

ADVANCED NONLINEAR AND NONSTATIONARY METHODS FOR THE
ANALYSIS OF FINANCIAL TIME SERIES

BY

ADİL YILMAZ

SUPERVISOR

PROF. DR. GAZANFER ÜNAL

PH.D. THESIS

INSTITUTE OF SOCIAL SCIENCES

DEPARTMENT OF FINANCIAL ECONOMICS

YEDİTEPE UNIVERSITY

JUNE, 2020

PLAGIARISM

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Date: 29.06.2020

Name/Surname: ADİL YILMAZ

Signature: 

ÖZET

FİNANSAL ZAMAN SERİLERİNİN ANALİZİNDE İLERİ LİNEER OLMAYAN VE DURAĞAN OLMAYAN YÖNTEMLER

Durağan olmayan ve doğrusal olmayan zaman serileri analizi üzerine geniş bir yelpazedeki ileri konular ve yöntemler bu tezin içeriğini oluşturmaktadır. Bu araştırmalar ve içerikleri şöyle özetlenmektedir. Wavelet coherence metodu FTSE 100 ve S&P 500 ile birlikte çeşitli başlıca Asya borsaları ve gayrimenkul piyasaları üzerinde uygulanarak incelenmektedir. Gelişmiş ekonomi borsalarının Asya borsaları üzerinde güçlü etkileri olduğu ortaya çıkmaktadır. Bu ilişkiler, bireysel zaman serileri ile oluşturulan modellerden daha az tahmin hatası gösteren VARMA esaslı bir tahmin modeli geliştirmek için kullanılmaktadır. Ayrıca, üç zaman serisi arasındaki dinamik korelasyonların analizi için multiple wavelet coherence metodu uygulanmaktadır. FIGARCH(p, d, q) proseslerinin kaosa ilgili özellikleri, doğrudan stokastik fark denklemleri ve simüle zaman serileri üzerinden mutual information, false nearest neighbor metodu, Wolf algoritması, Kantz algoritması ve Jacobian algoritması ile correlation dimension ve en büyük Lyapunov üssü hesaplanarak incelenmektedir. Kaotik gürültülü Duffing harita serileri ve daha sonra aynı özelliklere sahip finansal zaman serileri için yeni bir stokastik modelleme ve tahmin yöntemi geliştirilmektedir. Bu telegraf proses ve fraksiyonel Gauss gürültüsünün kombinasyonu, rejim değiştirme ve long-range bağımlı özelliklere sahip gürültülü Duffing harita serisine ve daha sonra aynı özelliklere sahip gerçek dünya finansal zaman serisine fit edilmektedir. Multiscale correlation dimension (MSCD), zaman serilerinin kaotik özelliklerinin incelenmesinde yeni bir yöntem olarak geliştirilmektedir. Kaotik yapıyı ölçmek için correlation dimension algoritması ve ölçekleme aracı olarak mean filtre ve Gauss filtresi kullanılarak çok ölçekli karakteristikler yakalanabilmektedir. MSCD yöntemine ile, correlation dimension ve Hurst üssü ilişkisi, çeşitli zaman serileri üzerinde çoklu ölçeklerde incelenmektedir.

ABSTRACT

ADVANCED NONLINEAR AND NONSTATIONARY METHODS FOR THE ANALYSIS OF FINANCIAL TIME SERIES

A broad range of advanced subjects and methods for nonstationary and nonlinear time series analysis comprises this thesis. These researches and their contents are summarized as follows. The wavelet coherence is investigated and applied on FTSE 100 and S&P 500 with several major Asian markets and major real estate markets. It is revealed that the developed economy stock markets have strong influences over Asian stock markets. These relationships are used to develop a VARMA based forecast model that shows less forecast errors than the models constructed with individual time series. Moreover, the multiple wavelet coherence is added for the analysis of dynamic correlations between three time series together. The chaos related properties of FIGARCH(p , d , q) processes are examined by computing the correlation dimension and the largest Lyapunov exponent from the stochastic difference equations directly and the simulated time series through mutual information, false nearest neighbour method, Wolf's algorithm, Kantz's algorithm and Jacobian algorithm. A new stochastic modelling and forecasting method for the chaotic noisy Duffing map series and for the financial time series with identical characteristics is introduced. The combination of telegraph process and the fractional Gaussian noise is fitted to the noisy Duffing map series that possesses regime switching and long-range dependent characteristics and then the real-world financial time series with identical features. Multiscale correlation dimension (MSCD) is proposed as a new method for chaotic properties of time series. Based on MSCD method, fractal dimension and Hurst exponent relationship is investigated in multiple scales on several time series.

ACKNOWLEDGEMENTS

I am most grateful to my supervisor Prof. Dr. Gazanfer Unal for his guidance, inspiration and friendship.

I would like to thank my contributors and the members of my committee, Prof. Dr. Ilyas Akhisar and Associate Prof. Dr. Sema Dube.

Finally, I would like to express my gratitude to my dear family for their endless love and support.

TABLE OF CONTENTS

PLAGIARISM	i
ÖZET	ii
ABSTRACT	iii
ACKNOWLEDGEMENTS	iv
LIST OF TABLES	ix
LIST OF FIGURES	x
LIST OF SYMBOLS/ABBREVIATIONS	xv
1. INTRODUCTION	1
2. BACKGROUND	7
2.1. Time Series Analysis	7
2.1.1. Linear Time Series Analysis	8
2.1.1.1. Random Variables.....	8
2.1.1.2. Stochastic Processes.....	8
2.1.1.3. Stationarity	8
2.1.1.4. Autocovariance Function	9
2.1.1.5. Autocorrelation Function	10
2.1.1.6. Gaussian Processes	10
2.1.1.7. Linear Processes.....	11
2.1.1.8. Moving Average Processes.....	11
2.1.1.9. Autoregressive Processes	12
2.1.1.10. ARMA Processes	13
2.1.1.11. ARIMA Process	13
2.1.1.12. ARFIMA Process	14
2.1.2. Nonlinear Time Series Analysis.....	14
2.1.2.1. ARCH	14
2.1.2.2. GARCH.....	15
2.1.2.3. IGARCH.....	16
2.1.2.4. FIGARCH.....	16
2.1.2.5. Long Memory Processes.....	16
2.1.2.6. Fractals.....	18

2.1.2.7. Chaos	21
2.1.2.8. Wavelet Analysis	21
2.2. Objectives of the thesis	22
2.2.1. Structure of the Thesis	23
2.2.2. Published Papers	23
3. WAVELET COHERENCE ANALYSIS	25
3.1. Wavelet Analysis	26
3.1.1. The Cross Wavelet Transform	27
3.1.2. Phase Angle	28
3.1.3. The Cone of Influence	28
3.1.4. Wavelet Coherence	29
3.2. Co-movement Analysis of Asian Stock Markets against FTSE100 and S&P 500	29
3.2.1. The Wavelet Coherence Results	31
3.3. Co-movement Analysis of Major Real-estate Markets	47
3.3.1. The Wavelet Coherence Results	48
3.3.2. Forecast with Multivariate ARMA Process	49
3.4. Co-movement Analysis of Major Real-estate Markets by Multiple Wavelet Coherence	54
3.4.1. Multiple Wavelet Coherence	54
3.4.2. Multiple Wavelet Coherence Results	58
3.4.3. The Comparison of MWC and WTC Results	60
3.4.4. Forecast with Multivariate ARMA Process	62
3.5. Final Remarks	63
4. NONLINEAR TIME SERIES ANALYSIS	65
4.1. The Chaoticity Properties of FIGARCH	65
4.1.1. FIGARCH	66
4.1.2. Chaotic Behavior and FIGARCH	68
4.1.3. Phase Space Reconstruction	68
4.1.4. Mutual Information	69
4.1.5. False Nearest Neighbor Method	70
4.1.6. Correlation Dimension	72

4.1.7. Lyapunov Exponent.....	76
4.1.7.1. Wolf's Algorithm	76
4.1.7.2. Kantz's Algorithm	77
4.1.7.3. Direct Calculation with the 2-Dimensional Map.....	80
4.1.8. Final Remarks	82
4.2. Stochastic Duffing Equation in Modelling Financial Time Series	83
4.2.1. The Noisy Duffing Map	87
4.2.1.1. Wavelet Decomposition and Wavelet Scalogram	88
4.2.1.2. MFDFA Analysis	90
4.2.1.3. The Noisy Duffing Map and Its Regime Switching Properties	92
4.2.2. Modelling with Telegraph Process and Fractional Gaussian Noise	93
4.2.2.1. The Noisy Duffing Series Estimation by Telegraph Process	93
4.2.2.2. Fractional Gaussian Noise Process Estimation	95
4.2.3. The Demonstration on the Real-World Time Series.....	98
4.2.3.1. Wavelet Scalogram and Multifractality of the Smoothed Probabilities for Excess Returns.....	99
4.2.3.2. Telegraph Process and fGn Estimation	101
4.2.3.3. The Comparison of the Results of the Analyses of the Noisy Duffing Series and the Smoothed Probabilities for Excess Returns	104
4.2.4. Forecast Results.....	108
4.2.4.1. The Noisy Duffing Series Forecast Results for the Range of 50-step ahead	108
4.2.4.2. The Smoothed Probabilities of Excess Returns Forecast Results for the Range of 50-step ahead	110
4.2.4.3. The Noisy Duffing Series Forecast Results for the Range of 100-step ahead	112
4.2.4.4. The Smoothed Probabilities of Excess Returns Forecast Results for the Range of 100-step ahead.....	112
4.2.4.5. The Noisy Duffing Series Forecast Results for the Range of 25-step ahead	113

4.2.4.6. The Smoothed Probabilities of Excess Returns Forecast Results for the Range of 25-step ahead	114
4.2.5. Final Remarks	115
4.3. Multiscale Correlation Dimension Method	115
4.3.1. The Details of the Method	117
4.3.1.1. Scaling Filters	117
4.3.1.2. Correlation Dimension	123
4.3.2. Applications.....	126
4.3.2.1. Chaotic Systems	126
4.3.2.2. Stochastic Processes.....	127
4.3.2.3. Financial Time Series Models	128
4.3.3. The Relationship between Correlation Dimension and Hurst Exponent in Multiscale	132
4.3.4. Final Remarks	138
5. CONCLUSIONS	139
REFERENCES	143

LIST OF TABLES

3.1	Details of stock exchange pairs	30
4.1	The embedding delay for each FIGARCH model	70
4.2	The largest Lyapunov exponent results for all the FIGARCH models with Wolf's algorithm.....	78
4.3	The largest Lyapunov exponent results for all the FIGARCH models with Kantz's algorithm	80
4.4	LLE results for all the FIGARCH models with the three different method .	82

LIST OF FIGURES

3.1	Wavelet coherence results of S&P 500 - AXJO and FTSE - AXJO	32
3.2	Wavelet coherence results of S&P 500 - AXPJ and FTSE - AXPJ	33
3.3	Wavelet coherence results of S&P 500 - BIST and FTSE - BIST	35
3.4	Wavelet coherence results of S&P 500 - HSI and FTSE - HSI.....	36
3.5	Wavelet coherence results of S&P 500 - IDX and FTSE - IDX	37
3.6	Wavelet coherence results of S&P 500 - KL and FTSE - KL	38
3.7	Wavelet coherence results of S&P 500 - KLSE and FTSE - KLSE.....	40
3.8	Wavelet coherence results of S&P 500 - KS11 and FTSE - KS11	41
3.9	Wavelet coherence results of S&P 500 - N225 and FTSE - N225	42
3.10	Wavelet coherence results of S&P 500 - RTS and FTSE - RTS	43
3.11	Wavelet coherence results of S&P 500 - Shenzen and FTSE - Shenzen ...	44
3.12	Wavelet coherence results of S&P 500 - TW and FTSE - TW	46
3.13	Wavelet coherence result of IYR-REF UN TO.	48
3.14	Wavelet coherence result of IYR-IFEU.	49
3.15	Wavelet coherence result of IYR-IFGL.	50
3.16	Forecast result of IYR-IFEU following the most coherent area.....	52
3.17	Forecast result of IYR-IFEU following the least coherent area.	52
3.18	Forecast result of IYR-IFGL following the most coherent area.....	53
3.19	Forecast result of IYR-IFGL following the least coherent area.	54

3.20	MWC results of IYR with IFGL and LINK.	58
3.21	MWC results of IFGL with IYR and LINK.	59
3.22	MWC results of LINK with IYR and IFGL.	60
3.23	The comparison of WTC (IYR-IFGL) and MWC (LINK-(IYR-IFGL)) results	61
3.24	The comparison of WTC (IYR-LINK) and MWC (IFGL-(IYR-LINK)) results	61
3.25	The comparison of WTC (IFGL-LINK) and MWC (IYR-(IFGL- LINK)) results.....	62
3.26	The forecast results of IYR-IFGL-LINK after the most coherent region ..	63
3.27	The forecast results of IYR-IFGL-LINK after the least coherent region...	64
4.1	Mutual information result for FIGARCH $d=0.90$	71
4.2	FNN embedding dimension result for FIGARCH $d=0.90$	73
4.3	Correlation dimension results of FIGARCH $d=0.90$ for embedding delays (1-20) and embedding dimensions (1-20)	74
4.4	Correlation dimension results for each FIGARCH model according to the embedding delays determined by mutual information.....	75
4.5	The best converged lines of correlation dimension results for each FIGARCH model.....	75
4.6	The largest Lyapunov exponent result for FIGARCH $d=0.90$	77
4.7	The largest Lyapunov exponent result with Kantz's algorithm for FIGARCH $d=0.90$	79
4.8	The largest Lyapunov exponent result by directly with the 2- dimensional map for FIGARCH $d=0.80$	81

4.9	The noisy Duffing map	88
4.10	The noisy Duffing map wavelet scalogram and decomposed scales	89
4.11	MFDFA results of chaotic maps. (Henon = 0.3077, ikeda = 0.1436, chirikov = 0.0067, multifractal = 1.0304, monofractal = 0.7642, white noise = 0.4787).....	90
4.12	The noisy Duffing series with standard deviation of 0.01 and the 8000 number of data points. (Scale 2 = 0.3783, scale 4 = 0.7531, scale 6 = 1.4271, scale 8 = 2.3105, multifractal = 1.0304, monofractal = 0.7642, white noise = 0.4787)	91
4.13	The noisy Duffing series versus the modified series as to telegraph process	94
4.14	Telegraph process prediction for the noisy Duffing series	95
4.15	The average value of all estimated telegraph process paths for the noisy Duffing series.	95
4.16	The noisy section of the noisy Duffing series and fGn estimation.	96
4.17	The estimated series as a combination of telegraph process and fGn for the noisy Duffing series.....	97
4.18	The probability distributions of the original Duffing series and the series estimated with telegraph process and fGn.....	98
4.19	Smoothed probabilities of the medium variance regime for excess stock returns with its telegraph process representation.	99
4.20	Wavelet scalogram and decomposed scales of the smoothed probabilities of excess returns.	100
4.21	Multifractality results of smoothed probabilities for stock returns. (scale 2 = 1.2170, scale 4 = 1.1635, scale 6 = 2.1206, scale 8 = 3.0002, multifractal = 1.0304, monofractal = 0.7642, white noise = 0.4787).....	101

4.22	Telegraph process prediction for the smoothed probabilities of excess returns.....	101
4.23	The average value of all estimated telegraph process paths for the smoothed probabilities of excess returns.	102
4.24	The noisy section of the smoothed probabilities of excess returns and fGn estimation.....	102
4.25	The estimated series as a combination of telegraph process and fGn for the smoothed probabilities of excess returns.	103
4.26	The probability distributions of the original smoothed probabilities of excess returns and the series estimated with telegraph process and fGn. ..	103
4.27	Comparison of the results of probability distributions of estimated noisy Duffing series and the smoothing probabilities of excess return series.	104
4.28	MFDEFA results of the noisy Duffing series and the smoothed probabilities for excess returns	105
4.29	Telegraph process representations of the noisy Duffing series and the smoothed probabilities of excess returns with the original series.	106
4.30	The noise-like section with the estimated fGn results for the noisy Duffing series and the smoothed probabilities of excess returns.....	106
4.31	Telegraph process and fGn estimations of the noisy Duffing series and the smoothed probabilities of excess returns.	107
4.32	Histogram results of the noisy Duffing series and the smoothed probabilities of excess returns.	107
4.33	The average path of the 50 random processes generated for the noisy Duffing series.	109
4.34	Telegraph process forecast of the noisy Duffing series.	109

4.35	Telegraph process and fGn estimation through forecast range with the original noisy Duffing series.	110
4.36	The average path of the 50 random processes generated for the smoothed probabilities of excess returns.	110
4.37	Telegraph process forecast of the smoothed probabilities of excess returns.	111
4.38	Telegraph process and fGn estimation through forecast range with the smoothed probabilities of excess return series.	111
4.39	Telegraph process and fGn estimation for the forecast range of 100-step ahead for the noisy Duffing series.	112
4.40	Telegraph process and fGn estimation for the forecast range of 100-step ahead for the smoothed probabilities of excess returns.	113
4.41	Telegraph process and fGn estimation for the forecast range of 25-step ahead for the noisy Duffing series.	113
4.42	Telegraph process and fGn estimation for the forecast range of 25-step ahead for the smoothed probabilities of excess returns.	114
4.43	Multiscale overlapping coarse-graining algorithm	118
4.44	MSCD results of All Filters	124
4.45	Multiscale correlation dimension results of chaotic maps	128
4.46	Multiscale correlation dimension results of stochastic processes	129
4.47	Multiscale Correlation Dimension Results of FARIMA	130
4.48	Multiscale correlation dimension results of FIGARCH.	132
4.49	MSCD with Gaussian filter results for specific time series.	136
4.50	The relationship between D and H for specific time series.	137

LIST OF SYMBOLS/ABBREVIATIONS

t	Time
ε	Gaussian White Noise
μ	The Mean
σ	Standard Deviation
σ^2	Variance
γ	Autocovariance Function
N	Normal Distribution
E	Expected Value
S	Sample Space
W	Wavelet Transform
ω	Frequency
$*$	Complex Conjugate
I	Imaginary Part of Power Spectrum
R	Real Part of Power Spectrum
S	Smoothing Operator
I_m	Identity Matrix
M	Matrix
S	Wavelet Spectra
$\det$	Determinant of a Matrix
τ	Embedding Delay
d	Embedding Dimension
P	Probability Distribution
Θ	Heaviside Step Function
$ x $	Absolute Value
$\ x - y\ $	Euclidean Distance
V	Velocity
λ	The Rate of a Poisson Process
H	Hurst Exponent

D	Fractal Dimension
ARCH	Autoregressive Conditionally Heteroscedasticity
GARCH	Generalized Autoregressive Conditionally Heteroscedasticity
IGARCH	Integrated Generalized Autoregressive Conditionally Heteroscedasticity
FIGARCH	Fractionally Integrated Generalized Autoregressive Conditionally Heteroscedasticity
ACF	The Autocorrelation Function
LRD	Long-range Dependence
fBm	Fractional Brownian Motion
fGn	Fractional Gaussian Noise
MA	Moving Average Process
AR	Autoregressive Process
ARMA	Autoregressive Moving Average Process
ARIMA	Autoregressive Integrated Moving Average Process
ARFIMA	Autoregressive Fractionally Integrated Moving Average Process
R/S	Rescaled Adjusted Range Statistic
MFDFA	Multifractal Detrended Fluctuation Analysis
CWT	Continuous Wavelet Transform
WTC	Wavelet Coherence
XWT	Cross Wavelet Transform
MWT	Multiple Wavelet Transform
PWC	Partial Wavelet Coherence
REIT	Real Estate Investment Trust
ETF	Exchange Traded Fund
IYR	iShares US Real Estate Exchange Traded Fund
IFGL	iShares International Developed Real Estate Exchange Traded Fund
IFEU	iShares Europe Developed Real Estate Exchange Traded Fund
CREIT	Canadian Real Estate Investment Trust
LINK	Real Estate Investment Trust based on Hong Kong and China

MLE	Maximal Lyapunov Exponent
LLE	The largest Lyapunov Exponent
Ht	Local Hurst Exponent
Ph	The Probability Distribution of Local Hurst Exponent
Dh	Multifractal Spectrum
Pdf	Probability Density Function
MSCD	Multiscale Correlation Dimension



1. INTRODUCTION

”When the weather changes, nobody believes the laws of physics have changed. Similarly, I don’t believe that when stock market goes into terrible gyrations, its rules have changed.”

Benoit Mandelbrot

Nonlinear time series analysis and modelling have been attracting attention in recent decades. Advanced methods have been developed in finance as well as in wide range of other disciplines with the discovery of the dynamic properties of time series. However, the application of these models in real-life problems is still not very mature. Nonlinear time series analysis algorithms are relatively harder to apply than linear time series analysis algorithms. Nonlinear methods which are still young necessitate more knowledge and experience due to complexity and variety of methods and data sets (Michael, 2005). Also, because nonlinearity area is huge, it is quite a challenge to build general theories of it.

Uncertainty in financial markets and economy affects all aspects of life and the people by wrong decisions, inefficiencies and crisis (Hafner, 2013). Modelling and prediction of economic and financial data and time series have been trying to solve these problems and new methods have constantly been developed. Even though, there are many methods developed for different types of data, the success of predictions through them has long been in question since Louis Bachelier’s ground breaking work in which financial time series were described as the collection of independent and identically distributed random variables (Bachelier, 1900).

Normal distribution assumption is the base of many fundamental methods. This assumption is unrealistic but necessary for applicable methods. Besides, it maybe the weakest block in the prediction success of those methods (Hafner, 2013). Taking into account the observations in real-world, it may not be too much to say that nonstationarity comes as

default, so the stationarity becomes exception. That leads to that the tools developed with the assumption of stationarity loses its sense for the analysis of the nonstationary time series.

Linear methods read steady and symmetrical structure of the data. Therefore, any inconsistency in the system is regarded as being caused by a random effect outside. However, it was discovered that even the very simple deterministic equations can create irregularity in itself as result of nonlinear and chaotic influences. So, irregularities in a system can be both attributed to random or/and nonlinear effects (Kantz and Schreiber, 2004).

Linear models almost always depend on the assumption of that returns of financial assets exhibit random walk type behavior. Also, time series are accepted to be stationary or transformed to be stationary. However, many time series are not be able to transformed straightforwardly (De Gooijer et al., 2017). Financial time series mostly show nonlinear features like nonstationarity, aperiodicity, time irreversibility, asymmetric cycles and so on require nonlinear models. These are observed in financial time series like occurrence of large returns more often than expected, large stock markets returns being often negative, the tendency of large returns existing in clusters or the appearance of large volatility preceded by large negative returns (Franses et al., 2000).

One of the nonlinear model named autoregressive conditionally heteroscedasticity (ARCH) model came to existence with the realization of importance of volatility in financial markets in 1980s. Deficiency of linear methods depending on Gaussian processes to model high variability gave rise to dominance of ARCH based models which succeeded in modelling and forecast of volatility. Then, in 1986, generalised ARCH (GARCH) model which adds previous lags of the conditional variance to the equation was introduced (Bollerslev, 1986).

The variety of GARCH models has been developed since in order to fit in with and depict different features of financial time series like excess kurtosis, fat-tailedness as well as volatility clustering. So, these GARCH-type models have found great interest among practitioners in fast-growing finance sector. Several GARCH extensions which namely are EGARCH, QGARCH, TGARCH, GJR-GARCH and so on have been developed to

detect attributes like the asymmetric effect of positive and negative shocks on volatility and potential relation between the return and volatility which the standard GARCH model are not capable of identifying (Franses et al., 2000).

Furthermore, fractionally integrated GARCH (FIGARCH) was introduced in 1996 to capture long memory and persistence properties in volatility which is observed in economy and capital markets (Baillie et al., 1996). The model allows to mirror the observed autocorrelation patterns of empirical time series like hypergeometrical decay. The model's practical implementation also was showed by applying it to exchange rates, stock returns and option prices.

In capital markets, long memory effects were recognized and after 1990s, the research on the long-range dependence (LRD) properties of financial time series has grown quickly. Financial assets like stock prices, commodity prices, foreign exchange rates have been shown as having the characteristics of fractional process, long-range dependence or self-similarity. The long-range dependence feature exist in a process in case that the autocorrelations fade very slowly, without their sum converging.

Quantifying the Hurst exponent (H) which was introduced by Hurst (Hurst, 1951) is the main method for the analysis of LRD. H has a link to the notion of Brownian motion or Wiener process which have independent increments. But, in Hurst's groundbreaking work, it was shown that successive increments were not independent. As a result, the fractional Brownian motion (fBm) as a generalization of Brownian motion was developed. It is an example of self-similar process whose increments exhibit LRD (long-term memory or persistency) (Tarnopolski, 2016).

ARFIMA has also been a popular process for modelling and analysis of LRD. ARFIMA was introduced by Granger and Joyeux (1980) as a generalization of ARMA model (Granger and Joyeux, 1980). As LRD processes also closely related to fractional calculus (Liu et al., 2017), based on Hurst's analysis, ARFIMA and FIGARCH has been developed for certain financial time series and found quite success among reseachers and practitioners.

Long memory effect suggests that what happens today in the market may effect the future actions and move for a long period of time. Even, it may cause nonperiodic cycles lasting for many years like in nonlinear dynamical systems or deterministic chaotic systems. Deterministic chaos describes irregular temporal behavior and anomalies in nonstochastic systems (Kantz and Schreiber, 2004). Nonlinear, chaotic systems possess very irregular data which are not periodic or not totally independent but consist of unique mathematical structure which can be revealed for enhanced prediction. Chaos theory has found strong use for the analysis of real-world time series in terms of dynamical systems over the recent decades, especially, whenever the traditional methods were not able to capture complex structures of time series among wide range of fields.

In deterministic chaotic systems, even the smallest errors are amplified by making long term forecasts beyond the predictability horizon impossible. Despite that, the predictions in short term to some degree is achievable with the developed methods (Kantz and Schreiber, 2004). One popular example on that is the weather forecasting. It wouldn't matter however many weather stations around the world is built, it would still be impossible to expand the accepted three day forecast horizon. First, Lorenz in 1963 studied the unpredictable nature of the weather and showed the sensitivity to initial conditions in the atmosphere as well as the existing structure in chaos.

By using numerical and geometrical analysis, plotting trajectories in three dimension and revealing a complicated set with the infinite complex of surfaces which is regarded as an example of an fractal (Lorenz, 1963). Fractals have self-similar structures at all scales. When zoomed, patterns in smaller scales look similar to the ones in larger scales at some degree which can be calculated by statistical methods like fractal dimension, correlation dimension, lyapunov exponents etc. There is also multifractal theory dealing with fractal objects which are not be able to be characterized by a single fractal dimension.

