

2020-01-01

## Self-Similar Models: How close the diffusion entropy analysis and the detrended fluctuation analysis are from other models

William Kubin  
*University of Texas at El Paso*

Follow this and additional works at: [https://scholarworks.utep.edu/open\\_etd](https://scholarworks.utep.edu/open_etd)



Part of the [Mathematics Commons](#)

---

### Recommended Citation

Kubin, William, "Self-Similar Models: How close the diffusion entropy analysis and the detrended fluctuation analysis are from other models" (2020). *Open Access Theses & Dissertations*. 2990.  
[https://scholarworks.utep.edu/open\\_etd/2990](https://scholarworks.utep.edu/open_etd/2990)

This is brought to you for free and open access by ScholarWorks@UTEP. It has been accepted for inclusion in Open Access Theses & Dissertations by an authorized administrator of ScholarWorks@UTEP. For more information, please contact [lweber@utep.edu](mailto:lweber@utep.edu).

SELF-SIMILAR MODELS: HOW CLOSE THE DIFFUSION ENTROPY ANALYSIS  
AND THE DETRENDED FLUCTUATION ANALYSIS  
ARE FROM OTHER MODELS

WILLIAM KUBIN

Master's Program in Computational Science

APPROVED:

---

Maria C. Mariani, Ph.D., Chair

---

Thompson Sarkodie-Gyan, Ph.D.

---

Granville Sewell, Ph.D.

---

Elsa Villa, Ph.D.

---

Stephen Crites, Ph.D.  
Dean of the Graduate School

©Copyright

by

William Kubin

2020

*to my*

*FATHER and MOTHER*

*with love*

SELF-SIMILAR MODELS: HOW CLOSE THE DIFFUSION ENTROPY ANALYSIS  
AND THE DETRENDED FLUCTUATION ANALYSIS  
ARE FROM OTHER MODELS

by

WILLIAM KUBIN

THESIS

Presented to the Faculty of the Graduate School of  
The University of Texas at El Paso  
in Partial Fulfillment  
of the Requirements  
for the Degree of

MASTER OF SCIENCE

Computational Science Program

THE UNIVERSITY OF TEXAS AT EL PASO

May 2020

# Acknowledgements

I would like to pay my special regards to my parents, whom I am greatly indebted for my upbringing, support and encouragement to this stage. I also feel obliged to extend my sincere thanks to my advisor, Dr. Maria C. Mariani (Chair, The Math Department at The University of Texas at El Paso), for her advice, encouragement, enduring patience and constant support. Next, I extend special thanks to the members of my dissertation committee Dr. Thompson Sarkodie-Gyan, Dr. Granville Sewell and Dr. Elsa Villa for making the time out of your busy schedules to be part of this project.

Additionally, I want to thank The University of Texas at El Paso Computational Science Department and Mathematics Department professors and staff for all their hard work and dedication, providing me the means to complete my degree and prepare for a career as a data scientist.

NOTE: This thesis was submitted to my Supervising Committee on the May 31, 2020.

# Abstract

Financial and seismic data, like many other high frequency data are known to exhibit memory effects. In this research, we apply the concepts of Lévy processes, Diffusion Entropy Analysis (DEA) and the Detrended Fluctuation Analysis (DFA) to examine long-range persistence (long memory) behavior in time series data. Lévy processes describe long memory effects. In other words, Lévy process (where the increments are independent and follow the Lévy distribution) is self-similar. We examine the relationship between the Lévy parameter  $\alpha$  characterizing the data and the scaling exponent of DEA ( $\delta$ ) and that of DFA ( $H$ ) characterizing the self-similar property of the respective models. We investigate how close this model is to a self-similar model and prove the numerical relationship.

# Table of Contents

	<b>Page</b>
Acknowledgements . . . . .	v
Abstract . . . . .	vi
Table of Contents . . . . .	vii
<b>Chapter</b>	
1 Introduction . . . . .	1
1.1 Our Result . . . . .	1
2 Literature Review . . . . .	2
2.1 Lévy Flight Model . . . . .	3
2.2 Detrended Fluctuation Analysis (DFA) . . . . .	4
2.3 Diffusion Entropy Analysis (DEA) . . . . .	4
3 Self - Similarity . . . . .	5
3.1 Definition . . . . .	5
4 Lévy Flight Model . . . . .	6
5 Scaling Methods . . . . .	9
5.1 Detrended Fluctuation Analysis (DFA) . . . . .	9
5.2 Diffusion Entropy Analysis (DEA) . . . . .	10
6 Data . . . . .	14
6.1 Financial Market Data . . . . .	14
6.2 Geophysical Time Series Data . . . . .	14
6.3 Stationarity Test of Data . . . . .	15
6.3.1 Financial Market Data . . . . .	15
6.3.2 Geophysical Time Series Data . . . . .	18
6.3.3 Augmented Dickey-Fuller (ADF) . . . . .	20
7 Results . . . . .	22



7.1	Numerical Results . . . . .	22
7.1.1	Financial Market Data . . . . .	23
7.1.2	Geophysical Time Series Data . . . . .	24
7.1.3	Remarks on the Table of Numerical Results . . . . .	24
7.2	Analytical Relations . . . . .	25
7.2.1	Detrended Fluctuation Analysis (DFA) . . . . .	25
7.2.2	Diffusion Entropy Analysis (DEA) . . . . .	26
8	Figures . . . . .	28
8.1	Financial Market Data . . . . .	28
8.1.1	Lévy Flight Model . . . . .	28
8.1.2	Detrended Fluctuation Analysis (DFA) . . . . .	31
8.1.3	Diffusion Entropy Analysis (DEA) . . . . .	34
8.2	Geophysical Time Series Data . . . . .	37
8.2.1	Lévy Flight Model . . . . .	37
8.2.2	Detrended Fluctuation Analysis (DFA) . . . . .	39
8.2.3	Diffusion Entropy Analysis (DEA) . . . . .	41
9	Concluding Remarks . . . . .	43
9.1	Significance of the Result . . . . .	43
9.2	Future Work . . . . .	43
	References . . . . .	48
	Curriculum Vitae . . . . .	52

# Chapter 1

## Introduction

The detrended fluctuation analysis (DFA) invented by Peng et al. (1993) and the diffusion entropy analysis (DEA) are very important methods for detecting long-memory behavior in time series. They have been applied to diverse fields of interest including economic time series, weather recordings, cloud structure, DNA, traffic analysis, bioengineering etc. [1]. The DFA, similar to the Hurst rescaled-range analysis, is based on the random walk theory [2, 3]. It permits the detection of intrinsic self-similarity embedded in a seemingly non-stationary time series. The DEA, on the other hand, can be used to determine if the characterization of a time series follows a Gaussian or Lévy distribution, as well as establish the existence of long-range correlations in the time series. We also study truncated Lévy flight distributions and self-similarity. For the set of financial market data and data from recordings of volcanic eruptions used in this paper, we compare and quantify the scaling and long-range persistence in time series using DFA and DEA. From the resulting scaling parameters, we conclude that the time series exhibits long-range correlations (or long-memory behavior) if  $0.5 < H, \delta < 1$ , and anti-persistent if  $0 < H, \delta < 0.5$ .

### 1.1 Our Result

We will show that a relationship exists between the Lévy parameter  $\alpha$  characterizing the data and the resulting  $H$  or  $\delta$  parameters of the DFA and DEA characterizing the self-similar property respectively. Finally, we will show analytically the relations we obtain from numerical simulations between the Lévy model, DFA and the DEA.

# Chapter 2

## Literature Review

### **Self-Similar Models: How close the Diffusion Entropy Analysis and the Detrended Fluctuation Analysis are from other models: A Literature Review**

In the 1960s, Benoit B. Mandelbrot described the phenomenon of self-similarity, beginning with his article, "How Long is the Coast of Britain? Statistical Self-Similarity and Fractional Dimension," published in *Science* (1967). In this article he introduced fractals as part of the solution to a problem that had occupied his attention for some time: How to measure a curve as complex as a geographic coastline? Geographical curves are so involved in their detail that their lengths are often infinite or, rather, undefinable. However, many are statistically "self-similar," meaning that each portion can be considered a reduced-scale image of the whole.

He discussed two salient characteristics of fractals that applied to this problem: self-similarity and "fractional" dimensionality.

Self-similarity referred to the persistence of patterns as an observer zoomed in or out of the visualization of a fractal set. Fractional geometry described the quality these sets had mathematically of being "fuzzier" than a line but never completely filling a plane.

In 1975, Mandelbrot coined the term "fractal" to describe such mathematical sets. Over the course of time, scholars were encouraged to investigate the application of fractal geometry to fields ranging from engineering and medicine to finance, climate study, and art.

This study is not different as we apply fractals and self-similarity to data from financial markets and volcanic eruptions.

Below is an overview of scholarly materials on the Lévy flight model and the scaling methods, Detrended Fluctuation Analysis (DFA) and Diffusion Entropy Analysis (DEA).

## 2.1 Lévy Flight Model

The Lévy flight is a random walk in which the step-lengths have a probability distribution that is heavy-tailed. Today, researchers have extended the use of the term "Lévy flight" to include cases where the random walk takes place on a discrete grid rather than on a continuous space. [14][15] Meanwhile, the particular case for which Mandelbrot (1982, p. 294) used the term "Lévy flight" is defined by the survival function of the distribution (particularly Pareto distribution) of step-sizes. For general distributions of step-sizes, satisfying the power-like condition, the distance from the origin of the random walk approximates to a stable distribution after a large number of steps. This enables many processes to be modeled using Lévy flights and this is particularly as a result to the generalized central limit theorem.

The definition of a Lévy flight stems from chaos theory and is useful in stochastic measurement and simulations for random or pseudo-random natural phenomena. Researchers, after analyzing over 12 million movements recorded from 14 ocean predator species in the Atlantic and Pacific Oceans, discovered that when sharks and other ocean predators cannot find food, they abandon Brownian motion, the random motion seen in swirling gas molecules, for Lévy flight - a mix of long trajectories and short, random movements found in turbulent fluids. The data showed that Lévy flights interspersed with Brownian motion can describe the animals' hunting patterns [16][17][17][18][19]. Also, other animals (including birds and humans) when searching for food follows paths that have been modeled using Lévy flight [20][21][22].

Lévy Flight model, since then, has been applied to several fields including earthquake data analysis, financial mathematics, cryptography, signals analysis as well as many applications in astronomy, biology, and physics.

## 2.2 Detrended Fluctuation Analysis (DFA)

Detrended Fluctuation Analysis (DFA) is a method for determining the statistical self-similarity of a signal in stochastic processes, chaos theory and time series. It comes in handy when analysing time series that exhibit long-memory behavior. DFA is related to estimates based on spectral techniques like autocorrelation and Fourier transform.

To represent an extension of the fluctuation analysis (FA) which is affected by non-stationarities, Peng et al (1994) introduced DFA in a paper. Except the fact that DFA may be applicable to signals whose underlying statistics are non-stationary (changing with time), the scaling exponent obtained is similar to the Hurst exponent.

Researchers have applied the DFA to numerous systems including DNA sequences, [23][24] neural oscillations, [25] speech pathology detection, [26] and heartbeat fluctuation in different sleep stages.[27]

## 2.3 Diffusion Entropy Analysis (DEA)

Diffusion Entropy Analysis (DEA) is a method based on the numerical evaluation of variance for determining the scaling exponent of a complex dynamic process described by a time series. It focuses on the scaling exponent evaluated through the Shannon Entropy of the diffusion generated by the fluctuations of time series. N. Scafetta in 2003, argued the DEA is the only method that correctly quantifies the scaling exponent of the time series of a complex process.[28]

### Discussion and Conclusion

These methods of scaling have been used to perform analysis on time series data from several fields throughout the years. For this reason, we believe they produce consistent scaling parameters in the analysis of the financial market data and geophysical (volcanic) time series data.

# Chapter 3

## Self - Similarity

In this chapter, we provide a brief introduction to the theory of self-similarity and provide the definition necessary to accurately describe what we mean when we say that a model is self-similar.

In mathematics, a self-similar object is exactly or approximately similar to a part of itself. For instance, coastlines, snow flakes and several objects in the real world are statistically self-similar: parts of them show the same statistical properties at many scales. Self-similarity is a typical property of fractals. Fractals exhibit similar patterns at increasingly small scales.

