am most grateful to Prof. Leung for his guidance, advice, and availability on all days of the year. During my Ph.D. study, Prof. Leung always shares with me his deep insight, valuable research experiences, and percipient interdisciplinary vision, which are great inspiration in training me for my future professional development in scientific study. I still remember how Prof. Leung revised my manuscripts word by word by hand. The way to find and form research problems, to capture key points of a problem, and to find solutions for them are things Prof. Leung has taught me and will be of great benefit to my future research career. The most important lesson learnt during my Ph.D. study is to recognize the importance of research ethics, trust, and keeping promise, traits crucial to not only being a researcher but also a responsible person. For me, Prof. Leung is an exemplary example of being a distinguished researcher and a respectable person. Frankly, I am not an easy graduate student to direct. I am grateful to Prof. Leung's understanding and forgiveness when I made mistakes, particularly a serious one. His saying: 'One might gain something in the short run by breaking the code of practice, but one would lose out in the long run' will always sit in my mind.

I am deeply indebted to Prof. Zu-Guo Yu (Xiangtan University and Queensland University of Technology), my M.Sc. supervisor, for his continuous encouragements when I was upset and help when I needed. It is my luck to have such a nice supervisor. Special thanks go to Prof. Lung-Sang Chan at The University

of Hong Kong for his provision of data and valuable suggestions on earthquake study in my thesis, to Prof. Qiang Zhang (Zhongshan University) for letting me know the importance of efficiency in doing research, and to Prof. Qi-Hua Ran (Zhejiang University) for making me believe that in current society full of kinds of 'relationships' one can still be recognized only if working hard.

I have to express my deep appreciation of the valuable comments made by my internal committee members, Prof. Bo Huang and Prof. Hui Lin, and my external examiners, Prof. Nina Siu-Ngan Lam (Louisiana State University) and Prof. Dong-Gen Wang (Hong Kong Baptist University), on improving the quality of my thesis.

Technical supports for my research, such as the computational environment and hardware, from Mr Hon-Tat Lee should also be specially thanked.

Ph.D. study is not an easy life. I can not brave through the whole thing without the supports of my parents and friends. Thanks also go to my beloved officemates in Room 221A. Particular thanks go to Miss Xue-Hua Gao, Chun-Lan Guo, and Mr Si Gao, Yun Li, Ming Luo, Yong Xu and Wei Zhang. In the summer of 2011, it is time to say goodbye. Wish you all good luck and all the best. I hope we can continue our friendship in the years to come.

Contents

Abstract i

摘要 iv

Acknowledgements vi

Table of contents viii

List of figures xi

List of tables xvi

Abbreviation xvii

1 Introduction 1

1.1 Basic Concepts and Introduction of Fractal and Multifractal . . . 1

1.2 Research Problems 8

1.3 Research Objectives 12

1.4 Research Significance 13

1.5 Organization of the Thesis 16

2 Research Framework 17

2.1 Introduction 17

2.2 The Research Framework 18

2.3 Summary 21

3 Review of Relevant Methodology and Discussions 22

3.1 Methods of Fractal Analysis . . . 23

3.1.1 Box-Counting Dimension . . . 26

3.1.2 Rescaled Range (R/S) Analysis . . . 27

3.1.3 Two-Dimensional (2D) R/S analysis . . . 28

3.1.4 Fluctuation Analysis (FA) . . . 29

3.1.5 Detrended Fluctuation Analysis (DFA) . . . 29

3.1.6 Detrended Moving Average (DMA) . . . 31

3.1.7 Fourier Power Spectral Analysis . . . 32

3.1.8 Structure Fluctuation (SF) . . . 33

3.1.9 Preferred Method . . . 33

3.2 Methods of Multifractal Analysis . . . 34

3.2.1 Partition Function-Based Formalism . . . 35

3.2.2 Multifractal Detrended Fluctuation Analysis (MF-DFA) . 37

3.2.3 2D MF-DFA . . . 40

3.3 Several Problems in DFA and MF-DFA . . . 41

3.3.1 Strong Fluctuation in Scaling Behavior . . . 41

3.3.2 Influence of Periodic Trend . . . 43

3.3.3 Problematic Relationship in MF-DFA . . . 45

3.3.4 Problematic Relationship Between H and $h(q)$ in the 2D MF-DFA . . . 51

3.4 Summary . . . 54

4 Methodological Investigation of DFA and MF-DFA 57

4.1 Improvement of DFA . . . 57

4.1.1 Strong Fluctuation in Scaling Behavior . . . 58

4.1.2 Influence of Periodic Trend . . . 70

4.2 Correction of Problematic Relationships in MF-DFA . . . 85

4.2.1 Problematic Relationship Between $h(q)$ and $\tau(q)$ in MF-DFA 85

4.2.2 Problematic Relationship Between H and $h(q)$ in the 2D MF-DFA . . . 90

4.3 Summary 91

5 Application to Real-life Geographical Examples **93**

5.1 Performance of Multifractal Temporally-Weighted Detrended Fluctuation Analysis (MF-TWDFA): Application in Air Temperature Study 93

5.1.1 Results 96

5.1.2 Interpretations and Discussions 104

5.2 Performance of EMD-Based Pre-Detrending Processing: Application in Sunspot Series 108

5.2.1 Basic Results 109

5.2.2 Further Discussion 116

5.3 Summary 124

6 Application to Earthquake Problems **126**

6.1 Introduction 126

6.2 Epicenter Migration 130

6.3 Summary 141

7 Conclusion **143**

7.1 Main Results 143

7.1.1 TWDFA and MF-TWDFA 144

7.1.2 EMD-Based Method 146

7.1.3 Problematic Relationship in 1D MF-DFA 147

7.1.4 Problematic Relationship in 2D MF-DFA 148

7.1.5 Universal Scaling Structure of Epicenter Migration 149

7.2 Research Limitations and Suggestions for Further Research 150

Bibliography **152**

List of Figures

1.1 von Koch curve. 2

1.2 Numerical examples of non-stationary and stationary cases, fBm (upper panel) and fGn (bottom panel). 3

1.3 Daily mean temperature in Hong Kong from 2005 to 2007 (upper panel), and the monthly number of sunspots during the period 1749 to 2009 (with 3123 months) (bottom panel). 5

1.4 Distribution of epicenters in south China from 1970 to 1995. 6

1.5 Magnitude series of earthquake records in south China from 1970 to 1995. . . 6

2.1 The framework for multifractal analysis in this thesis. 19

3.1 $h(q)$ for fractional Gaussian noise and temperature record against q. 38

3.2 An enlarged plot of part of the fGn ($x - < x >$), for s=3. 48

3.3 Relationships between $h(q)$ and $\tau(q)$ for numerical examples with H=0.1 to 0.9. (from top left to bottom right). 50

3.4 Results of 2D R/S analysis using 2D fBm with size 256×256 and known Hurst exponent $H = 0.2$(left panel), 0.5(center panel) and 0.8(right panel). 52

3.5 Results of the 2D R/S analysis using 2D fBm with size 512×512 and $H = 0.2$ (left panel) $H = 0.5$ (center panel) $H = 0.8$ (right panel). 52

3.6 Results of 2D MF-DFA for 2D fGn with $H = 0.1$ to 0.9 (from top left to bottom right). 54

3.7 Results of 2D MF-DFA for 2D fBm with $H = 0.1$ to 0.9 (from top left to bottom right). 55

4.1 Two local moving windows (MW) along the air-temperature series and the corresponding fitting polynomial $\hat{Y}$ in these moving-windows 59
4.2 $h(q)$ of the binomial multifractal series with length 10^{12} and a ranging from 0.7 (upper left-hand panel) to 0.85 (lower right-hand panel). 63
4.3 Results of MF-TWDFA and MF-DFA using binomial multifractal series with length 10^{12} and a ranging from 0.7 (upper left-hand panel) to 0.85 (lower right-hand panel). The solid line is the theoretical line derived from Eq. (4.12) and is presented here as benchmark. 65
4.4 Results of MF-TWDFA and MF-DFA using fGn (upper panel) fBm (bottom panel) with length 2000 and Hurst exponent $H = 0.8$ 66
4.5 The constructed series (upper panel) and its scaling behavior (bottom panel), the marked slope 0.92 is the estimated Hurst exponent. 72
4.6 The superimposed series (upper panel) and its scaling behavior (bottom panel). 73
4.7 The EMD results of the superimposed series (upper panel) and the scaling behavior from removing $\sum_{i=5}^{6} \mathrm{IMF}_i$ and $\sum_{i=5}^{8} \mathrm{IMF}_i + r$, e.g. $\sum_{i=1}^{4} \mathrm{IMF}_i$ (bottom panel). 79
4.8 The scaling behavior of Fourier truncation (upper panel) and the adaptive detrending method (bottom panel). 80
4.9 The EMD results of superimposed series (upper panel) and the scaling behavior of removing $\sum_{i=6}^{7} \mathrm{IMF}_i$ and $\sum_{i=6}^{9} \mathrm{IMF}_i + r$, e.g. $\sum_{i=1}^{5} \mathrm{IMF}_i$ (bottom panel). 81
4.10 The scaling behavior of Fourier truncation (upper panel) and the adaptive detrending method (bottom panel) with different parameters. 82
4.11 Selection of K by the variance method in Hu et al. (2009) 84
4.12 Relationships between $h(q)$ and $\tau(q)$ for numerical examples with H=0.1 to 0.9 (from top left to bottom right). 89

5.1 Daily mean temperature in Hong Kong from 2005 to 2007 96
5.2 MF-DFA and MF-TWDFA results of Hong Kong daily mean temperature from 2005 to 2007 98

5.3 Detrended daily temperature series in Hong Kong from 2005 to 2007 (upper panel) and its profile with the local trends estimated by DFA and TWDFA (bottom panel) 99

5.4 Power spectrum of the profile of detrended daily temperature series, the trend estimated by TWDFA and the detrended profile. For better illustration, the power spectrum of the profile of daily temperature series and the detrended profile have been shifted vertically upwards by 3 (from setting $\log_{10} S(f)$ as $\log_{10}(10^3 \cdot S(f))$ which is equal to $\log_{10} S(f) + 3$) and similarly downwards by 4 at $\log_{10}$ scale respectively. 100

5.5 DFA and TWDFA results of the detrended daily temperature series in Hong Kong from 2005 to 2007 101

5.6 The power spectrum of the original series and the annual-cycle-removed series. The power spectrum of the original series have been shifted vertically upwards by 3 at $\log_{10}$ scale for convenience of comparison and observation. 103

5.7 $h(q)$ of the annual-cycle-removed series against q (upper panel) and the singularity spectrum of this series (bottom panel) 105

5.8 The monthly sunspot time series. 109

5.9 Results of the EMD on the original sunspot series, IMF_1 to r (from top to bottom). 110

5.10 Scaling behavior of sunspot time series by removing the components one by one. 111

5.11 Original data and the fitted series using $\sum_{i=5}^{8} \text{IMF}_i + r$ (upper panel) and the residual $x - (\sum_{i=5}^{8} \text{IMF}_i + r) = \sum_{i=1}^{4} \text{IMF}_i$ (bottom panel) 112

5.12 Compare the power spectra of the detrended series ($\sum_{i=1}^{4} \text{IMF}_i$) and the original series. The vertical straight line corresponds to the frequency of the 11-year cycle, 0.0076. 113

5.13 Effect of each component on the scaling behavior of the sunspot time series. . 114

5.14 The series after eliminating IMF_5 and IMF_6 from the original sunspot series. 114

5.15 The scaling behaviors of the original and the detrended series using DFA. . . 115

5.16 Power spectra of the IMFs of the original sunspot series using Fourier transform, IMF_1 to IMF_4 (from top to bottom). 116

5.17 Power spectra of the IMFs of the original sunspot series using Fourier transform, IMF_5 to IMF_8 (from top to bottom), the bottom panels are the locally enlarged plots of the corresponding upper panels. 117

5.18 Scaling behaviors of the original and the Fourier truncation series (removing the first 50 terms of the Fourier transform of the sunspot time series) using MF-DFA. 119

5.19 Fitted series using Fourier truncation method in comparison with original series (upper panel) and the residual of the fit (bottom panel). 121

5.20 Scaling behavior of the original and Fourier truncation series (removing the first 70 and 18th~70th terms of the Fourier transform of the sunspot time series) using MF-DFA. 124

6.1 Distribution of epicenters in south China from 1970 to 1995. 132

6.2 Distribution of epicenters in southern California during the period 1975 to 1995, which is part of our study period. (This figure is adopted from Godano et al. (1999).) 133

6.3 The scalings of the catalogues with different threshold magnitudes of south China (the upper panel) and southern California (the bottom panel) were obtained by DFA. The selection of threshold magnitudes takes into consideration that there should be sufficient earthquake events left to ensure the statable results of DFA. The solid straight lines are estimated by linear regression. These scalings are shifted vertically for the sake of better illustration. 134

6.4 The scalings of the catalogues with events randomly removed occupying different percentages of south China (the upper panel) and southern California (the bottom panel) obtained by DFA. The scalings without any event being removed are also presented for comparison and are marked by the legend corresponding to '0%'. The solid straight lines are estimated by linear regression. These scalings are shifted vertically for the sake of better illustration. . . . 136

6.5 The scalings of the catalogues with their area zooming in from the boundary by different percentages of south China (the upper panel) and southern California (the bottom panel) were obtained by DFA. The solid straight lines are estimated by linear regression. These scalings are shifted vertically for the sake of better illustration. 137

6.6 The scalings of the epicenter of microshocks of the rock fracture process projecting on different surfaces (the upper panel) and in different analysis conditions were obtained by DFA. The solid straight lines are estimated by linear regression. These scalings are shifted vertically for the sake of better illustration. 138

6.7 The scalings of $\{\theta_i\}$ obtained from shuffled $\{\vec{X}_i\}$ for south China and southern California are obtained by DFA. These scalings are shifted vertically for the sake of better illustration. 139

List of Tables

3.1 Averaged $D(q)$ of vertical shifted fGns for different H and q values 49

3.2 Slopes and intercepts of the linear fits of the curve of $\tau(q)$ and $qh(q)$ for different H values in Fig.3.3 51

3.3 Averaged values of $h(2)$ of the two-dimensional MF-DFA using 2D fGn and 2D fBm with different H values 55

4.1 Averaged $\Delta h(2)$ of MF-TWDFA and MF-DFA using fGns and fBms with different values of the Hurst exponent H and lengths 65

4.2 Calculated right side of the new relationship Eq. (4.38) for numerical examples with different H values in Fig. 4.12 89

5.1 Root mean squared fluctuation around the regression line depicting the scaling behavior by DFA and TWDFA on the original and annual-cycle-removed temperature series 102

5.2 Summary of experimental results. 122

Abbreviation

1D: One-Dimensional

2D: Two-Dimensional

ADM: Adaptive Detrending Method

CSR: Completed Spatial Randomness

DFA: Detrended Fluctuation Analysis

DMA: Detrended Moving Average

EMD: Empirical Mode Decomposition

EMD-HSA: Empirical Mode Decomposition Based Arbitrary Order Hilbert Spectral Analysis

FA: Fluctuation Analysis

fBm: fractional Brwonian motion

fGn: fractional Gaussian noise

G-R Law: Gutenberg-Richter Law

GTWR: Geographically and Temporally Weighted Regression

GWR: Geographically Weighted Regression

HHT: Hilbert-Huang Transform

IMF: Intrinsic Mode Function

MF-DFA: Multifractal Detrended Fluctuation Analysis

MF-MWDFA: Multifractal Detrended Moving Window Detrended Fluctuation Analysis

MF-TWDFA: Multifractal Temporally-Weighted Detrended Fluctuation Analysis

MFA: Multifractal Analysis

MWDFA: Moving Window Detrended Fluctuation Analysis

OLR: Ordinary Linear Regression

R/S Analysis: Rescaled Range Analysis

RTL: Region-Time-Length

TWDFA: Temporally-Weighted Detrended Fluctuation Analysis

QBO: Quasi-Biannual Oscillation

WTMM: Wavelet Transform Modulus Maxima

Chapter 1

Introduction

1.1 Basic Concepts and Introduction of Fractal and Multifractal

Ever since the development of fractal by Mandelbrot (1982), this concept has been used in a range of applications in a large variety of geographical phenomena exhibiting complexity. Fractals makes the study of highly irregular and complex structures feasible, it also enables the study of those processes that defy traditional mathematical analysis (Feder, 1988; Mandelbrot, 1982). Although a strict definition of a fractal is hard to give, Falconer (1990) suggested that it seems best to regard a fractal as a set **F** that has properties such as:

- **F** has a fine structure (i.e. is detailed on arbitrarily small scales);
- It is too irregular to be easily described in traditional Euclidean geometric language;
- **F** often has some form of self-similarity (perhaps approximate or statistical);
- Usually, the 'fractal dimension' of **F** (defined in some way) is greater than its topological dimension; and,
- In most cases of interest, **F** is defined in a very simple way (perhaps recursively).

Among these properties, self-similarity is the most important concept which is used in fractal. This concept means that a fractal is a shape made of parts similar to the whole in some way (Mandelbrot, 1982). It is not difficult to find examples of fractal-like behavior in nature. In some examples self-similarity in nature can be detected by human eyes; for example, in a von Koch curve, which is shown below in Fig. 1.1: However, the self-similarity of observed objects is not usually

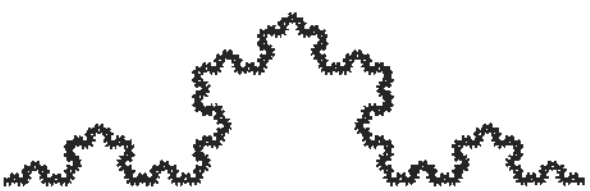

Figure 1.1: von Koch curve.

able to be detected by sight alone. In these cases the self-similarity exists in the statistical sense. Two well-known numerical examples of this sort of self-similarity are illustrated in Fig. 1.2. The first is fractional Brownian motion (fBm) and the second is fractional Gaussian noise (fGn) (they are given as non-stationary and stationary examples, respectively).

In the geographical world there are a large number of temporal and spatial examples which exhibit self-similarity. This thesis will study self-similarity in a number of geographical cases (such as temperature variation, sunspot variability, and the epicenter distribution) in addition to the numerical examples. Some of the research objects are presented in the illustrations below without a detailed description in order to give some intuitional impression of the nature of this re-

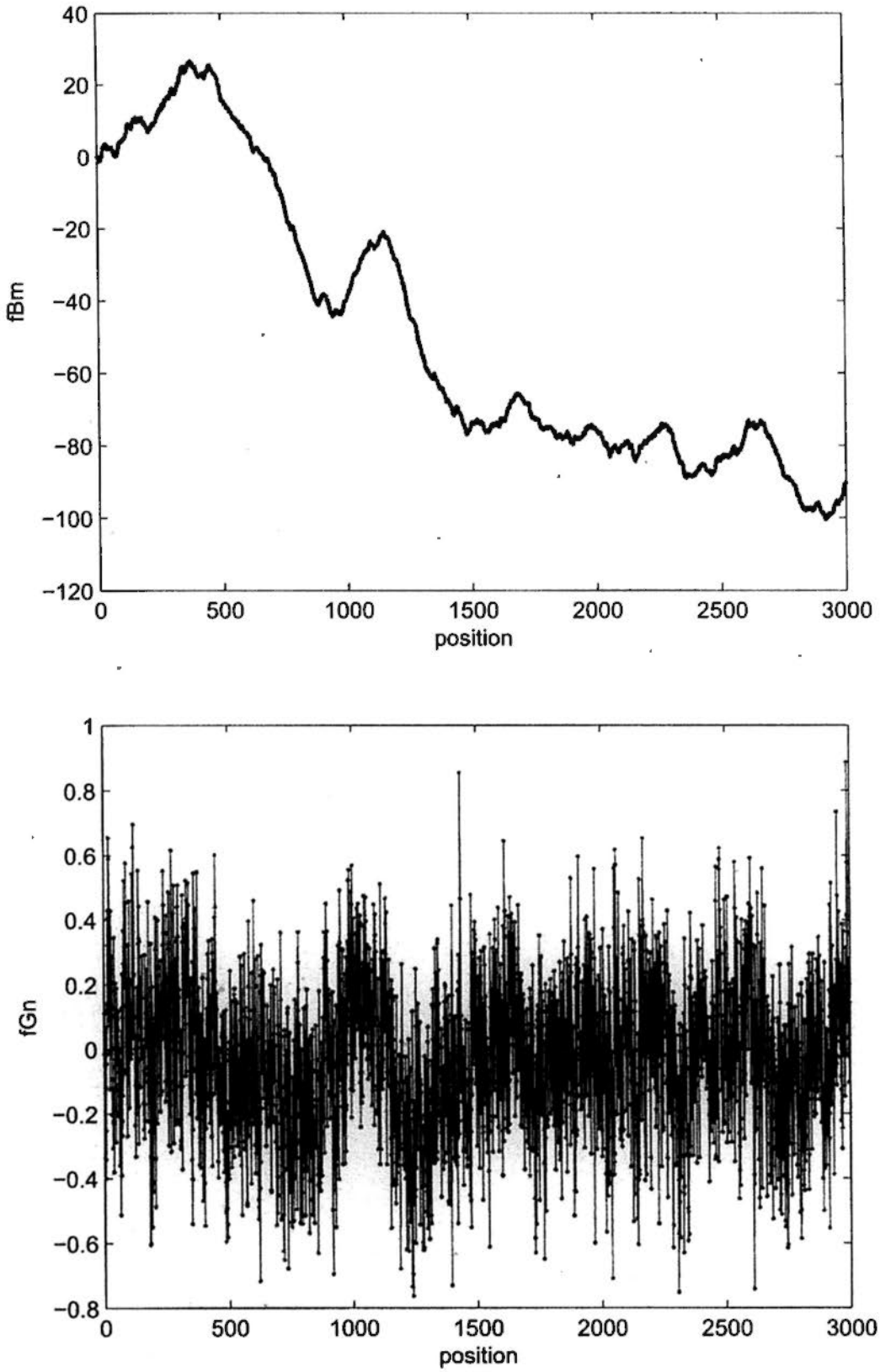

Figure 1.2: Numerical examples of non-stationary and stationary cases, fBm (upper panel) and fGn (bottom panel).

search project. The daily mean temperature records and monthly mean sunspots are shown in Fig. 1.3. The seismic examples are also offered in the way of spatial distribution and time series in Figures 1.4 and 1.5.

Theoretically, if the self-similarity exists (at least in a statistical sense) then there is a corresponding formula which is given as:

$$M_s(\mathbf{F}) \sim c \cdot s^{-D_f}, \tag{1.1}$$

Where s is the considered scale and M_s is the corresponding measurement. This formula is called the power-law and it is represented as a straight line in the double logarithmic plot. Then D_f is determined from the invariance across multiple scales in the power-law, and it is named as a fractal dimension of explored structures or processes. Three common fractal dimensions are: the Hausdroff dimension, the box-counting dimension, and the packing dimension (Falconer, 1990; Feder, 1988). Given a set $\mathbf{F}$, covering it by 'balls' whose sizes could be different but have to be less than s, then the corresponding d-measure $M_{d,s}$ is defined as the infimum (e.g. the minimal values of all possibly obtainable coverings). The Hausdroff dimension is the value of d leading the limit of $M_{d,s}$, corresponding to the vanishing diameter or size s of the 'balls' used to cover the set in changing from the infinite to zeros (Falconer, 1990; Feder, 1988). With regard to the box-counting dimension: by using boxes with size s to cover the set $\mathbf{F}$, and setting the measure M_s as the number, N_s, of balls with a non-empty intersection of the set $\mathbf{F}$, then the box-counting dimension can be obtained from the power-law $N_s \sim s^{-D_f}$ (Feder, 1988). Alternatively, the packing dimension could be obtained by locating the centroid of the non-overlapped boxes which are used to cover $\mathbf{F}$, the points of $\mathbf{F}$ (similarly to the definition of Hausdroff dimension) (Falconer, 1990). All three fractal dimensions are defined to measure the degree of complexity. Generally speaking, the larger values of fractal dimensions indicate a higher degree of complexity. Among the three fractal dimensions, the box-counting dimension has become the most popular fractal dimension which is used in real computation due to the simplicity of its calculation. Compared

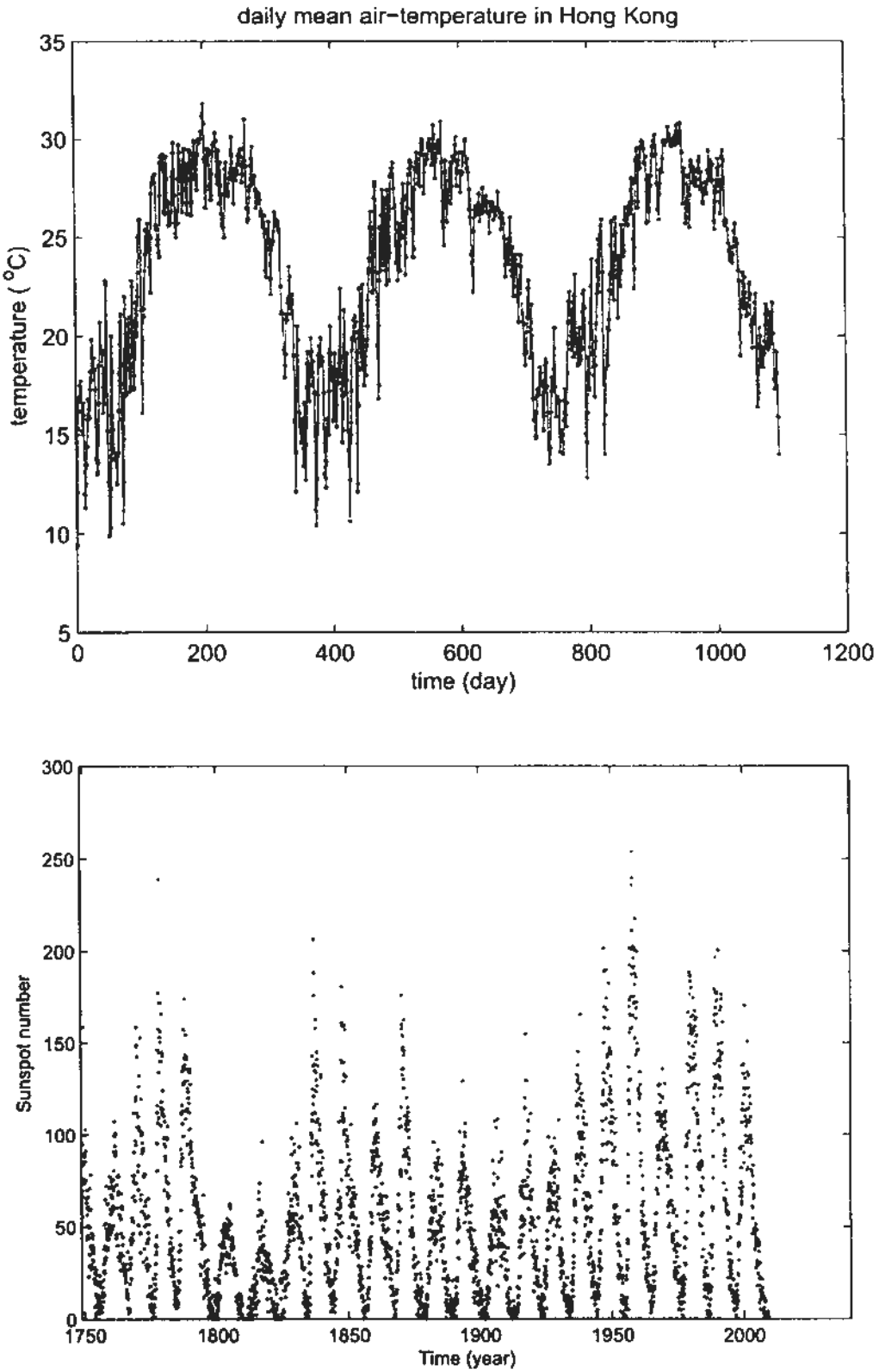

Figure 1.3: Daily mean temperature in Hong Kong from 2005 to 2007 (upper panel), and the monthly number of sunspots during the period 1749 to 2009 (with 3123 months) (bottom panel).

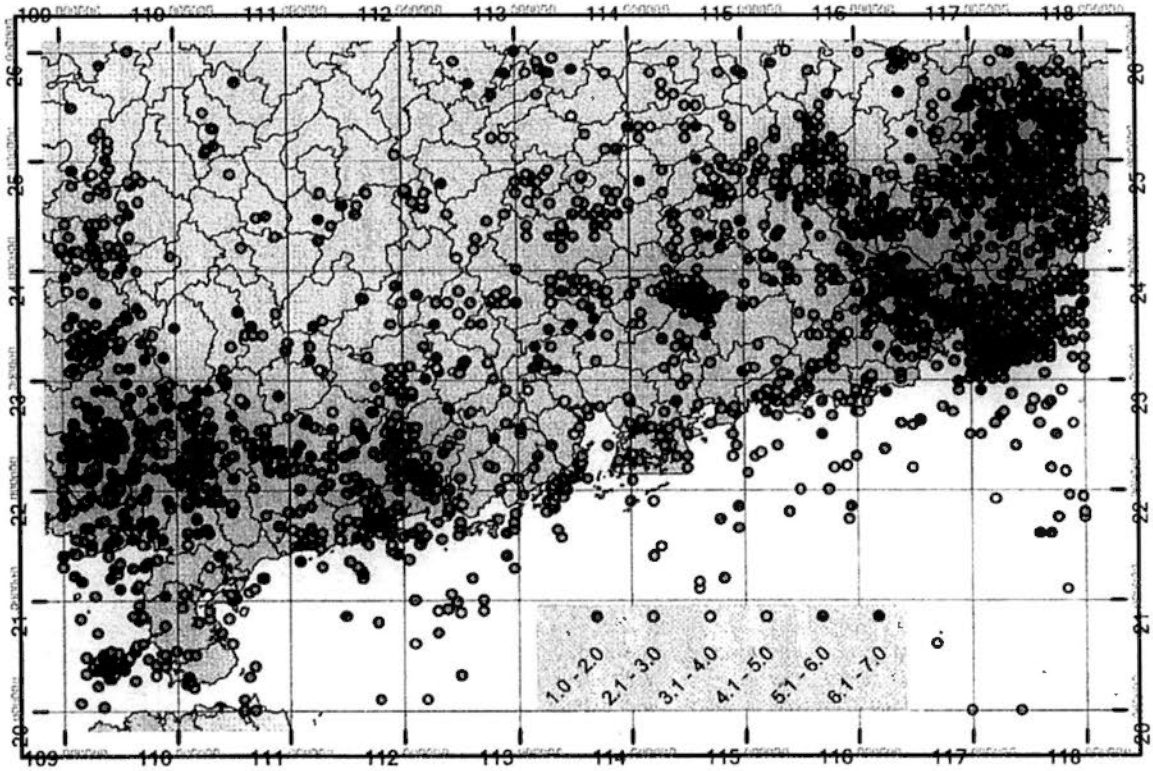

Figure 1.4: Distribution of epicenters in south China from 1970 to 1995.

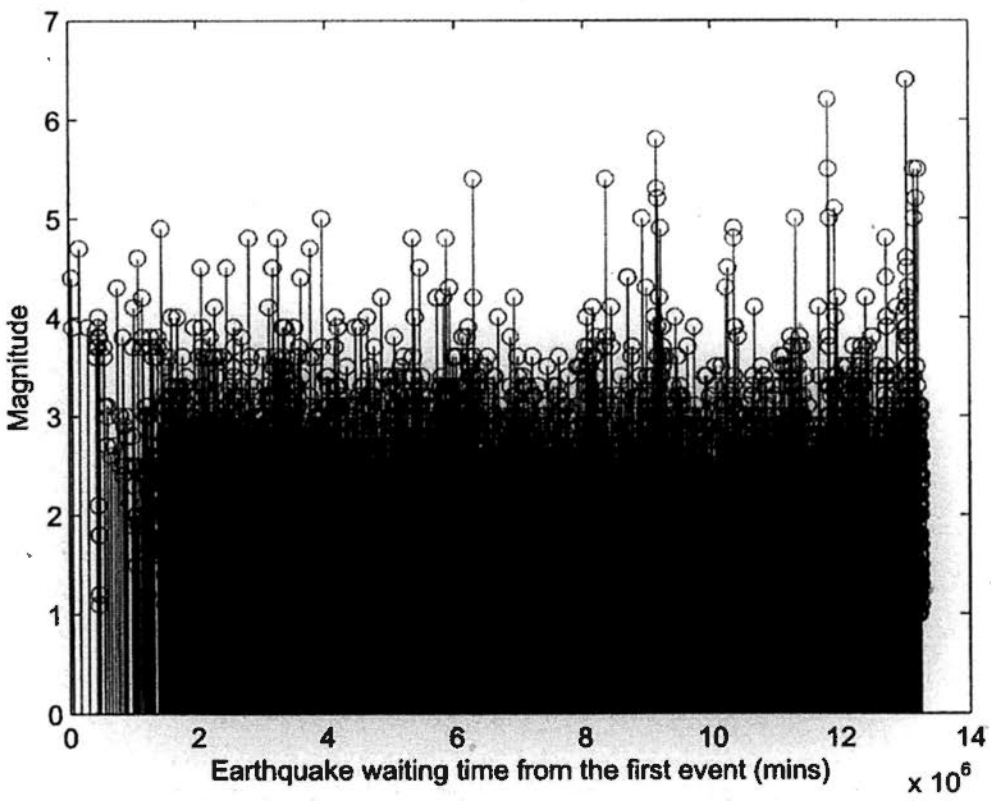

Figure 1.5: Magnitude series of earthquake records in south China from 1970 to 1995.

with the traditional topological dimension D_T, only taking the integer values D_f (which can be either integer or fraction) could give a more accurate measurement of the investigated processes. The larger value of D_f corresponds to a higher degree of complexity and irregularity. Thus D_f provides a route to handle complex and irregular structures and processes.

There are some other related concepts which enable us to understand complexity and irregularity besides the fractal dimension. For example, the long-range correlation was proposed by Hurst (1951) and it can also be determined by the concerned power-law. This correlation is related to the auto-correlation function which decays following a power-law rather than decaying exponentially (Rybski et al., 2006, 2008). The positive or negative long-range correlation means that the current pattern would be more likely to be maintained or broken in the future. The Hurst exponent (H) is an important exponent which was first proposed by (Hurst, 1951) to quantify this long-range correlation. Furthermore, a relationship between D_f and H has been established as (Falconer, 1990):

$$D_f = D_T + 1 - H, \tag{1.2}$$

Here D_T is the topological dimension of the support of the fractal.

However, the fractal dimension falls short in capturing scale-invariant geographical phenomena (such as temperature variations, sunspot activities, and the changed direction between successive earthquake events) which require an infinite number of scaling exponents instead of one single dimension for the characterization of their non-linear variation across a wide range of scales. Thus, it is necessary to introduce the concept of multifractal by considering the multiple qth order moments of the concerned measure (or fluctuation) rather that just analyzing a single order moment in the fractal analysis. Consequently, in this generalized conceptual framework the fractal dimension D_f and Hurst exponent H have been extended to a generalized fractal dimension $D(q)$ (Feder, 1988) and a generalized Hurst exponent $h(q)$ by Kantelhardt et al. (2002). In a one-dimensional situation (e.g. for time series study) Kantelhardt et al. (2002)

established the relationship connecting $D(q)$ to $h(q)$ as:

$$\tau(q) = D(q) \cdot (1 - q) = q \cdot h(q) - 1, \tag{1.3}$$

Where $\tau(q)$ is the mass exponent in the partition function based multifractal formalism (Halsey et al., 1986). Hence, the multifractal analysis (MFA) can be performed with the generalized Hurst exponent $h(q)$ as a bridge.

The research problems involved in this thesis will be concluded and presented in the following section based on the discussion above.

1.2 Research Problems

This thesis has two sections, which are: conceptual discussions and applications. The core aim of this thesis is to analyze the different fractal and multifractal properties of the numerical examples, geographical structures, and processes. The research problems are presented below.

On a conceptual level, five problems will be discussed:

1. **Selection of the most appropriate method for exploring the fractal and multifractal properties of geographical problems.** There are many algorithms which have been developed for fractal and multifractal analyses. Each method has its advantages and disadvantages. Selection of which method is to be used should be based on a complete review and careful comparison of the existing popular methods. Since the long-range correlation and generalized Hurst phenomena are related to fractal and multifractal analysis, they should be included in this thesis. Among the fractal and multifractal analysis, the most important task is to study the power-law; however, there are some trends in geographical phenomena which have a negative influence on the scaling behavior of fractal and multifractal analysis (e.g. the global warming in temperature records and the periodic trend in some of the geographical records such as the 11-year-cycle in sunspot series and the annual cycle in a streamflow series). The

capability of Detrended Fluctuation Analysis (DFA) (Peng et al., 1994) to eliminate the influence of these kinds of non-stationary trends makes it the preferred method to be used to study the long-range correlation of geographical processes in this thesis. A description of the detailed comparison between DFA and other algorithms, together with relevant discussions, will be given in the following chapters. From this comparison it was found that the Multifractal Detrended Fluctuation Analysis (MF-DFA) (Kantelhardt et al., 2002) and its basis (i.e. DFA) are the preferred methods for MFA in this thesis. However, despite their numerous advantages DFA and MF-DFA are not entirely perfect methods. After careful investigation two significant disadvantages of DFA and MF-DFA have been identified in this thesis, they are:

2. **Handling of oscillations in the fluctuation function.** One disadvantage of DFA and MF-DFA are the oscillations in the fluctuation function and significant errors in crossover positioning which are introduced in actual implementations due to the removal of local estimated discontinuous polynomial trends in DFA and MF-DFA. Such strong fluctuations can lead to difficulty in estimating the scaling exponents from power-law and detecting the possible crossover points. Usually, the crossover points indicate a phase change in the underlying dynamics across the scales (Hu et al., 2001). Therefore detection of the crossover points is very helpful for us to understand the studied processes.

3. **Handling of periodic and quasi-periodic trends.** The other disadvantage of DFA and MF-DFA is the negative influence of periodic trend on the scaling behavior of DFA. Previous studies have shown the possibility of spurious crossover points which are caused by the periodic or quasi-periodic trends (Hu et al., 2001). To obtain the genuine scaling behavior, it is necessary to handle these kinds of trends.

4. **Correcting the classical relationship connecting $h(q)$ to $\tau(q)$.** A

potential problem in a classical relationship connecting $h(q)$ to $\tau(q)$ as expressed in Eq.(1.3) in MF-DFA for one-dimensional case. Through this relationship, the conventional MFA based on $\tau(q)$ can be performed from an MF-DFA aspect. A large number of studies have employed this relationship in their MFA (see for example in Kimiagar et al. (2009); Movahed and Hermanis (2008); Movahed et al. (2006); Telesca et al. (2004c, 2005)); however, in this thesis our investigation shows that this relationship has a potential problem in that a correction of this important relationship is expected.

5. **Correcting two relationships of two-dimensional MF-DFA.** Potential problems in two relationships of MF-DFA for study in two-dimensional space: $h(q) \equiv H$ for fBm and, generally, $h(q = 2) \equiv H$. Since the inception of MF-DFA, the focus has been placed on signals in one-dimensional space. Gu and Zhou (2006) recently extended MF-DFA to higher dimensions and also related the generalized Hurst exponent, $h(q)$, in higher dimensions to the original Hurst exponent H. These two relationships have also been extended by Gu and Zhou (2006) to two-dimensional versions; however, these relationships again appear to be invalid from the empirical and numerically experimental perspectives.

There are three problems with regard to the role of applications, temperature variations, sunspots activities, and the general scaling law in the directional analysis of epicenter migration which are briefly introduced as follows:

1. **The long-range temporal correlation and critical scales of temperature variation.** Recently, much attention has been given to the long-range temporal correlations of atmospheric phenomena (Pattantyús-Ábrahám et al., 2004). Temperature, as an important indicator reflecting changes in the atmosphere, has also attracted considerable interest (Eichner et al., 2003; Fatichi et al., 2009; Fraedrich and Blender, 2003; Koscielny-Bunde et al., 1998; Lennartz and Bunde, 2009; Orun and Kocak, 2009;

Pattantyús-Ábrahám et al., 2004; Talkner and Weber, 2000). In terms of the critical scales of temperature activities, there are some well-known scales in climatology (e.g. the 'general weather regimes' or 'Grosswetterlagen', the monthly and seasonal scale corresponding to the usual scale of El Niño, the climate anomaly, and seasonal scales); however, only the weekly scale has been uncovered by the DFA in previous studies. Does this then mean that the other scales in the temperature variation are unimportant? Or, is this due to the drawback of the use of DFA and MF-DFA?

2. **The long-range correlation of sunspots activities without the influence of 11-year cycle.** Due to the negative influence of periodic trend on the scaling behavior obtained by DFA, Movahed et al. (2006) and Hu et al. (2009) have derived two totally different results using DFA for the same sunspot time series. This conflict has arisen because different methods have been selected to handle the effect of the 11-year cycle. The question then is, whose results capture the nature of the sunspots activities?

3. **A systematic analysis of the temporal and spatial patterns of earthquake process.** Most previous studies have almost always focused on the event-betweenness temporal and spatial distance; however, in comparison the directional information has attracted considerably less interest. From this perspective, we would like to see if any natural mechanism could be found for the better understanding of earthquake processes. Three catalogues of directional information will be investigated in this thesis (i.e. south China, southern California, and the experimental microshock database).

Among these three problems the first two are actually employed to test the efficiency and capability of the modified DFA and MF-DFA when handling the real-life data and, in particular, those corresponding to the second and third research problems in the conceptual part.

1.3 Research Objectives

On the basis of the above discussions on the research problems, this thesis can be seen to render a rigorous and systematic study of geographical phenomena in multiple temporal and spatial scales. In order to do this the discussions on conceptual level will be used as the basis of the case studies. On the conceptual level, this study will aim at making some improvements in methodology and clarifying some problematical relationships. The objectives of the case studies are to test the validation of the modified methods and better understand the complicated temporal-spatial geographical structures and processes from some new research aspects.

The objectives of this thesis can be outlined as:

1. This thesis will select the most appropriate methods to be used for the fractal and multifractal analyses according to the specialities of the geographical problems (such as the influence of noises or trends). The preferred methods which are used are expected to capture the natural complexity and characteristics of the concerned geographical processes and structures, which are usually covered by unknown noises or trends.

2. This thesis will attempt to reduce the strong fluctuation in scaling behavior of DFA and MF-DFA in order to obtain a better scaling law for estimating the scaling exponents and positioning the crossover points. Modified methods are expected to improve the performance of DFA and MF-DFA. Then the improved DFA and MF-DFA are going to be used to detect the critical scales in temperature variations. Although these scales are already known as important scales in climatology, they are not revealed by conventional DFA and MF-DFA.

3. This thesis will develop a new method to not only remove the periodic or quasi-periodic trends but also to maintain the intrinsic scalings of the studied processes. This newly-developed method will be compared with existing methods in handling a challenging problem (i.e. the long-range

correlation of sunspot activities without the influence of the 11-year cycle). The results obtained by the proposed method are going to be compared with the inconsistent results which are reported in several previous studies (Hu et al., 2009; Movahed et al., 2006). The real long-term memory of sunspots records may then be determined.

4. This thesis will demonstrate the invalidity of the relationships in MF-DFA in one- and two-dimensional space from empirical, experimental, and theoretical aspects. The corrected relationships will be proposed via the formal and empirical derivations. Their validation will then be tested by constructed numerical experiments. This study's hope is that these corrected relationships will provide more accurate routes to grasp the nature(s) of processes.

5. This thesis will perform a directional analysis in a more systematic case study of epicenter migration. The directional information here is particularly noteworthy in epicenter migration, from either seismological or mathematical aspects; however, to date little academic attention has been given to this topic. This thesis's study of this aspect will be able to offer a supplement for a better understanding of complicated earthquake processes.

1.4 Research Significance

In this thesis the methods for the fractal and multifractal analysis, especially the DFA and MF-DFA, are going to be carefully investigated. It will first make some improvements and corrections to the development of a methodology. It will then detail a number of examples of the application of this theory in real-life geographical problems.

The following achievements exemplify the research significance of this thesis:

- Conceptual aspect:

  1. This thesis will thoroughly review the mainstream algorithms for frac-

tal and multifractal analyses. The advantages of the DFA and MF-DFA for analyzing the geographical structures and processes will also described.

2. An enhanced version is proposed against the strong fluctuations, especially at large scales, in the scaling behavior of the DFA and MF-DFA. A better scaling law can be obtained using this new method. This improvement can ensure the better performance of DFA in the computation of the scaling exponents and detection of crossover points. These exponents and crossover points are helpful for understanding the effects of the complicated dynamics across multiple scales.

3. An adaptive method for estimating and eliminating the periodic trends is employed in this thesis in order to avoid the negative influence of such trends on the outcome of DFA and MF-DFA. When compared to the existing popular methods the proposed method is equivalent in removal of the periodic trends, but it is considerably more easy to implement. Its good performance in analyzing the numerical examples is shown in this thesis.

4. The incorrectness of a classical relationship connecting $h(q)$ to $\tau(q)$ in one-dimensional MF-DFA is pointed out. Such a relationship has been applied in a significant number of previous studies; however, its validation has been found to be limited in some special situations. This study will formally propose a new relationship and demonstrate its correctness using numerical experiments and empirical studies. This corrected relationship extends the classical relationship to a more general version and, therefore, it could be of great benefit to the MFA using MF-DFA.

5. Extending the one-dimensional MF-DFA, Gu and Zhou (2006) gave two relationships: $h(q) \equiv H$ for fBm in the two-dimensional space and $h(q = 2) \equiv H$, in their development of the two-dimensional MF-DFA; however, these generalized relationships are inconsistent with that in

one-dimensional space. With the aid of numerical experiments, this thesis will suggest the more logical and reasonable relationships which can well calculate the Hurst exponent using the two-dimensional MF-DFA.

- Applications:

  1. In the application of improved DFA and MF-DFA in exploring the long-range correlation of temperature variations, compared with the results of conventional DFA and MF-DFA, a better scaling law is obtained and consequently two more critical scales which have been recognized in climatology are detected. Such performance, prior to that of conventional DFA and MF-DFA, substantiates the improvement of the modified versions.

  2. The adaptive periodic-trend-remove method facilitates the elimination of the negative influence of an 11-year cycle in the study of sunspot activities. The real aspect of long-range correlation of sunspots activities is uncovered on the basis of this method. In addition, the incorrect selection of parameters in the application of Movahed et al. (2006) is also found.

  3. In terms of the earthquake process, although some work has already been done by others about the epicenter motion focus on the waiting time and jump distance between two successive earthquake events, the temporal and spatial distance alone can not determine the temporal-spatial position of the epicenters if we consider the epicenter motion under the polar coordinates. Thus, we can take the changed motion direction between two events into consideration as a supplement work to the current studies in earthquake process; this could shed a light on our understanding of the epicenter motion. By exploring the long-range correlation of the changed direction series a universal scaling behavior can be found to exist in the first scaling range, while non-

universal behavior in the second scaling range reflects the different geological structures.

1.5 Organization of the Thesis

This thesis is organized as follows: the research framework is given in chapter 2. In chapter 3, DFA and MF-DFA are selected as preferred methods to be used on the basis of the review and comparison of the exiting methods for fractal and multifractal analysis (their disadvantages and incorrectness are also pointed out here). Against the weakness and problems of DFA and MF-DFA, the corresponding improvements and correctness are proposed in chapter 4. The numerical experiments are also constructed in chapter 4 in order to test these improved methods and corrected relationships. Applications in two real-life examples, namely temperature variation and sunspots activities, are presented in chapter 5 in order to show how the improved DFA and MF-DFA work. Chapter 6 systematically analyzes the occurrence of earthquakes. Finally, this thesis will be summarized in chapter 7.

Chapter 2

Research Framework

Following the introduction of the research problems, objectives, and significance in the previous chapter, the framework for this study is proposed in this chapter.

2.1 Introduction

Almost all geographical structures and processes are highly complicated and irregular and, therefore, they go beyond the capabilities of conventional mathematical analysis. This study intends to analyze these geographical structures and processes, and to do so it will introduce the concepts of fractal and multifractal analysis. Consequently, this study will render a rigorous and systematic study of geographical phenomena at multiple temporal and spatial scales. This study will develop an overall framework which integrates both: discussions on the concepts of fractal and multifractal analysis, and the application of fractal and multifractal analyses in a number of case studies. The structure of this framework is presented in Fig. 2.1.

In this thesis the discussion of the concepts and case studies is organized according to the structure of this framework. After selecting the most appropriate method to be used for the fractal and multifractal analyses of the geographical problems, this study then worked on the weakness and problematic relationships

of these methods. Numerical experiments, theoretical discussions, and empirical derivations are employed to analyze and handle these weaknesses and problems. The improved methods are then applied to two challenging geographical problems to test their efficacy. The application part includes, but is not limited to, the above two examples. A systematic study on earthquake migration is presented in this thesis as part of the application of the theory in a number of case studies.

2.2 The Research Framework

The research framework which is illustrated in Fig. 2.1 shows that two main sections are included in this study, namely: 'concepts' which focuses on the discussions on methodology, and the 'application' of the geographical problems in the case studies.

In terms of concepts, the discussion will firstly focus on how to select the appropriate algorithms for fractal and multifractal analyses by taking the special features of the geographical problems into account. On the basis of this, DFA and MF-DFA have been selected as the preferred methods to be used in this thesis for fractal and multifractal analyses. However, two disadvantages and two problems of the use of these methods have been identified. As shown in Fig. 2.1, there are four further sections corresponding to the problems in DFA and MF-DFA (the improvements and corrections are proposed in these sections, respectively).

The first disadvantage of the use DFA and MF-DFA are the strong fluctuations in their scaling behavior. This is especially problematic at large scales due to insufficient statistical samples. This study will employ the idea of the moving window technique to ensure that there are enough samples to provide stable results. In addition, the autocorrelation which is commonly seen in geographical processes are also considered to enhance the performance of DFA and MF-DFA for better scaling behavior using the idea of Geographical Weighted Regression (GWR). Since what is studied here is time series, the correlation considered is temporal and the modified DFA and MF-DFA are called Temporal Detrended Weighted Fluctuation Analysis (TWDFA) and Multifractal Temporal Detrended Weighted

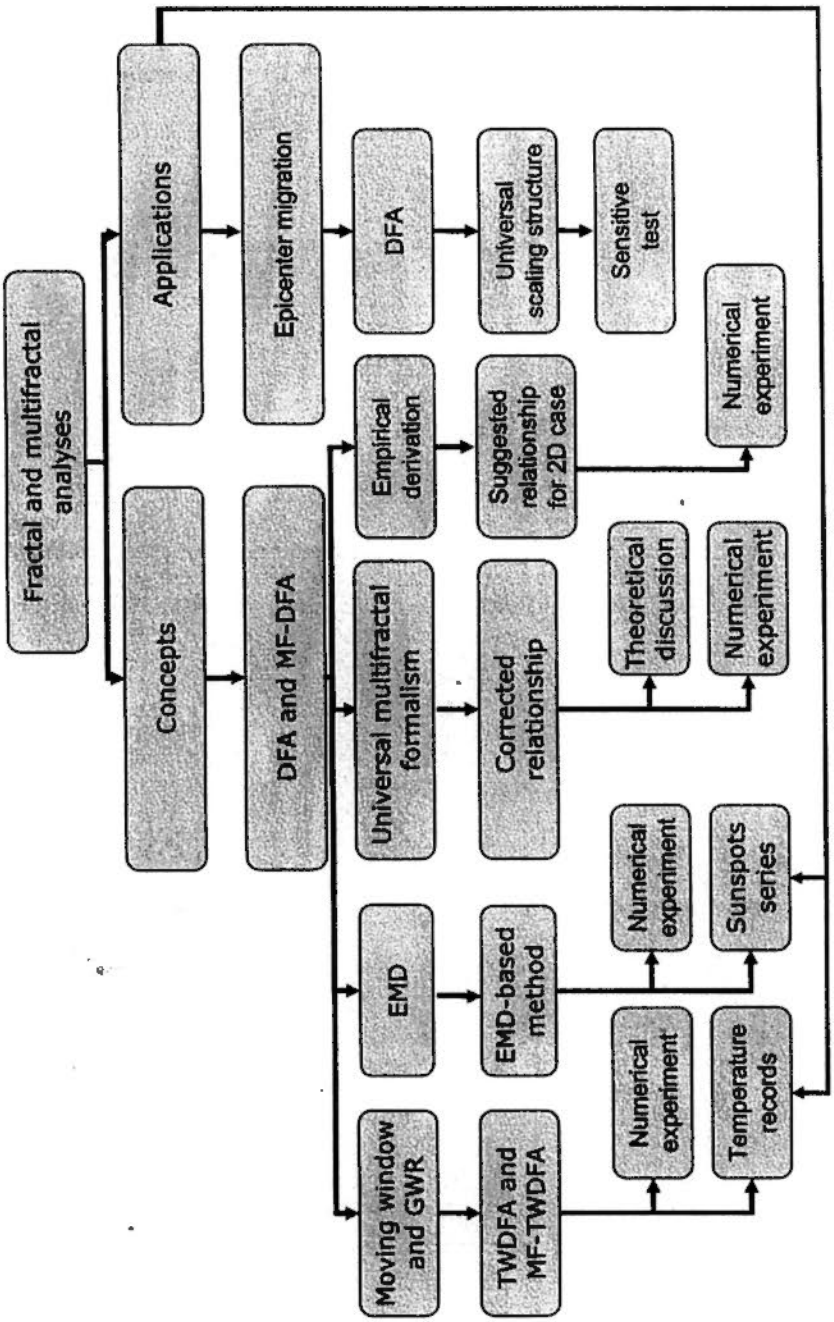

Figure 2.1: The framework for multifractal analysis in this thesis.