It is obvious that chaos is a genuine characteristic of a natural systems and fractals and chaos relate to each other which are special notions changed the view of the world. Reseachers and practioners sometimes perceive that the financial markets behave quite random and sometimes act rather deterministic. When the forecasts become extremely difficult with

the commonly used linear tools, total randomness is accepted as an explanation. Nonlinear methods and chaos theory could bring new approaches for understanding how markets and economies function and for identification, modelling and forecasting certain type of financial time series where the traditional linear methods are inadequate to.

Wavelet analysis also had a great advancement in mid-1980s with the works of Daubechies (Daubechies, 1988) who introduced a smooth, compactly supported and orthonormal wavelet basis and Grossmann and Morlet (Grossmann and Morlet, 1984) who studied continuous wavelet transform and applied it on geological data. Meyer constructed second orthogonal wavelet in 1985 (Daubechies et al., 1986) and the concept of multiresolution with Mallat (Mallat, 1989) who also introduced new approach to understanding and computing orthonormal basis.

As a result of these developments in wavelet theory, it became possible to analyze, transform and decompose signals both in time domain and frequency domain. Wavelet applications are especially beneficial over conventional methods because of not having stationarity requirements. The real-world systems are rarely stationary. To be able to deliver reliable forecast outcomes, empirical research must take into consideration nonstationarity. So, researchers and practitioners in many different areas started making use of these useful features of wavelet techniques for the applications in image compressing, control systems, finance, geophysics, biomedical, mechanical engineering and so on. In recent years, in finance and economy, wavelet applications notably become very valuable for exploring dynamic correlations and causality relationships between multiple time series.

Mathematical models are composed to illustrate physical world through observation and data gathering. Cutting-edge mathematical modelling methods and technology have continuously been researched, developed and produced. There has been rising interest and need for development of useful nonlinear techniques and effective nonlinear time series models.

Increases in computer power has been allowing research in nonlinear modelling and applications accelerate in recent decades. Also, advancements in software such as R,

Matlab, Mathematica, Python etc. and availability of extensive data collections supported the application of nonlinear time series analysis.

These advancements in mathematical modelling and algorithms especially in financial markets have been replacing many financial transactions in a large scale in most developed countries. Nowadays, 80% of transactions are executed by advanced algorithm in humanly impossible speeds in FTSE and NYSE.



2. BACKGROUND

2.1. TIME SERIES ANALYSIS

Physical world constantly changes throughout time, as observed serially. Information is gathered in a wide variety of time scales measured at successive time intervals. This finite set of values, called time series plays an significant role in wide range of research areas in modern science. In time series notation, given N number of observations, $(x_1, x_2, \dots, x_N)$ usually is used to represent scalar data points at $(t = 1, 2 \dots N)_{t \in \mathbb{R}}$ instances.

Time series analysis intends to expose fundamental characteristics of observed systems. These systems are usually considered stochastic process and modelled by linear Gaussian stochastic processes. Most recently, deterministic dynamical systems and models have started drawing more attention and interest as a result of increasing awareness of importance of nonlinearity and need for a model for describing natural dynamical systems and available robust methods developed like phase space reconstruction, fractal dimension, lyapunov exponents etc. In fact, in real-world time series, stochastic time series often are seen comprising deterministic element like in deterministic time series contaminated with a measurement noise.

The theory of time series was built on the acceptance of different type of stationarity. But again, physical world data rarely satisfies this assumption. In a stationary process, the mean and the variance stay the same over time. Empirical evidence shows that this is seldom the case for most of the economic, financial and other real-world time series observed. Systems always evolve and change at different scales.

For the forecast of future events, from dependent past and present events, first option is the traditional probabilistic approach which uses theory of probability and mathematical statistics. Another approach is dynamical one which is based on the theory of determinism. As, forming on regularity and causation, each event or cause lead to move generate stimulate the instance of another event or reaction (Bezruchko and Smirnov, 2010). In

following section, we introduce the most commonly used definitions, concepts and models of probabilistic approach.

2.1.1. Linear Time Series Analysis

2.1.1.1. Random Variables

Random variable is a variable which has a connection with a set of possible events or the sample space S . Random variables can be grouped as discrete random variables and continuous random variables. Discrete random variable is a function $X : S \rightarrow R$ supposing that for each interval $I \subseteq R$, $\{X \in I\}$ is an event (Ghahramani, 2018).

Continuous random variable is a random variable which have a probability density of;

$$P(X \leq x) = \int_{-\infty}^x f(x)dx \quad (2.1)$$

where $f(x) \geq 0$, $X \in R$ and $P(X < \infty) = \int_{-\infty}^{\infty} f(x)dx = 1$ (Bremaud, 2020).

2.1.1.2. Stochastic Processes

Stochastic process is set of random elements $\{X_t : t \in T\}$ which change over time, on the probability space $(\Omega, \mathcal{F}, P)$ with the values in measurable space (X, Σ) . Here, Ω , $\mathcal{F}$ and P denotes sample space, σ -field and probability measure on $\mathcal{F}$, consecutively (Douc et al., 2014). Stochastic processes can be in discrete time or continuous time. Besides, variables of X_t can have values with a definite probability in a discrete or continuous states as comprising four possible groups in total (Cochard, 2019).

2.1.1.3. Stationarity

As mentioned previously, although the real-world events are usually not stationary, the stationarity is the foundation of time series analysis. The weak stationarity or second-order stationarity that is usually default use for stationarity in many sources is implied if

its expected value or mean and autocorrelation function stay the same when time is shifted.

If for all $t, h \in \mathbb{Z}$;

$$E(X_t) = \mu \quad (2.2)$$

$$\text{cov}(X_t, X_{t+h}) = \gamma_t$$

where μ is constant and γ_t is independent of t .

A process is strictly stationary if for all $n \in \mathbb{N}^*$, $(X_{t_1}, \dots, X_{t_n})$ and $(X_{t_1+h}, \dots, X_{t_n+h})$ have the same distribution for all sets of time points $t_1, \dots, t_n \in T^n$ and $h \in \mathbb{Z}$.

2.1.1.4. Autocovariance Function

The autocovariance function is used to express the magnitude of linear relationship between random variables. It is the series of covariances of a stationary process. Letting γ_x is the autocovariance function of a weak stationary stochastic process $\{X_t, h, t \in \mathbb{Z}\}$;

$$\gamma_x(t+h, t) = \text{Cov}(X_{t+h}, X_t) = \text{Cov}(X_h, X_0) = \gamma_x(h, 0) \quad (2.3)$$

As it is seen here, autocovariance function is not dependent on t but on the time difference h . So, the autocovariance function $\gamma_x(t+h, t)$ is usually written with one variable as $\gamma_x(h)$.

There are some essential properties of autocovariance function as;

- 1) $\gamma_x(0) \geq 0$
- 2) $\gamma_x(h) = \gamma_x(-h) \forall h \in \mathbb{Z}$
- 3) $|\gamma_x(h)| \leq \gamma_x(0)$
- 4) $\gamma_x(\cdot)$ is non-negative function.

5) $\sum_{i=1}^n \sum_{j=1}^n \alpha_i \gamma_x(i-j) \alpha_j \geq 0$ for all positive integers n and vectors $\alpha = (\alpha_1, \dots, \alpha_n)' \in \mathbb{R}^n$ and time points $(t_1, \dots, t_n) \in \mathbb{Z}^n$ (Douc et al., 2014).

2.1.1.5. Autocorrelation Function

The autocorrelation function for a weakly stationary time series X_t is defined as;

$$\rho_h = \frac{Cov(X_t, X_{t-h})}{\sqrt{Var(X_t)Var(X_{t-h})}} = \frac{Cov(X_t, X_{t-h})}{Var(X_t)} = \frac{\gamma_x(h)}{\gamma_x(0)} \quad (2.4)$$

The autocovariance function can be estimated by sample autocorrelation function;

$$\hat{\gamma}_x(h) = \frac{1}{n} \sum_{t=1}^{n-h} (X_{t+h} - \bar{X})(X_t - \bar{X}) \quad (2.5)$$

where the sample mean $\bar{X} = \frac{1}{n} \sum_{t=1}^n X_t$ and $h = 0, 1, \dots, n-1$. Therefore, the sample autocorrelation function can similarly be written as;

$$\hat{\rho}_h = \frac{\hat{\gamma}_x(h)}{\hat{\gamma}_x(0)} \quad (2.6)$$

2.1.1.6. Gaussian Processes

A time series $\{X_t, t \in \mathbb{Z}\}$ is called Gaussian process if the joint probability distribution of its vectors with the n dimension $\{X_{t_1}, \dots, X_{t_n}\}$ for any time points $(t_1, \dots, t_n)$ and for any $\{n \in \mathbb{N}^*\}$ is multivariate normal distribution.

The multivariate density function is described as;

$$p_n(x; \mu, V) = \frac{1}{\sqrt{(2\pi)^n |V_n|}} \exp\left(-\frac{1}{2}(x_n - \mu_n)' \cdot V_n^{-1} \cdot (x_n - \mu_n)\right) \quad (2.7)$$

where $n \times n$ mean vector is $\mu = E(X) = (\mu_{t_1}, \dots, \mu_{t_n})$, $n \times n$ the autocovariance function

is $V_n = Cov(X) = \{\gamma(t_i, t_j); i, j = 1, \dots, n\}$, $|\cdot|$ is determinant of a matrix. In a case of weakly stationarity, the vector μ and V_n are independent of time because $\mu_t = \mu$ and $\gamma(t_i, t_j) = \gamma(|t_i - t_j|)$.

2.1.1.7. Linear Processes

A linear process $\{X_t, t \in \mathbb{Z}\}$ is defined to be a linear combination of white noise which is a series of independent and identically distributed random variables with finite mean and variance. If $\epsilon_t \sim iid N(0, \sigma^2)$ then ϵ_t is called Gaussian white noise. The autocorrelation function (ACF) of a white noise series is zero. Linear time series can be written as;

$$X_t = \mu + \sum_{i=0}^{\infty} a_i \epsilon_{t-i} \quad (2.8)$$

where μ is the mean of X_t , a_i denotes the sequence of coefficients and $\epsilon_t \sim iid N(0, \sigma^2)$. If X_t is weakly stationary time series, its mean is $E(X_t) = \mu$, its variance is $Var(X_t) = \sigma_\epsilon^2 \sum_{i=0}^{\infty} a_i^2$ and its autocorrelation function can be calculated with $\gamma_x(h) = \sigma_\epsilon^2 \sum_{j=0}^{\infty} a_j a_{j+h}$. Financial time series are rarely linear or Gaussian (Tsay and Chen, 2018).

2.1.1.8. Moving Average Processes

Moving average (MA) model which simply is a linear sequence of weighted white noise process is written as;

$$X_t = \mu + \sum_{i=1}^q \theta_i \epsilon_{t-i} + \epsilon_t \quad (2.9)$$

where ϵ_t is a white noise process with the properties of $E(\epsilon_t) = \mu$ and $var(\epsilon_t) = \sigma^2$.

Some main properties such as the mean, the variance and autocorrelation function of moving average process of order q , $MA(q)$ are;

$$E(X_t) = \mu \quad (2.10)$$

$$Var(X_t) = \sigma^2 + \sum_{j=1}^q \theta_j^2 \sigma^2 \quad (2.11)$$

$$\gamma_h = \begin{cases} \sigma^2 \sum_{i=0}^{q-|h|} \theta_i \theta_{i+|h|}, & |k| \leq q \\ 0, & |k| > q \end{cases} \quad (2.12)$$

2.1.1.9. Autoregressive Processes

Autoregressive process of order p , in short $AR(p)$ is defined as;

$$X_t = \sum_{i=1}^p \alpha_i X_{t-i} + \epsilon_t \quad (2.13)$$

where ϵ_t is the white noise with mean zero and the variance σ_ϵ^2 .

The equation which is the simple regression on the lag of X_t tells the name of autoregression.

The mean of general $AR(p)$ model is calculated by;

$$E(X_t) = \frac{\alpha_i}{1 - \alpha_1 - \alpha_2 - \dots - \alpha_p} \quad (2.14)$$

If the X_t is stationary, then the autocorrelation function is;

$$\rho_h = \sum_{i=1}^p \alpha_i \rho_{h-i} \quad (2.15)$$

where $\rho_h = \frac{\gamma_h}{\gamma_0}$ and $\gamma_h = \sum_{i=1}^p \alpha_i \gamma_{h-i}$ for $h > 0$. There is also a stationary condition for autoregressive processes. For any complex plane z , the roots of autoregressive polynomial $1 - \alpha_1 z - \alpha_2 z^2 - \dots - \alpha_p z^p = 0$ lie outside the unit circle in the complex plane (Brooks, 2008).

2.1.1.10. ARMA Processes

General autoregressive moving average process $ARMA(p, q)$ is given as;

$$X_t = \sum_{i=0}^p \alpha_i X_{t-i} + \sum_{j=0}^q \beta_j \epsilon_{t-j} \quad (2.16)$$

where ϵ_t is white noise process with $N \sim (0, \sigma^2)$, the autoregressive and moving average orders $\{p, q \in \mathbb{Z}^{+}\}$.

If the ARMA process is weakly stationary which is dependent on the condition of characteristic equation solutions to be less than 1, then the unconditional mean of the model is obtained by (Tsay and Chen, 2018);

$$E(X_t) = \frac{\alpha_0}{1 - \alpha_1 - \dots - \alpha_p} \quad (2.17)$$

2.1.1.11. ARIMA Process

Autoregressive integrated moving average model (ARIMA) is a modification of autoregressive process (AR) and moving average process (MA) models which allows differencing. $ARIMA(p, d, q)$ as a combination of $AR(p)$ and $MA(q)$ models is described as follows (Box et al., 2015);

$$(1 - \sum_{i=1}^p \phi_i B^i)(1 - B)^d (X_t - \mu) = (1 + \sum_{i=1}^q \theta_i B^i) \epsilon_t \quad (2.18)$$

where $(1 - B)^d$ is the difference operator as d takes integer values. B represents the backshift operator. It works in the notation as $B^i X_t = X_{t-i}$. ARIMA model is especially powerful when it comes to short-range dependence (Sheng et al., 2011).

2.1.1.12. ARFIMA Process

When d is allowed to take fractional values, it is suggested that the model becomes better at capturing long-range dependence (Granger and Joyeux, 1980). Then, as denoting the autoregressive order, the difference coefficient and the moving average order with p , d and q , in general form ARFIMA(p, d, q) process which can be described as;

$$\phi(B)(1 - B)^d X_t = \theta(B)\epsilon_t \quad (2.19)$$

where $d \in (-0.5, 0.5)$, $\phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$, $\theta(B) = 1 + \theta_1 B + \dots + \theta_q B^q$ and $(1 - B)^d$ is the fractional difference operator.

2.1.2. Nonlinear Time Series Analysis

2.1.2.1. ARCH

Volatility that is conditional variance of the asset value is a very important element of financial time series in terms of modelling and forecasting (Tsay and Chen, 2018). The autoregressive heteroscedastic conditional model of order p , ARCH(p) of Engle (1982)(Engle, 1982) is defined as;

$$\begin{aligned} X_t &= \sigma_t \epsilon_t \\ \sigma_t^2 &= \alpha_0 + \sum_{i=1}^p \alpha_i X_{t-i}^2 \end{aligned} \quad (2.20)$$

where $\alpha_0, \alpha_p > 0, \alpha_1, \dots, \alpha_{p-1}, t \in \mathbb{Z}, \epsilon_t \sim N(0, 1)$, independent and identically distributed.

ARCH model is alike a version of ARMA models on the variance of the time series. It

has several weaknesses like lack of proper way of determining the value of p, likeliness to overpredict the volatility etc. which limited its practical usage. But, it opened a way for its extended version, more robust GARCH models (Brooks, 2008).

2.1.2.2. GARCH

Bollerslev (1986) (Bollerslev, 1986) introduced the generalized ARCH (GARCH) model which allows depending on the lagged conditional variances. GARCH(p,q) is defined as;

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 \quad (2.21)$$

where $\alpha_i \geq 0$ and $\beta_j \geq 0$ are the parameters and σ_t^2 is the conditional variance. The conditional variance of error term of a model ϵ_t with the properties of $\epsilon_t \sim N(0, \sigma_t^2)$ is written as;

$$\sigma_t^2 = \text{var}(\epsilon_t | \epsilon_{t-1}, \epsilon_{t-2}, \dots) = E[(\epsilon_t)^2 | \epsilon_{t-1}, \epsilon_{t-2}, \dots] \quad (2.22)$$

The GARCH(p, q) model can be expressed in a form that shows that it is effectively an ARMA(m, p) model for the conditional variance.

$$v_t = \epsilon_t^2 - \sigma_t^2 \quad (2.23)$$

$$\epsilon_t^2 [1 - \alpha_1(B) - \beta_1(B)] = \alpha_0 + [1 - \beta_1(B)]v_t$$

where v_t is mean zero serially uncorrelated and $m \equiv \max\{p, q\}$.

There are also several difficulties of the model. For instance, if the length of time series is not large enough, the likelihood has a tendency of being flat. And the model may overpredict the volatility in certain situations. Because of these issues, various versions of GARCH models have been developed over the years to deal with financial time series which are seldom conditionally normal (Douc et al., 2014).

2.1.2.3. IGARCH

The GARCH process is defined to be integrated in variance. So, IGARCH can be written in the same notation as;

$$[1 - \alpha_1(B) - \beta_1(B)](1 - B)\epsilon_t^2 = \alpha_0 + [1 - \beta_1(B)]v_t \quad (2.24)$$

when,

$$\sum_{i=1}^p \beta_i + \sum_{j=1}^q \alpha_j = 1 \quad (2.25)$$

2.1.2.4. FIGARCH

If fractional difference operator d is added to IGARCH(p, q), then FIGARCH(p, d, q) is obtained and described as;

$$\phi(B)(1 - B)^d u_t^2 = \omega + [1 - \beta(B)]v_t \quad (2.26)$$

where $0 < d < 1$ and $\phi(B) \equiv [1 - \alpha_1(B) - \beta_1(B)]$ is of order $m - 1$ (Baillie et al., 1996). FIGARCH process has got fractional and long memory properties which are very useful to capture and model these features in volatility in economic and financial time series with hypergeometrical decaying autocorrelation patterns. In following sections, this particular model will be presented and investigated elaborately.

2.1.2.5. Long Memory Processes

Long memory or long range dependence is an important phenomenon in finance. Its literature starting with the introduction of Hurst phenomenon in hydrology was built with the first stationary model fGn and then models like ARFIMA and FIGARCH in finance over decades.

The first method developed for long memory was rescaled range or R/S analysis by H. E. Hurst (Hurst, 1951). Letting X_n be the inflow of to the dam in the original problem or any time series in period $n = 1, 2, \dots, k$, the rescaled adjusted range statistic is;

$$R/S(k) = \frac{\max_{0 \leq j \leq k} \left\{ \sum_{n=1}^j X_n - \frac{j}{k} \sum_{n=1}^k X_n \right\} - \min_{0 \leq j \leq k} \left\{ \sum_{n=1}^j X_n - \frac{j}{k} \sum_{n=1}^k X_n \right\}}{\sqrt{\frac{1}{k} \sum_{j=1}^k \left(X_j - \frac{1}{k} \sum_{j=1}^k X_j \right)^2}} \quad (2.27)$$

where the numerator is called the adjusted range and the denominator is the sample standard deviation. After examination of many different time series, Hurst found that $R/S(k) \approx k^H c$ for large k . In this equation, c denotes some constant and H determines if the time series is long range dependence ($H > 0.5$) or short range dependence ($H = 0.5$).

Fractional Brownian Motion (fBm) (Mandelbrot and Van Ness, 1968) is defined of Hurst parameter $0 < H \leq 1$ for $t > 0$ as a centered Gaussian process with the covariance function;

$$\mathbb{E}[B_t^H, B_s^H] = \frac{1}{2}(t^{2H} + s^{2H} - |t - s|^{2H}) \quad (2.28)$$

FBm is the only self-similar continuous-time Gaussian process with independent increments. Based on H values, *fBm* is arranged into three different groups. If $H = 0.5$, the series is independent, if it is in the interval $[0, 0.5)$, the series is antipersistent and $H \in [0.5, 1)$ indicates the persistent series. Also, discrete time increment process of *fBm* is called fractional Gaussian noise which has the covariance function;

$$\gamma(t, s) = \frac{\sigma^2 |t - s - 1|^{2H} - 2|t - s|^{2H} + |t - s + 1|^{2H}}{2} \quad (2.29)$$

Regime Switching

One of the popular nonlinear time series models Markov switching model (Hamilton, 1989) also known as regime switching model. As the name of the model hints, the model aims to capture time series features in different regimes. A simple regime switching model can be

described with the first order AR specifications as;

$$y_t = c_{s_t} + \phi_{s_t} y_{t-1} + \epsilon_t \quad (2.30)$$

where $\epsilon_t \sim N(0, \sigma^2)$, c and ϕ are coefficients and s_t denotes the regime indicator. Then, the dynamics of the y_t of the underlying regime s_t can be determined by a Markov chain.

$$Pr(s_t = j \mid s_{t-1} = i, s_{t-2} = k, \dots, y_{t-1}, y_{t-2}, \dots) = Pr(s_t = j \mid s_{t-1} = i) = p_{ij} \quad (2.31)$$

This model described with these two equations permits number of dynamic systems at different stage controlled by the regime indicator s_t (Hamilton and Raj, 2002). Regime switching models are very useful at capturing nonlinear dynamics of many financial return series and explaining behaviors like alternating and subsequently appearing persistent low and high volatility periods, fat-tails, skewness, time-varying correlations.

2.1.2.6. *Fractals*

The concept of a fractal was first introduced by Mandelbrot in 1975 (Mandelbrot, 1977). According to Mandelbrot's definition, a fractal is structure comprised of parts that are similar to the whole of its structure. A universe of fractal systems has been explored and new mathematical analyses and modelling techniques have been implemented since.

Hausdorff Dimension

Fractal objects or systems are identified by fractal dimension. There are several fractal dimension computation methods. One of the very first fractal dimension method is the Hausdorff dimension;

$$D_H := \lim_{r \rightarrow 0} \frac{\log N(\varepsilon)}{\log(\frac{1}{\varepsilon})} \quad (2.32)$$

where $\varepsilon > 0$ and $N(\varepsilon)$ is the smallest number of r -diameter balls needed to cover the fractal

(Schroeder, 2009).

The Correlation Dimension

One of the most used fractal dimensions is the correlation dimension introduced by Grassberger and Procaccia (Grassberger and Procaccia, 1983a) which counts how many points have a smaller Euclidean distance than some given distance ε .

$$D := \lim_{\varepsilon \rightarrow 0} \frac{\log C(\varepsilon)}{\log(\varepsilon)} \quad (2.33)$$

where $C(\varepsilon)$ is correlation sum or correlation integral (Schroeder, 2009).

The correlation summation at scale ε for a given embedding dimension is defined as;

$$C(\varepsilon) = \frac{2}{N(N-1)} \sum_{j=1}^N \sum_{i=j+1}^N \Theta(\varepsilon - \|X_i - X_j\|) \quad (2.34)$$

where $\Theta(\cdot)$ is the Heavyside function;

$$\Theta(x) = \begin{cases} 0, & x \leq 0 \\ 1, & \text{otherwise} \end{cases} \quad (2.35)$$

and $N = T - (m-1)\tau$ is the number of embedding points in an m -dimensional space, T is the length of time series.

Higuchi's fractal dimension

Higuchi's algorithm is described as follows (Higuchi, 1988). Giving a N -length one-dimensional time series with equal intervals $x(1), x(2), \dots, x(n)$, a new time series X_k^m is constructed as;

$$X_k^m : x(m), x(m+k), x(m+2k), \dots, x(m + \text{int}[\frac{N-m}{k}].k) \quad (2.36)$$

where k as an integer is the time interval and number of new time series sets and $m = 1, 2, \dots, k$. Therefore, Eq. 2.36 gives k number of new time series. The length of each new time series obtained by X_k^m is defined as;

$$L_m(k) = \left(\frac{1}{k} \left\{ \left[\sum_{i=1}^{\text{int}[\frac{N-m}{k}]} |x(m+ik) - x(m+(i-1).k)| \right] \frac{N-1}{\text{int}[\frac{N-m}{k}.k]} \right\} \right) \quad (2.37)$$

where $\frac{N-1}{\text{int}[\frac{N-m}{k}.k]}$ is the normalization factor for the curve length of k sets of constructed time series. By Eq. 2.37, the length of the curve for the interval k is obtained by computing the average value over k series of $L_m(k)$;

$$L(k) = \frac{1}{k} \sum_{m=1}^k L_m(k) \quad (2.38)$$

Then, based on Eq. 2.38, the fractal dimension D_f is described by;

$$L(k) \sim k^{D_f} \quad (2.39)$$

So, the complexity measure D_f can be calculated by the least-squares linear best fit procedure as finding the slope of the curve on the graph of $\ln(L(k)) - \ln(1/k)$ based on Eq. 2.39. D_f takes values ranging between 1 and 2.

MF DFA

Self-similarity, invariance in relation to scaling is one of the main characteristics of nonlinear systems and many fractal time series. Multifractal detrended fluctuation analysis (MF DFA) is very effective for the identification of fractal scaling properties and the revelation of long-range correlations in nonstationary time series. It has extensively been used for the analyses of physiological time series as well as the time series in various areas (Ihlen, 2012).

Many time series exhibit multifractal scaling behavior. So, it can be useful to learn about the scaling properties of data sets under examination. Multifractal spectrum can determine the deviations in fractal structure in time periods with large and small fluctuations (Ihlen,

2012).

Given S_α be the subset of multifractal A containing all points with pointwise dimension α , then $f(\alpha)$ denoting the S_α is called multifractal spectrum of A . The maximum value of $f(\alpha)$ equals to the box dimension (Halsey et al., 1986).

2.1.2.7. Chaos

Even though fractals and chaotic dynamical systems have various interrelationships, they are completely different concepts. As mentioned in previous section, fractals are self-similar objects which are often the output of basic dynamical systems. On the other hand, chaos is a sort of deterministic but unpredictable dynamical behavior (Feldman, 2012).

Deterministic means that it actually based on certain rules. Despite of these rules, chaotic behavior that is unpredictable and sensitive to initial conditions in the long term comes into existence in a aperiodic orbits which are bounded. Because of such interesting and unique properties, chaotic dynamical systems like fractal objects have been attracting a growing interest among a wide range of scientific disciplines. Various class of data sets are analyzed with mixed well-developed methods like phase space reconstruction, fractal dimension, correlation dimension, lyapunov exponent and so on. These methods will be visited and explored in much detail throughout the following related sections.