Self-similar processes are types of stochastic processes that exhibit the phenomenon of self-similarity. They can exhibit long-range dependency, also called long-memory behavior. A phenomenon is usually considered to have long-range dependence if the dependence decays more slowly than an exponential decay, typically a power-like decay.

### 3.1 Definition

*A stochastic process  $X(t)$  is said to be self-similar if there exists a constant  $H > 0$  such that for any scaling factor  $a > 0$ , the process  $\{X_{at}\}_{t \geq 0}$  and  $\{a^H X_t\}_{t \geq 0}$  have the same law in the sense of finite dimensional distributions. The constant  $H$  is called the self-similarity exponent of the process  $X$ .*

# Chapter 4

## Lévy Flight Model

Lévy and Khintchine [4] solved the problem of determining the functional form that all the stable distribution must follow. They found that the most general representation is through the characteristic functions  $\varphi(q)$ , that are defined by the following equation:

$$\ln(\varphi(q)) = i\mu q - \gamma |q|^\alpha \left[ 1 - i\beta \frac{q}{|q|} \tan\left(\frac{\pi\alpha}{2}\right) \right] \quad (4.1)$$

if  $\alpha \neq 1$ , and

$$\ln(\varphi(q)) = \mu q - \gamma |q| \left[ 1 + i\beta \frac{q}{|q|} \frac{2}{\pi} \log(q) \right] \quad (4.2)$$

if  $\alpha = 1$ ,

where  $0 < \alpha \leq 2$  (i.e., the same parameter mentioned before) is called the scaling exponent;  $\gamma$  is a positive scale factor,  $\mu$  is a real number, called the location parameter and  $\beta$  is an asymmetry parameter ranging from -1 to 1, that is called the skewness parameter. Also, the  $\mu$  is observed as the mean and that  $\gamma^2$  will coincide with the variance  $\sigma^2$ .

The analytical form of the Lévy stable distribution is known only for a few values of  $\alpha$  and  $\beta$ .

We consider the symmetric distribution ( $\beta = 0$ ) with a zero mean ( $\mu = 0$ ). In this case the characteristic function takes the form:

$$\varphi(q) = \exp(-\gamma |q|^\alpha) \quad (4.3)$$

Since the characteristic function of a distribution is Fourier transform, the stable distribution of index  $\alpha$  and scale factor  $\gamma$  is

$$P_L(x) = \frac{1}{\pi} \int_0^\infty \exp(-\gamma |q|^\alpha) \cos(qx) dq \quad (4.4)$$

The stable Lévy processes (Lévy flight) have independent increments but are designated as long memory processes [5]. In order to avoid the problems arising in the infinite second moment (i.e. the fact that stable Lévy processes with  $\alpha < 2$  have infinite variance), we consider a stochastic process with finite variance that follows scale relations called Truncated Lévy Flight (TLF) (Mantegna and Stanley (1994)).

The TLF distribution is defined by

$$T(x) = cP(x)\chi_{(-l,l)}(x) \tag{4.5}$$

with  $P(x)$  a symmetric Lévy distribution. The TLF distribution is not stable, but it has finite variance. However, convergence may be slow depending on the size of the cut off length parameter  $l$ . (Mantegna and Stanley (1994)). If the parameter  $l$  is small (so that the convergence is fast) the cut that it presents in its tails is very abrupt. In order to have continuous tails, the author in Koponen (1995) considered a TLF in which the cut function is a decreasing exponential characterized by a parameter  $l$ . The characteristic function of this distribution is defined as:

$$\varphi(q) = \exp \left[ c_0 - c_1 \frac{(q^2 + 1/l^2)^{\alpha/2}}{\cos(\pi\alpha/2)} \cos(\alpha \arctan(l | q |)) \right] \tag{4.6}$$

with scale factors:

$$c_1 = \frac{2\pi \cos(\pi\alpha/2)}{\alpha\Gamma(\alpha) \sin(\pi\alpha)} At \tag{4.7}$$

and

$$c_0 = \frac{l^{-\alpha}}{\cos(\pi\alpha/2)} c_1 = \frac{2\pi}{\alpha\Gamma(\alpha) \sin(\pi\alpha)} Al^{-\alpha} t \tag{4.8}$$

We obtain  $T = N\Delta t$  if we discretize in time with steps  $\Delta t$ . Thus, we must calculate the sum of  $N$  stochastic variables that are independent and identically distributed (i.i.d) at each interval.

For small  $N$  the probability will be very similar to the stable Lévy distribution. The model can be improved by standardizing if the variance is given by:

$$\sigma^2 = -\frac{\partial^2 \varphi(q)}{\partial q^2} \Big|_{q=0} \tag{4.9}$$



we have that

$$-\frac{\partial^2 \varphi(q/\sigma)}{\partial q^2} \Big|_{q=0} = -\frac{1}{\sigma^2} \frac{\partial^2 \varphi(q)}{\partial q^2} \Big|_{q=0} = 1 \quad (4.10)$$

Therefore, the standardized model is:

$$\ln \varphi_s(q) = \ln \varphi \left( \frac{q}{\sigma} \right) = c_0 - c_1 \frac{((q/\sigma)^2 + 1/l^2)^{\alpha/2}}{\cos(\pi\alpha/2)} \cos \left( \alpha \arctan \left( l \frac{|q|}{\sigma} \right) \right) \quad (4.11)$$

$$= \frac{2\pi A l^{-\alpha} t}{\alpha \Gamma(\alpha) \sin(\pi\alpha)} \left[ 1 - ((q/\sigma)^2 + 1)^{\alpha/2} \cos \left( \alpha \arctan \left( \frac{ql}{\sigma} \right) \right) \right]. \quad (4.12)$$

This equation represents the normalized Lévy model applied in this study. In the numerical simulations, we adjust the values of A, the arbitrary scale parameter  $l$  and the characteristic exponent  $\alpha$  to best fit the cumulative function.

# Chapter 5

## Scaling Methods

There are several methods used to estimate scaling exponents of time series data. In this study, we limit ourselves to Detrended Fluctuation Analysis (DFA) and the Diffusion Entropy Analysis (DEA) methods. We compare the DFA and the DEA scaling exponents in the case of Lévy statistics and generate a relation between  $\alpha$  characterizing the normalized Lévy distribution and  $(H, \delta)$  characterizing the DFA and DEA respectively.

### 5.1 Detrended Fluctuation Analysis (DFA)

The Detrended Fluctuation Analysis (DFA) method is an important technique in revealing long range correlations in non-stationary time series.

The advantages of DFA over conventional methods are that, it permits the detection of intrinsic self-similarity embedded in a seemingly non-stationary time series (i.e. time series whose means, variances and covariances change over time.), and also avoids the spurious detection of apparent self-similarity, which may be an artifact of extrinsic trends. Thus, the obtained exponents of the DFA is similar to the Hurst exponent, except that DFA may also be applied to signals whose underlying statistics (such as mean and variance) or dynamics are non-stationary (changing with time).

The numerical procedure to estimate the DFA exponent is presented below.

Let  $N$  be the length of time series  $(y_1, y_2, y_3, \dots, y_N)$ . The logarithmic ratio of the time series is obtained. The length of the new time series  $M(t)$  will be  $N - 1$ .

$$M(t) = \log \left( \frac{y_{t+1}}{y_t} \right) \quad (5.1)$$

for  $t = 1, 2, \dots, N - 1$

The absolute value of  $M(t)$  is integrated:

$$y(t) = \sum_{i=1}^t |M(i)| \quad (5.2)$$

Then the integrated time series of length  $N$  is divided into  $m$  boxes of equal length  $n$  with no intersection between them. As the data is divided into equal length intervals, there may be some left over at the end. In order to take account of these leftover values, the same procedure is repeated but beginning from the end, obtaining  $2N/n$  boxes. Then, a least squares line is fitted to each box, representing the trend in each box, thus obtaining  $y_n(t)$ . Finally the root mean square fluctuation is calculated by using the formula:

$$F(n) = \sqrt{\frac{1}{2N} \sum_{t=1}^{2N} [y(t) - y_n(t)]^2} \quad (5.3)$$

This computation is repeated over all box sizes to characterize a relationship between the box size  $n$  and  $F(n)$ . A linear relationship between the  $F(n)$  and  $n$  (i.e. box size) in a log-log plot reveals that the fluctuations can be characterized by a scaling exponent  $H$ , the slope of the line relating  $\log F(n)$  to  $\log n$ . This generates the mathematical relation:

$$F(n) \propto n^H \quad (5.4)$$

For data series with no correlations or short-range correlation,  $H$  is expected to be 0.5. For data series with long-range power law correlations,  $H$  would lie between 0.5 and 1 and for power law anti-correlations  $H$  would lie between 0 and 0.5. This method is used to measure correlations in financial market time series and in the daily evolution of some of the most relevant indices.

Next we discuss the Diffusion Entropy Analysis (DEA).

## 5.2 Diffusion Entropy Analysis (DEA)

The DEA is another common method used to analyze and detect the scaling properties of time series. Using the DEA, one can determine if the characterization of a time series

follows a Gaussian or Lévy distribution, as well as establish the existence of long-range correlations in the time series.

A function  $\Phi(r_1, r_2, \dots)$  is said to be invariant to scale changes if it fulfills the property:

$$\Phi(r_1, r_2, \dots) = \gamma^a \Phi(\gamma^b r_1, \gamma^c r_2, \dots). \quad (5.5)$$

This equation shows that if we scale all coordinates  $r$  with an appropriate choice of exponents  $a, b, c, \dots$ , then we always recover the original function. In the case of a time series we can interpret the sequence of the numbers that form it as the generators of a diffusion process and, therefore, to study the relevant probability distribution function  $\rho(x, t)$  where  $x$  denotes the variable that collects the fluctuations of the diffusive process. If the time series is stationary, the scale property takes the shape

$$\rho(x, t) = \frac{1}{t^\delta} F\left(\frac{x}{t^\delta}\right), \quad (5.6)$$

where the  $\delta$  is the scaling exponent.

It is relatively easy to determine the exponent using the following procedures:

- Transform the series into a diffusion process.
- Calculate Shannon's entropy of the said process.

$$S(t) = - \int_{-\infty}^{\infty} \rho(x, t) \ln[\rho(x, t)] dx \quad (5.7)$$

Suppose that  $\rho(x, t)$  satisfies the scale condition (5.6), replacing it in the previous equation we obtain (after a few simple steps)

$$S(t) = A + \delta \ln(t) \quad (5.8)$$

where

$$A = - \int_{-\infty}^{\infty} F(y) \ln[F(y)] dy \quad (5.9)$$

Equation (5.8) describes the fact that the entropy grows linearly with  $\ln(t)$  and that the slope of the linear function is the scaling coefficient.

Below is the algorithm of Diffusion:

Consider a sequence of  $M$  numbers

$$\epsilon_i, i = 1, \dots, M \tag{5.10}$$

The purpose of the DEA algorithm is to establish the possible existence of scalability in an efficient way without altering the series with any form of detrending. To do this we do the following:

1. Select an integer  $l$  such that  $1 \leq l \leq M$ . This integer will be our "weather".
2. For each time we can find  $M - l + 1$  subseries of length  $l$  defined such that

$$\epsilon_i^s \equiv \epsilon_{i+s}, i = 1, \dots, l \tag{5.11}$$

with  $s = 0, \dots, M - l$ .

3. For each subseries we construct a distribution diffusion path by the position

$$x^{(s)}(l) = \sum_{i=1}^l \epsilon_i^s = \sum_{i=1}^l \epsilon_{i+s} \tag{5.12}$$

We can imagine this position as the Brownian path, at regular intervals of time, of which has jumped towards forward and back according to what is determined by each subseries.

4. We are now in a position to assess the entropy of the diffusion. For this we partition the  $x - axis$  into size  $\eta(l)$  cells and we tell how many parts there are in each of them at a time  $l$  dice. Let us call this number  $N_i(l)$ .

5. With  $N(l)$  we determine the probability  $p_i(l)$  that the particle is find in the  $i$ -th cell in time  $l$ .

$$p_i(l) = \frac{N_i(l)}{M - l + 1} \quad (5.13)$$

6. The entropy of the diffusion process at time  $l$  is given by

$$S_d(l) = - \sum_i p_i(l) \ln[p_i(l)] \quad (5.14)$$

The sub-index  $d$  wants to sign that the entropy evaluation was made from a discrete process.

7. A good starting point for choosing the size of the cell  $\eta(l)$  is to assume independence of  $l$  and determine it as a suitable fraction of the square root of the variance of the fluctuation  $\epsilon_i$ .