Fluctuation Analysis (MF-TWDFA). Numerical experiments with known fractal and multifractal properties are constructed in this thesis for testing the validation of TWDFA and MF-TWDFA. An insignificant difference to the expected results in this study shows that the TWDFA and MF-TWDFA work very well. Furthermore, the performance of TWDFA and MF-TWDFA in this study's analysis of temperature records is better than that of DFA and MF-DFA. Consequently, scaling law with less fluctuations could be obtained. As a result, two more critical scales, which are masked by the strong oscillations in the outcome of DFA, are detected from the better scaling law of TWDFA.

The second weakness of DFA and MF-DFA occurs when analyzing the series with periodic trends. Empirical Mode Decomposition (EMD) is applied here as the pre-processing to eliminate these trends for DFA and MF-DFA. EMD is an adaptive and data-driven method which can decompose the series into many components with different dominant frequencies. The denoised series can be obtained by removing the component(s) corresponding to the targeted periodic trend(s). The EMD-based method is also tested in this study with the designed experimental data. It is then employed to study the controversial issue of long-range correlation of sunspot time series arising from the use of different pre-detrending methods employed to handle the prominent 11-year cycle. By comparing the results of this study with that of previous studies, it is suggested that the inconsistent results could contribute to the incorrect selection of the parameters for the implementation of the detrending algorithm. The conclusion of the long-range correlation of sunspots activities should be about 0.72.

This study will demonstrate the incorrectness of using numerical experiments and theoretical discussions with regard to the two problematic relationships in MF-DFA established to connect different fractal exponents in one- and two-dimensional space. It will then present the corrected relationship for the one-dimensional case through a formal study taking the universal multifractal formalism as a bridge. A suggested relationship for the problem in two-dimensional space is given on the basis of empirical derivation. Several kinds of numerical

experiments are then constructed to test these improvements and their correctness. The corrected relationships are of great benefit to profile the fractal and multifractal properties of the complicated processes using DFA and MF-DFA.

In addition to the applications in the above geographical problems, this thesis will complete a systematic analysis of the temporal and spatial patterns of the earthquake process in order to better understand the earthquake process. For this purpose, a new analysis perspective is proposed in order to study the scaling behavior of epicenter migration in south China, southern California, and a acoustic emission database using DFA. The obtained scaling behavior shows a basic structure which contains both the common property of earthquake process and the local seismic characteristics of the study areas. Different conditions are set to test the sensitivity of the obtained scaling structure, considering the effects of threshold magnitudes, random removal, and the boundary of study area.

It should be noted from the framework that numerical experiments will be employed frequently in this thesis. This is due to the fact that we can construct experimental examples with known or expected properties in these numerical experiments. The performance and capability of the studied methods can then be evaluated by comparing the calculated results with the expected results.

2.3 Summary

This chapter has illustrated the framework which will be used in this thesis. The framework itself has two main sections: the first section describes the theoretical concepts and the second section describes their application in a number of case studies. Some interactions can be found between these two main sections from the framework which is illustrated in Fig. 2.1. These discussions on the research issues can build the basis for the case studies. Meanwhile the applications of proposed methods in geographical problems can be used to show their efficiency and to benefit our understanding of the nature of the geographical world. The next chapter will expand on these ideas by following the research framework which is proposed here.

Chapter 3

Review of Relevant Methodology and Discussions

This chapter will first review the relevant methods for fractal and multifractal analysis. Through comparing these methods, this thesis will give DFA and MF-DFA as preferred methods for fractal and multifractal analysis of geographical problems. In this thesis MF-DFA, and its basis DFA are going to be investigated systematically from several aspects, including two of the disadvantages of DFA and two classical relationships in MF-DFA. The corresponding improvements and corrections are proposed after the discussions of methodology in chapter 4.

From the inception of MF-DFA and rescaled range (R/S) analysis the focus of their academic study has been placed on signals in one-dimensional (1D) space. Recently, MF-DFA has been extended to higher dimensions and it has also been related to the generalized Hurst exponent, $h(q)$, in higher dimensions to Hurst exponent H (Gu and Zhou, 2006). Alvarez-Ramirez et al. (2008) have also extended R/S analysis into two-dimensional (2D) space. Alvarez-Ramirez et al. (2008) employed numerical experiments to compare the two-dimensional MF-DFA and two-dimensional R/S analysis, and they also extended the relationship between H and the scaling exponent obtained from R/S analysis of the two-dimensional case. To avoid confusion, this thesis will call the one-dimensional (MF-)DFA and

R/S analysis (MF-)DFA and R/S analysis, and that in the two-dimensional space the two-dimensional (MF-)DFA (2D (MF-)DFA) and two-dimensional R/S (2D R/S) analysis.

3.1 Methods of Fractal Analysis

This section will first describe the fractal analysis, including calculation of box-counting dimension and the algorithms for calculating the Hurst exponent for long-range correlation analysis.

Fractal analysis aims at characterizing the self-similarity (at least in a statistical sense) of the analyzed processes. The self-similarity exhibits in the formula as the power-law between the measures $M_s(\mathbf{F})$ and corresponding scales s, $M_s(\mathbf{F}) \sim c \cdot s^{-D_f}$. The most common method to calculate the fractal dimension D_f is the box-counting algorithm, which considers the power-law between the number of boxes in the fractal supports having non-empty intersection to the fractal $\mathbf{F}$ and the corresponding box size.

The concept of fractal was proposed by Mandelbrot (1982) for studying the length of a coast. Since the inception of this concept, fractal properties have been found in a large variety of geographical phenomena and appears to be a natural structures of many things in the world.

In general, the fractal nature of geography have been studies by Goodchild and Mark (1987), Lam and De Cola (1993) and Gao and Xia (1996). The complexity of structures and processes in physical geography, and the way to solve the spatial scale problems have also been investigated via the concept of fractals (Atkinson and Tate, 2000), and Richards (2002).

In particular, fractal has been employed to study many real-life phenomena, such as the extraction of the multiscale features of remotely sensed images (Emerson et al., 1999; Lovejoy et al., 2001; Marghany et al., 2009; Myint, 2003; Pachepsky and Ritchie, 1998; Parrinello and Vaughan, 2002; Qiu et al., 1999), especially the integrated system called Image Characterization and Modeling System (ICAMS) developed for fractal analysis of remote sensing data (Quattrochi

et al., 1997), the structure, shape and the size of the cities (Batty, 2008; Batty and Longley, 1994), the fractality of the earthquake process (Goltz, 1997; Harte, 1998; Hirabayashi et al., 1992; Hirata, 1989; Kagan, 1981; Kagan and Knopoff, 1980; Lei and Kusunose, 1999; Sadovskiy et al., 1984; Takayasu, 1990), the heterogeneity and complexity of the spatial point pattern (Buczkowski et al., 1998; Cola, 1991; Kyriacos et al., 1994; Vere-Jones, 1999), and the fractal properties of rainfall process (Kantelhardt et al., 2006; Leung, 2010; Lilley et al., 2006; Olsson, 1995; Peters et al., 2001; Tessier et al., 1996), hydrological process (Koscielny-Bunde et al., 2006; Leung, 2010; Neuman, 2010; Pandey et al., 1998; Tessier et al., 1996; Zhang et al., 2008, 2009).

The concept of long-range correlation is also given here as a relative property to fractal. One of the most important purposes of time series analysis is to develop suitable models and to obtain accurate prediction based on the known record. To achieve this task, one core ingredient is to capture the self-dependence or auto-correlation in the series (Doukhan et al., 2003). The long-range correlation is a very important and common dependency concept in time series, from which the understanding of time series for prediction can be improved. Given a time series $\{x_k\}_{k=1}^N$, after extracting the mean, $< x >= \frac{1}{N}\sum_{k=1}^N x_k$, as $\tilde{x}_k = x_k - < x >$, the auto-correlation function of $\{\tilde{x}_k\}_{k=1}^N$ separated by s, $C(s)$, can be considered as (Bashan et al., 2008):

$$C(s) = \frac{< \tilde{x}_k \tilde{x}_{k+s} >}{< \tilde{x}_k^2 >} = \frac{1}{(N-s) < \tilde{x}_k^2 >} \sum_{k=1}^{N-s} \tilde{x}_k \tilde{x}_{k+s}. \tag{3.1}$$

Then if $\{x_k\}$ is an uncorrelated series, $C(s) = 0$ for positive s. If there is some finite t_x making $C(s)$ to decrease exponentially as $C(s) \sim \exp(-s/t_x)$, $\{x_k\}_{k=1}^N$ can be described by short-range correlation. For those series which have diverges $t_x = \int_0^\infty C(s)\mathrm{d}x$, long-range correlations can be defined if the power-law can be found in the scaling behavior of $C(s)$ as:

$$C(s) \sim s^{-\gamma}, \tag{3.2}$$

The scaling exponent γ lies from 0 to 1. Although long-range correlation can be described from the auto-correlation function, the potential power-law Eq.(3.2) might fail to be found due to the unknown noises or trends. Thus, the scaling exponent γ might be difficult to determine directly from auto-correlation function. Consequently, it is necessary to develop indirect but effective methods to uncover the long-range correlation and Hurst phenomenon.

Many methods have been proposed to handle the long-range correlated process. For example, Hurst (1950) put forward rescaled range analysis to study the water levels of reservoir on the Nile River. More importantly, he gave an important exponent, generally known as the Hurst exponent, H, to quantify long-range correlation of signal series. Fluctuation Analysis (FA) was proposed by Peng et al. (1992) to study the long-range correlations in nucleotide sequences on the basis of the numerical representation, random walk model. Following this, Peng et al. (1994) developed DFA to determine the fractal scaling properties and long-range correlation in both stationary and non-stationary time series. Detrending is one key step when implementing DFA, whereby the polynomial regression is employed to estimate the local trend. According to the different order, m, employed, DFA can be denoted as DFAm. Taqqu et al. (1995) subsequently gave the direct relationship between the scaling exponent of DFA and the Hurst exponent H for fGn. Movahed et al. (2006) and Movahed and Hermanis (2008) later proved the relationship between h and H for the fBm and fGn. Studies have also been made on some properties of DFA, such as: the effect of non-stationarities (Chen et al., 2002), trends (Hu et al., 2001), and extreme data loss (Ma et al., 2010) on DFA, the relation between DFA and power spectral density analysis (Heneghan and McDarby, 2000) and the comparative study of DFA and some other correlation analysis methods (Bashan et al., 2008; Xu et al., 2005). In addition, inspired by the idea of R/S analysis and DFA and utilizing the capability of moving average to capture the low-frequency trends of the signals, Alessio et al. (2002) presented the Detrending Moving Average (DMA) technique to reveal the scaling behaviors by estimating H. The power spectrum technique, which was employed to analyze

the scaling properties of the rainfall series (Fraedrich and Larnder, 1993) and temperature series (Talkner and Weber, 2000), is another method which is devoted to estimating the Hurst exponent. The Structure Function (SF) (also called variogram in spatial statistics) and its many modified versions are important tools which are used in the geophysical or geographical fields (Balankin et al., 2009; Davis et al., 1994) when some scaling exponents related to H are to be estimated.

Finally, this thesis will make the comparisons among these methods. The preferred method to be employed in this thesis can then be determined on the basis of these discussions. In order to facilitate this discussion some popular methods for fractal and multifractal analyses are briefly introduced in the next section.

3.1.1 Box-Counting Dimension

As mentioned in the first chapter, the box-counting dimension is the most popular fractal dimension which is used in real computation because of its simplicity of calculation. The box-counting dimension for fractal $\mathbf{F}$ with support $\mathbf{R^n}$ can be defined as follows:

Given a scale s, the support $\mathbf{R^n}$ can be covered by a set of $\mathbf{R^n}$ boxes as:

$$(m_1 s, (m_1+1)s] \times \cdots \times (m_n s, (m_n+1)s], \tag{3.3}$$

Where $m_1, \cdots, m_n$ are integers (for example, the boxes are intervals for $\mathbf{R}$ while squares for $\mathbf{R^2}$). Using $N_s(\mathbf{F})$ to denote the number of boxes with non-empty intersection to fractal $\mathbf{F}$.

The upper limit $\overline{D}_f(\mathbf{F})$ and inferior limit $\underline{D}_f(\mathbf{F})$ can then be calculated as:

$$\overline{D}_f(\mathbf{F}) = \limsup_{s\to 0} \frac{\log N_s(\mathbf{F})}{-\log s} \tag{3.4}$$

$$\underline{D}_f(\mathbf{F}) = \liminf_{s\to 0} \frac{\log N_s(\mathbf{F})}{-\log s}. \tag{3.5}$$

If the limit of $\frac{\log N_s(\mathbf{F})}{-\log s}$ exists, we have $\overline{D}_f(\mathbf{F}) = \underline{D}_f(\mathbf{F})$. Then the D_f can be

obtained.

When compared with another two common fractal dimensions (i.e. the Hausdroff dimension (Falconer, 1990; Feder, 1988) and packing dimension (Falconer, 1990; Feder, 1988)) the box-counting dimension is much easier to be used when handling the practical problems by coding and computing with the aid of computer.

3.1.2 Rescaled Range (R/S) Analysis

The R/S analysis and the Hurst exponent, H, are firstly proposed to explore the persistence of the storage capacity of water reservoirs on the Nile River by Hurst (1951). Based on the description of R/S analysis given by Hurst (1951), for a given time series $\{x_k\}_{k=1}^{N}$, $Y_M = \{y_k\}$ can be constructed as the M-dimensional sample subvectors of $\{x_k\}_{k=1}^{N}$, here $M = sN$ and $s \in (0,\ 1)$. Then define

$$\bar{y}_s = \frac{1}{M}\sum_{k=1}^{M} y_k, \tag{3.6}$$

$$z_i = \sum_{k=1}^{i}(y_k - \bar{y}_s), \tag{3.7}$$

$$(R/S)_s = \frac{\max z_i - \min z_i}{[\sum_{k=1}^{M}(y_k - \bar{y}_s)^2/M]^{1/2}}. \tag{3.8}$$

For $s \in (s_{\min},\ s_{\max})$, if we have

$$(R/S)_s \sim s^H, \tag{3.9}$$

then we can obtain H as the Hurst exponent. According to the value of H, $\{x_k\}$ is considered as long-range anti-correlated if $0 < H < 0.5$; uncorrelated if $H = 0.5$; and long-range correlated if $H > 0.5$. After obtaining the Hurst exponent, the degree of the prediction of studied series could be assessed. For example, if a series is characterized by $H = 0.5$, then nothing can be done for the prediction due to the total randomness indicated by the value of the Hurst exponent. However, if $H = 1$, the studied series is a definite linear process and

the future of the series is under control and is completely dependent on what happened in the past. Thus the study of the long-range correlation measured by the Hurst exponent is very helpful for establishing a prediction model.

Many applications of R/S in different research fields can be found: such as in economical studies by Cajueiro and Tabak (2008) and Muniandy et al. (2001), in bioinformatics by Wang et al. (2008), in environment problems by Tarafdar and Harper (2008), in geophysics by Chamoli et al. (2007); Jimenez et al. (2006); Li et al. (2002); Peters et al. (2001), and in theoretical discussions by Chamoli et al. (2007) and Rangarajan and Ding (2000).

3.1.3 Two-Dimensional (2D) R/S analysis

Recently, the R/S analysis has been generalized to the 2D space by Alvarez-Ramirez et al. (2008). For the 2D R/S analysis, we just need to construct a sub-surface instead of the one-dimensional subsequence. Given a two-dimensional surface X with size $N_r \times N_c$ and scale s. Consider a sub-matrix $Y_{M_r,M_c} = \{y_{i,j}\}$ with size $M_r \times M_c$, $M_r = sN_r$ and $M_c = sN_c$ respectively. We can obtain the following R/S statistic (Alvarez-Ramirez et al., 2008):

$$\bar{y}_s = \frac{1}{M_r M_c} \sum_{i=1}^{M_r} \sum_{j=1}^{M_c} y_{i,j}, \tag{3.10}$$

$$(R/S)_s = \frac{\max_{i,j} \sum_{k=1}^{i} \sum_{l=1}^{j} (y_{k,l} - \bar{y}_s) - \min_{i,j} \sum_{k=1}^{i} \sum_{l=1}^{j} (y_{k,l} - \bar{y}_s)}{[\sum_{i=1}^{M_r} \sum_{j=1}^{M_c} (y_{k,l} - \bar{y}_s)^2 / (M_r M_c)]^{1/2}}. \tag{3.11}$$

According to Alvarez-Ramirez et al. (2008), as s increases from 0.025 to 0.5, the R/S statistic follow the power law:

$$(R/S)_s \sim s^{2H}. \tag{3.12}$$

The Hurst exponent H can then be derived for the two-dimensional situation (Alvarez-Ramirez et al., 2008). In 2010, Raoufi (2010) employed the 2D R/S analysis to study the fractal property of indium tin oxide (ITO) thin films.

3.1.4 Fluctuation Analysis (FA)

Peng et al. (1992) proposed FA to study the long-range correlations of nucleotide sequences based on random walk model. In FA, firstly the 'net displacement' of the walker, $\{x_k\}_{k=1}^{N}$, after s steps is computed as (Peng et al., 1992):

$$y(s) = \sum_{k=1}^{l} x_k. \tag{3.13}$$

The fluctuation function $F(s)$ can be defined as the difference between the average of the square and the square of the average,

$$\begin{aligned} F^2(s) &= \overline{[\Delta y(s) - \overline{\Delta y(s)}]^2} \\ &= \overline{[\Delta y(s)]^2} - \overline{[\Delta y(s)]}^2, \end{aligned} \tag{3.14}$$

here $\Delta y(s)$ equals $y(s_0+s) - y(s_0)$. The bars means the s_0 indicating all possible positions of $\{x_k\}_{k=1}^{N}$. The Hurst exponent, H, can be obtained if the power-law,

$$F(s) \sim s^H \tag{3.15}$$

can be found.

However, since there usually are some noises in the DNA sequences, Peng and the co-workers improved the detrending procedure into the FA for better estimating the long-range correlation of the DNA sequences (Buldyrev et al., 1993; Peng et al., 1994).

3.1.5 Detrended Fluctuation Analysis (DFA)

To handle the 'nucleotide heterogeneity', a key step, detrending, was introduced into FA and then DFA was consequently proposed by Peng et al. (1994). DFA can be performed in the following steps:

- *Step 1*: Determine the 'profile'

$$Y(i) \equiv \sum_{k=1}^{i} [x_k - < x >], \ i = 1, \cdots, N. \tag{3.16}$$

Here, $< x >$ means the average of $\{x_k\}$. It is remarked in Kantelhardt et al. (2002) that the subtraction of the mean $< x >$ is not compulsory since it would be eliminated by the detrending procedure in *Step 3*.

- *Step 2*: For each given positive integer s, the 'profile' can be divided into $N_s \equiv \text{int}(N/s)$ non-overlapping local windows/segments with equal length s. Here $\text{int}(\cdot)$ is a function which takes the integer part of a number. Since N/s may not be integer, there might be a short part of the 'profile' remained uncovered. To make use of the information containing in this slack, the same procedure can be repeated starting from the opposite end of the series. Hence, $2N_s$ local windows are obtained altogether.

- *Step 3*: For the vth of the $2N_s$ local windows, the variance can be determined as:

$$\begin{aligned} F^2(v,s) \equiv \quad & \frac{1}{s} \sum_{i=1}^{s} \{Y[(v-1)s + i] - y_v(i)\}^2, \\ & \text{for } v = 1, \cdots, N_s; \end{aligned} \tag{3.17}$$

$$\begin{aligned} F^2(v,s) \equiv \quad & \frac{1}{s} \sum_{i=1}^{s} \{Y[N - (v - N_s)s + i] - y_v(i)\}^2, \\ & \text{for } v = N_s + 1, \cdots, 2N_s, \end{aligned} \tag{3.18}$$

where y_v is the fitting polynomial representing the local trend in the vth local window. If we use m to denote the order of y_v, then the one-dimensional DFA can be denoted as one-dimensional DFAm.

- *Step 4*: Obtain the fluctuation function as:

$$F(s) \equiv \{\frac{1}{2N_s}\sum_{v=1}^{2N_s} F^2(v,s)\}^{1/2}. \tag{3.19}$$

- *Step 5*: Determine the scaling behavior of $F(s)$ by analyzing the log-log plots of $F(s)$ versus s. If

$$F(s) \sim s^h. \tag{3.20}$$

Here, h can be related to the classical Hurst exponent H by:

$$H = \begin{cases} h, & \text{for 1D stationary case;} \\ h-1, & \text{for 1D non-stationary case.} \end{cases} \tag{3.21}$$

According to the value of H, $\{x_k\}$ is considered as long-range anti-correlated if $0 < H < 0.5$; uncorrelated if $H = 0.5$; and long-range correlated if $H > 0.5$. Thus, H is a very useful index for extracting the feature of a time series from the perspective of long-range dependence.

In reality, DFA has been successfully applied in many fields such as sunspot series (Hu et al., 2009; Movahed et al., 2006), earthquake processes (Balasco et al., 2002; Telesca et al., 2001a, 2005), temperature series Eichner et al. (2003); Koscielny-Bunde et al. (1998); Pattantyús-Ábrahám et al. (2004), DNA sequence (Peng et al., 1994), streamflow (Matsoukas et al., 2000; Zhang et al., 2008, 2009) and the geomagnetic storm and solar flare indices (Yu et al., 2009).

3.1.6 Detrended Moving Average (DMA)

Moving average technique is a classical way to estimate the low-frequency trend. Based on this characteristic, Alessio et al. (2002) proposed the DMA in which the trend estimated by the moving average procedure at different scales takes the place of that calculated by polynomial regression in DFA. Generally, the procedure of DMA consists of the following main steps (Alessio et al., 2002; Xu et al., 2005):

- *Step 1*: Detect the local trend of the profile $Y(i)$ with scale s in series employing moving average

$$\tilde{Y}_s(i) = \frac{1}{s}\sum_{k=0}^{s-1} Y(i-k). \tag{3.22}$$

- *Step 2*: Subtract the estimated trend $\tilde{Y}_s(i)$ from the integrated profile $Y(i)$, then the residual $C_s(i)$

$$C_s i = Y(i) - \tilde{Y}_s(i). \tag{3.23}$$

- *Step 3*: The fluctuation function of DMA $F(s)$ then can be determined as

$$F(s) = \sqrt{\frac{1}{N-s+1}\sum_{i=s}^{N} C_s^2(i)}. \tag{3.24}$$

- *Step 4*: Determine the scaling behavior of $F_q(s)$ by analyzing the log-log plots of $F_q(s)$ versus s if

$$F(s) \sim s^H. \tag{3.25}$$

DMA attracted much attention covering the financial time series (Carbone et al., 2004a), biological analysis (Shiogai et al., 2010), and theoretical discussion (Arianos and Carbone, 2007; Carbone et al., 2004b; Serletis, 2008; Xu et al., 2005).

3.1.7 Fourier Power Spectral Analysis

Since the signals considered in real computation is discrete, the Fourier power spectral of given series $\{x_k\}_{k=1}^N$ can then be expressed as:

$$\hat{x}(f) = N^{-1/2}\sum_{k=0}^{N-1} x_{k+1} e^{-2\pi i f k}. \tag{3.26}$$

Then

$$S(f) = |\hat{x}(f)|^2 \tag{3.27}$$

is the power spectrum of $\{x_k\}$. If the power-law can be found as:

$$S(f) \sim 1/f^{\beta}, \tag{3.28}$$

the scaling exponent β is determined. Generally, the relationship between β and h of DFA is given by Heneghan and McDarby (2000) as $\beta = 2h - 1$. Hence, β could be related to the long-range correlation analysis.

3.1.8 Structure Fluctuation (SF)

In the turbulent study, SF analysis is a very popular method to explore self-similarity. Given a non-stationary series $\{x_k\}_{k=1}^{N}$ with stationary increment, the SF is defined as:

$$< |\Delta x(s)|^2 >=< |x_{k+s} - x_k|^2 > \sim s^{\zeta(2)}. \tag{3.29}$$

The scaling exponent $\zeta(2)$ is related to H for fBm as $\zeta(2) = 2H$ (Davis et al., 1994).

3.1.9 Preferred Method

Generally, the R/S method and box-counting algorithm can be used only for the analysis of the stationary time series. The SF is also subject to the stationary assumption. In fact, however, most of the geophysical series are non-stationary and are contaminated by various trends. Some of the trends still remain unknown because of the complicated intrinsic and extrinsic dynamic systems that the series has formed.

Fortunately, the detrending step included into the DFA procedure enables it to partly handle the non-stationary series as well as the stationary series. With respect to the performance of the power spectrum method in estimation of H exponent, significant fluctuation of power spectrum, in particular that of low-frequency range, can usually mask the scaling regions and lead to difficulty when

directly estimating of the slope in the double-logarithmic plots (Matsoukas et al., 2000; Tsonis et al., 1998; Weber and Talkner, 2001). In contrast to the spectral analysis, FA is a more systematic procedure (Talkner and Weber, 2000) with small finite size effects and a smoother fluctuation function (Tsonis et al., 1998). Although FA works better than spectral analysis to estimate the long-range correlation, it cannot avoid the effect of the trend (such as global warming) as pointed out by Koscielny-Bunde et al. (1998). As to the DMA, Xu et al. (2005) compared the performance of DFA and DMA, the scaling curves obtained by DFA1 are claimed to be more stable over much broader scale ranges when compared with DMA, which suggests a better fitting range can be obtained to quantify the correlation property. In addition, Bashan et al. (2008) pointed out that the DMA has the similar performance with the DFA technique if the trends within the studied series are weak. However, the DFA method is still the best choice, particularly when the trends within the studied series remain unknown before analysis. Actually, the trends of the series are often unknown. In this sense, the DFA method has the potential to fully recognize and diagnose the trends within the series by using different orders of detrending polynomial (Bashan et al., 2008). Due to the foregoing advantages and its simplicity to implement, DFA is the prominent priority in the analysis of the scaling properties of the series amongst the alternative methods for analyzing geographical processes.

3.2 Methods of Multifractal Analysis

Because of the insufficiency of fractal analysis to capture the heterogeneity of complex and irregular processes with a single fractal dimension (Grassberger and Procaccia, 1983), multifractal analysis is employed in this thesis in order to give a full description of complicated scaling behaviors over multiple time scales. By taking into the qth order moment of the measure or fluctuations account, the fractal dimension D_f and Hurst exponent H are extended to a generalized fractal dimension $D(q)$ (Mandelbrot, 1982) and a generalized Hurst exponent $h(q)$ (Kantelhardt et al., 2002). The D_f and H are two special cases of $D(q)$ and $h(q)$.

$D(q)$ and $h(q)$ with different values of q reflect the measures and fluctuations with different degrees. The parts with high density measure or a large fluctuation will dominate the qth order moment when the positive values of q are taken into consideration. However, for the negative values of q, what is reflected in the scaling behavior is the information related to the small density, or fluctuation parts of structures and processes. In an extreme situation, the $D(\pm\infty)$ corresponds to the information of those parts with largest or smallest density and fluctuations. Usually, the $D(q)$ and $h(q)$ monotonous decreases with q increases (Kantelhardt et al., 2002).

3.2.1 Partition Function-Based Formalism

Halsey et al. (1986) proposed the partition function-based formalism technique, which is the most commonly used method, to extract the multifractality of studied objects. The basic procedures can be enumerated as:

Given a measure μ with support $E \subset \mathbf{R}^n$, the partition function can be represented as:

$$Z_s(q) = \sum_{\mu(B_m)\neq 0} [\mu(B_m)]^q, \ q \in \mathbf{R}, \tag{3.30}$$

Where B_m is a box of a given side ϵ covering the support E:

$$B_m = (m_1 s, (m_1+1)s] \times \cdots \times (m_n s, (m_n+1)s], . \tag{3.31}$$

If power-law $Z_s(q) \propto s^{\tau(q)}$ is existing, we can get the $\tau(q)$ using following equation:

$$\tau(q) = \lim_{s\to 0} \frac{\log Z_s(q)}{\log s}. \tag{3.32}$$

And generalized fractal dimension $D(q)$ can be calculated by

$$D(q) = \frac{\tau(q)}{q-1}, \quad \text{for } q \neq 1, \tag{3.33}$$

$$D(q) = \lim_{s\to 0} \frac{Z_{1,s}}{\ln s}, \quad \text{for } q = 1, \tag{3.34}$$

where $Z_{1,s} = \sum_{\mu(B)\neq 0} \mu(B) \ln \mu(B)$. The multifractal spectrum can then be expressed as:

$$\alpha(q) = \frac{\mathrm{d}}{\mathrm{d}q}\tau(q), \tag{3.35}$$

$$f(\alpha) = q\alpha(q) - \tau(q). \tag{3.36}$$

Usually, in the real calculation, $\alpha(q)$ can be obtained by:

$$\alpha(q) = \lim_{s\to 0} \frac{\sum_{\mu(B)\neq 0}[\mu(B)]^q \ln \mu(B)}{Z_s(q) \ln s} \tag{3.37}$$

$D(q)$ for a large positive q reflects the fractal property of the dense regions, while $D(q)$ for negative q quantities reflects the situation for the sparse regions. $D(+\infty)$ and $D(-\infty)$ are two extreme dimensions indicating the most and least clustering of distribution, respectively. Specifically, $D(0)$ is the conventional fractal dimension D_f. $D(1)$ and $D(2)$ corresponds to the information dimension and correlation dimension (Falconer, 1990). $D(0)$, as the geometric measure, shows the extent of fractal distributing in the support. $D(1)$ and $D(2)$ depicts the non-uniformity and global clustering degree of data, respectively (Enescu et al., 2005). MFA is a useful way to characterize the spatial heterogeneity of both theoretical and experimental fractal patterns (Grassberger and Procaccia, 1983). In fact, some applications in the spatial analysis can be found. Vere-Jones (1999) has discussed the fractal dimension of the point pattern in fixed spatial region and fixed interval of time respectively theoretically. Buczkowski et al. (1998) used the modified box-counting method designed by Kyriacos et al. (1994) to measure the regular and random distributed objects. While Cola (1991) analyzed the multiscale spatial autocorrelation of point data.

3.2.2 Multifractal Detrended Fluctuation Analysis (MF-DFA)

Given a time series, $\{x_k\}_{k=1}^{N}$ of length N with compact support, the MF-DFA method consists of five main steps (Kantelhardt et al., 2002), all steps except the 4th one are the same as those of DFA. In MF-DFA, the 4th step takes the qth instead of 2nd order moment of fluctuations into consideration as:

- *Step 4*: Obtain the qth order fluctuation function as:

$$F_q(s) \equiv \{\frac{1}{2N_s}\sum_{v=1}^{2N_s}[F^2(v,s)]^{q/2}\}^{1/q}, \tag{3.38}$$

  where q can be any real number.

Then in step 5, $h(q)$ can be considered in the following way:

- *Step 5*: Determine the scaling behavior of $F_q(s)$ by analyzing the log-log plots of $F_q(s)$ versus s for each q. Generally, we only use s varying from $m+2$ to $N/4$. If

$$F_q(s) \sim s^{h(q)}, \tag{3.39}$$

  $h(q)$ called the generalized Hurst exponent could be obtained. Specifically, for $q=2$, the MF-DFA are actually DFA and $h(q=2)$ is h in DFA.

The values of $h(q)$ for different q's can be employed to quantify the property of the parts of series with different degrees of fluctuations. The $h(q)$ corresponding to large q reflects the property of large fluctuations while that to small q connects to small fluctuations. Specifically, the $h(q)$ of the positive and negative infinite q correspond to the maximum and minimum fluctuations respectively. According to the dependence of the $h(q)$ on q, the multifractality could be detected. If $h(q)$ is independent/dependent on q, then there is the monofractality/multifractlity. Fig. 3.1 exhibits the dependence of $h(q)$ on q for two examples, namely fractional Gaussian noise and temperature record. It is shown that $h(q)$ almost remains constant for all qs for the former while varies significantly for the latter, which indicates their different multifractal properties.

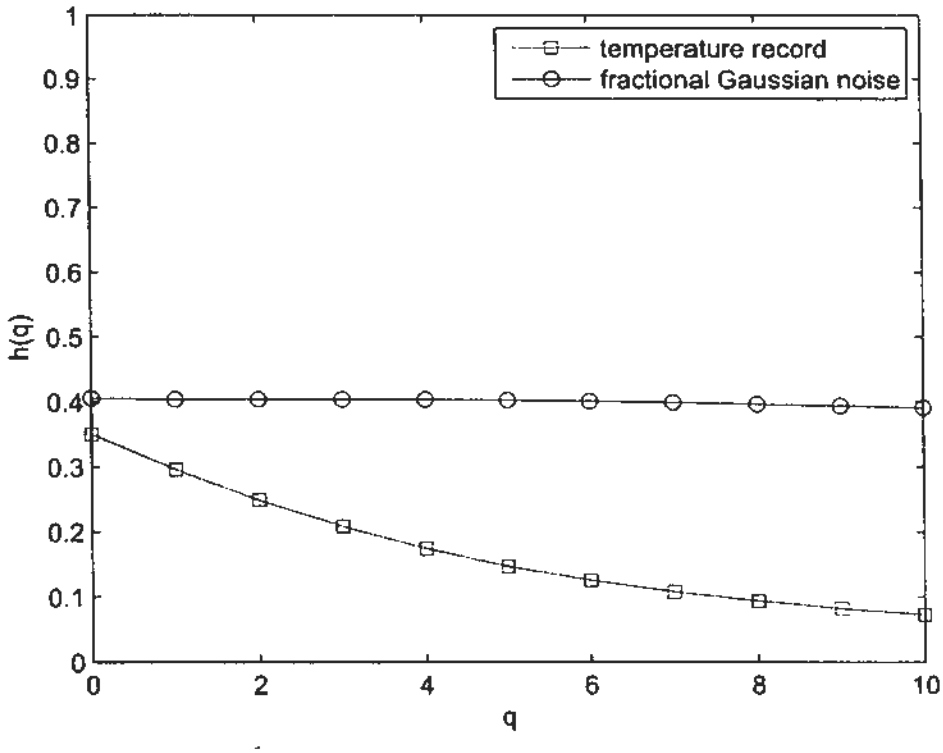

Figure 3.1: $h(q)$ for fractional Gaussian noise and temperature record against q.

According to Kantelhardt et al. (2002), if the time series $\{x_k\}_{k=1}^{N}$ is stationary, positive and normalized, then it is not necessary to detrend in step 3. Thus the DFA can be replaced by the standard fluctuation analysis as:

$$F_{FA}^2(v,s) \equiv [Y(vs) - Y((v-1)s)]^2. \tag{3.40}$$

Then, combining Eq.(3.38) and Eq.(3.39), we can obtain:

$$\sum_{v=1}^{2N_s} |Y(vs) - Y((v-1)s)|^q \sim s^{qh(q)-1}. \tag{3.41}$$

To relate the one-dimensional MF-DFA to the box-counting formalism, Kantelhardt et al. (2002) gave the partition sum of the analyzed time series as:

$$Z_q(s) \equiv \sum_{v=1}^{N/s} |p_s(v)|^q \sim s^{\tau(q)}, \tag{3.42}$$

where

$$p_s(v) \equiv \sum_{k=(v-1)s+1}^{vs} x_k = Y(vs) - Y((v-1)s). \tag{3.43}$$

From the above two equations, they established the relationship to relate the partition function formalism and MF-DFA together as:

$$\tau(q) = qh(q) - 1. \tag{3.44}$$

Based on $\tau(q)$, the generalized fractal dimension $D(q)$ and the singular spectrum can be derived as follows:

$$D(q) \equiv \frac{\tau(q)}{q-1}, \tag{3.45}$$

$$\alpha = \tau'(q) \quad \text{and} \quad f(\alpha) = q\alpha - \tau(q). \tag{3.46}$$

With Eq.(3.44), $D(q)$ and the singular spectrum can be expressed via $h(q)$ as:

$$D(q) = \frac{qh(q) - 1}{q-1}, \tag{3.47}$$

$$\alpha = h(q) + qh'(q) \quad \text{and} \quad f(\alpha) = q[\alpha - h(q)] + 1. \tag{3.48}$$

Such relationship (Eq. 3.44) has been employed to study the scaling behaviors of time series in different fields of research (Kimiagar et al., 2009; Movahed and Hermanis, 2008; Movahed et al., 2006; Telesca et al., 2004c, 2005).

Actually, there are some other methods for multifractal analysis, such as the wavelets transform modulus maxima (WTMM) (Muzy et al., 1994, 1991) and empirical mode decomposition based arbitrary order Hilbert spectral analysis (EMD-HSA) (Huang et al., 2008). Compared to WTMM and EMD-HSA, the algorithm of MF-DFA is much easier to implement. Moreover, the systematic discussion and comparison of the performance of WTMM and MF-DFA on fractal and multifractal analysis was made in Oswiecimka et al. (2006). Simulated numerical series and real-life examples were employed for comparison. Then, Oswiecimka et al. (2006) claimed that MF-DFA works in a more automatic way while WTMM should be applied with care when applied to analyze real data.

3.2.3 2D MF-DFA

The procedure of the 2D MF-DFA is very similar to that of the MF-DFA. Based on the description by Gu and Zhou (2006), for each given s, we can first divide the two-dimensional surface, denoted by a matrix $X(i,j)$ with size $M \times N$, into $M_s \times N_s$ non-overlapping sub-surfaces $X_{v,w}$ with size $s \times s$, where $M_s = \text{int}(M/s)$ and $N_s = \text{int}(N/s)$. The profile $u_{v,w} = \sum\sum X_{v,w}(i,j)$ can then be obtained in each $X_{v,w}$. Subsequently, we can eliminate local trend from every $u_{v,w}$ by using one of the following local surfaces as the pre-specified polynomial detrending function $\tilde{u}$:

$$\tilde{u}_{v,w}(i,j) = ai + bj + c, \tag{3.49}$$

$$\tilde{u}_{v,w}(i,j) = ai^2 + bj^2 + c, \tag{3.50}$$

$$\tilde{u}_{v,w}(i,j) = aij + bi + cj + d, \tag{3.51}$$

$$\tilde{u}_{v,w}(i,j) = ai^2 + bj^2 + ci + dj + e, \tag{3.52}$$

$$\tilde{u}_{v,w}(i,j) = ai^2 + bj^2 + cij + di + ej + f, \tag{3.53}$$

here a, b, c, d, e and f are free parameters to be determined. Then for each sub-surface, we can have a residual matrix:

$$\varepsilon_{v,w}(i,j) = u_{v,w}(i,j) - \tilde{u}_{v,w}(i,j). \tag{3.54}$$

The variance of the residual matrix is obtained as:

$$F^2(v,w,s) = \frac{1}{s^2}\sum_{i=1}^{s}\sum_{j=1}^{s}\varepsilon^2_{v,w}(i,j). \tag{3.55}$$

Subsequently, the two-dimensional qth order fluctuation function, $F_q(s)$, becomes

$$F_q(s) = \{\frac{1}{M_s N_s}\sum_{v=1}^{M_s}\sum_{w=1}^{N_s}F^q(v,w,s)\}^{1/q}. \tag{3.56}$$

By varying s from 6 to $\min(M, N)/4$, if the power law $F_q(s) \sim s^{h(q)}$ exists, we can obtain the generalized Hurst exponent $h(q)$ for the two-dimensional surface. Gu and Zhou (2006) also gave without proofs the relationship between $h(q)$ and H for the 2-dimensional fBm as:

$$h(q) \equiv H, \tag{3.57}$$

and the relationship between $h(q = 2)$ and H as:

$$h(q = 2) \equiv H. \tag{3.58}$$

3.3 Several Problems in DFA and MF-DFA

Although the DFA is the preferred method for the fractal analysis, and DFA and its extended version, MF-DFA, have been applied to solve problems in many fields such as astronomy (Anh et al., 2007, 2008; Movahed et al., 2006; Yu et al., 2009), hydrology (Movahed and Hermanis, 2008; Zhang et al., 2008, 2009), meteorology (Talkner and Weber, 2000), seismology (Telesca et al., 2004a,c, 2005; Varotsos et al., 2002, 2003) and electrics (Kimiagar et al., 2009), we still have to investigate them systematically before applying to the geographical processes. Firstly, we study the basis, DFA, since its disadvantages could also bring negative influence on MF-DFA. And then the incorrectness in MF-DFA would be discussed in this section.

3.3.1 Strong Fluctuation in Scaling Behavior

Detrending is the key step in DFA. As to the concept of trend, it is hard to give an exact definition. However, there are generally two approaches for estimating the trend: one employs regression and the other uses the moving mean of data (Wu et al., 2007). In DFA, it treats the polynomial fits of the time series in local windows as the local trends to be handled in the detrending step of the algorithm. It should be noted that ordinary linear or high order regression can be used in

the detrending step. In finding the best-fit curve, all points in a local window take on the same regression parameters regardless of their location in the window. However, it has been argued, particularly in geographical studies (Tobler, 1970), that points near in space should be more related than points some distance away. The same argument can also be applied to points in time series. It is natural to expect points near in time in a time series to be more related, e.g. have similar values, than points some distance apart. Change of temperature over time is a typical example. As pointed out by Eichner et al. (2003), 'the persistence of weather states on short terms is a well-known phenomenon: a warm day is more likely to be followed by a warm day'. With respect to DFA, points locating at the end (beginning) of a local window of a time series should have stronger correlation with points closer to them, even though they are located at the beginning (end) part of the local window next (before) to it. That is, with reference to a point, points at greater distance within the same local window might not be as related as those in another local window right next to it. The implication then is that points should be weighted according to their position in the time series. Furthermore, Alvarez-Ramirez et al. (2005) pointed out in their study that as the removal of local trends in DFA is based on discontinuous polynomial fitting, oscillations in the fluctuation function and significant errors in crossover locations might be introduced. It is very common to see that the linear relationship in the $\log - \log$ plot of the fluctuation function $F(s)$ versus the scale s is very poor when s is large.

Besides, the non-overlapping window considered in the step 2 of DFA might lead to the insufficient samples for calculating the fluctuation function of DFA, especially for the large scales. For example, if $s = N/4$, only the fluctuation in four window could be taken into account. As a results, the static on the basis of the insufficient samples may attribute to the unstable scaling behavior with strong fluctuation, particularly at large scales.

3.3.2 Influence of Periodic Trend

In the geographical phenomena, the periodic or quasi-periodic trend is common to see, such as the annual cycle in temperature and streamflow records, and the famous 11-year-cycle in sunspot series. Usually, the existence of crossover in the scaling behavior of DFA can imply a transition from one to another state of underlying correlation. In the analysis of the real-life examples, the crossover points generally indicate the change of the scaling behavior of a process. And scales marked by such crossover points usually have a critical and significant physical meanings. For example, the crossover points in the study of temperature records could be detected around one week, one month and one season, which uncover different properties of persistence in the climate systems, such as more or less similar temperature records found in successive days but significant difference in the days separated by seasons. The significance of these detected scales can generally be confirmed by the climatological studies. However, some trends can lead to the crossover which has nothing to do with state transition of dynamical properties of the underlying process, for example the spurious crossover time scales due to periodic or quasi-periodic trend (Hu et al., 2001; Nagarajan and Kavasseri, 2005). Therefore, handling this kind of trend is very much critical for obtaining the genuine scaling behavior of a time series.

In fact, many methods have been designed to remove the periodic or quasi-periodic trends from time series. According to Wu et al. (2007), there are two commonly used methods for estimating the trend of time series/signals. They are the linear regression method and the moving average method. Other than that, there are some more complicated trend extraction methods, like the higher order regression analysis, Fourier-based filtering (Nagarajan and Kavasseri, 2004), singular-value decomposition (SVD) (Nagarajan and Kavasseri, 2005) and the adaptive detrending methods (Hu et al., 2009). Details about these methods (except the adaptive detrending in Hu et al. (2009)) and their comparison can be found in Kantelhardt (2008) and Bashan et al. (2008).

The linear regression method may be inappropriate or physically meaningless

for real-life problems, especially under nonlinearity and non-stationarity. With regard to the moving average methods, a pre-determined time scale, like that in Alessio et al. (2002) and Alvarez-Ramirez et al. (2005), is required before the trends are estimated. Higher order regression analysis and Fourier-based filtering are often based on the stationarity and linearity assumptions, and some pre-determined function forms, such as polynomial, sine or cosine functions, for higher order regression and Fourier-based filtering. Besides this, there are no physically justifiable foundations supporting the pre-specified function forms (Wu et al., 2007). Using the SVD method one has to determine the number of singular values to be removed. The Fourier-based method is faced with similar problem, that is how many of the frequencies should be removed. The adaptive detrending in Hu et al. (2009) is based on the regression analysis. Thus the order of polynomial and the time scale have to be determined when implementing these trend extracting methods.

What should be emphasized is that the definition of trend is debatable. For example, the periodic trend in financial series is considered an important component rather than noise which should be removed during the analysis (Wu et al., 2007). If the periodic trend dominates the series, the scaling information of other components would be masked. To investigate the scaling behavior of those non-periodic components, the dominant part has to be removed. Thus, to give a complete description of the scaling behavior, in the study of the temperature variations, we analyze the scaling behavior of series with annual cycle and series after removal of the annual cycle. However, for the same sunspot time series, Movahed et al. (2006) and Hu et al. (2009) derived two totally different results using DFA on the basis of two different ways to remove the effect of the 11-year cycle. Specifically, Movahed et al. (2006) applied the Fourier truncation method whereas Hu et al. (2009) developed an adaptive detrending method. Furthermore, Hu et al. (2009) claimed that the Hurst exponent obtained in Movahed et al. (2006) is incorrect because an inappropriate detrending method, namely Fourier truncation, was used. Therefore results in Movahed et al. (2006) are con-

sidered an artificial outcome of Fourier truncation rather than a real property of the sunspot series. Unlike Hu et al. (2009), Nagarajan and Kavasseri (2004) showed that the Fourier-based method is a conventional but effective method for removing periodic trends. The question then is which method is appropriate for handling the 11-year cycle? On the basis of the above discrepancy, how to remove the 11-year cycle thus becomes a key issue in the study of the sunspot time series using DFA. It triggered us to adopt another angle to get to the bottom of this problem.

3.3.3 Problematic Relationship in MF-DFA

Though the MF-DFA have become popular over the years, there seem to be some problems in the relationships of the exponents given in these papers. In the development of MF-DFA, Kantelhardt et al. (2002) established a relationship, which has been extensively applied ever since, between MF-DFA and the standard partition-function-based multifractal formalism as: $\tau(q) = qh(q) - 1$ for positive and stationary time series. This relationship has been employed as a basis for multifractal analysis in a good number of studies (see for example (Kimiagar et al., 2009; Movahed and Hermanis, 2008; Movahed et al., 2006; Telesca et al., 2004c, 2005)). However, our investigation shows that such relationship has a potential problem. It might at least be incorrect for some signal like fGn.

Since the concept of the Hurst exponent, H, and the relationship between $h(q)$ and H are the critical issues, they are briefly reviewed here first. The scaling exponent β, which is determined by the Fourier power spectrum $\mathrm{E}(f) = f^{-\beta}$, is employed to describe many processes. Thus Hurst exponent can also be defined using β. On one hand, a definition of the Hurst exponent, H, of fBm and fGn requires different relations to connect these exponents according to the underlying processes as $H = (\beta - 1)/2$ for fBm and $H = (\beta + 1)/2$ for fGn (Barton and Poor, 1988; Mandelbrot and Ness, 1968). Then H lies between 0 and 1 for both fBm and fGn. On the other hand, there is a different definition of the Hurst exponent, H, of fBms and fGns given as $H = (\beta - 1)/2$ (Flandrin, 1992; Mandelbrot, 1982). In

this definition, H varies from -1 to 0 for fGn and from 0 to 1 for fBm. We denote this definition of H as H_1 in the following discussion. The following statement gives the reasons that the former definition ought to be the one adopted in MF-DFA. As mentioned above, MF-DFA becomes DFA if $q = 2$. Heneghan and McDarby (2000) pointed out that $H = (\beta + 1)/2$ for fGn and $H = (\beta - 1)/2$ for fBm and proved that the relationship between the DFA exponent $2h(2)$ and the power spectrum scaling exponent, β, is $\beta = 2h(2) - 1$. That is to say $h(2) = H$ for fGn and $h(2) = H + 1$ for fBm. With this definition of H, the relationships between H and DFA scaling exponent $h(2)$ for fGn and fBm have been confirmed by analytical calculation. Taqqu et al. (1995) analytically calculated that the DFA scaling exponent h equals H for fGn. Furthermore, Movahed et al. derived the relationship $h(2) = H + 1$ for fBm (Movahed et al., 2006) and $h(2) = H$ for fGn (Movahed and Hermanis, 2008). These research results can be employed to support that the scaling exponent $h(2)$ obtained by MF-DFA should be related to H rather than H_1. In addition, Kantelhardt et al. (2002) gave the relationship between H and β as $H = (\beta+1)/2$ for the stationary long-range correlated series. So in MF-DFA, the former definition of the Hurst exponent is adopted. And in this thesis, we just consider the former definition of Hurst exponent.

Here, we would like to point out the incorrectness of relationship between $\tau(q)$ and $h(q)$ from the theoretical and numerical points of view.

The Theoretical Issue

It should be observed that for positive, stationary, and normalized time series $\{x_k\}_{k=1}^N$, if no trend has to be eliminated, then *Step 3* will not be performed. In such case, *Step 1* of MF-DFA should then be compulsory. If so,

$$Y(vs) - Y((v-1)s) = \sum_{k=(v-1)s+1}^{vs} (x_k - < x >), \tag{3.59}$$

instead of Eq.(3.43). And even if $\{x_k\}_{k=1}^{N}$ is a positive series, $x_k - < x >$ might still be negative. Without loss of generality, we can suppose:

$$\sum_{k=(v-1)s+1}^{vs} (x_k - < x >) > 0, \tag{3.60}$$

$$\sum_{k=vs+1}^{(v+1)s} (x_k - < x >) < 0; \tag{3.61}$$

and

$$|\sum_{k=(v-1)s+1}^{vs} (x_k - < x >)| > |\sum_{k=vs+1}^{(v+1)s} (x_k - < x >)|. \tag{3.62}$$

Then, even if

$$\{x_k - < x >\}_{k=(v-1)s+1}^{(v+1)s} \supseteq \{x_k - < x >\}_{k=(v-1)s+1}^{vs}, \tag{3.63}$$

it is still possible to have

$$|\sum_{k=(v-1)s+1}^{(v+1)s} (x_k - < x >)| < |\sum_{k=(v-1)s+1}^{vs} (x_k - < x >)|. \tag{3.64}$$

Just take the fGn as an example. Fig.3.2 depicts an enlarged plot of part of the fGn. We transform $\{x\}$ upwards for $|\min(x)| + 1$ to ensure that the new series, $\{\tilde{x}\}$, is positive, and then replace it using $\{\tilde{x} - < \tilde{x} >\}$. Mathematically, $\{\tilde{x} - < \tilde{x} >\} = \{x - < x >\}$. Then we obtain a positive, stationary series. Given $s = 3$, it is easy to see that Eq.(3.60)-Eq.(3.64) are valid. However as is well known, if μ is a measure defined on sets A and B, then

$$A \supseteq B \tag{3.65}$$

must be followed by

$$\mu(A) \geq \mu(B). \tag{3.66}$$

That is to say $|p_s(v)|$ in Eq.(3.42) cannot be treated as a measure defined on the times series $\{x_k\}_{k=1}^{N}$ since it does not possess the basic property of a measure.