Availability of fast computers has been helping the development of still young nonlinear deterministic modelling which is now being applied many sorts of problems in variety of areas. Our aim is to apply nonlinear dynamical models to explain complicated behaviors and determine appropriate methods for modelling and forecasting complex systems like financial time series with nonlinear and nonstationary features.

2.1.2.8. Wavelet Analysis

As briefly mentioned in introduction, wavelet methods are especially favorable in comparison to conventional methods by means of not requiring stationarity. Wavelets are

wavelike functions localized with zero mean both in frequency and time domain.

In Fourier analysis, a signal is decomposed into a series of basis functions. These functions differ in shape from one another. They are comprised of sinewaves with diverse frequencies, phases and amplitudes. In contrast, basis functions in wavelet decomposition possess equal shapes. They are acquired by shifting and stretching a prototype wavelet. In comparison to Fourier analysis, in wavelet analysis, signals are locally characterized in both the time and the frequency domain and gathering original signal by inverse operation is easier (Lowen and Teich, 2005).

Frequency domain analysis has got disadvantages for assuming that the signal is stationary as the most linear time domain analysis. Therefore, wavelet analysis favors for not requiring stationarity as operating on both time and frequency domain.

The convolution of a wavelet function with the signal under investigation is called wavelet transform. In wavelet transform, by stretching and squeezing wavelet while moving it on different locations on a signal, the local matches between the shape of the wavelet and the shape of the signal are evaluated at a certain scale and location and then placed on a transform plane according to the correlation values between them. Wavelet transforms are grouped into continuous wavelet transform and discrete wavelet transform (Addison, 2017).

There are particularly several wavelet based methods we found to be very useful for the analysis of financial time series namely wavelet spectrum, cross wavelet transform (XWT), wavelet coherence (WTC). We will be covering these wavelet methods in more detail, in subsequent chapters.

2.2. OBJECTIVES OF THE THESIS

The fundamental objective of the thesis is to investigate effective nonlinear and nonstationary methods for analysing, modelling, forecasting financial time series. The research subjects of the thesis are quite broad. All the nonlinear and nonstationary concepts and methods introduced in previous section are investigated and explored in various scales.

2.2.1. Structure of the Thesis

The rest of the thesis is organized briefly as following. In chapter 3, the wavelet coherence methods for the nonstationary time series analysis are covered. Firstly, in section 3.2, the wavelet coherence method is investigated and applied on the major international stock markets. Then, in section 3.3, interconnected relationships between the major international real estate markets are analysed and the effectiveness of the wavelet coherence and VARMA forecast method is examined for the determined time periods. Section 3.4 also introduces multiple wavelet coherence method in addition to wavelet coherence and VARMA, and the interrelationships between three time series together are investigated with the same methodology. Chapter 4 presents the researches on nonlinear time series analysis. In section 4.1, the chaos related properties of FIGARCH process are examined by employing various methods. Subsequently, a new stochastic modelling and forecasting method for the chaotic noisy Duffing map series and then for the financial time series with identical characteristics is introduced in section 4.2. After that a new method, multiscale correlation dimension (MSCD) is proposed as a new measure for chaotic properties of time series in section 4.3. Finally, in chapter 5, all findings are summarized with concluding comments.

2.2.2. Published Papers

The published papers emerge as an output of the researches of the thesis are listed as follows;

- (i) A. Yilmaz and G. Unal, “Co-movement analysis of Asian stock markets against FTSE100 and S&P 500: Wavelet-based approach,” *International Journal of Financial Engineering*, vol. 3, no. 04, p. 1650033, 2016. (Yilmaz and Unal, 2016b)
- (ii) A. Yilmaz, G. Unal, and C. Karatasoglu, “Wavelet based analysis of major real estate markets,” *Journal of Advanced Studies in Finance*, vol. 7, no. 2, pp. 107–116, 2016. (Yilmaz et al., 2016)
- (iii) C. Karatas, G. Unal, and A. Yilmaz, “Co-movement and forecasting analysis of major real estate markets by wavelet coherence and multiple wavelet coherence,” *Chinese Journal of Urban and Environmental Studies*, vol. 5, no. 02, p. 1750010, 2017. (Karatas et al., 2017)

- (iv) A. Yilmaz and G. Unal, “Chaoticity properties of fractionally integrated generalized autoregressive conditional heteroskedastic processes,” *Bulletin of Mathematical Sciences and Applications*, vol. 15, pp. 69–82, 2016. (Yilmaz and Unal, 2016a)
- (v) A. Yilmaz and G. Unal, “Stochastic Duffing equation in modelling of financial time series,” *International Journal of Dynamics and Control*, vol. 7, no. 4, pp. 1173–1194, 2019. (Yilmaz and Unal, 2019)
- (vi) A. Yilmaz and G. Unal, “Multiscale correlation dimension method,” *International Journal of Modern Physics C (IJMPC)*, vol. 31, no. 03, pp. 1–17, 2020. (Yilmaz and Unal, 2020)



3. WAVELET COHERENCE ANALYSIS

Concerning the forecast of time series, the characteristics of interconnections between financial assets as a result of their nature and dependent parameters influencing them together produce important information. The intensity of these relationships change over time as to constantly altering market situations and different degree of parameter' effects.

For the purpose of determining these influential relationships between time series, wavelet analysis becomes exceptionally useful. Wavelet analysis differentiates from traditional methods like the most used Fourier analysis by operating in both time domain and frequency domain. As a result of such properties, it permits exploration of co-movements on both domains explicitly. Therefore, wavelet analysis especially emerged a extensively used method for the investigation of time-frequency relationships between different time series. It has been utilized by many scientists from different disciplines such as geophysics, biomedical engineering and so on.

Particularly, the works of Lau and Weng (1995) (Lau and Weng, 1995) who gave comprehensive introduction to wavelet analysis, and Torrence and Compo (1998) (Torrence and Compo, 1998) who provided helpful advanced significance tests by setting up significance levels and confidence intervals with depicting cross wavelet spectra and coherence on atmospheric time series, rose the practical capability of wavelet analysis.

Then, the applications on wavelet coherence and cross wavelet transform are widened by Grinsted et al. (2004) (Grinsted et al., 2004) who focused on wavelet analysis for two time series. Besides, they expanded wavelet software packages with the phase angle statistics to be able to test confidence in casual relationships and demonstrated on geophysical time series. Following these developments, wavelet analysis have been applied by many researchers from wide range areas to analyze dynamic correlations between time series. Among them, wavelet analysis was applied for revealing the causality relationship between government spending and revenue (Almasri and Shukur, 2003), on monetary policy with wavelet coherency (Aguilar-Conraria et al., 2008), for the co-movement of international

stock market returns (Rua and Nunes, 2009). And lately, it was utilized for the analysis of the integration between emerging stock markets with the U.S. market (Graham et al., 2012), for the examination of the relationship between oil prices, industrial production and inflation in German economy and interest rates and share prices in Indian economy by the cross wavelet analysis (Tiwari, 2012), for the investigation of co-movement of energy commodities, traded assets such as gold, oil and stocks and European stock markets (Barunik and Vacha, 2013).

In following sections, we apply wavelet methods to analyse stock exchange pairs of FTSE 100 and S&P 500 and various Asian markets and international real estate investment trusts due to their interlinked quality. We utilize wavelet coherence to uncover clear interlinked relationships between pairs and how these relationships change in the time-frequency space. Moreover, supplementary information about casual relationship is gathered by phase angle statistics.

Following the introduction of the main theories of wavelet analysis, the continuous wavelet transform, the cross wavelet transform and phase, we continue with empirical analyses and wavelet coherency results.

3.1. WAVELET ANALYSIS

As mentioned earlier, the wavelet analysis have strong benefits over traditional methods like Fourier analysis because of having zero mean and localizing in frequency and time space. It allows to analyse the time series having nonstationary power at various different frequencies. There is a trade-off between localization in time and frequency according to Heisenberg uncertainty principle. But, wavelet transform methods can minimize this uncertainty.

For the wavelet-based analysis, we employ the wavelet package provided by Grinsted et al. (Grinsted et al., 2004) in which Morlet wavelet is chosen because it provides good balance between time and frequency. Phase relationships of two time series that are not normally distributed is explored in time-frequency space in subsequent sections by the wavelet methods based on continuous wavelet transform (CWT) by which time series can be thoroughly decomposed and then reconstructed. The cross wavelet transform (XWT) which

is obtained by CWT displays the common power and relative phase of two time series in time-frequency space. The causality links can be exposed by phase relationships between these time series.

CWT that performs as a band pass filter to the time series $x(t)$ is defined by the convolution as follows;

$$W_x(\tau, s) = \frac{1}{\sqrt{s}} \sum_{t=1}^N x(t) \varphi^*\left(\frac{t - \tau}{s}\right) \quad (3.1)$$

As written before, we utilize the Morlet wavelet which is introduced in (Goupillaud et al., 1984) and defined as;

$$\varphi(\eta) = \pi^{-\frac{1}{4}} e^{i\omega_0\eta} e^{-\frac{1}{2\eta^2}} \quad (3.2)$$

where η denotes time and ω_0 represents frequency which is selected as 6 because it yields a satisfactory balance between time and frequency localization (Grinsted et al., 2004).

There are number of particular reasons for favoring Morlet wavelet over others such as numerical advantages, scale to frequency conversion facility, low Heisenberg box area and exceptional compromise between time and frequency concentration.

3.1.1. The Cross Wavelet Transform

The cross wavelet transform of two time series is defined as:

$$W_n^{XY}(s) = W_n^X(s) \cdot W_n^{Y*}(s) \quad (3.3)$$

where $W_n^X(s)$ and $W_n^Y(s)$ are wavelet transforms of the time series X and Y . $W_n^{Y*}(s)$ is complex conjugate of $W_n^Y(s)$, and the cross wavelet power is $|W_n^{XY}(s)|$ which takes values between 0 and 1 and displays regions in the time-frequency space where the two time series share a high common power.

The theoretical distribution of the cross wavelet power of two time series is defined as;

$$D\left(\frac{|W_n^X(s).W_n^{Y*(s)}|}{\sigma_X\sigma_Y} < p\right) \Rightarrow \frac{Z_v(p)}{v} \sqrt{P_k^X P_k^Y} \quad (3.4)$$

where P_k^X and P_k^Y the background wavelet spectra of two time series, $Z_v(p)$ is the confidence level linked with the probability p for a pdf described through the square root of the product of two χ^2 distributions (Torrence and Compo, 1998).

3.1.2. Phase Angle

The wavelet coherence phase difference is defined as;

$$\varphi_{xy}(u, s) = \tan^{-1}\left(\frac{I\{S(s^{-1}W_{xy}(u, s))\}}{R\{S(s^{-1}W_{xy}(u, s))\}}\right) \quad (3.5)$$

where $\varphi_{xy} \in [\pi, -\pi]$, I and R denote the imaginary and the real parts of the power spectrum consecutively.

What phase difference shows is how two time series move together at different scales. So, supposing that two time series are in phase or positive correlated then the arrows point right. If the phase difference is between 0 and $-\pi/2$, the arrows point upwards, it shows that the first time series leads the second time series. Similarly, if the phase difference is between 0 and $\pi/2$, the arrows point downward then the second time series leads the first one. Besides, when the directions of arrows are in-between, the connection can be described as a mix of the implication of the both main direction to a degree of less than 100%.

3.1.3. The Cone of Influence

The cone of influence is the area of edges which appears because of the fall in the wavelet power by a factor of e^{-2} . The wavelet power drops at the edges because of the discontinuity which occurs as a result of bringing the number of data points of the time series to the power of two and padded with zeros. Particularly, the more zeroes are needed as the scale

increases.

3.1.4. Wavelet Coherence

The wavelet coherence (WTC) of two time series is defined as;

$$R_n^2(s) = \frac{|S(s^{-1}W_n^{XY}(s))|^2}{S(s^{-1}|W_n^X(s)|^2).S(s^{-1}|W_n^Y(s)|^2)} \quad (3.6)$$

where S denotes the smoothing operator. WTC equation gives values between 0 and 1 which represent the low correlation and high correlation between two time series consecutively.

Monte Carlo method is used for estimating the significance level of wavelet coherence. The statistically significant regions of co-movement is found by simulating a great number of white noise pairs which have the same length as the original time series with which the squared coherence values are compared then.

3.2. CO-MOVEMENT ANALYSIS OF ASIAN STOCK MARKETS AGAINST FTSE100 AND S&P 500

Firstly, we experiment wavelet coherence method to analyse the relationship between FTSE100 and S&P 500 and various Asian markets of S&P/ASX 200 (AXJO) (Australia), S&P/ASX200 A-REIT (AXPJ) (Australia), BIST (Turkey), HIS (Hong Kong), IDX (Indonesia), KLSE (Malaysia), 1368.KL (Malaysia), KOSPI (Korea), N225 (Japan), RTS (Russia), Shenzhen (China), 0050.TW (Taiwan). All the time series datum were collected by Yahoo Finance website.

For the analysis, the time series are arranged as pairing the markets. FTSE100 and S&P 500 time series are individually and one by one paired with Asian market time series and prepared by arranging data date by date. 01/01/2000 is taken as a reference date for each time series. Because of the availability of data, time series pairs set up have different time periods which are given in detail in Table 3.1.

Table 3.1: Details of stock exchange pairs

Stock Exchange Pairs	Time Period	No of data
FTSE-AXJO and S&P500-AXJO	1/4/2000 to 3/11/2016	4099
FTSE-AXPJ and S&P500-AXPJ	3/4/2013 to 3/10/2016	694
FTSE-BIST and S&P500-BIST	1/3/2000 to 3/11/2016	4131
FTSE-HSI and S&P500-HSI	1/3/2000 to 3/11/2016	4035
FTSE-IDX and S&P500-IDX	1/20/2009 to 3/11/2016	1743
FTSE-KL and S&P500-KL	1/3/2000 to 3/11/2016	4028
FTSE-KLSE and S&P500-KLSE	1/3/2000 to 3/11/2016	3874
FTSE-KS11 and S&P500-KS11	1/4/2000 to 3/11/2016	3898
FTSE-N225 and S&P500-N225	1/4/2000 to 3/11/2016	3823
FTSE-RTS and S&P500-RTS	1/3/2000 to 3/11/2016	3881
FTSE-Shenzhen and S&P500-Shenzhen	1/4/2000 to 12/17/2015	3842
FTSE-TW and S&P500-TW	12/31/2007 to 3/11/2016	1946

3.2.1. The Wavelet Coherence Results

By wavelet coherence, it is intended to detect how correlations change over certain frequency and time period. In a three-dimensional contour plot which displays the wavelet coherence, colors from red to blue presents the level of significant dependence. The red color represents the most dependent relationship as the blue color portrays the least dependent one. The region influenced by the edge effects where the number of data points are not sufficient is shown by black lines on each sides.

Black arrows illustrates the phase relationships. If the direction of arrow is right, then the time series are positively correlated. If opposite is the case, then the time series are not in phase. Additionally, if the arrow points down, it refers that the first time series leads the second time series. Horizontal axis and vertical axis of the plot represent the frequency and the time, consecutively.

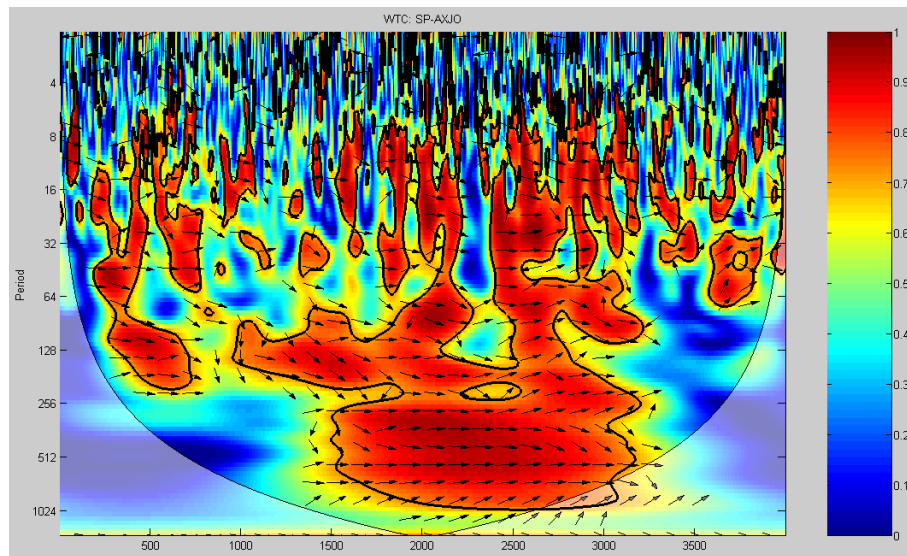
(a) S&P 500 – AXJO and FTSE – AXJO

S&P500 - AXJO and FTSE - AXJO wavelet coherence analysis results for the time period of 1/4/2000 and 3/11/2016 are given in Fig. 3.1. Both plots display that there are significant correlations mostly between the dates of 11/28/2005 and 10/20/2015 and at the frequencies of 8 through 1024. The co-movement for the frequencies smaller than 4 appears less strong. All the arrows point right indicates the positive correlations between time series.

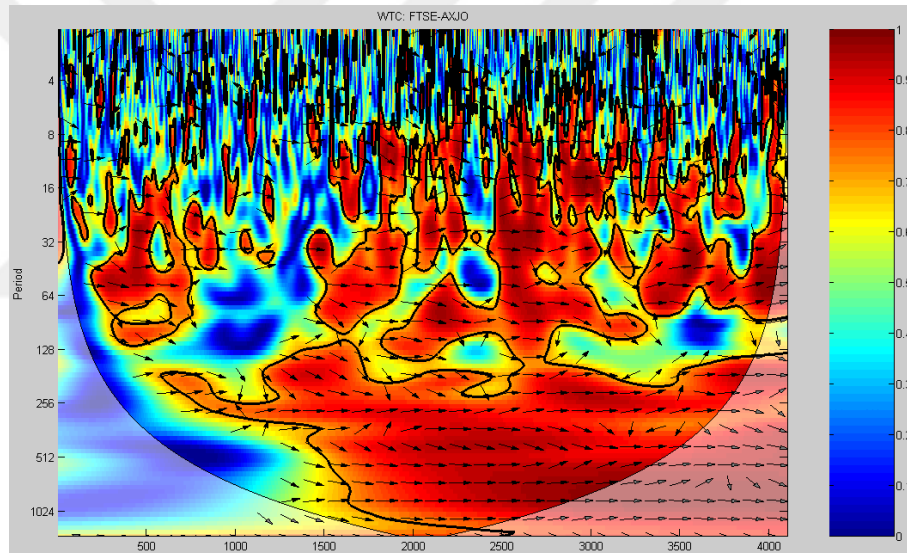
(b) S&P 500 – AXPJ and FTSE – AXPJ

Through the time period between 3/4/2013 and 3/10/2016, the most correlated region is detected in Fig. 3.2 in the frequency of 64 and on the time period starting from 10/28/2013 to 8/14/2015. The correlations are mostly positive. But the correlations appear weak in other frequencies.

(c) S&P 500 – BIST and FTSE – BIST

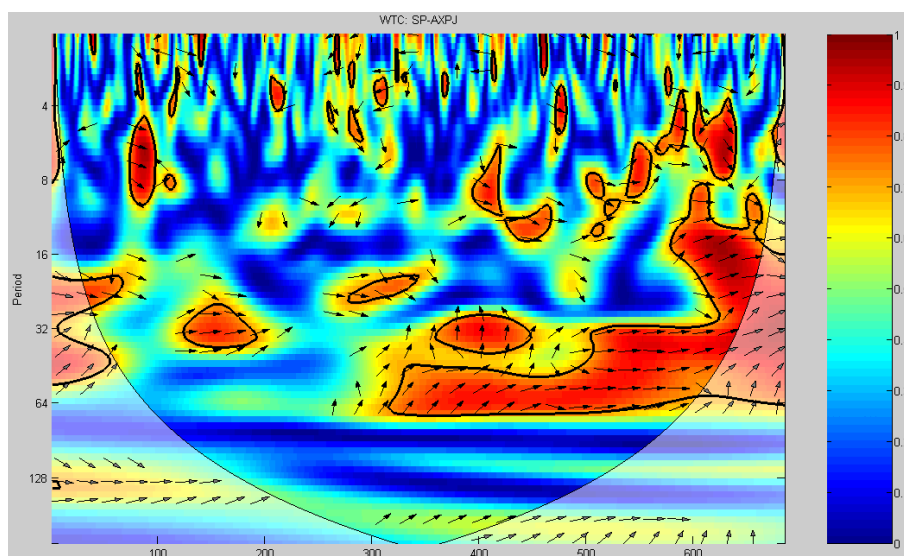


a. S&P 500 - AXJO

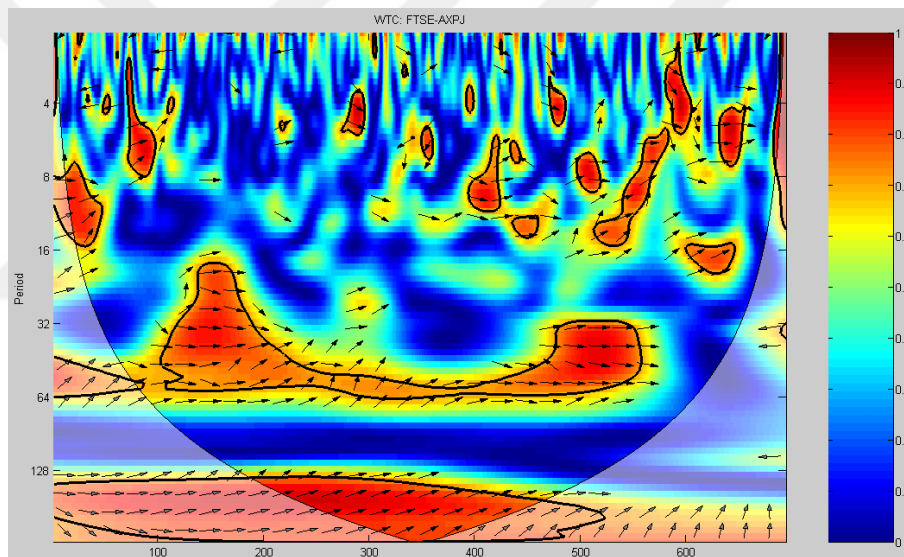


b. FTSE - AXJO

Figure 3.1: Wavelet coherence results of S&P 500 - AXJO and FTSE - AXJO



a. S&P 500 - AXPJ



b. FTSE - AXPJ

Figure 3.2: Wavelet coherence results of S&P 500 - AXPJ and FTSE - AXPJ

In Fig. 3.3, it is shown that particularly, in the time interval of 2/17/2006 and 8/14/2012 and in the time frequencies of 0 to 64, there are clear co-movements. Also, in the time interval of 9/30/2005 to 9/4/2009 and the frequencies between 512 to 1024, co-movement indicators exist. The correlations are positive according to arrows' directions.

(d) S&P 500 – HSI and FTSE – HSI

Fig. 3.4 shows that for S&P 500-HSI and FTSE-HSI pairs, the frequency of 64 is the most correlated region at the time interval of 1/26/2006-1/8/2013. In addition, at the frequencies from 0 to 16, many correlated parts appear for the duration. In the frequency of 512, co-movement is also persistent and clear. Arrows are right pointed in all the significant regions suggesting that the correlations are positive.

(e) S&P 500 – IDX and FTSE – IDX

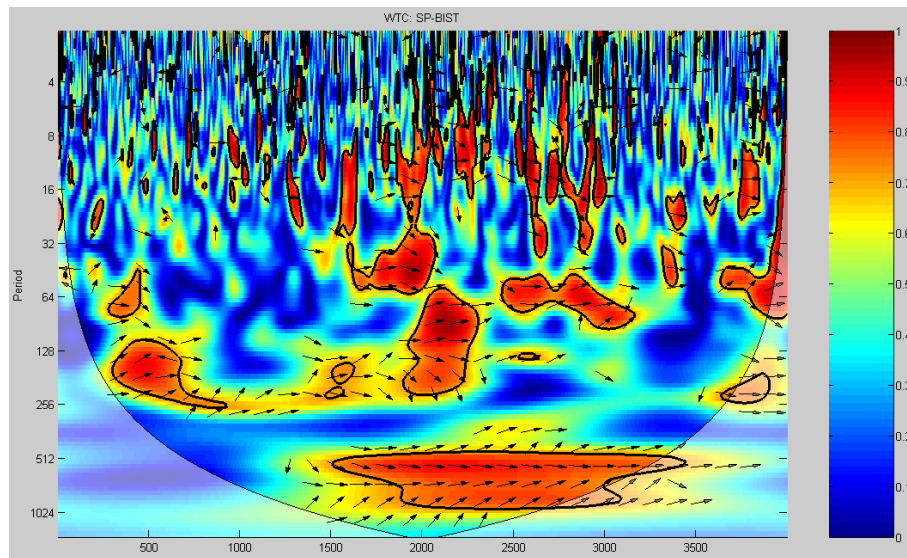
In plots of S&P 500-IDX and FTSE-IDX shown in Fig. 3.5, the correlations appears to be accumulated irregularly in the time interval between 10/26/2009 and 9/13/2012, mostly between the frequencies of 0 and 32. Additionally, the directions of the arrows indicates the positive correlations between time series.

(f) S&P 500 – KL and FTSE – KL

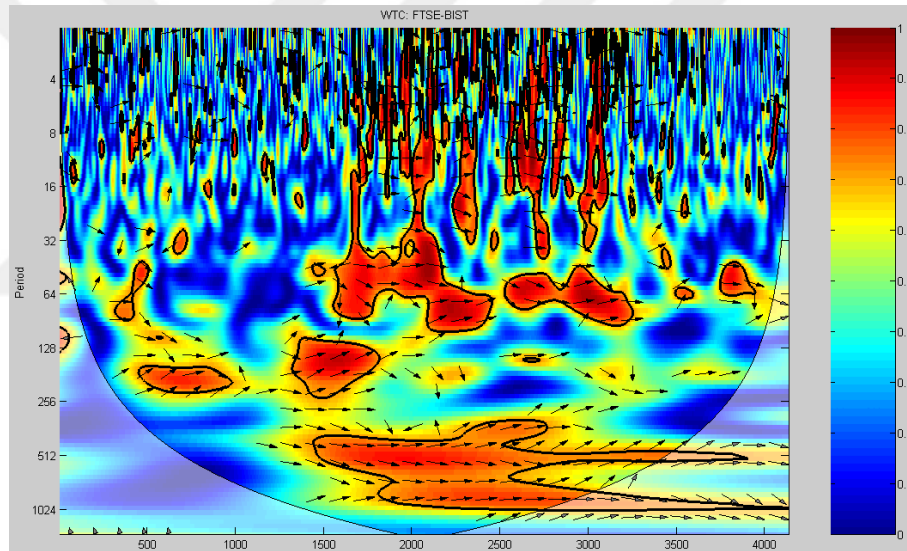
It is shown in Fig. 3.6, there are number of small correlated regions spread between the frequencies 0 to 16. The most correlated area appears to be at the frequencies between 32 and 64 and in the time interval of 6/20/2008-11/9/2010. Also, right-pointed arrows suggest the positive correlation between the time series of FTSE and KL. But, compared to the results of other pairs, co-movements look relatively small.