The DEA technique is able to determine the correct scaling exponent even when the statistical properties of the time series, as well as the dynamic properties, are anomalous.

# Chapter 6

## Data

In this study, we use data from two (2) different fields namely Financial Stock Markets and Geophysics.

### 6.1 Financial Market Data

The financial data used for the analysis in this paper was obtained from yahoo finance. All data points collected are daily close values. Below are the names of countries or cities where data was collected as well as the respective start and end dates of data collection: Mexico (MXX), from November 8 1991 to October 22 2001; Brazil (BOVESPA), from April 27 1993 to June 24 2005; Argentina (MERVAL), from October 8 1996 to June 24 2005; Hong Kong (HSI), from January 2 1991 to June 24 2005; Phillipines (PSI), from 1997 to 2001; Thailand (SETI), from 1997 to 2001; New York (SP500) from January 3 1950 to June 23 2005; SPC, from 1991 to 2001 and USA (NASDAQ), from 1997 to 2001.

### 6.2 Geophysical Time Series Data

The Bezymianny Volcano Campaign Seismic Network (PIRE) collected the volcanic data at different eruption times at two different seismic stations namely BEZB and BELO. Data used in this paper were requested for 10 days before and 5 days after the published time of the volcanic eruptions. Volcanic eruptions 1 and 2 were from BEZB and Volcanic eruptions 3-8 were from BELO.

## 6.3 Stationarity Test of Data

The term ‘stationary’ is a fundamental part of time series analysis. It is imperative to check for stationarity of the data before any model is applied especially when using time series data for empirical financial research . A time series is said to be stationary if its joint probability, mean and variance do not change over time and it does not follow any trend.

We use the Augmented Dickey-Fuller (ADF) test to examine the stationarity of both the returns of the financial market data and volcanic eruption data.

Before performing stationarity test on the data, we take a closer look at how the data looks like on graphs.