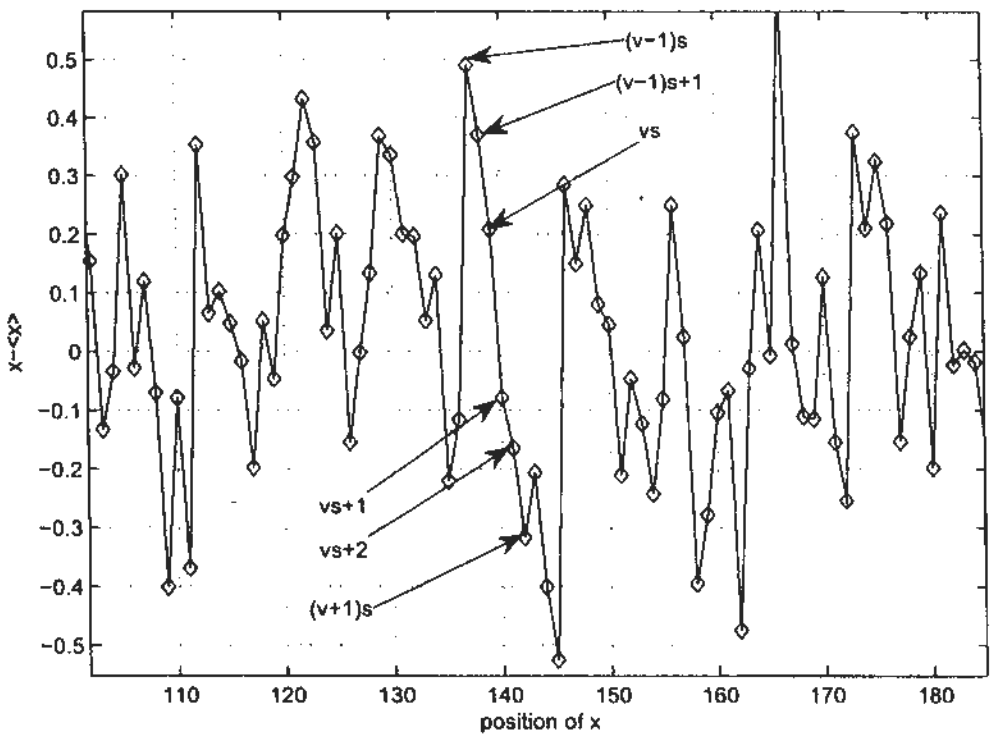

Figure 3.2: An enlarged plot of part of the fGn ($x - < x >$), for s=3.

However, when the box-counting algorithm, for example, is used, it is necessary to have a measure defined on a set. Then, the partition-sum function based on $|p_s(v)|$ is not proper. Now that Eq.(3.43) is incorrect, the relationship in Eq.(3.44) which is obtained by combining Eq.(3.41), Eq.(3.42) and Eq.(3.43), might also be incorrect, at least for fGns. In the following subsection, we employ numerical experiments to support our arguments.

Numerical Experiments

We first substantiate our view on the problematic relationship between $\tau(q)$ and $h(q)$ established in Kantelhardt et al. (2002) with a series of numerical experiments. In order to numerically obtain the relationship between $\tau(q)$ and $h(q)$, we have to calculate $\tau(q)$ and $h(q)$ first. To be consistent, we consider in here the stationary, positive, and normalized time series following the assumption in Kantelhardt et al. (2002). As it is well-known, fGn is the stationary increment of fBm. Thus we construct our numerical examples using fGn to calculate $\tau(q)$

and $h(q)$ by the fixed-size box-counting algorithm in Halsey et al. (1986) and MF-DFA in Kantelhardt et al. (2002) respectively. The procedure to construct our examples is presented in detail as follows:

In the experiment, we first used MatLab to synthesize 10 fBms with length $2^{15}+1$ for each assigned Hurst exponent H whose values range from 0.1 to 0.9 with step size 0.1. For each fBm, we obtained a corresponding fGn, $\{G_k\}$, with the same H. As the fixed-size box-counting algorithm is based on the partition-sum function, it is necessary to keep all elements of the series non-negative. Thus, we replaced the original $\{G_k\}$ with $\{G_k - \min(G) + 1\}$, where $\min(G)$ denotes the minimum value of series $\{G_k\}$, so that we could transform $\{G_k\}$ into a positive series which still is stationary. Moreover, the detrending step in MF-DFA ensures that the vertical translation holds the same H as that of $\{G_k\}$. Without confusion, we still use $\{G_k\}$ to denote the vertically shifted fGns in the following description. To calculate $\tau(q)$ and $h(q)$, we let q increase from -5 to 5 with step size 1. Applying the box-counting algorithm (Halsey et al., 1986), we computed $\tau(q)$ for each series $\{G_k\}$. After normalizing the transformed series, we applied the fixed-size box-counting algorithm to get $\tau(q)$. The almost unvaried values of averaged $D(q)$ shown in Table 3.1 indicate the monofractality of these vertically shifted fGns. Whereafter, we calculated $h(q)$ using MF-DFA. It should be noted

Table 3.1: Averaged $D(q)$ of vertical shifted fGns for different H and q values

	q=-5	q=-4	q=-3	q=-2	q=-1	q=0	q=1	q=2	q=3	q=4	q=5
H=0.1	1.0310	1.0310	1.0310	1.0310	1.0310	1.0310	1.0310	1.0310	1.0310	1.0310	1.0310
H=0.2	1.0310	1.0310	1.0310	1.0310	1.0310	1.0310	1.0310	1.0310	1.0310	1.0310	1.0310
H=0.3	1.0311	1.0311	1.0310	1.0310	1.0310	1.0310	1.0310	1.0310	1.0310	1.0310	1.0310
H=0.4	1.0311	1.0311	1.0311	1.0311	1.0311	1.0310	1.0310	1.0310	1.0310	1.0310	1.0310
H=0.5	1.0311	1.0311	1.0311	1.0311	1.0311	1.0310	1.0310	1.0310	1.0310	1.0310	1.0310
H=0.6	1.0311	1.0311	1.0311	1.0311	1.0311	1.0310	1.0310	1.0310	1.0310	1.0310	1.0310
H=0.7	1.0314	1.0314	1.0313	1.0312	1.0311	1.0310	1.0310	1.0309	1.0308	1.0307	1.0307
H=0.8	1.0321	1.0319	1.0317	1.0315	1.0313	1.0310	1.0308	1.0306	1.0304	1.0302	1.0300
H=0.9	1.0349	1.0341	1.0334	1.0326	1.0318	1.0310	1.0303	1.0295	1.0287	1.0280	1.0272

that these normalized transformed fGns are stationary, positive, and normalized series satisfying totally the assumption in (Kantelhardt et al., 2002) to in the derivation of the relationship between $\tau(q)$ and $h(q)$, e.g. Eq.(3.44). Since the Hurst exponents of fGn here are all positive as emphasized in the introduction, MF-DFA is directly performed without employing the double summation tech-

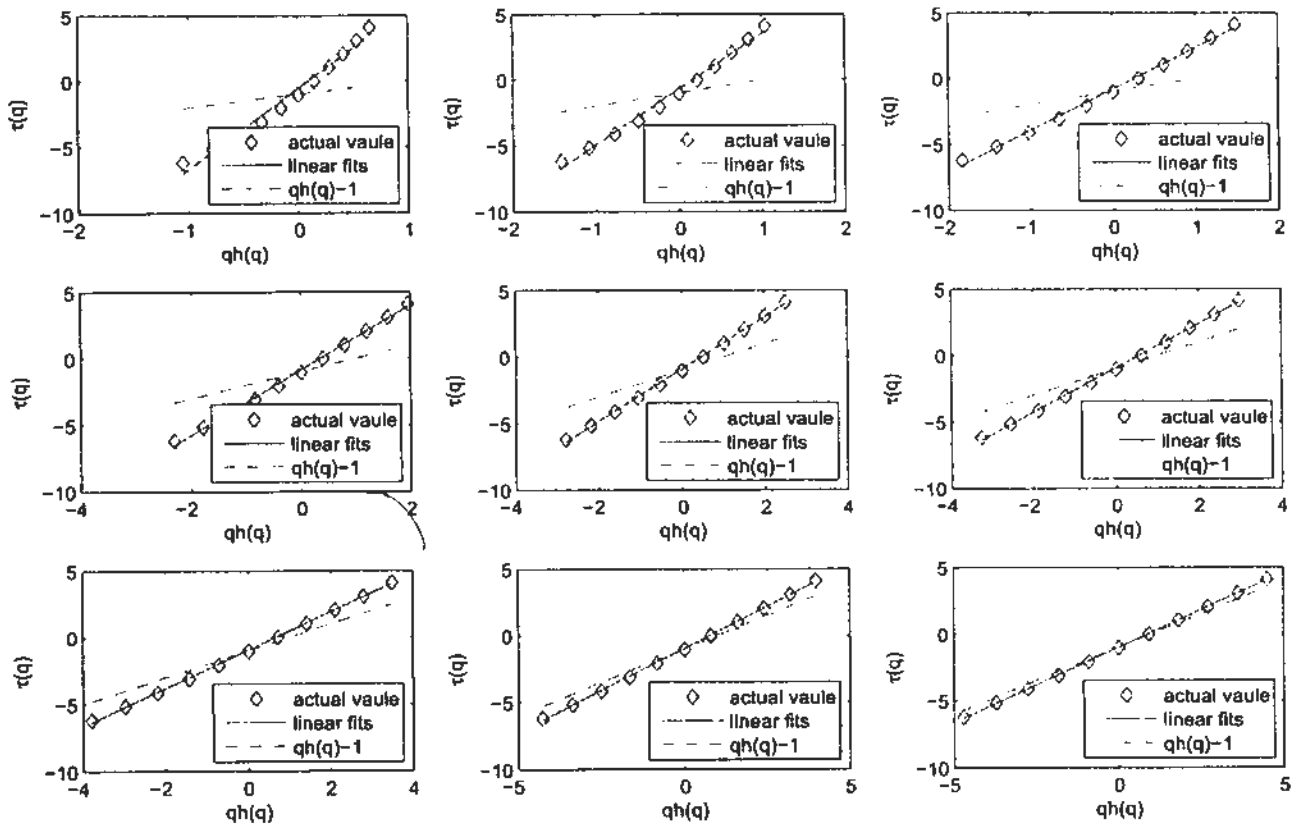

Figure 3.3: Relationships between $h(q)$ and $\tau(q)$ for numerical examples with H=0.1 to 0.9. (from top left to bottom right).

nique as suggested by Kantelhardt et al. (2002). Step 3, detrending, of MF-DFA is able to keep $h(q)$ of $\{G_k\}$ a constant when a constant is added to each element G_k of the series. Now, we could obtain $\tau(q)$ and $h(q)$ for each H. And then the average values of ten $D(q)$ and $h(q)$ for each H are compared. The figures are shown in Fig.3.3. Apparently, the actual curves of $\tau(q)$ versus $qh(q)$ do not follow the expression stipulated by Eq.(3.44). The deviations are actually rather drastic.

From Fig.3.3, one can observe that the curves poorly fit the linear relationship when $H \leq 0.3$. As H increases, however, the curve of $\tau(q)$ versus $qh(q)$ is getting more and more like the straight line: $\tau(q) = qh(q) - 1$. Table 3.2 gives the slopes and intercepts of the linear fits of these actual curves. From this table, we can see that the linear fit is close to $\tau(q) = qh(q) - 1$ only when $H = 0.9$. However, there is a significant difference between the actual curve and $\tau(q) = qh(q) - 1$

Table 3.2: Slopes and intercepts of the linear fits of the curve of $\tau(q)$ and $qh(q)$ for different H values in Fig.3.3

H	slopes of the linear fits	intercepts of linear fits
0.1	6.0839	-0.5593
0.2	4.2370	-0.7178
0.3	3.1375	-0.8149
0.4	2.4383	-0.8663
0.5	1.9604	-0.9184
0.6	1.6609	-0.9382
0.7	1.4240	-0.9504
0.8	1.2560	-0.9656
0.9	1.1264	-0.9995

when $H \leq 0.8$, especially for small H. This numerical counter example shows that the relationship between $\tau(q)$ and $h(q)$ stipulated by Eq.(4.2) established by Kantelhardt et al. (2002) appears to be incorrect.

3.3.4 Problematic Relationship Between H and $h(q)$ in the 2D MF-DFA

Gu and Zhou (2006) claimed that H can be obtained via Eq.(3.57) for two-dimensional fBm and Eq.(3.58). For the two-dimensional R/S analysis, however, H is calculated using the power law (see Eq.(3.12)). Alvarez-Ramirez et al. (2008) claimed that the two-dimensional R/S analysis is equivalent to the two-dimensional DFA1 (two-dimensional DFA which eliminates local trend using planar sub-surface), e.g. $2H = h(2)$. They also generated two-dimensional fBm to carry out numerical experiments using both the two-dimensional R/S analysis and two-dimensional MF-DFA to support their claims. However, the relationships obtained in these two separate studies are contradictory: i.e. $H = h(2)$ in Eq.(3.58) versus $2H = h(2)$ in Alvarez-Ramirez et al. (2008). Since H is a constant for a given two-dimensional signal, either one of these relationships has to be wrong or they are both wrong. Such contradiction compels us to numerically investigate the relationship and make a conjecture about the real one.

Problem of the 2D R/S Analysis in Alvarez-Ramirez et al. (2008)

The 2D R/S is studied in Alvarez-Ramirez et al. (2008) via the two-dimensional fBm. Since the one-dimensional R/S analysis is only suitable for the stationary time series, and the two-dimensional fBm used in Alvarez-Ramirez et al. (2008) is a well-known non-stationary process, it is questionable that R/S can handle the non-stationary fBm just because it is generalized to the two-dimensional space. To satisfy our curiosity, we repeated the experiment of Alvarez-Ramirez et al. (2008) and found that the same results cannot be obtained although the same software FRACLAB 2.03 (http://fraclab.saclay.inria.fr/) was employed to synthesize the two-dimensional fBm. Fig.3.4 and Fig.3.5 show our results obtained by the 2D R/S analysis.

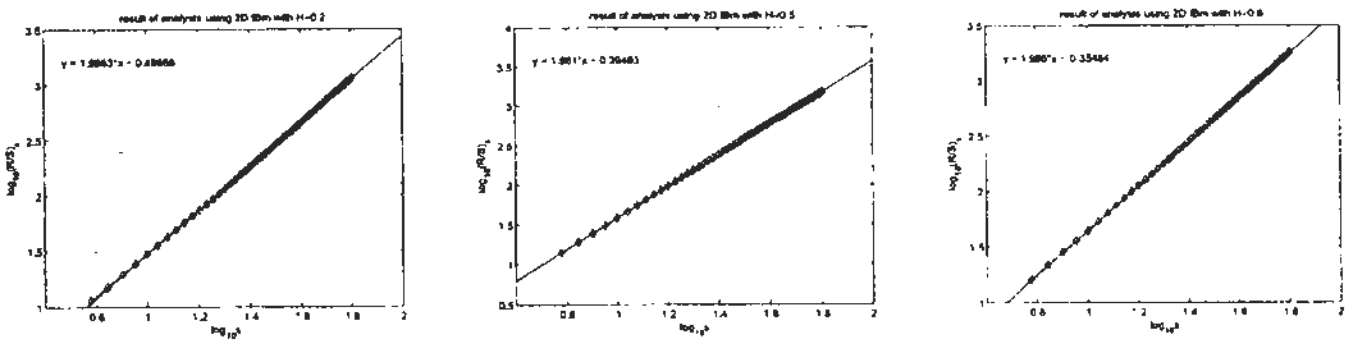

Figure 3.4: Results of 2D R/S analysis using 2D fBm with size 256×256 and known Hurst exponent $H = 0.2$(left panel), 0.5(center panel) and 0.8(right panel).

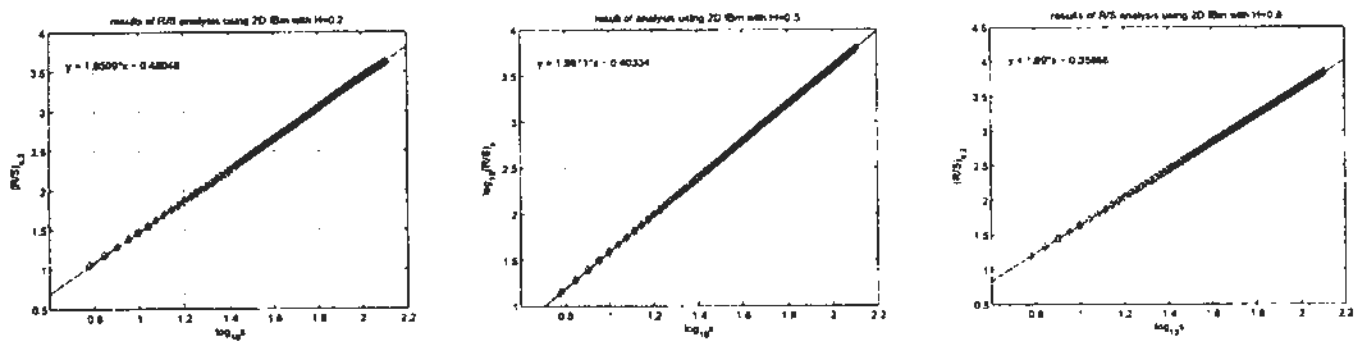

Figure 3.5: Results of the 2D R/S analysis using 2D fBm with size 512×512 and $H = 0.2$ (left panel) $H = 0.5$ (center panel) $H = 0.8$ (right panel).

It can easily be observed that no matter what the size of the surface is, 256×256 or 512×512, the slopes of all these plots are approximately 2 instead

of the pre-specified H values, i.e. 0.2, 0.5 and 0.8; or $2H$. Thus H obtained from the power law in Eq.(3.12) Alvarez-Ramirez et al. (2008) should be incorrect, at least for the two-dimensional fBm. Thus, the two-dimensional R/S analysis should still be an improper method to estimate H on non-stationary surfaces.

Problem of the Relationship of H and $h(q)$ in Gu and Zhou (2006)

The relationship between H and $h(q)$ for the one-dimensional MF-DFA is well-known. In brief, $H = h(q = 2)$ for stationary time series, whereas $H = h(q = 2) - 1$ for non-stationary series as shown in Eq.(3.21). The problem is when MF-DFA is extended to the two-dimensional space, can the relationships, like Eq.(3.57) for the two-dimensional fBm and Eq.(3.58), be obtained without considering the property of stationarity? We provide in here some numerical experiments for their verifications.

As mentioned above, two-dimensional fBm and two-dimensional fGn are famous non-stationary and stationary process respectively. Thus we used these two processes to investigate the relationship between H and $h(q = 2)$ for stationary and non-stationary processes in two-dimensional space.

We generated 20 two-dimensional fBms with size 512×512 and assigned value of H ranging from 0.1 to 0.9 with step size 0.1. According to McGaughey and Aitken (2002), the mixed second partial derivative $\partial^2/\partial x \partial y F(x, y)$ of the two-dimensional fBm is the two-dimensional fGn. Therefore, we can get one two-dimensional fGn from each two-dimensional fBm using the following approximation equation:

$$\frac{\partial^2}{\partial x \partial y} B_H(x, y) = [B_H(x, y) - B_H(x - 1, y)] - [B_H(x, y - 1) - B_H(x - 1, y - 1)], \tag{3.67}$$

where B_H denotes two-dimensional fBm, and x and y are integers. For this numerical study, we obtained 20 two-dimensional fBms and 20 two-dimensional fGns for each H, which varies from 0.1 to 0.9 with step size 0.1. We performed the experiments following exactly the procedure given by Gu and Zhou (2006). We

employed Eq.(3.49) and Eq.(3.53) as the detrending function. We found it difficult to establish any relation from applying Eq.(3.49). By detrending via Eq.(3.53), however, we obtained certain relationship depicted in Fig.3.6 and Fig.3.7.

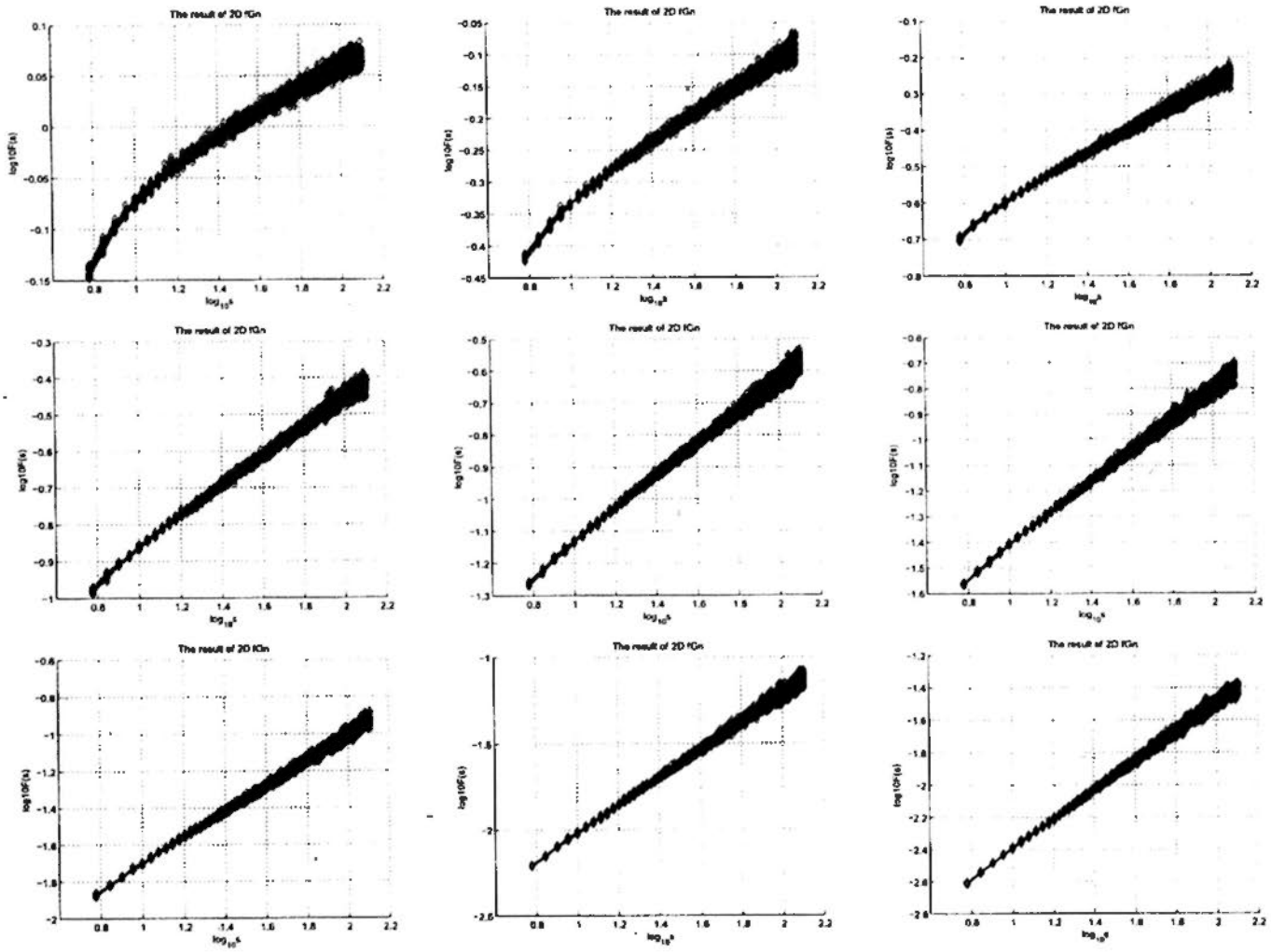

Figure 3.6: Results of 2D MF-DFA for 2D fGn with $H = 0.1$ to 0.9 (from top left to bottom right).

Table 3.3 gives the average values of $h(2)$ for the 2D fGn and fBm for each H. It is easy to read the invalidation of Eq.(3.57) for the two-dimensional fBm and Eq.(3.58) from Table 3.3.

3.4 Summary

This chapter first reviewed the popular methods which are used in fractal analysis. DFA has been selected for use in this study based on the results of

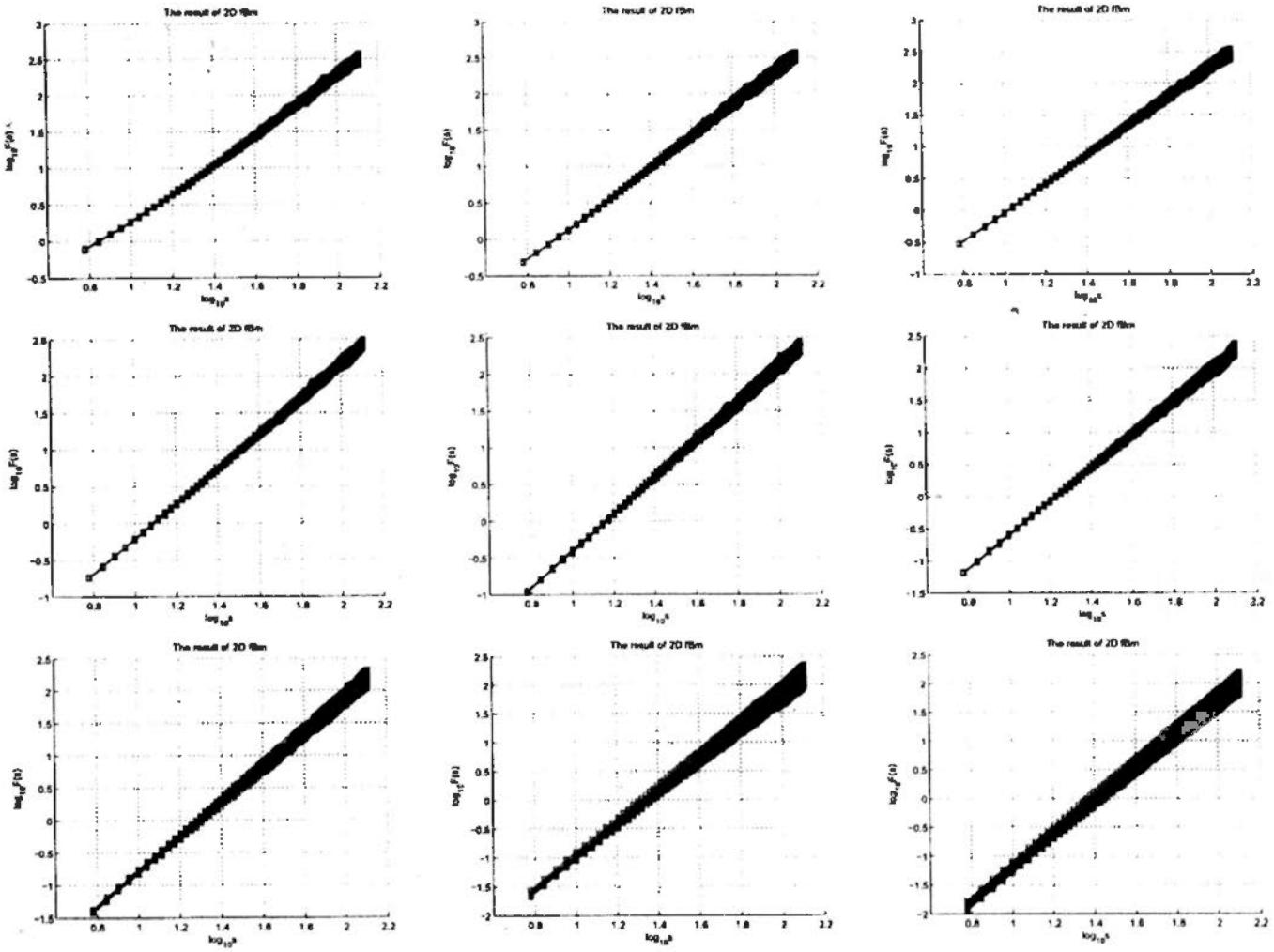

Figure 3.7: Results of 2D MF-DFA for 2D fBm with $H = 0.1$ to 0.9 (from top left to bottom right).

Table 3.3: Averaged values of $h(2)$ of the two-dimensional MF-DFA using 2D fGn and 2D fBm with different H values

H	averaged $h(2)$ of 2D fGn	averaged $h(2)$ of 2D fBm
0.1	0.12856729	2.02598325
0.2	0.21972847	2.15956748
0.3	0.31424202	2.27278913
0.4	0.40323210	2.38832299
0.5	0.50254354	2.50397095
0.6	0.60270936	2.59705262
0.7	0.69889756	2.69256248
0.8	0.80237260	2.78993299
0.9	0.89712406	2.89163698

this review. The concepts of multifractal analysis have been introduced for use in this study because of the limitations of fractals in the analysis of the real-world data; therefore, a generation of DFA (i.e. MF-DFA) is chosen for the MFA.

Two weaknesses of DFA have been identified by previous studies, namely: the strong fluctuation in the scaling behavior of DFA and MF-DFA, and the negative influence of periodic and quasi-periodic trends. In addition, two problematic relationships of MF-DFA are also claimed on the basis of theoretical discussions, the numerical experiments, and the empirical analysis. Therefore, these theoretical issues will need to be discussed in further detail before we can apply the DFA and MF-DFA in the geographical problems.

Chapter 4

Methodological Investigation of DFA and MF-DFA

This chapter will focus on the methodology issue. On the basis of review in chapter 3, there are two disadvantages in DFA and two potential problematic relationships appearing in MF-DFA. In addition, introducing the ideas from other research fields should be a appropriate way to improve the performance of the methods in analyzing the geographical problems. This chapter is organized by describing the corresponding improvements and corrections first, and then testing their capability using numerical experiments.

4.1 Improvement of DFA

Two disadvantages of DFA, strong fluctuation in scaling behavior and influence of periodic trend, have been pointed out in chapter 3. Now we would like to give our proposal to handle them in this section as follows.

4.1.1 Strong Fluctuation in Scaling Behavior

Moving Window Detrended Fluctuation Analysis (MWDFA)

Detrending is a key step in DFA. However, the current way for detrending as described in step 3 in DFA procedure may result in strong fluctuations in scaling behavior as introduced in chapter 3.

To avoid such problem, we propose a method that can take into consideration local relationships of points in time series in the detrending procedure of DFA. On the basis of this principle, the moving-window DFA (MWDFA) and the more general temporally-weighted DFA (TWDFA) are proposed in the following of this section to improve the detrending procedure in DFA, and the performance of DFA in general. The proposed model is then evaluated through numerical simulations. Note that the relationship between the Hurst exponent and h of DFA1 has been proved (Movahed and Hermanis, 2008; Movahed et al., 2006; Taqqu et al., 1995), the local trend estimated by linear regression in the detrending step of DFA is our main focus in the following discussion.

Suppose that the ith point belongs to the vth local window and the scale is still s. In step 3 of DFA, as pointed out above, the local trend at the ith point $Y(i)$ is determined by $y_v(i)$. However $y_v(i)$ is determined by the points in its local window rather than by those points, which might be outside the window, nearer to it. To account for such local effects, it is thus more reasonable to determine $y_v(i)$ using the nearby points. To achieve this, we need another way to estimate the local trend instead of using $y_v(i)$ in the conventional DFA. We propose to estimate $Y(i)$ by using the fitting polynomial $\hat{Y}(i)$ in each local window consisting of points $\{j : |i-j| <= s\}$. Then for $Y(i)$, we can get the corresponding $\hat{Y}(i)$ from using the moving window (MWi), defined as $\{j : |i-j| <= s\}$. Fig.4.1 gives a graphical illustration of the concept. That is, to capture the local effect, we move the fixed-size window along the series for detrending. Therefore, $y_v(i)$ by $\hat{Y}(i)$ in step 3 of DFA (Eq.(3.17) and (3.18)), the modified $F^2(v,s)$ and $F(s)$ can be obtained. That is, Eq.(3.17) and (3.18) can be respectively redefined as

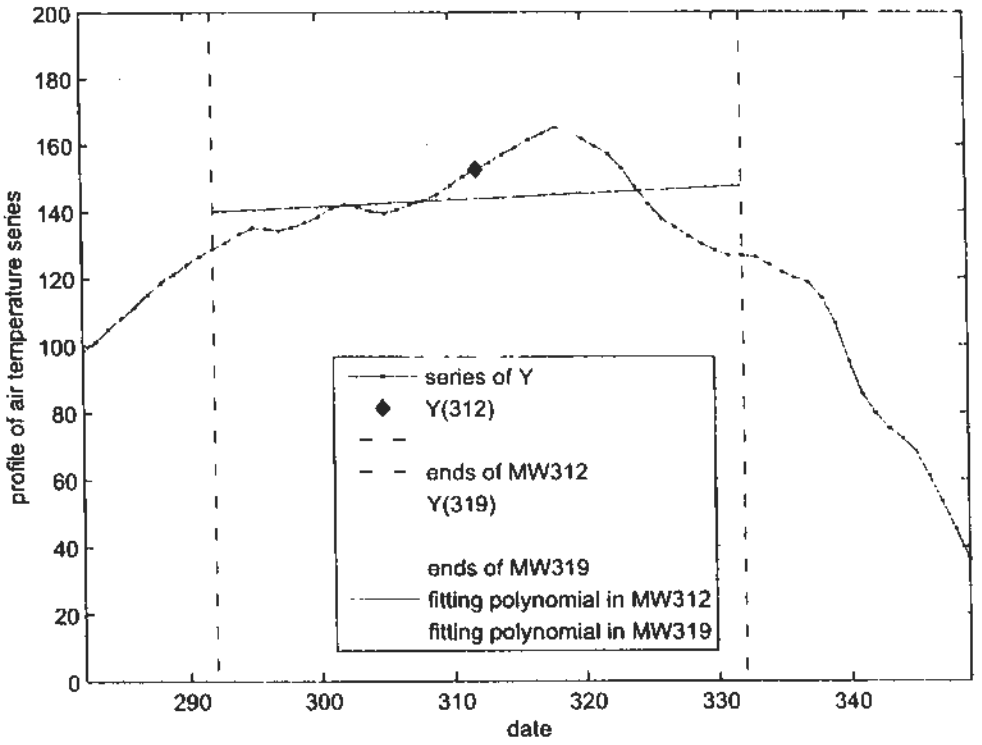

Figure 4.1: Two local moving windows (MW) along the air-temperature series and the corresponding fitting polynomial $\hat{Y}$ in these moving-windows

equations:

$$F^2(v,s) \equiv \frac{1}{s}\sum_{i=1}^{s}\{Y[(v-1)s+i] - \hat{Y}((v-1)s+i)\}^2, \quad \text{for } v = 1, \cdots, N_s; \tag{4.1}$$

$$F^2(v,s) \equiv \frac{1}{s}\sum_{i=1}^{s}\{Y[N-(v-N_s)s+i] - \hat{Y}(N-(v-N_s)s+i)\}^2, \quad \text{for } v = N_s+1, \cdots, 2N_s; \tag{4.2}$$

where $\hat{Y}(i)$ is the value of the fitting polynomial of MWi at point i. Similar to MF-DFA, from varying s we can also see if the power-law exits between $F(s)$ and s so that h can likewise be obtained. We call this modified method the moving-window detrended fluctuation analysis (WMDFA).

Multifractal Temporally-Weighted Detrended Fluctuation Analysis (MF-TWDFA)

In ordinary linear regression (OLR), parameters of the regression model are universal over the whole data space. Given n observations $(y_i; x_{i1}, x_{i2}, \ldots, x_{ip})$,$(i = 1, 2, \ldots, n)$, the OLR can be expressed as:

$$y_i = \beta_0 + \sum_{k=1}^{p} \beta_k x_{ik} + \varepsilon_i, \quad i = 1, 2, \ldots, n, \tag{4.3}$$

here $\beta_i (i = 1, 2, \ldots, p)$ are the parameters. However, this uniformity may not reflect the spatial or temporal non-stationarity in real-life situations. That is, the parameters might not be able to capture local effects due to proximity in space or time. With respect to space, some methods have been proposed to consider spatial non-stationarity or spatial drift into OLS, see for example Anselin (1988, 1990). In more general terms, Anselin (1988) have also studied the following varying-parameter regression model:

$$y_i = \beta_{i0} + \sum_{k=1}^{p} \beta_{ik} x_{ik} + \varepsilon_i, \quad i = 1, 2, \ldots, n, \tag{4.4}$$

where β_{ik} are parameters which can vary in the data space. Geographically weighted regression (GWR) is actually a kind of varying-parameter regression model taking the same form as that of Eq.(4.4) (Leung et al., 2000). By the varying parameters, the local uniformity in the spatial distribution can be well captured. Very recently, the GWR has been extended to the geographically and temporally weighted regression (GTWR) to study the heterogeneity in the spatial and temporal processes (Huang et al., 2010). Parameters in GWR can be estimated using the weighted least-squares approach. To take into consideration the local effect in time series, the idea of the varying-parameter model in general and the GWR and GTWR in particular is employed to handle local effects in DFA in our analysis (Zhou and Leung, 2010b). Moving window regression (MWR) can be treated as the special case of GWR (Lloyd, 2007). When estimating $Y(i)$

with a given MWi, MWR gives equal weight (generally, 1) to each point of MWi, while GWR allocates different weights (w_{ij}) according to the distance between i and j ($j \in$WMi). Of course, for the general varying-parameter regression model in Eq.(4.4), w_{ij} can be defined in different ways according to the situations of the geographical problems. To estimate the varying parameters for the model in Eq.(4.4) in the context of time series, we give the expressions for the moving-window method and the GWR method as follows (Lloyd, 2007):

In the MWi, the fitting polynomial can be expressed as:

$$\hat{Y}(i) = \beta_0(i) + \beta_1(i)i + \varepsilon, \tag{4.5}$$

and $\beta(i) = (\beta_0(i), \beta_1(i))^T$ can be obtained by solving:

$$\beta(i) = (T'W(i)T)^{-1}T'W(i)Y, \tag{4.6}$$

where T is a $2 \times N$ matrix:

$$T = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ \vdots & \vdots \\ 1 & N \end{pmatrix}, \tag{4.7}$$

and

$$W(i) = \begin{pmatrix} w_{i1} & 0 & \cdots & 0 \\ 0 & w_{i2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & w_{iN} \end{pmatrix}_{N \times N} \tag{4.8}$$

If we choose MWR model at scale s, w_{ij} can be defined as:

$$w_{ij} = \begin{cases} 1, & \text{if } |i-j| \leq s, \\ 0, & \text{otherwise.} \end{cases} \tag{4.9}$$

For GWR, on the other hand, the width of the weighted function affects the re-

gression model more significantly, whereas the type of the weighted function does not seem to have as an important effect. Thus, we can define w_{ij} as follows (Lloyd, 2007):

$$w_{ij} = \begin{cases} [1 - (\frac{i-j}{s})^2]^2, & \text{if } |i-j| \leq s, \\ 0, & \text{otherwise.} \end{cases} \tag{4.10}$$

Similar to the procedure of MWDFA, $h(q)$ can likewise be obtained under these two schemes. Since we define w_{ij} in terms of distance between two points in time i and j in a time series, the proposed method can aptly be called temporally-weighted detrended fluctuation analysis (TWDFA).

Actually, if we let the order of the fitting polynomial be zero, then what we obtain is the mean in each moving window. Comparing with the detrending moving average (DMA) (Alessio et al., 2002) which uses the mean in each moving window as the trend of corresponding point (the description of DMA can be referred to the relative content in chapter 3), MWDFA and TWDAF are actually one kind of combination of DFA and DMA too. One advantage of MWDFA and TWDFA taking the moving window technique is that sufficient windows for calculating the fluctuation function can be obtained. Thus a more stable scaling behavior could be obtained in the statistical sense. Consequently, the strong fluctuation could be much relieved. The performance of the modified DFAs are checked by the experimental numerical examples in the succeeding sections. Similar to the extension of DFA to MF-DFA, the MWDFA and TWDFA can also be generalized to multifractal MWDFA (MF-MWDFA) and multifractal TWDFA (MF-TWDFA) by considering the qth order moment of fluctuations. It is easy to notice that MWDFA and TWDFA should correspond to the special case $q = 2$. Thus the testing of MWDFA and TWDFA could be included in the testing of their multifractal versions.

Numerical Experiments

Since MF-WMDFA can be treated as a special case of MF-TWDFA, we just need to compare the performance of MF-TWDFA to that of MF-DFA, whose

stability and validation have been tested, in our experiments.

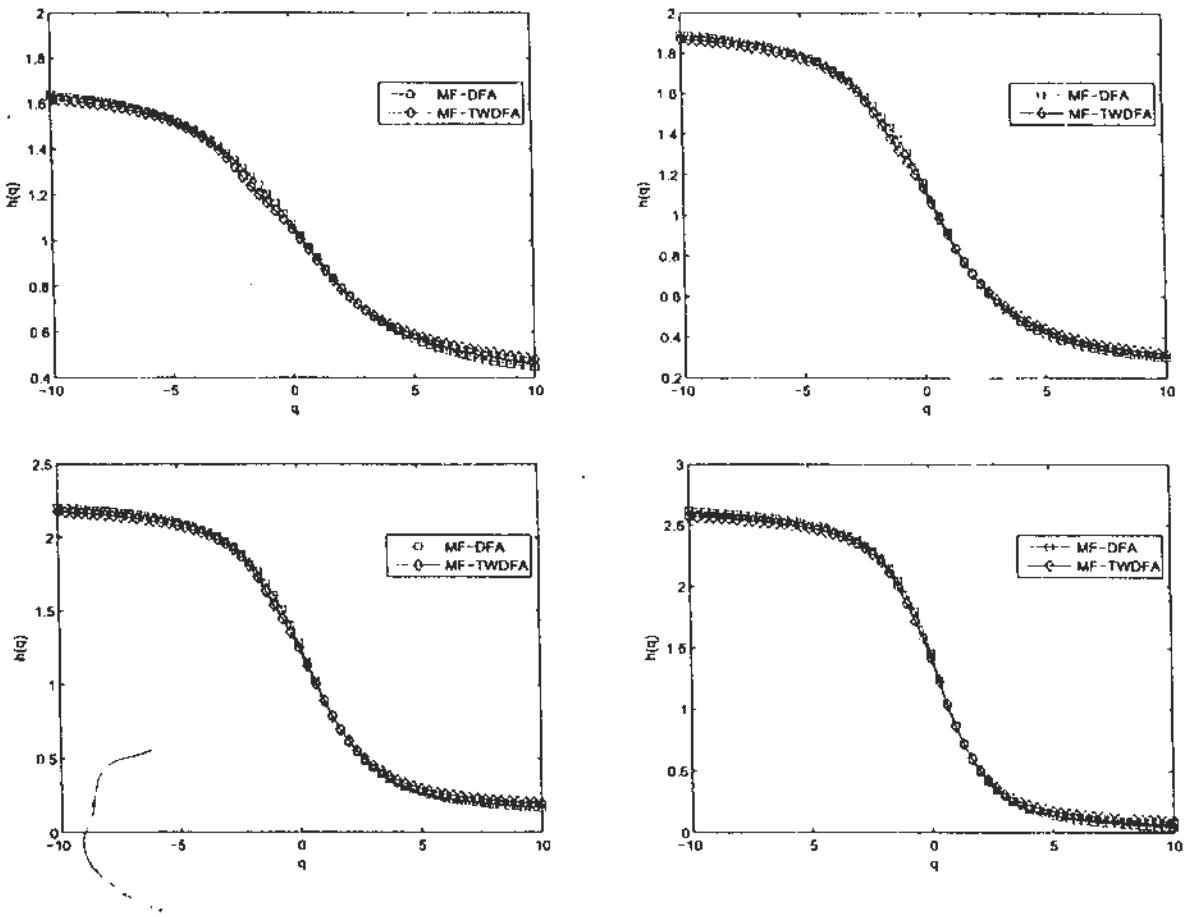

Figure 4.2: $h(q)$ of the binomial multifractal series with length 10^{12} and a ranging from 0.7 (upper left-hand panel) to 0.85 (lower right-hand panel).

Firstly, the simulated multifractal series are employed for testing the performance of MF-TWDFA by calculating $h(q)$ and studying the multifractality, as well as comparing these results with MF-DFA. We generated the binomial multifractal model (Feder, 1988), which was also employed by Kantelhardt et al. (2002), $\{x_k\}_{k=1}^{N=2^{n_{max}}}$ as:

$$x_k = a^{n(k-1)}(1-a)^{n_{max}-n(k-1)}, \qquad (4.11)$$

where the parameter a and satisfies $0.5 < a < 1$, and $n(k)$ denotes the number of digits equal to 1 in the binary representation of k. Theoretically, $h(q)$ can be expressed as:

$$h(q) = \frac{1}{q} - \frac{\ln[a^q + (1-a)^q]}{q\ln(2)}. \qquad (4.12)$$

Based on Eq. (4.12), the singularity spectrum $f(\alpha)$ vs. α can be obtained. In this thesis, we set $a = 0.7,\ 0.75,\ 0.8,\ 0.85$ with length $N = 10^{12}$. $h(q)$ obtained by MF-DFA and MF-TWDFA is presented in Fig. 4.2, from which it can be observed that two methods can obtain almost the same $h(q)$ for binomial multifractal series.

Fig. 4.3 depicts the singularity spectra of these binomial series. We can see these spectrum obtained by MF-DFA and MF-TWDFA are very similar though the multifractality is a little overestimated, if the spread of these singularity spectra is employed to measure the multifractality of series.

It is well known that fGn and fBm are stationary and non-stationary time series respectively. To compare the performances of MF-DFA and MF-TWDFA for long-range correlation analysis (e.g. here we just to consider the special case $q = 2$, MWDFA and TWDFA), we used FRACLAB 2.03 developed by INRIA (http://www.irccyn.ec-nantes.fr/hebergement/FracLab) to generate 50 fGns and 50 fBms with length 500, 1000, 1500, 2000, 2500, 3000 and set the Hurst exponent ranging from 0.1 to 0.9 with step size 0.1 to calculate $h(2)$ for testing. The numerical results of one fBm and one fGn with $H = 0.8$ are depicted in Fig. 4.4. The close agreement with the straight line with expected slope around 0.8 confirms the validity in using the TWDFA to study the scaling behavior of stationary and non-stationary series. The less strong fluctuation in the scaling behavior of TWDFA than that of DFA could be observed in Fig. 4.4. Since DFA has already been tested as a stable method for the correlation analysis, we compare the results of TWDFA with DFA directly. Defining $\Delta h(q)$ as the difference between $h(q)$ obtained by DFA and TWDFA, the $\Delta h(2)$ results listed in Table 4.1 demonstrate that for fGns and fBms with different Hurst exponents and lengths, TWDFA and DFA have equivalent performance in the correlation analysis. The high similarity in results of the two methods indicates that our proposed method is as effective as the extensively used MF-DFA in calculating H. However, we can observe from Fig. 4.4 that the plots of MF-TWDFA are smoother than that of MF-DFA. It can thus help us to find the hidden crossover point(s), particularly at large scale s, that cannot be found by conventional MF-DFA because of the

Table 4.1: Averaged $\Delta h(2)$ of MF-TWDFA and MF-DFA using fGns and fBms with different values of the Hurst exponent H and lengths

	H	$\Delta h(2) \pm \sigma(\Delta h(2))$					
		500	1000	1500	2000	2500	3000
fGn	0.1000	-0.0366 ± 0.0091	-0.0260 ± 0.0048	-0.0217 ± 0.0034	-0.0198 ± 0.0030	-0.0186 ± 0.0027	-0.0177 ± 0.0024
	0.2000	-0.0343 ± 0.0101	-0.0255 ± 0.0067	-0.0205 ± 0.0062	-0.0179 ± 0.0055	-0.0158 ± 0.0042	-0.0159 ± 0.0044
	0.3000	-0.0331 ± 0.0134	-0.0213 ± 0.0096	-0.0192 ± 0.0082	-0.0157 ± 0.0074	-0.0180 ± 0.0058	-0.0153 ± 0.0056
	0.4000	-0.0305 ± 0.0181	-0.0199 ± 0.0114	-0.0169 ± 0.0096	-0.0148 ± 0.0086	-0.0152 ± 0.0072	-0.0136 ± 0.0078
	0.5000	-0.0255 ± 0.0228	-0.0227 ± 0.0125	-0.0194 ± 0.0132	-0.0156 ± 0.0102	-0.0173 ± 0.0093	-0.0165 ± 0.0077
	0.6000	-0.0314 ± 0.0180	-0.0215 ± 0.0163	-0.0159 ± 0.0165	-0.0148 ± 0.0108	-0.0160 ± 0.0094	-0.0153 ± 0.0102
	0.7000	-0.0297 ± 0.0188	-0.0186 ± 0.0127	-0.0175 ± 0.0129	-0.0132 ± 0.0131	-0.0149 ± 0.0098	-0.0146 ± 0.0106
	0.8000	-0.0308 ± 0.0233	-0.0225 ± 0.0154	-0.0181 ± 0.0135	-0.0177 ± 0.0140	-0.0155 ± 0.0097	-0.0148 ± 0.0121
	0.9000	-0.0377 ± 0.0296	-0.0243 ± 0.0214	-0.0210 ± 0.0189	-0.0167 ± 0.0123	-0.0197 ± 0.0135	-0.0140 ± 0.0121
fBm	0.1000	-0.0046 ± 0.0268	0.0007 ± 0.0200	0.0012 ± 0.0150	-0.0022 ± 0.0128	-0.0031 ± 0.0122	0.0006 ± 0.0126
	0.2000	-0.0184 ± 0.0289	-0.0149 ± 0.0220	-0.0119 ± 0.0165	-0.0102 ± 0.0143	-0.0092 ± 0.0153	-0.0081 ± 0.0165
	0.3000	-0.0230 ± 0.0266	-0.0218 ± 0.0216	-0.0209 ± 0.0168	-0.0166 ± 0.0170	-0.0158 ± 0.0134	-0.0164 ± 0.0139
	0.4000	-0.0358 ± 0.0275	-0.0313 ± 0.0183	-0.0243 ± 0.0188	-0.0201 ± 0.0180	-0.0206 ± 0.0166	-0.0190 ± 0.0148
	0.5000	-0.0387 ± 0.0281	-0.0311 ± 0.0222	-0.0266 ± 0.0191	-0.0253 ± 0.0189	-0.0240 ± 0.0170	-0.0247 ± 0.0156
	0.6000	-0.0445 ± 0.0318	-0.0360 ± 0.0236	-0.0288 ± 0.0208	-0.0257 ± 0.0142	-0.0288 ± 0.0158	-0.0222 ± 0.0161
	0.7000	-0.0506 ± 0.0279	-0.0450 ± 0.0247	-0.0345 ± 0.0219	-0.0347 ± 0.0176	-0.0308 ± 0.0164	-0.0282 ± 0.0158
	0.8000	-0.0596 ± 0.0284	-0.0436 ± 0.0182	-0.0361 ± 0.0134	-0.0337 ± 0.0155	-0.0350 ± 0.0154	-0.0315 ± 0.0181
	0.9000	-0.0689 ± 0.0335	-0.0490 ± 0.0189	-0.0428 ± 0.0162	-0.0391 ± 0.0128	-0.0338 ± 0.0116	-0.0359 ± 0.0155

strong fluctuation at large s. Such advantage is demonstrated and quantitatively measured in the real-life problem to be discussed in the next chapter.