(g) S&P 500 – KLSE and FTSE – KLSE

The observations in Fig. 3.7 refer that co-movement is steady in the time interval of 12/1/2005-7/10/2012 and in the frequencies of 64-128. In addition, there are positive

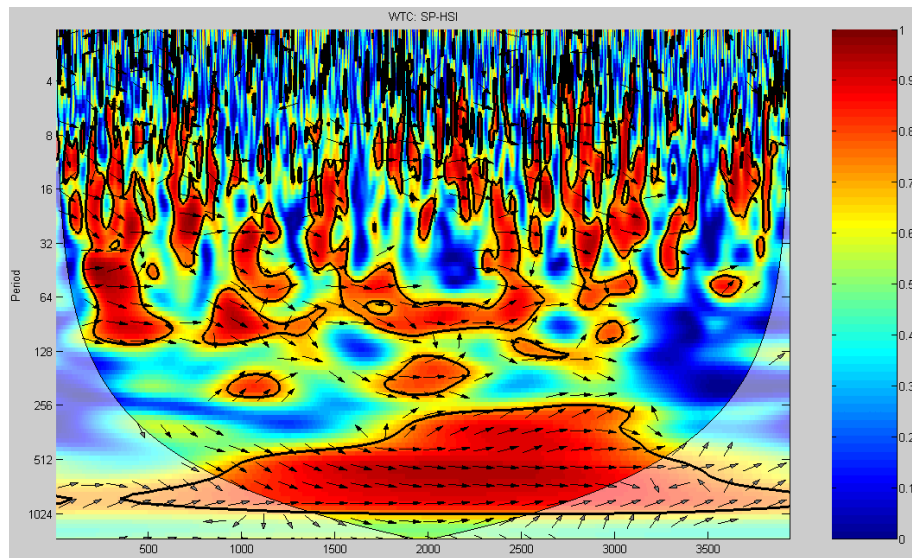


a. S&P 500 - BIST

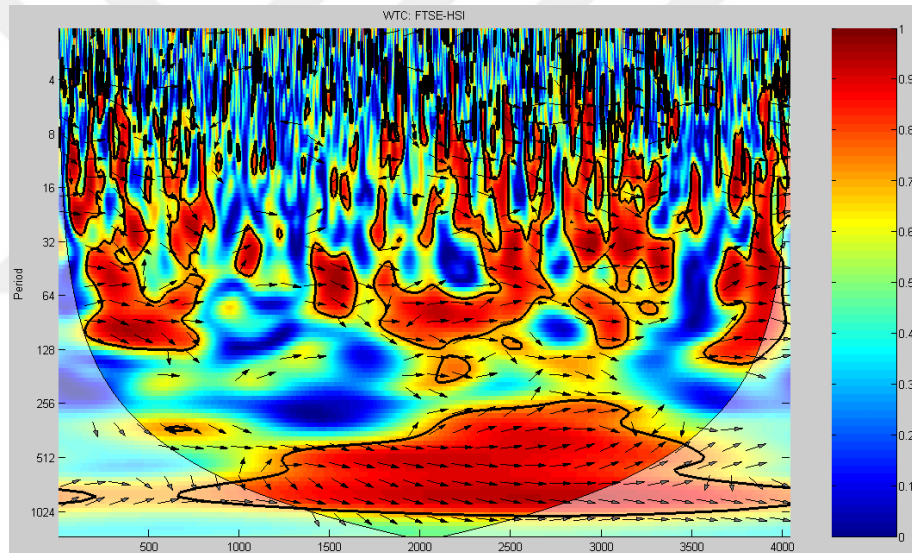


b. FTSE - BIST

Figure 3.3: Wavelet coherence results of S&P 500 - BIST and FTSE - BIST

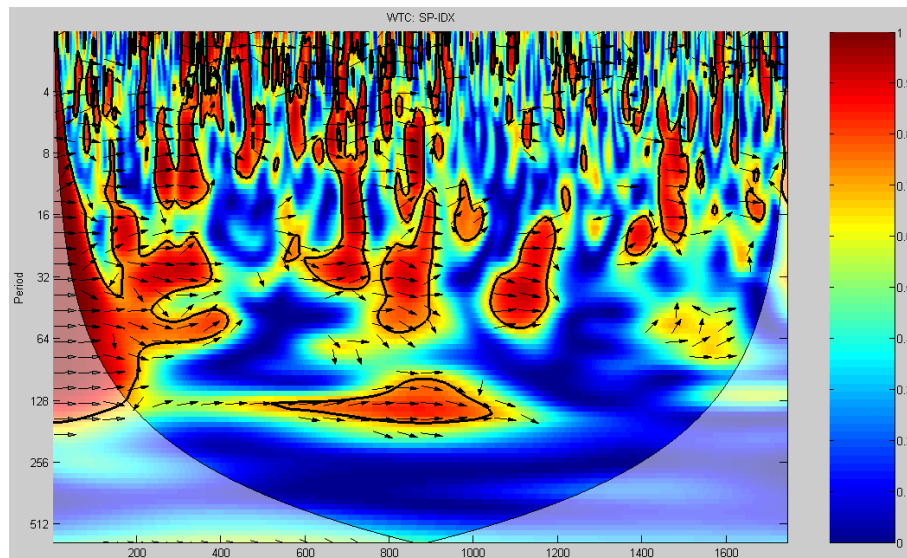


a. S&P 500 - HSI

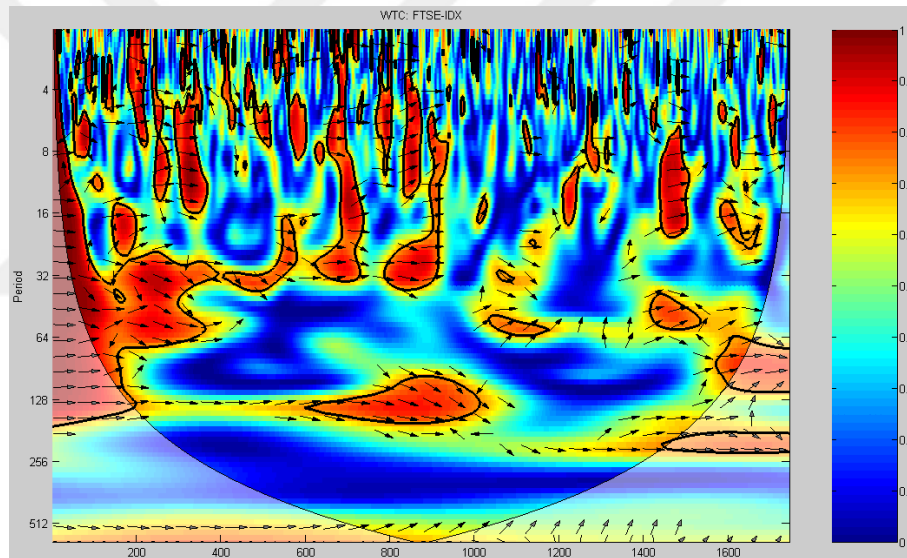


b. FTSE - HSI

Figure 3.4: Wavelet coherence results of S&P 500 - HSI and FTSE - HSI

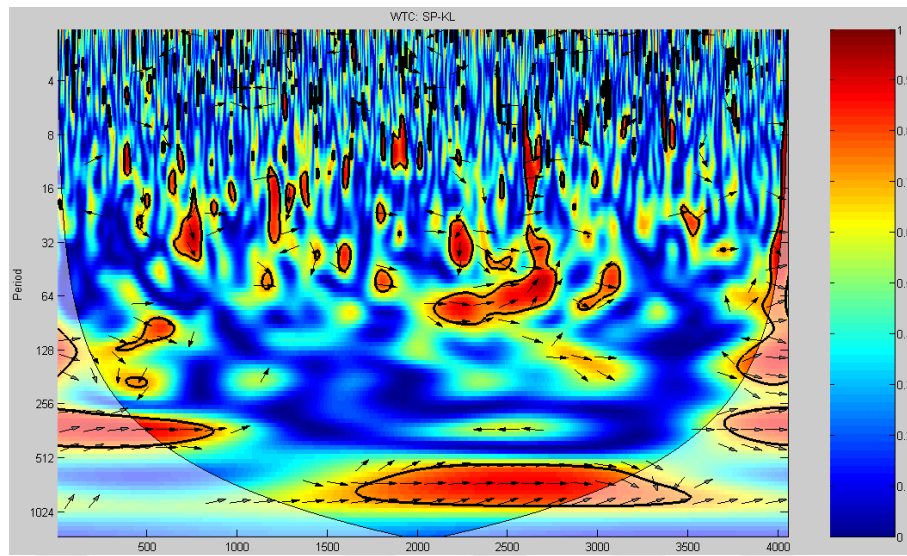


a. S&P 500 - IDX

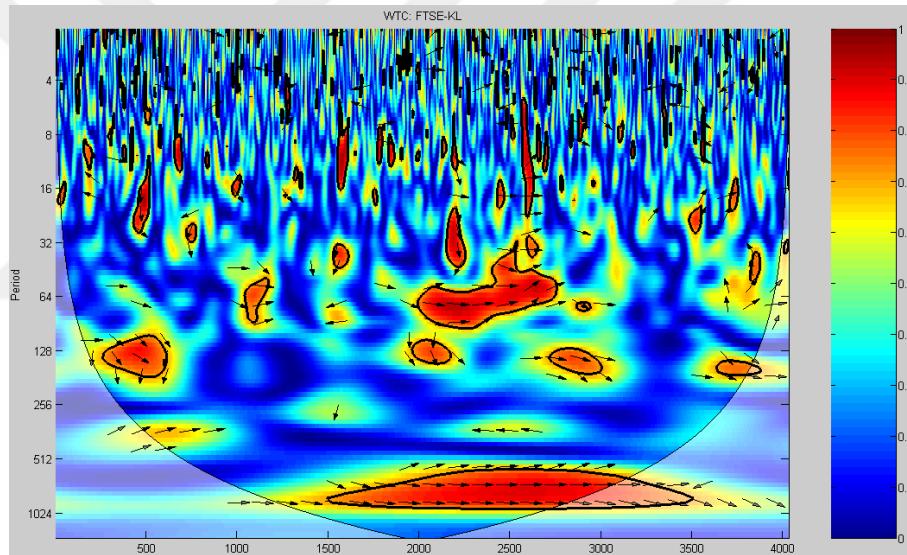


b. FTSE - IDX

Figure 3.5: Wavelet coherence results of S&P 500 - IDX and FTSE - IDX



a. S&P 500 - KL



b. FTSE - KL

Figure 3.6: Wavelet coherence results of S&P 500 - KL and FTSE - KL

correlations in the same interval and in about all the frequencies accumulating mostly in the frequency of 512-1024.

(h) S&P 500 – KS11 and FTSE – KS11

Fig. 3.8 indicates that particularly, between 1/13/2005 and 4/8/2013, at different frequencies from 8 to 256, there are fairly correlated regions. Taking into account phase relationships, it appears that around all arrows are right-pointed suggesting that the time series pairs are in phase.

(i) S&P 500 – N225 and FTSE – N225

As indicated in Fig. 3.9, S&P 500-N225 and FTSE-N225 wavelet coherence analysis generates similar but stronger results in similar the frequency interval 8-128 and the time interval 12/30/2005-7/24/2013 with FTSE-KS11 and S&P 500-KS11. Additionally, there are positive co-movement in the frequency 512 and in all the time periods.

(j) S&P 500 – RTS and FTSE – RTS

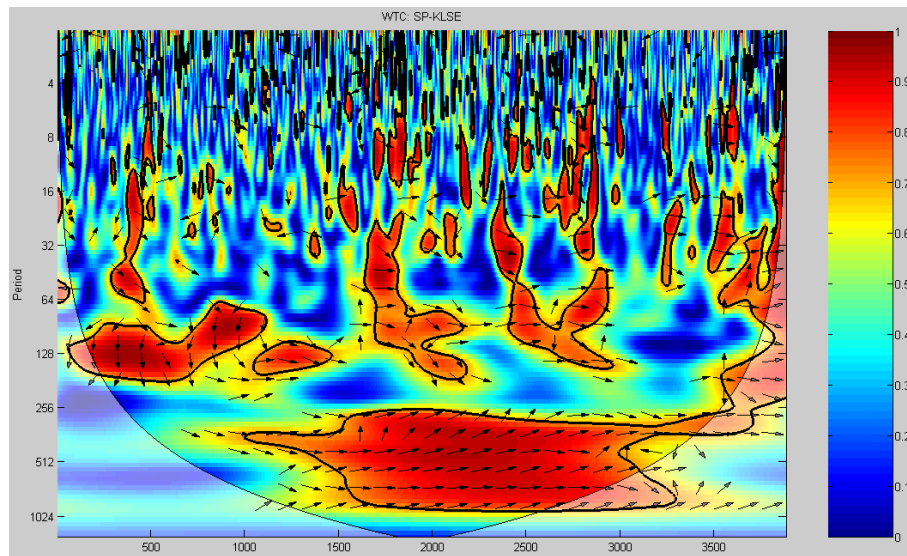
In Fig. 3.10 subplots of S&P 500-RTS and FTSE-RTS shows that significantly correlated regions appear to be in the time period between 1/26/2006 and 6/25/2016 in around all periods in different degrees. In addition, in those regions, as the directions of arrows refer, the correlations are positive.

(k) S&P 500 – Shenzhen and FTSE – Shenzhen

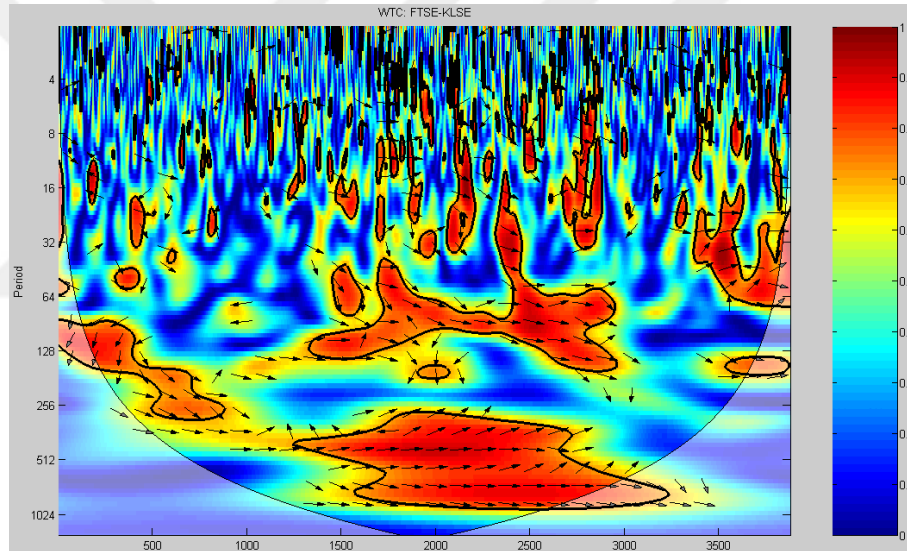
As indicated in Fig. 3.11 there are not considerable co-movement relations in the time series pairs of FTSE-Shenzen and S&P-Shenzen referring that they are not interrelated.

(l) S&P 500 – TW and FTSE – TW

As shown in Fig. 3.12, according to wavelet coherence analysis results, FTSE-TW and S&P-