### 6.3.1 Financial Market Data

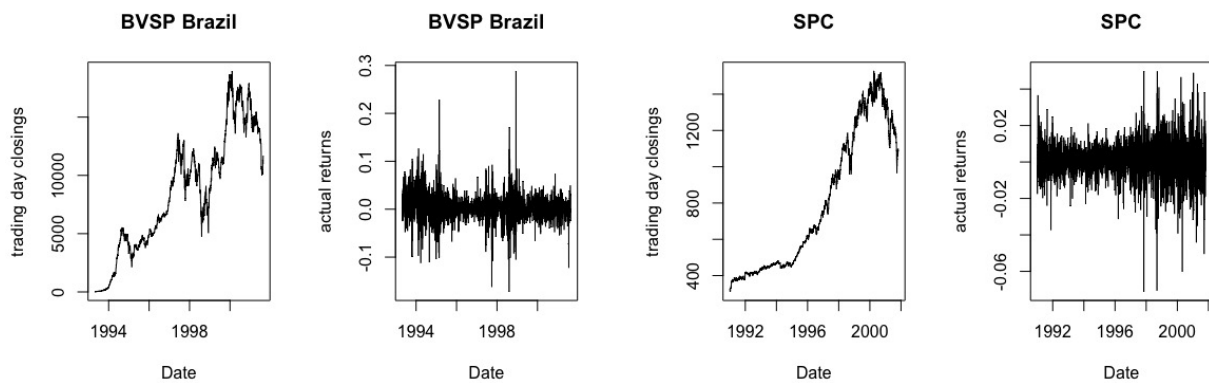


Figure 6.1: Plots of trading days closings and actual returns of BVSP and SPC



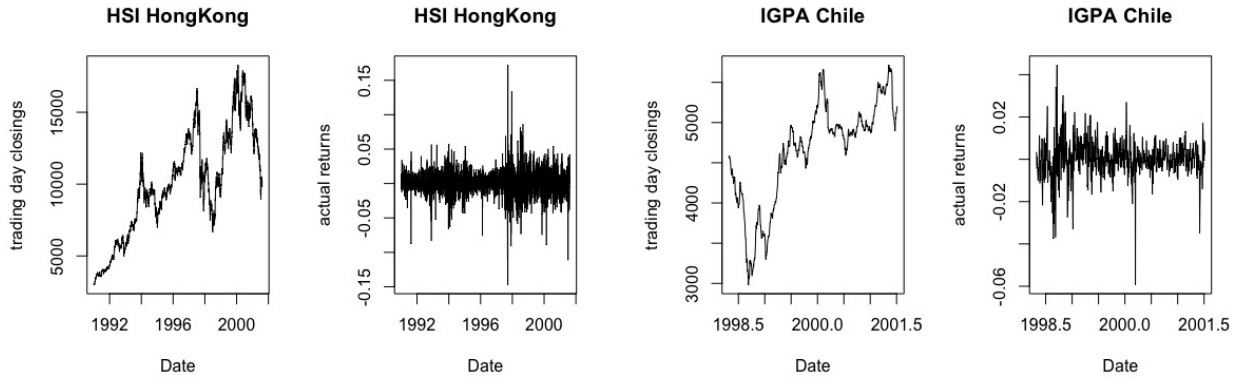


Figure 6.2: Plots of trading days closings and actual returns of HSI and IGPA

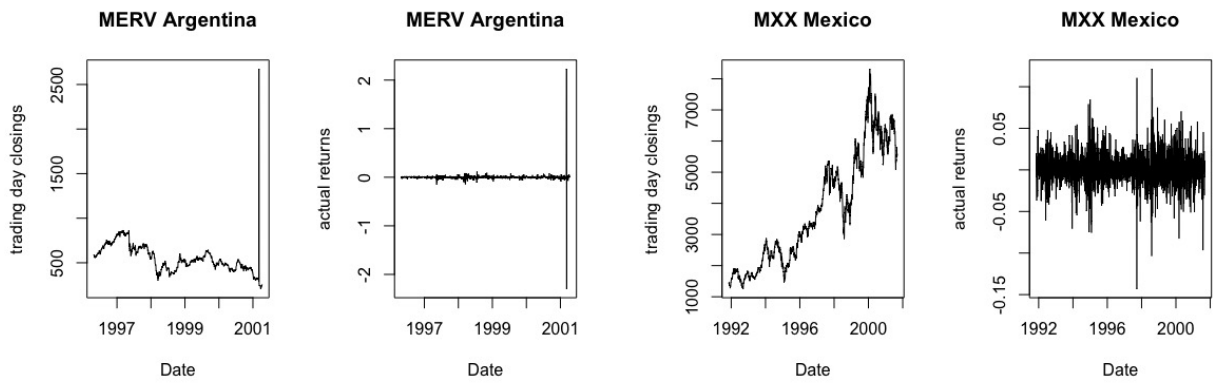


Figure 6.3: Plots of trading days closings and actual returns of MERV and MXX

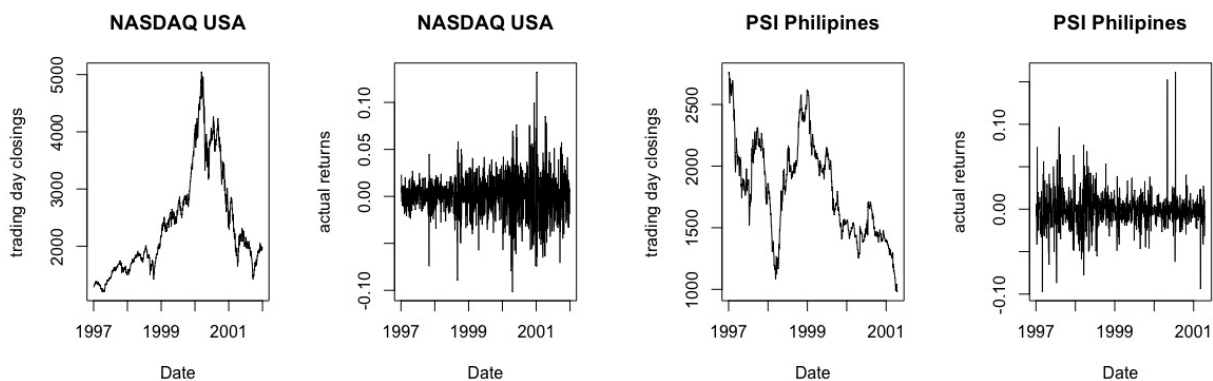


Figure 6.4: Plots of trading days closings and actual returns of Nasdaq and PSI

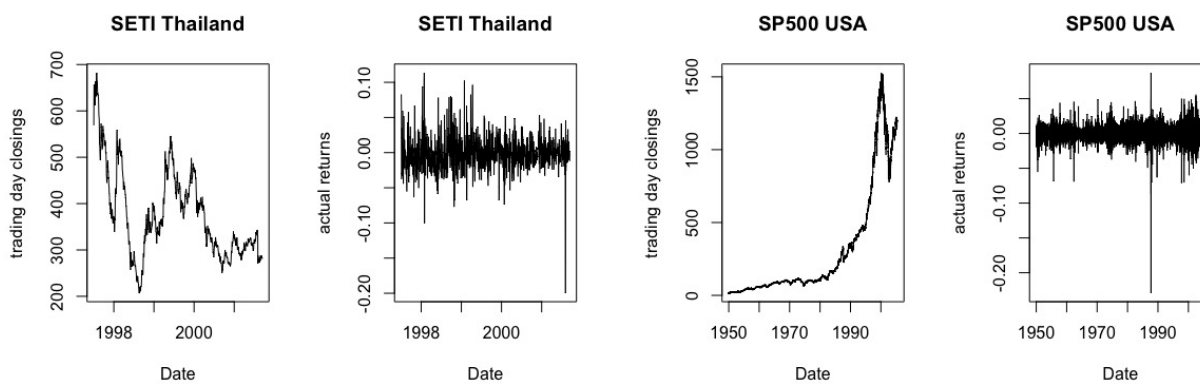


Figure 6.5: Plots of trading days closings and actual returns of SETI and SP500

### 6.3.2 Geophysical Time Series Data

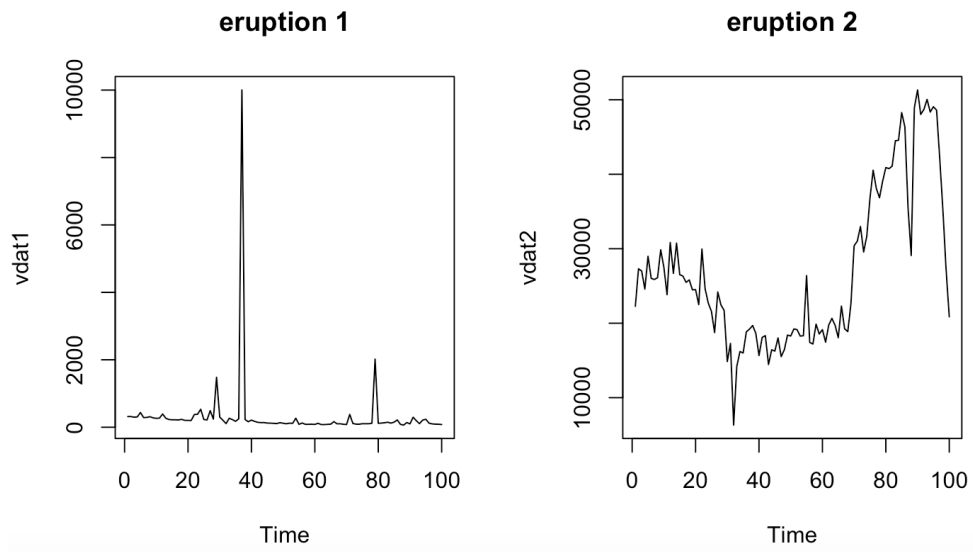


Figure 6.6: Plots of volcanic eruption numbers 1 and 2

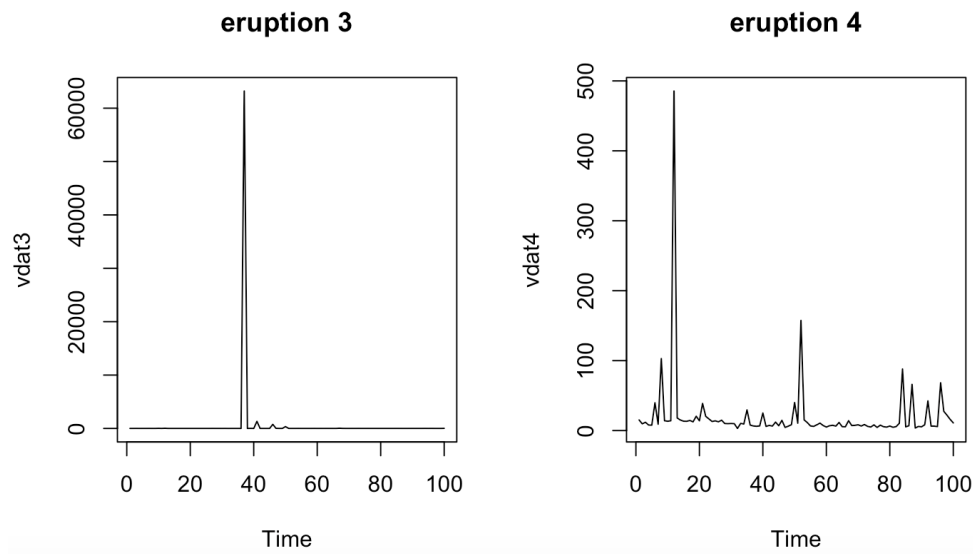


Figure 6.7: Plots of volcanic eruption numbers 3 and 4

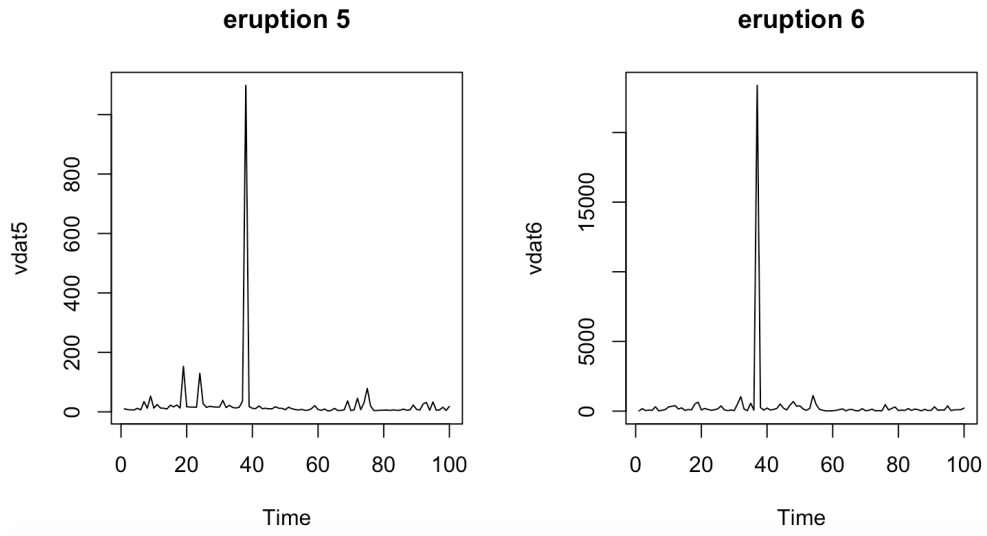


Figure 6.8: Plots of volcanic eruption numbers 5 and 6

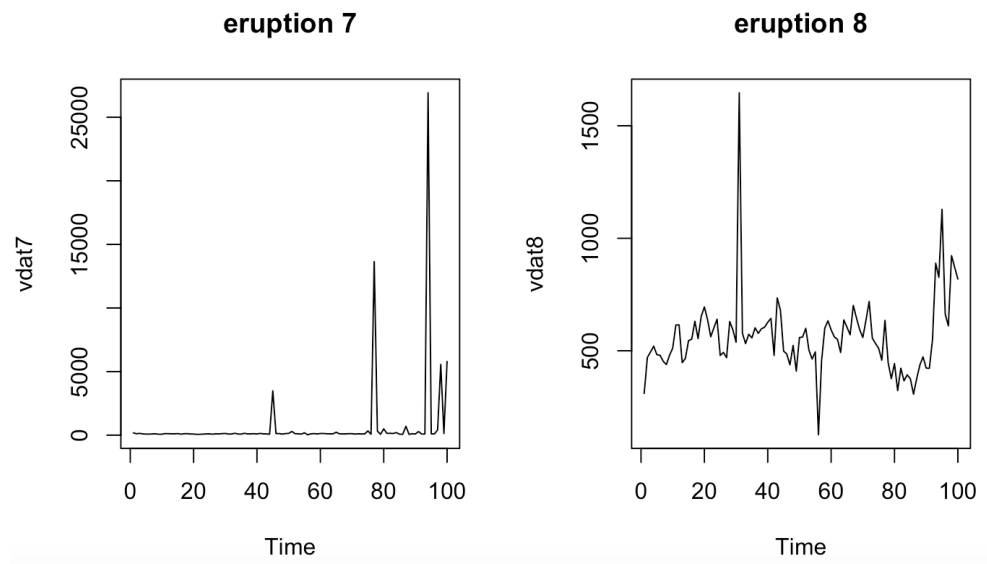


Figure 6.9: Plots of volcanic eruption numbers 7 and 8

### 6.3.3 Augmented Dickey-Fuller (ADF)

The ADF statistic tests the null hypothesis that the time series has a unit root against the alternative that it is stationary with the assumption that the data exhibits an Auto-Regressive Moving Average (ARMA) structure.

$H_0$  : A unit root is present in a time series data.

$H_a$  : Time series is stationary.

Market	Test statistic	Lag order	p-value
BVSP	-16.571	12	0.01
SPC	-19.992	13	0.01
HSI	-18.937	13	0.01
IGPA	-10.87	9	0.01
MERV	-13.298	10	0.01
MXX	-12.496	13	0.01
Nasdaq	-7.4649	10	0.01
PSI	-10.207	10	0.01
SETI	-8.4743	10	0.01
SP500	-30.856	24	0.01

Table 6.1: ADF test results for returns of various financial markets

Station	Eruption	Test statistic	Lag order	p-value
BEZB	1	-3.155	7	0.10
	2	-1.946	7	0.60
BELO	3	-3.188	8	0.09
	4	-3.181	8	0.09
	5	-3.342	8	0.07
	6	-3.366	7	0.06
	7	-2.784	7	0.25
	8	-2.092	7	0.54

Table 6.2: ADF test results for different eruption times of Volcano at the two seismic stations

### Comments

The test statistics are used to compute the p-values in the tables above. The p-values suggest whether the null hypothesis that the time series has unit root is acceptable or not, at the specified lag parameters estimated from the data.

We observe that, at a significant level of  $\alpha = 5\%$ , the p-values of returns of the financial market data is less than the significance level, so we *reject* the null hypothesis that the time series has a unit root and conclude that the returns of financial market is stationary.

We also observe that the p-values are greater than the significant level of  $\alpha = 5\%$  for the volcanic eruption data, so we *fail to reject* the null hypothesis that the time series has a unit root and conclude that the volcanic eruption data is non-stationary.

# Chapter 7

## Results

### 7.1 Numerical Results

Below are table of results for both the financial markets data and volcanic eruption data. There are six (6) columns in each table where the first column indicates the particular financial market or volcanic eruption numbers, the second - fourth columns show the scaling exponents of the DFA and the DEA as well as the best Lévy parameter  $\alpha$  characterizing the data. The last two columns are basically product of the respective scaling exponents and the Lévy parameter  $\alpha$ .

### 7.1.1 Financial Market Data

Markets	DFA ( $H$ )	DEA ( $\delta$ )	Lévy ( $\alpha$ )	$H.\alpha$	$\delta.\alpha$
BVSP	0.72	0.57	1.34	0.96	0.76
SPC	0.62	0.60	1.40	0.87	0.84
HSI	0.70	0.60	1.40	0.98	0.84
IGPA	0.65	0.53	1.40	0.91	0.74
MERV	0.62	0.56	1.40	0.87	0.78
MXX	0.66	0.59	1.34	0.90	0.79
Nasdaq	0.72	0.56	1.12	0.96	0.66
PSI	0.71	0.55	1.40	0.99	0.77
SETI	0.70	0.54	1.34	0.94	0.72
SP500	0.66	0.65	1.40	0.92	0.91

Table 7.1: Scaling exponents of various financial markets



### 7.1.2 Geophysical Time Series Data

Seismic Station	Eruption Number	DFA ( $\alpha$ )	DEA( $\delta$ )	Lévy ( $\alpha$ )	$H\alpha$	$\delta\alpha$
BEZB	1	0.74	0.68	1.12	0.83	0.76
	2	0.92	0.68	1.34	1.23	0.91
BELO	3	0.85	0.68	1.12	0.95	0.76
	4	0.66	0.68	1.40	0.92	0.95
	5	0.76	0.68	1.12	0.85	0.76
	6	0.67	0.68	1.34	0.90	0.91
	7	0.81	0.68	1.40	1.13	0.95
	8	0.75	0.68	1.34	1.01	0.91

Table 7.2: Scaling exponents of Volcanic Data from the two seismic stations

### 7.1.3 Remarks on the Table of Numerical Results

Empirical evidence presented in Tables 6.1 and 6.2 show that:

$$\delta \approx \frac{1}{\alpha} \tag{7.1}$$

$$H \approx \frac{1}{\alpha} \tag{7.2}$$

## 7.2 Analytical Relations

In this section, we prove analytically the numerical relations obtained from the previous section.

### 7.2.1 Detrended Fluctuation Analysis (DFA)

From (5.4), we obtain a mathematical relation that reveals a linear relationship between the root mean square fluctuation  $F(n)$  and the box size  $n$  in a log-log plot. i.e.

$$\log F(n) = \log K + H \log(n) \quad (7.3)$$

where  $K$  is a positive constant of proportionality.

$$\log F(n) - \log K = H \log(n) \quad (7.4)$$

**property 1:** Given a random variable  $X$ , if we denote with  $Law(X)$  its probability density function (for example, for a Gaussian random variable, we write  $Law(X) = N(\mu, \sigma^2)$ ) then we will say that the random variable  $X$  is stable, or that it has a stable distribution if for any  $n \geq 2$  there exists a positive number  $C_n$  and a number  $D_n$  so that:

$$Law(X_1 + X_2 + \dots + X_n) = Law(C_n X + D_n) \quad (7.5)$$

where  $X_1, X_2, \dots, X_n$  are independent random copies of  $X$ , this means that  $Law(X_i) = Law(X)$  for  $i = 1, 2, \dots, n$ . If  $D_n = 0$ , then  $X$  is said to be a strictly stable variable. It can be shown (see Samorodnitsky and Taqqu (1994)) that

$$C_n = n^{1/\alpha} \quad (7.6)$$

for some parameter  $\alpha$ ,  $0 < \alpha \leq 2$ .

This implies that

$$\log C_n = \frac{1}{\alpha} \log n \quad (7.7)$$

Subtracting (7.3) and (7.7) gives the relation:

$$H \approx 1/\alpha$$

if

$$\lim_{n \rightarrow \infty} \left| \frac{\log F(n) - \log K - \log C_n}{\log n} \right| \rightarrow 0 \quad (7.8)$$

## 7.2.2 Diffusion Entropy Analysis (DEA)

From the PDF in (5.6), one can calculate the Shannon entropy [5, 7],

$$S(t) = -\int p(x, t) \ln p(x, t) dx$$

Suppose the PDF satisfies (5.6), then after a few simple steps (Mariani *et al* [7]), we obtain

$$S(t) = A + \delta \ln(t) \quad (7.9)$$

where

$$A = -\int F(y) \ln[F(y)] dy \quad (7.10)$$

Since the Lévy distribution is stable, for any  $n \geq 2$  there exists a positive number (using property 1)  $C_n$  and a number  $D_n$  such that:

$$Law(X_i) = Law(C_n X + D_n) \quad (7.11)$$

for  $i = 1, 2, \dots, n$  and  $Law(X)$  denotes the probability density function of  $X$ . If  $D_n = 0$ , then  $X$  is said to be a strictly stable variable. It can be shown (see Samorodnitsky and Taqqu (1994)) that

$$C_n = n^{1/\alpha} \quad (7.12)$$

for some parameter  $\alpha$ ,  $0 < \alpha \leq 2$ .

Equation (7.9) becomes

$$S(t) - S(1) = \delta \ln(t) \quad (7.13)$$

Then, for some positive number  $C_n$

$$|S(t) - S(1) - \ln(C_n)| = 0, \text{ where } t \geq 2 \quad (7.14)$$

such that

$$\delta \approx \frac{1}{\alpha}$$

# Chapter 8

## Figures

Figures obtained from the numerical simulations of the various Financial markets and Geophysical time series data are shown below.