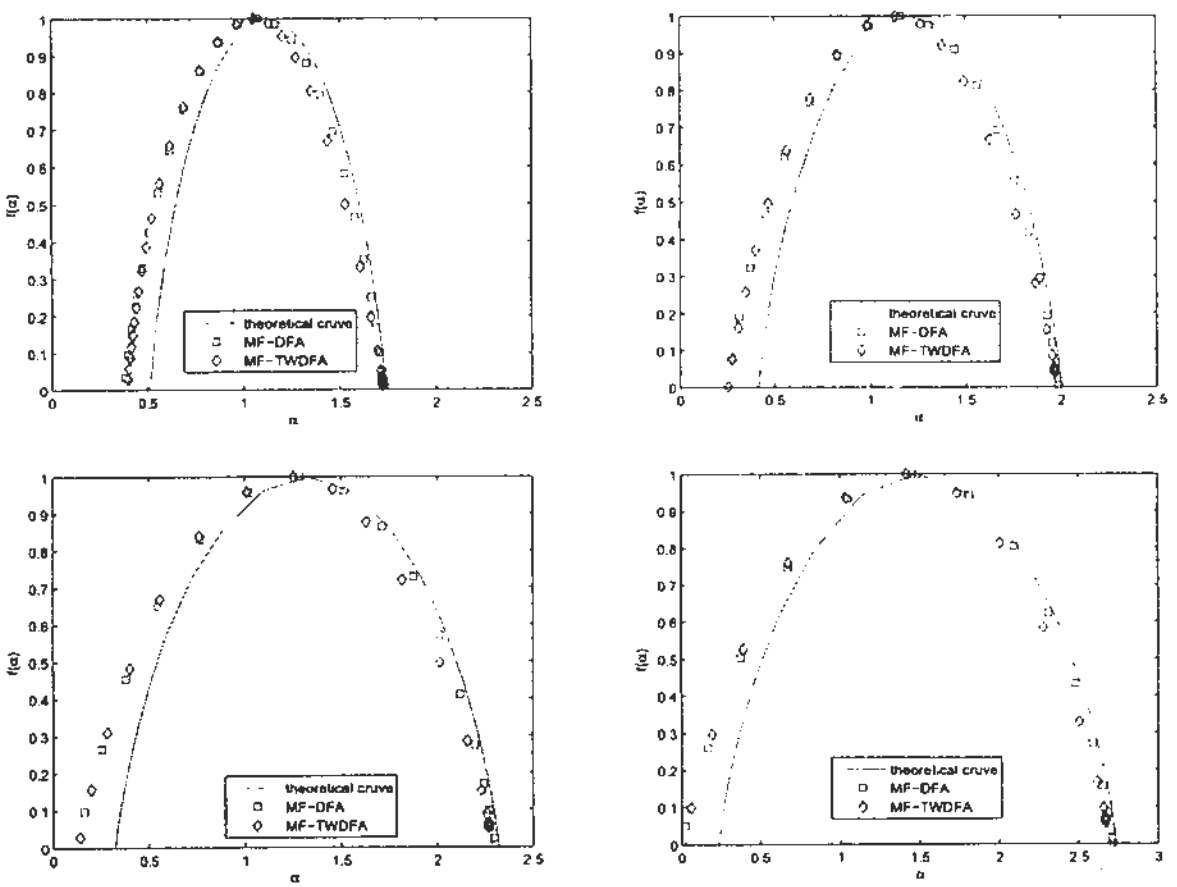

Figure 4.3: Results of MF-TWDFA and MF-DFA using binomial multifractal series with length 10^{12} and a ranging from 0.7 (upper left-hand panel) to 0.85 (lower right-hand panel). The solid line is the theoretical line derived from Eq. (4.12) and is presented here as benchmark.

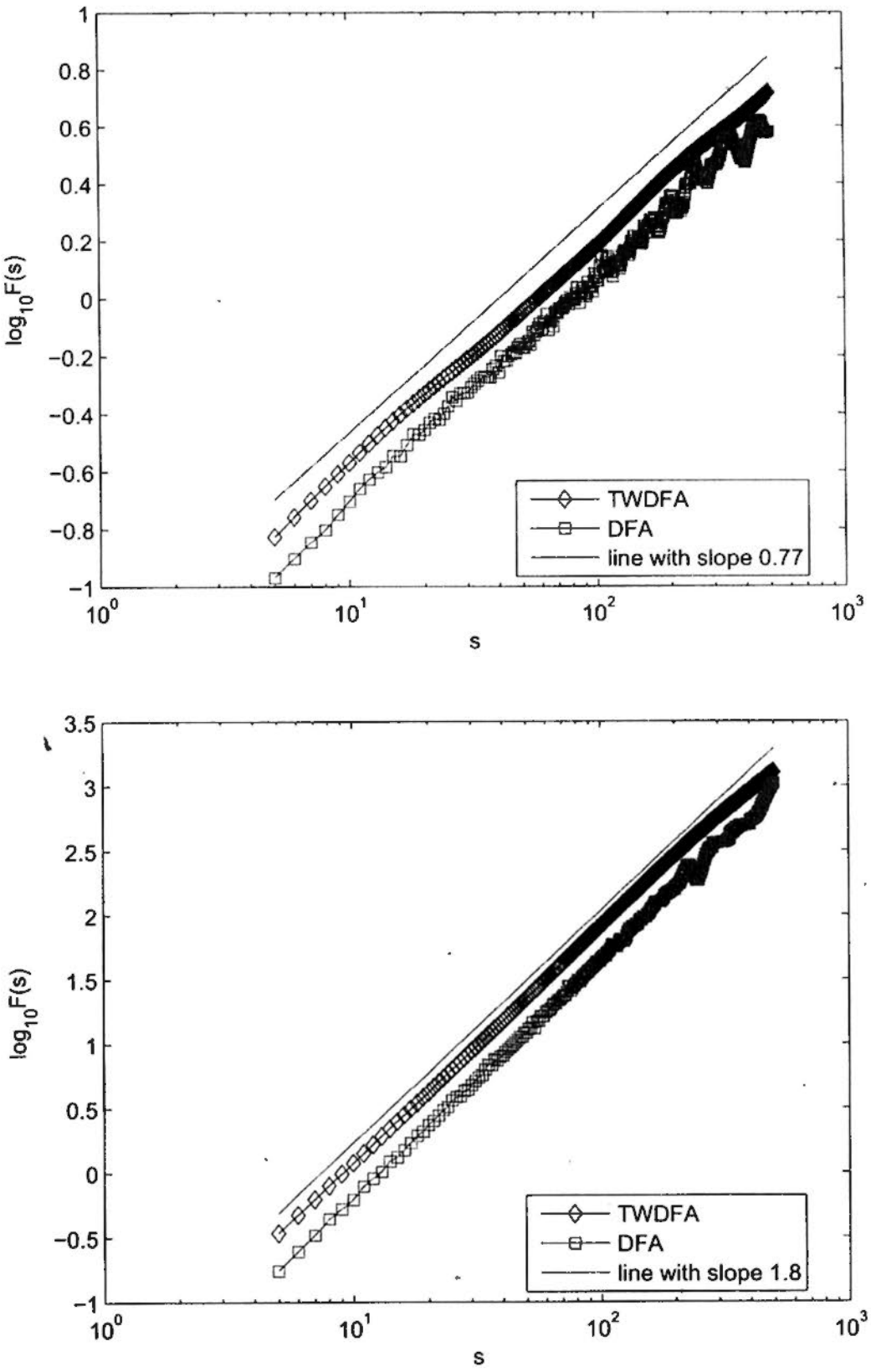

Figure 4.4: Results of MF-TWDFA and MF-DFA using fGn (upper panel) fBm (bottom panel) with length 2000 and Hurst exponent $H = 0.8$.

Analytical Calculation

To give a solid basis of the validation of our proposed modification, the analytical calculation is given in this subsection. The relation between $h(2)$ obtained by MF-TWDFA and Hurst exponent of fGns, H, is derived in this subsection using the basic ideas in Arianos and Carbone (2007); Movahed and Hermanis (2008); Movahed et al. (2006); Taqqu et al. (1995). We show that the exponent $h(2)$ of the power-law $F^2(s) \sim s^{h(2)}$ equals H of the analyzed fGns (e.g. the example $q = 2$ of the Eq. (3.39)). It is shown that $F^2(s)$ behaves as:

$$\begin{aligned} F^2(s) &= \frac{1}{N_s}\sum_{v=1}^{Ns} F^2(v,s) \\ &\sim C_H s^{2H}, \end{aligned} \tag{4.13}$$

where $F^2(v,s)$ is defined as:

$$F^2(v,s) = \frac{1}{s}\sum_{i=1}^{s}[Y_v(i) - \hat{Y}_v(i)]^2, \tag{4.14}$$

and C_H is a function of H, Y_v and $\hat{Y}_v$ denote the value of Y in the vth segment and the estimated value in the MWi respectively. Then for $q = 2$, what we have to prove actually is:

$$\begin{aligned} F^2(s) &= \frac{1}{N_s}\sum_{v=1}^{Ns} F^2(v,s) \\ &= \frac{1}{N_s}\sum_{v=1}^{Ns}\frac{1}{s}\sum_{i=1}^{s}[Y_v(i) - \hat{Y}_v(i)]^2 \\ &= \frac{1}{sN_s}\sum_{v=1}^{Ns}\sum_{i=1}^{s}[Y_v(i) - \hat{Y}_v(i)]^2 \\ &= \langle [Y(i) - \hat{Y}(i)]^2 \rangle_i \\ &\sim C_H s^{2H}. \end{aligned} \tag{4.15}$$

It should be noted that for fGn x_k, the profile $Y(i)$ is a fBm signal.

In TWDFA, the fitting function $\hat{Y}$ in MWi can be expressed as $\hat{Y}(i) = a(i) +$

$b(i)i$, here $a(i)$ and $b(i)$ can be calculated from Eq. (4.6) as:

$$\begin{aligned} b(i) &= \frac{-\sum_{j=i-\theta_1 s}^{i+\theta_2 s} j w_{ij} \sum_{j=i-\theta_1 s}^{i+\theta_2 s} w_{ij} Y(j) + \sum_{j=i-\theta_1 s}^{i+\theta_2 s} w_{ij} \sum_{j=i-\theta_1 s}^{i+\theta_2 s} j w_{ij} Y(j)}{\sum_{j=i-\theta_1 s}^{i+\theta_2 s} w_{ij} \sum_{j=i-\theta_1 s}^{i+\theta_2 s} j^2 w_{ij} - (\sum_{j=i-\theta_1 s}^{i+\theta_2 s} j w_{ij})^2}, \\ a(i) &= \frac{\sum_{j=i-\theta_1 s}^{i+\theta_2 s} j^2 w_{ij} \sum_{j=i-\theta_1 s}^{i+\theta_2 s} w_{ij} Y(j) - \sum_{j=i-\theta_1 s}^{i+\theta_2 s} j w_{ij} \sum_{j=i-\theta_1 s}^{i+\theta_2 s} j w_{ij} Y(j)}{\sum_{j=i-\theta_1 s}^{i+\theta_2 s} w_{ij} \sum_{j=i-\theta_1 s}^{i+\theta_2 s} j^2 w_{ij} - (\sum_{j=i-\theta_1 s}^{i+\theta_2 s} j w_{ij})^2}. \end{aligned} \tag{4.16}$$

To simplify the problem, we set $w_{ij} = 1$ and $\theta_1 = \theta_2 = \theta$. Then TWDFA changes to MWDFA. We believe the scaling behavior of the fluctuation function $\log(F^2(s))$ verse $\log(s)$ should be mainly maintained since the scale s bring more influence to the regression than the weight w_{ij} (Lloyd, 2007). Therefore, for example, $a(\theta s+1)$ and $b(\theta s+1)$ in MW$(\theta s+1)$ is equal to the normal regression parameters in this local window and can be expressed as (Movahed and Hermanis, 2008; Movahed et al., 2006; Taqqu et al., 1995):

$$\begin{aligned} b(\theta s+1) &= \frac{-\sum_{j=1}^{2\theta s+1} j \sum_{j=1}^{2\theta s+1} Y(j) + (2\theta s+1) \sum_{j=1}^{2\theta s+1} jY(j)}{(2\theta s+1) \sum_{j=1}^{2\theta s+1} j^2 - (\sum_{j=1}^{2\theta s+1} j)^2} \\ &\approx \frac{\sum_{j=1}^{2\theta s+1} Y(j) j - \frac{1}{2}(2\theta s+1) \sum_{j=1}^{2\theta s+1} Y(j) \sum_{j=1}^{2\theta s+1} j}{(2\theta s+1)^3/12}, \\ a(\theta s+1) &= \frac{\sum_{j=1}^{2\theta s+1} j^2 \sum_{j=1}^{2\theta s+1} Y(j) - \sum_{j=1}^{2\theta s+1} j \sum_{j=1}^{2\theta s+1} jY(j)}{(2\theta s+1) \sum_{j=1}^{2\theta s+1} j^2 - (\sum_{j=1}^{2\theta s+1} j)^2} \\ &\approx \frac{1}{2\theta s+1} \sum_{j=1}^{2\theta s+1} Y(j) - \frac{b(\theta s+1) \cdot (2\theta s+1)}{2}. \end{aligned} \tag{4.17}$$

Using the above two equations and taking the self-similarity of fGn and fBm into consideration, $\langle [Y(i) - \hat{Y}(i)]^2 \rangle$ can be written as follows:

$$\begin{aligned} \langle [Y(i) - \hat{Y}(i)]^2 \rangle &= \langle [Y(\theta s+1) - \hat{Y}(\theta s+1)]^2 \rangle \\ &= \langle [Y(\theta s+1) - a(\theta s+1) - b(\theta s+1) \cdot (\theta s+1)]^2 \rangle \\ &\approx \langle Y^2(\theta s+1) - \frac{2Y(\theta s+1) \sum_{j=1}^{2\theta s+1} Y(j)}{2\theta s+1} \\ &\quad -Y(\theta s+1) b(\theta s+1) + \frac{(\sum_{j=1}^{2\theta s+1} Y(j))^2}{(2\theta s+1)^2} \end{aligned}$$

$$+\frac{(\sum_{j=1}^{2\theta s+1} Y(j))b(\theta s+1)}{(2\theta s+1)}+\frac{1}{b^2(\theta s+1)}\rangle. \qquad (4.18)$$

In the proof of the relation between the DFA scaling exponent and H in Movahed and Hermanis (2008) and Taqqu et al. (1995), they assume that s is large enough. Thus, we make the same assumption in our proof also. It is then easy to see $\forall x>0$, $(C_1\theta s+C_2)^x \sim C_1^x s^x$. Employing this relation and inserting the expression of b into Eq. (4.18), it becomes:

$$\begin{aligned}
\langle[Y(i) - \hat{Y}(i)]^2\rangle \approx\ & \langle Y^2(\theta s+1)\rangle \\
& +\langle\frac{9}{16}\frac{(\sum_{j=1}^{2\theta s+1} Y(j)j)^2}{\theta^6 s^6}\rangle \\
& +\langle(\frac{(-18\theta s+12\theta^2 s^2)}{16\theta^6 s^6})\sum_{j=1}^{2\theta s+1} Y(j)\sum_{j=1}^{2\theta s+1} Y(j)j\rangle \\
& -\langle\frac{3}{2}\frac{Y(\theta s+1)}{\theta^3 s^3}\sum_{j=1}^{2\theta s+1} Y(j)j\rangle \\
& +\langle\frac{(-16Y(\theta s+1)\theta^5 s^5+24Y(\theta s+1)\theta^4 s^4)}{16\theta^6 s^6}\sum_{j=1}^{2\theta s+1} Y(j)\rangle \\
& +\langle\frac{(-12\theta^3 s^3+4\theta^4 s^4+9\theta^2 s^2)}{16\theta^6 s^6}(\sum_{j=1}^{2\theta s+1} Y(j))^2\rangle \\
=\ & \langle Y^2(\theta s+1)\rangle \\
& +\frac{9}{16}\frac{\langle(\sum_{j=1}^{2\theta s+1} Y(j)j)^2\rangle}{\theta^6 s^6} \\
& +(\frac{(-18\theta s+12\theta^2 s^2)}{16\theta^6 s^6})\langle\sum_{j=1}^{2\theta s+1} Y(j)\sum_{j=1}^{2\theta s+1} Y(j)j\rangle \\
& -\langle\frac{3}{2}\frac{Y(\theta s+1)}{\theta^3 s^3}\sum_{j=1}^{2\theta s+1} Y(j)j\rangle \\
& +\langle\frac{(-16Y(\theta s+1)\theta^5 s^5+24Y(\theta s+1)\theta^4 s^4)}{16\theta^6 s^6}\sum_{j=1}^{2\theta s+1} Y(j)\rangle \\
& +\frac{(-12\theta^3 s^3+4\theta^4 s^4+9\theta^2 s^2)}{16\theta^6 s^6}\langle(\sum_{j=1}^{2\theta s+1} Y(j))^2\rangle. \qquad (4.19)
\end{aligned}$$

It should be noted that $Y(i)$ is the fBm produced by the fGn. Then there are

some properties as follows Movahed and Hermanis (2008); Taqqu et al. (1995):

$$Y(i) \approx i^H, \tag{4.20}$$

$$\langle \sum_{i=1}^{2\theta s+1} Y^2(i) \rangle \approx \frac{(2\theta s+1)^{2H+1}}{2H+1} - \frac{(2\theta s)^{2H+1}}{2H+1}, \tag{4.21}$$

$$\begin{aligned} \langle (\sum_{j=1}^{2\theta s+1} Y(j)j)^2 \rangle &\approx \frac{(2\theta s+1)^{2H+4}}{4(H+1)}(1 - \frac{1}{(H+2)(2H+1)}) \\ &= \frac{(2\theta s)^{2H+4}}{4(H+1)}(1 - \frac{1}{(H+2)(2H+1)}), \end{aligned} \tag{4.22}$$

$$\begin{aligned} \langle \sum_{j=1}^{2\theta s+1} Y(j) \sum_{j=1}^{2\theta s+1} Y(j)j \rangle &\approx \frac{(2\theta s+1)^{2H+3}}{4}(\frac{2}{H+1} - \frac{1}{2H+1}) \\ &= \frac{(2\theta s)^{2H+3}}{4}(\frac{2}{H+1} - \frac{1}{2H+1}), \end{aligned} \tag{4.23}$$

$$\langle \sum_{j=1}^{2\theta s+1} Y(j)j \rangle \approx \frac{(2\theta s+1)^{H+2}}{H+2} = \frac{(2\theta s)^{H+2}}{H+2}, \tag{4.24}$$

$$\langle \sum_{j=1}^{2\theta s+1} Y(j) \rangle \approx \frac{(2\theta s+1)^{H+1}}{H+1} = \frac{(2\theta s)^{H+1}}{H+1}, \tag{4.25}$$

$$\langle (\sum_{j=1}^{2\theta s+1} Y(j))^2 \rangle \approx \frac{(2\theta s+1)^{2H+2}}{2H+2} = \frac{(2\theta s)^{2H+2}}{2H+2}. \tag{4.26}$$

Then with the assumption s is large enough, we can obtain:

$$\langle [Y(i) - \hat{Y}(i)]^2 \rangle \sim C_H(\theta s)^{2H} \sim C_H s^{2H}, \tag{4.27}$$

here C_H is:

$$C_H = \frac{2^{2H}}{2H+2} - \frac{2^{H+1}}{H+1} + 1. \tag{4.28}$$

Then the validation of the MF-TWDFA has been tested above through the numerical experiments and analytical calculation.

4.1.2 Influence of Periodic Trend

The negative influence of the periodic trend on the scaling behavior in DFA has been mentioned in the chapter 3. The effect of periodic trend can be shown in the following example. A stationary series with Hurst exponent $H = 0.9$ with

length 3123, is constructed by applying the Fourier filtering method described in Makse et al. (1996). This series and its scaling behavior are shown in Fig. 4.5. After superimposing a quasi-periodic trend, a typical scaling behavior of DFA under the influence of this kind of trend as indicated by Hu et al. (2001) and the superimposed series could be found in Fig. 4.6.

Since only the polynomial trend is considered in the detrending step in DFA, the influence of periodic trend cannot be eliminated totally. Then some pre-detrending processing seems to be necessary before DFA. There are some methods proposed to handle the periodic trend, such as higher order regression analysis, Fourier-based filtering (Nagarajan and Kavasseri, 2004), SVD (Nagarajan and Kavasseri, 2005) and the adaptive detrending methods (Hu et al., 2009). However, their disadvantages and the different results from them (for example difference between Movahed et al. (2006) and Hu et al. (2009) using DFA on the same sunspot series) triggered us to adopt another angle to get to the bottom of this problem.

In Hu et al. (2009), empirical mode decomposition (EMD) (Huang et al., 1998) is considered an inappropriate method to extract trends from sunspot time series. Nevertheless, EMD is proposed to handle nonlinear and non-stationary signals. EMD is a totally data-driven adaptive method without any a priori assumptions about the function forms, time scale, stationarity or linearity (Huang et al., 1998; Wu et al., 2007). It can extract many basic components, called intrinsic mode functions (IMFs), from the original signal to detail the structural information. It looks like the Fourier decomposition or wavelet methods. However, the IMFs with different dominant frequencies are with physical meanings that cannot be conveyed by the Fourier or wavelet methods (Rilling et al., 2003; Sinclair and Pegram, 2005). In fact, EMD has been employed to study the cycle of sunspot time series of the monthly record from 1848~1992 and 1894~2003 respectively (Li et al., 2007; Xu et al., 2008). Compared with the conventional analysis, many periodic components with real physical meanings can be found using EMD. It is claimed in Li et al. (2007); Xu et al. (2008) that the IMFs can correspond to

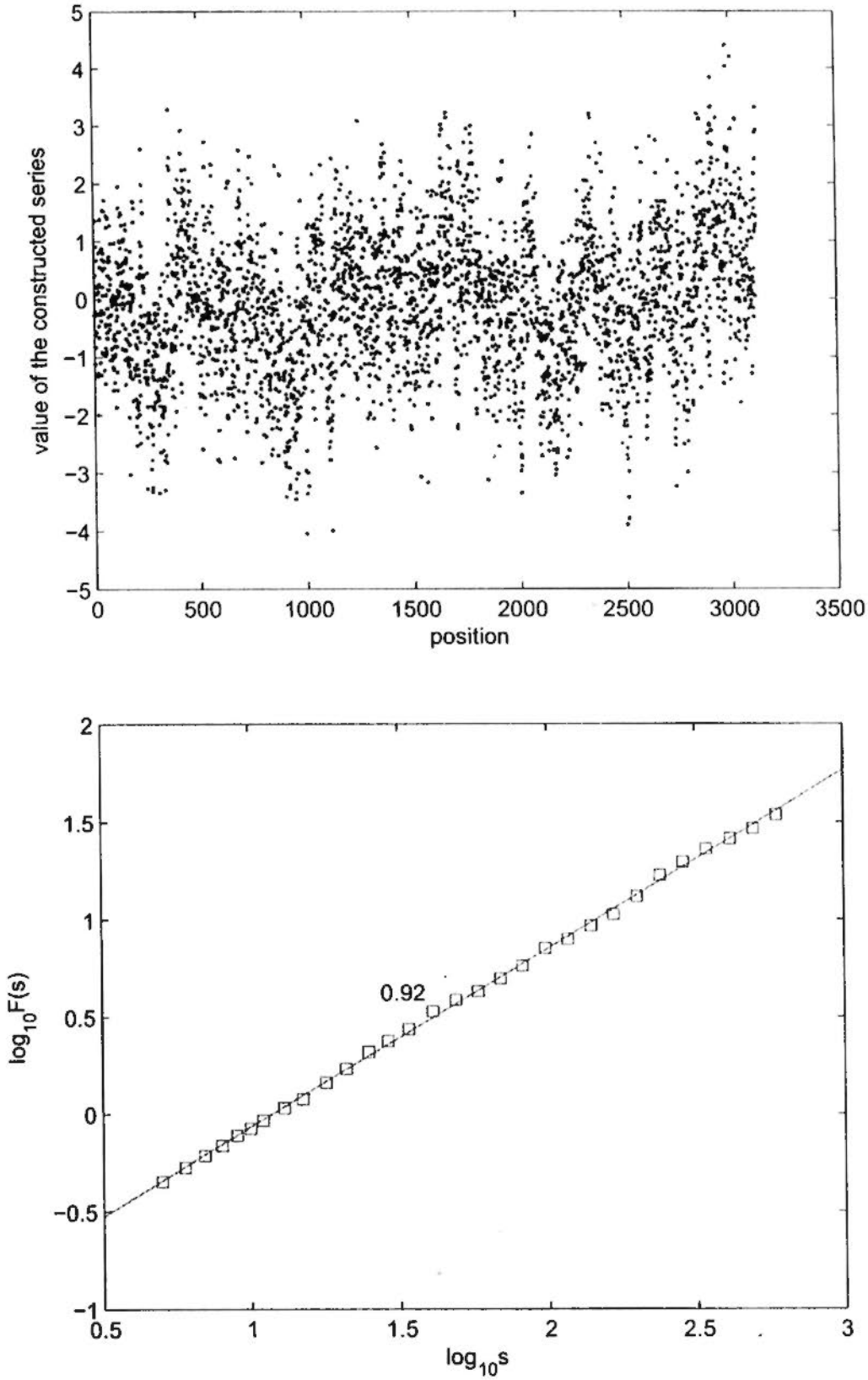

Figure 4.5: The constructed series (upper panel) and its scaling behavior (bottom panel), the marked slope 0.92 is the estimated Hurst exponent.

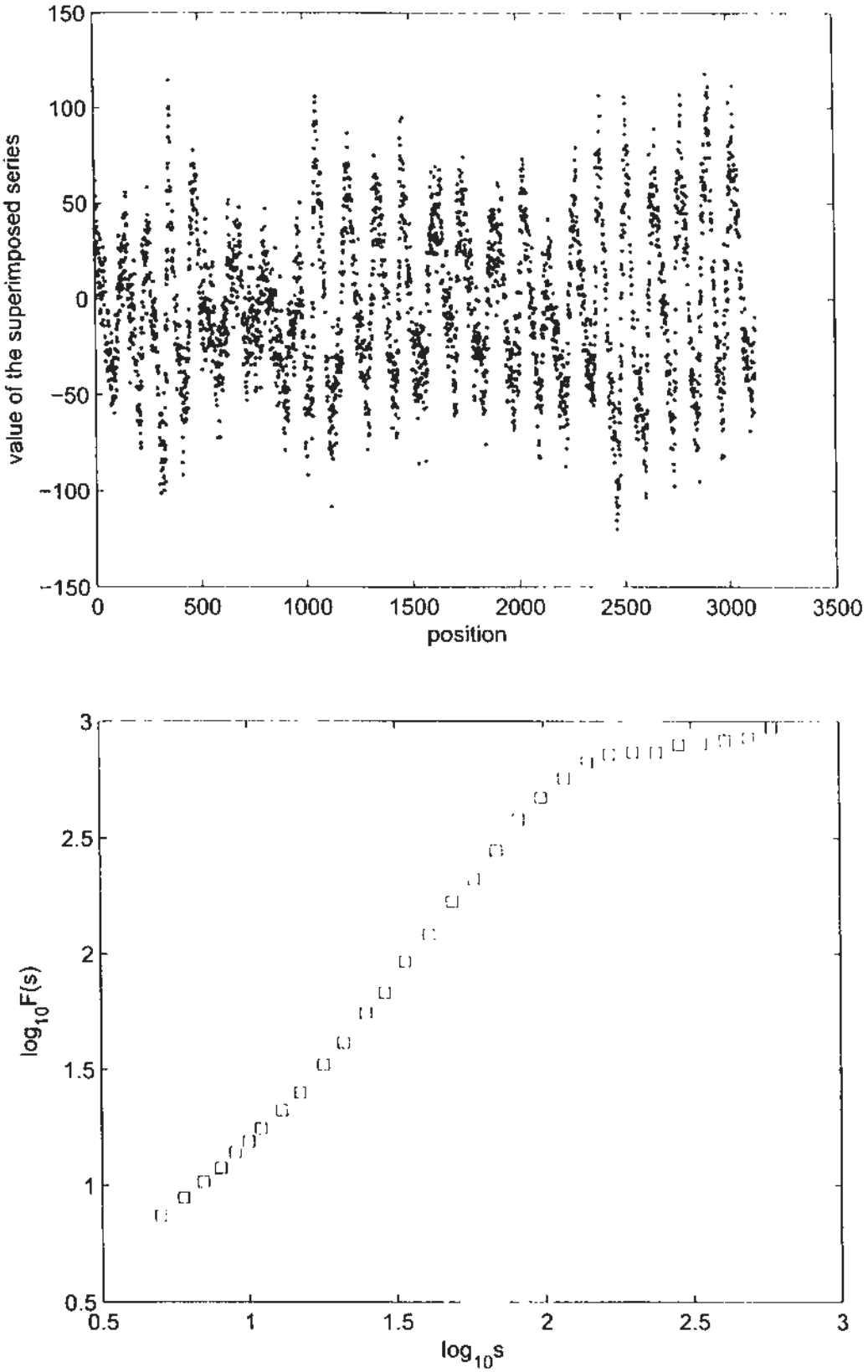

Figure 4.6: The superimposed series (upper panel) and its scaling behavior (bottom panel).

some well-known cycles of solar activities, such as the cycle observed in solar wind with quasi-period of about 1.3−1.4 years (Richardson et al., 1994), quasi-biannual oscillation (QBO) about 25−30 months (Naujokat, 1986), 11-year cycle, Hale period of 20-24 years (Wilson, 1988), double Hale period (Fairbridge and Hillaire-Marcel, 1977) and possible Gleissberg period of around 80 years (Gleissberg, 1944).

In order to compare the performances of different periodic-trend-removing methods, we firstly introduce them briefly, especially the Fourier truncation and the adaptive detrending methods which cause the contradictory results by Movahed et al. (2006) and Hu et al. (2009), and the EMD-based methods proposed by us in following.

Fourier Truncation Method

Generally, one series can be treated as an combination of many components in the sinus and cosinus forms with different frequencies. Fourier transform can decompose the given series to these components. These components can be employed to reconstruct the original series using inverse Fourier transform. If perform the Fourier transform on the series $\{x_k\}_{k=1}^N$, then a series with equal length $\{\hat{x}_j\}_{j=1}^N$ can be obtained as:

$$\hat{x}_j = \sum_{i=k}^{N} x_k \omega_N^{(k-1)(j-1)}. \tag{4.29}$$

And $\{\hat{x}_j\}_{j=1}^N$ can be employed to reconstruct the original series via the inverse Fourier transform as below:

$$x_k = (1/N) \sum_{j=1}^{N} \hat{x}_j \omega_N^{-(j-1)(k-1)}. \tag{4.30}$$

Here $\omega_N = e^{(-2\pi i)/N}$, is an Nth root of unity. For each j, $\hat{x}_j$ corresponds certain component with certain frequency. The larger j corresponds the higher frequency. Then if eliminate $\hat{x}_j$, e.g. set $\hat{x}_j = 0$, the information of this component will be

removed in the reconstructed series. Movahed et al. (2006) removed the first 50 terms of the Fourier transforms of sunspots series. That means all the frequencies smaller than the one corresponding to $\hat{x}_{50}$ was eliminated. Then Movahed et al. (2006) found the crossover points existing in the scaling behavior of the original sunspots series disappeared in the log – log plot of the detrended series. And they claimed the affect of the around 11-year cycle of sunspots series was removed.

Adaptive Detrending Method

In 2009, Hu et al. proposed this data-driven method to estimate the local trend of given series. And then apply it on removing the around 11-year cycle from the sunspots series to obtain the scaling plot without the influence of this cycle. Brief description of the adaptive detrending method can be given as (Hu et al., 2009):

- *Step 1*: Divide $\{x_k\}_{k=1}^{N}$ into segments of length $2n+1$. The continuous segments share $n+1$ common points. The last segments might consist of points less than $2n+1$.

- *Step 2*: Then for the jth segment, a fitting polynomial of order K can be obtained and denoted as $x_l^{(j)}$ with $l = 1, 2, \cdots, 2n+1$.

- *Step 3*: Now, the trend for the overlapped region can be defined as:

$$x_l^{(c)} = (1 - \frac{l-1}{n})x_{l+n}^{(i)} + \frac{l-1}{n}x_l^{(i+1)}, \quad l = 1, 2, \cdots, n+1. \tag{4.31}$$

Here, overlapped segments and temporal dependence of the neighboring points are considered to avoid any jumps or discontinuities around the ends of neighboring segments. In addition, two parameters, the segment length, $2n+1$, and the order of fitting polynomial, K, should be consider to estimate the local trend. As Hu et al. (2009) claimed, the length of segment can be set according to the scale of considered trend. And K can be determined by the variance plot (detailed information can be referred to the Fig. 2 in Hu et al. (2009)).

Empirical Mode Decomposition (EMD)-Based Method

EMD is an empirical, intuitive, data-driven and adaptive method that makes no requirements on predetermined basis function and stationarity or linearity assumption about time series (Huang et al., 1998; Wu et al., 2007). The concept of IMF plays a key role in EMD analysis. An IMF is a function satisfying: 1) the difference between the number of extrema and that of zero-crossings is one at most, and 2) the mean value of the envelope of either the local maxima or local minima is zero (Wu et al., 2007). Given an original time series/signal, $X = \{x(t)\}$, the method of EMD can be summarized as follows (Huang et al., 1998; Wu et al., 2007):

- *Step 1*: Identify all local extrema and connect all local maxima and minima with cubic spline to obtain the upper and lower envelope, ENV_{max} and ENV_{min}, respectively;
- *Step 2*: Calculate the average envelope M using $M = (\text{ENV}_{max}+\text{ENV}_{min})/2$ and obtain h through $h = X - M$;
- *Step 3*: Determine whether or not h is a IMF. If it is not a IMF, repeat step 1 and step 2 on h until the envelops have zero-mean under certain stopping criteria, which is the name sifting process by Huang et al. (1998). If h is a IMF, take h as the first component, IMF_1;
- *Step 4*: Perform step 1 and step 2 on $X - \text{IMF}_1$ to extract the second component IMF_2. Repeat the procedure until all k IMFs are extracted from the original series X.

Based on the algorithm above, we have $X = \sum_{i=1}^{k} \text{IMF}_i + r$, where r is the residual from which no more IMFs can be extracted. In the present study, the stopping criteria of the sifting process proposed by Rilling et al. (2003) is employed to guarantee globally small fluctuations in the mean while taking into account locally large excursions. The EMD implemented in this thesis employed the Matlab codes written by G. Rilling and P. Flandrin from Laboratoire de Physique CNRS & ENS Lyon (France): http://perso.ens-lyon.fr/patrick.flandrin/emd.html.

These extracted IMFs are usually physical meaningful. Then it enable us to consider the influence of these components and their combinations on the scaling behavior of analyzed series. Then, the procedure of EMD-based method can be summarized as follows (Zhou and Leung, 2010a):

- *Step 1*: Employ EMD to decompose the analyzed series to IMFs and r, with increased dominant frequencies which can be estimated by Fourier power spectrum;
- *Step 2*: Select the corresponding components according to the frequencies of periodic trends and requirements of discussion, like studying the effect of different components on the scaling behavior;
- *Step 3*: Remove the selected components and sum the others remained to obtain the detrended series.

Numerical Experiments

To test the reliability of our proposed methods, two well-controlled numerical experiments with some known properties are designed to check and compare the EMD-based method, Fourier truncation, and the adaptive detrending method. By comparing the outputs of these three methods with the expected results, their performances can be quantitatively evaluated. We first construct a stationary series with known Hurst exponent without crossover points in its scaling behavior. Different trends are then superimposed onto this series. These three detrending methods are subsequently performed in order to remove the effects of these trends. By studying the scaling behavior of the detrended series, we can evaluate the capabilities of the methods by evaluating their differences with the expected Hurst exponent.

The first experiment is the example shown in Figs. 4.5 and 4.6. The superimposed periodic trend is actually the cycle extracted by EMD which corresponds to the 11-year cycle, from sunspots series. We then employ Fourier truncation, adaptive detrending method and EMD-based methods with settings determined

by the added periodic trend, e.g. removing the first 70 and 18th~70th terms of Fourier transforms coefficients, setting length of segments, $2n+1$, and choosing order of detrending polynomial, K, as $2n+1=61$ and $K=2$ respectively Hu et al. (2009), and setting the removed IMFs as the 5th and 6th ones. The visualization of EMD results, and the scaling behavior of the series after removing $\sum_{i=5}^{6}\mathrm{IMF}_i$ and $\sum_{i=5}^{8}\mathrm{IMF}_i + r$ are shown in Fig. 4.7. The results of Fourier truncation and adaptive detrending method are shown in Fig. 4.8.

This numerical experiment shows that none of the three methods can totally remove the effect of the added trends. Besides, a crossover point appears at the position similar to the sunspots series. All slopes estimated for the first scaling range are around 0.92, which is the expected value. The fluctuations of the series by removing the first 70 Fourier transform coefficients and by the adaptive detrending method on the larger scales remain almost unchanged, which means the information of scales larger than the crossover points is lost in the detrending procedure. However, the exponents obtained from the EMD-based method and Fourier truncation in the other setting are 0.94 (Fig. 4.7) and 1.08 (Fig. 4.8) respectively, which are roughly close to the expected value, 0.92. It can be observed that considerable, although not all, information on the larger scales are maintained after detrending by these two methods.

In the first numerical experiment, the cycle of the superimposed periodic trend is not very exact. In the second numerical experiment, a periodic trend with exact cycle equalling 500 (first row of plots in the upper panel of Fig. 4.9) is superimposed onto the constructed series. Results obtained by the similar experimental procedure are depicted in Fig. 4.9. It shows the EMD results and the comparison between the scaling behaviors of the original and detrended series. And, Fig. 4.10 depicts the results of Fourier truncation and the adaptive detrending methods.

Again, all three detrending methods are unable to completely remove the effect of the added periodic trend since there are the crossover points in all results. And the scaling behavior on scales larger than the crossover points are different from the expected pattern. We conjecture that the original structure of the series on

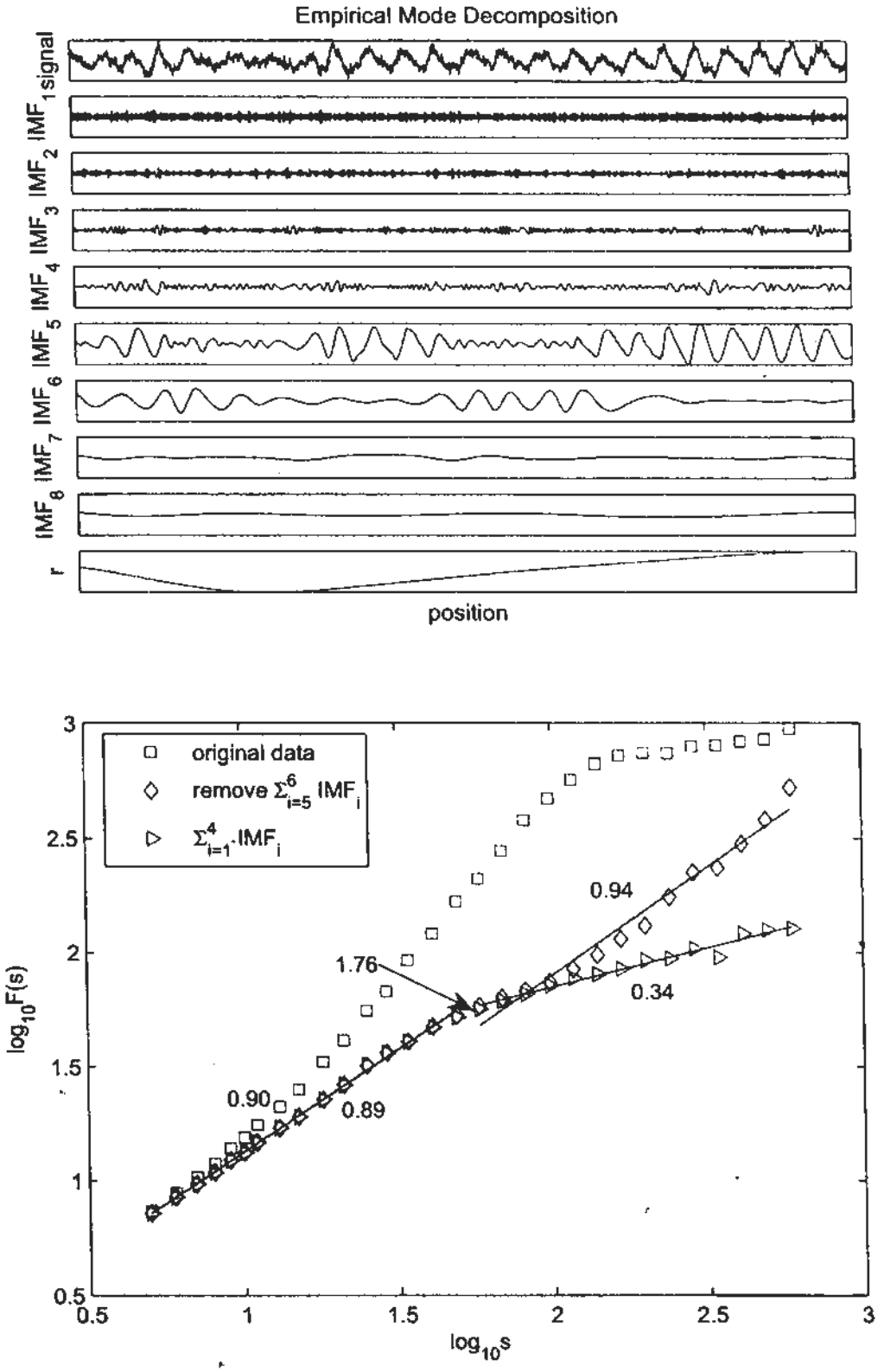

Figure 4.7: The EMD results of the superimposed series (upper panel) and the scaling behavior from removing $\sum_{i=5}^{6}\mathrm{IMF}_i$ and $\sum_{i=5}^{8}\mathrm{IMF}_i + r$, e.g. $\sum_{i=1}^{4}\mathrm{IMF}_i$ (bottom panel).

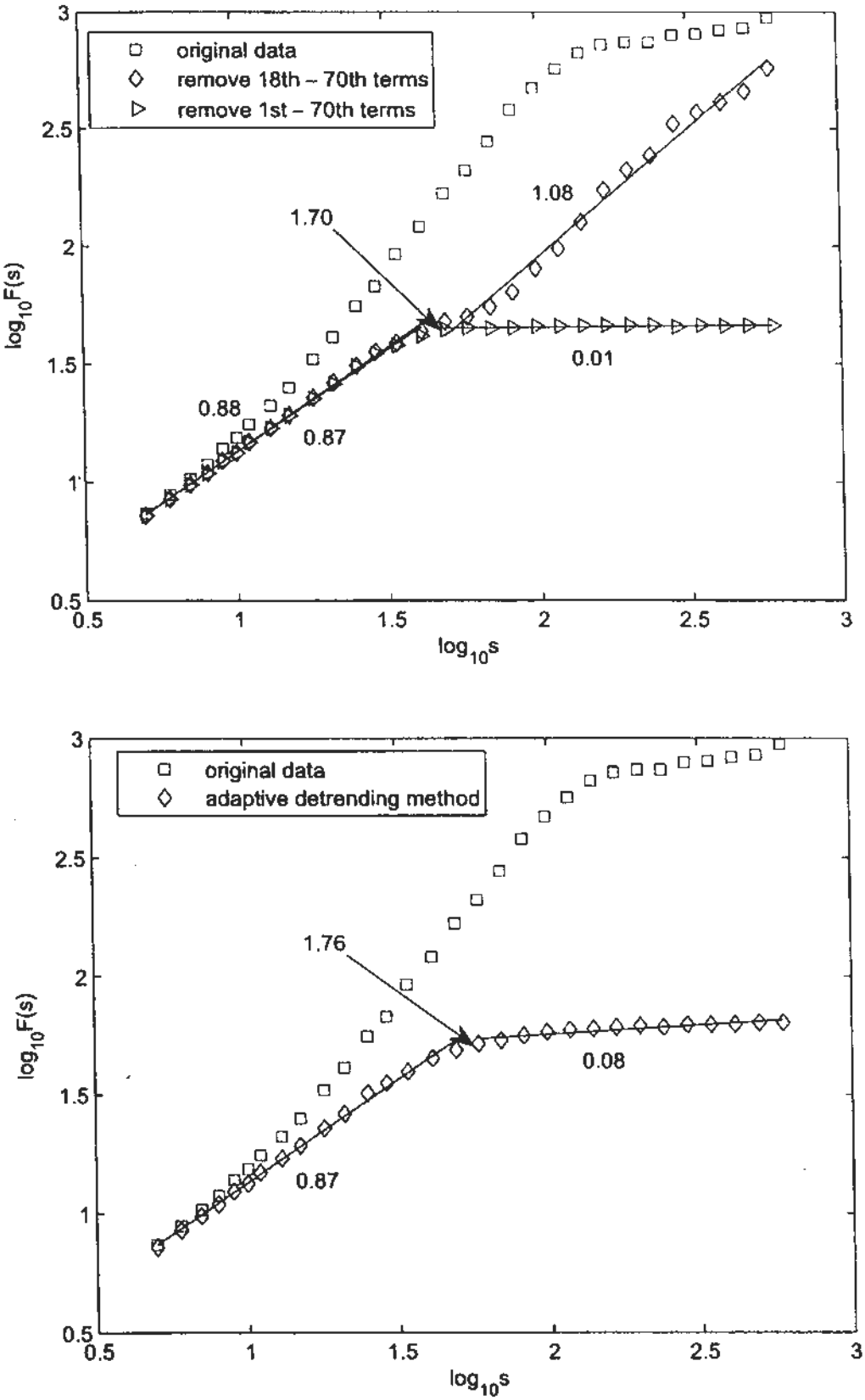

Figure 4.8: The scaling behavior of Fourier truncation (upper panel) and the adaptive detrending method (bottom panel).

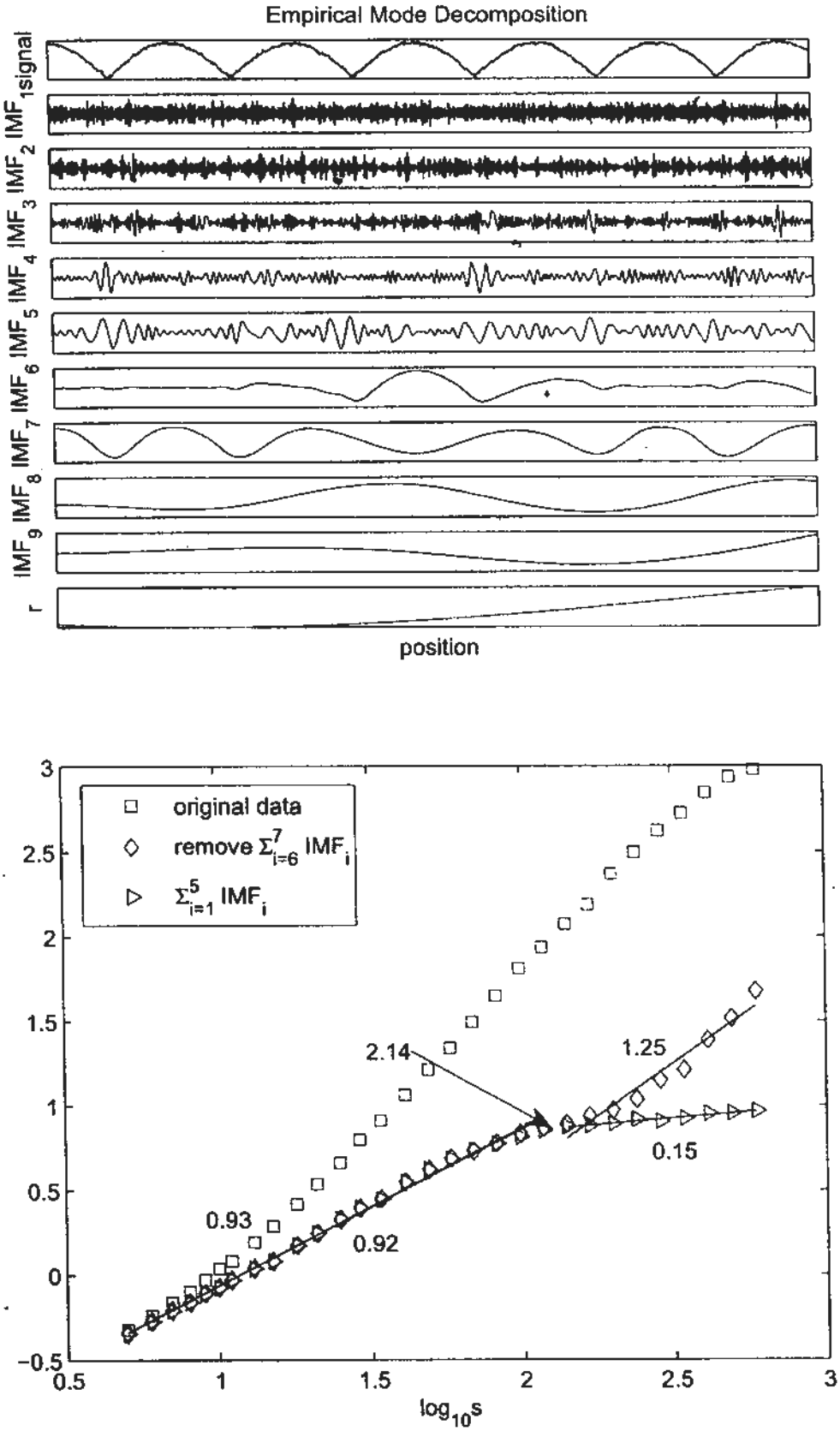

Figure 4.9: The EMD results of superimposed series (upper panel) and the scaling behavior of removing $\sum_{i=6}^{7}\mathrm{IMF}_i$ and $\sum_{i=6}^{9}\mathrm{IMF}_i + r$, e.g. $\sum_{i=1}^{5}\mathrm{IMF}_i$ (bottom panel).

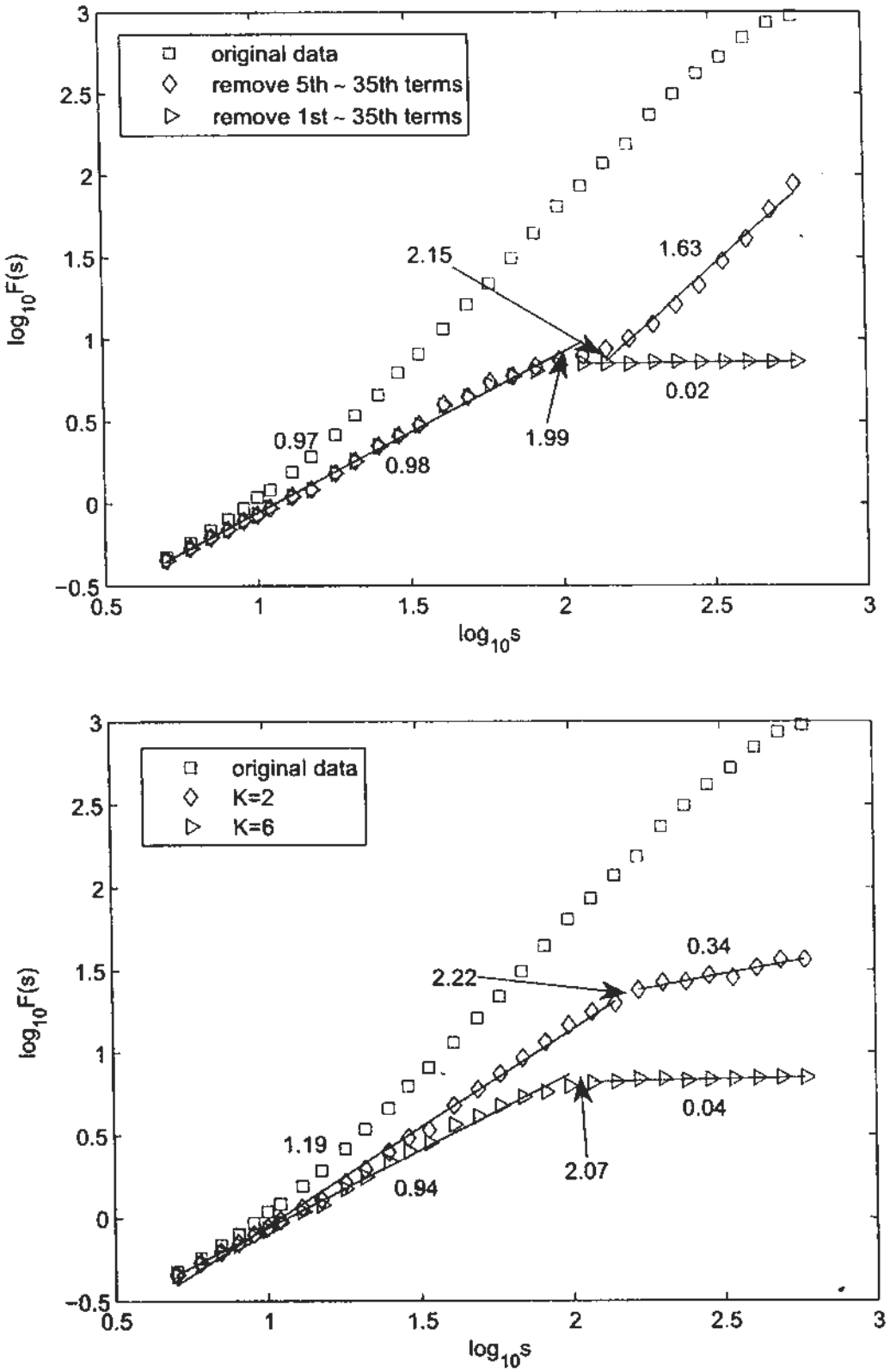

Figure 4.10: The scaling behavior of Fourier truncation (upper panel) and the adaptive detrending method (bottom panel) with different parameters.

relative large scales might be destroyed during the detrending procedure. With the appropriately selected parameters, however, the three methods can obtain the results close to the expected value, 0.92, for the relatively small scaling range. Thus, selection of the relevant parameters becomes significant. For EMD-based methods, we only choose the components with relevant frequencies mostly from $11 \sim 13$ IMFs. However, for Fourier truncation method, we have to face large amount of the Fourier coefficients. In this numerical experiment, the dominant frequency of the added trend should be $1/500$ month^{-1}, which corresponds to the 6th and 7th coefficients because $3123/500 \approx 6.4$. However, after removing the $6 \sim 35$th or $1 \sim 35$th coefficients, the Fourier method can result in the correct scaling exponent. With regards to the adaptive detrending method, Hu et al. (2009) gave their rules to select the parameters, n and K. In this experiment, n can be obtained by $2n + 1 = 251$ depending on the cycle of the trend. According to the variance method suggested in Hu et al. (2009), K should be selected as 2 (Fig. 4.11). However, it can be seen in the right panel of Fig. 4.10 that the result, 1.19, when $K = 2$ is significantly different from 0.92. While the exponent 0.94, which is closer to 0.92, is obtained when $K = 6$.