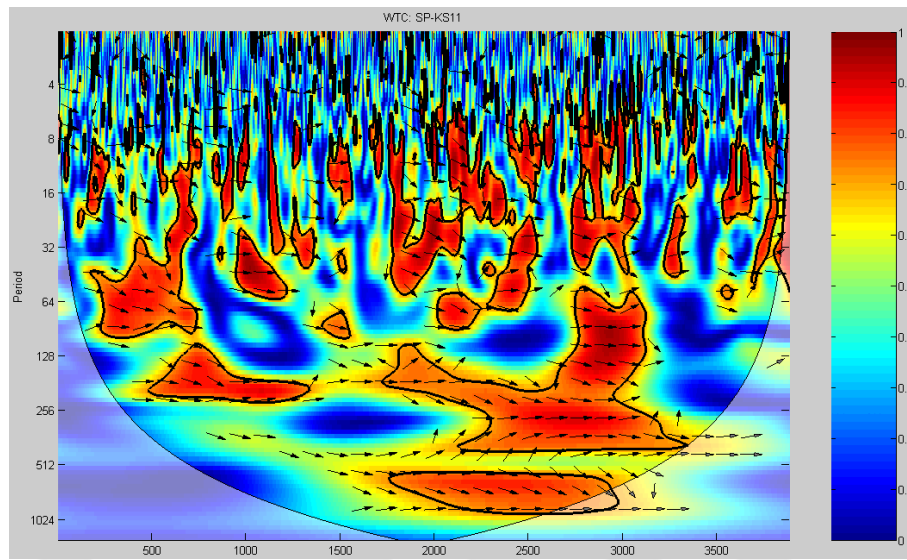


a. S&P 500 - KLSE

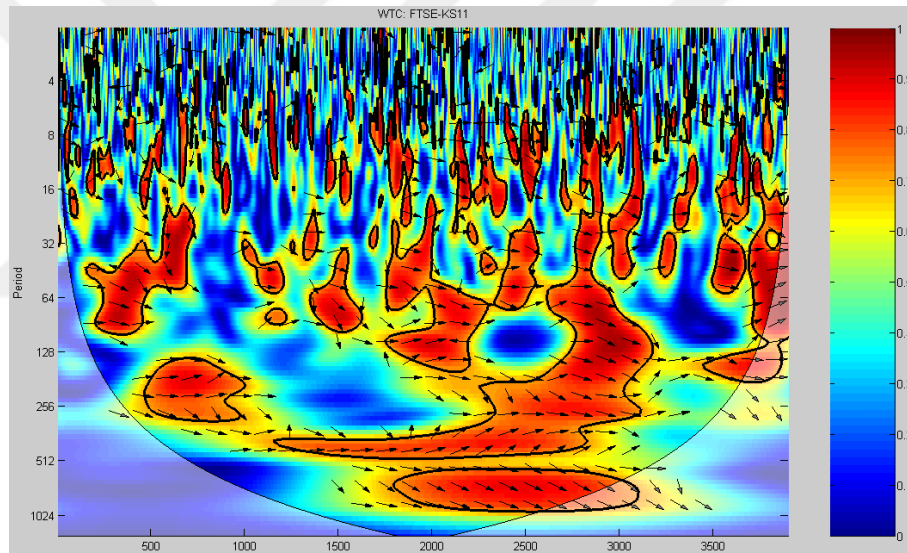


b. FTSE - KLSE

Figure 3.7: Wavelet coherence results of S&P 500 - KLSE and FTSE - KLSE

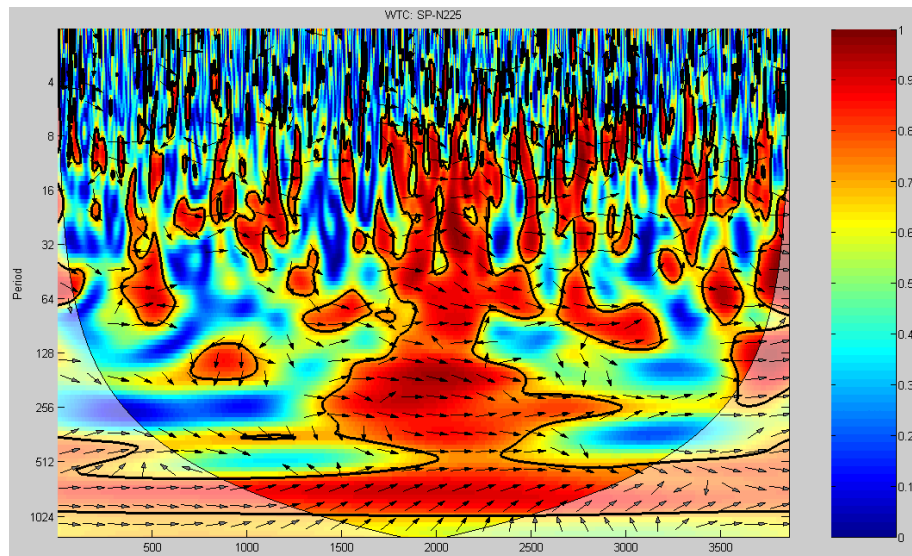


a. S&P 500 - KS11

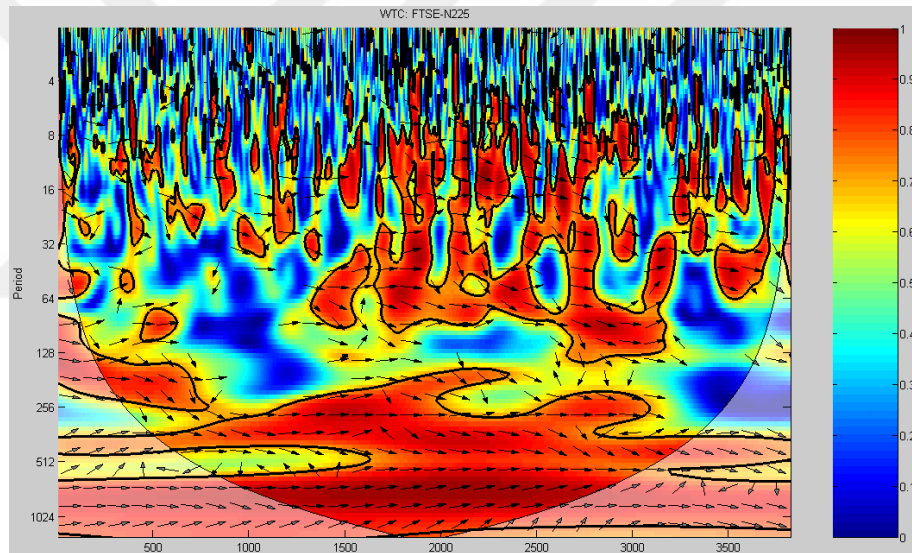


b. FTSE - KS11

Figure 3.8: Wavelet coherence results of S&P 500 - KS11 and FTSE - KS11

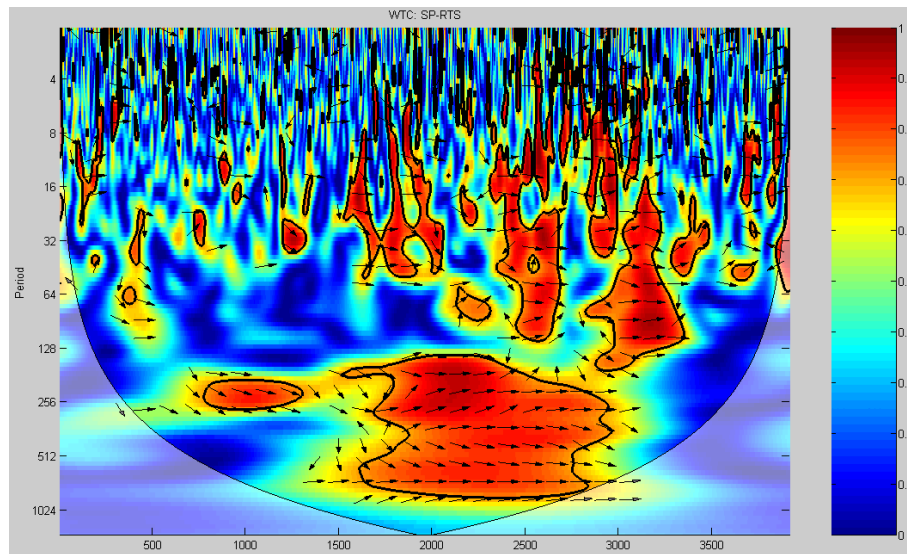


a. S&P 500 - N225

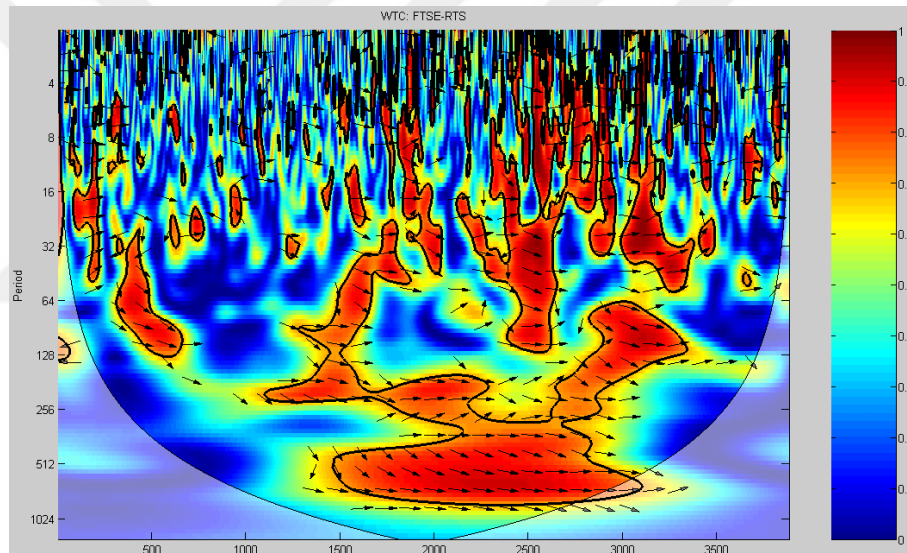


b. FTSE - N225

Figure 3.9: Wavelet coherence results of S&P 500 - N225 and FTSE - N225

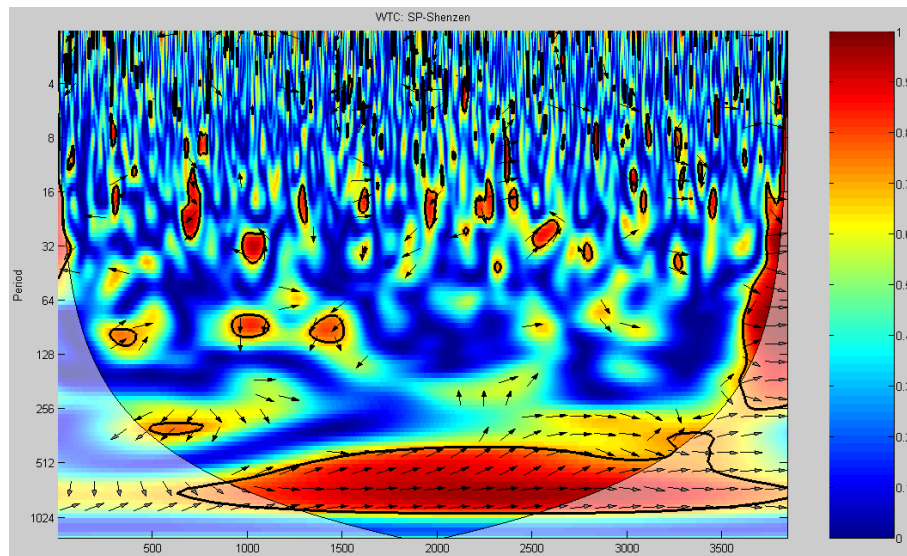


a. S&P 500 - RTS

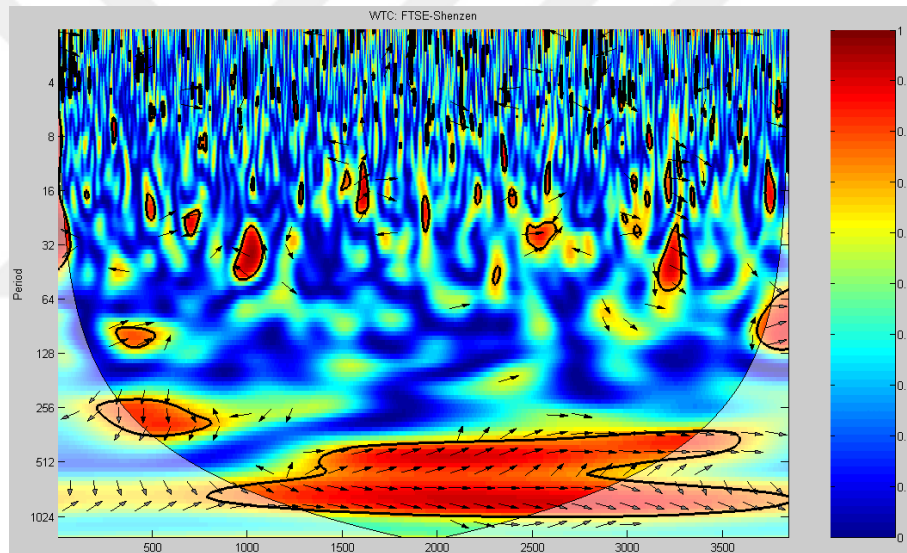


b. FTSE - RTS

Figure 3.10: Wavelet coherence results of S&P 500 - RTS and FTSE - RTS



a. S&P 500 - Shenzhen



b. FTSE - Shenzhen

Figure 3.11: Wavelet coherence results of S&P 500 - Shenzhen and FTSE - Shenzhen

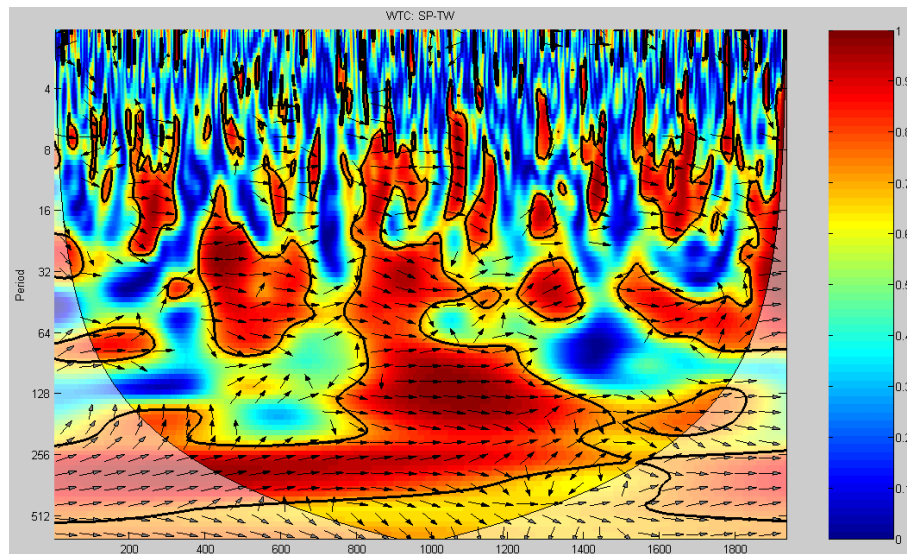
TW show strong relationships at the 9/9/2009-10/28/2013 and at the frequencies of 32-128 with the positive correlations considering right-pointed arrows.

In general, the wavelet results for both S&P 500 and FTSE pairs generated similar results as expected considering the strong relationships of these two developed economy and their stock markets. Similarly, the stock markets of the most developed economies of the Asian markets of AXJO, HSI, N225 and TW paired come out as the most correlated markets with S&P 500 and FTSE. And then, the stock markets of the developing economies of KS11, RTS, BIST, IDX, KLSE and AXPJ come as relatively less correlated markets. Lastly, KL and Shenzhen are appeared to be the least correlated stock markets in the time interval the analysis are undertaken. As a consequence, the different degree of interconnections can be identified by wavelet coherence analysis for each paired time series in different time interval and frequencies.

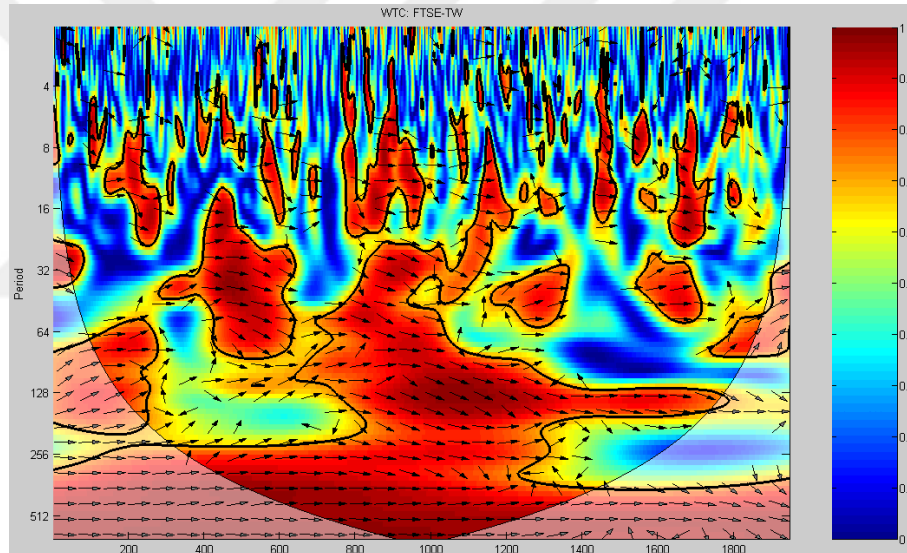
In summary, the most correlated time interval and the frequencies are come out for the most correlated pairs as follows; for S&P 500-AXJO and FTSE-AXJO as 11/28/2005-10/20/2015 and 8 to 1024 days, for S&P 500-HSI and FTSE-HSI as 1/26/2006-1/8/2013 and 64 to 512 days, for S&P 500-N225 and FTSE-N225 as 12/30/2005-7/24/2013 and 8-128 days, for S&P 500-TW and FTSE-TW as 9/9/2009-10/28/2013 and 32-128 days.

Also, the relatively less correlated pairs occurred can be listed with the time interval and the frequencies as; for S&P 500-KS11 and FTSE-KS11 between 1/13/2005 and 4/8/2013, 8 days to 256 days; for S&P 500-RTS and FTSE-RTS between 1/26/2006 and 6/25/2013 in almost all periods; S&P 500-BIST and FTSE-BIST between 2/17/2006 and 8/14/2012, 1 day to 64 days; S&P 500-IDX and FTSE-IDX, between 10/26/2009 and 9/13/2012, 1 day to 32 days; S&P 500-KLSE and FTSE-KLSE, between 12/1/2005 and 7/10/2012, 64 days and 128 days; S&P 500-AXPJ and FTSE-AXPJ, between 10/28/2013 and 8/14/2015 at 64 days.

Lastly, the least correlated pairs in comparison to other pairs occurred as S&P 500-KL and FTSE-KL in the time period of 6/20/2008-11/9/2010 between the frequencies of 32-64 days and S&P 500-Shenzhen and FTSE-Shenzhen with no considerable correlations in any time



a. S&P 500 - TW



b. FTSE - TW

Figure 3.12: Wavelet coherence results of S&P 500 - TW and FTSE - TW

interval or frequency.

In consequence, by grouping daily and weekly frequencies as short term frequencies and monthly and yearly ones as long term, the wavelet coherence analysis displayed short term and long term co-movements of selected Asian stock markets in various degrees with S&P 500 and FTSE. Especially, it is found that AXJO, RTS, HSI are interconnected in all frequencies whilst N225, BIST, KS11, IDX are interrelated in weekly and yearly frequencies and TW, KLSE, AXPJ are only correlated in monthly frequencies.

This part of the research is also published (Yilmaz and Unal, 2016b).

3.3. CO-MOVEMENT ANALYSIS OF MAJOR REAL-ESTATE MARKETS

In our second empirical study of wavelet coherence analysis, we examine the relationship between Real Estate Markets through major REITs. We use the daily prices of Creit Canadian REIT (REF-UN.TO)–Canada, LINK REIT–Hong Kong and China, iShares US Real Estate ETF(IYR)–USA, iShares International Developed Real Estate ETF(IFGL)–Japan, USA, developed Europe, Canada, iShares Europe Developed Real Estate ETF(IFEU)–Developed Europe.

A real estate investment trust (REIT) is a trust that owns, and mostly operates, many types of income-producing commercial real estate. Due to higher liquidity, the lower transaction costs, the lower capital requirements, investors generally choose to invest indirect ways like real estate investment trusts (REITs), real estate exchange traded funds (ETFs) and so on. Therefore, we employ REIT stock prices that track the monthly real estate benchmark indices to inspect daily co-movements between the real estate market prices.

Also, after identifying the dynamic dependencies and interconnections through wavelet coherence analysis, we demonstrate effective forecast method based on the information about interdependent time period and frequency by VARMA model which generates linear forecasts and takes advantage of the information of the past values of two time series productively (Lütkepohl, 2006). Then, when the forecast results of the intervals selected

as to their degree of correlations are compared, it is observed that the intervals with the most coherent regions generated better forecast results than the intervals with the least coherent regions which shows the strong performance of the applied method.

Daily prices for Creit Canadian REIT (REF-UN.TO), LINK REIT (0823.HK), iShares US Real Estate ETF (IYR), iShares International Developed Real Estate ETF (IFGL), iShares Europe Developed Real Estate ETF (IFEU) were gathered Yahoo Finance website. For the analysis, the period of time series become around 6.5 year as between 20/11/2009 and 28/6/2016.

3.3.1. The Wavelet Coherence Results

(a) IYR-CREIT

IYR-CREIT wavelet coherence results in Fig. 3.13 indicates that during the the time period, there are significant regions in various frequencies. Particularly, the most coherent area appears to be consistent at 8-32 days frequencies and two REIT move together in small time periods spread through the time period. The right-pointed phase arrows indicates that the time series pair is positively correlated.

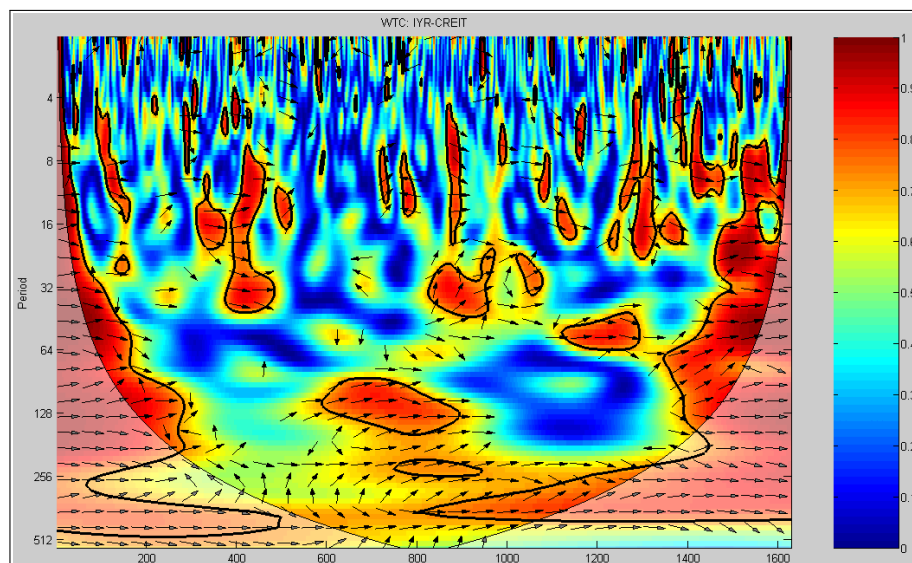


Figure 3.13: Wavelet coherence result of IYR-REF UN TO.

(b) IYR-IFEU

Fig. 3.14 shows wavelet coherence analysis results for IYR and IFEU time series pair. There are large significantly coherent regions, especially at the frequencies between 4 and 64 in the time interval of 6/9/2011-11/11/2013. And the least coherent regions are located in the time period between 6/28/2010 and 1/31/2011. In addition, all the arrows point right which suggests that the two time series are in phase.

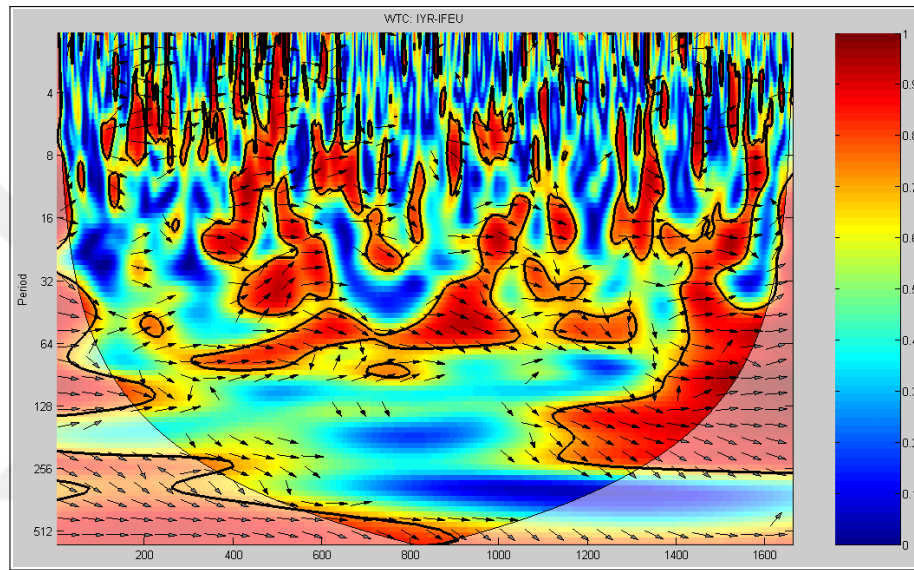


Figure 3.14: Wavelet coherence result of IYR-IFEU.

(c) IYR-IFGL

It is shown in Fig. 3.15 IYR-IFGL time series pair are correlated in almost all the time intervals and in the frequencies, as especially appearing mostly at in the time interval of 9/8/2010-4/10/2012. Also, the least coherent area is situated at the time interval between 4/11/2012 and 1/28/2013. Similar to previous time series pairs, the arrows are right-pointed indicating the positive correlation between the time series.

3.3.2. Forecast with Multivariate ARMA Process

We apply multivariate ARMA process (VARMA) for experimenting forecasting performance with the information about the interconnected relationships between REIT time

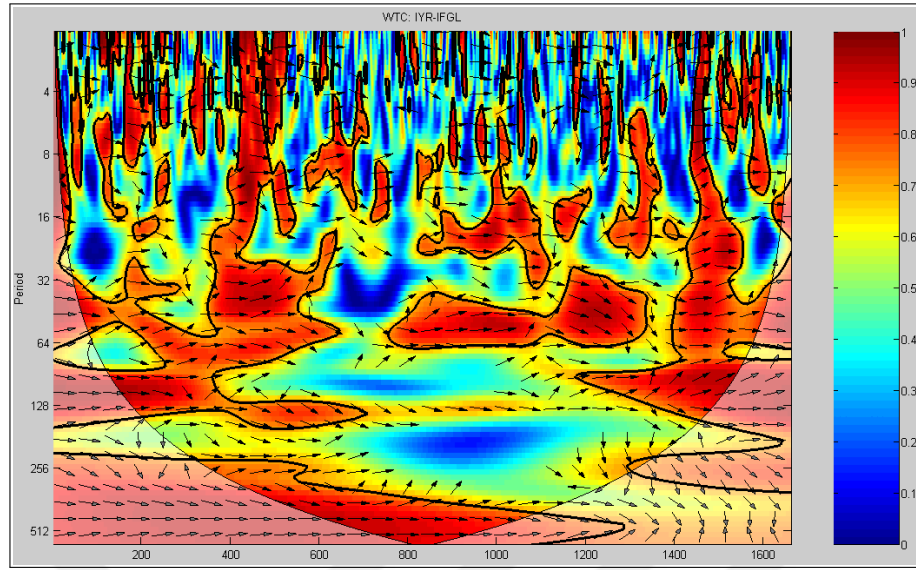


Figure 3.15: Wavelet coherence result of IYR-IFGL.

series pairs uncovered by wavelet coherence analysis.

It is expected that VARMA process enhance the forecast efficiency by including more interconnected variables (Granger, 1969). VARMA enables more precise forecast performance by handling the information located in the past values of interlinked variables at the same time by constructing a vector. Applying VARMA became harder historically. However, for two or three variable cases, it demonstrated to generate better results (Lütkepohl, 2006).

VARMA process is defined as;

$$A_0 y_t = A_1 y_{t-1} + A_2 y_{t-2} + \dots + A_p y_{t-p} + M_0 \varepsilon_t + M_1 \varepsilon_{t-1} + \dots + M_q \varepsilon_{t-q} \quad (3.7)$$

where p and q are the orders of VARMA process, y_t is vector of m dimensional time series, $A_0, A_1, \dots, A_p$ and $M_0, M_1, \dots, M_q$ are the matrices of order $m * m$ and ε_t is a noise vector.

It also can be written lag operator notation for more precise representation as;

$$A(L)y_t = M(L)\varepsilon_t \quad (3.8)$$

where $A(z) = A_0 - A_1z - \dots - A_pz^p$ and $M(z) = M_0 + M_1z + \dots + M_qz^q$ are matrix valued polynomials. The zero order matrices are usually be identical and they are equal to the identity matrix, $A_0 = M_0 = I_m$ (Lütkepohl, 2006).

(a) IYR-IFEU

In previous section, as a result of wavelet coherence analysis, the most coherent region was determined at the time interval of 6/9/2011-11/11/2013. So, following this region we made the forecasting for the 30-day time period of 12/11/2013-24/12/2013.

In Fig. 3.16, orange line, blue line and green line represent IFEU time series, IYR time series and 30 day forecast time series successively. As it is displayed, the forecast results follow the actual result very closely. Particularly for IFEU, the method which is the combination of VARMA and wavelet coherence demonstrates very efficient forecast results in very small error bands represented with red color line.

After the examination of the most coherent area, the forecast results for the least coherent area are also made for the comparison. So, the forecast time interval appears between the time interval of 2/1/2011-3/15/2011 as following the least coherent area.

As expected, according to the information based on the least coherent area revealed by wavelet coherence for the time series of IYR and IFEU, the forecast results for both time series show the significant divergence from the actual result presented in Fig 3.17.

(a) IYR-IFGL

For IYR and IFGL time series, we apply the same forecasting method. So, the 30 day forecast is made for the time interval of 4/11/2012-5/22/2012 following the most coherent

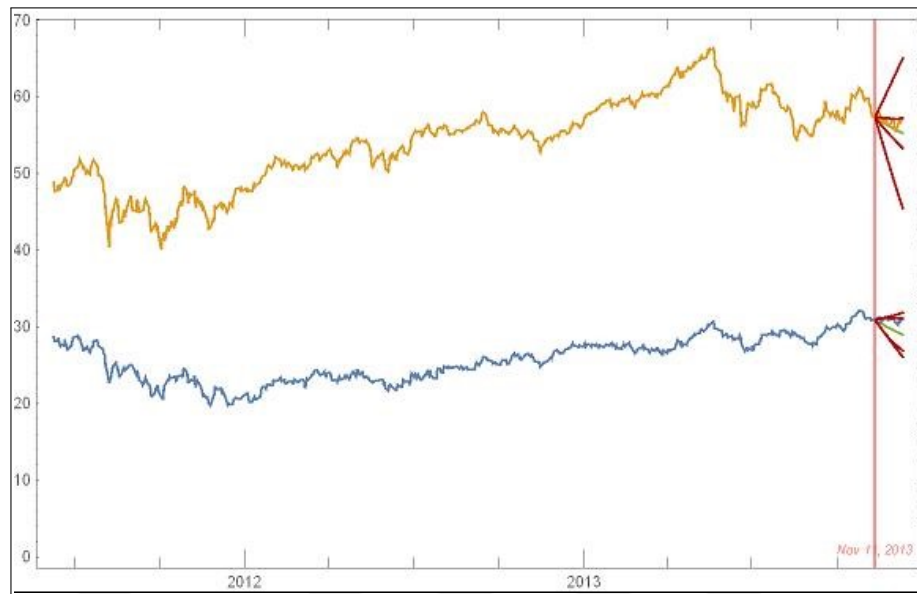


Figure 3.16: Forecast result of IYR-IFEU following the most coherent area.

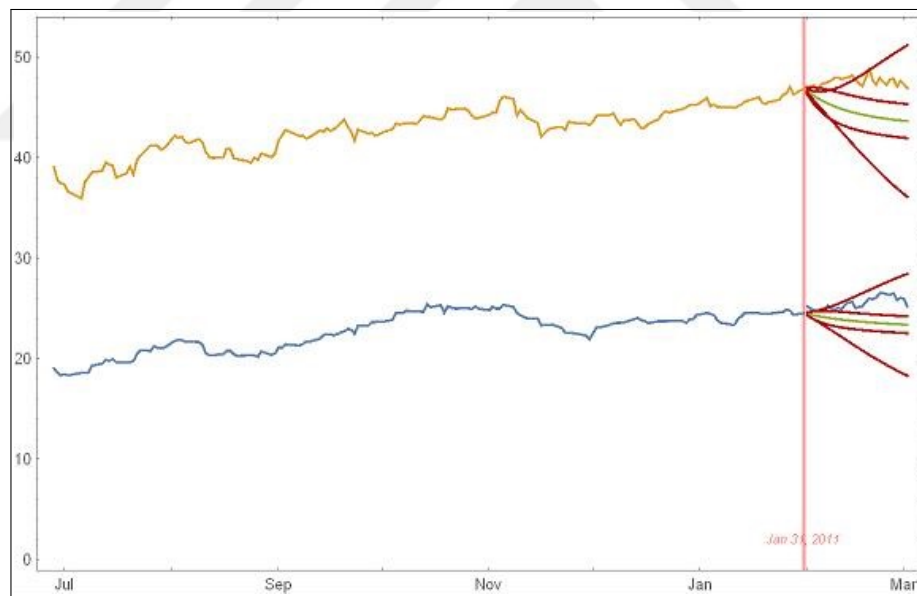


Figure 3.17: Forecast result of IYR-IFEU following the least coherent area.

region appeared at the time interval between 9/8/2010 and 4/10/2012.

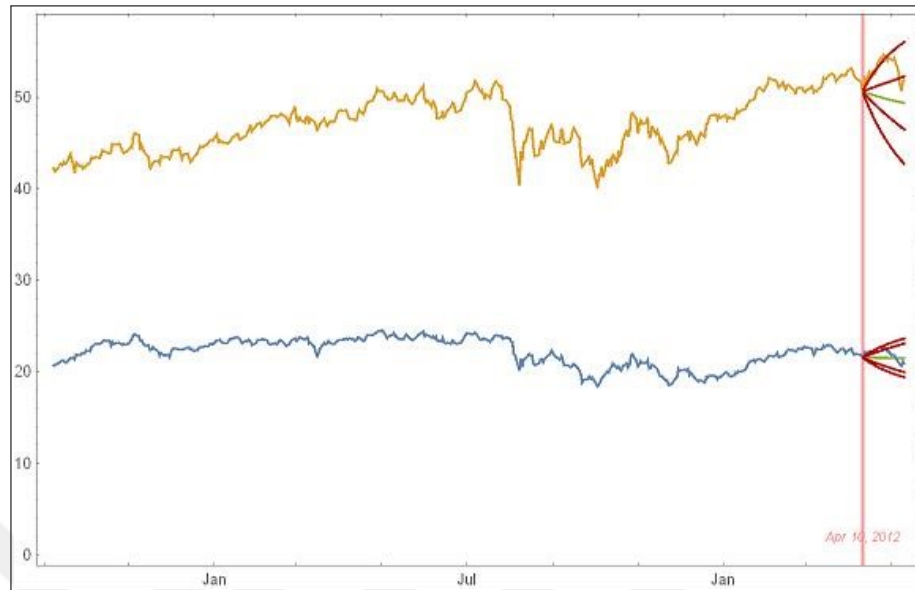


Figure 3.18: Forecast result of IYR-IFGL following the most coherent area.

IYR time series, IFGL time series and the forecast result are portrayed in blue color, in orange color and in green color consecutively in Fig. 3.18. It is shown that the forecast result for IYR appears very close to the actual line.

The forecast results depending on the least coherent region appearing on the time interval of 4/11/2012-1/28/2013 and run for the time period of 1/29/2013-3/13/2013 apparently show that the forecast line and the actual line differs considerably taking into account the scale.

In conclusion, these results strongly refer that the degree of correlations between time series uncovered by wavelet coherence analysis directly affect the efficiency of the forecast results when combined with the VARMA method. They not only asserts the validity of the wavelet coherence results but also the correct modeling of time series for the carefully determined time intervals.

This part of the research is also published (Yilmaz et al., 2016).