### 8.1 Financial Market Data

#### 8.1.1 Lévy Flight Model

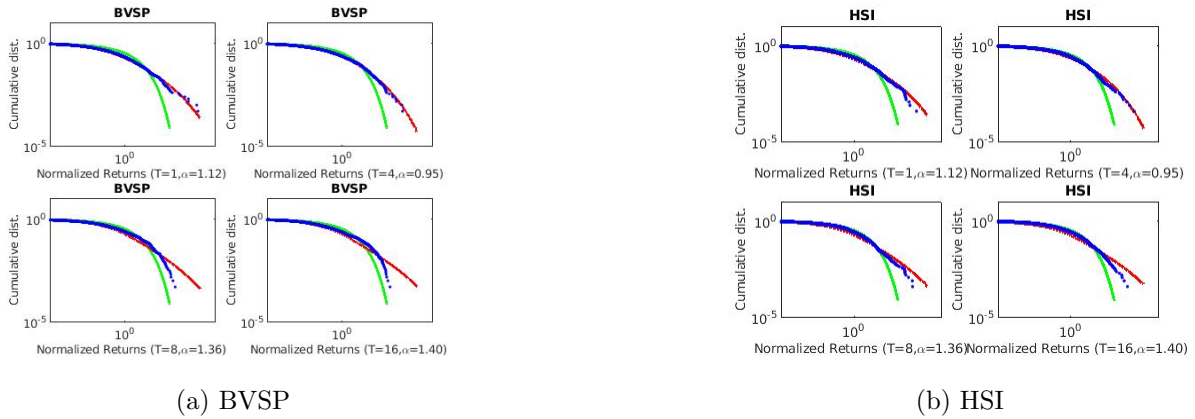
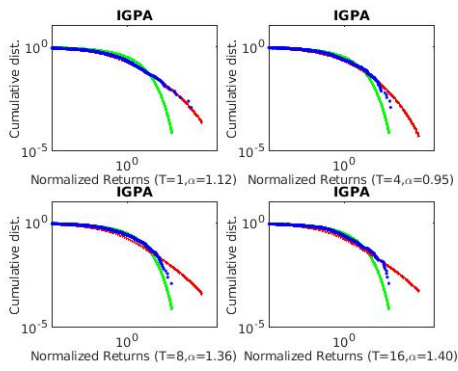
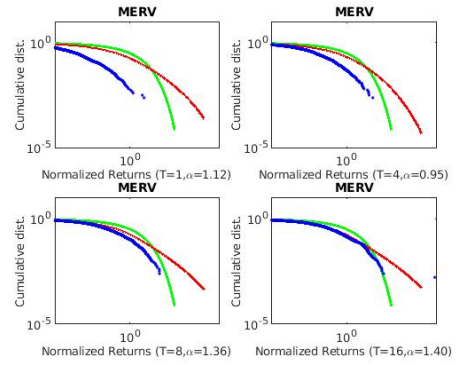