As a summary of these two numerical experiments, the three detrending methods appear to be effective in detrending the effect of periodic trends on the relatively small scales for the further analysis of MF-DFA if the parameters are appropriately selected. For larger scaling range, however, it is difficult to remove the effect of periodic trends completely. The pseudo-crossover-points might appear because of the possible changes of the scaling structure in large scaling range during the detrending process. Generally, the IMFs do have meaning, physically. Thus the crossover points unraveled by the EMD-based method are usually intrinsic to the structure of a series. However, the crossover points obtained by Fourier truncation and the adaptive detrending method are generally artifacts of the selected parameters, such as frequencies to be removed and the length of segment for detrending. Although with appropriate parameters, these two methods can obtain results similar to that of the EMD-based method, it seems that

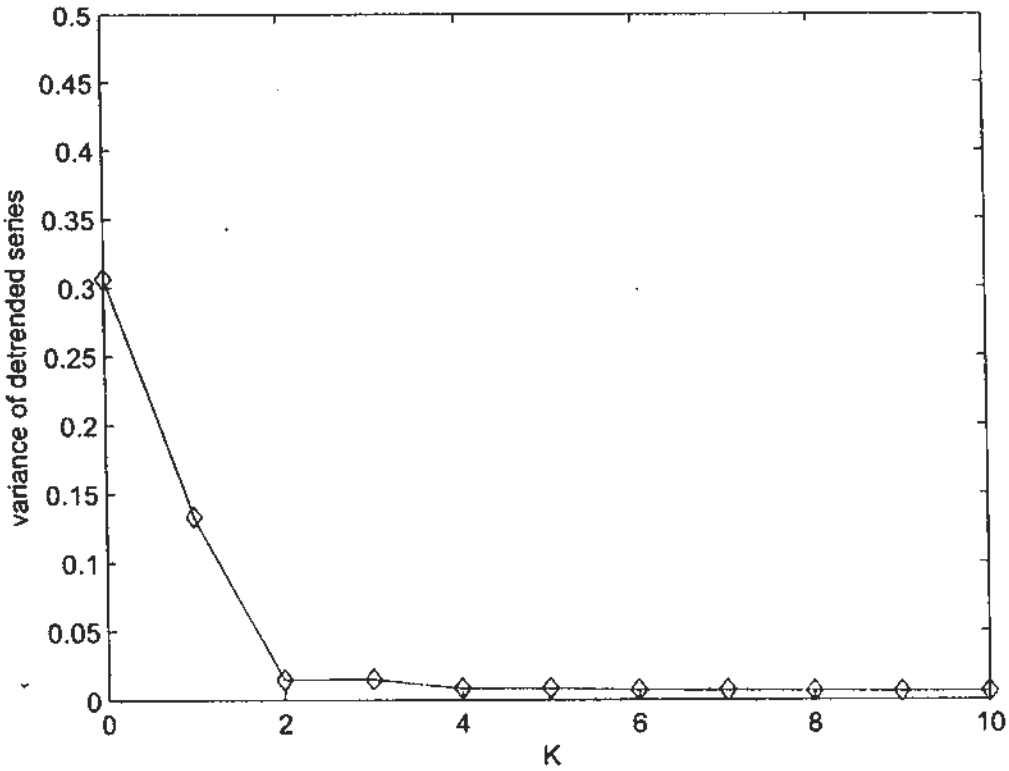

Figure 4.11: Selection of K by the variance method in Hu et al. (2009)

the EMD-based method is a more natural way for the study of intrinsic scaling behaviors. Furthermore, its parameters selection procedure is much easier. With regard to the situation on scales larger than the crossover points, the adaptive detrending method almost wipe out all information, while the EMD-based and the Fourier methods manage to keep considerable information even though the results are a bit off the expected pattern. It should be noted that by comparing Fig. 4.7, Fig. 4.9 with Fig. 4.8, Fig. 4.10, the EMD-based method leads to results, including the slopes and crossover points, similar to that of the Fourier method. However, the results obtained by the adaptive detrending method, though still acceptable, differ rather significantly from that of the other two methods.

4.2 Correction of Problematic Relationships in MF-DFA

(i) In the development of the one-dimensional MF-DFA, Kantelhardt et al. (2002) established a relationship, which has been extensively applied ever since, between the one-dimensional MF-DFA and the standard partition-function-based multifractal formalism as: $\tau(q) = qh(q) - 1$. However, its incorrectness has been shown in our investigation from theoretical and experimental aspects in chapter 3. Based on the universal multifractal formalism proposed and developed by Lavallee et al. (1993); Lovejoy et al. (2008); Schertzer and Lovejoy (1987), the relationship between $h(q)$ and $\tau(q)$ can be established as $h(q) = qh(q) - qH' - 1$, which was firstly suggested by Yu et al. (2009) without detailed analysis in their study of the daily rainfall in the Pearl River basin of China. We would like to discuss this relationship in detail in this section.

(ii) Extending on the one-dimensional MF-DFA, Gu and Zhou (2006) gave two relationships: $h(q) \equiv H$ for fBm in the two-dimensional space and $h(q = 2) \equiv H$, in their development of the two-dimensional MF-DFA. The invalidation of their relationships has also been pointed out in chapter 3. Instead, we propose a postulated relationship as follows: $H = h(2)$ for the 2D fGn, and $h(2) - 2$ for the 2D fBm. in this section.

4.2.1 Problematic Relationship Between $h(q)$ and $\tau(q)$ in MF-DFA

In this subsection, we focus on the stationary, positive, and normalized time series in our discussion. In the universal multifractal formalism, there are three parameters α, C_1 and H', called the multifractal index, the codimension and the nonconservation parameter respectively (Lovejoy et al., 2008). Comparing the codimension multifractal formalism based on $K(q)$ and $c(\gamma)$ (Schertzer and Lovejoy, 1987) with the dimension multifractal framework based on $\tau(q)$ and $f(\alpha)$ (Halsey et al., 1986), we can observe that some relationships have already been

established, particularly

$$\tau(q) = (q-1)D - K(q), \tag{4.32}$$

where D is the dimension of the observing space (for time series, $D = 1$) (Halsey et al., 1986; Lovejoy et al., 2008). Besides, there is an interesting relationship for $q \geq 0$ in Seuront et al. (1999):

$$\zeta(q) = qH' - \frac{C_1}{\alpha - 1}(q^\alpha - q), \tag{4.33}$$

with

$$K(q) = \frac{C_1}{\alpha - 1}(q^\alpha - q), \tag{4.34}$$

and $\zeta(q)$ is the exponent obtained from the structure function $S_q(s) \equiv < |x_{k+s} - x_k|^q > \sim s^{\zeta(q)}$ (see Davis et al. (1994) for more details about the structure function and $\zeta(q)$). According to Koscielny-Bunde et al. (2006),

$$\zeta(q) = qh(q) - q, \tag{4.35}$$

and then

$$qh(q) = q + \zeta(q) = (H' + 1)q - K(q). \tag{4.36}$$

It should be noted that if $q = 1$, then according to Eq.(4.34) $K(q = 1) \equiv 0$, Eq.(4.36) becomes

$$h(1) = H' + 1. \tag{4.37}$$

Combining Eq.(4.32) and Eq.(4.36), we can obtain for $q \geq 0$ (Zhou et al., 2011)

$$\tau(q) = qh(q) - qH' - 1. \tag{4.38}$$

Thus, the original relationship in Eq.(3.44), $\tau(q) = qh(q) - 1$, differs from the one we derived in Eq.(4.38) by $-qH'$.

On one hand, if the two equations are compared using the vertically shifted

fGn, which were employed to check the original relationship Eq.(3.44), with Eq.(3.45) we can then find for $q = 0$, both Eq.(3.44) and Eq.(4.38) lead to $\tau(0) = -1$ so that $D(0) = \frac{\tau(0)}{0-1} = \frac{-1}{-1} = 1$. Since experimental examples are monofractal, $D(q) \equiv$ constant and $h(q) \equiv H$ (H is the Hurst exponent of the vertically shifted fGn here, which could be any value of 0.1, 0.2, $\cdots$, 0.9) so that $D(q) = D(0) = 1$ and $\tau(q) = D(q)(q-1) = q - 1$. However, the right hand side of Eq.(3.44), $qh(q) - 1$, equals $qH - 1$. Thus, Eq.(3.44) is valid only for $H = 1$. Based on Eq.(4.37), we obtain $H' = H - 1$. Then we can calculate the right hand side of Eq.(4.38), $qh(q) - qH' - 1$, and it indeed equals $\tau(q)$ since $qH - q(H-1) - 1 = qH - qH + q - 1 = q - 1$. Thus Eq.(4.38) should be the correct relationship for the vertically shifted fGn.

On the other hand, we investigate the two equations using the binomial cascade model. Lovejoy et al. (2008) described the nonconservation parameter H' as $H' = 0$ meaning that the set can be modeled as a pure multiplicative process, i.e. the set can be conserved from scale to scale. According to Lavallee et al. (1993), the multiplicative processes were first developed as models of turbulent cascades. An example of such multiplicative process in two dimensional space is depicted in Figure 8.6 in Lavallee et al. (1993) which is actually a cascade model in two dimensional space. Therefore, the binomial cascade model is just the one which satisfies the $H' = 0$ condition. Actually, the equation (19) and (20) in Kantelhardt et al. (2002) ensure $h(1) = 1/q - \ln[a^q + (1-a)^q]/(q\ln(2))|_{q=1} = 1 - 0 = 1$, which indicates $H' = h(1) - 1 = 0$. Under this situation, Eq.(3.44) and Eq.(4.38) become identical.

Though our proposed relationship in Eq.(4.38) is established on the basis of non-negative q, the analysis and conclusion of the two examples above are actually independent of the sign of q. That means the new relationship is also valid for negative q in terms of the above two examples. Thus we think the generalized relationship in Eq.(4.38) should be valid for all q. We would like to discuss this problem from the empirical points of view.

Take the vertically shifted fGn with Hurst exponent H_0, which could be any

value of 0.1, 0.2, $\cdots$, 0.9, as an example, it is known that the vertically shifted fGn is monofractal (Mandelbrot, 1982). Then its generalized fractal dimension $D(q) = D \equiv 1$ and $h(q) = H_0$. As given by Halsey et al. (1986), $\tau(q) = D(q) \cdot (q-1) = q - 1$. However, $qh(q) - 1 = qH_0 - 1$. Then Eq.(3.44) does not hold unless $H_0 = 1$. Therefore, the validity of the relationship Eq.(3.44) is questionable. Empirically, $\tau(q)$ should be 0 for $q = 1$ to ensure $D(1)$, which comes from $D(1) = \tau(q)/(q-1)|_{q=1}$, finite. Therefore, there should be one additional term, $A(q)$, in Eq.(3.44) to ensure $\tau(1) = 0$. It is easy to obtain $A(1) = h(1) - 1$ for $q = 1$ if set $\tau(1) = h(1) - A(1) - 1 = 0$. Then we can further derive the formula of $A(q)$ from the special situation, the fGn, as mentioned above. It is already known that $\tau(q) = q - 1$. Then if $qH_0 - A(q) - 1$ is expected to be equal to $q - 1$, it can be obtained that $A(q) = q \cdot (H_0 - 1)$. Taking into account that $h(q) \equiv H_0$ for fGn and $A(1) = h(1) - 1$, the new relationship can be established empirically as $\tau(q) = qh(q) - q \cdot (h(1) - 1) - 1 = qh(q) - qH' - 1$. Therefore, it is reasonable to say that the original relationship, Eq.(3.44) proposed in Kantelhardt et al. (2002), only holds for the special situation, $H' = 0$ or $h(1) = 1$, of the general relationship in Eq.(4.38). And Lovejoy et al. (2008) pointed out that there is always the case in geophysical phenomena and turbulence that the observable have $H' \neq 0$, like $H' > 1$ for temperature and pressure in the verticals and $H' \approx -0.35$ for rain rate in time on (climate) scales larger than one month (Tessier et al., 1996).

In addition, the numerical examples are employed again to check the new relationship we just obtained as Eq.(4.38). The illustrative figures of $\tau(q)$ versus $qh(q) - qH'$ are shown in Fig.4.12. The linear fits of actual curve conform with the expected relationship $\tau(q) = qh(q) - qH' - 1$ very well. The slopes and intercepts of the linear fits in Table 4.2 are very close to the 1 and -1 respectively which solidly support our new relationship.

Since H' seems to be similar to the Hurst exponent, H, its meaning needs to be discussed here. What should be emphasized here is that although for some cases, H' and H are numerically equal, these two exponents have different physical meanings. It is already known that $\zeta(2) = \beta - 1$ Davis et al. (1994).

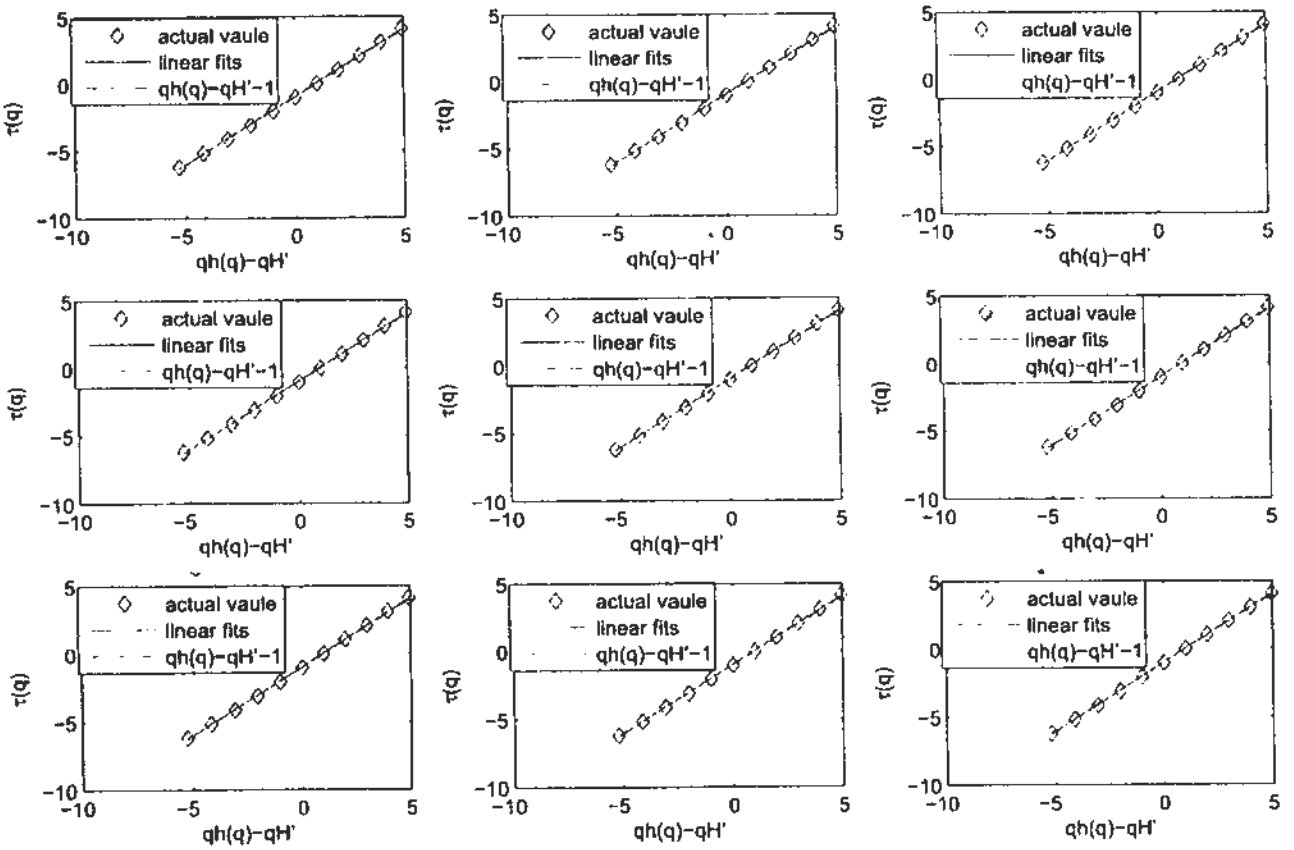

Figure 4.12: Relationships between $h(q)$ and $\tau(q)$ for numerical examples with H=0.1 to 0.9 (from top left to bottom right).

Table 4.2: Calculated right side of the new relationship Eq. (4.38) for numerical examples with different H values in Fig. 4.12

H	coefficient of $qh(q) - qH'$	intercepts of linear fits
0.1	1.0114	-0.9526
0.2	1.0126	-0.9562
0.3	1.0134	-0.9612
0.4	1.0129	-0.9626
0.5	1.0156	-0.9727
0.6	1.0157	-0.9743
0.7	1.0166	-0.9737
0.8	1.0172	-0.9785
0.9	1.0203	-1.0032

For monofractal example, such as fBm, there is $\tau(q) = D \cdot (q-1)$ and $h(q) \equiv H$, then from Eq.(4.32), we can obtain $K(q) \equiv 0$. Thus according to Eq.(4.33) and Eq.(4.38), there is $\zeta(q) = qH'$. Specifically, for $q = 2$, we have $2H' = \beta - 1$. Then $H' = \frac{\beta-1}{2}$ which is the Hurst exponent from another definition in Flandrin (1992); Mandelbrot (1982), H_1. However, for the multifractal examples, $K(q) = D \cdot (q-1) - \tau(q)$ should not always equal zero. Then H' should be $\frac{\beta-1}{2} + \frac{K(2)}{2} = H_1 + \frac{K(2)}{2}$. For the case $K(q) \neq 0$, H' is different from H numerically. Actually, H' is the nonconservation parameter in the universal multifractal formalism, and $H' = 0$ means that the set can be modeled as a pure multiplicative process, i.e. the set can be conserved from scale to scale. While H is the Hurst exponent which can be employed to measure long-range correlation of series, and $H = 0$ indicates the strong long-range anti-correlation. It is not uncommon to see they have different values. Take the binomial multifractal model as an example, $H' = 0$ as mentioned above. However, $H = h(2) = 1/2 - \ln(a^2 + (1-a)^2)/(2\ln 2)$ according to Eq. (20) in Kantelhardt et al. (2002). Then it obviously is not equal to 0.

Based on this new relationship (Eq.(4.38)), instead of Eq.(3.48), we can further obtain the singular spectrum, α and $f(\alpha)$, as follows:

$$\alpha = \frac{\mathrm{d}}{\mathrm{d}q}\tau(q) = h(q) + q\frac{\mathrm{d}}{\mathrm{d}q}h(q) - H', \tag{4.39}$$

$$f(\alpha) = q\alpha - \tau(q) = q(\alpha - h(q) + H') + 1. \tag{4.40}$$

Therefore, H' should be estimated before we calculate the multifractal spectrum using $h(q)$ of the MF-DFA.

4.2.2 Problematic Relationship Between H and $h(q)$ in the 2D MF-DFA

In chapter 3, the problem in the relationship between H and $h(q)$ in the two-dimensional MF-DFA has been shown. It has been recognized that the 2D R/S analysis should still be inappropriate for 2D non-stationary process. And the Eq.(3.57) for the two-dimensional fBm and Eq.(3.58) given by Gu and Zhou

(2006) is also problematic.

Actually, some relationship could be read from Table 3.3. Looking carefully into Table 3.3, it could be found that except for the situation $H = 0.1$ for the 2D fBm, the empirical results appear to suggest the following relationship between H and $h(2)$:

$$H = \begin{cases} h(2), & \text{for the 2D fGn;} \\ h(2) - 2, & \text{for the 2D fBm.} \end{cases} \tag{4.41}$$

According to Movahed and Hermanis (2008), 1D DFA gives inaccurate results for strongly anti-correlated time series when H is close to zero. Thus the result obtained for $H = 0.1$ for 2D fBm may be due to the weakness of the method in capturing strong anti-correlation of signals. Besides, 2D fGns and fBms are the most common classical stationary and non-stationary examples respectively. In summary, for the 2D situation, we postulate that the relationship between the Hurst exponent H and $h(q = 2)$ should be as follows:

$$H = \begin{cases} h(2), & \text{for the 2D stationary signal;} \\ h(2) - 2, & \text{for the 2D non-stationary signal.} \end{cases} \tag{4.42}$$

Such relationship is actually in synchrony with that of the 1D situation (Movahed et al., 2006) (see Eq.(3.21)). This reinforce our belief in the validity of the above conjecture. This, however, has to be verified by further theoretical investigation.

4.3 Summary

In this chapter, we present the results of the methodological issues. Against the mentioned two disadvantages of DFA and MF-DFA, corresponding modifications have been proposed. For the strong oscillations of scaling law, the moving window techniques and the idea of GWR are employed to develop the TWDFA and MF-TWDFA. They have been tested by the numerical experiments. Furthermore, to give a solid basis of the developed modifications, the analytical calculations have been presented too. The negative influence of the periodic and

quasi-periodic trends on the scaling behavior of DFA and MF-DFA are overcome by introducing the EMD. A EMD-based method is suggested to estimate and remove the periodic trend before DFA and MF-DFA. Its good performance, especially at the small scales, has been demonstrated by the numerical examples. With regard to the problematic relationships, the corrected relationships are suggested on the basis of the formal study and empirical analysis. The results based on them closing to the expected confirm the correctness of our corrections.

Chapter 5

Application to Real-life Geographical Examples

As shown in the discussions in chapter 4, the modified DFA and pre-detrending processing is indeed able to improve the performance of DFA and MF-DFA when handling the constructed numerical examples. In this chapter, they are applied in the real-life geographical examples to study the temperature variation and long-range correlation in sunspots series. Some discussion on and interpretation of the results are also presented. Through the applications in the geographical examples, the efficiency of the modified DFA and pre-detrending processing could be confirmed.

5.1 Performance of Multifractal Temporally-Weighted Detrended Fluctuation Analysis (MF-TWDFA): Application in Air Temperature Study

One of the most important tasks in climatology is to understand the nature and limits of climate variability (Pelletier, 1997; Tsonis et al., 1998, 1999). In general, the climate dynamic process is nonlinear (Tsonis et al., 1999).

In recent decades, especially the last ten years, nonlinear analysis of climate

data has attracted much interest in the academic community. Tsonis et al. (1999) investigated the extratropical atmospheric circulation from a unique insights of "random walk" on the basis of height anomalies. Peters et al. (2001) claimed that the rainfall event is similar to a variety of nonequilibrium relaxation processes. Yu et al. (2011) proposed a multifractal framework to study the daily rainfall series by different multifractal methods. As to the river streamflow or runfall, they have also analyzed them through different nonlinear techniques, such as multifractal analysis (Kantelhardt et al., 2006; Koscielny-Bunde et al., 2006; Pandey et al., 1998; Tessier et al., 1996) and the frequency distribution (Zhang and Singh, 2007). The technique of embedding time series to the phase space, power spectrum and the return map were employed by Tziperman et al. (1995) to examine the chaotic structure in the series from the El Niño-Southern Oscillation prediction model of Zebiak and Cane (1987).

Temperature is one of the most common and important measures to signify the climate (Ashkenazy et al., 2008). Therefore, many researches focus on the temperature records. Ashkenazy et al. (2003) showed the nonlinearity of temperature series for the time scales 1-100 kyr by the magnitudes correlation analysis of temperature increments, which was first proposed and applied to the analysis of human heartbeat series by Ashkenazy et al. (2001). The nonlinearity of temperature records was also pointed out by Bartos and Janosi (2006) and confirmed by Ashkenazy et al. (2008) through studying the asymmetry of temperature. Another hot topic is the determination of the trend in temperature record (Fatichi et al., 2009; Rybski et al., 2006, 2008), which is closely related to global warming.

Since the complicated dynamics of climate are effective on a large variety of time scales (Weber and Talkner, 2001), the study of the dynamics process on different scales is meaningful. The exact relationship between the scales involved can be studied by the scaling law (Tsonis et al., 1999). One of the most significant scaling behaviors on different scales is the scale invariance, which is a law that incorporates variability and transitions over the whole scaling range and usually is a result of nonlinear dynamics (Tsonis et al., 1998). Long-range correlation

proposed by Hurst (1951), as one property of time series, can be determined by the scaling law. This correlation refers to the auto-correlation function which decays following a power law rather than exponentially (Rybski et al., 2006, 2008). The positive or negative long-range correlation means the current pattern would be more likely to be maintained or broken in the future. The Hurst exponent (H) is an important index for measuring long-range correlation of a series (Hurst, 1951). An advantage of the MF-DFA is that it can estimate H even if the time series is affected by non-stationary trends. With regard to temperature variation, short-time persistence of weather is well known in the literature (Eichner et al., 2003; Orun and Kocak, 2009). The trend is more likely to be maintained in the short run. Besides, a weekly scale corresponding to the average duration of the so-called 'general weather regimes' or 'Grosswetterlagen' is a typical timescale for weather change (Eichner et al., 2003). Although the relevant physical process of climate systems on short-term scales are well understood, much less is known about the temperature fluctuation on larger time scales, such as larger than a month (Talkner and Weber, 2000). Then they indicated that the long-range correlations should be detected by the analysis of correlation structure as the first step to improve our understanding of climate on large scales. However, the long-range correlation of the temperature series at larger scale is more difficult to define because of the influence of different processes and trends such as the circulation patterns, global warming (Kurnaz, 2004), urban growth (Eichner et al., 2003), and the El Niño southern oscillation (one of the most pronounced phenomena, whose variation is on the scale of months to seasons (Bunde et al., 2002)). Although the difficulty to define the long-range correlation in temperature series, understanding long-range correlations in atmosphere is of fundamental interest physically and practically (Fraedrich and Blender, 2003; Pattantyús-Ábrahám et al., 2004).

Generally, the crossover points appearing in the scaling behavior of DFA and MF-DFA are expected to indicate the critical scale range of the analyzed process. Besides the crossover point at about 10 days found by Eichner et al. (2003), the long-range correlation on the scale longer than 10 days is measured by H around

0.65 according to the universal persistence law (Koscielny-Bunde et al., 1998), the power law for the continental land air temperature (Lennartz and Bunde, 2009) and the results for coastal cities, including Hong Kong (Eichner et al., 2003). As a contrast, the point of view of non-universal atmospheric persistence is held by some other researchers (Orun and Kocak, 2009; Pattantyús-Ábraháin et al., 2004). Somewhat related, DFA is also applied to study the problem of global warming (Lennartz and Bunde, 2009).

5.1.1 Results

As discussed in chapter 3, temperatures of adjacent dates are generally more similar than that at points distant in time. The detrending process should thus take this natural phenomenon into account. To verify our arguments, we employed the mean daily temperature data obtained by the Hong Kong Observatory (http://www.hko.gov.hk/hko) from January 1, 2005 to December 31, 2007. The resulting time series is of length 1095 (see Fig.5.1).

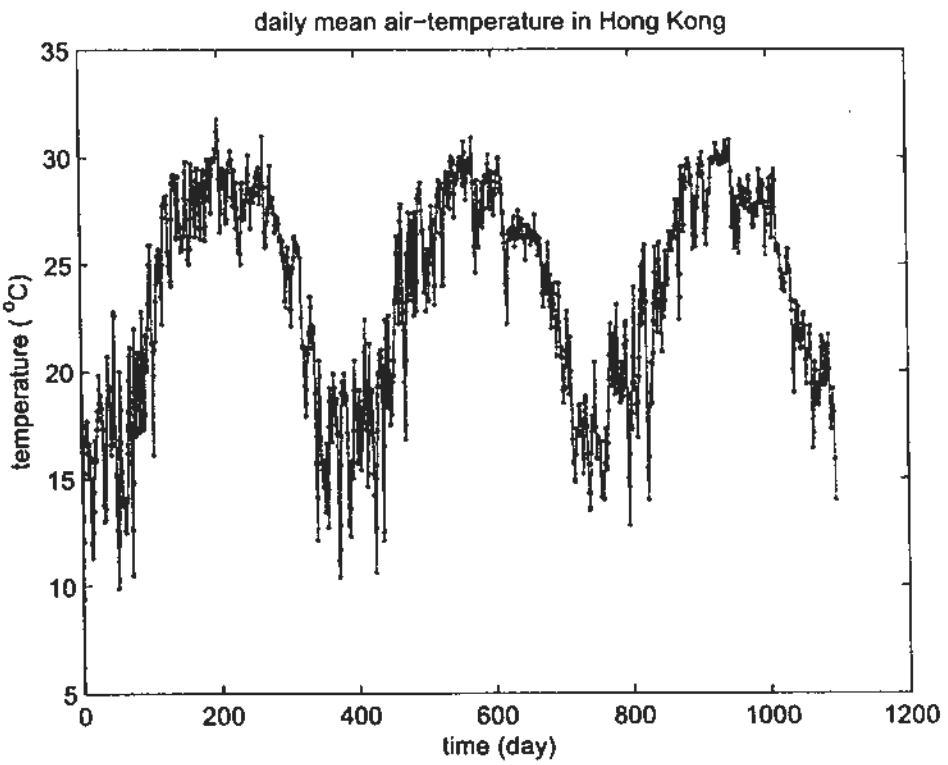

Figure 5.1: Daily mean temperature in Hong Kong from 2005 to 2007

The results obtained by the DFA and the TWDFA are shown in Fig.5.2. Plots of the results obtained by the two methods exhibit similar scaling behaviors. Generally, the scaling behavior of DFA at large scale, such as around $s = N/4$, is hard to determine because of the great fluctuation. However, TWDFA has less fluctuation, even when the scale range is increased to around $s = N/2$, where the scaling behavior seems to be maintained quite well. To illustrate this advantage of TWDFA, we make the semi-log plot of its fluctuation function up to the scale $s = N/2 \approx 548$ while only to $s = N/4 \approx 273$ for DFA. Based on the same reasons, Fig. 5.5 is plotted in the similar way. As expected, it can be observed that the TWDFA obtains a much smoother plot of $\log_{10} F(s)$ vs $\log_{10} s$, especially when s is large (see the part with s larger than 100). It is impossible to detect the crossover points by the DFA in this example. However, in the plot of the TWDFA, three crossover points can be respectively unraveled at $s = 30$; at $s = 190$; and at $s = 360$. In real life, the three crossover points correspond to the time scales of around "one month", "half a year" and the "whole year" respectively. The slopes, which mean the Hurst exponents, in different time regions are 0.9197, 1.5966, 0.7651 and 0.0175 for the respective time scales: "under thirty days", "one to six months", "six months to one year" and "over one year", indicating different properties of long-range correlation.

As pointed out by many researchers and us in chapter 3 that periodic trends will generally affect the analysis results of the DFA and MF-DFA (Hu et al., 2009, 2001; Kantelhardt et al., 2003; Movahed and Hermanis, 2008; Movahed et al., 2006). Periodic trend and seasonal trend, for example, might affect the scaling behavior of river runoff and precipitation in hydrology. In this empirical analysis, we employed the method applied to remove the seasonal trend in the hydrological series (Kantelhardt et al., 2003) and the study of the temperature series (Eichner et al., 2003; Koscielny-Bunde et al., 1998; Pattantyús-Ábraháin et al., 2004) to first eliminate the annual trend of temperature. Denoting the temperature series as $\{T_i\}$, we consider the departure $\Delta T_i = T_i - \bar{T}_i$ where $\bar{T}_i$ is calculated for each calendar year i by averaging over all years. Then the detrended temperature

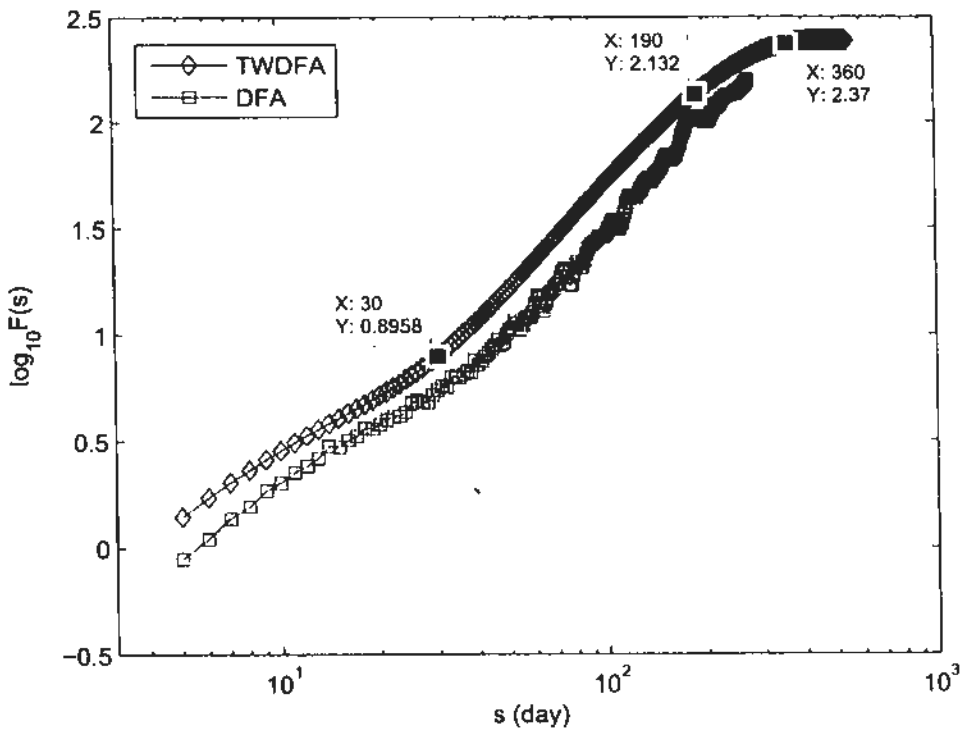

Figure 5.2: MF-DFA and MF-TWDFA results of Hong Kong daily mean temperature from 2005 to 2007

series can be obtained (see the upper panel of Fig. 5.3). The bottom panel of Fig. 5.3 illustrates the profile of detrended series with the trends estimated by DFA and TWDFA. It should be observed that obvious discontinuous points mentioned by Alvarez-Ramirez et al. (2005) as irregular jumps can be found at the ends of the fitted straight line of each segment of DFA (see step 2 of DFA for this kind of segment). However, this kind of irregular jumps can be avoided by the moving window technique of TWDFA, as shown by the solid smooth curve in the bottom panel of Fig. 5.3. Since the TWDFA trend is an obvious signal with relatively long time period, TWDFA ought to be a high-pass filter as DFA and DMA. This conclusion is supported by the power spectrum of the signal before and after removing the TWDFA trend (Fig. 5.4). It is confirmed that the estimated trend is dominated by low frequencies. In addition, for relatively high frequencies, the power density of residuals is almost the same as those of the detrended series. Because of the strong fluctuation of the power spectra, the crossover points are difficult to be detected. Then only two scaling exponents,

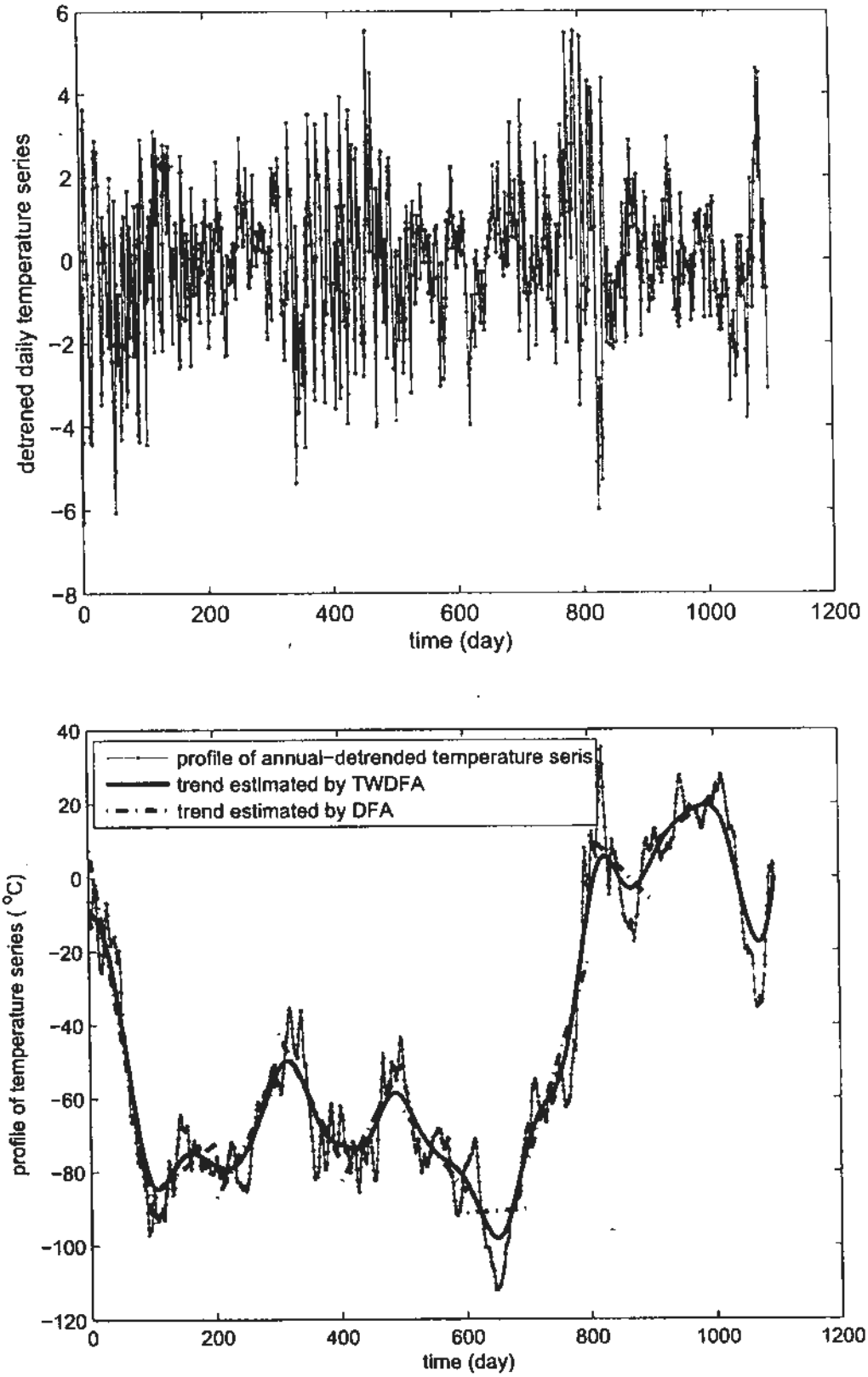

Figure 5.3: Detrended daily temperature series in Hong Kong from 2005 to 2007 (upper panel) and its profile with the local trends estimated by DFA and TWDFA (bottom panel)

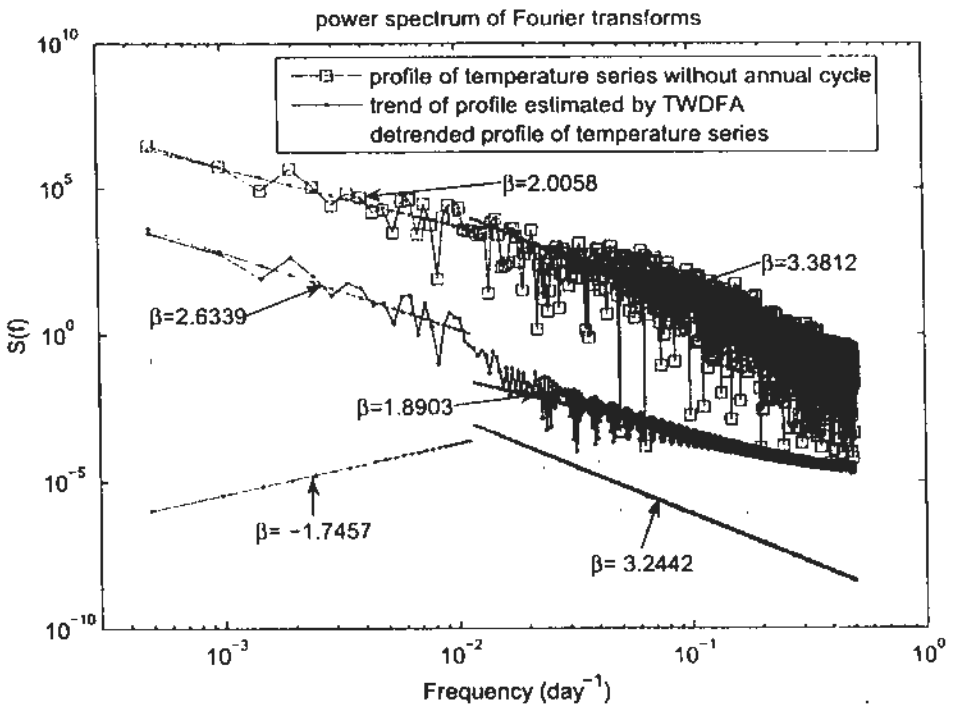

Figure 5.4: Power spectrum of the profile of detrended daily temperature series, the trend estimated by TWDFA and the detrended profile. For better illustration, the power spectrum of the profile of daily temperature series and the detrended profile have been shifted vertically upwards by 3 (from setting $\log_{10} S(f)$ as $\log_{10}(10^3 \cdot S(f))$ which is equal to $\log_{10} S(f) + 3$) and similarly downwards by 4 at $\log_{10}$ scale respectively.

β, of each power spectrum for relatively low and high frequencies are estimated by $S(f) \sim f^{-\beta}$ ($S(f)$ is the power spectrum density) and accordingly denoted in Fig. 5.4. It can be observed that the β corresponding to the relatively low frequencies are relatively similar for the profile of detrended daily temperature and the TWDFA-estimated trend (2.0058 vs. 2.6339). While the β's of the profile and detrended profile series estimated by the slope of relatively high frequencies are much closer (3.3812 vs. 3.2442). The experimental results support our above conjecture that the TWDFA method should work as a high-pass filter. However, since the great fluctuation of the power spectra can usually be found in real-life applications such as the temperature example in this thesis, the estimated slope should be very uncertain. Then DFA and TWDFA are recommended for the calculation of the scaling exponent.

We then performed the TWDFA and the DFA on this detrended series, and the result is shown in Fig. 5.5. It can be observed that the results of DFA

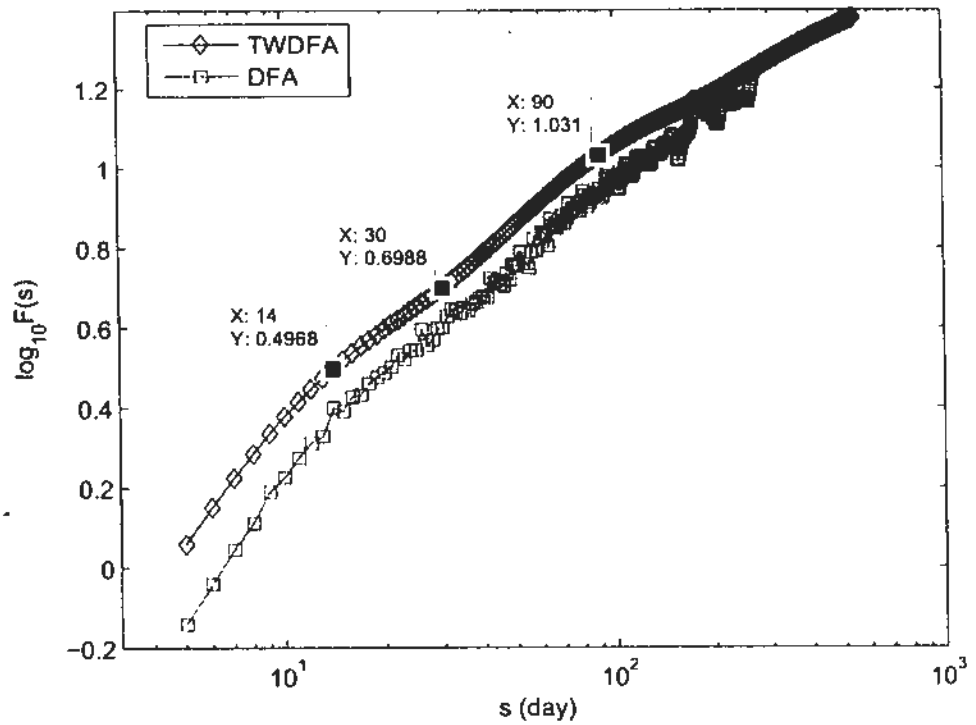

Figure 5.5: DFA and TWDFA results of the detrended daily temperature series in Hong Kong from 2005 to 2007

and TWDFA exhibit similar scaling behaviors. However, besides the common crossover point at the location around $s = 14$, TWDFA manages to find another two crossover points locating at $s = 30$ and $s = 90$, which can hardly be deciphered by the DFA that exhibits much stronger fluctuation. The two additional crossover points correspond to one month and one season in real time. However, the half-year and one-year crossover points detected without removing the annual trend disappeared when such trend is removed. The slopes in the four time-scale regions divided by these three crossover points are 0.9752, 0.5698, 0.7113 and 0.4469 respectively. In addition, the TWDFA result again seems to be smoother than that of DFA. In order to quantitatively confirm that the TWDFA method indeed obtain a better linear relationship for estimating the power-law of scaling behavior, we employ the root mean squared fluctuation (RMSF) around the regression line to measure the goodness of linear fit. Obviously, the perfect fitting corresponds to RMSF=0. The RMSFs in the scaling ranges divided by the crossover points detected by TWDFA for DFA and TWDFA are listed in Ta-

ble 5.1. The symbol "−" in Table 5.1 indicates that RMSF at this range cannot be

Table 5.1: Root mean squared fluctuation around the regression line depicting the scaling behavior by DFA and TWDFA on the original and annual-cycle-removed temperature series

	original temperature series				annual-cycle-removed temperature series			
	<30 days	30~190 days	190~360 days	>360 days	<14 days	14~30 days	30~90 days	>90 days
TWDFA	0.01997	0.01958	0.02281	0.00542	0.02171	0.00207	0.00644	0.00601
DFA	0.05074	0.08597	0.03273	-	0.02698	0.02800	0.03580	0.05230

obtained because of the maximum studied scales $N/4 = 250 < 360$. It is obvious that all RMSFs of TWDFA are much less than that of DFA. Therefore, TWDFA is able to obtain much better linear relationship for estimating the correlation properties of time series and can reduce the uncertainty due to great fluctuation around the regression line. The results are generally reasonable. If temperature series is observed at the scale less than two weeks, the slope (0.9752) indicates that the variation of temperature more likely maintains the same local trend. At a slightly larger scale, the variation of temperature appears to be maintaining the same pattern, but with the possibility of having random sudden changes. This is an interesting discovery, since the Hurst exponent in this scale region, 0.5698, is even smaller than that at the region from one month to three months (i.e. one season), 0.7113. This discovery indicates that the variation of temperature is more stable at the one-month-to-one-season scale region than that at the two-to-four-week region. On the other hand, the Hurst exponent at the scale longer than one season is 0.4469, suggesting a anti-correlation in this time region. This means that the variation of temperature is most distinguishable at scale longer than a season when the annual trend has been removed. From the numerical examples, the fluctuation function of MF-DFA and MF-TWDFA generally exhibit similar scaling behavior. Thus periodic trends should have similar effect on MF-TWDFA method in general. This effect can be unraveled by comparing the locations of the crossover points and the corresponding slopes in Fig. 5.2 and Fig. 5.5. The crossover points located at 190 days and 360 days of the original series disappear in the annual-cycle-removed results. These two scales just correspond to the half and one cycle of the annual periodic trend. The 14-days

crossover point can be newly detected in the annual-detrended series. In order to further discuss the effect of the annual trend on the scaling behavior of fluctuation function of TWDFA, the power spectrum (see Fig. 5.6) of the original series and the annual-cycle-removed series are employed. The frequency of the annual

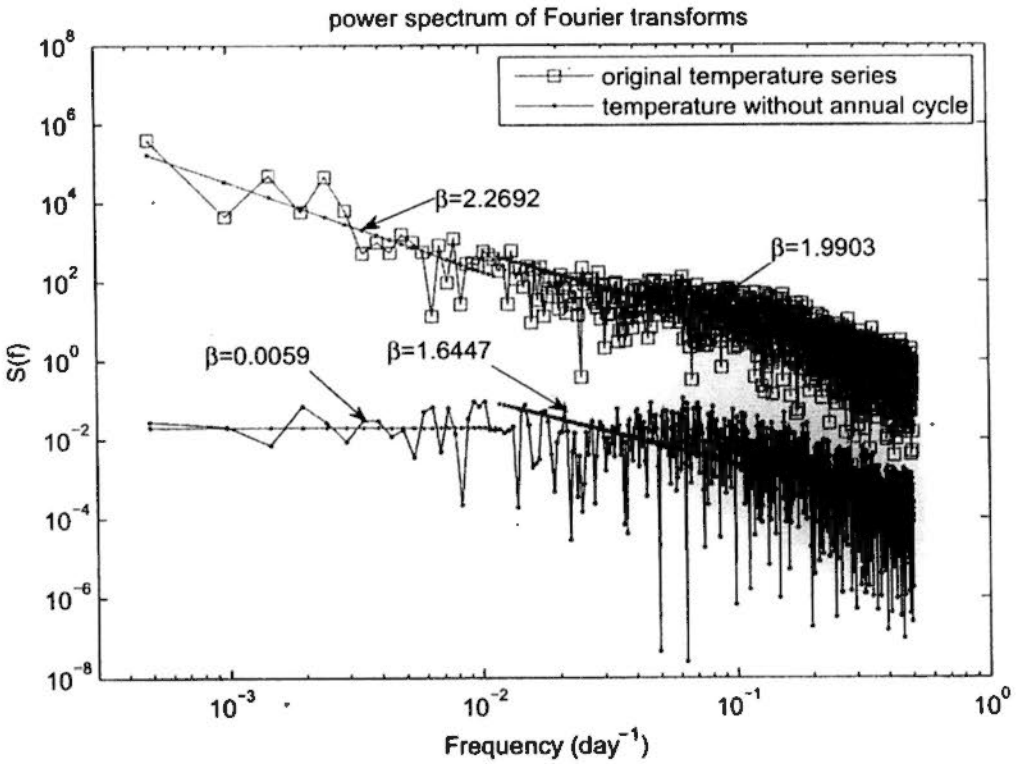

Figure 5.6: The power spectrum of the original series and the annual-cycle-removed series. The power spectrum of the original series have been shifted vertically upwards by 3 at $\log_{10}$ scale for convenience of comparison and observation.

cycle corresponds to $1/365 \approx 0.0027(\text{day}^{-1})$, which indicates approximately the 5th and 6th terms in the power spectrum. In Fig. 5.6, it is clearly seen that the power spectrum density of the first several, especially the first 10, terms marked by squares are much larger than the dotted ones. However, for higher frequencies, the two spectra have similar values. The difference of the β's for the lower and higher frequency ranges of these two series indicates that the scaling behavior at higher range are more similar to each other (1.9003 vs. 1.6447) compared with their obvious different behavior at the lower range (2.2692 vs. 0.0059). Based on the above observations, it is safe to draw the conclusion that the change of the location of the crossover points at relatively large time scale are mainly due to the annual-cycle-removing procedure. Because of the complicated properties

of real-life temperature series, it cannot be ensured that effect of periodic trend leads to the undetectability of the newly-appeared 14-day crossover points in the scaling behavior of the original series.

Furthermore, a clear linear relationship can be maintained even when s is nearly $\text{int}(N/2) = \text{int}(1095/2) = 547$. The experiment thus shows that by the proposed method, we can obtain more information relating to the scaling behavior of the temperature time series. And, it supports our theoretical arguments.

As to the multifractality of the temperature series, MF-TWDFA is preformed on the annual-cycle-removed series. Similar to MF-DFA, if $h(q)$ obtained by MF-TWDFA depends obviously on q, then it indicates that the series is multifractal. Fig. 5.7 illustrates that the $h(q)$ varies with q, and its significant dependence on q suggests that the multifractality exists in the temperature series. The bottom panel of Fig. 5.7 also supports the multifractality suggested by the $h(q)$-q plot.