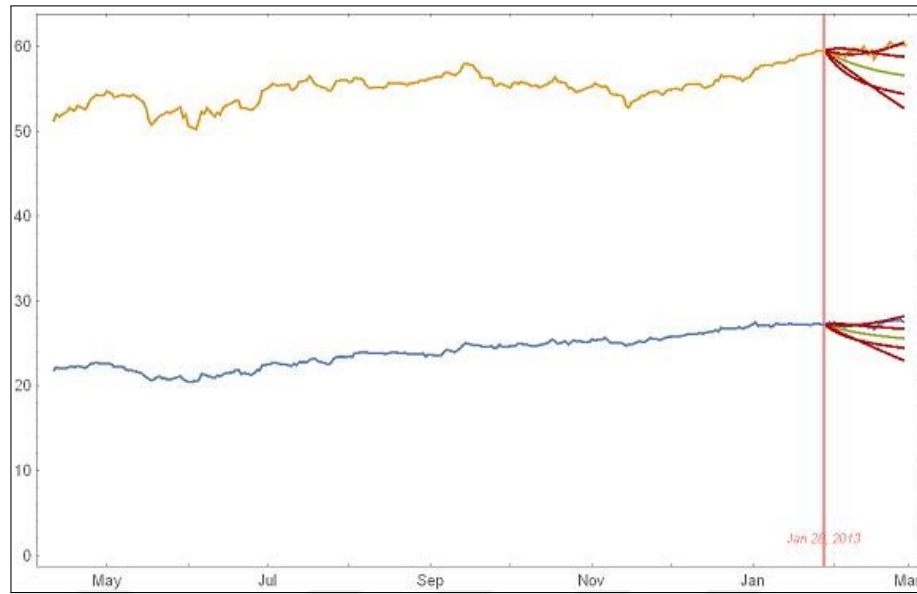


Figure 3.19: Forecast result of IYR-IFGL following the least coherent area.

3.4. CO-MOVEMENT ANALYSIS OF MAJOR REAL-ESTATE MARKETS BY MULTIPLE WAVELET COHERENCE

In this section, we extend wavelet coherence analysis with multiple wavelet coherence (MWC) to be able to show the correlation between three time series. For this purpose, as the time series we again continue with three REITs, China's Hong Kong LINK REIT (0823.HK), iShares US Real Estate ETF (IYR), iShares International Developed Real Estate ETF (IFGL) which are found to be suited by producing valid results by MWC. These time series lay over almost 6.5 year time period from 20/11/2009 to 28/6/2016.

3.4.1. Multiple Wavelet Coherence

Multiple wavelet coherence as an extension of bivariate version of wavelet coherence which was given in Eq. 3.6, introduced and demonstrated in previous section allows to examine the coherent relationships of multiple individual time series to one another. Given $x_1, x_2, \dots, x_p$ as the number of series, the squared multiple wavelet coherence between the series x_1 and

all the other series $x_2, x_3, \dots, x_p$ is defined as (Aguiar-Conraria and Soares, 2014);

$$R_{1(2,3,\dots,p)}^2 = R_{1(q)}^2 = 1 - \frac{M^d}{S_{11}M_{11}^d} \quad (3.9)$$

where M represents the $p \times p$ matrix of all smoothed cross wavelet spectra S_{ij} given as;

$$S_{ij} = S(W_{x_i x_j}) = S_{ji}^* = S(|W_{x_i}|^2) \quad (3.10)$$

$$M = \begin{bmatrix} S_{11} & \dots & S_{1p} \\ \vdots & \ddots & \vdots \\ S_{p1} & \dots & S_{pp} \end{bmatrix} \quad (3.11)$$

The cofactor of the element in the position (i, j) is described as;

$$M_{ij}^d = (-1)^{i+j} \det(M_i^j) \quad (3.12)$$

where M_i^j denotes the submatrix that is obtained by dropping the i th row and j th column of the matrix M .

Eventually, the complex partial wavelet coherency of x_1 and x_j where $2 \leq j \leq p$ can be defined as;

$$\rho_{1j.qj} = -\frac{M_{j1}^d}{\sqrt{M_{11}^d M_{jj}^d}} \quad (3.13)$$

Partial wavelet coherence (PWC) allows to get WTC of two time series by eliminating the effect of the third one. As WTC is analogous to the classic correlation coefficient, PWC is similar to the partial correlation (Grinsted et al., 2004).

The complex coherence matrix which is a $p \times p$ matrix consisting of all smoothed wavelet

coherences ρ_{ij} presented as;

$$C = \begin{bmatrix} 1 & \rho_{12} & \dots & \rho_{1p} \\ \rho_{21} & 1 & \dots & \rho_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{p1} & \rho_{p2} & \dots & 1 \end{bmatrix} \quad (3.14)$$

where $\rho_{ij} = \rho_{ji}^*$ and defined as;

$$\rho_{ij} = \frac{S(W_{ij})}{\sqrt{S(|W_i|^2)S(|W_j|^2)}} \quad (3.15)$$

Then, the squared wavelet coherency is written as;

$$R_{1q}^2 = 1 - \frac{C^d}{C_{11}^d} \quad (3.16)$$

And the complex partial wavelet coherency is represented as;

$$\rho_{1j.q_j} = -\frac{C_{j1}^d}{\sqrt{C_{11}^d C_{jj}^d}} \quad (3.17)$$

As PWC is the absolute value of quantities given in Eq. 3.17, the squared partial wavelet coherency is written as;

$$r_{1j.q_j}^2 = \frac{|C_{j1}^d|^2}{C_{11}^d C_{jj}^d} \quad (3.18)$$

Also, the partial phase difference of x_1 over x_j given all the other series is obtained as;

$$\phi_{1j.q_j} = \text{ArcTan}\left(\frac{I(\rho_{1j.q_j})}{R(\rho_{1j.q_j})}\right) \quad (3.19)$$

Then, in terms of partial coherency, the multiple coherence can be written as (Aguiar-

Conraria and Soares, 2014);

$$R_{1(2,3,\dots,p)}^2 = 1 - R_{1(q)}^2 = (1 - r_{12}^2)(1 - r_{13,2}^2) \dots (1 - r_{1p,23\dots(p-1)}^2) \quad (3.20)$$

So, for three time series, if we arrange the coherence matrix, we get a 3×3 matrix of all the smoothed wavelet coherencies ρ_{ij} ;

$$C = \begin{bmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{21} & 1 & \rho_{23} \\ \rho_{31} & \rho_{32} & 1 \end{bmatrix} \quad (3.21)$$

Then, C_{11}^d can be obtained as;

$$C_{11}^d = \begin{bmatrix} 1 & \rho_{23} \\ \rho_{32} & 1 \end{bmatrix} = 1 - \rho_{23}\rho_{32} \quad (3.22)$$

For an easier representation if we write $\rho_{ij}\rho_{ji} = R_{ij}^2$, C_d can be calculated as;

$$C_d = C_{11}^d - \rho_{12}C_{12}^d + \rho_{13}C_{13}^d$$

$$C_d = 1 - R_{23}^2 - \rho_{12}(\rho_{21} - \rho_{23}\rho_{31}) + \rho_{13}(\rho_{21}\rho_{32} - \rho_{31}) \quad (3.23)$$

$$C_d = 1 - R_{23}^2 - R_{12}^2 - R_{13}^2 + \rho_{12}\rho_{23}\rho_{31} + \rho_{13}\rho_{21}\rho_{32}$$

$$C_d = 1 - R_{23}^2 - R_{12}^2 - R_{13}^2 + 2R(\rho_{12}\rho_{23}\rho_{13}^*)$$

Ultimately, we can write the squared multiple wavelet coherence formula as;

$$R_{1(23)}^2 = 1 - \frac{C_d}{C_{11}^d} = 1 - \frac{1 - R_{23}^2 - R_{12}^2 - R_{13}^2 + 2R(\rho_{12}\rho_{23}\rho_{13}^*)}{1 - R_{23}^2} \quad (3.24)$$

By utilizing the formula in Eq. 3.24, we demonstrate MWC by investigating the relationships

between the three REITs and how they influence each other in next section.

3.4.2. Multiple Wavelet Coherence Results

MWC results for IYR, IFGL and LINK

The plot given in Fig. 3.25 shows the multiple wavelet coherence results of IYR with IFGL and LINK. The x axis and the y axis represent time and frequency consecutively. The vertical bar on the right side of the figure guides about the degree of power of coherence through colors from yellow through blue there in the degree of power decreases in the same order. While, the pale area around the edges indicates the cone of influence, the coherent regions which are outlined by thick black line marks the 5% significance level. As seen,

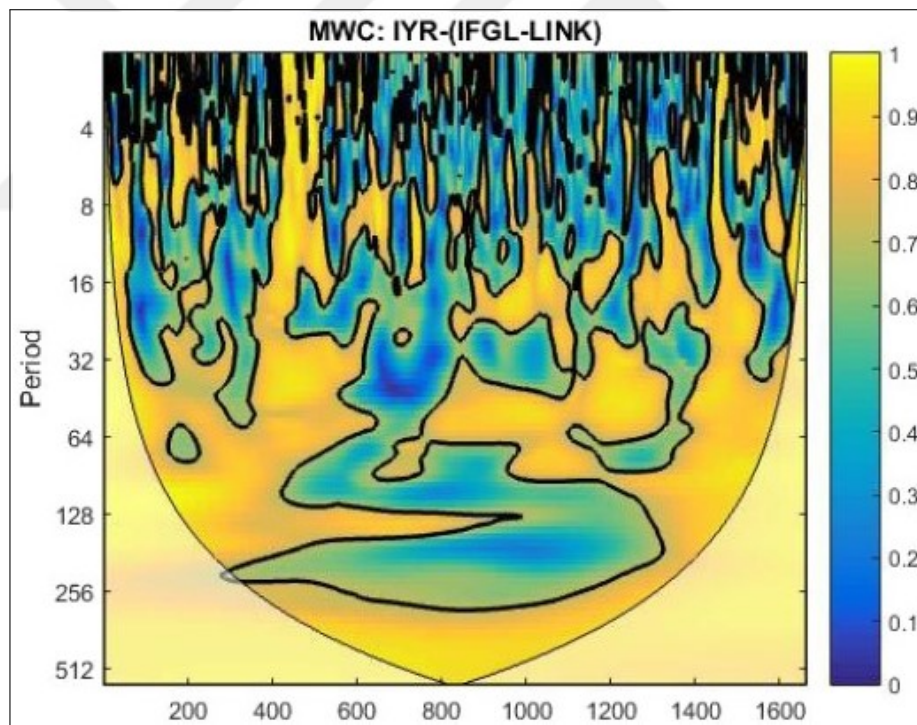


Figure 3.20: MWC results of IYR with IFGL and LINK.

there are various areas where three time series are correlated. The most coherent areas appear starting from the date of 09/08/2010 to 04/10/2012 at the frequencies of 8 to 128 and at time interval between 01/28/2013 and 08/28/2014 at the frequencies of 8-64. On the other hand, the least coherent areas are observed at the time interval of 04/10/2012-01/28/2013

and at the frequencies of 16-64. These dates are located at the following data positions on x-axis of the figure; 09/08/2010-04/10/2012 on x-axis at 200-600, 01/28/2013-08/28/2014 on x-axis at 800-1200, 04/10/2012-01/28/2013 on x-axis at 600-800. Fig. 3.24 shows the

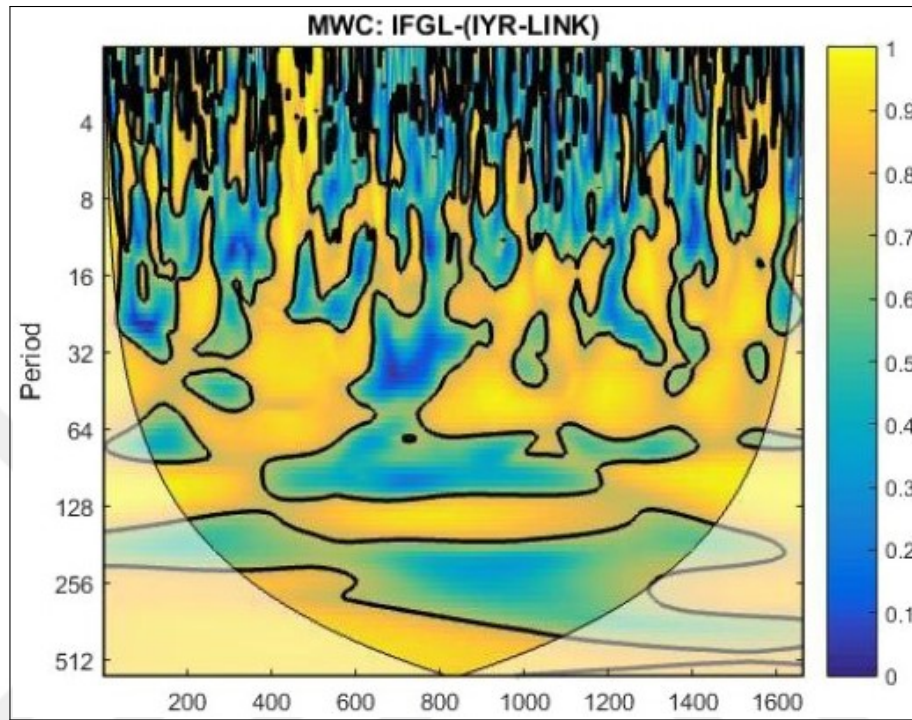


Figure 3.21: MWC results of IFGL with IYR and LINK.

MWC results of IFGL with IYR and LINK. It is observed that the most coherent regions are at time period between 09/08/2010 and 04/10/2012 at the frequencies of 16 to 64 and at time period of 01/28/2013-11/05/2015 at the frequencies of 8 to 64. Also, the lowest power of coherency is appeared to be at the time interval of 04/10/2012-01/28/2013 at almost all the frequencies. These dates are located at the following data positions on x-axis of the figure; 09/08/2010-04/10/2012 on x-axis at 200-600, 01/28/2013-11/05/2015 on x-axis at 800-1500, 04/10/2012-01/28/2013 on x-axis at 600-800. The MWC results of LINK with IYR and IFGL are given in Fig. 3.23. When the most coherent region is observed at the time interval of 01/28/2013-08/28/2014, the least coherent area appears to be at the time period between 04/10/2012 and 01/28/2013. Data positions of the dates are; 01/28/2013-08/28/2014 on x-axis at 800-1200, 04/10/2012-01/28/2013 on x-axis at 600-800.

In summary, as IYR appears to be the most correlated time series with other two time

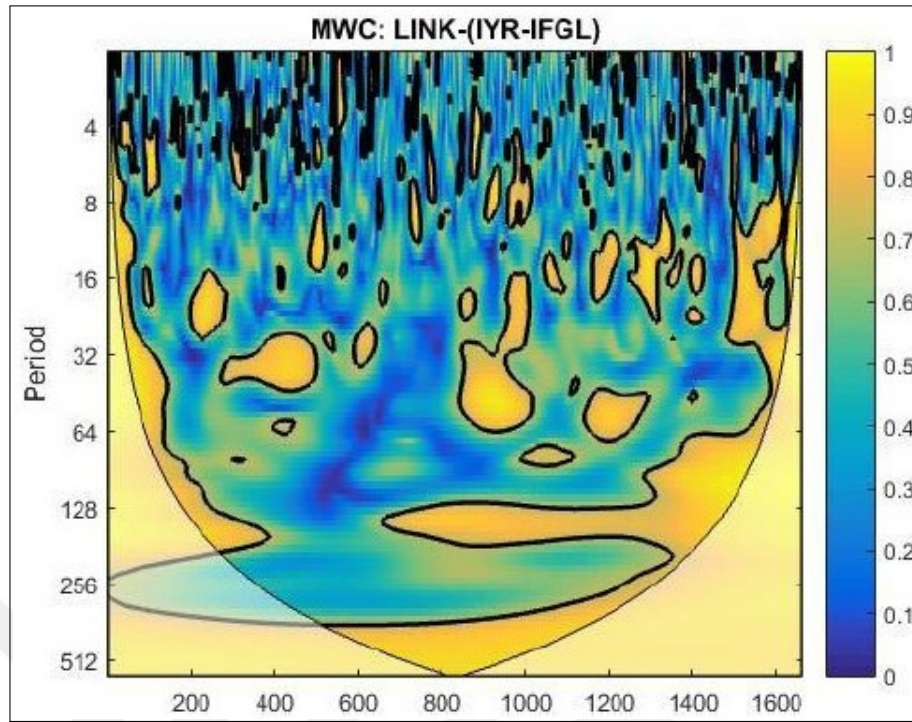


Figure 3.22: MWC results of LINK with IYR and IFGL.

series, LINK seems to be the least correlated one with the other two time series. There are shared correlation regions observed. The most coherent regions occur at the time interval between 01/28/2013-08/28/2014 at the frequencies of 8-128. Besides, the least coherent regions appears in the time interval between 04/10/2012 and 01/28/2013. We utilize these findings for efficient forecasting with VARMA method as demonstrated in previous section with WTC.

3.4.3. The Comparison of MWC and WTC Results

In this section, we compare WTC and MWC result side by side to show how different and similar information is gathered by these two methods. Firstly, in Fig. 3.23 we present the comparison of WTC (IYR-IFGL) and MWC (LINK-(IYR-IFGL)) results. It is observed that the affect of IFGL over IYR is dominated by LINK at the time interval of 08/28/2014-04/01/2016 (x-axis 1200-1600) at short frequencies of 4-64 days and at the time interval of 06/23/2011-04/10/2012 (x-axis 400-600) at the frequencies of 4-32 days. Besides, following the LINK's involvement, IFGL's influence on IYR vanishes during the time interval of

01/28/2013-08/28/2014 (x-axis 800-1200) at the frequency of 128 days. Fig. 3.24 shows the

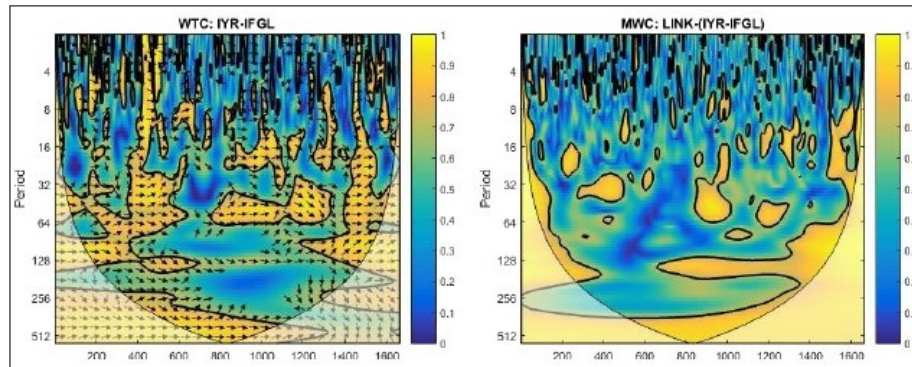


Figure 3.23: The comparison of WTC (IYR-IFGL) and MWC (LINK-(IYR-IFGL)) results

comparison of WTC (IYR-LINK) and MWC (IFGL-(IYR-LINK)) results. It is observed that after inserting IFGL, LINK's influence on IYR drops during all the time period and all the frequencies except 256-512 day cycles. So, it implies that IFGL has high power correlations with IYR in comparison to LINK.

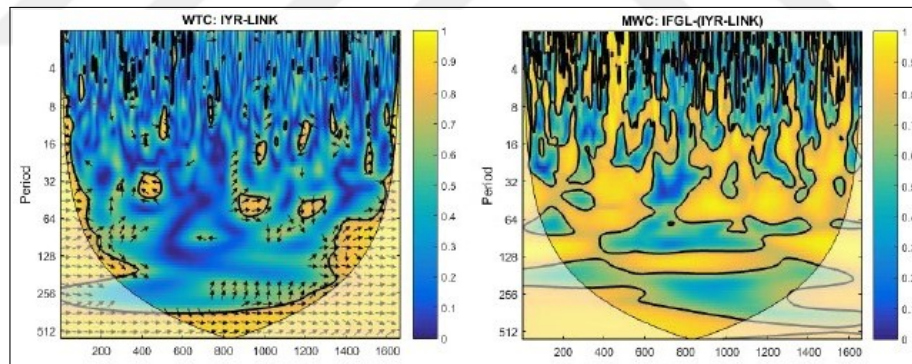


Figure 3.24: The comparison of WTC (IYR-LINK) and MWC (IFGL-(IYR-LINK)) results

The comparison of WTC (IFGL-LINK) and MWC (IYR-(IFGL-LINK)) results is given in Fig. 3.25. It is seen that the influence of LINK on IFGL is lost after the addition of IYR which especially dominates at the time period 01/28/2013-08/28/2014 (x-axis 800-1200) at the frequency of 128.

In conclusion, it can be stated that IYR has the most dominant effect on other two time series of IFGL and LINK which has the least effect in comparison.

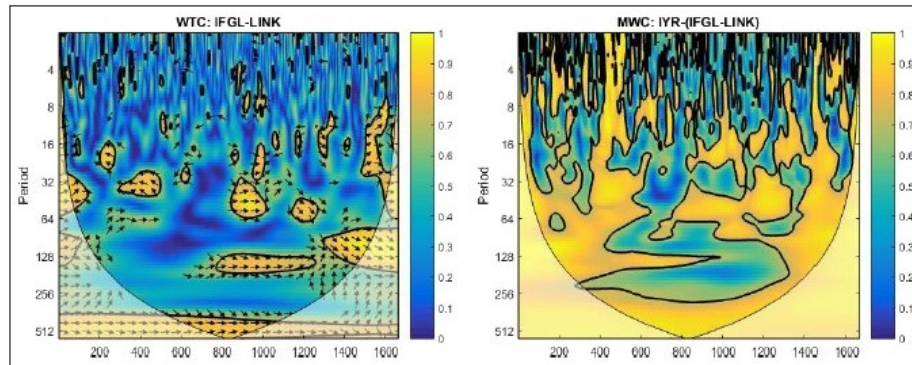


Figure 3.25: The comparison of WTC (IFGL-LINK) and MWC (IYR-(IFGL-LINK)) results

3.4.4. Forecast with Multivariate ARMA Process

In this section, we demonstrate the efficient applicability of the MWC outcomes with VARMA method.

IYR-IFGL

For the forecast demonstration of IYR-IFGL-LINK, as applied in previous section, we make use of the information revealed by MWC in the most coherent region emerged in between the dates of 01/28/2013 and 8/28/2014 in order to determine the following 30 day forecast period which is the time interval between 08/29/2014 and 09/10/2014.

The forecast results given in Fig. 3.26 for three time series of IYR, IFGL and LINK shows that the forecast line follows the actual line quite closely in error bands.

After the demonstration of the forecast method for the most coherent region, now we apply it for the least coherent region detected by MWC as in the time interval of 4/11/2012-01/28/2013. Therefore, the 30 day-forecast period arises in the time interval of 01/29/2013-03/12/2013. As presented in Fig. 3.27, the forecast results are not as good as the forecast made after the most coherent region by diverging from the actual line and appearing outside of the error bands.

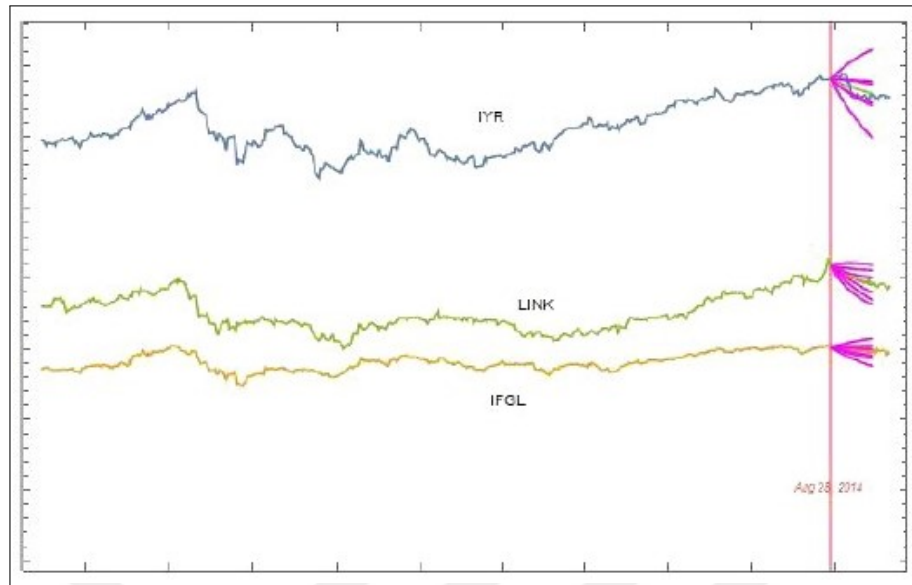


Figure 3.26: The forecast results of IYR-IFGL-LINK after the most coherent region

Therefore, in the light of these results, once again we can verify the efficiency of the forecast method that is the combination of the wavelet coherence and VARMA.

This part of the research is also published (Karatas et al., 2017).

3.5. FINAL REMARKS

Overall, with the analysis of wavelet coherence methods, we explored the new ways of observing various exciting relationships of nonstationary time series in time-frequency space in a more detailed way than the conventional methods. We have constructed a new and efficient forecasting method by combining the wavelet coherence methods and VARMA and demonstrated it in a number of real-world time series.

We analysed S&P 500 and FTSE100 stock markets with several major Asian stock markets by wavelet coherence method and find out how they influence each other and how they change over time. Then, we applied the wavelet coherence to paired REIT time series to determine the time intervals and the frequencies in which dynamic correlations vary in different degrees. And then we added the multiple wavelet coherence for the analysis of

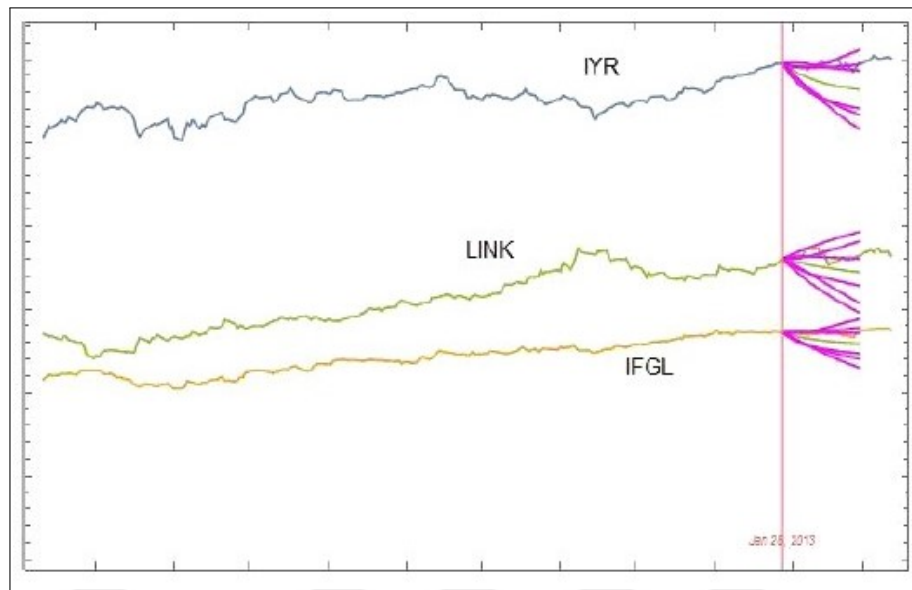


Figure 3.27: The forecast results of IYR-IFGL-LINK after the least coherent region

dynamic correlations between more than two time series all together. By the guidance of these findings, we utilized VARMA for the forecast of certain time periods and demonstrated the efficiency of the method by comparing it for the strongly coherent and weakly coherent regions. We demonstrated the forecast efficiency by experimenting the wavelet coherence-VARMA method on comparing the forecast results with the actual outcomes for the strongly correlated and weakly correlated regions revealed by wavelet coherence. So, these favorable results affirmed the validity of the wavelet coherence revelations and the appropriate modeling of the time series for the carefully determined time intervals.