Figure 8.1: Lévy model fits for BVSP and HSI

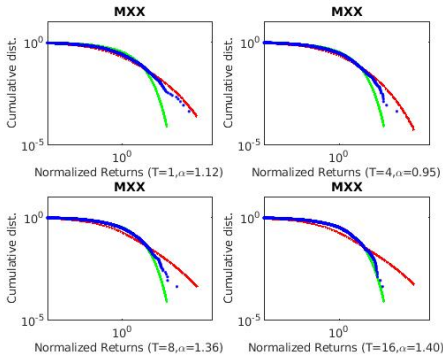


(a) IGPA

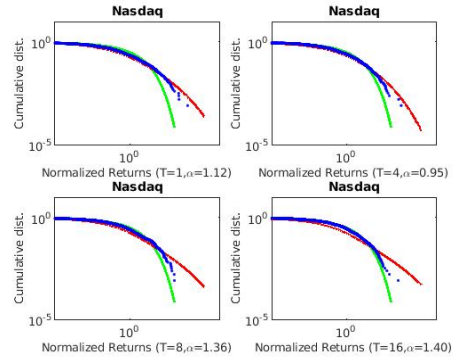


(b) MERV

Figure 8.2: Lévy model fits for IGPA and MERV

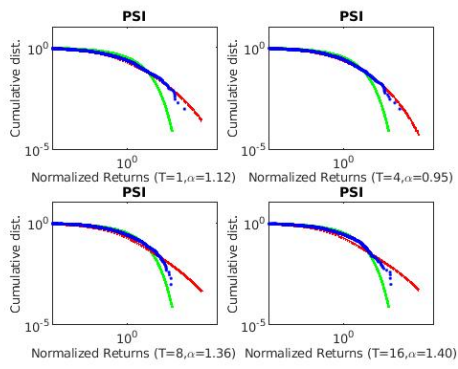


(a) MXX

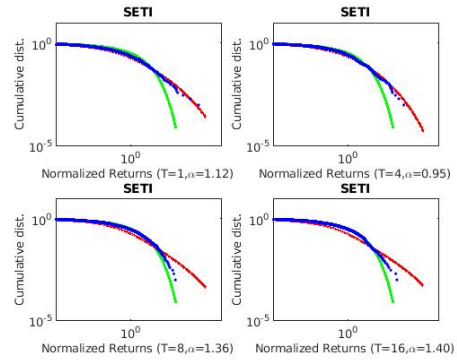


(b) Nasdaq

Figure 8.3: Lévy model fits for MXX and Nasdaq

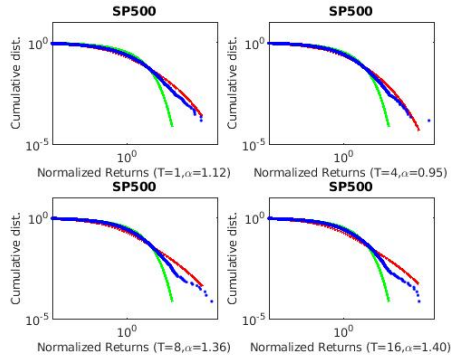


(a) PSI

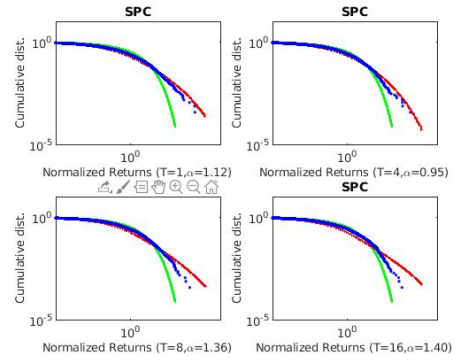


(b) SETI

Figure 8.4: Lévy model fits for PSI and SETI



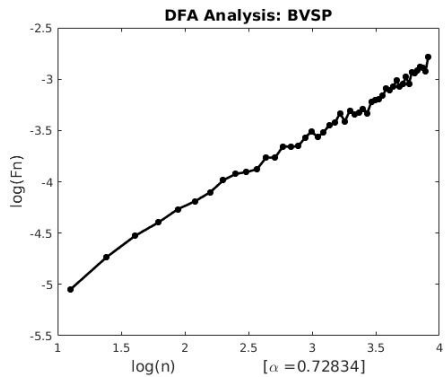
(a) SP500



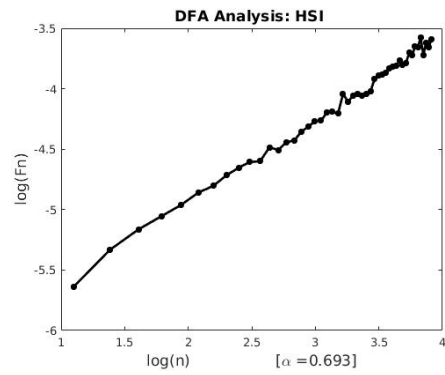
(b) SPC USA

Figure 8.5: Lévy model fits for SP500 and SPC USA

## 8.1.2 Detrended Fluctuation Analysis (DFA)

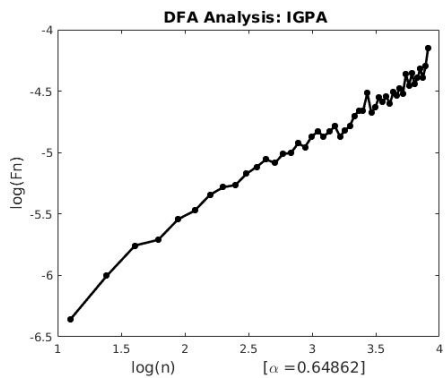


(a) BVSP

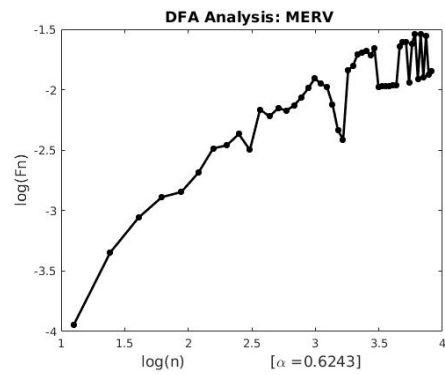


(b) HSI

Figure 8.6: DFA for BVSP and HSI



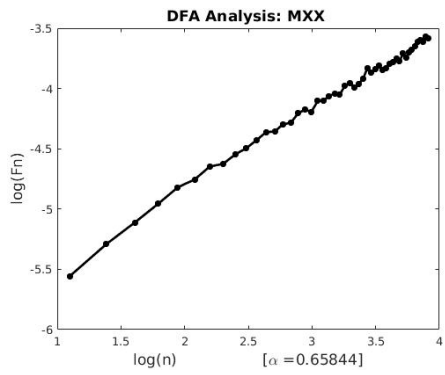
(a) IGPA



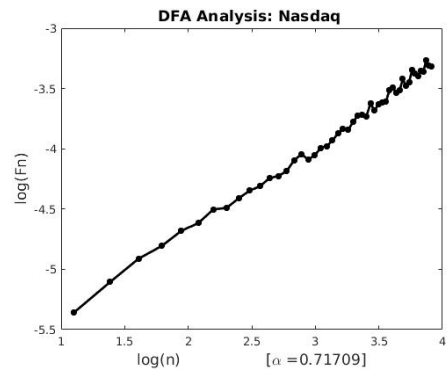
(b) MERV

Figure 8.7: DFA for IGPA and MERV



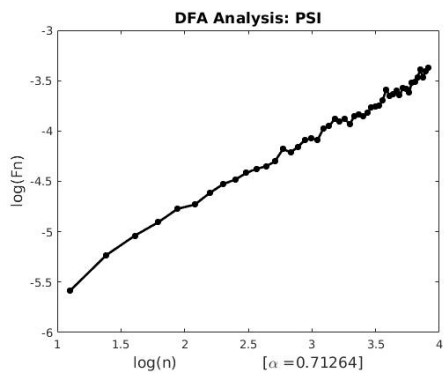


(a) MXX

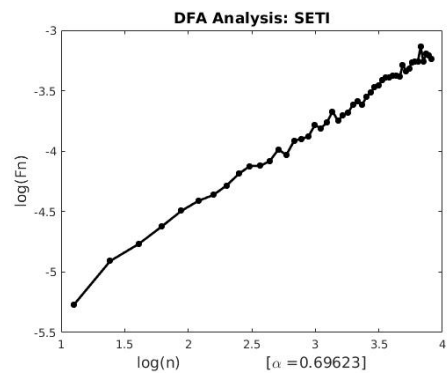


(b) Nasdaq

Figure 8.8: DFA for MXX and Nasdaq

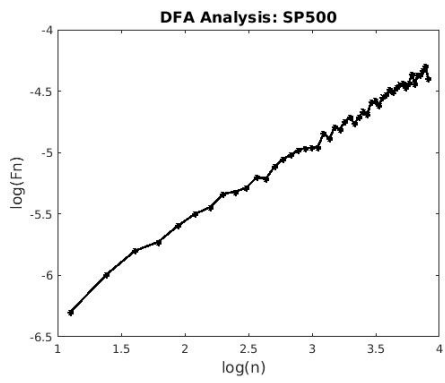


(a) PSI

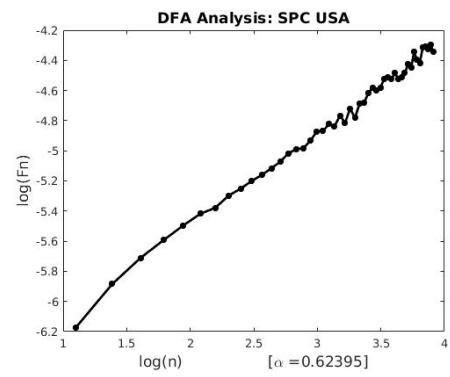


(b) SETI

Figure 8.9: DFA for PSI and SETI



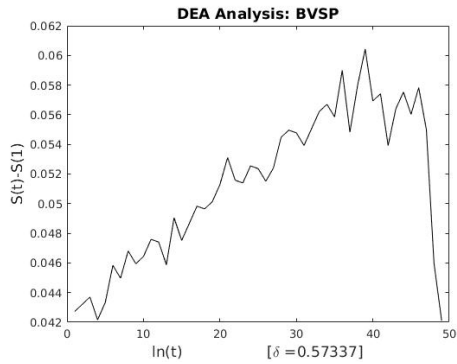
(a) SP500



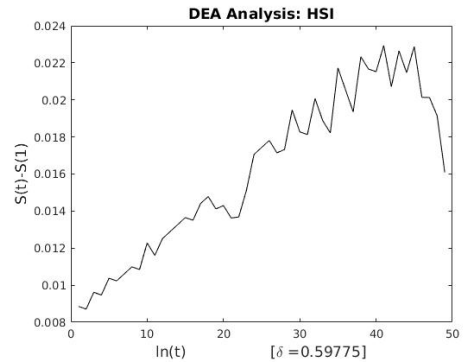
(b) SPC USA

Figure 8.10: DFA for SP500 and SPC USA

### 8.1.3 Diffusion Entropy Analysis (DEA)

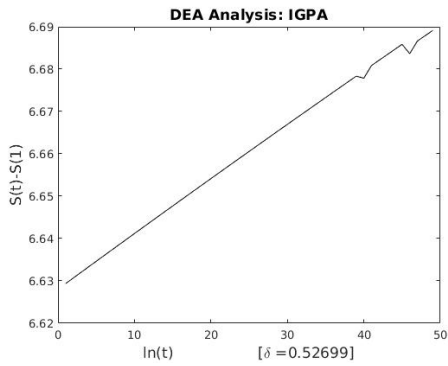


(a) BVSP

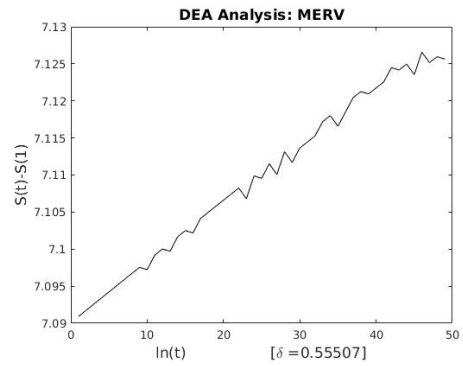


(b) HSI

Figure 8.11: DEA for BVSP and HSI

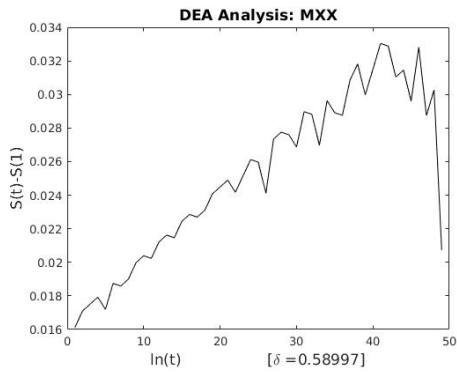


(a) IGPA

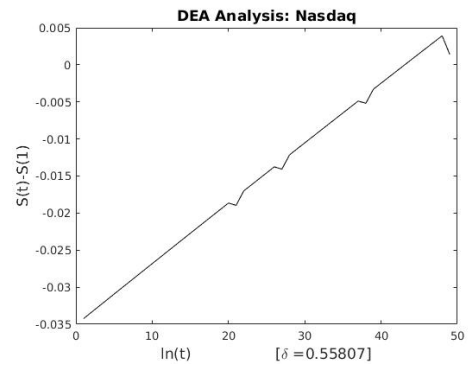


(b) MERV

Figure 8.12: DEA for IGPA and MERV

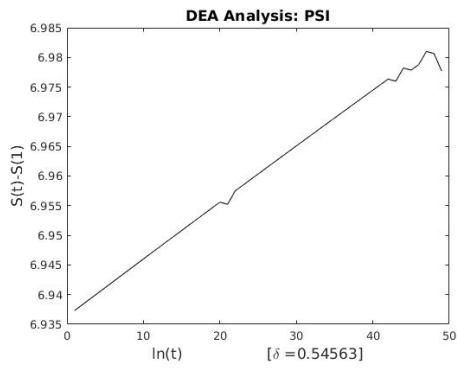


(a) MXX

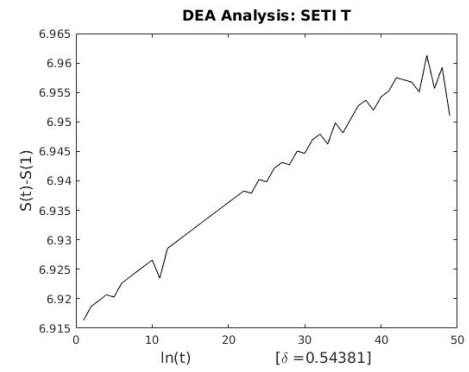


(b) Nasdaq

Figure 8.13: DEA for MXX and Nasdaq

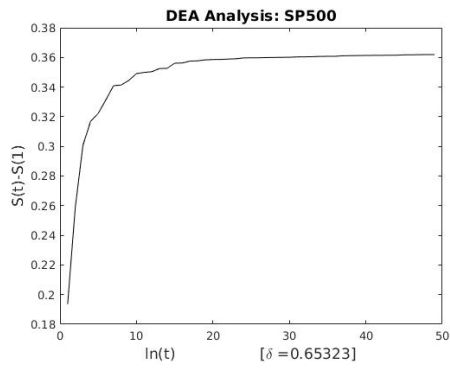


(a) PSI

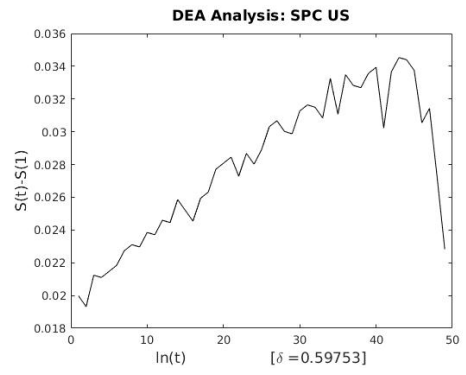


(b) SETI

Figure 8.14: DEA for PSI and SETI



(a) SP500



(b) SPC USA

Figure 8.15: DEA for SP500 and SPC USA

## 8.2 Geophysical Time Series Data

### 8.2.1 Lévy Flight Model

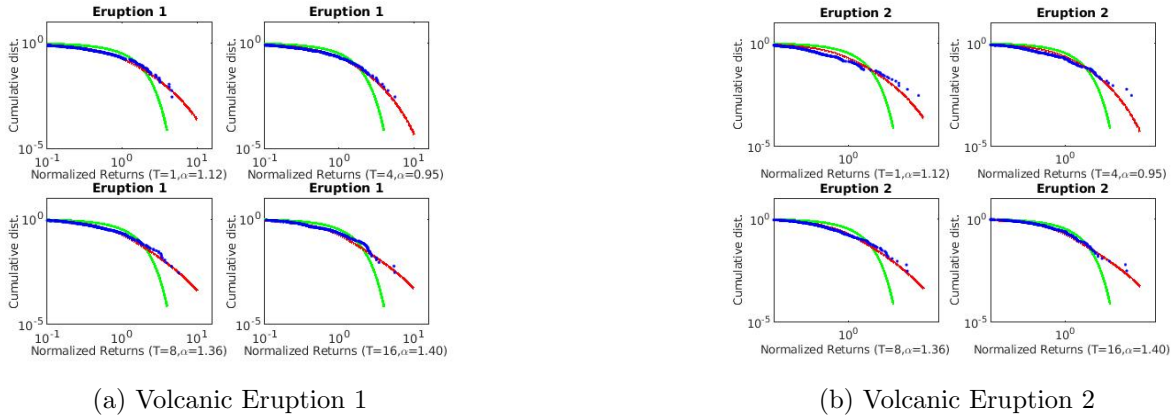


Figure 8.16: Lévy model fit for Volcanic Eruptions 1 and 2

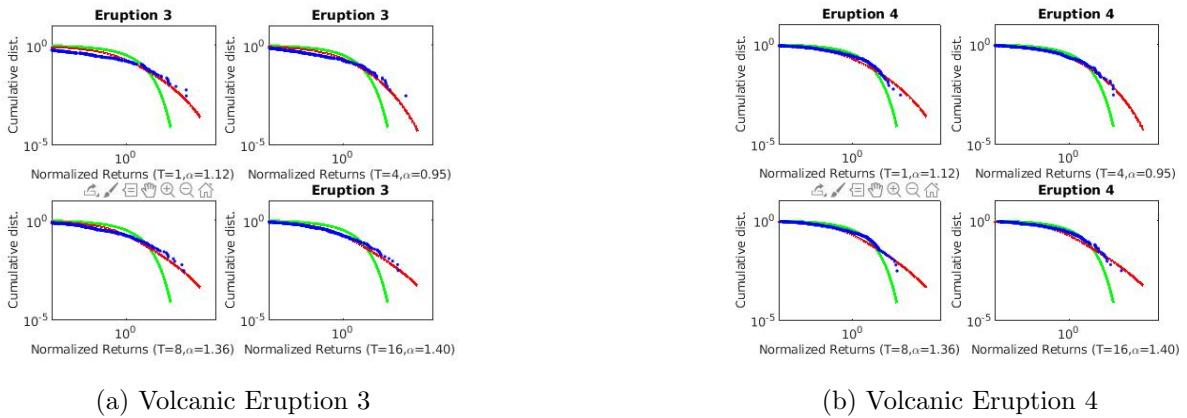
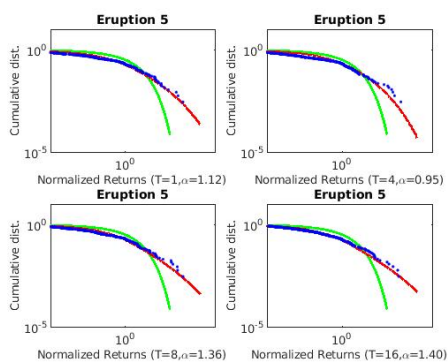
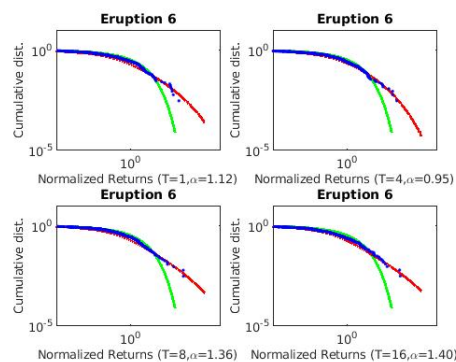