5.1.2 Interpretations and Discussions

The climatological interpretations of the crossover points are that the weekly scale should correspond to the average duration of the so-called "general weather regimes" or "Grosswetterlagen" which is the typical time scale for weather changes, e.g. about 1 week (Eichner et al., 2003). The two-week crossover point can be found by both the MF-DFA and the MF-TWDFA. In fact, Eichner et al. (2003) found the 10-day crossover point using DFA0. They also indicated that the order of the regression polynomial used for detrending could affect the location of the crossover points. This location will increase, i.e. larger crossover time scales, as the order m increases. Given that we employed the linear fitting ($m = 1 > 0$) in the detrending step in MF-DFA and MF-TWDFA, our crossover point appears to coincide with the finding of Eichner et al. (2003), which means that both the two-week and 10-day crossover points actually correspond to a typical Grosswetterlagen. At the time scale of less than two-weeks, the short-term persistence dominates so that the Hurst exponent is almost 1 (exactly, 0.9752) which coincides with the statement made in Eichner et al. (2003); Orun and Kocak (2009)

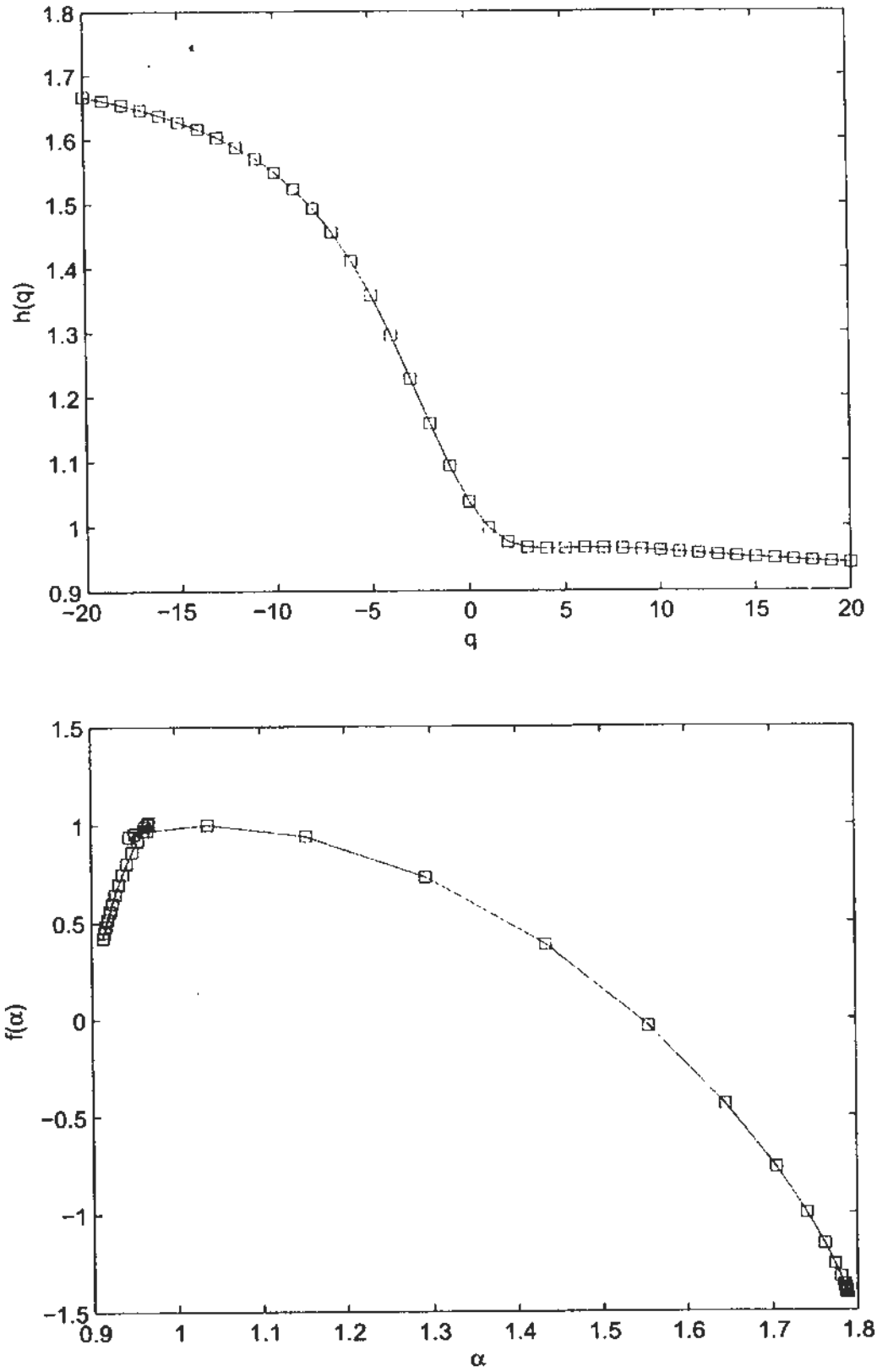

Figure 5.7: $h(q)$ of the annual-cycle-removed series against q (upper panel) and the singularity spectrum of this series (bottom panel)

about the 'minimum skill' forecast. With regard to the one-month and one-season crossover points which are only found by the MF-TWDFA, but failed to be unraveled by the (MF-)DFA studies of such temperature series (Eichner et al., 2003; Fatichi et al., 2009; Fraedrich and Blender, 2003; Koscielny-Bunde et al., 1998; Lennartz and Bunde, 2009; Orun and Kocak, 2009; Pattantyús-Ábrahám et al., 2004; Talkner and Weber, 2000), since the weather changes after the Grosswetterlagen, the air-temperature, an important indicator of weather change, should most likely change compared to that in the one-day-to-two-week scales. Therefore it is reasonable to find a weaker long-range correlation in this regime ($H = 0.5698$ vs 0.9752). However, the variation of temperature should maintain the same trend within the same season. Thus, it is natural that we can find the the exponent H increases to 0.7113 in the one-month-to-one-season scale region indicating persistence in temperature variation.

Comparing with the universal persistence law found by Koscielny-Bunde et al. (1998) and the study of Eichner et al. (2003) which both suggested the exponent should be around 0.65 at the scale larger than 10 days, we can find 0.7113 is slightly larger than 0.65. It should be noted that on the scale of months to seasons, one of the most pronounced phenomena is the El Niño southern oscillation event which strongly affects the weather over the tropical Pacific (Bunde et al., 2002; Eichner et al., 2003). Ding et al. (2002) pointed out that temperature anomaly occurred in Hong Kong during the El Niño years. Besides, El Niño could also lead to drier conditions in parts of Southeast Asia and Northern Australia which might also affect the temperature in these area. As 2006 and 2007 are the El Niño years, our discovery from the series suggests that the long-range correlation of temperature is slightly stronger than the average situation on the scale of months to seasons, corresponding to the typical scale of El Niño (Bunde et al., 2002; Eichner et al., 2003). It is plausible that the warming process because of the El Niño events might keep the temperature on a stable trend. On the even larger time scale, the plots of the MF-DFA become too fluctuated to estimate its Hurst exponent which however can still be estimated by our proposed MF-

TWDFA method. In larger time scales, temperature is influenced by different processes and trends, like the circulation patterns and global warming, the long-term correlation becomes more difficult to define (Kurnaz, 2004). In the time scale covering seasons, we find the scaling behavior of temperature changes to slightly anti-correlated ($0.4469 < 0.5$) but is not much different from 0.5 that indicates a random process. It means that the variations of temperature is most distinguishable in this regime which corresponds to the variation of temperatures among seasons. On the other hand, results obtained by (MF-)DFA suggest that the scaling behavior of the temperature series on the scale region larger than one season still follows the previous trend even after the annual trend has been removed (Eichner et al., 2003; Koscielny-Bunde et al., 1998; Orun and Kocak, 2009; Pattantyús-Ábrahám et al., 2004). This clearly contradicts our conventional understanding of seasonal variation of temperature. The difference might be due to different ways of detrending. However, our conclusion appears to be more reasonable and comprehensible. Actually, the importance of the monthly and seasonal crossover points has already been pointed out by Ding et al. (2002) using several different statistical methods supporting their significance.

In terms of prediction, the value of the Hurst exponent, almost 1, indicates that the future temperature in weekly scales is predicable and reliable. Actually, the weather forecast is with great accuracy in practical implementation. However, in the weeks-to-month scales and the scales larger than one season, it is hard to perform the prediction due to the randomness indicated by the Hurst exponent, close to 0.5. The prediction of temperature in the one-month-to-one-season scales is also possible though it is more difficult compared with that in the weekly scales. Thus, the study of the long-range correlation is very helpful in recognizing the underling dynamics of process and is beneficial to evaluating the possibility in establishing a accurate prediction model.

5.2 Performance of EMD-Based Pre-Detrending Processing: Application in Sunspot Series

Sunspot is the relatively lower temperature part of the solar surface that looks like black regions (Liou, 2002). The number of sunspots varies greatly during different periods. It might stay relatively high over some periods but drop to almost zero a few years later. Since the famous 11-year cycle appearance of sunspots was discovered by Schwabe in 1841, it has been confirmed by many researchers through long-term observations and studies (Gnevyshev, 1977; Moussas et al., 2005; Schatten and Sofia, 1987). Besides, some other solar cycles have been found in succession, such as the cycle of solar wind, quasi-biannual oscillation (QBO), Hale period and possible Gleissberg period (Li et al., 2007; Xu et al., 2008). Many geophysical phenomena and atmospheric processes are significantly influenced by the sunspot activity (Liou, 2002; Moussas et al., 2005). Thus it is meaningful to study the property of the sunspot time series.

As mentioned in chapter 3, some debates on the long-range correlation of the sunspot series exist. Different pre-detrending processing techniques for handling the periodic trend, such as the dominate 11-year cycle, lead to the different long-range correlations property (Hu et al., 2009; Movahed et al., 2006). We proposed the EMD-based method to handle this kind of trend., Its validation and the comparison to the Fourier truncation and adaptive detredning methods have done by numerical experiments. In the numerical experiments in chapter 4, for the small scales the three pre-detrending methods appear to be effective in detrending the effect of periodic trends if the parameters are appropriately selected. For larger scaling range, however, some difficulties exist. Generally, there are definitive physical meaning in the IMFs. Thus it seems to be more intrinsic the crossover points unraveled by the EMD-based method. In contrast, Fourier truncation and the adaptive detrending methods work in a artificial way, especially for the selection of parameters, such as frequencies to be removed and the length of segment for detrending. Besides, the EMD-based methods parameters

selection procedure is much easier. With regard to the situation on scales larger than the crossover points, the adaptive detrending method almost wipe out all information, while the EMD-based and the Fourier methods manage to keep considerable information even though the results are a bit off the expected pattern. To compare these pre-detrending processing in the real-life example to get the bottom of the discrepancy between Movahed et al. (2006) and Hu et al. (2009), all of them are also considered in the following discussion.

5.2.1 Basic Results

We analyzed the monthly number of sunspots during the period 1749~2009 with 3123 months from SIDC's website (http://sidc.oma.be/sunspot-data/) used in Movahed et al. (2006) (to the year of 2006) and Hu et al. (2009). The downloaded data is our original time series, henceforth referred to as "original data" or "original series". From Fig. 5.8, the 11-year cycle is very obvious. Applying the

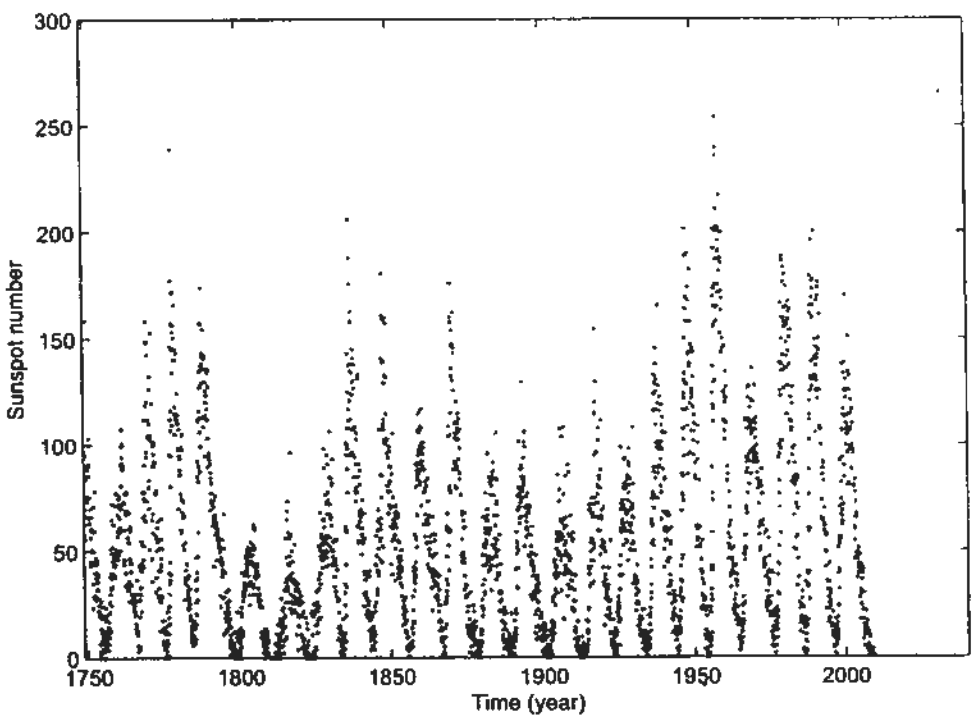

Figure 5.8: The monthly sunspot time series.

EMD method to the sunspot time series, eight IMFs and one r can be obtained

(Fig. 5.9). As discussed in the introduction, different IMFs correspond to differ-

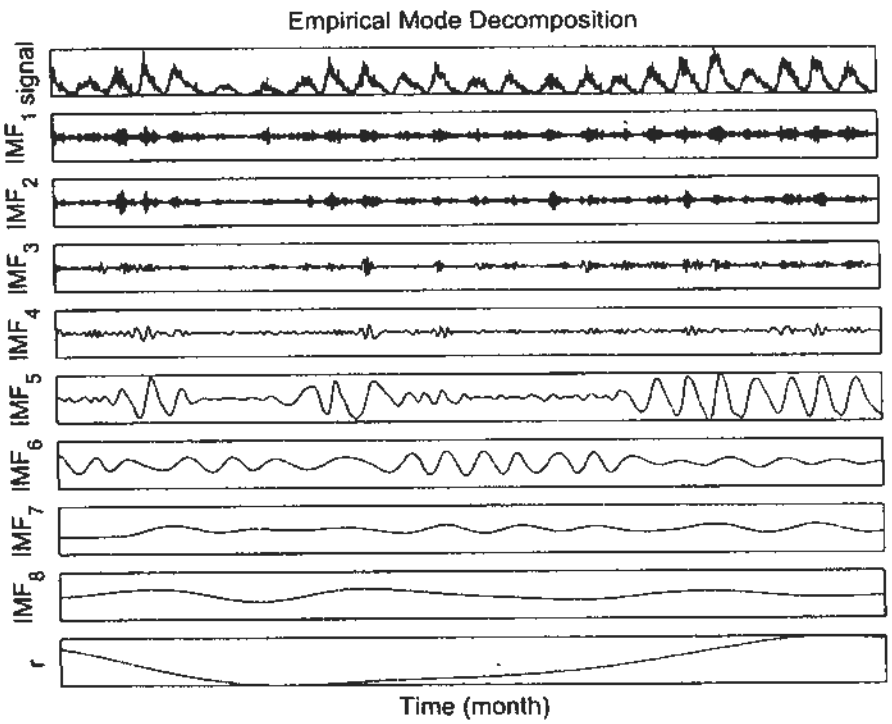

Figure 5.9: Results of the EMD on the original sunspot series, IMF_1 to r (from top to bottom).

ent well-known cycles of sunspot activities. Therefore, we apply DFA to analyze the series by reducing the IMFs or r one by one so that we can study the influence of each of them on the scaling behavior of the sunspot time series. As depicted in Fig. 5.10, there are roughly two groups of scaling behavior. We denote the series, including the original data, with relatively steep slopes (slope$\geq$1) group 1, and the others group 2. The scaling properties of the series within each group are similar. We discover that IMF_5 is the critical component that decides the scaling property of the sunspot time series. When IMF_5 is removed, the series becomes group 2 right away and exhibits totally different scaling behavior from the original data. For the series, $\sum_{i=1}^{4} \mathrm{IMF}_i$, a crossover point at $s = 10^{1.76} = 57.5440 \approx 60$ can be identified. The slopes to the left and right of this point are 0.69 and 0.28 respectively. The position of the crossover point and the left slope are very similar to the result obtained by Hu et al. (2009). While the series $\sum_{i=1}^{5} \mathrm{IMF}_i^*$ has property similar to the original series with Hurst exponent larger than 1 (see

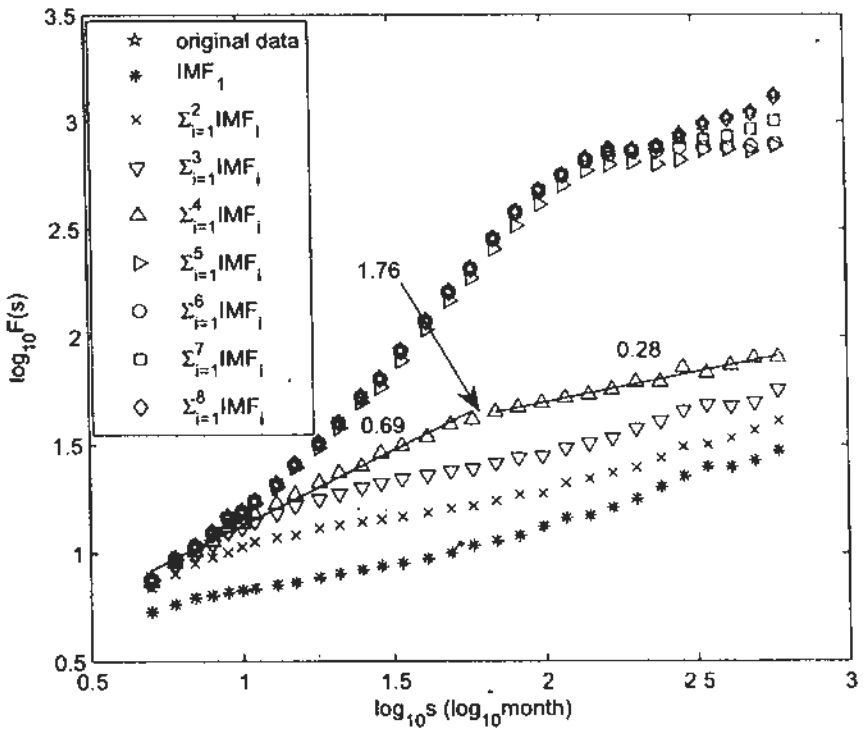

Figure 5.10: Scaling behavior of sunspot time series by removing the components one by one.

Movahed et al. (2006) for more discussion on the results obtained from the original series). As observed in from Fig. 5.11 (upper panel), $\sum_{i=5}^{8} \mathrm{IMF}_i + r$ fits the original data, x, very well. Besides, the 11-year cycle should be also eliminated in $x - (\sum_{i=5}^{8} \mathrm{IMF}_i + r) = \sum_{i=1}^{4} \mathrm{IMF}_i$. To confirm it, we calculate the power spectrum of the detrended series and compare it with the original series. From Fig. 5.12, we can see that the powers of the detrended series, $\sum_{i=1}^{4} \mathrm{IMF}_i$, for the frequencies around the 11-year-cycle frequency, e.g. $\frac{1}{11\times 12} = 0.0076$ month^{-1}, are significantly less than that of the original series. Such difference quantitatively indicates that the 11-year cycle has been removed. As a whole, our results, up to here, are very similar to that obtained by Hu et al. (2009). This is a validation of their results from the angle of EMD.

In addition, we remove each component from the original time series once, and study the respective scaling behavior afterwards. From Fig. 5.13, we can also discover the remarkable effect of IMF_5. Though not as strong as the effect of IMF_5, IMF_6 also affects the scaling behavior of the series. The others have

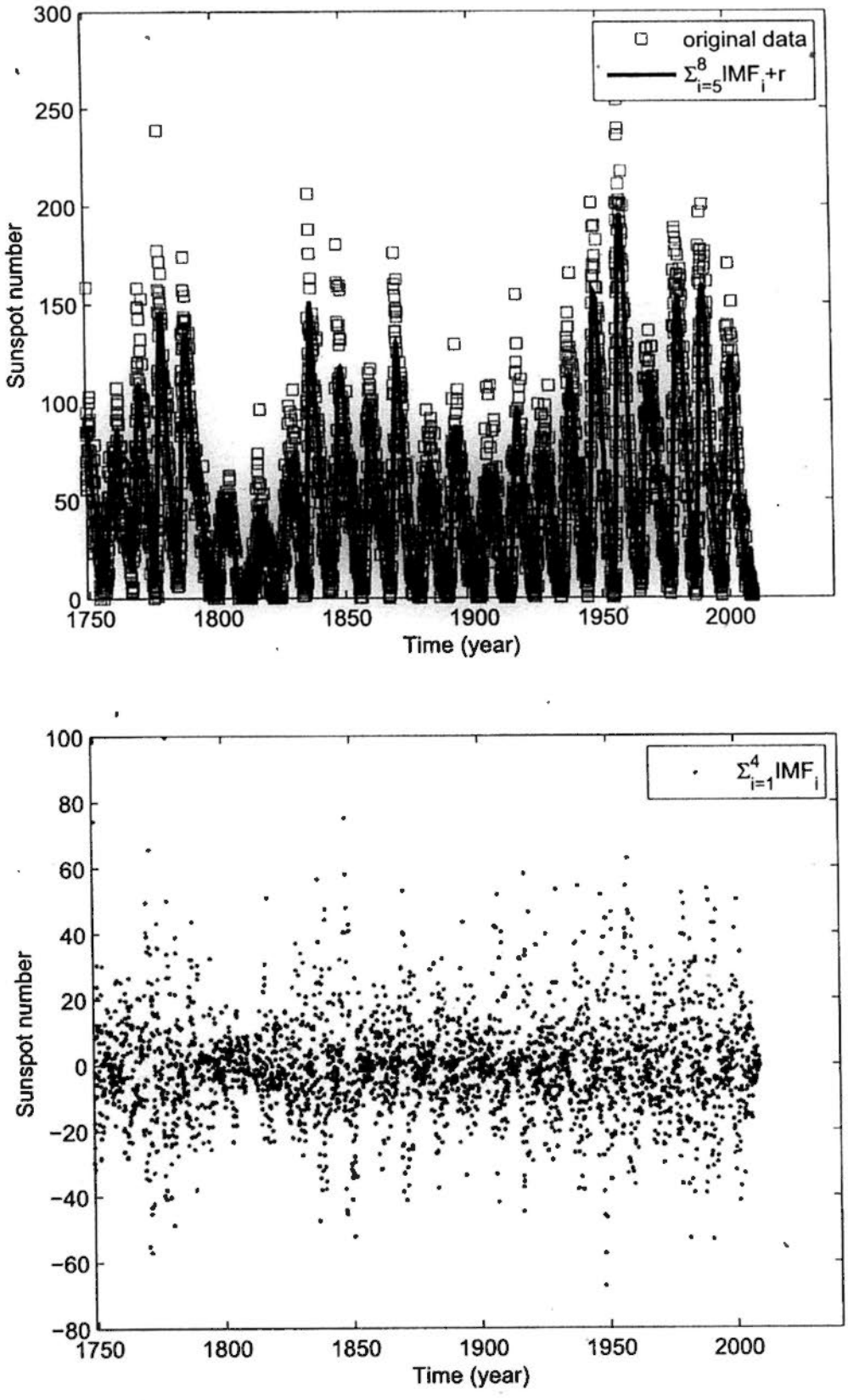

Figure 5.11: Original data and the fitted series using $\sum_{i=5}^{8} \mathrm{IMF}_i + r$ (upper panel) and the residual $x - (\sum_{i=5}^{8} \mathrm{IMF}_i + r) = \sum_{i=1}^{4} \mathrm{IMF}_i$ (bottom panel)

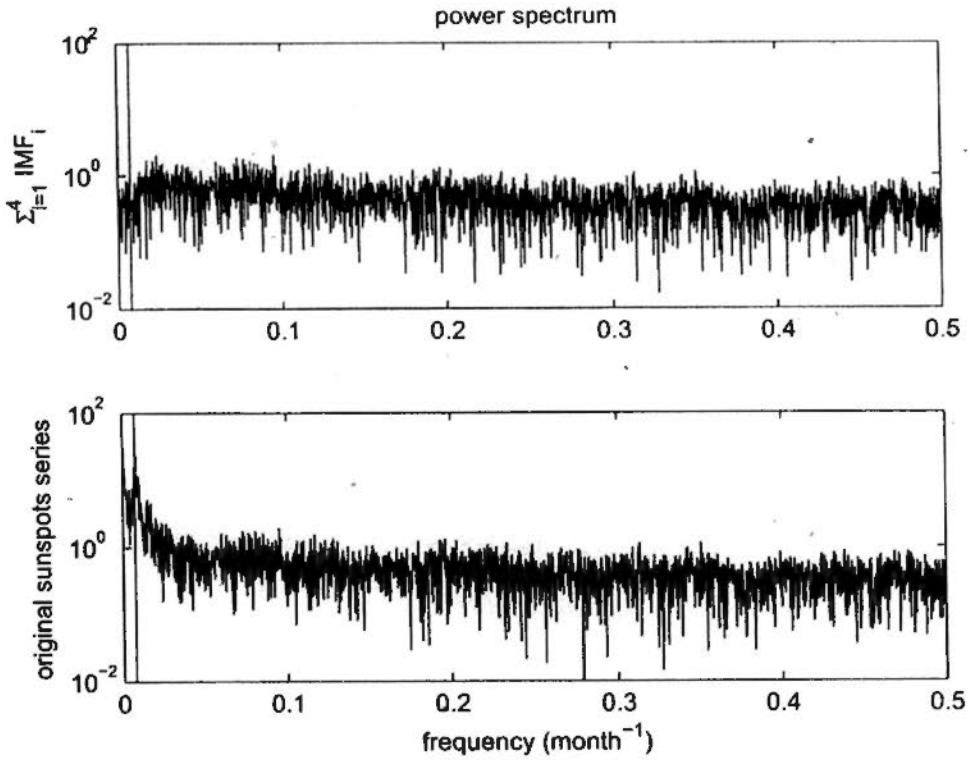

Figure 5.12: Compare the power spectra of the detrended series ($\sum_{i=1}^{4}$ IMF$_i$) and the original series. The vertical straight line corresponds to the frequency of the 11-year cycle, 0.0076.

almost indistinguishable influences on the scaling behavior. Two crossover points, $10^{1.46} \approx 29$ and $10^{2.07} \approx 117$, can be identified, as shown in Fig. 5.13.

However, by removing these two most remarkable components, i.e. removing IMF$_5$ and IMF$_6$ from the sunspot time series, the original time series (Fig. 5.8) is transformed into the time series in Fig. 5.14. Here the 11-year cycle becomes less obvious, although some other irregular cycles can still be found.

To compare with the effect of the 11-year cycle on the fractal property of sunspot time series, we employ DFA to study the original and this detrended series, the results are depicted in Fig. 5.15. As the DFA results of the original series have already been discussed in Movahed et al. (2006) and Hu et al. (2009), we now focus our analysis on the detrended series. A crossover point of the time series can be found in Fig. 5.15. Its position is $10^{1.76} \approx 58$, which is very close to 60 obtained by Hu et al. (2009). Two different slopes are discovered on the two sides of the crossover point. The left slope is 0.72 which is similar to the result, 0.74, of Hu et al. (2009). However, the right slope, 1.49 is much larger than that

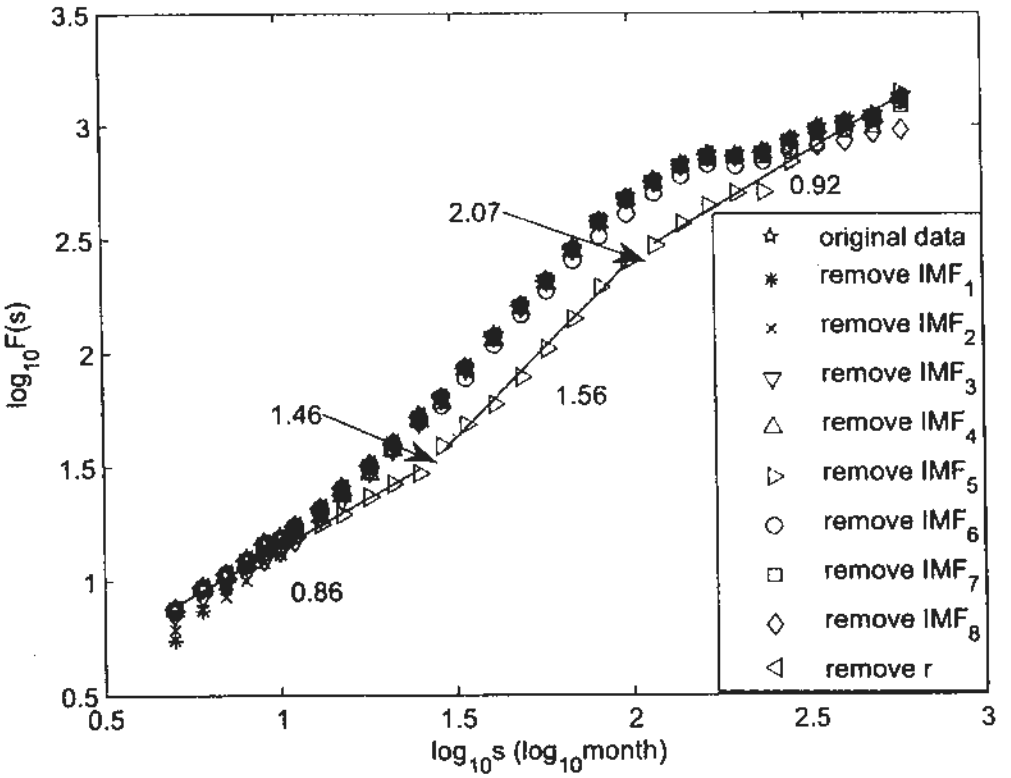

Figure 5.13: Effect of each component on the scaling behavior of the sunspot time series.

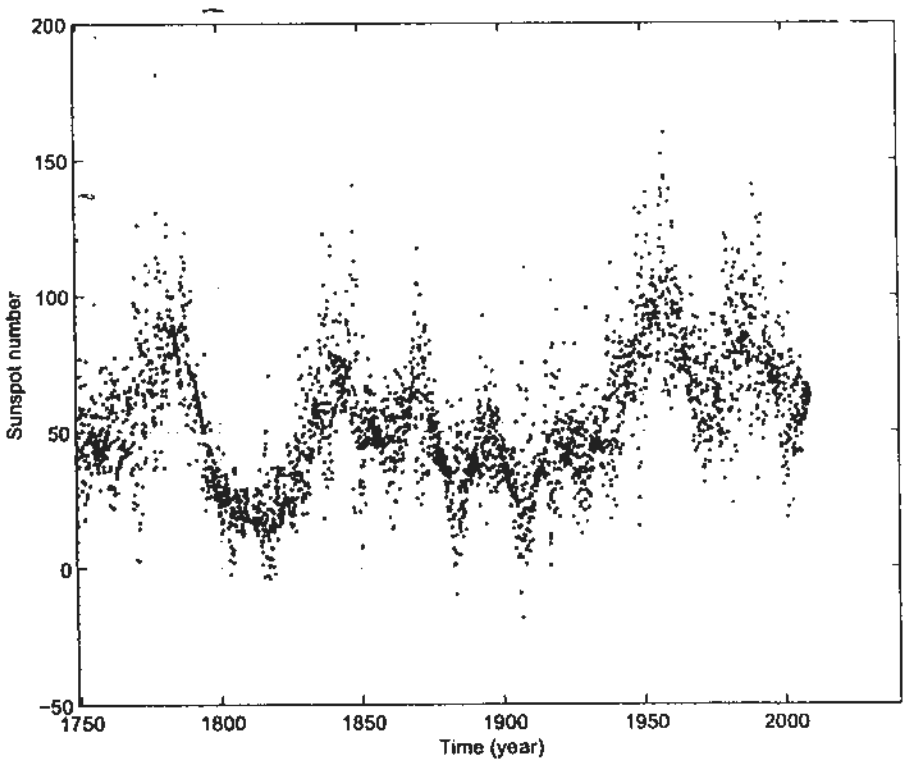

Figure 5.14: The series after eliminating IMF_5 and IMF_6 from the original sunspot series.

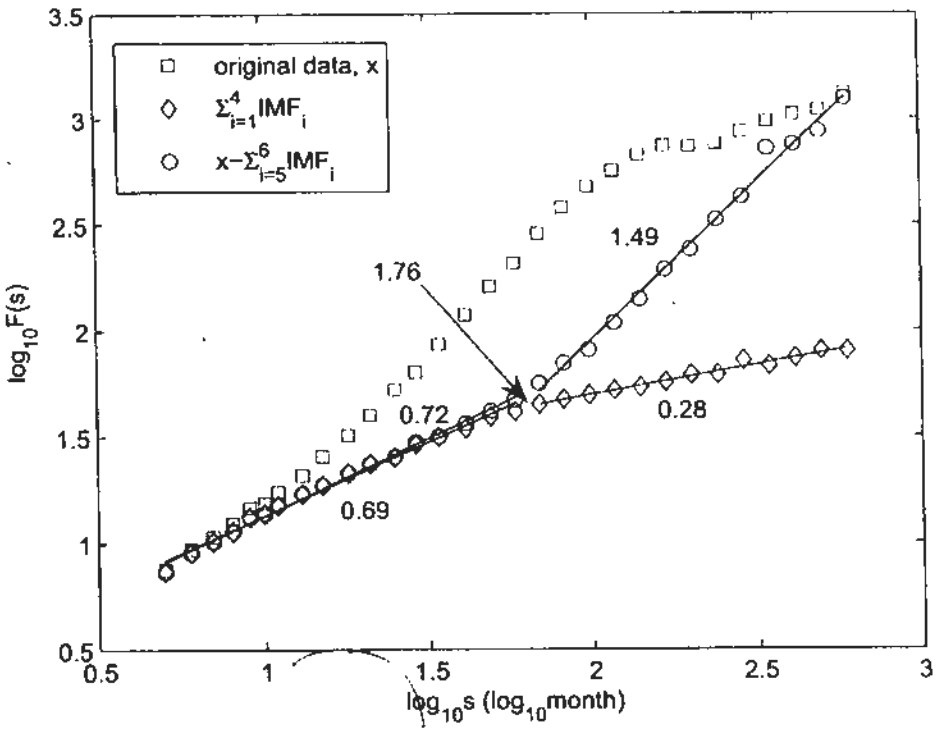

Figure 5.15: The scaling behaviors of the original and the detrended series using DFA.

obtained by Hu et al. (2009). It can also be observed that the scaling behaviors of $\sum_{i=1}^{4}$ IMF$_i$ and $x-\sum_{i=5}^{6}$ IMF$_i$ at the time scale $s < 60$ are very close, 0.69 vs 0.72, to each other. However, if s increases beyond this crossover point, two different behaviors are observed, 0.28 vs 1.49. Generally, the maximum scale on which the scaling behavior can be studied is $N/4$ (N is the length of series) since the fluctuation usually becomes too strong to obtain the reliable scaling behavior on even the larger scales because of the insufficient local windows (Kantelhardt et al., 2002). It is interesting to find in this real-life example that the fluctuations of $x-\sum_{i=5}^{6}$ IMF$_i$ and the original data, x, approach the same value at the maximum time scale, $s = N/4 = 3123/4 \approx 781$ months. Then the information of the sunspots series on the larger scales seems to be maintained considerably. Take the difference between $\sum_{i=1}^{4}$ IMF$_i$ and $x - \sum_{i=5}^{6}$ IMF$_i$ into consideration, it is reasonable to draw the conclusion that the components with dominant frequencies lower than those of IMF$_5$ and IMF$_6$ determine the fluctuation of the series at scales larger than 60 months.

5.2.2 Further Discussion

To affirm our results above, we study the problem from the Fourier-based angle in this subsection. We employed the Fourier transforms to investigate the eight IMFs. Since the concerned 11-year cycle corresponds to the frequency $1/132 \approx 0.0076$ per month and the interval of the two successive discrete frequencies is $1/3123 \approx 0.0003$, consider what we would like to find is the rough correspondence, we think it ought to be fine to transform these components to the coefficients with equal length to extract the relevant frequencies for our further analysis. The respective power spectra of them are depicted in Fig. 5.16 and Fig. 5.17.

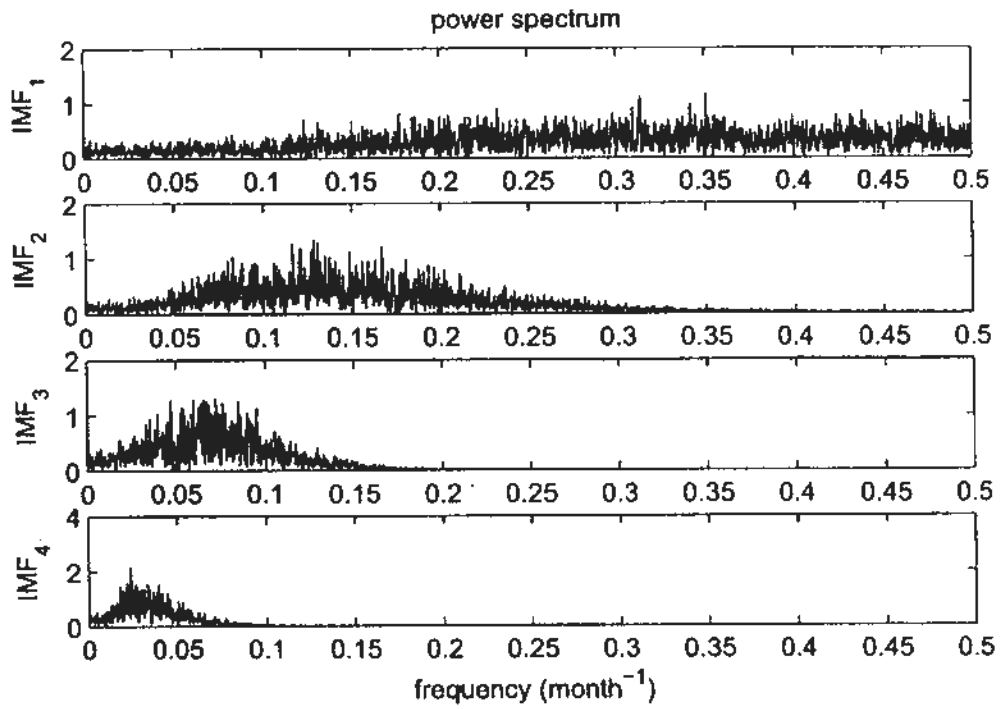

Figure 5.16: Power spectra of the IMFs of the original sunspot series using Fourier transform, IMF_1 to IMF_4 (from top to bottom).

It should be noted that the abscissa of Fig. 5.16 and Fig. 5.17 is the frequency expressed in number per month. Generally, the components, IMFs, decomposed by EMD have their physical meaning. In the sunspot series study, Li et al. (2007) systematical studied some relations between these IMFs and some solar active cycles. On the basis of the power spectra, these IMFs obtained in this thesis are attempted to correspond to certain solar cycles. IMFs are analyzed one by one as follows: For IMF_1, it is hard to observe any kind of cycle because

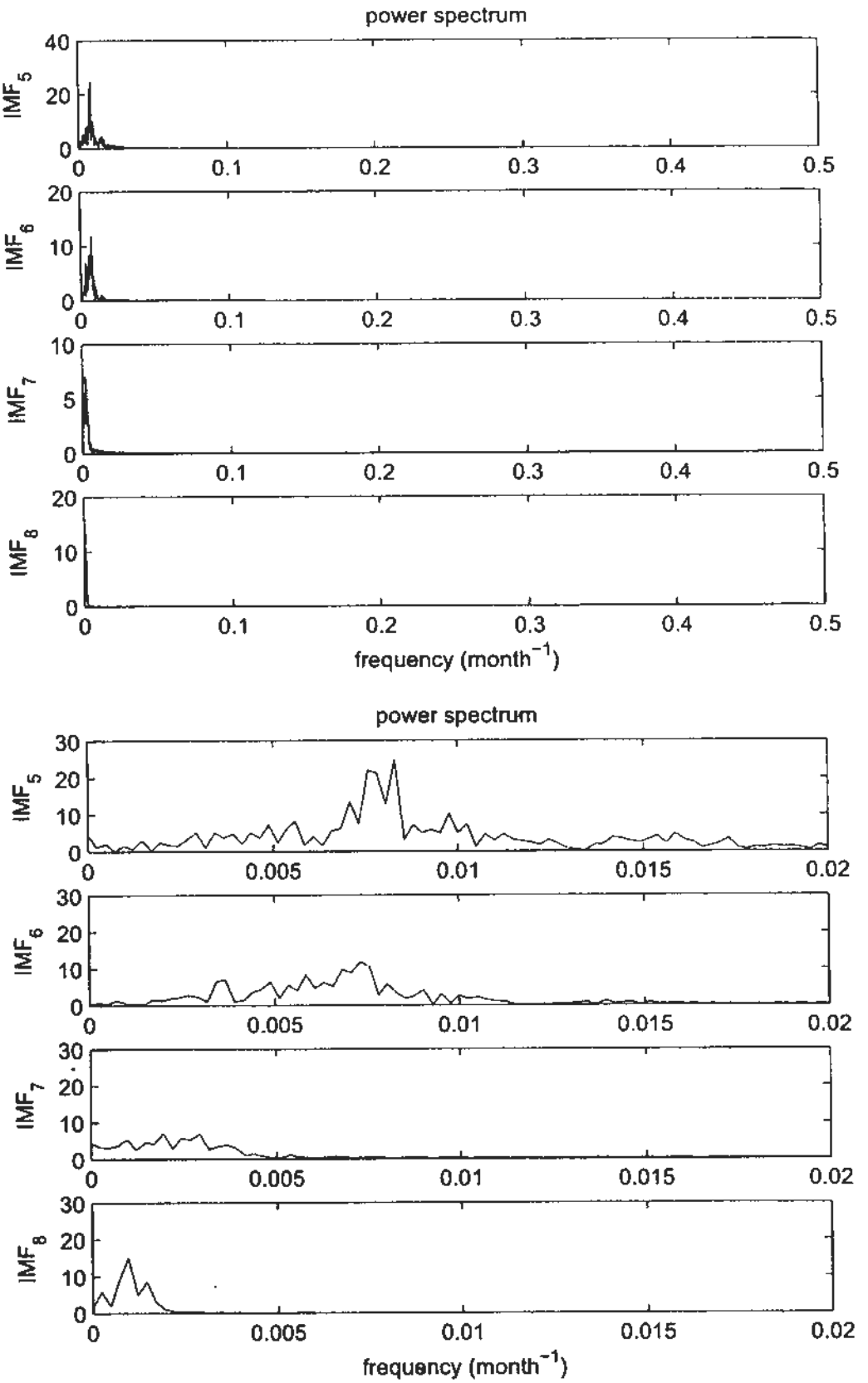

Figure 5.17: Power spectra of the IMFs of the original sunspot series using Fourier transform, IMF_5 to IMF_8 (from top to bottom), the bottom panels are the locally enlarged plots of the corresponding upper panels.

of its uniform distributed power spectrum density. Since the dominant frequencies distribute around 5−10 months, i.e. 150−300 days, IMF_2 should correspond to the around 153-day cycle, which is found by Rieger et al. (1984) in γ-ray flares and Cane et al. (1998) from interplanetary magnetic fields. With regard to IMF_3 and IMF_4, the dominant frequencies corresponding to 10−20 months and 25−30 months suggest that the 1.3−1.4 years periods and QBO might be included in these two IMFs. The 11-year cycle should be denoted as $\frac{1}{11\times12} = 0.0076$ per month. Obviously, the dominant frequency of IMF_5 and IMF_6 is around 0.0076. Thus they should be both related to the 11-year cycle of sunspots. In addition, some frequencies with relative large amplitude can be found between 0.003 and 0.004, i.e. the periods about 21−27 years, in power spectrum of IMF_6. That means IMF_6 might also include the Hale period with 22-years cycle. The relative dominant frequencies of IMF_7 around 0.002, i.e. 42 months, ought to include the double Hale period (Fairbridge and Hillaire-Marcel, 1977). These discovers are similar to results of Li et al. (2007). However, the length of series considered in this study is more than 260 years, which is much longer than 110 and 55 years researched by Li et al. (2007). Then much longer cycles might be found in our study. Actually, the power spectrum of IMF_8 has a peak at around 0.001. Then it should indicate the Gleissberg period with nearly 80 years cycle (Gleissberg, 1944), those frequencies with less amplitudes around it should be related to some other kinds of solar centenary period varying in the range from 65 to 130 years (Nagovitsyn, 1997). Therefore, IMF_2 to IMF_8 should correspond to the 153-day cycle of γ-ray flares and interplanetary magnetic fields, the 1.3−1.4 years periods, QBO, 11-years cycle, Hale, double Hale and Gleissberg period respectively. And the effect of those periods on the scaling behavior of sunspots series can be studied through the summation of selected IMFs.

According to many studies, the effect of the 11-year cycle has to be eliminated to avoid the production of spurious crossover points. Coincide with our above analysis, IMF_5 and IMF_6 are also the most influential components manifesting the effect of the 11-year cycle. Then what should be removed from the original series?

$\sum_{i=5}^{8} \mathrm{IMF}_i + r$ or just $\sum_{i=5}^{6} \mathrm{IMF}_i$? In our opinion, if we only consider the effect of the 11-year cycle, we should remove only IMF_5 and IMF_6 from the original data. However, if we want to study the scaling behavior without the irregular cycle, which should be the combination of such periods mentioned above and be shown in Fig. 5.14, we should just keep $\sum_{i=1}^{4} \mathrm{IMF}_i$. What we would like to emphasize is that the EMD-based method can preserve the scaling behavior at larger scale. On the other hand, the adaptive detrending method removes all information at scale larger than a pre-determined time scale.

What is puzzling is the invalidation of the Fourier-based method, like the Fourier truncation employed by Movahed et al. (2006), since the Fourier-based method should be a simple but effective method to handle periodic and quasi-periodic trends. Theoretically, the 11-year cycle should be eliminated using the combination of some sine and cosine function with some frequencies. The dif-

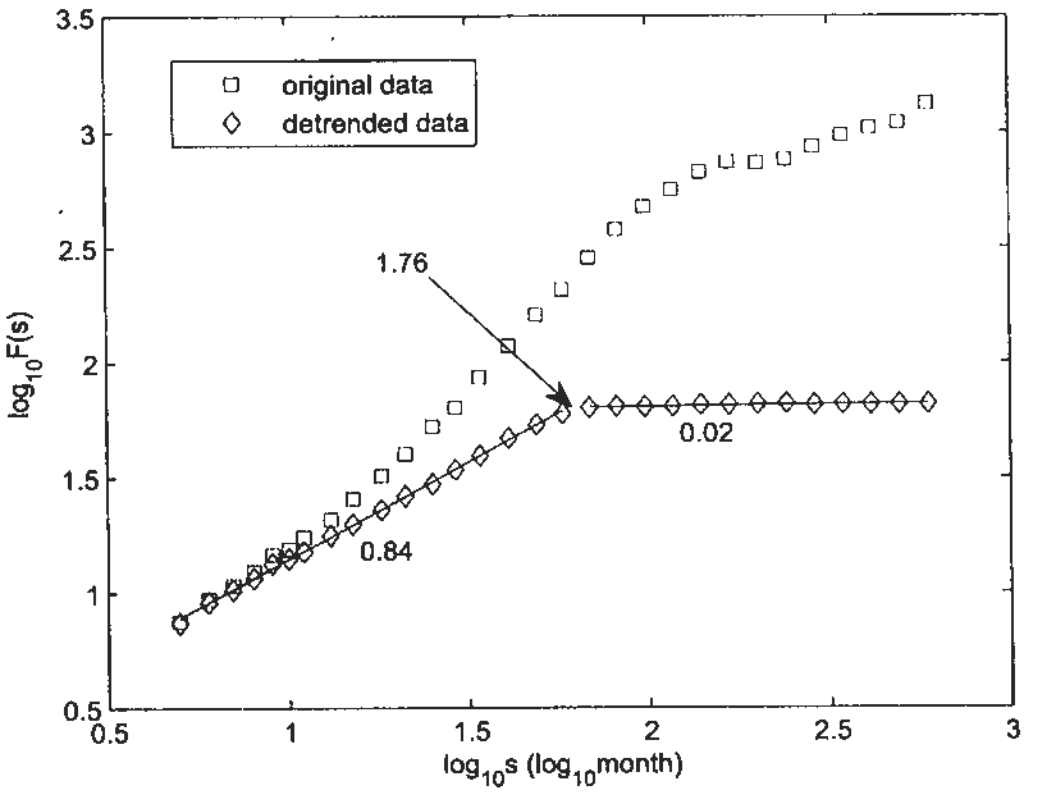

Figure 5.18: Scaling behaviors of the original and the Fourier truncation series (removing the first 50 terms of the Fourier transform of the sunspot time series) using MF-DFA.

ficulty is to decide how many sine and cosine function terms should be chosen.

Thus, the significant differences among the scaling behaviors obtained by the Fourier method (Movahed et al., 2006), the adaptive detrending method (Hu et al., 2009), and the EMD-based method need further cross comparison. To compare, we repeated the Fourier truncation and also truncated the first 50 lowest frequencies. In the practical implement of Fourier transform, what we can obtain are the coefficients corresponding to the discrete frequencies. Specifically, these first lowest frequencies removed should be $i \cdot \frac{1}{3123}$ month $^{-1}$, $i = 1, \cdots, 50$. Interestingly, the scaling behavior discovered is so different from that presented in Movahed et al. (2006), as shown in Fig. 5.18. The crossover point at $10^{1.76} \approx 60$, which corresponds to the 50th frequency of the Fourier transform of the sunspot time series, is the same as that obtained by Hu et al. (2009) and our EMD-based method. The scaling behavior on the left of this point with slope 0.85 is totally different from what Movahed et al. (2006) obtained, and slightly different from the EMD-based method obtained in this study. The fluctuation at larger scales remains almost constant, with slope near zero. This is reasonable since information beyond 60 months has already been removed by the Fourier truncation. Fig. 5.19 is a supplement to the explanation of the Fourier truncation. It can be observed that Fourier truncation provides a close fit of the original data and there are no obvious cycles in the residuals.