These relationships yield very useful information, particularly, for the portfolio diversification and risk elimination. From investor's perspective, as a result of such relationships, it can be beneficial to diversify assets like these developed economy market stocks with developing Asia economy market stocks or REITs by taking into account time periods and correlation frequencies and forecast results accordingly.

4. NONLINEAR TIME SERIES ANALYSIS

4.1. THE CHAOTICITY PROPERTIES OF FIGARCH

There have been a number of scientific studies which examine the existence of chaotic behavior in financial and economic time series (Dechert and Gençay, 2000; Das and Das, 2006, 2007; Moeini et al., 2007; Günay, 2015). Learning about the chaotic features in a time series helps to understand the short-term predictability and controllability of the underlying difference equations (Abarbanel, 2012). Therefore, chaos theory has been drawing the attention of science for decades. Generalized autoregressive conditional heteroscedasticity (GARCH) has also been very popular in finance and economy. It is a set of nonlinear stochastic difference equations and so, it is very hard to associate it with the deterministic chaos.

Fractionally integrated generalized autoregressive conditional heteroscedasticity is a modified version of GARCH model. It was introduced by Baillie, Bollerslev and Mikkelsen in 1996 (Baillie et al., 1996) to provide more persistence on the conditional variance. It allows slow hyperbolic rate of decay for the innovations in the conditional variance and has an ability to estimate long memory of conditional volatility. There have been many researches that show strong evidences of long-range dependence in a variety of financial assets and long memory in volatility (Cajueiro and Tabak, 2008). So, in this section, the chaotic characteristics of FIGARCH model is examined by utilizing the correlation dimension and Lyapunov exponent methods.

As mentioned before, since its introduction four decades ago, chaos theory has been appealing to many researchers and the existence of chaotic behavior has been studied in variety of scientific areas, ranging from atmospheric dynamics (Lorenz, 1963; Essex et al., 1987), geophysics (Sivakumar, 2004), medicine (West, 2012), turbulence (Manneville, 2010) to financial markets (Trippi et al., 1994) and so on.

Chaotic systems are deterministic systems which are unpredictable in the long term because

of their sensitivity to even a very small change of initial conditions. Irregularity can be autonomously generated by the nonlinearity of the intrinsic dynamics in the deterministic frame. The most direct link between chaos theory and the real world is the analysis of time series in terms of nonlinear dynamics. Chaos theory has influenced a new set of effective time series tools and supplies a new language to formulate time series problems (Schreiber, 1999).

In this section, the presence of chaoticity in nonlinear FIGARCH model is investigated by using simulated time series and nonlinear difference equation directly. Firstly, as assuming that FIGARCH is a deterministic chaotic system, the phase space is reconstructed by utilizing delay coordinate embedding method by Takens' theorem. The reconstructed phase space is the one-to-one picture of the original time series and hold the same mathematical features. Takens' theorem validates that with proper delay time and embedding dimension.

So, in order to estimate the proper time delay, the mutual information method is exercised. Then, the embedding dimension is determined with the false nearest neighbor method. After that the correlation dimension which quantifies the self-similarity is measured by using the Grassberger and Procaccia's procedure. Following this, Lyapunov exponent is calculated to quantify the sensitivity to initial conditions that is the most fundamental characteristic of chaos. For this aim, there are two algorithms, Wolfs' and Kantz's algorithms are used to the simulated FIGARCH time series with employing a more direct technique by forming a dimensional map from FIGARCH difference equation.

4.1.1. FIGARCH

As mentioned in introduction, Generalized Autoregressive Conditional Heteroscedastic (GARCH) model and Integrated version (IGARCH) model were introduced by Bollerslev (1986) (Bollerslev, 1986) and Engle and Bollerslev (1986) (Engle and Bollerslev, 1986) consecutively. There are various problems with GARCH model like non-negativity and leverage effects. Also, it doesn't let for any straight feedback between the conditional variance and the conditional mean. And for the IGARCH model, it appears to be too restrictive while it involves infinite persistence of a volatility shock. As a solution of some

of these problems, fractionally integrated GARCH model was introduced by Baillie et al. in 1996 (Baillie et al., 1996). FIGARCH allows persistence in the conditional variance and is more flexible on the conditional variance and so more capable of describing and representing the observed temporal dependencies in financial market volatility.

FIGARCH(p, d, q) model can be obtained by replacing the first difference operator with fractional differencing operator in GARCH (p, q) and defined as;

$$\phi(L)(1-L)^d u_t^2 = \omega + [1 - \beta(L)]\varepsilon_t \quad (4.1)$$

where $0 < d < 1$ and all the roots of $\phi(L)$ and $[1 - \beta(L)]$ lie outside the unit circle. It also can be expressed after rearranging the conditional variance as $\varepsilon = u_t^2 - \sigma_t^2$;

$$[1 - \beta(L)]\sigma_t^2 = \omega + [1 - \beta(L)] - \phi(L)(1-L)^d u_t^2 \quad (4.2)$$

$$(1 - \sum_{j=1}^q \beta_j L^j)\sigma_t^2 = \omega + (1 - \sum_{j=1}^q \beta_j L^j - (\sum_{k=1}^p \phi_k L^k)(1-L)^d)u_t^2 \quad (4.3)$$

Then the conditional variance of u_t is presented as;

$$\sigma_t^2 = \omega[1 - \beta(L)]^{-1} + \{1 - [1 - \beta(L)]^{-1}\phi(L)(1-L)^d\}u_t^2 \quad (4.4)$$

$$\sigma_t^2 \equiv \omega[1 - \beta(L)]^{-1} + \lambda(L)u_t^2$$

where $\lambda(L) = \lambda_1 L + \lambda_2 L^2 \dots$ and for the FIGARCH(p, d, q) in Eq. 4.1 to be well-defined and the conditional variance to be positive absolutely for all t , all the coefficients in the infinite ARCH representation in Eq. 4.4 must be non-negative (Baillie et al., 1996). Later, a set of conditions which secures the non-negativity of the conditional variance in all situations were provided (Conrad and Haag, 2006). Also, it is shown that FIGARCH model have more memory than a GARCH or IGARCH model (Davidson, 2004).

4.1.2. Chaotic Behavior and FIGARCH

Basically, chaotic behavior can be described as irregularity of motion, unpredictability and sensitivity to initial conditions. FIGARCH nonlinear stochastic difference equations are examined in following sections to determine whether or not it possess any chaotic properties to be able to learn about its predictability horizon and forecasting efficiency.

Lyapunov exponent is a measure of the average speed with which infinitesimally close states separate. It is used to measure the presence of chaos in a time series. Therefore, in order to calculate Lyapunov exponent for FIGARCH, firstly it is necessary to convert scalar time series to the multivariate state or phase space that is needed for chaotic motions to exist in the first place.

For the analysis, a number of FIGARCH(1, d , 1) simulations are produced. In order to be able to generate such number of time series with varying d values as many as possible, α , β and ω coefficients are fixed at 0.01. As d values varying between 0.05 and 0.9, total 50 number of random samples with the length of 8192 were generated by employing MFE toolbox (Sheppard, 2006) and Oxmetrics GARCH package (Oxmetrics, 2006).

4.1.3. Phase Space Reconstruction

The embedding theorem provides solution to conversion from scalar to multivariate state (Takens, 1981). Letting $\{x_0, x_1, \dots, x_i, \dots, x_n\}$ a time series, according to Takens embedding theorem, the reconstructor of the original system is written as a vector sequence;

$$p(i) = (x_i, x_{i+\tau}, x_{i+2\tau}, \dots, x_{i+(d-1)\tau}) \quad (4.5)$$

where τ and d denote the embedding delay and the embedding dimension successively. For a large enough d , many important features of the original system can be reproduced without uncertainty in the new space of vectors as Takens asserted. Specifically, the attractor constructed will possess the same mathematical features as the original system such as dimension, Lyapunov exponents and so on. Therefore, determination of the appropriate

values of the embedding delay and the embedding dimension is very crucial to be able to reconstruct the attractor.

4.1.4. Mutual Information

Embedding delay τ value is found by observing the first minimum of the nonlinear correlation function called the mutual information (Fraser and Swinney, 1986) between x_i and $x_{i+\tau}$ as a appropriate quantity for determining τ . There are two essential bases for the computation of the embedding delay. Firstly, τ should be large enough to provide that the information between $x_{i+\tau}$ and x_i is significantly different. Secondly, τ shouldn't be too large that $x_{i+\tau}$ and x_i are totally independent statistically.

Let the sets of $A = \{a_i\}$ and $B = \{b_j\}$, then the mutual information between them in bits can be written as;

$$I_{AB} = \log_2 \left[\frac{P_{AB}(a_i, b_j)}{P_A(a_i)P_B(b_j)} \right] \quad (4.6)$$

where $P_A(a_i)$ and $P_B(b_j)$ denotes the distinctive probability densities. Also, $P_{AB}(a_i, b_j)$ represents the joint probability density of A and B which is zero with the mutual information, if a_i and b_j are completely independent. The average of all the calculation of information statistic between A and B is given as;

$$I_{AB} = \sum_{a_i, b_j} P_{AB}(a_i, b_j) \log_2 \left[\frac{P_{AB}(a_i, b_j)}{P_A(a_i)P_B(b_j)} \right] \quad (4.7)$$

Mutual information calculates the mutual dependence of two sets dependent on the notion of the information between them. Therefore, for the calculation of $s(t)$ at time t that are connected to the calculation of $s(t + \tau)$, if it is rewritten;

$$I(\tau) = \sum_{s(n), s(n+\tau)} P(s(n), s(n + \tau)) \log_2 \left[\frac{P(s(n), s(n + \tau))}{P_A(s(n))P_B(s(n + \tau))} \right] \quad (4.8)$$

Therefore, when $\tau \rightarrow \infty$, then $I(\tau) \rightarrow 0$ considering that the correlation between $s(n)$ and $s(n + \tau)$ fades (Abarbanel, 2012). It is suggested that because $I(\tau)$ is a type of autocorrelation function, so it is convenient to select time delay value at first minimum of the mutual information. However, it sometimes might be ambiguous taking into account the linear nature of the method (Fraser and Swinney, 1986).

Table 4.1: The embedding delay for each FIGARCH model

FIGARCH (d) value	Embedding delay
0.05	2
0.15	1
0.25	2
0.35	4
0.45	6
0.55	7
0.65	8
0.75	8
0.80	6
0.90	8

So, in order to determine the convenient embedding delay, the mutual information for each model is calculated like as shown in Fig. 4.1 and outcomes are presented in Table 4.1.

4.1.5. False Nearest Neighbor Method

Given that $y(i)$ is a reconstructed state space in dimension d with data vectors by the time delay determined by mutual information.

$$y(i) = (x_i, x_{i+\tau}, x_{i+2\tau}, \dots, x_{i+(d-1)\tau}) \quad (4.9)$$

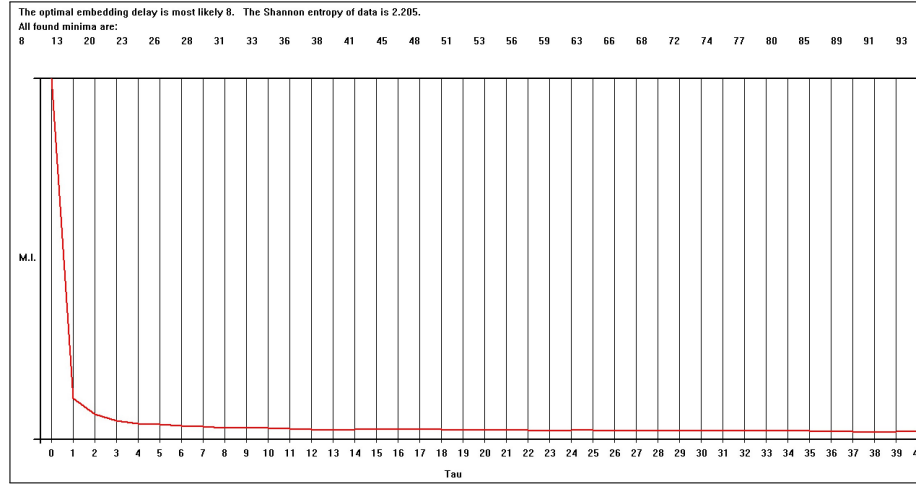


Figure 4.1: Mutual information result for FIGARCH $d=0.90$

Then the nearest neighbor in phase space is a vector given as;

$$y^{NN}(i) = (x_i^{NN}, x_{i+\tau}^{NN}, x_{i+2\tau}^{NN}, \dots, x_{i+(d-1)\tau}^{NN}) \quad (4.10)$$

Supposing that the vector $y^{NN}(i)$ is a false neighbor of $y(i)$ having reached its neighborhood by projection from a higher dimension considering the current dimension d doesn't unfold the attractor, then this false neighbor might be eliminated from the neighborhood of $y(i)$ by moving to the next dimension $d + 1$.

So, by observing every data point of $y(i)$ and noting that at what dimension, all the false neighbors are eliminated, consecutively the intersections of orbits of lower and lower dimensions are eliminated until the last point intersections are removed. At that juncture d is determined where the attractor is unfolded.

By comparing the distance between the vectors $y(i)$ and $y^{NN}(i)$ in dimension d with the distance between the same vectors in dimension $d + 1$, true and false neighbors can be determined. The comparison of $x_{i+d\tau} - x_{i+d\tau}^{NN}$ with the Euclidian distance $|y_i - y_i^{NN}|$ between the nearest neighbors in dimension d will be sufficient. A false neighbor is identified if the additional distance is greater compared to the distance in dimension d between the nearest neighbors. The relationship of the square of the Euclidian distance

between the nearest neighbors in dimension d and $d + 1$ is;

$$R_d^2(i) = \sum_{m=1}^d [x_{i+(m-1)\tau} - x_{i+(m-1)\tau}^{NN}]^2$$

$$R_{d+1}^2(i) = \sum_{m=1}^{d+1} [x_{i+(m-1)\tau} - x_{i+(m-1)\tau}^{NN}]^2 \quad (4.11)$$

$$R_{d+1}^2(i) = R_d^2(i) + |x_{i+d\tau} - x_{i+d\tau}^{NN}|^2$$

The false neighbor is found when;

$$\sqrt{\frac{R_{d+1}^2(i) - R_d^2(i)}{R_d^2(i)}} = \frac{|x_{i+d\tau} - x_{i+d\tau}^{NN}|}{R_d(i)} > r_{tol} \quad (4.12)$$

where r_{tol} is some threshold (Abarbanel, 2012).

When the percentage of false neighbors are illustrated, the unfolded geometry and where there is no unfolding any more can be observed. With the adequate selection of d dimension, modelling the system in d number of dynamical degrees of freedom will be sufficient to obtain the properties of the original system.

Fig. 4.2 presents the minimum embedding dimension where the percentage of the nearest neighbors drops zero with the threshold r_{tol} that is false neighbor Euclidian distance tolerance and a_{tol} that is the neighbor tolerance based on the attractor size. When the ratio of the Euclidian distances between neighbor candidates in consecutive embedding dimensions surpass r_{tol} , the neighbors are affirmed as false neighbors.

4.1.6. Correlation Dimension

Correlation dimension is one of the most accepted tools to be used for the analysis of degree of complexity. There is a very efficient method developed by Grassberger and Procaccia in 1983 (Grassberger and Procaccia, 1983b) that measure an attractor based on a contraction

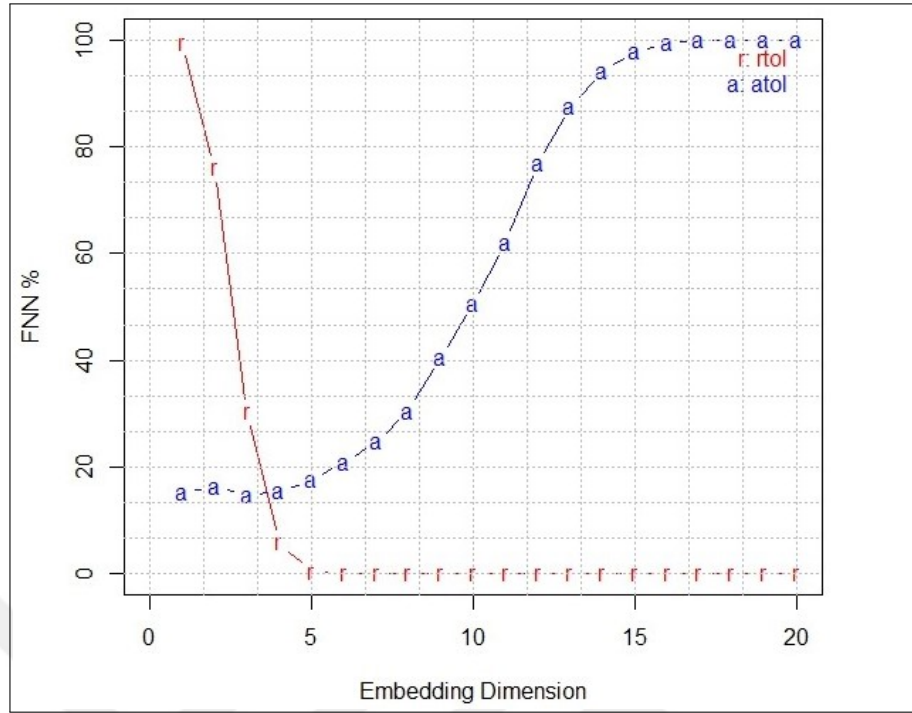


Figure 4.2: FNN embedding dimension result for FIGARCH d=0.90

rate of a fractal measure in some phase space in a given set. They describe the correlation sum that approximates the probability of possessing pair of points with separation distance less than a given size of ε as;

$$C(\varepsilon) = \frac{1}{N^2} \sum_{i,j}^N \Theta(\varepsilon - \|x_i - x_j\|) \quad (4.13)$$

where Θ is the Heaviside step function shown as;

$$\begin{cases} \Theta(\varepsilon - \|x_i - x_j\|) = 1 & (\varepsilon - \|x_i - x_j\|) \geq 0 \\ 0 & (\varepsilon - \|x_i - x_j\|) < 0 \end{cases} \quad (4.14)$$

For small values of ε , when $N \rightarrow \infty$, C follows a power law;

$$C(\varepsilon) \propto \varepsilon^{D_C} \quad (4.15)$$

where D_C denotes the correlation dimension which further can be written as;

$$D_C = \lim_{\varepsilon \rightarrow 0} \lim_{N \rightarrow \infty} \frac{\partial \ln C(\varepsilon)}{\partial \ln(\varepsilon)} \quad (4.16)$$

For the estimation of correlation dimension, the least square fit in a plot of $\ln C(\varepsilon) - \ln(\varepsilon)$ is applied for the estimation of the slope due to the slow convergence.

For each FIGARCH simulation, for embedding delays ranging from 1 to 20 and embedding dimensions ranging from 1 to 20 correlation dimension calculations are made and results are given in Fig. 4.3. Also, correlation dimension values for the embedding dimensions

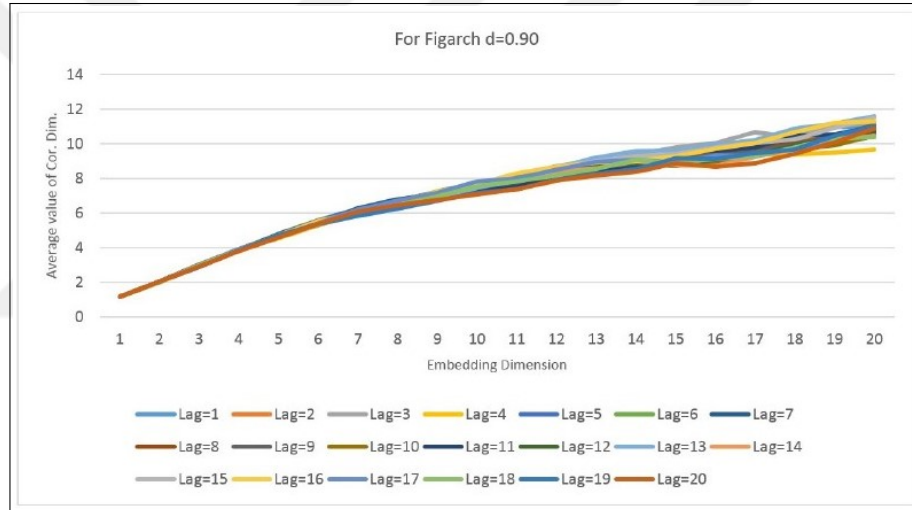


Figure 4.3: Correlation dimension results of FIGARCH d=0.90 for embedding delays (1-20) and embedding dimensions (1-20)

(1-20) for each FIGARCH simulation is calculated as to the embedding delay determined by mutual information and presented in Fig. 4.4. It is seen that there is no convergence at any of the embedding dimension values for any of the model. So, embedding dimensions can't be determined. The best converged lines for each FIGARCH model is also given in Fig. 4.5. Moreover, to examine if there is any convergence in correlation dimension values for different embedding delay values, for several embedding delay values (1 to 20) and for embedding dimensions (1 to 20) calculations are made but again there is no convergence detected.

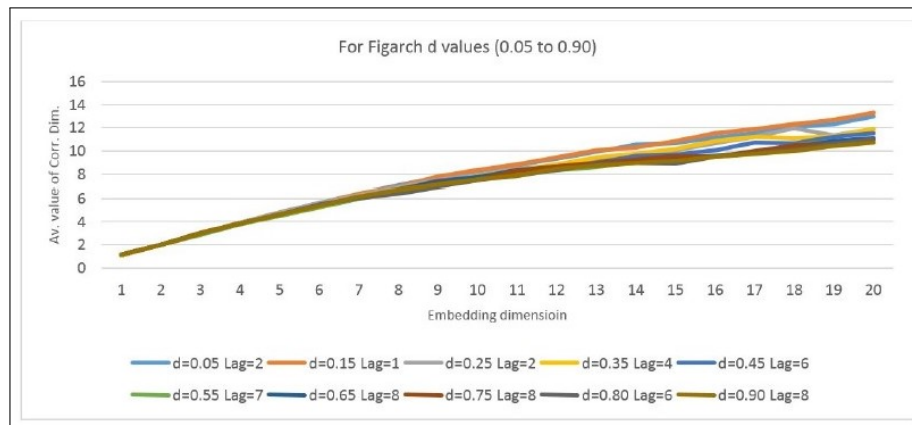


Figure 4.4: Correlation dimension results for each FIGARCH model according to the embedding delays determined by mutual information

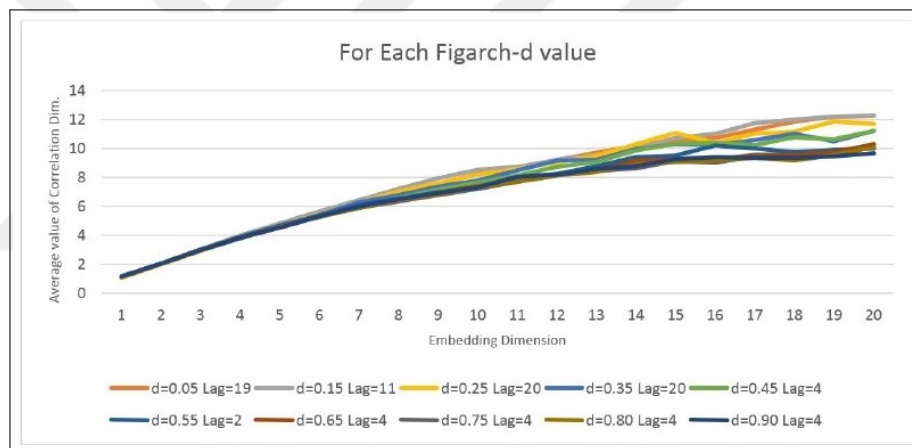


Figure 4.5: The best converged lines of correlatin dimension results for each FIGARCH model

It is suggested that for the determination of embedding dimension, the application of correlation dimension and the false nearest neighbor method work well for low dimensional chaotic attractors. On the other hand, for high dimensional attractors, the convergence of the false nearest method might provide better results (Fraser and Swinney, 1986). Therefore, assuming that there are high dimensional attractors, the false nearest neighbor results are accepted as sufficient for further analysis.

4.1.7. Lyapunov Exponent

As a parameter that describes the behavior of a dynamical system, the Lyapunov exponent provides the average rate of the exponential divergence from nearby initial conditions. A positive Lyapunov exponent indicates the chaotic system. The system with negative Lyapunov exponent converges to a periodic state and if the Lyapunov exponent is zero, there is a bifurcation.

The largest Lyapunov exponent is one of the most widely used tool to investigate the chaotic behavior. If the attractor is chaotic, on average, the trajectories deviate, at an exponential rate described by the largest Lyapunov exponent.

4.1.7.1. Wolf's Algorithm

For the calculation of the largest Lyapunov exponent in a direct way from time series, Wolf et al. in 1985 (Wolf et al., 1985) introduced an algorithm given as;

$$\Lambda_{max} = \frac{1}{Mt_{evolv}} \sum_{i=0}^M \ln \frac{L_{evolv}^{(i)}}{L_0^{(i)}} \quad (4.17)$$

where M denotes the total number of replacement steps, t_{evolv} denotes the fixed evolution time with the same order of magnitude as the embedding delay, L_{evolv} represents the final distance between the evolved points and L_0 represents the Euclidian distance between the nearest neighbors of initial point.