Figure 8.17: Lévy model fit for Volcanic Eruptions 3 and 4

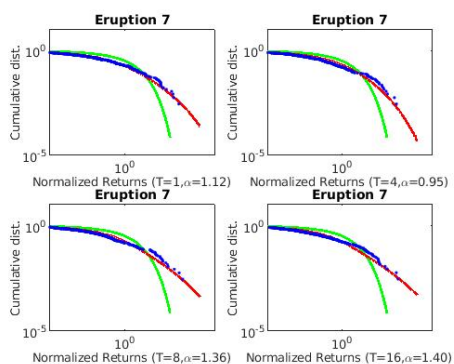


(a) Volcanic Eruption 5

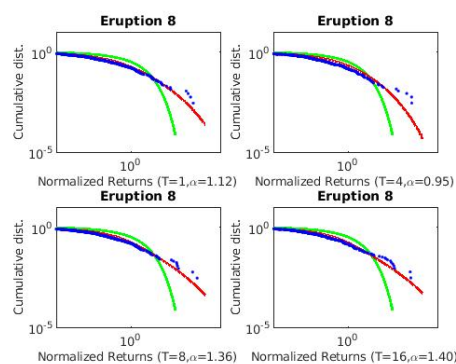


(b) Volcanic Eruption 6

Figure 8.18: Lévy model fit for Volcanic Eruptions 5 and 6



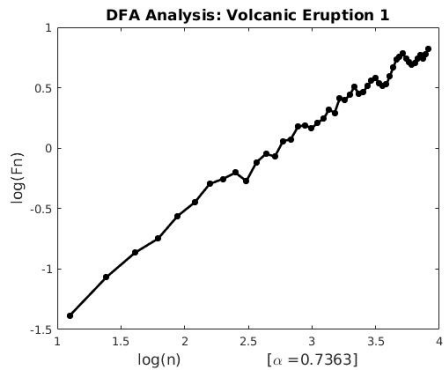
(a) Volcanic Eruption 7



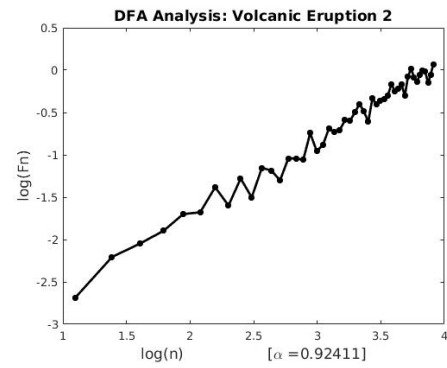
(b) Volcanic Eruption 8

Figure 8.19: Lévy model fit for Volcanic Eruptions 7 and 8

## 8.2.2 Detrended Fluctuation Analysis (DFA)

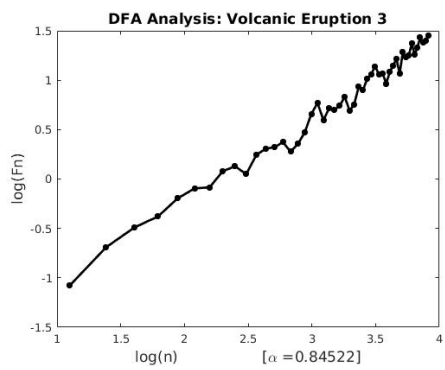


(a) Volcanic Eruption 1

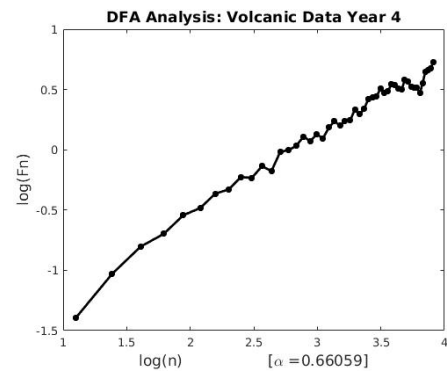


(b) Volcanic Eruption 2

Figure 8.20: DFA for Volcanic Eruption 1 and 2



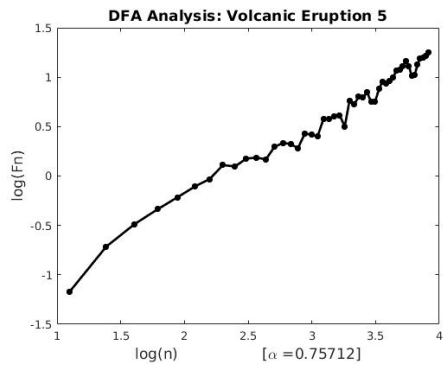
(a) Volcanic Eruption 3



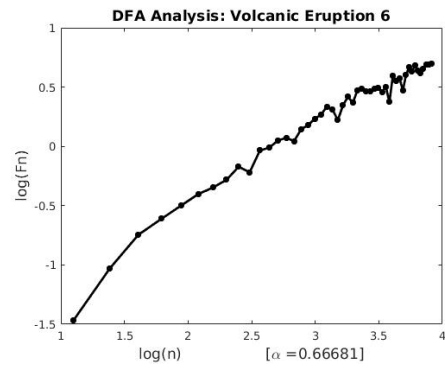
(b) Volcanic Eruption 4

Figure 8.21: DFA for Volcanic Eruption 3 and 4



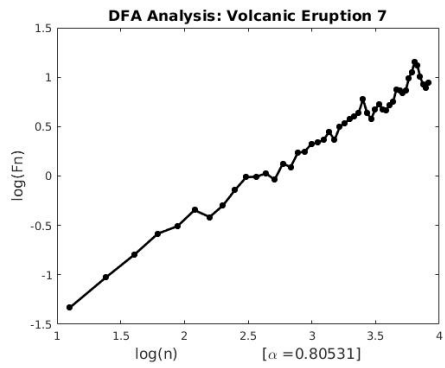


(a) Volcanic Eruption 5

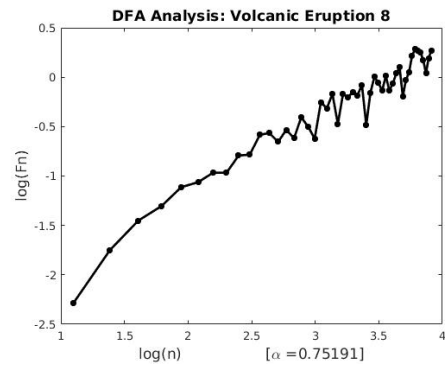


(b) Volcanic Eruption 6

Figure 8.22: DFA for Volcanic Eruption 5 and 6



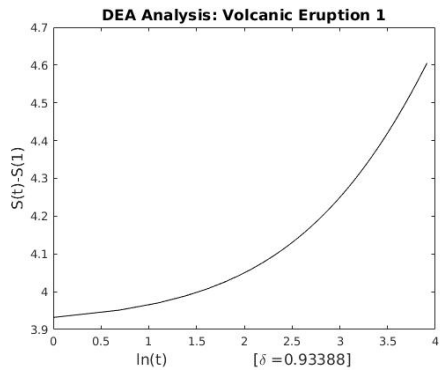
(a) Volcanic Eruption 7



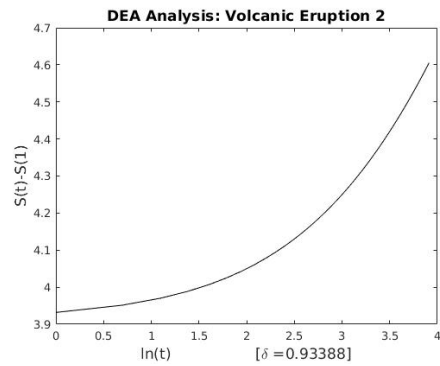
(b) Volcanic Eruption 8

Figure 8.23: DFA for Volcanic Eruption 7 and 8

### 8.2.3 Diffusion Entropy Analysis (DEA)

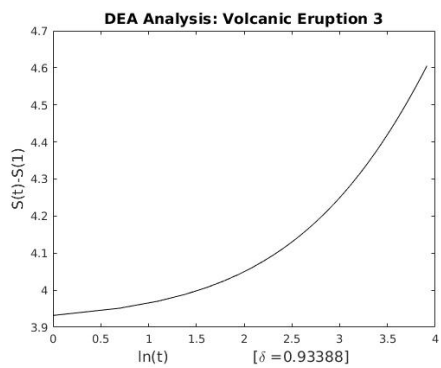


(a) Volcanic Eruption 1

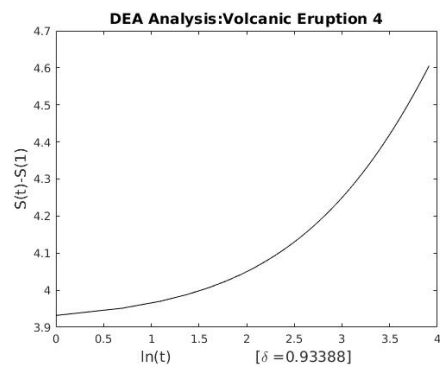


(b) Volcanic Eruption 2

Figure 8.24: DEA for Volcanic Eruption 1 and 2

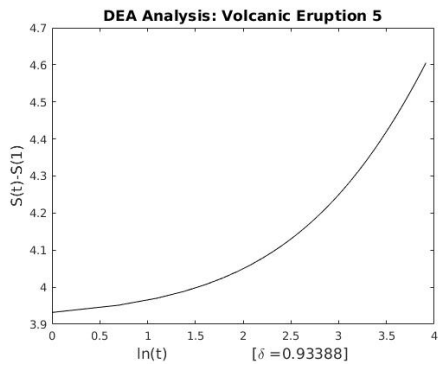


(a) Volcanic Eruption 3

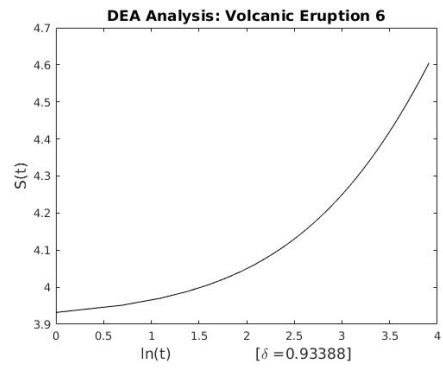


(b) Volcanic Eruption 4

Figure 8.25: DEA for Volcanic Eruption 3 and 4

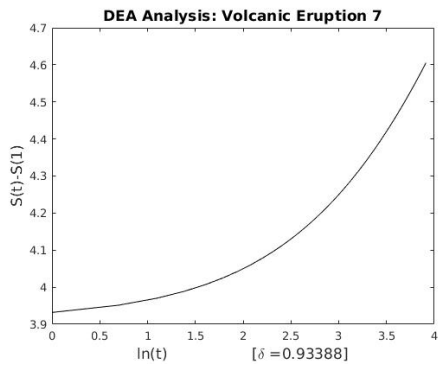


(a) Volcanic Eruption 5

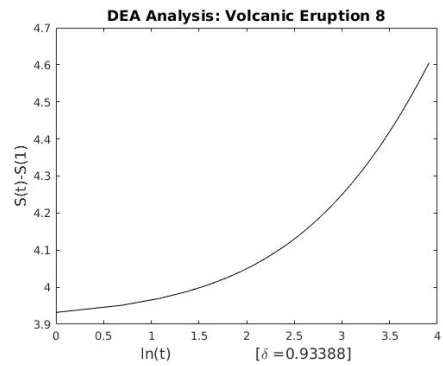


(b) Volcanic Eruption 6

Figure 8.26: DEA for Volcanic Eruption 5 and 6



(a) Volcanic Eruption 7



(b) Volcanic Eruption 8

Figure 8.27: DEA for Volcanic Eruption 7 and 8

# Chapter 9

## Concluding Remarks

### 9.1 Significance of the Result

Now that we have shown that there exists an inverse relation between the Lévy Flight model parameter and the parameters of the self-similar models, DFA and DEA, what does it mean to the financial market and seismic data analysis community?

In the event where the scaling exponent of either of the parameter estimators is not known or is difficult or expensive to estimate, we are able to approximate it using the inverse relationship we generated in the numerical and analytical results in this study.

### 9.2 Future Work

#### Aims

Given the financial market data and geophysical time series, I seek to

1. investigate the robustness of the Detrended Fluctuation Analysis (DFA) and the Diffusion Entropy Analysis (DEA). Given an additive time series model  $Y_t$ , s.t

$$Y_t = T_t + S_t + \epsilon_t$$

where the time series can be decomposed into three (3) main components namely: Trend =  $T_t$ , Seasonal =  $S_t$  and Random(or noise) =  $\epsilon_t$ . The trend component of the

series shows the general tendency of the data to increase or decrease during a long period of time. I estimate the trend components using a Moving Average (MA) filter. Seasonality, on the other hand, measures the presence of variations, contrasted with cyclical patterns, that occur at specific regular intervals less than a year, such as weekly, monthly, or quarterly. I compute seasonality by averaging over all periods in the time series for each time unit and centering the observed values. Finally, by removing both trend and seasonal components of the original time series  $Y_t$ , I obtain the random (error) component  $\epsilon_t$  which is mostly noise.