What remains is the question why there is still a difference between the results obtained the Fourier truncation and that by Hu et al. (2009) and us (the EMD-based method)? It should be observed that the slope to the left of the crossover point, 0.84, in Fig. 5.18 is close to the slope of the sunspot series, 0.86, by removing IMF_5 at smaller scale. Thus, the reason for the difference should be that the EMD-based method and method of Hu et al. (2009) are adaptive so that the trends removed might contain both larger and smaller frequencies. On the other hand, the Fourier method can completely remove all selected frequencies. Since the residuals of the EMD-based and the method of Hu et al. (2009) contain less information of higher frequencies, the slopes are slightly less than that obtained by Fourier truncation. To substantiate this argument, we remove

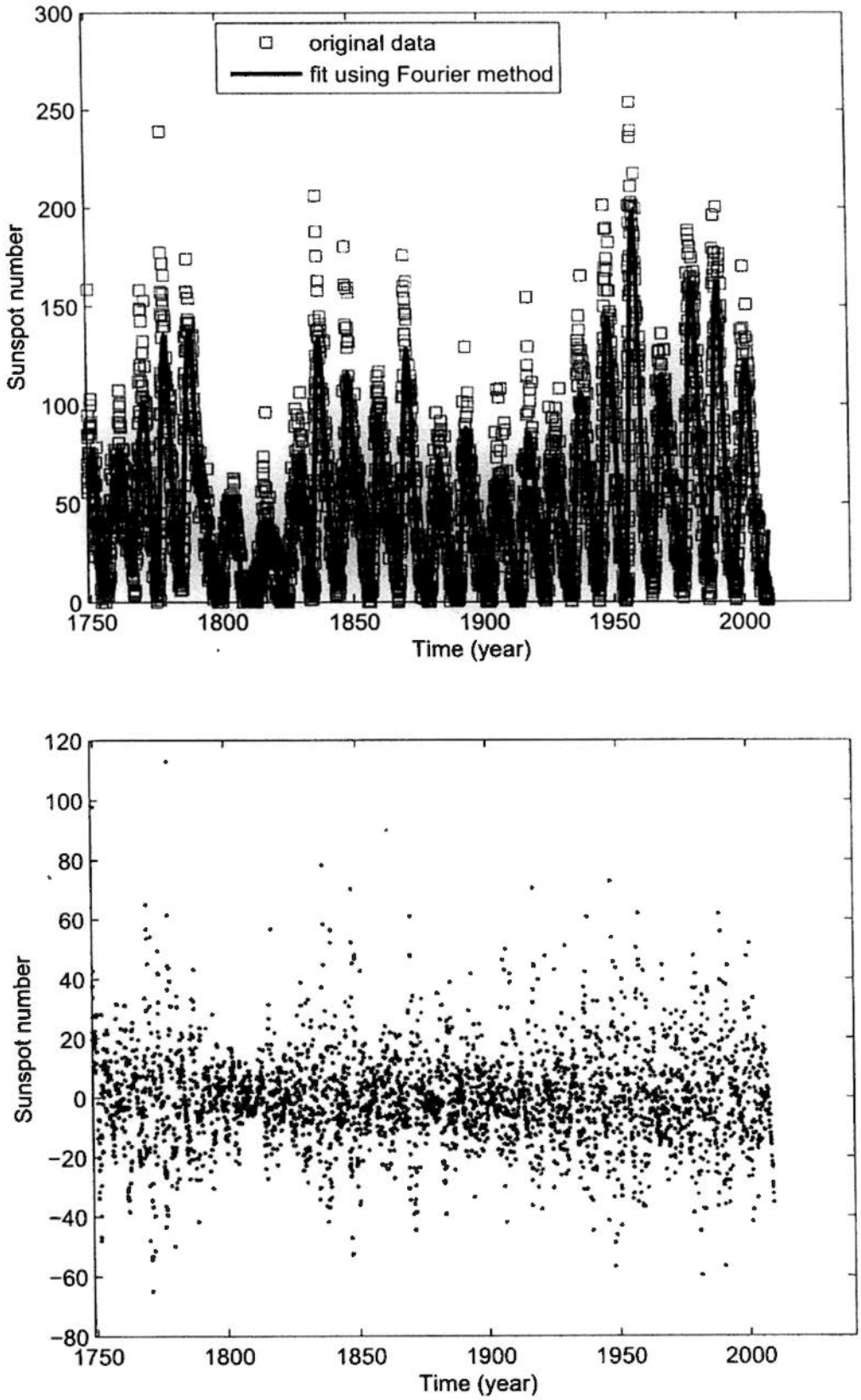

Figure 5.19: Fitted series using Fourier truncation method in comparison with original series (upper panel) and the residual of the fit (bottom panel).

20 more frequencies, e.g. the first 70 frequencies of the Fourier transform of the sunspot series. Then the fluctuation at scale range larger than 45 months should remain constant. The result can be observed in Fig. 5.20. The slope is 0.69 and the crossover point is at $10^{1.69} \approx 48$ months as expected. Fourier truncation is further employed to deal with the situation by considering only the effect of the 11-year cycle. Thus the period (frequency^{-1}) around 11-year corresponding to the 23rd term of the Fourier transform should be removed. Here, we also try to remove the 18~70 terms. The results is also illustrated in Fig. 5.20 visually. The crossover point is $10^{1.76} \approx 60$ months which is similar to what is shown in Fig. 5.10 and Fig. 5.15. Furthermore, the scaling behaviors on both sides of the crossover point are also similar to the results of the EMD-based method which just removes IMF_5 and IMF_6 in Fig. 5.15, 0.72 vs 0.72 and 1.47 vs 1.49 for the left and right sides of the crossover point respectively. Our above conclusion that fluctuation at larger time scale is mainly determined by the lower frequencies is affirmed again. Here, in comparison with the behavior of the series by removing the first 70 terms, we can see the first 17 terms of the Fourier transform increase the fluctuation at the same time scale (can be compared in Fig. 5.20). Nevertheless, the choice of the removed terms of Fourier truncation is a problem which cannot be avoided.

To recapitulate, we summarize the experimental results concerning the slopes and crossover points in Table 5.2.

Table 5.2: Summary of experimental results.

analyzed series	left slopes	position of crossovers	right slopes	corresponding figure
removing IMF_5	0.86	29 (months)	1.56	Fig. 5.13
$\sum_{i=1}^{4} IMF_i$	0.69	60 (months)	0.28	Fig. 5.10, Fig. 5.15
removing IMF_5 and IMF_6	0.72	60 (months)	1.49	Fig. 5.15
removing 1~50th Fourier terms	0.84	60 (months)	0.02	Fig. 5.18
removing 1~70th Fourier terms	0.69	48 (months)	0.01	Fig. 5.20
removing 18~70th Fourier terms	0.72	60 (months)	1.47	Fig. 5.20
result in Hu et al. (2009)	0.74	60 (months)	$\approx$0	

By means of DFA, the influence of these components or their combinations on the scaling behavior of the sunspot time series has been studied. For the scale range less than 60 months, we have discovered property, $H \approx 0.73$, similar to

that of Hu et al. (2009), inconsistent with that of Movahed et al. (2006) when the dominant frequencies of less than 60^{-1} month^{-1} is removed. However, our result show that if we just want to study the fractal property of the sunspot series without the effect of the 11-year cycle, then the slope of this time range changes to 0.69. For time range larger than 60 months, as all information is eliminated by adaptive detrending, the fluctuation maintains constant, i.e. the slope is close to zero. However the EMD-based method obtains the slopes 0.28 and 1.48 respectively. Then, we have repeated the Fourier truncation methods, which should be effective to periodic or quasi-periodic trends, but invalidated by Hu et al. (2009). Interestingly, we have discovered that this method is actually an effective mean to handle time series with periodic and quasi-periodic trends, such as the sunspot series in this study. However, some difficulties are still encountered in choosing the appropriate frequencies. Applying Fourier truncation with the chosen frequency, we can obtain results very close to what are attained by the EMD-based approach under two detrending situations. And it is reasonable that one frequency might contain both noise and useful information, which cannot be extracted by the Fourier-based method. Like what is shown above, the removal of the first 50 and 70 terms leads to different results. Another discovery in this study is the position of the crossover point. We have found out that the crossover point changes with different number of removed terms in the Fourier-based method. Thus, it is the artifact rather than the true reflection of the intrinsic situation. We have further shown that the adaptive detrending is also a good method although it requires a priori determined function forms and time scale for detrending. And the information of the scale larger than the pre-specified scale is all reduced. Even there is useful information contained in these frequencies, we cannot extract them by the Fourier-based methods. Besides, compared with what is attained by adaptive detrending and Fourier filtering, the components obtained by the EMD method are usually reasonable with real physical meaning. Although, as it was obtained through the numerical experiments, this exponent might be inaccurate, the EMD-based method can at least partly reflect information contained in the

series on large scales, which, as suggested by the first experiment in chapter 4, can be of considerable use in the study of the scaling behavior of the series.

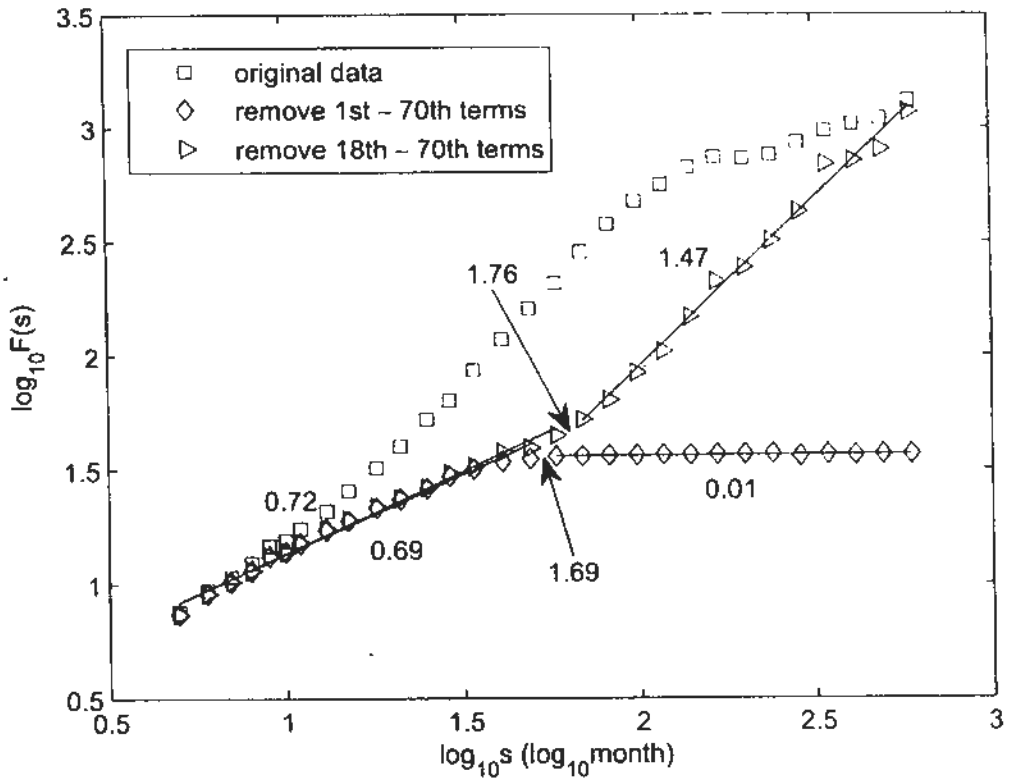

Figure 5.20: Scaling behavior of the original and Fourier truncation series (removing the first 70 and 18th~70th terms of the Fourier transform of the sunspot time series) using MF-DFA.

5.3 Summary

This chapter showed how the TWDFA, MF-TWDFA and EMD-based method work in handling the real-life problems. TWDFA and MF-TWDFA could ensure the better scaling behaviors of the temperature records compared with conventional DFA and MF-TWDFA. Consequently, two critical scales, monthly and seasonal scales, which cannot be found by the conventional DFA, have been found in the annual-detrended temperature series by the proposed model. The EMD-based method handles an challenging problem, the long-range correlation of sunspots activities. On removing the effects of these periods obtained by EMD, the natural long-range correlation of the sunspot time series can be revealed. Particularly,

with the removal of the 11-year cycle, a crossover point located at around 60 months is discovered to be a reasonable point separating two different time scale ranges, $H \approx 0.72$ and $H \approx 1.49$. And on removing all cycles longer than 11 years, we have $H \approx 0.69$ and $H \approx 0.28$. The three cycle-removing methods – Fourier truncation, adaptive detrending and the proposed EMD-based method — are further compared, and possible reasons for the different results are given as the incorrect selection of parameters when employing the Fourier truncation. It is seen that three methods should be equivalent when eliminating the periodic trends, especially at small scales. However, EMD-based method is preferred because of its capability to maintain the information at large scales and the convenience when determining the parameters.

Chapter 6

Application to Earthquake Problems

6.1 Introduction

The occurrence of earthquakes in a given area is a complicated spatio-temporal process (Bak et al., 2002). A better understanding of the pattern of earthquake occurrence is conducive to the study of the dynamics of earthquake processes. Fractal analysis and MFA are known as the mathematical formalism which is able to: handle complex dynamic structures (Feder, 1988; Mandelbrot, 1982), and characterize the heterogeneity of fractal patterns (Grassberger and Procaccia, 1983). Actually, Takayasu (1990) pointed out that 'earthquakes have so many different fractal properties that they belong to the most interesting fractal phenomena'. The fractality and multifractality of the earthquake process has been discovered and measured by many researchers (Goltz, 1997; Harte, 1998; Hirabayashi et al., 1992; Hirata, 1989; Kagan, 1981; Kagan and Knopoff, 1980; Lei and Kusunose, 1999; Sadovskiy et al., 1984; Takayasu, 1990). Therefore, this study will focus on the applications of fractal and multifractal analysis in this research field.

The spatial distributions of epicenters or hypocenters are known to be fractal

(Hirabayashi et al., 1992; Hirata, 1989; Kagan, 1981; Kagan and Knopoff, 1980; Sadovskiy et al., 1984). For example, Sadovskiy et al. (1984) gave the capacity dimension using the box-counting method. Kagan and Knopoff (1980), Hirata (1989) applied the correlation function or integral method (Grassberger and Procaccia, 1983) to calculate the correlation dimension. Multifractal analysis using a fixed-size box-counting algorithm (Halsey et al., 1986) has been performed by Hirabayashi et al. (1992). Meanwhile, Harte (1998) used the Hill estimator to estimate the correlation dimension of epicenters and hypocenters from four earthquake catalogues. Harte (1998) found that the point pattern of shallow and deeper earthquakes in Japan were more tightly clustered than those in New Zealand. In this paper Harte (1998) discussed the affect of the amount of data and boundary. A characteristic scale of around 13 km was found by Lei and Kusunose (1999), who suggested that it is a common feature of the heterogeneity of the crust. Some laws about epicenter location are also relative to fractal statistics, such as the famous Gutenberg-Richter (GR) Law (Gutenberg and Richter, 1949; Stein and Wysession, 2003; Turcotte, 1997) which states that the size distribution of earthquake is scale-free. Davidsen and Paczuski (2005) and Abe and Suzuki (2003) studied the power-law of the distribution of the distances between two successive events using finite size-scaling and a modified Zipf-Mandelbort Law, respectively. And, Molchan and Kronrod (2005) discussed the spatial scaling of the seismicity rate using the California data.

With regard to the temporal aspect, there have been many studies of the fractality, multifractality, and nonlinearity from a number of different perspectives (Balankin et al., 2009; Console and Murru, 2001; Gardner and Knopoff, 1974; Gasperini and Mulargia, 1989; Lennartz et al., 2008; Li et al., 2002; Livina et al., 2005; Marsan et al., 2000; Mega et al., 2003; Shcherbakov et al., 2005; Shlien and Nafi Toksoz, 1970; Telesca et al., 2004a,b, 2005). Shcherbakov et al. (2005) stated that, 'the occurrence of an earthquake is an outcome of complex nonlinear threshold dynamics in the brittle part of Earth's crust.' Omori's Law (Omori, 1894) is used to express the correlation of the main shock and aftershocks. Mean-

while, earthquake frequency has been studied by many researchers. The Generalized Poisson (GP) model is an example of this kind of model (Console and Murru, 2001; Gardner and Knopoff, 1974; Gasperini and Mulargia, 1989; Shlien and Nafi Toksoz, 1970). However, Mega et al. (2003) discovered the inter-cluster correlation by studying a catalog of seismic events in California using Diffusion Entropy (DE), which is used to study the memory in time series (Grigolini et al., 2001). The waiting times between earthquakes (also called the return time or the inter-occurrence time) is another popular topic in the field of temporal analysis (see for example the review by Saichev and Sornette (2007)). The scaling function fitting by a generalized gamma distribution was found by Corral (2004) to be relative to the statistic of the return time in each local bin. Lindman et al. (2005) showed that a characteristic kink in observed waiting time distribution could not separate the correlated and uncorrelated earthquakes. The strong dependence of the recurrence time on the previous times was found by Livina et al. (2005). In the same year, Carbone et al. (2005) found the unified law for the inter-occurrence time. Long-range correlation (also called long-run correlation or long(-term) memory) was found in waiting time using R/S analysis (Goltz, 1997; Jimenez et al., 2006) or by using MF-DFA (Balankin et al., 2009; Lennartz et al., 2008; Marsan et al., 2000; Telesca et al., 2004b). R/S analysis and MF-DFA have also been applied to analyze the temporal distribution of earthquakes in North China (Li et al., 2002) and in the earthquake-related series (Telesca et al., 2004a, 2005). In terms of the waiting times of aftershocks, a model based on a non-homogeneous Poisson process has been established to quantify the scaling behavior by Shcherbakov et al. (2005). Attempts at finding the earthquake precursors are to be found in the study by Goltz (1997). In addition, the precursors in two practical cases were explored by Huang et al. (2001) and Huang (2008) in the M=7.2 Kobe earthquake and the M=8.0 Wenchuan earthquake using the Region-Time-Length (RTL) algorithm.

Generally, earthquakes which happen in same seismic zone ought to be related to each other and follow the same generating mechanism. Sorting the earthquake

events according to the time when they strike, the earthquake spatial and temporal distribution can be regarded as the results of motion of a point. It should be noticed from the above reviews that in order to simplify the earthquake problems many researchers have expressed the earthquake process in the form of a time series by extracting the spatial and temporal distance between two successive earthquake events. The series which they subsequently obtain contains the corresponding earthquake information. This kind of ordered series can be treated and analyzed as a special time series. Many recent studies have focused on the general rules governing these series, for example: the idea of universal jumps by Corral (2006), the betweenness spatial distribution by Davidsen and Paczuski (2005), the power-law distribution of large earthquake times by Mega et al. (2003), their universality by Corral (2004), Bottiglieri et al. (2010), Davidsen and Goltz (2004), and earthquake memory by Livina et al. (2005). Taking into account the spatial and temporal dimensional perspectives simultaneously, Bak et al. (2002); Christensen et al. (2002) proposed a unified scaling law to depict the multidimensional nature of earthquakes and suggested a universal mechanism which is followed by all earthquake processes. The long-range correlation can also be found in Abe and Suzuki (2003); Corral (2006); Davidsen and Paczuski (2005).

It is not enough to wholly profile the occurrence of the earthquakes only considering the events-betweenness distance and time. On one hand, seismologically speaking it is more likely to see the epicenters locating beside the fault lines, which are usually along some directions; hence, the direction which the epicenters migrate along should contain meaningful seismological information. On the other hand, mathematically speaking, with aid of the idea of polar coordinate it can be seen that if we want to determine the position of an event temporally and spatially then the relative information of the waiting time, distance and direction of this earthquake event to the previous event are required. However, despite the seismological and mathematical importance, no work on the direction of the epicenter's migration could be found within the search of the previous literature which was conducted at the start of this study and, therefore, this study aims to

analyze this undeveloped property.

In addition to the real-life earthquakes catalogues, there are some simulative experiments which are designed to facilitate the understanding of the earthquakes; such as, the Acoustic Emission (AE) of rock fracture being analogous to the real-life microshocks. With regard to the probability density function of the waiting time of AE series using different materials, Davidsen et al. (2007) found that they are indistinguishable from that of earthquakes in California which was presented by Corral (2004). Thus, Davidsen et al. (2007) claimed that the rock fracture process should be the same as the earthquake process. In order to draw a comparison between these two views, the data from the two real-life earthquake catalogues (i.e. south China and southern California covering the period from 1970 to 1995) as well as that from the previous rock fracture experiments (Ying et al., 2009) are employed in this study for the changed direction analysis. The rock fracture experimental data was obtained from R. P. Young's research group at the University of Toronto. Experiments were performed by W. Ying and presented in her thesis Laboratory Simulation of Reservoir-induced Seismicity https://tspace.library.utoronto.ca/bitstream/1807/24919/1/Ying_Winnie_W-L_201006_PhD_thesis.pdf. Somewhat similar to the concept of epicenter, the projection of locations of microshocks in rock fracture recorded in three dimensional space on one selected two-dimensional surface are considered in this study, which will still name them as epicenters.

6.2 Epicenter Migration

Given a series $\{\vec{X}_i\}$ recording the locations of N epicenters sorted in the natural temporal order, the motion direction from the $(i-1)$th to ith epicenter can be represented as $\vec{r}_{i-1} = \vec{X}_i - \vec{X}_{i-1}$. Although the locations of epicenters are represented by the latitude and the longitude, this would bring little difference if we treat them under the Cartesian coordinate system for the local area (such as south China and southern California). The changed direction θ can then be obtained from every three successive epicenters as $\theta_{i-1} = \arccos(\vec{r}_i \cdot \vec{r}_{i-1}/(||\vec{r}_i|| \cdot ||\vec{r}_{i-1}||))$.

In this way $\{\theta_i\}$ with length $N-2$ can be obtained for a given series $\{X_i\}$. This study will consider the foreshocks, mainshocks, and aftershocks together. This is the same technique that is used in the analyses of the studies by Bak et al. (2002); Christensen et al. (2002); Corral (2004, 2006); Davidsen and Goltz (2004); Davidsen and Paczuski (2005); Livina et al. (2005).

For the obtained series $\{\theta_i\}$, the Detrended Fluctuation Analysis (DFA) is chosen in order to study its scaling. DFA is proposed to be employed as a technique to investigate the long-term correlation of the non-stationary, as well as the stationary, time series (Peng et al., 1994). This capability is the reason why the DFA is so popularly used in scaling analysis of time series. In their investigation of atmospheric variability, Koscielny-Bunde et al. (1998) reported a universal scaling law of long-term temperature records by analyzing fourteen meteorological stations around the world employing DFA. Following this study, Govindan et al. (2002) and Fraedrich and Blender (2003) used DFA to study the temperature series generated by the climate models under different scenarios and found a different scaling pattern from that of the observational records. In addition, DFA has been employed in seismology Currenti et al. (2005); Telesca et al. (2001a,b, 2003, 2008).

The 1970-1995 earthquake catalogue of south China contains 13,653 M$\geq$ 1 events. In this study the events with M$\geq$ 1 which were recorded in southern California during the same period have been downloaded from http://www.data.scec-.org. The basic patterns of the epicenter distribution in these two study areas are visually presented below in Figs. 6.1 and 6.2:

The scalings curves for the two areas which were obtained by using DFA (Fig. 6.3) exhibit a similar pattern consisting of two power-law parts, as can be seen in the figures below: The first part yields a slope around 0.5, which is indicative of a random process. The second part exhibits a significant positive long-range correlation with a slope equal about 0.75. A slight difference exists between the two scaling behaviors in the location of the crossover point separating the two scaling ranges. Since the lengths of the analyzed series for the two areas

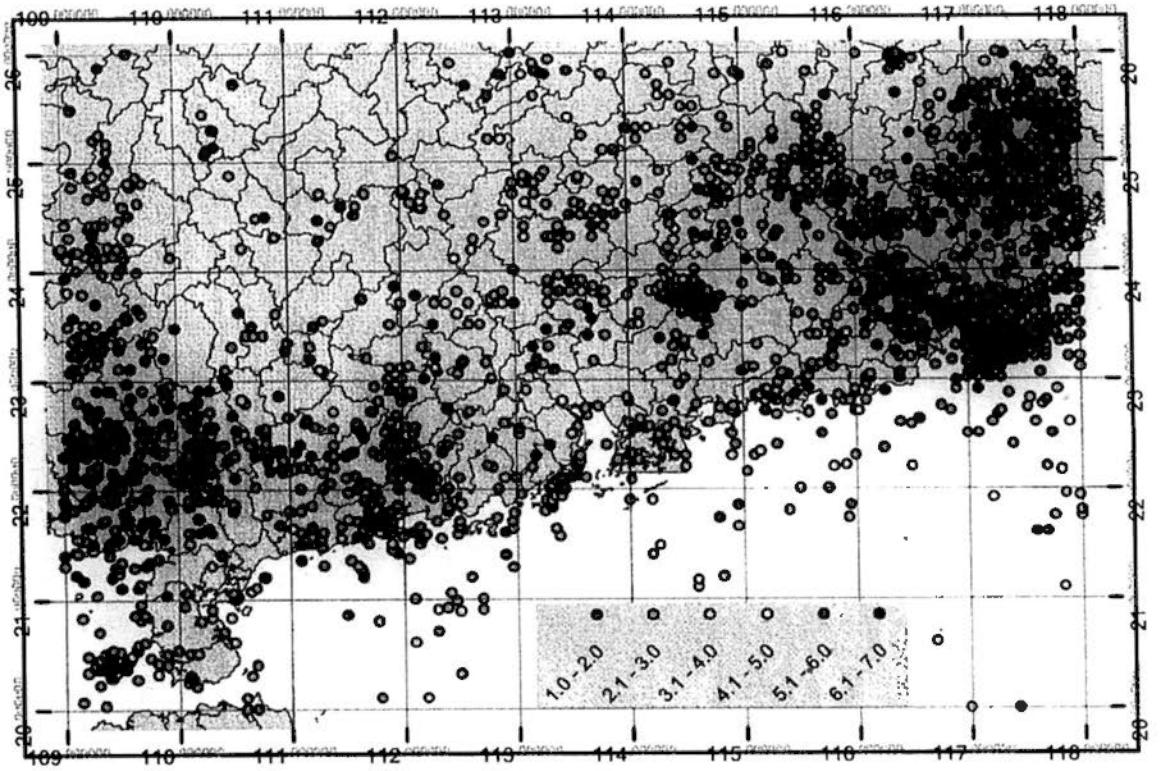

Figure 6.1: Distribution of epicenters in south China from 1970 to 1995.

are different, the positions of the crossover points are represented by their relative position in the series (with the investigated scaling range of each catalogue divided into 60 equal intervals at the logarithmic scales). The crossover points are located the 37th and 45th interval for south China and southern California, respectively.

After obtaining the double power-law scaling behavior, the sensitivity analysis is performed from three aspects, namely: the effects of threshold magnitude, incompleteness, and the edge of the study area. This study is first concerned with the possible effects of the magnitude range used in the analysis, data incompleteness and boundary of the studied area on the DFA results. Sensitivity tests were conducted to test whether these factors should substantially modify the scaling behavior. In order to determine the effect of the threshold magnitude, events with magnitude less than a threshold value are removed from the catalogues and subjected to the DFA. The scaling curves using different threshold magnitudes are also shown in Fig. 6.3. The resulting curves retain the two-part scaling structure for all threshold values. The slope of the first scaling range remains at 0.5,

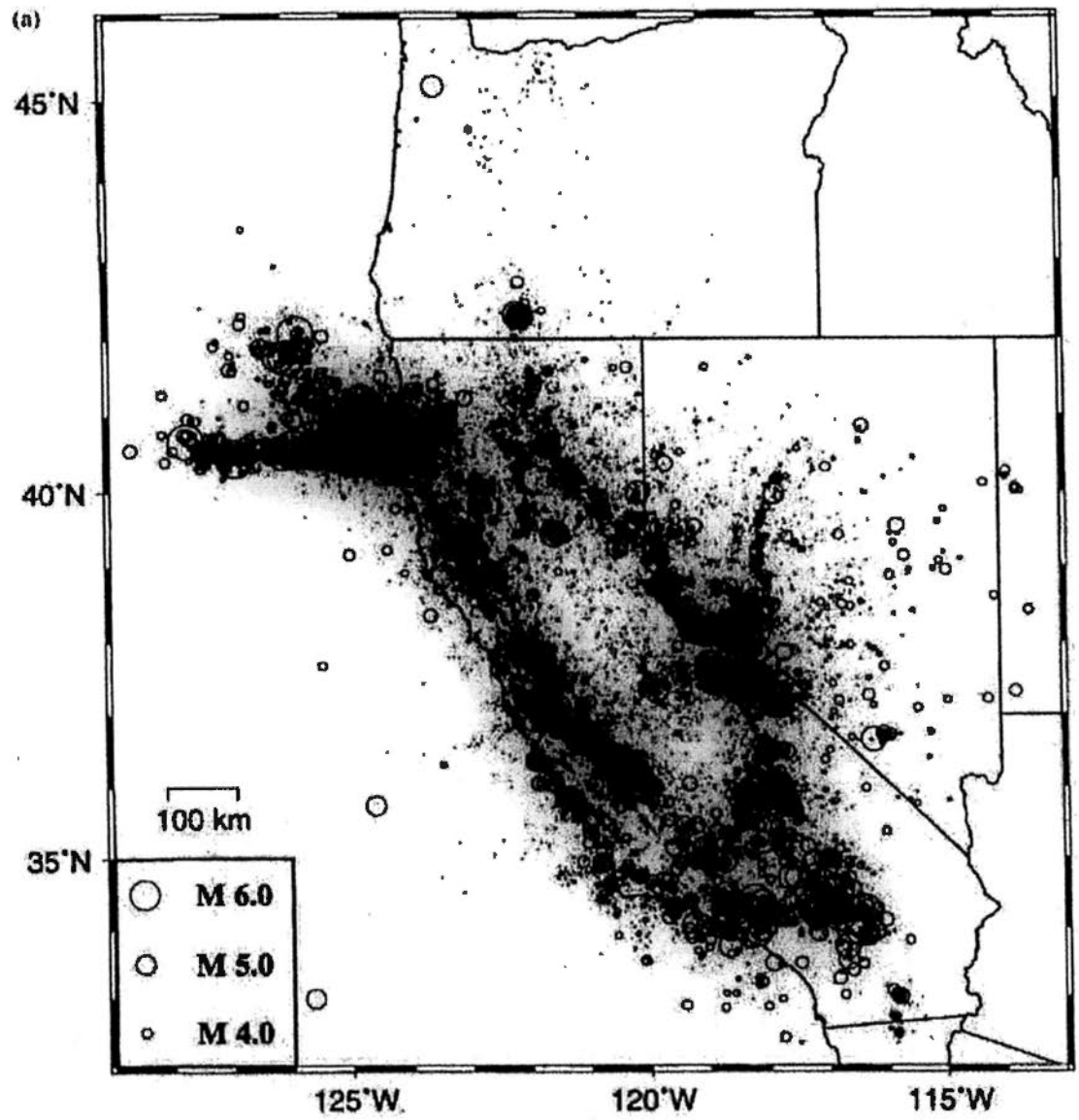

Figure 6.2: Distribution of epicenters in southern California during the period 1975 to 1995, which is part of our study period. (This figure is adopted from Godano et al. (1999).)

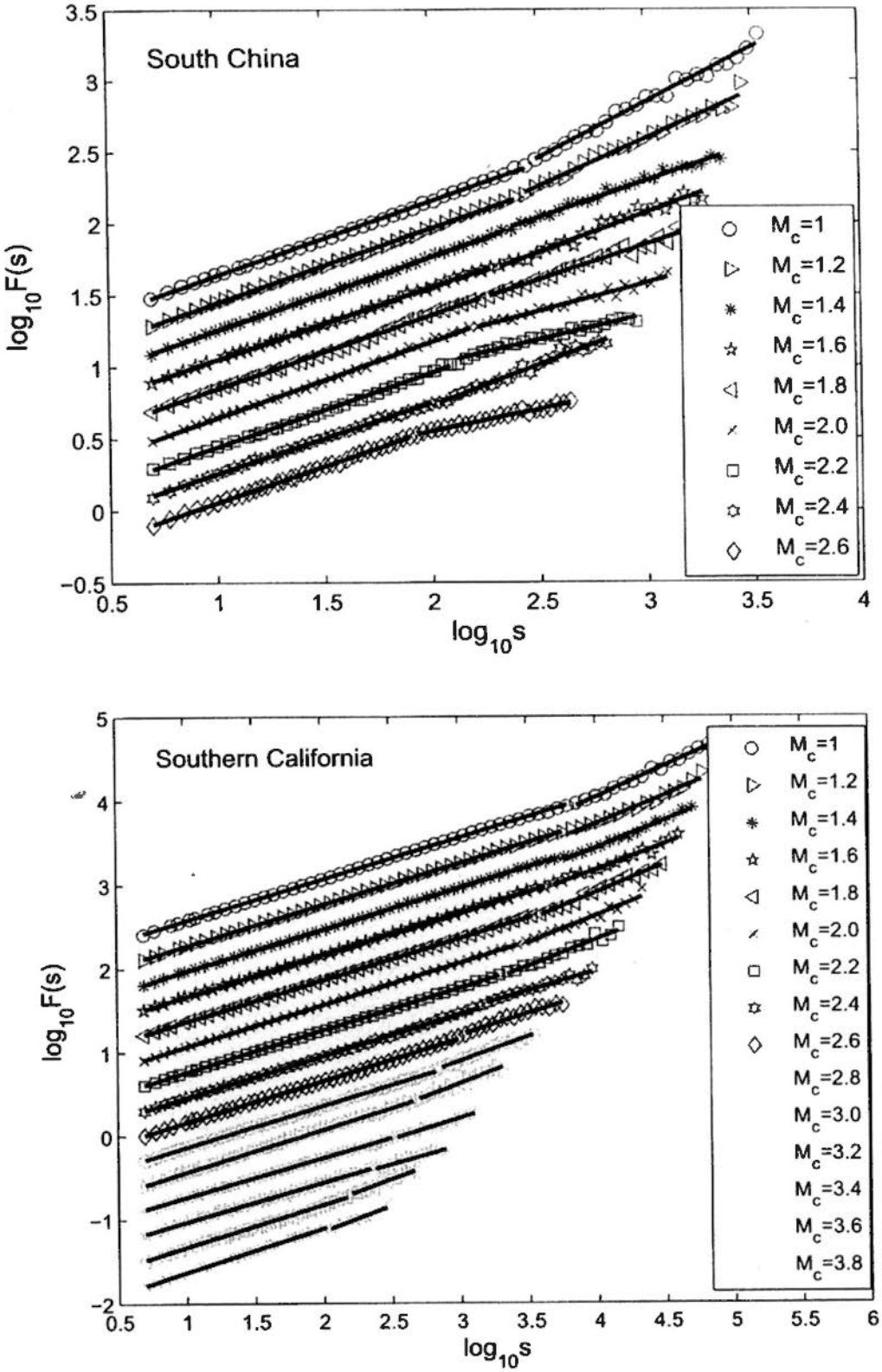

Figure 6.3: The scalings of the catalogues with different threshold magnitudes of south China (the upper panel) and southern California (the bottom panel) were obtained by DFA. The selection of threshold magnitudes takes into consideration that there should be sufficient earthquake events left to ensure the statable results of DFA. The solid straight lines are estimated by linear regression. These scalings are shifted vertically for the sake of better illustration.

while the slope of the second scaling range decreases with an increasing threshold magnitude approaching and fluctuating around the value 0.5. Thus, in this study the double power-law scaling behavior seems more apparent for small events.

For testing the effects of data incompleteness, a portion of events were randomly removed and the reduced catalogue was re-analyzed using DFA. As shown in Fig. 6.4, no significant modification to the scaling pattern is noted when up 15% of the events are removed. When a greater percentage of events was removed, the slope of the second scaling range begins to decrease.

The effect of area boundary on the scaling behavior was also studied by confining the data set to events located beyond a certain distance from the edge of the study area in the DFA. The threshold distances used were set at 5%, 10%, 15% and 20% of the width or length of the area. As shown in Fig. 6.5, the confinement of data to within a certain distance from the edge reveals no significant modification in the scaling pattern. The above studies show that the scaling behavior is highly insensitive to the threshold magnitude, data incompleteness and proximity of events to the boundary of the study area. The scaling behavior is, therefore, a property of the seismicity of the area and possibly bears pertinent information on the dynamics of the seismicity of the area.

Since Davidsen et al. (2007) suggested that rock fracture is similar to the earthquake process, this study will perform the DFA on the series $\{\theta_i\}$ obtained from the rock fracture experiment on sandstone. The upper panel of Fig. 6.6 shows that the scaling is independent of the choice of the surface of projection, and no crossover point can be found in the rock fracture process (e.g. the slopes for the epicenters on three surfaces are about 0.5 in the whole scaling range). We may then select their projections on one of three surfaces and preform the randomly-removal and boundary-zooming-in analysis. The right panel of Fig. 6.6 shows that the scalings are maintained the slopes about 0.5 in all different analysis conditions. Hence, the scaling pattern of rock fracture is different from that of the epicenter in two real-life earthquake catalogues. Only the power-law with slope 0.5 could be found for rock fracture processes comparing with the two-part-scaling

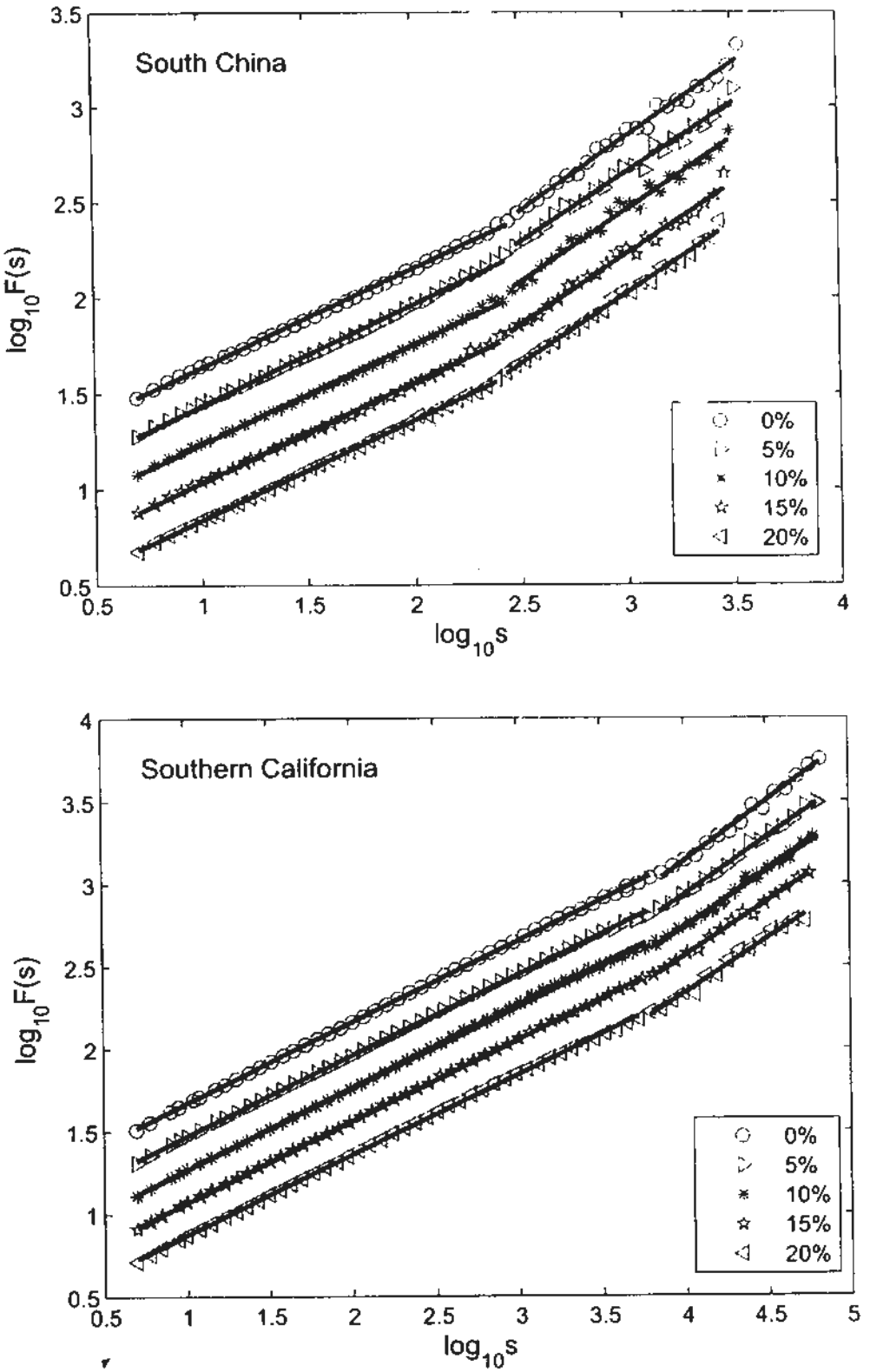

Figure 6.4: The scalings of the catalogues with events randomly removed occupying different percentages of south China (the upper panel) and southern California (the bottom panel) obtained by DFA. The scalings without any event being removed are also presented for comparison and are marked by the legend corresponding to '0%'. The solid straight lines are estimated by linear regression. These scalings are shifted vertically for the sake of better illustration.

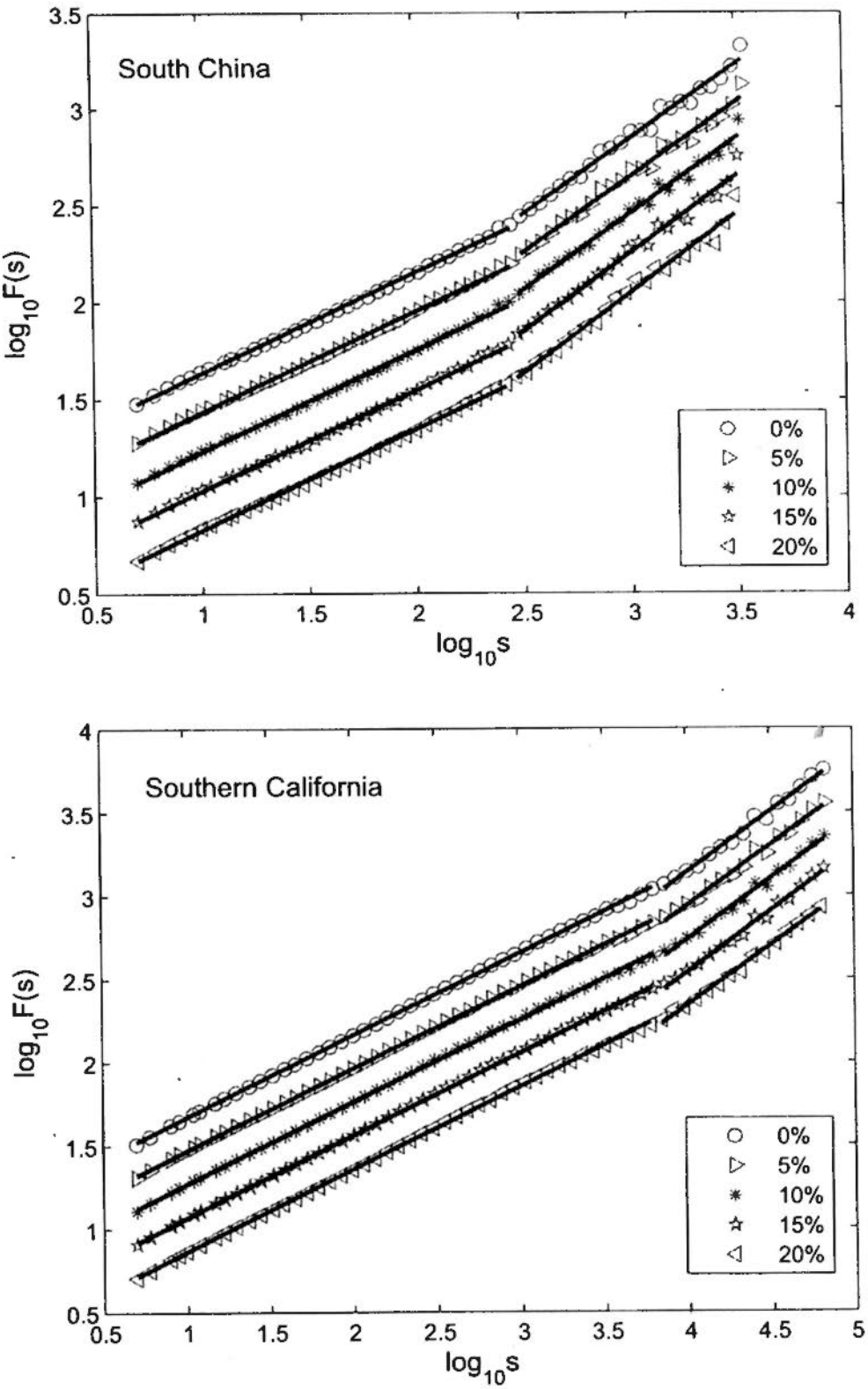

Figure 6.5: The scalings of the catalogues with their area zooming in from the boundary by different percentages of south China (the upper panel) and southern California (the bottom panel) were obtained by DFA. The solid straight lines are estimated by linear regression. These scalings are shifted vertically for the sake of better illustration.

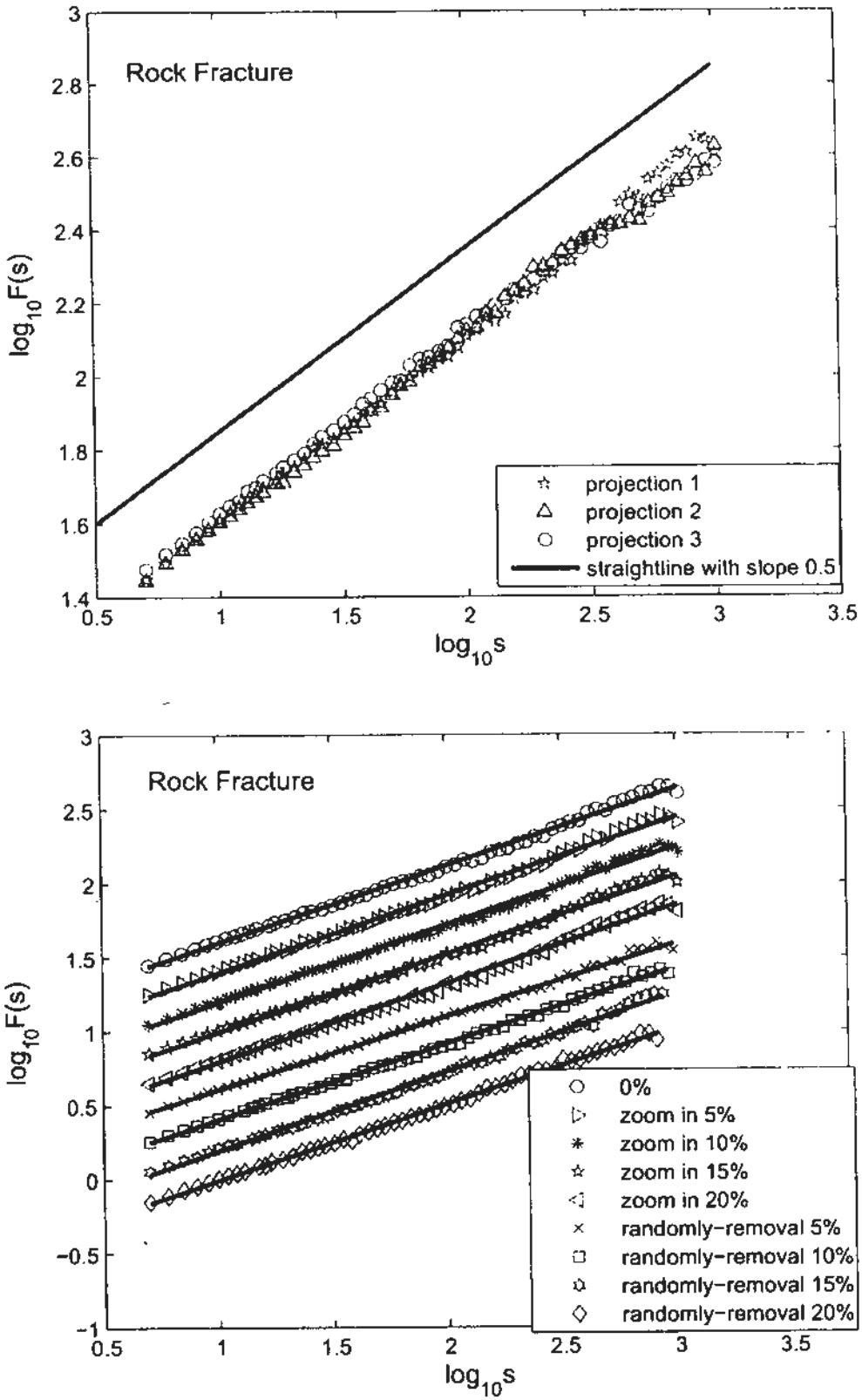

Figure 6.6: The scalings of the epicenter of microshocks of the rock fracture process projecting on different surfaces (the upper panel) and in different analysis conditions were obtained by DFA. The solid straight lines are estimated by linear regression. These scalings are shifted vertically for the sake of better illustration.

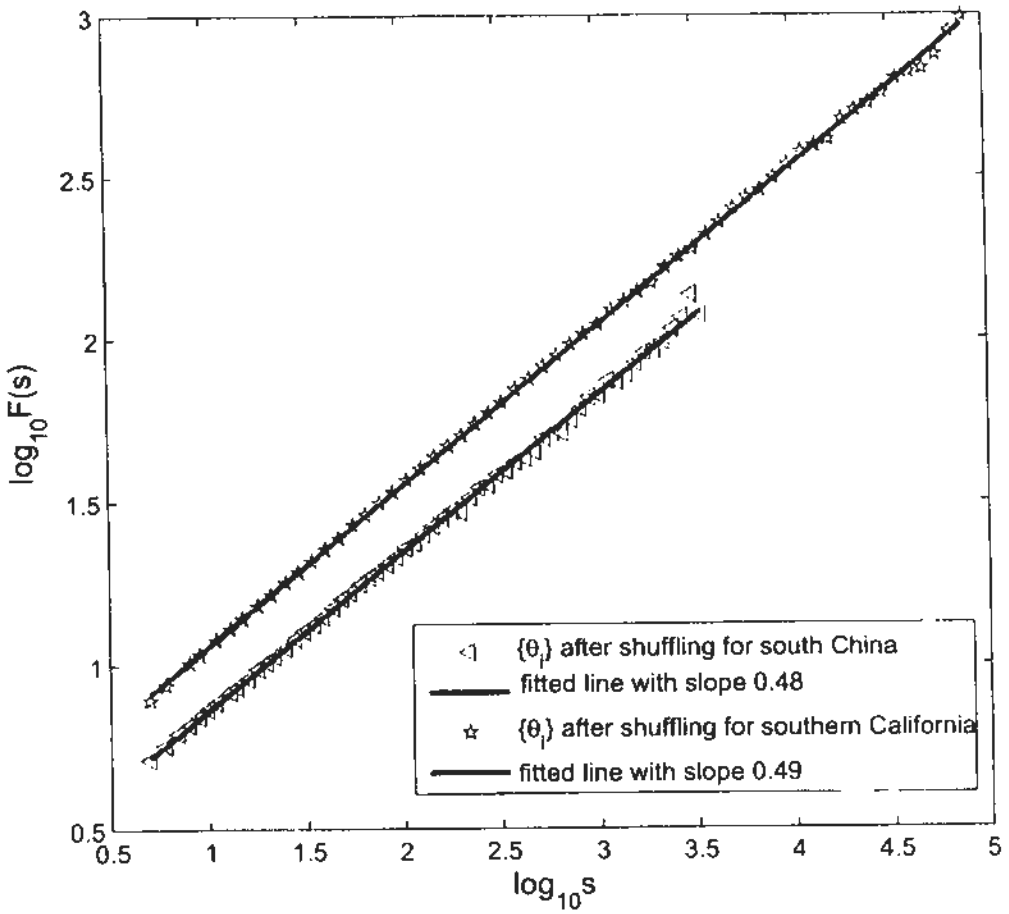

Figure 6.7: The scalings of $\{\theta_i\}$ obtained from shuffled $\{\vec{X}_i\}$ for south China and southern California are obtained by DFA. These scalings are shifted vertically for the sake of better illustration.

pattern with slope about 0.5 and 0.75 for south China and southern California, respectively.

The series $\{\theta_i\}$, describing the change in the migration direction of epicenter location, theoretically provides information different from that of the spatial and temporal distance ($\{T\}$ and $\{D\}$) between successive earthquake events. This study has quantitatively studied the relationship among the three series $\{\theta\}$, $\{T\}$, and $\{D\}$. First to be investigated was the dependence of $\{\theta_i\}$ on the temporal order of earthquake events. By shuffling the temporal order of $\{\vec{X}_i\}$ the epicenter distribution is still kept; however, as shown in Fig. 6.7, the two-part scaling pattern of $\{\theta\}$ is found to diminish and a single power-law with slope about 0.5 appears for both study areas. Therefore, the two-part scaling pattern is closely

related to temporal order of events but it has little relationship to the spatial distribution of epicenters. The correlation coefficient, C, of the pairs of $\{\theta\}$, $\{T\}$, and $\{D\}$ are calculated. For south China the computed $C_{T,\theta}$, $C_{T,D}$, and $C_{\theta,D}$ are about 0.02, 0.12, and 0.2, respectively. For southern California the results are around 0.001, 0.07, and 0.2 and for rock fracture they are nearly 0.03, 0.004, and 0.26, correspondingly. Hence, $\{\theta\}$, as a new perspective of analysis is almost independent of the waiting time, but it is sightly correlated to the betweenness distance. This researcher feels that this should be happening because both $\{\theta\}$ and $\{D\}$ are extracted from the locations of the epicenters, even though the correlation coefficient between them is still low. Therefore, this study holds that the $\{\theta\}$ contains the new information which can be reflected by neither $\{T\}$ nor $\{D\}$.

As a supplement to the previous studies on the spatial and temporal distance between two successive earthquake events, this thesis will propose a new analysis perspective focusing on the direction of earthquake migration. The changed direction is first extracted from the locations of epicenters and is denoted by the angles, $\{\theta_i\}$, between two migration directions. The DFA is then employed to investigate $\{\theta_i\}$ for the catalogues of south China and southern California with earthquake events holding minimum magnitude 1. The same scaling structure consisting of two ranges is found for both catalogues.

The analysis above reveals the characteristic scaling behavior of the change in direction of earthquake migration. From this the following inferences can be made. Firstly, the scaling curves for both catalogues comprise two parts: the first range is dominated by the power-law slope 0.5 while the second range is about 0.75. The positions of the crossover points are different for the two catalogues. These differences probably contain information pertaining to local seismicity. Secondly, only the slope of the second scaling range changes with the threshold magnitude used in the analysis, the slope of the first scaling range remains at 0.5 for all threshold magnitudes. Thirdly, the location of the crossover points is found to depend on the length of analyzed series $\{\theta\}$. For further comparison, a simi-

lar analysis is made on simulated microshock process. Consequently, the whole scaling range is dominated by the power-law scaling with slope 0.5. The simple composition of the material used in the rock fracture experiment may be able to explain this difference.

The scaling structure reveals a basic pattern of earthquake migration with respect to the temporal order of earthquake events. From an average view, the crossover point indicates a critical scale for the study area. The earthquake events within this scale migrate randomly. However, if the observable scale is beyond the critical one then the memory of migration appears. From the directional analysis of earthquake migration, the pattern can be interpreted to consist of several clusters of earthquake events with an average size which is quantified by the critical scale. Within a cluster, the migration direction of earthquake events can be considered random; however, the migration direction of the clusters exhibits a control by some underlying factors. These factors may be related to the fault pattern in an area and the tectonic factors governing the recurrence of earthquakes. Such proportion and the randomness in each cluster is maintained under all conditions considered here, only the slope in the second scaling range decreases with the increasing threshold magnitudes approaching the value 0.5. However, the extent of the decrease for two catalogues is different. Such difference may be attributed to the difference in fault pattern in the two areas: the fault system in south China shows a rectangular pattern while that in southern California shows a anastomosing pattern. The earthquakes are more likely to happen on fault lines. Conceivably, the change in epicentral migration direction is relatively larger in a rectangular fault system than that in an anastomosing system.

6.3 Summary

In summary, a universal scaling pattern of earthquake migration is revealed by directional analysis. The analysis in this study suggests that the real-life earthquake migration at scales less than the critical scale actually follows a random

process; however, a positive long-range correlation appears at large scales. Such correlation and the critical scale is highly insensitive to the random-removal or the effect of boundary of the analyzed area but it is related to the threshold magnitudes. However, the positive long-range correlation at large scales approaches to randomness as the threshold magnitudes increase, even though there are some fluctuations around the random status.

With regard to earthquake migration, the migration direction, spatial and temporal between-event distance depict the seismic process from different perspectives. Should the unified scaling law be established on the basis of combination of the these three parameters, which is similar to what have been done by Bak et al. (2002) to combine the G-R Law, Omori Law, and the epicenter distribution together? Actually, as we have shown, since the new parameter, migration direction, the waiting time and the betweenness distance have very little dependence on one another, it would not make too much of a difference even if we combine them together. Then, in future study, we would like to investigate these parameters separately to ensure a complete description of the earthquake migration and the simplicity to perform the analysis.

Chapter 7

Conclusion

The results of this thesis are summarized in this chapter. The main results will be presented first.

7.1 Main Results

This thesis aims to render a rigorous and systematic study of geographical phenomena in multiple temporal and spatial scales with the aid of fractal and multifractal analyses. To achieve this the study first worked on a conceptual level. A number of theoretical issues have been discussed, and the academic debate on these has been reviewed. The corresponding works have been proposed to refine the framework of the fractal and multifractal analyses which will be conducted in this study. After building a solid basis, this study applied the fractal and multifractal analyses in the geographical case studies. Some of the significant properties of these geographical processes has been uncovered.

On the basis of the reviews of the mainstream algorithms for fractal and multifractal analyses, taking the speciality of the geographical problems into consideration, the DFA and MF-DFA were selected as the preferred methods to be used in this study due to: their simple procedure, small finite size effect, stable scaling behavior, and their capability to study the structures and processes disturbed by unknown trends or noises.

Despite their many advantages, DFA and MF-DFA are not entirely prefect methods. In this study, after elaborative investigation two weakness of DFA and MF-DFA, and two problematical relationships of MF-DFA have been identified, and corresponding improvements and corrections have been suggested. This study has proposed TWDFA and MF-TWDFA to obtain better scaling behavior with less fluctuations. An EMD-based method was also developed as a preprocessing in order to eliminate the negative influence of periodic trends on the results of DFA and MF-DFA. Two problematic relationships in one- and two-dimensional space were corrected based on the previous studies, which are both theoretical and experimental in nature. These corrections can be employed to reveal accurate relationships in the formalism of multifractal analysis for general situations.

DFA, MF-DFA and their modified versions are applied in this study of geographical problems. The better performance of improved DFA and MF-DFA is shown in two of the case studies, which are: temperature variation and sunspot activity. A natural mechanism of earthquake process has been revealed in this study based on the scaling analysis from a new perspective using DFA. The conclusions of this analysis are briefly described below.

7.1.1 TWDFA and MF-TWDFA

It has been argued that points nearer in time are more related than points some distance apart. Correspondingly, we have proposed a locally detrended mechanism for MF-DFA, and the model is called MF-TWDFA in general and MF-MWDFA in particular. The theoretical arguments have been supported by numerical simulations and a real-life problem. Specifically, this study has demonstrated that MF-TWDFA can indeed improve the linear relationship between the fluctuation function $F_q(s)$ and the scale s. Therefore, crossover point(s) can be effectively detected. In particular, it is easier to detect scaling behavior for large-range scales which cannot be detected by the conventional MF-DFA otherwise. With this technical advantage, we managed to find the crossover of timescale

from one month and one season in the temperature series elusive as regards detection by the conventional MF-DFA. The weekly scale seems to correspond to the 'general weather regimes' or 'Grosswetterlagen' in climatology. The monthly and seasonal scale might correspond to the usual scale of El Niño, the climate anomaly. The significance of these two crossover points is echoed by the results obtained by other researchers via the statistical and wavelet methods (Ding et al., 2002). On the basis of these two crossover points, we further found that sudden changes of temperature are more likely to happen at the weekly rather than the monthly scale. And the effect of the El Niño events on the long-range correlation of temperature of Hong Kong in the corresponding timescale region is also reflected. Besides, the scaling behavior of the temperature series of Hong Kong on a scale longer than one season unraveled by the MF-TWDFA appears to be more reasonable and interpretable.

To recapitulate, a locally detrended mechanism for MF-DFA has been proposed, namely the MF-TWDFA in general and the MF-MWDFA in particular. The proposed methods can also be treated as a kind of combination of MF-DFA and DMA since the moving-window technique has been adopted for improving the estimation of local trends. The validation of the relation between the MWDFA and TWDFA scaling exponent and the Hurst exponent has been shown in the appendix. Besides, simulated numerical experiments are also employed to validate the proposed methods. Equivalent performances of MF-TWDFA and MF-DFA on the correlation property and multifractality nature of these constructed series are confirmed. Besides, we argue that a better linear relationship can be obtained using MF-TWDFA. Therefore, the crossover points, which indicate changes in the correlation properties of time series at different scales (Hu et al., 2001) and draw much attention in the applications of MF-DFA (Eichner et al., 2003; Talkner and Weber, 2000), are more easily detected and located. In addition, the uncertainty due to the large fluctuations around the regression line (Lennartz and Bunde, 2009) can be reduced by the proposed methods. Compared against the results of conventional MF-DFA, this kind of uncertainty, especially at large scales, as

shown in the real-life application, can indeed be reduced. RMSF is employed to quantitatively support this advantage. Another benefit gained from this technical advantage is that the scaling behavior, which usually cannot be detected because of the great fluctuation for a range larger than $N/4$, at the scale range up to $N/2$ can still be studied in this application. Furthermore, the conclusion drawn on the basis of the newly detected crossover points seems to bear more physical meaning. The use of MF-MWDFA and MF-TWDFA is in particular more reasonable when linear regression is employed as the basis for detrending.

7.1.2 EMD-Based Method

The EMD method has been employed to decompose the sunspot time series into eight IMFs and one r. As regards their physical meanings, IMF_2 to IMF_8 should correspond to the 153-day cycle of γ-ray flares and interplanetary magnetic fields, the 1.3-1.4 year periods, QBO, the 11-year cycle, and the Hale, double Hale and Gleissberg periods respectively. This study explored the influence of these components or their combinations on the scaling behavior of the sunspot time series using DFA. For the scale ranges less than 60 months, the finding is similar to that of Hu et al. (2009) while is inconsistent with that of Movahed et al. (2006), provided the dominant frequencies of less than 60^{-1}/month is removed. For time range larger than 60 months, compared with the complete loss of the information in the results of Hu et al. (2009), the EMD-based method obtains the slopes 0.28 and 1.49 respectively. This result indicates that the scaling behavior can still be found for the series with the 11-year cycles removed on the larger scales. The slope greater than 1 suggests that this detrended series seems to be non-stationary and with Hurst exponent 0.49. Although, as it was obtained through a numerical experiment, this exponent might be inaccurate, the EMD-based method can at least partly reflect information contained in the series on large scales, which, as suggested by the first experiment, can be of considerable use in the study of the scaling behavior of the series.

Two numerical experiments have also been performed to test these three de-

trending methods. None of them can completely remove the influences of the added periodic trends. Although spurious crossover points appear, the effect of the periodic trends on relatively small scales can be satisfactorily eliminated by the three methods. While almost all information on the large scales is removed by the adaptive detrending method, the other two methods can partly maintain — even considerably in the first experiment — the information contained in the series. It should be noted that if appropriate parameters are selected, the results, including the scaling behavior and positioning of the crossover points, obtained by the EMD-based and Fourier methods can be very similar, which can also be found in the sunspot series analysis. However, the determination of the setting for the EMD-based method is much easier and more natural than for the other two methods.

In summary, the EMD-based method is strongly recommended since it is a totally data driven and adaptive method. In general, mostly 11-13 IMFs are obtained by EMD. It is very convenient to study the effect of each component compared with the large number of terms of the Fourier-based method. And it is reasonable to draw the conclusion that $H \approx 0.72$ and 1.49 for two different time scale ranges divided by a crossover point located at about 60 months if the 11-year cycle is removed, and $H \approx 0.69$ and 0.28 for time ranges divided by the crossover point of 60 months with all cycles longer than 11 years removed.

7.1.3 Problematic Relationship in 1D MF-DFA

MF-DFA is a generalized version of DFA developed for the purpose of characterizing the fractal and multifractal properties of stationary and non-stationary signals. Kantelhardt et al. (2002) established the relationship for stationary, positive series in Eq.(3.44) which connects $h(q)$ to $\tau(q)$ for MF-DFA. This relationship has been used in a large number of researches. We have shown in this study that Eq.(3.44) is in fact valid only under the special situation $H' = 0$. Our argument has also been substantiated by numerical experiments and theoretic derivation which suggest that the generalized relationship between $h(q)$ and $\tau(q)$ should be

$\tau(q) = qh(q) - qH' - 1$ instead of the one stipulated in Eq.(3.44) for the stationary, positive time series. The singular spectrum, α and $f(\alpha)$, based on $h(q)$ have also been proposed in Eq.(4.39) and Eq.(4.40) respectively to replace the existing ones stipulated in Eq.(3.48).

To recapitulate, this study has conceptually argued and experimentally demonstrated one problem in MF-DFA. That is the relationship between $h(q)$ and $\tau(q)$ in MF-DFA. A modified relationship on the plausible form about this problem has been suggested.

7.1.4 Problematic Relationship in 2D MF-DFA

In the extension of MF-DFA to two-dimensional space by Gu and Zhou (2006), some corresponding relationships of the exponents have also been extended.

In this thesis, we have casted doubt on the validity of two relationships in Gu and Zhou (2006), which connect the Hurst exponent H to $h(q = 2)$ in Eq.(3.58) in Gu and Zhou (2006). Such relationships not only lack the theoretical derivations, they also contradict some empirical understandings. Specifically, (i) establish the correct relationship between H and $h(q = 2)$ is critical for obtaining accurate Hurst exponent from the scaling exponent $h(q)$ of 2D MF-DFA. However, the original one in Eq.(3.58) is not in synchrony with the 1D situation. Based on our numerical experiments, we have brought forth a new relationship, Eq.(4.42), which appears to be more natural and logical, and is in synchrony with the 1D situation. In addition, we have also shown by numerical experiments that it is incorrect to use 2D R/S analysis Alvarez-Ramirez et al. (2008) to analyze non-stationary surfaces.

In summary, we have conceptually argued and experimentally and empirically demonstrated the problematical relationship of the exponents in 2D MF-DFA. A more logical and reasonable relationship was conjectured by us.

7.1.5 Universal Scaling Structure of Epicenter Migration

This study has proposed a new parameter, which is the θ independence of the previous event-betweenness spatial and temporal distance, in order to capture the directional information of earthquake migration. Common findings for two of the case studies (i.e., south China and southern California) can be used to uncover the natural mechanisms of the earthquake process. Firstly, a double power-law scaling pattern is revealed through the directional analysis using DFA. Secondly, the scaling exponents of two scaling range are equal 0.5 and 0.75, respectively. Meanwhile, the whole scaling range is dominated by the slope 0.5 for the microshock data. A critical scale could be indicated by the crossover point dividing two ranges. Our analysis in this study suggests that some clusters are formed during the earthquake migration, with averaged size characterized by the critical scale. The directional independent events can be identified within each cluster; however, a positive correlation exists outside the clusters. Thirdly, this correlation and the critical scale is highly insensitive to the randomly-removal or the effect of boundary of the analyzed area, but it is related to the threshold magnitudes. The double power-law scaling pattern becomes less apparent when threshold magnitude increases. Fourthly, as the threshold value increases, the slopes at large scales approach 0.5 even though there are still some fluctuations around 0.5. The slight differences between the results of the two areas are the relative locations of crossover points and the different extent of the decrease of the slopes in the second scaling range, hinting at the different geological structure of south China and southern California (e.g. in the differing structures of the two fault systems). For southern California, the crossover point appears at the scale 0.75=45/60, with scaling range normalized to 1. In the same study area, a kink could be found in Fig. 4 of Bak et al. (2002), and a crossover point separating the power-law is to be found at the left side to another scaling range in Figs. 1 and 2 of Davidsen and Paczuski (2005). Interestingly, the scales indicated by the kink and crossover point are both about 0.75 provided the scaling range normalized to 1. Thus, this scale should contain the special seismical information for southern

California.

7.2 Research Limitations and Suggestions for Further Research

In spite of great efforts which have been made in this study to improve the DFA and MF-DFA and so better understand the earthquake process, there still are some limitations to this research. Some suggestions for further research based on these limitations are also listed here.

1. Although the improvement of the TWDFA and MF-TWDFA is apparent they are, however, computationally expensive because the time distance matrix W_t needs to be computed. How to develop a fast algorithm for the proposed analysis is one problem which needs to be addressed by further research. In addition, choosing a suitable w_{ij} for MF-TWDFA, although not unique to the proposed model, is another topic that needs to be investigated further.

2. Although the scaling behavior at small scales could be maintained well during the procedure of the EMD-based method to eliminate the periodic trend, its original scaling pattern at large scales is affected (as shown in the numerical experiments). Therefore, a better algorithm to keep the information at all scales is required in order to better reduce the influence of the periodic trend.

3. In the discussions on the relationship in 2D MF-DFA, a conjecture was made in this thesis as to the corrected relationship; however, this new relationship was suggested on the basis of numeric experiments and empirical derivation. Solid mathematical study, such as the formal study or analytical calculation, is so far lacking in this study because of the research limitations. It is to be recommended, therefore, that further research should conduct an in-depth investigation from the theoretical aspect.

4. With regard to the exploration of the earthquake process, a very interesting pattern of the scaling behavior has been found for different earthquake catalogues, which suggests the a natural mechanism of earthquakes has been revealed. Some interpretations of the geophysical aspects of this has been offered; however, the seismical meaning of the findings of this study deserves further careful study. Further research should be conducted in cooperation with a geophysicist. In addition, we plan to apply the directional analysis to areas other than south China and southern California to check the universality of the scaling behavior obtained in this study.

5. Considering the fractal nature of the earthquake process, it is recommended to perform further fractal and multifractal analyses on the earthquake wave records, which should contain more instantaneous information. Such information is expected to provide a more subtle profile of earthquake activity. In particular, some special features are expected to be extracted from this information that is recorded before strong events which will shed some light on short-time earthquake prediction.

Bibliography

Abe, S. and Suzuki, N. (2003). Law for the distance between successive earthquakes. *Journal of Geophysical Research*, 108:2113.

Alessio, E., Carbone, A., Castelli, G., and Frappietro, V. (2002). Second-order moving average and scaling of stochastic time series. *The European Physical Journal B*, 27(2):197 200.

Alvarez-Ramirez, J., Echeverria, J. C., and Rodriguez, E. (2008). Performance of a high-dimensional R/S method for hurst exponent estimation. *Physica A*, 387(26):6452 - 6462.

Alvarez-Ramirez, J., Rodriguez, E., and Echeverria, J. C. (2005). Detrending fluctuation analysis based on moving average filtering. *Physica A*, 354:199 219.

Anh, V., Yu, Z.-G., and Wanliss, J. A. (2007). Analysis of global geomagnetic variability. *Nonlinear Processes in Geophysics*, 14(6):701–708.

Anh, V. V., Yong, J. M., and Yu, Z. G. (2008). Stochastic modeling of the auroral electrojet index. *Journal of Geophysical Research*, 113(A10):A10215.

Anselin, L. (1988). *Spatial econometrics: methods and models*. Springer.

Anselin, L. (1990). Spatial dependence and spatial structural instability in applied regression analysis. *Journal of Regional Science*, 30(2):185–207.

Arianos, S. and Carbone, A. (2007). Detrending moving average algorithm: A closed-form approximation of the scaling law. *Physica A*, 382(1):9 - 15.

Ashkenazy, Y., Baker, D. R., Gildor, H., and Havlin, S. (2003). Nonlinearity and multifractality of climate change in the past 420,000 years. *Geophysical Research Letters*, 30:2146.

Ashkenazy, Y., Feliks, Y., Gildor, H., and Tziperman, E. (2008). Asymmetry of daily temperature records. *Journal of the Atmospheric Sciences*, 65(10):3327 3336.

Ashkenazy, Y., Ivanov, P., Havlin, S., Peng, C., Goldberger, A., and Stanley, H. (2001). Magnitude and sign correlations in heartbeat fluctuations. *Physical Review Letters*, 86(9):1900–1903.

Atkinson, P. M. and Tate, N. J. (2000). Spatial scale problems and geostatistical solutions: A review. *The Professional Geographer*, 52(4):607–623.

Bak, P., Christensen, K., Danon, L., and Scanlon, T. (2002). Unified scaling law for earthquakes. *Physical Review Letters*, 88(17):178501.

Balankin, A. S., Matamoros, D. M., Ortiz, J. P., Ortiz, M. P., Leon, E. P., and Ochoa, D. S. (2009). Scaling dynamics of seismic activity fluctuations. *EPL (Europhysics Letters)*, 85(3):39001 (6pp).

Balasco, M., Lapenna, V., and Telesca, L. (2002). $1/f^{\alpha}$ fluctuations in geoelectrical signals observed in a seismic area of southern italy. *Tectonophysics*, 347(4):253 268.

Barton, R. J. and Poor, H. V. (1988). Signal detection in fractional gaussian noise. *IEEE Transactions on Information Theory*, 34(5 Part 1):943 959.

Bartos, I. and Janosi, I. M. (2006). Nonlinear correlations of daily temperature records over land. *Nonlinear Processes in Geophysics*, 13(5):571 -576.

Bashan, A., Bartsch, R., Kantelhardt, J. W., and Havlin, S. (2008). Comparison of detrending methods for fluctuation analysis. *Physica A: Statistical Mechanics and its Applications*, 387(21):5080 5090.

Batty, M. (2008). The size, scale, and shape of cities. *Science*, 319(5864):769 771.

Batty, M. and Longley, P. (1994). *Fractal cities*. Academic press New York.

Bottiglieri, M., de Arcangelis, L., Godano, C., and Lippiello, E. (2010). Multiple-time scaling and universal behavior of the earthquake interevent time distribution. *Physical Review Letters*, 104(15):158501.

Buczkowski, S., Hildgen, P., and Cartilier, L. (1998). Measurements of fractal dimension by box-counting: a critical analysis of data scatter. *Physica A*, 252(1):23–34.

Buldyrev, S., Goldberger, A., Havlin, S., Peng, C., Stanley, H., Stanley, M., and Simons, M. (1993). Fractal landscapes and molecular evolution: Modeling the myosin heavy chain gene family. *Biophysical Journal*, 65(6):2673–2679.

Bunde, A., Kropp, J., and Schellnhuber, H.-J. (2002). *The Science of Disasters*. Springer, New York.

Cajueiro, D. O. and Tabak, B. M. (2008). Testing for long-range dependence in world stock markets. *Chaos, Solitons & Fractals*, 37(3):918 - 927.

Cane, H. V., Richardson, I. G., and von Rosenvinge, T. T. (1998). Interplanetary magnetic field periodicity of 153 days. *Geophysical Research Letters*, 25(24):4437–4440.

Carbone, A., Castelli, G., and Stanley, H. (2004a). Time-dependent hurst exponent in financial time series. *Physica A: Statistical Mechanics and its Applications*, 344(1-2):267 – 271.

Carbone, A., Castelli, G., and Stanley, H. E. (2004b). Analysis of clusters formed by the moving average of a long-range correlated time series. *Physical Review E*, 69(2):026105.

Carbone, V., Sorriso-Valvo, L., Harabaglia, P., and Guerra, I. (2005). Unified scaling law for waiting times between seismic events. *EPL (Europhysics Letters)*, 71(6):1036–1042.

Chamoli, A., Bansal, A. R., and Dimri, V. (2007). Wavelet and rescaled range approach for the hurst coefficient for short and long time series. *Computers & Geosciences*, 33(1):83 - 93.

Chen, Z., Ivanov, P. C., Hu, K., and Stanley, H. E. (2002). Effect of nonstationarities on detrended fluctuation analysis. *Physical Review E*, 65(4):041107.

Christensen, K., Danon, L., Scanlon, T., and Bak, P. (2002). Unified scaling law for earthquakes. *Proceedings of the National Academy of Sciences of the United States of America*, 99(Suppl 1):2509–2513.

Cola, L. D. (1991). Fractal analysis of multiscale spatial autocorrelation among point data. *Environment and Planning A*, 23:545–556.

Console, R. and Murru, M. (2001). A simple and testable model for earthquake clustering. *Journal of Geophysical Research*, 106(85):8699–8711.

Corral, A. (2004). Long-term clustering, scaling, and universality in the temporal occurrence of earthquakes. *Physical Review Letters*, 92(10):108501.

Corral, A. (2006). Universal earthquake-occurrence jumps, correlations with time, and anomalous diffusion. *Physical Review Letters*, 97(17):178501.

Currenti, G., Negro, C. D., Lapenna, V., and Telesca, L. (2005). Scaling characteristics of local geomagnetic field and seismicity at etna volcano and their dynamics in relation to the eruptive activity. *Earth and Planetary Science Letters*, 235(1-2):96 - 106.

Davidsen, J. and Goltz, C. (2004). Are seismic waiting time distributions universal? *Geophysical Research Letters*, 31:L21612.

Davidsen, J. and Paczuski, M. (2005). Analysis of the spatial distribution between successive earthquakes. *Physical Review Letters*, 94(4):048501.

Davidsen, J., Stanchits, S., and Dresen, G. (2007). Scaling and universality in rock fracture. *Physical Review Letters*, 98(12):125502.

Davis, A., Marshak, A., Wiscombe, W., and Cahalan, R. (1994). Multifractal characterizations of nonstationarity and intermittency in geophysical fields: Observed, retrieved, or simulated. *Journal of Geophysical Research*, 99(D4):8055–8072.

Ding, X., Zheng, D., and Yang, S. (2002). Variations of the surface temperature in hong kong during the last century. *International Journal of Climatology*, 22(6):715–730.

Doukhan, P., Oppenheim, G., and Taqqu, M. (2003). *Theory and applications of long-range dependence*. Birkhauser.

Eichner, J. F., Koscielny-Bunde, E., Bunde, A., Havlin, S., and Schellnhuber, H.-J. (2003). Power-law persistence and trends in the atmosphere: A detailed study of long temperature records. *Physical Review E*, 68(4):046133.

Emerson, C. W., Siu-Ngan Lam, N., and Quattrochi, D. A. (1999). Multi-scale fractal analysis of image texture and patterns. *Photogrammetric Engineering and Remote Sensing*, 65:51–62.

Enescu, B., Ito, K., Radulian, M., Popescu, E., and Bazacliu, O. (2005). Multifractal and chaotic analysis of vrancea (romania) intermediate-depth earthquakes: Investigation of the temporal distribution of events. *Pure and Applied Geophysics*, 162(2):249–271.

Fairbridge, R. W. and Hillaire-Marcel, C. (1977). An 8,000-yr palaeoclimatic record of the 'Double-Hale' 45-yr solar cycle. *Nature*, 268(5619):413–416.

Falconer, K. (1990). *Fractal geometry: Mathematical foundations and applications*. Wiley New York.

Fatichi, S., Barbosa, S. M., Caporali, E., and Silva, M. E. (2009). Deterministic versus stochastic trends: Detection and challenges. *Journal of Geophysical Research*, 114:D18121.

Feder, J. (1988). *Fractals*. Plenum Press New York.

Flandrin, P. (1992). Wavelet analysis and synthesis of fractional brownian motion. *IEEE Transactions on Information Theory*, 38(2 Part 2):910–917.

Fraedrich, K. and Blender, R. (2003). Scaling of atmosphere and ocean temperature correlations in observations and climate models. *Physical Review Letters*, 90(10):108501.

Fraedrich, K. and Larnder, C. (1993). Scaling regimes of composite rainfall time series. *Tellus*, 45(4):289–298.

Gao, J. and Xia, Z. (1996). Fractals in physical geography. *Progress in physical geography*, 20(2):178.

Gardner, J. K. and Knopoff, L. (1974). Is the sequence of earthquakes in southern california, with aftershocks removed, poissonian? *Bulletin of the Seismological Society of America*, 64(5):1363–1367.

Gasperini, P. and Mulargia, F. (1989). A statistical analysis of seismicity in Italy: The clustering properties. *Bulletin of the Seismological Society of America*, 79(4):973.

Gleissberg, W. (1944). A table of secular variations of the solar cycle. *Terrestrial Magnetism and Atmospheric Electricity*, 49(4):243–244.

Gnevyshev, M. N. (1977). Essential features of the 11-year solar cycle. *Solar Physics*, 51(1):175–183.

Godano, C., Tosi, P., Derubeis, V., and Angliera, P. (1999). Scaling properties of the spatio-temporal distribution of earthquakes: A multifractal approach applied to a Californian catalogue. *Geophysical Journal International*, 136(1):99-108.

Goltz, C. (1997). *Fractal and chaotic properties of earthquakes*, volume 77. Springer, Verlag Berlin Heidelberg.

Goodchild, M. and Mark, D. (1987). The fractal nature of geographic phenomena. *Annals of the Association of American Geographers*, 77(2):265–278.

Govindan, R. B., Vyushin, D., Bunde, A., Brenner, S., Havlin, S., and Schellnhuber, H.-J. (2002). Global climate models violate scaling of the observed atmospheric variability. *Physical Review Letters*, 89(2):028501.

Grassberger, P. and Procaccia, I. (1983). Measuring the strangeness of strange attractors. *Physica D: Nonlinear Phenomena*, 9(1-2):189–208.

Grigolini, P., Palatella, L., and Raffaelli, G. (2001). Asymmetric anomalous diffusion: An efficient way to detect memory in time series. *Fractals*, 9(4):439–449.

Gu, G.-F. and Zhou, W.-X. (2006). Detrended fluctuation analysis for fractals and multifractals in higher dimensions. *Physical Review E*, 74(6):061104.

Gutenberg, B. and Richter, C. (1949). *Seismicity of the Earth*. Princeton University, Princeton, NJ.

Halsey, T. C., Jensen, M. H., Kadanoff, L. P., Procaccia, I., and Shraiman, B. I. (1986). Fractal measures and their singularities: The characterization of strange sets. *Physical Review A*, 33(2):1141-1151.

Harte, D. (1998). Dimension estimates of earthquake epicentres and hypocentres. *Journal of Nonlinear Science*, 8(6):581–618.

Heneghan, C. and McDarby, G. (2000). Establishing the relation between detrended fluctuation analysis and power spectral density analysis for stochastic processes. *Physical Review E*, 62(5):6103–6110.

Hirabayashi, T., Ito, K., and Yoshii, T. (1992). Multifractal analysis of earthquakes. *Pure and Applied Geophysics*, 138(4):591-610.

Hirata, T. (1989). A correlation between the b value and the fractal dimension of earthquakes. *Journal of Geophysical Research*, 94(B6):7507–7514.

Hu, J., Gao, J., and Wang, X. (2009). Multifractal analysis of sunspot time series: the effects of the 11-year cycle and fourier truncation. *Journal of Statistical Mechanics*, 2009(02):P02066.

Hu, K., Ivanov, P. C., Chen, Z., Carpena, P., and Stanley, H. E. (2001). Effect of trends on detrended fluctuation analysis. *Physical Review E*, 64:011114.

Huang, B., Wu, B., and Barry, M. (2010). Geographically and temporally weighted regression for modeling spatio-temporal variation in house prices. *International Journal of Geographical Information Science*, 24(3):383–401.

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N.-C., Tung, C. C., and Liu, H. H. (1998). The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London. Series A*, 454(1971):903–995.

Huang, Q. (2008). Seismicity changes prior to the Ms8.0 Wenchuan earthquake in Sichuan, China. *Geophysical Research Letters*, 35:L23308.

Huang, Q., Sobolev, G. A., and Nagao, T. (2001). Characteristics of the seismic quiescence and activation patterns before the M=7.2 Kobe earthquake, January 17, 1995. *Tectonophysics*, 337(1-2):99–116.

Huang, Y. X., Schmitt, F. G., Lu, Z. M., and Liu, Y. L. (2008). An amplitude-frequency study of turbulent scaling intermittency using empirical mode decomposition and Hilbert spectral analysis. *EPL (Europhysics Letters)*, 84(4):40010.

Hurst, H. (1951). Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, 116:770–799.

Hurst, H. E. (1950). Long-term storage capacity of reservoirs. *American Society of Civil Engineering*, 76.

Jimenez, A., Tiampo, K. F., Levin, S., and Posadas, A. M. (2006). Testing the persistence in earthquake catalogs: The Iberian Peninsula. *EPL (Europhysics Letters)*, 73(2):171–177.

Kagan, Y. (1981). Spatial distribution of earthquakes: The three-point moment function. *Geophysical Journal International*, 67:697–717.

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

Kantelhardt, J. (2008). Fractal and multifractal time series. *Arxiv preprint arXiv:0804.0747*.

Kantelhardt, J. W., Koscielny-Bunde, E., Rybski, D., Braun, P., Bunde, A., and Havlin, S. (2006). Long-term persistence and multifractality of precipitation and river runoff records. *Journal of Geophysical Research*, 111:D01106.

Kantelhardt, J. W., Rybski, D., Zschiegner, S. A., Braun, P., Koscielny-Bunde, E., Livina, V., Havlin, S., and Bunde, A. (2003). Multifractality of river runoff and precipitation: Comparison of fluctuation analysis and wavelet methods. *Physica A*, 330(1-2):240–245.

Kantelhardt, J. W., Zschiegner, S. A., Koscielny-Bunde, E., Havlin, S., Bunde, A., and Stanley, H. E. (2002). Multifractal detrended fluctuation analysis of nonstationary time series. *Physica A*, 316(1-4):87–114.

Kimiagar, S., Movahed, M. S., Khorram, S., Sobhanian, S., and Tabar, M. R. R. (2009). Fractal analysis of discharge current fluctuations. *Journal of Statistical Mechanics*, 2009(03):P03020.

Koscielny-Bunde, E., Bunde, A., Havlin, S., Roman, H. E., Goldreich, Y., and Schellnhuber, H.-J. (1998). Indication of a universal persistence law governing atmospheric variability. *Physical Review Letters*, 81(3):729–732.

Koscielny-Bunde, E., Kantelhardt, J. W., Braun, P., Bunde, A., and Havlin, S. (2006). Long-term persistence and multifractality of river runoff records: Detrended fluctuation studies. *Journal of Hydrology*, 322(1-4):120–137.

Kurnaz, M. (2004). Application of detrended fluctuation analysis to monthly average of the maximum daily temperatures to resolve different climates. *Fractals*, 12(4):365–373.

Kyriacos, S., Buczkowski, S., Nekka, F., and Cartilier, L. (1994). A modified box-counting method. *Fractals*, 2(2):321–324.

Lam, N. and De Cola, L. (1993). *Fractals in geography*. PTR Prentice Hall.

Lavallee, D., Lovejoy, S., Schertzer, D., and Ladoy, P. (1993). *Fractals in Geography*, chapter Nonlinear variability of landscape topography: Multifractal analysis and simulation, pages 158–192. PTR Prentice Hall, Englewood Cliffs, NJ.

Lei, X. and Kusunose, K. (1999). Fractal structure and characteristic scale in the distributions of earthquake epicentres, active faults and rivers in Japan. *Geophysical Journal International*, 139(3):754–762.

Lennartz, S. and Bunde, A. (2009). Trend evaluation in records with long-term memory: Application to global warming. *Geophysical Research Letters*, 36:L16706.

Lennartz, S., Livina, V. N., Bunde, A., and Havlin, S. (2008). Long-term memory in earthquakes and the distribution of interoccurrence times. *EPL (Europhysics Letters)*, 81(6):69001.

Leung, Y. (2010). *Knowledge Discovery in Spatial Data*. Springer Verlag.

Leung, Y., Mei, C., and Zhang, W. (2000). Statistical tests for spatial nonstationarity based on the geographically weighted regression model. *Environment and Planning A*, 32(1):9–32.

Li, J., Chen, Y., and Mi, H. (2002). $1/f^{\beta}$ temporal fluctuation: detecting scale-invariance properties of seismic activity in north china. *Chaos, Solitons & Fractals*, 14(9):1487 - 1494.

Li, Q., Wu, J., Xu, Z.-W., and Wu, J. (2007). Extraction of the periodic components of solar activity with the emd method. *Chinese Astronomy and Astrophysics*, 31(3):261 - 269.

Lilley, M., Lovejoy, S., Desaulniers-Soucy, N., and Schertzer, D. (2006). Multifractal large number of drops limit in rain. *Journal of Hydrology*, 328(1-2):20 37.

Lindman, M., Jonsdottir, K., Roberts, R., Lund, B., and Bödvarsson, R. (2005). Earthquakes descaled: On waiting time distributions and scaling laws. *Physical Review Letters*, 94(10):108501.

Liou, K.-N. (2002). *An introduction to atmospheric radiation*. Academic Press, second edition.

Livina, V. N., Havlin, S., and Bunde, A. (2005). Memory in the occurrence of earthquakes. *Physical Review Letters*, 95(20):208501.

Lloyd, C. (2007). *Local models for spatial analysis*. CRC Press.

Lovejoy, S., Schertzer, D., and Allaire, V. (2008). The remarkable wide range spatial scaling of trmm precipitation. *Atmospheric Research*, 90(1):10 -32.

Lovejoy, S., Schertzer, D., Tessier, Y., and Gaonac'h, H. (2001). Multifractals and resolution-independent remote sensing algorithms: the example of ocean colour. *International Journal of Remote Sensing*, 22(7):1191- 1234.

Ma, Q. D. Y., Bartsch, R. P., Bernaola-Galván, P., Yoneyama, M., and Ivanov, P. C. (2010). Effect of extreme data loss on long-range correlated and anticorrelated signals quantified by detrended fluctuation analysis. *Physical Review E*, 81(3):031101.

Makse, H. A., Havlin, S., Schwartz, M., and Stanley, H. E. (1996). Method for generating long-range correlations for large systems. *Physical Review E*, 53(5):5445 -5449.

Mandelbrot, B. (1982). *The fractal geometry of nature*. W. H. Freeman, New York.

Mandelbrot, B. B. and Ness, J. W. V. (1968). Fractional brownian motions, fractional noises and applications. *SIAM Review*, 10(4):422 437.

Marghany, M., Cracknell, A. P., and Hashim, M. (2009). Comparison between radarsat-1 SAR - different data modes for oil spill detection by a fractal box counting algorithm. *International Journal of Digital Earth*, 2(3):237- 256.

Marsan, D., Bean, C. J., Steacy, S., and McCloskey, J. (2000). Observation of diffusion processes in earthquake populations and implications for the predictability of seismicity systems. *Journal of Geophysical Research*, 105:28081–28094.

Matsoukas, C., Islam, S., and Rodriguez-Iturbe, I. (2000). Detrended fluctuation analysis of rainfall and streamflow time series. *Journal of Geophysical Research*, 105:29165–29172.

McGaughey, D. R. and Aitken, G. J. M. (2002). Generating two-dimensional fractional brownian motion using the fractional gaussian process (FGp) algorithm. *Physica A*, 311(3-4):369 – 380.

Mega, M. S., Allegrini, P., Grigolini, P., Latora, V., Palatella, L., Rapisarda, A., and Vinciguerra, S. (2003). Power-law time distribution of large earthquakes. *Physical Review Letters*, 90(18):188501.

Molchan, G. and Kronrod, T. (2005). On the spatial scaling of seismicity rate. *Geophysical Journal International*, 162(3):899 909.

Moussas, X., Polygiannakis, J., Preka-Papadema, P., and Exarhos, G. (2005). Solar cycles: A tutorial. *Advances in Space Research*, 35(5):725 738.

Movahed, M. S. and Hermanis, E. (2008). Fractal analysis of river flow fluctuations. *Physica A*, 387(4):915 - 932.

Movahed, M. S., Jafari, G. R., Ghasemi, F., Rahvar, S., and Tabar, M. R. R. (2006). Multifractal detrended fluctuation analysis of sunspot time series. *Journal of Statistical Mechanics*, 2006(02):P02003.

Muniandy, S. V., Lim, S. C., and Murugan, R. (2001). Inhomogeneous scaling behaviors in malaysian foreign currency exchange rates. *Physica A*, 301(1-4):407 - 428.

Muzy, J., Bacry, E., and Arneodo, A. (1994). The multifractal formalism revisited with wavelets. *International Journal of Bifurcation and Chaos in Applied Sciences and Engineering*, 4(2):245-302.

Muzy, J. F., Bacry, E., and Arneodo, A. (1991). Wavelets and multifractal formalism for singular signals: Application to turbulence data. *Physical Review Letters*, 67(25):3515 3518.

Myint, S. W. (2003). Fractal approaches in texture analysis and classification of remotely sensed data: comparisons with spatial autocorrelation techniques and simple descriptive statistics. *International Journal of Remote Sensing*, 24(9):1925 1947.

Nagarajan, R. and Kavasseri, R. (2004). Minimizing the effect of sinusoidal trends in detrended fluctuation analysis. *Arxiv preprint cond-mat/0411543*.

Nagarajan, R. and Kavasseri, R. G. (2005). Minimizing the effect of periodic and quasi-periodic trends in detrended fluctuation analysis. *Chaos, Solitons & Fractals*, 26(3):777 – 784.

Nagovitsyn, Y. (1997). A nonlinear mathematical model for the solar cyclicity and prospects for reconstructing the solar activity in the past. *Astronomy Letters*, 23(6):742 748.

Naujokat, B. (1986). An update of the observed quasi-biennial oscillation of the stratospheric winds over the tropics. *Journal of the Atmospheric Sciences*, 43(17):1873–1877.

Neuman, S. (2010). Apparent/spurious multifractality of data sampled from fractional brownian/lévy motions. *Hydrological Processes*, 24:2056 2067.

Olsson, J. (1995). Limits and characteristics of the multifractal behaviour of a high-resolution rainfall time series. *Nonlinear Processes in Geophysics*, 2(1):23-29.

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

Orm, M. and Kocak, K. (2009). Application of detrended fluctuation analysis to temperature data from turkey. *International Journal of Climatology*, 29(14):2130-2136.

Oswiecimka, P., Kwapien, J., and Drozdz, S. (2006). Wavelet versus detrended fluctuation analysis of multifractal structures. *Physical Review E*, 74(1):016103.

Pachepsky, Y. and Ritchie, J. (1998). Seasonal changes in fractal landscape surface roughness estimated from airborne laser altimetry data. *International Journal of Remote Sensing*, 19(13):2509 2516.

Pandey, G., Lovejoy, S., and Schertzer, D. (1998). Multifractal analysis of daily river flows including extremes for basins of five to two million square kilometres, one day to 75 years. *Journal of Hydrology*, 208(1-2):62 – 81.

Parrinello, T. and Vaughan, R. A. (2002). Multifractal analysis and feature extraction in satellite imagery. *International Journal of Remote Sensing*, 23(9):1799 1825.

Pattantyús-Ábrahám, M., Király, A., and Jánosi, I. (2004). Nonuniversal atmospheric persistence: Different scaling of daily minimum and maximum temperatures. *Physical Review E*, 69(2):021110.

Pelletier, J. D. (1997). Analysis and modeling of the natural variability of climate. *Journal of Climate*, 10(6):1331-1342.

Peng, C.-K., Buldyrev, S. V., Goldberger, A. L., Havlin, S., Sciortino, F., Simons, M., and Stanley, H. E. (1992). Long-range correlations in nucleotide sequences. *Nature*, 356(6365):168–170.

Peng, C.-K., Buldyrev, S. V., Havlin, S., Simons, M., Stanley, H. E., and Goldberger, A. L. (1994). Mosaic organization of dna nucleotides. *Physical Review E*, 49(2):1685-1689.

Peters, O., Hertlein, C., and Christensen, K. (2001). A complexity view of rainfall. *Physical Review Letters*, 88(1):018701.

Qiu, H. L., Lam, N. S., Quattrochi, D., and Gamon, J. (1999). Fractal characterization of hyperspectral imagery. *Photogrammetric engineering and remote sensing*, 65(1):63–71.

Quattrochi, D., Lam, N. S.-N., Qiu, H., and Zhao, W. (1997). *Scale in Remote Sensing and GIS*, chapter Image Characterization and Modeling System (ICAMS): A geographic information system for the characterization and modeling of multiscale remote sensing data, pages 295–308. CRC Press LLC.

Rangarajan, G. and Ding, M. (2000). Integrated approach to the assessment of long range correlation in time series data. *Physical Review E*, 61(5):4991–5001.

Raoufi, D. (2010). Fractal analyses of ito thin films: A study based on power spectral density. *Physica B: Condensed Matter*, 405(1):451 – 455.

Richards, A. (2002). Complexity in physical geography. *Geography*, 87(2):99–107.

Richardson, J. D., Paularena, K. I., Belcher, J. W., and Lazarus, A. J. (1994). Solar wind oscillations with a 1.3 year period. *Geophysical Research Letters*, 21(14):1559–1560.

Rieger, E., Share, G. H., Forrest, D. J., Kanbach, G., Reppin, C., and Chupp, E. L. (1984). A 154-day periodicity in the occurrence of hard solar flares? *Nature*, 312(5995):623–625.

Rilling, G., Flandrin, P., and Goncalves, P. (2003). On empirical mode decomposition and its algorithms. In *IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing NSIP*, volume 3, pages 8–11.

Rybski, D., Bunde, A., Havlin, S., and von Storch, H. (2006). Long-term persistence in climate and the detection problem. *Geophysical Research Letters*, 33(6):L06718.

Rybski, D., Bunde, A., and von Storch, H. (2008). Long-term memory in 1000-year simulated temperature records. *Journal of Geophysical Research*, 113(D2):D02106.

Sadovskiy, M. A., Golubeva, T. V., Pisarenko, V. F., and Shnirman, M. G. (1984). Characteristic rock dimensions and hierarchical properties of seismicity. *USSR Report Earth Sciences JPRS UES*, page 65.

Saichev, A. and Sornette, D. (2007). Theory of earthquake recurrence times. *Journal of Geophysical Research*, 112:B04313.

Schatten, K. H. and Sofia, S. (1987). Forecast of an exceptionally large even-numbered solar cycle. *Geophysical Research Letters*, 14(6):632–635.

Schertzer, D. and Lovejoy, S. (1987). Physical modeling and analysis of rain and clouds by anisotropic scaling multiplicative processes. *Journal of Geophysical Research*, 92:9693–9714.

Serletis, D. (2008). Effect of noise on fractal structure. *Chaos, Solitons & Fractals*, 38(4):921 – 924.

Seuront, L., Schmitt, F., Lagadeuc, Y., Schertzer, D., and Lovejoy, S. (1999). Universal multifractal analysis as a tool to characterize multiscale intermittent patterns: example of phytoplankton distribution in turbulent coastal waters. *Journal of Plankton Research*, 21(5):877–922.

Shcherbakov, R., Yakovlev, G., Turcotte, D. L., and Rundle, J. B. (2005). Model for the distribution of aftershock interoccurrence times. *Physical Review Letters*, 95(21):218501.

Shiogai, Y., Stefanovska, A., and McClintock, P. (2010). Nonlinear dynamics of cardiovascular ageing. *Physics Reports*, 488(2-3):51 – 110.

Shlien, S. and Nafi Toksoz, M. (1970). A clustering model for earthquake occurrences. *Bulletin of the Seismological Society of America*, 60(6):1765 –1787.

Sinclair, S. and Pegram, G. G. S. (2005). Empirical mode decomposition in 2-D space and time: a tool for space-time rainfall analysis and nowcasting. *Hydrology and Earth System Sciences*, 9(3):127 –137.

Stein, S. and Wysession, M. (2003). *An introduction to seismology, earthquakes, and earth structure*. Wiley-Blackwell.

Takayasu, H. (1990). *Fractals in the physical sciences*. Manchester University Press, Manchester, New York.

Talkner, P. and Weber, R. O. (2000). Power spectrum and detrended fluctuation analysis: Application to daily temperatures. *Physical Review E*, 62(1):150 –160.

Taqqu, M. S., Teverovsky, V., and Willinger, W. (1995). Estimators for long-range dependence: an empirical study. *Fractals*, 3:785 –798.

Tarafdar, S. and Harper, D. (2008). Anti-persistence in levels of lake naivasha: Assessing effect of human intervention through time-series analysis. *Physica A*, 387(1):296 – 302.

Telesca, L., Colangelo, G., Lapenna, V., and Macchiato, M. (2004a). Fluctuation dynamics in geoelectrical data: an investigation by using multifractal detrended fluctuation analysis. *Physics Letters A*, 332(5-6):398 – 404.

Telesca, L., Cuomo, V., Lapenna, V., and Macchiato, M. (2001a). Depth-dependent time-clustering behaviour in seismicity of Southern California. *Geophysical Research Letters*, 28:4323–4326.

Telesca, L., Cuomo, V., Lapenna, V., and Macchiato, M. (2001b). A new approach to investigate the correlation between geoelectrical time fluctuations and earthquakes in a seismic area of Southern Italy. *Geophysical Research Letters*, 28(23):4375 – 4378.

Telesca, L., Cuomo, V., Lapenna, V., and Macchiato, M. (2004b). Detrended fluctuation analysis of the spatial variability of the temporal distribution of Southern California seismicity. *Chaos, Solitons & Fractals*, 21(2):335 – 342.

Telesca, L., Lapenna, V., Lovallo, M., and Macchiato, M. (2004c). Long-range time-correlation properties of seismic sequences. *Chaos, Solitons & Fractals*, 21(2):387 – 393.

Telesca, L., Lapenna, V., and Macchiato, M. (2003). Spatial variability of the time-correlated behaviour in Italian seismicity. *Earth and Planetary Science Letters*, 212(3-4):279 – 290.

Telesca, L., Lapenna, V., and Macchiato, M. (2005). Multifractal fluctuations in earthquake-related geoelectrical signals. *New Journal of Physics*, 7:214.

Telesca, L., Lapenna, V., Macchiato, M., and Hattori, K. (2008). Investigating non-uniform scaling behavior in ultra low frequency (ULF) earthquake-related geomagnetic signals. *Earth and Planetary Science Letters*, 268(1-2):219 - 224.

Tessier, Y., Lovejoy, S., Hubert, P., Schertzer, D., and Pecknold, S. (1996). Multifractal analysis and modeling of rainfall and river flows and scaling, causal transfer functions. *Journal of Geophysical Research*, 101(D21):26427–26440.

Tobler, W. R. (1970). A computer movie simulating urban growth in the detroit region. *Economic Geography*, 46:234–240.

Tsonis, A. A., Roebber, P. J., and Elsner, J. B. (1998). A characteristic time scale in the global temperature record. *Geophysical Research Letters*, 25(15):2821 2823.

Tsonis, A. A., Roebber, P. J., and Elsner, J. B. (1999). Long-range correlations in the extratropical atmospheric circulation: Origins and implications. *Journal of Climate*, 12(5):1534 1541.

Turcotte, D. (1997). *Fractals and chaos in geology and geophysics*. Cambridge University Press.

Tziperman, E., Cane, M. A., and Zebiak, S. E. (1995). Irregularity and locking to the seasonal cycle in an ENSO prediction model as explained by the quasi-periodicity route to chaos. *Journal of the Atmospheric Sciences*, 52(3):293–306.

Varotsos, P. A., Sarlis, N. V., and Skordas, E. S. (2002). Long-range correlations in the electric signals that precede rupture. *Physical Review E*, 66(1):011902.

Varotsos, P. A., Sarlis, N. V., and Skordas, E. S. (2003). Electric fields that "arrive" before the time derivative of the magnetic field prior to major earthquakes. *Physical Review Letters*, 91(14):148501.

Vere-Jones, D. (1999). On the fractal dimensions of point patterns. *Advances in Applied Probability*, 31(3):643 663.

Wang, S.-C., Li, P.-C., and Tseng, H.-C. (2008). Long range correlation and possible electron conduction through DNA sequences. *Physica A*, 387(21):5159 - 5168.

Weber, R. O. and Talkner, P. (2001). Spectra and correlations of climate data from days to decades. *Journal of Geophysical Research*, 106(D17):20131–20144.

Wilson, R. M. (1988). Bimodality and the hale cycle. *Solar Physics*, 117(2):269–278.

Wu, Z., Huang, N. E., Long, S. R., and Peng, C.-K. (2007). On the trend, detrending, and variability of nonlinear and nonstationary time series. *Proceedings of the National Academy of Sciences of the United States of America*, 104(38):14889–14894.

Xu, L., Ivanov, P. C., Hu, K., Chen, Z., Carbone, A., and Stanley, H. E. (2005). Quantifying signals with power-law correlations: A comparative study of detrended fluctuation analysis and detrended moving average techniques. *Physical Review E*, 71(5):051101.

Xu, T., Wu, J., Wu, Z.-S., and Li, Q. (2008). Long-term sunspot number prediction based on emd analysis and ar model. *Chinese Journal of Astronomy and Astrophysics*, 8(3):337 -342.

Ying, W.-L., Benson, P. M., and Young, R. P. (2009). Laboratory simulation of fluid-driven seismic sequences in shallow crustal conditions. *Geophysical Research Letters*, 36(20):L20301.

Yu, Z.-G., Anh, V., and Eastes, R. (2009). Multifractal analysis of geomagnetic storm and solar flare indices and their class dependence. *Journal of Geophysical Research*, 114:A05214.

Yu, Z.-G., Leung, Y., Chen, Y. D., Zhang, Q., and Anh, V. (2011). Multifractal analyses of daily rainfall in the pearl river basin of China. *Journal of Hydrology*, submitted.

Zebiak, S. E. and Cane, M. A. (1987). A model El Nino-Southern Oscillation. *Monthly Weather Review*, 115(10):2262–2278.

Zhang, L. and Singh, V. P. (2007). Bivariate rainfall frequency distributions using archimedean copulas. *Journal of Hydrology*, 332(1-2):93 - 109.

Zhang, Q., Xu, C.-Y., Chen, Y. D., and Yu, Z.-G. (2008). Multifractal detrended fluctuation analysis of streamflow series of the yangtze river basin, China. *Hydrological Processes*, 22(26):4997–5003.

Zhang, Q., Xu, C.-Y., Yu, Z.-G., Liu, C.-L., and Chen, Y. D. (2009). Multifractal analysis of streamflow records of the east river basin (pearl river), China. *Physica A*, 388(6):927 - 934.

Zhou, Y. and Leung, Y. (2010a). Empirical mode decomposition and long-range correlation analysis of sunspot time series. *Journal of Statistical Mechanics*, 2010(12):P12006.

Zhou, Y. and Leung, Y. (2010b). Multifractal temporally weighted detrended fluctuation analysis and its application in the analysis of scaling behavior in temperature series. *Journal of Statistical Mechanics*, 2010(06):P06021.

Zhou, Y., Leung, Y., and Yu, Z.-G. (2011). A note on the relationships of exponents in multifractal detrended fluctuation analysis and conventional multifractal analysis. *Chinese Physics B*, 20(9):090507.