Now, I estimate long-memory behavior of the time series by removing, one at a time, the trends, seasonality and errors using the algorithms illustrated in Chapter 5, the DFA and DEA

2. investigate intra and inter correlations between the volcanic eruption times and estimate lead-lag effects in this geophysical time series using the Auto Correlation (ACF), Cross Correlation (CCF) and Cross Co-variance (CCV) analyses. The  $CCF(X, Y)$  is particularly affected by the structure of time series when the x-series and y-series exhibit "common" trends over some time period. This is when Pre-whitening is useful to estimate the CCF correctly by first, determining the residuals from the x-variable time series model. The estimated coefficients from the x-variable time series model are then used to filter the y-variable series to obtain an "estimated" y-variable series based on the x-variable time series model. Finally, we investigate the CCF between the residuals from the x-variable time series model and the filtered y-variable series.

I then examine whether one volcanic eruption time series is useful in forecasting another by performing a statistical hypothesis called Granger Causality Test. The Granger Causality Test estimates predictive causality based on the *post hoc fallacy* assumption that: "Since event Y followed event X, event Y must have been caused by event X." I will perform this analysis based on a series of t-tests and F-tests on lagged values of X (including lagged values of Y), that those X values provide statistically

significant information about future values of  $Y$ .

3. forecast long-memory behavior using one-step-ahead and two-step-ahead predictions in ARFIMA models. I also propose using forecasting method based on spectral analysis which assumes that a model consists of combination of trigonometric functions with a certain frequency. This is because many time series data including the ones used in this study are often characterized by cyclical variations and random fluctuations. I will characterize the time series data used in this study as a finite sum of time series of different frequencies  $\omega$  with corresponding oscillation periods  $\tau = \frac{2\pi}{\omega}$  using Discrete Fourier Transforms (DFT). This allows you to identify significant peaks and frequency components that explains the changes in the time series. Based on this, I will construct a forecast for future time periods.
4. estimate the stochastic volatility using Generalized Auto Regressive Conditional Heteroskedasticity (GARCH) model and also try out other extensions including GARCH, ARMA-GARCH, IGARCH to best fit the data used herein. I fit a GARCH( $p$ ,  $q$ ) model to the data where  $p$  and  $q$  denote the order of the fitted model, by computing the maximum-likelihood estimates of the conditionally normal time series model. GARCH uses Quasi-Newton optimizer to find the maximum likelihood estimates. The choice of extension model is dependent upon the type of model that the error variance in the time series data follow. These extensions are known to capture some characteristics of time series and other spurious time series such as time-varying volatility and volatility clustering.
5. perform a Bayesian estimation of stochastic volatility models via Markov Chain Monte Carlo (MCMC) scheme. Bayesian approaches are known to offer attractive alternative to statistical estimation which enables small sample results, robust estimation, model discrimination and model combination. In Bayesian analysis, the volatility parameters and the white noise process are both unknown random quantities.

I assume parameter vector  $\theta$  which denotes the vector of all the unknowns. I then

start with a *prior* probability distribution for  $\theta$  that provides a summary of all prior knowledge about  $\theta$ . The main Bayesian inference step is the use of Bayes' theorem. It is necessary for combining the prior information about  $\theta$  with that of the data by computing the conditional distribution of  $\theta$  given the data widely known as the *posterior distribution*. I use the MCMC numerical simulation approach to compute the posterior because it's nearly impossible to compute analytically. The resulting simulated data from the MCMC are used to compute the Bayesian estimates of our stochastic volatility model.

I examine how well these results compare to estimators of other well-known calibration methods of estimating stochastic volatility (SV) parameters including the GARCH models that makes use of maximum likelihood estimation. I will then use Statistical Information Criteria (such as AIC, AICc, BIC) to specify the best SV model for the financial and geophysical data used in this study.

<b>Time Span</b>	<b>Task</b>
Aug. - Dec 2020	<p>Investigate the robustness of the DFA and DEA by estimating long-memory behavior of time series less trends, seasonal and random components.</p> <p>Examine intra-and inter-correlations in the volcanic eruption times and estimate lead-lag effects in the eruptions using ACF, CCF and CCV.</p> <p>Examine causality in the different eruption times of the Bezymianny volcano using the post hoc hypothesis test, Granger Causality test.</p>
Jan. - May 2021	<p>Forecast long-memory behavior using <math>n</math>-step-ahead predictions in ARFIMA models.</p> <p>Propose forecasting method based on spectral analysis that uses the Discrete Fourier Transforms (DFT) which allows for the identification of significant peaks and frequency components in the time series.</p> <p>Estimate stochastic volatility of financial and geophysical time series using GARCH(p,q) which uses Quasi-Newton optimizer to find the Maximum Likelihood Estimates (MLE).</p>
Aug. - Dec. 2021	<p>Bayesian estimation of stochastic volatility (SV) models via MCMC numerical simulation scheme.</p> <p>Investigate Bayesian versus Maximum Likelihood Estimations of SV models.</p> <p>Develop methodology for systems of stochastic volatility (SV) models.</p>
Jan. - Feb. 2022	<p>Apply Statistical Information Criteria to review the model that best fit the financial and geophysical time series.</p>
Mar. 2022	<p>Submit first draft of dissertation</p>
Apr. 2022	<p>Revision of first draft of dissertation</p>
May 2022	<p>Dissertation defense, final revision and submission of dissertation.</p>

Table 9.1: Table of tasks to undertake in the months leading up to dissertation defense



# References

- [1] J. W. Kantelhardt and E. Koscielny-Bunde and H. H. A. Rego and S. Havlin and A. Bunde, "Detecting Long-range Correlations with Detrended Fluctuation Analysis," *Physica A*, 2001, pp. 441–454.
- [2] M. F. Shlesinger and B. J. West and J. Klafter, "Lévy dynamics of enhanced diffusion: Application to turbulence," *Physical Review Letters*, 1987, Vol. 58, pp. 11–16.
- [3] D. Ben-Avraham and S. Havlin, "Diffusion and Reactions in Fractals and Disordered Systems," *Cambridge University Press*, Cambridge, 2000.
- [4] A. Y. Khintchine and P. Lévy, "Sur les lois stables," *C.R. Acad. Sci.*, Paris, 1936, pp. 374–376.
- [5] E. Barany and M. P. Beccar Varela and I. Florescu and I. Sengupta, "Detecting market crashes by analyzing long-memory effects using high-frequency data," *Quantitative Finance*, 2012, Vol. 12, No. 4, pp. 623–634.
- [6] N. Scafetta and P. Grigolini, "Scaling detection in time series: diffusion entropy analysis," *Phys Rev E Stat Nonlin Soft Matter Phys.*, 2002, pp. 66.
- [7] M. C. Mariani and I. Florescu and M. P. Beccar Varela and E. Ncheuguim, "Long correlations and Lévy models applied to the study of memory effects in high frequency (tick) data," 2008.
- [8] M. P. Beccar Varela and H. Gonzalez-Huizar and M. C. Mariani and O. K. Tweneboah, "Lévy Flights and Wavelets Analysis of Volcano-Seismic Data," *Pure and Applied Geophysics, Online first*, 2019
- [9] P. Lévy, "Calcul des probabilités," Paris, 1925.

- [10] N. Scafetta and P. Hamilton and P. Grigolini, "Fractals," 2001, Vol. 9, pp. 193–208.
- [11] M. C. Mariani and M. A. M. Bhuiyan and O. K. Tweneboah and H. Gonzalez-Huizar and I. Florescu, "Volatility models applied to geophysics and high frequency financial market data," *Physica A*, 2018, Vol. 503, No. 1, pp. 304–321.
- [12] M. C. Mariani and H. Gonzalez-Huizar and M. A. M. Bhuiyan and O. K. Tweneboah, "Using Dynamic Fourier Analysis to Discriminate Between Seismic Signals from Natural Earthquakes and Mining Explosions," *AIMS Geosciences*, 2017, Vol. 3, No. 3, pp. 438–449.
- [13] M. C. Mariani and H. Gonzalez-Huizar and M. A. M. Bhuiyan and O. K. Tweneboah, "Forecasting the Volatility of Geophysical Time Series with Stochastic Volatility Models," *World Academy of Science, Engineering and Technology* 2017, Vol. 11, No. 10, pp. 393–399.
- [14] J. M. Kleinberg, "Navigation in a small world," *Nature*, 2000, Vol. 406(6798), pp. 845.
- [15] G. Li and S. D. S. Reis and A. A. Moreira and S. Havlin and H. E. Stanley and J. S. Andrade Jr, "Towards Design Principles for Optimal Transport Networks," *PRL*, 2010, Vol. 104, No. 1.
- [16] D. W. Sims and E. J. Southall and N. E. Humphries and G. C. Hays and C. J. A. Bradshaw and J. W. Pitchford and A. James and M. Z. Ahmed and A. S. Brierley and M. A. Hindell and D. Morritt and M. K. Musyl and D. Righton and E. L. C. Shepard and V. J. Wearmouth and R. P. Wilson and M. J. Witt and J. D. Metcalfe, "Scaling laws of marine predator search behaviour," *Nature*, 2008, Vol. 451 (7182), pp. 109–1102.
- [17] N. E. Humphries and N. Queiroz and J. R. M. Dyer and N. G. Pade and M. K. Musyl and K. M. Schaefer and D. W. Fuller and J. M. Brunnschweiler and T. K. Doyle and

- J. D. R. Houghton and G. C. Hays and C. S. Jones and L. R. Noble and V. J. Wearmouth and E. J. Southall and D. W. Sims, "Environmental context explains Lévy and Brownian movement patterns of marine predators," *Nature*, 2010, Vol. 465 (7301), pp. 106–1069.
- [18] A. Witze, "Sharks Have Math Skills," *discovery.com*, 2013
- [19] J. Dacey, "Sharks hunt via Lévy flights," *physicsworld.com*, 2013
- [20] G. M. Viswanathan and S. V. Buldyrev and S. Havlin and M. G. E. da Luz and E. P. Raposo and H. E. Stanley, "Optimizing the success of random searches," *Nature*, 1999, Vol. 401 (6756), pp. 910–914
- [21] G. Reynolds, "Navigating Our World Like Birds and some authors have claimed that the motion of bees," *The New York Times*, 2014
- [22] D. W. Sims and A. M. Reynolds and N. E. Humphries and E. J. Southall and V. J. Wearmouth and B. Metcalfe and R. J. Twitchett, "Hierarchical random walks in trace fossils and the origin of optimal search behavior," *Proceedings of the National Academy of Sciences*, 2014, Vol. 111 (30), pp. 1107–11078
- [23] Buldyrev et al, "Long-Range Correlation-Properties of Coding And Noncoding DNA-Sequences- Genbank Analysis," *Phys. Rev. E.*, 1995, Vol. 51 (5), pp. 508–5091
- [24] A. Bunde and S. Havlin, "Fractals and Disordered Systems," *Springer*, Berlin, Heidelberg, New York, 1996
- [25] R. Hardstone and S. S. Poil and G. Schiavone and R. Jansen and V. V. Nikulin and H. D. Mansvelder and K. Linkenkaer-Hansen, "Detrended Fluctuation Analysis: A Scale-Free View on Neuronal Oscillations," *Frontiers in Physiology*, January 2012, Vol. 3, pp. 450

- [26] M. Little and P. McSharry and I. Moroz and S. Roberts, "Nonlinear, Biophysically-Informed Speech Pathology Detection," *IEEE International Conference on Acoustics Speed and Signal Processing Proceedings*, 2006, Vol. 2, pp. 1080–1083
- [27] A. Bunde et al, "Correlated and uncorrelated regions in heart-rate fluctuations during sleep," *Phys. Rev. E.*, 2000, Vol. 85 (17), pp. 373–3739
- [28] N. Scafetta, "An entropic approach to the analysis of time series," *University of North Texas*, 2003, pp. 5906–5906
- [29] M. C. Mariani and P. K. Asante and M. A. M. Bhuiyan and M. P. Beccar-Varela and S. Jaroszewicz and O. K. Tweneboah, "Long-Range Correlations and Characterization of Financial and Volcanic Time Series," *Mathematics - MDPI*, 2020, Vol. 8, pp. 441

# Curriculum Vitae

William Kubin was born on June 4, 1990. He graduated The University of Cape Coast in Cape Coast - Ghana in 2015 where he pursued a Bachelor's degree in Mathematics-with-Economics and worked as a Teaching Assistant for an academic year.

He entered The University of Akron in the fall of 2016. He graduated with a Master's degree in Applied Mathematics where he was a Teaching Assistant for two (2) academic years during his studies.

In the fall of 2018, he entered the Graduate School of The University of Texas at El Paso to pursue the Computational Science Program (Ms/PhD).

Email: [wkubin@miners.utep.edu](mailto:wkubin@miners.utep.edu)