As the embedding delay and the embedding dimension are already determined, then the largest Lyapunov exponent (LLE) by employing Wolf's algorithm can be computed. Fig. 4.6

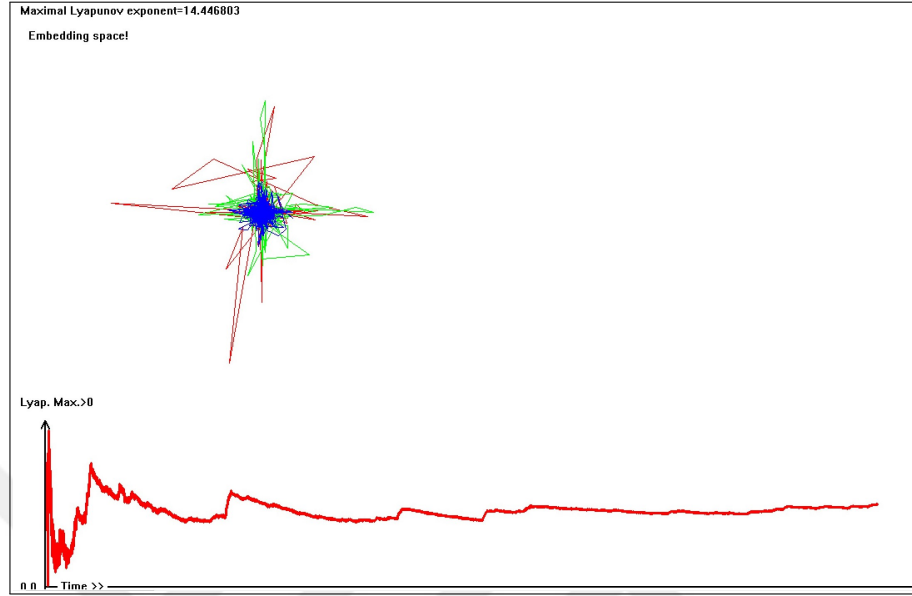


Figure 4.6: The largest Lyapunov exponent result for FIGARCH d=0.90

displays the largest Lyapunov result for FIGARCH d=0.90 which converges to positive values that is also the outcome for all the other FIGARCH models simulated given in Table 4.2. So, based on these results it can be suggested the existence of sensitivity to changes in initial conditions that is the implication of chaotic behavior.

4.1.7.2. *Kantz's Algorithm*

Another LLE algorithm was proposed by Kantz in 1994 (Kantz, 1994) and Rosenstein in 1993 (Rosenstein et al., 1993). To estimate LLE, all neighbors closer to a reference point than a given size ε is determined and then average distance of all trajectories to the reference trajectory is computed. The algorithm is described as;

$$S(t) = \frac{1}{N} \sum_{i=1}^N \ln \left[\frac{1}{|\Omega_i|} \sum_{j \in \Omega_i} |p_i - p_j| \right] \quad (4.18)$$

where p_i denotes the embedding vectors and $|\Omega_i|$ represents the number of neighbors in the neighborhood of Ω_i of the reference state p_i .

Table 4.2: The largest Lyapunov exponent results for all the FIGARCH models with Wolf's algorithm

FIGARCH (d) value	Embedding delay	Embedding dimension	LLE
0.05	2	6	9.1
0.15	1	7	2.27
0.25	2	6	11.85
0.35	4	8	2.38
0.45	6	6	7.6
0.55	7	6	24.5
0.65	8	6	22.17
0.75	8	6	20.8
0.80	6	6	21.1
0.90	8	6	14.4

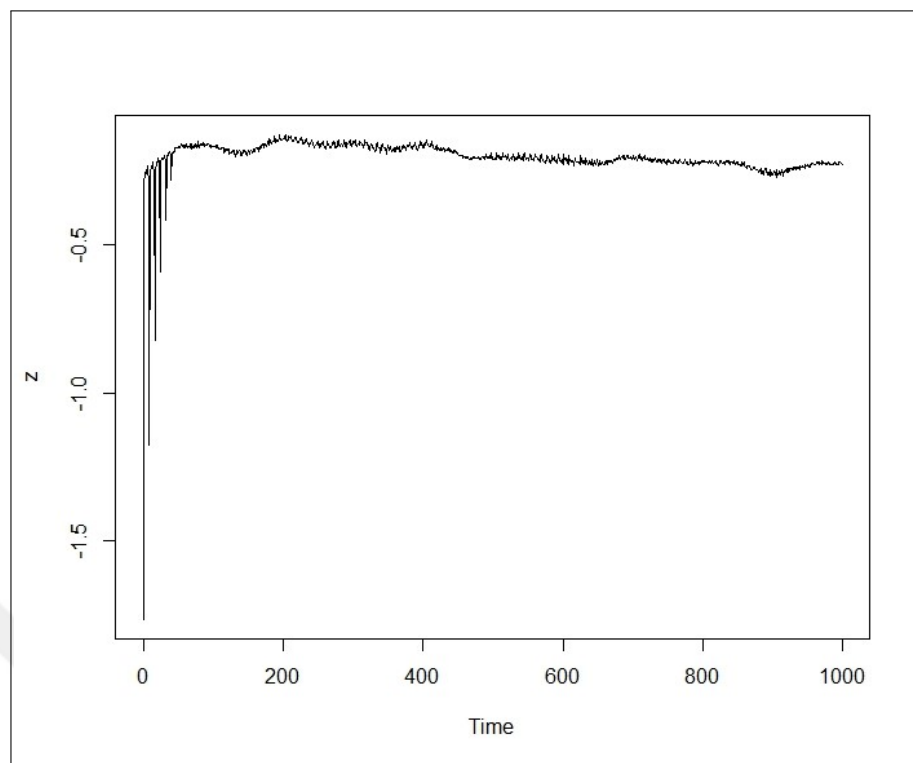


Figure 4.7: The largest Lyapunov exponent result with Kantz's algorithm for FIGARCH $d=0.90$

As seen in Fig. 4.7, the results of Kantz's algorithm in contrast to Wolf's algorithm results yield negative LLEs for all the FIGARCH models given in Table 4.3 disagreeing with that FIGARCH is a deterministic chaotic system.

Table 4.3: The largest Lyapunov exponent results for all the FIGARCH models with Kantz's algorithm

FIGARCH (d) value	Embedding delay	Embedding dimension	LLE
0.05	2	6	-3.38E-05
0.15	1	7	-6.40E-05
0.25	2	6	-7.47E-05
0.35	4	8	-7.93E-05
0.45	6	6	-8.19E-05
0.55	7	6	-3.50E-06
0.65	8	6	-1.14E-04
0.75	8	6	-1.19E-04
0.80	6	6	-6.19E-05
0.90	8	6	-1.24E-04

4.1.7.3. Direct Calculation with the 2-Dimensional Map

Lastly, as another method, the calculation of LLE is made by employing the dynamical rules of the 2-dimensional map directly from the difference equations rather than the simulated data that was used the calculations of Wolf's and Kantz's algorithm.

For this purpose, for FIGARCH $\{x_1(i), y_1(i)\}$ and for $i = 1$ to 5000, the phase trajectory is constructed. At every point, local Lyapunov exponent value is calculated along the trajectory and then all the local Lyapunov exponents are averaged. Each point on the trajectory with

its neighbor can be given as;

$$x_1(i), y_1(i) \rightarrow x_1(i+1), y_1(i+1) \quad (4.19)$$

$$x_2(i), y_2(i) \rightarrow x_2(i+1), y_2(i+1)$$

And local Lyapunov exponent defined as;

$$LE = \frac{\sqrt{(x_{2t}(i+1) - x_1(i+1))^2 + (y_{2t}(i+1) - y_1(i+1))^2}}{\sqrt{(x_2(i) - x_1(i))^2 + (y_2(i) - y_1(i))^2}} \quad (4.20)$$

Additionally, because employing $\{x_{2t}(i+1), y_{2t}(i+1)\}$ as the starting point for subsequent iteration causes that $\{x_2(i), y_2(i)\}$ trajectory separates more in each iteration, shrinking $\{x_{2t}(i+1) - x_1(i+1), y_{2t}(i+1) - y_1(i+1)\}$ by $d(i)/d_t(i)$ is necessary to retain the starting point of the subsequent iteration as close to the trajectory as it is on the previous one. As shown in Fig. 4.8, the outcome of negative Lyapunov exponent values agreeing with

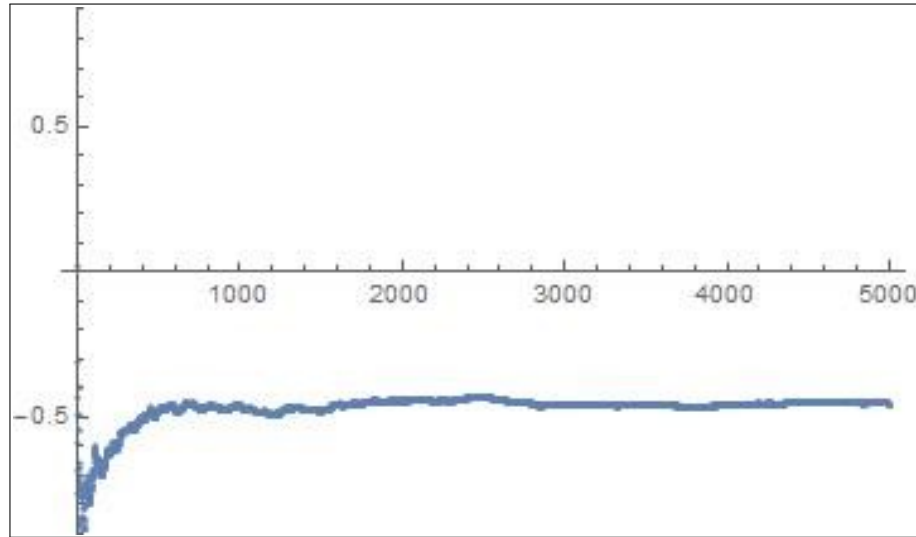


Figure 4.8: The largest Lyapunov exponent result by directly with the 2-dimensional map for FIGARCH $d=0.80$

the result of Kantz's algorithm supports that FIGARCH model is not deterministic chaotic. Also, all the results obtained by three different calculation method for LLE is summarized

in Table 4.4.

Table 4.4: LLE results for all the FIGARCH models with the three different method

FIGARCH (d) value	Wolf's algorithm	Kantz's algorithm	2-dimensional map
0.05	9.1	-3.38E-05	-1.05824
0.15	2.27	-6.40E-05	-0.37559
0.25	11.85	-7.47E-05	-0.19538
0.35	2.38	-7.93E-05	-0.15997
0.45	7.6	-8.19E-05	-0.18594
0.55	24.5	-3.50E-06	-0.24945
0.65	22.17	-1.14E-04	-0.32664
0.75	20.8	-1.19E-04	-0.40903
0.80	21.1	-6.19E-05	-0.45553
0.90	14.4	-1.24E-04	-0.61687

4.1.8. Final Remarks

The chaotic properties of FIGARCH stochastic difference equations has been investigated extensively by employing various well-established tools as well as comparing and verifying their results with the alternatives. Firstly, embedding delay and embedding dimension values are determined on numerical simulations of numerous FIGARCH models in order to continue with the computations of the correlation dimension and Lyapunov exponents. While, self-similar properties of FIGARCH has been identified by correlation dimension, positive LLE values found by Wolf algorithm implied the possible presence of the chaotic structure.

The latter result could have been the issue of the embedding method (Dechert and Gençay, 2000). In other words, the LLE may not be sustained under Taken's embedding theorem. Additionally, the Wolf algorithm fails to take advantage of all the available data for the

reason that it focuses on one fiducial trajectory and omits information (Rosenstein et al., 1993).

So, in addition, Kantz's algorithm and Jacobian derivative based on the stochastic difference equations were used to verify the result of Wolf's algorithm. This time negative LLE values appeared as a result of both algorithm. The latter implies the existence of some kind of nonlinear determinism in the data that originates from FIGARCH but rules out the possibility of deterministic chaos. In conclusion, when the data is modelled with FIGARCH stochastic difference equations, the irregularity and self-similarity shouldn't be associated to deterministic chaos.

This part of the research is also published (Yilmaz and Unal, 2016a).

4.2. STOCHASTIC DUFFING EQUATION IN MODELLING FINANCIAL TIME SERIES

The chaotic noisy Duffing map possesses unique multifractal characteristics contrary to various other chaotic maps like Henon map, Chirikov map and so on. There isn't any general method for modelling or forecasting multifractal systems. In order to do so, a solution processes for the noisy Duffing map can be decomposed into telegraph process and fractional Gaussian process (fGn).

It is a huge challenge to model and forecast chaotic time series due to its sensitivity to initial conditions. A tiniest error in calculating the current state is enlarged quickly, by the effect which although the action is predictable in the short term, it is unpredictable over its forecast horizon (Peitgen et al., 2006). Nonetheless, the uncertainty is enlarged at a limited rate, even for a chaotic system, so there is yet hope to achieve an acceptable short-term forecasts (Kantz and Schreiber, 2004).

In this section, a new modelling and forecasting technique for the chaotic noisy Duffing map and the time series possessing identical characteristics is introduced and demonstrated. As a starting point, the chaotic and fractal characteristics of the chaotic noisy Duffing map is

investigated. Next, this new method is exercised. Subsequently, the characteristics of the noisy Duffing map are researched in financial time series and a suitable sample is tested with the same method to experiment its applicability.

Two fundamental sort of dynamical systems exist which are continuous differential equations and discrete iterated maps or difference equations (Strogatz, 2018). Two dimensional chaotic discrete iterated Duffing map is used for this research. Due to their exceptional traits like sensitivity to alterations in initial conditions and randomness, numerous chaotic maps have been used in cryptography (Tang and Liu, 2011), networks (Budroni et al., 2017), image encryption (Wong, 2009).

Particularly, the Duffing map comes forward with its multifractal framework and its regime switching features in the class of accepted chaotic maps interpreting the deterministic behavior of periodically forced, nonlinear oscillator systems. For this analysis, the center is the nonlinear chaotic Duffing map, even though numerous other two dimensional chaotic maps are relevant to the identical methods introduced here. To inspect the Duffing map in the influence of extrinsic additive noise like real-world signals, some noise is added to it. Noisy deterministic nonlinear Duffing equation has been examined by physicists (Perkins, 2017), mathematicians (Mallick and Marcq, 2004) and engineers (Kapitaniak, 2012).

A specific suitable method for describing signals with fractal behavior is through the use of wavelets (Daubechies, 1992). For the analysis of signals, they consist of statistical copies of themselves like signals that indicate fractal features, because of the self-affine basis set of wavelets (Lowen and Teich, 2005). Therefore, in this research, wavelet analysis is used to decompose the chaotic Duffing map signal, scale by scale in the time domain. Wavelet decomposition is processed by basis function that are obtained from a mother wavelet through stretching and shifting. Especially, the wavelet power spectrum analysis characterize the signal energy on a time-scale domain. By this scale dependent energy distribution, powerful and weak scales can be distinguished. As the main focus of the first part of this thesis, wavelet analysis was introduced and researched in great detail. These methods are also become very useful for the analysis of the chaotic noisy Duffing map.

Additionally, multifractal de-trended fluctuation analysis (MFDFA) is used for further examination on multifractal features of separate scales that are accordingly isolated by wavelet transform, inspected by wavelet power spectrum and after that reconstructed by inverse wavelet transform. MFDFA is used for the measurement of fractal scaling characteristics and the revelation of the long range correlations in noisy non-stationary time series. It has been utilizing in variety of areas like hydrology (Rybski et al., 2011), traffic time series (Shang et al., 2008), geoelectrical signals (Telesca et al., 2005), heart rate dynamics (Dutta, 2010) and financial time series (Thompson and Wilson, 2016).

As a result of the investigation of the features by employing wavelet analysis and MFDFA analysis over various main chaotic maps, the Duffing map reveals interesting multifractal features which subsequently is demonstrated to be superimposed by telegraph process and fGn. In a unique way, the telegraph process is used to model for regime switching patterns firstly for the Duffing map data series accompanying white noise process and subsequently for the financial time series having similar features. A random motion is modeled by a steady velocity in the telegraph process. Solely, it takes either one negative value or one positive value on the real line at each location. So far, the telegraph process has been attempted to model the dynamics of the volatility of financial markets (Kolesnik and Ratanov, 2013; Ratanov, 2007), the dynamics of the price of risky assets (Di Crescenzo and Pellerey, 2002), animal movement in ecology (Holmes et al., 1994) and the probabilistic features of it have been examined (Foong and Kanno, 1994; Stadje and Zacks, 2004).

The Markov switching model (Hamilton, 1989), alias the regime switching model is very good at catching swift shifts of a course and the events which the recent dynamics of prices and fundamentals continue for number of periods following a variation. The complex dynamical patterns can be caught by permitting switching between multiple structures in various regimes. The regime switching models have particularly been used for the examination of financial and economic time series such as business cycles, interest rates, foreign exchange excess returns, equity excess returns and so on (Fallahi, 2011; Janczura and Weron, 2012; Zhang and Zhang, 2015). In this research, telegraph process is utilized to make predictions on the regime switching model to find out where the subsequent regime may emerge like the noisy Duffing map. It is then demonstrated on real-world time series.

Due to its capability over Markovian process on data series having long memory features, Fractional Gaussian noise is used as a modelling instrument for the volatile section with the chaotic and self-similar structure withdrawn from the regime switching section. Long-range dependency systems are modeled with stochastic processes which possess probability distributions that are constant to the changes of scale in time or space (Sørbye and Rue, 2018). Long-memory processes like fGn consist of a competent instrument to yield long-term forecasts in comparison to short-memory processes (Ferrara and Guegan, 2000).

Hurst et al. (Hurst, 1951; Mandelbrot and Van Ness, 1968) and Mandelbrot and Wallis (Mandelbrot and Wallis, 1969) introduced fractional (fBm) and fractional Gaussian noise (fGn). Hurst phenomenon which is an unique behavior observed firstly on hydrological and other geophysical time series was discovered by Hurst in 1951. After that in 1965, Mandelbrot presented the fGn process that is the discrete-time increment of fBm, the only self-similar continuous-time Gaussian process with stationary increments and has autocovariance function described by the self-similarity parameter $H \in (0, 1)$. H denotes the Hurst exponent that quantifies the Hurst phenomenon (Sørbye and Rue, 2018).

In this research, the characteristics of the Duffing map with additive white noise is investigated by wavelet analysis and MFDFA analysis. A new model is proposed for modelling and the forecast of the chaotic Duffing map with white noise and the financial time series with the identical characteristics of regime switching with long-range dependency structures. The efficiency of the developed method and the results of the demonstrations are given at the final section in detail. Moreover, this method can be applied to other variety of chaotic maps similarly for modelling and forecasting.

The research is continued with the introduction of the main focus that is the chaotic Duffing series with noise and the application of wavelet analysis and MFDFA analysis for the inspection of the associated characteristics. After that the combination of telegraph process and fGn is developed as a new modelling and forecasting tool for the noisy Duffing series and its possible application is set up as a model for the series with regime switching characteristics. Next, it is demonstrated on real world financial data series with identical characteristics with the application of the analyses introduced. The comparison of all the

results with the noisy Duffing series are given for each step side by side. Subsequently, for the number of time periods, variety of forecasts are made for both time series and effectiveness of the method is shown.

4.2.1. The Noisy Duffing Map

There are number of chaotic maps that are generally utilized in the chaos literature. Chaotic data is very challenging for estimation and forecast due to its sensitivity to initial conditions. In order to find a way to make efficient estimations and forecasts, these chaotic maps are analyzed with regard to their multifractal characteristics. Specific chaotic maps of Chirikov standard map (alias Chirikov-Taylor map), Ikeda map, discrete Duffing map and Henon map are selected considering the comparison effectiveness of the aimed estimation. As a result of multifractal analysis Duffing map emerges because of its distinctive characteristics identical to regime switching models which is also encountered in real world financial time series.

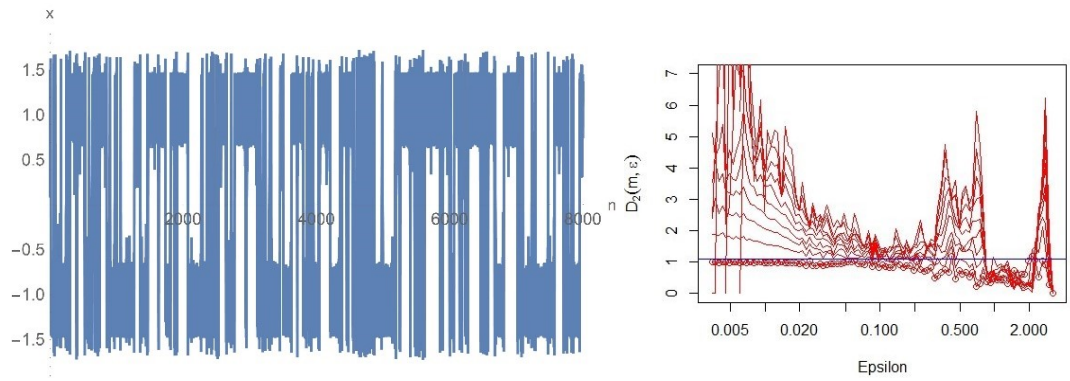
Considering there is no real world data without any noise, to adjust original Duffing map (Wiggins, 2003) accordingly, white noise is added to it. Therefore, the modified Duffing map with white noise is described as;

$$x_{n+1} = y_n \tag{4.21}$$

$$y_{n+1} = -bx_n + ay_n - y_n^3 + \varepsilon_n$$

where, ε denotes the Gaussian white noise with zero mean and standard deviation 1. To get chaotic characteristics, the values of 2.75 and 0.2 are assigned to a and b consecutively.

The noisy Duffing series is generated for 8000 number of points. The value of the standard deviation is fixed at 0.01 due to the swift disturbance effect that causes the points of the series to hit too large numbers. This is expected effect since the most unique characteristic of the chaotic map is the sensitivity to initial conditions. The final form of the Duffing map series produced by these equations is given in Fig. 4.9.a that displays a unique characteristic resembling a noisy pendulum. Besides, the correlation dimension measurement of the



a. The noisy Duffing map series

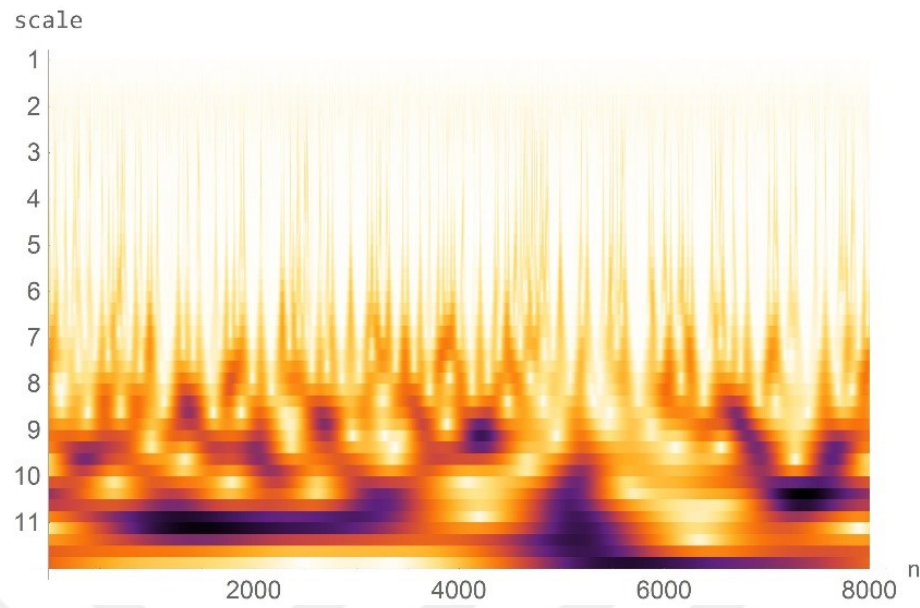
b. CD plot of the noisy Duffing map series

Figure 4.9: The noisy Duffing map

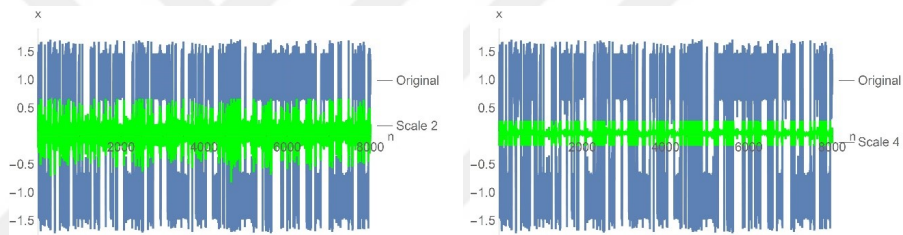
Duffing map made by Tisean software package (Hegger et al., 1999) is shown in Fig. 4.9.b in which the chaotic features appear to be less clear than the original Duffing map data series. As the horizontal line sloping the little plateau in the middle section. Apparently, the degree of the chaotic characteristics is weakened to some extent by the added white noise. The spikes through the line and disturbance on the plateau that is the indicator of chaos hint that the noisy Duffing series may display increased multifractal structure.

4.2.1.1. Wavelet Decomposition and Wavelet Scalogram

A multi-resolution illustration of a process can be obtained by wavelet methods which contains set of orthogonal basis functions. The same result can't be achieved by conventional methods like Fourier transform. The advantage of wavelet methods come from the flexibility of capturing the resolution in time and frequency space. Wavelet scalogram is acquired by using wavelet decomposition techniques with the continuous wavelet transform in which Gabor wavelet is employed. The outcome is displayed in Fig. 4.10.a with the scales ranging from 1 to 11 through the number of steps of 8000 with the energy levels are shown. Subsequently, four decomposed scales are withdrawn by inverse wavelet transform and comparisons of them with the noisy Duffing series are presented in Fig.4.10 which indicates that the scale 2 possesses the most energy.

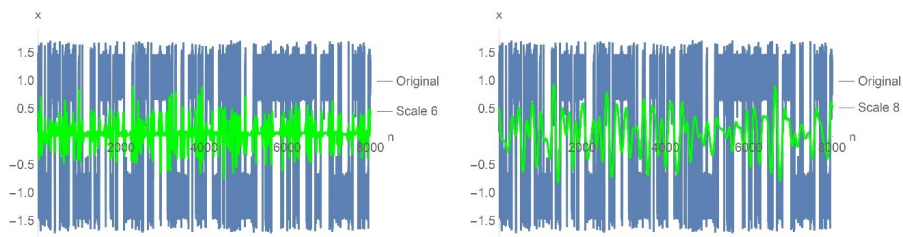


a. The wavelet scalogram result of the noisy Duffing series



b. Scale 2 decomposed vs. the noisy Duffing series

c. Scale 4 decomposed vs. the noisy Duffing series



d. Scale 6 decomposed vs. the noisy Duffing series

e. Scale 8 decomposed vs. the noisy Duffing series

Figure 4.10: The noisy Duffing map wavelet scalogram and decomposed scales

4.2.1.2. MFDFA Analysis

Multifractal detrended fluctuation analysis (MFDFA) is quite useful in identification of fractal scaling characteristics and the disclosing of long range correlations in nonstationary time series. Therefore, MFDFA is used for further examination of decomposed scales. Particularly, multifractal spectrum provides very helpful information for the determination of the deviations in fractal structure.

The multifractal toolbox software is used for MFDFA (Ihlen, 2012) that estimates multifractal spectrum of power law exponents to determine whether or not the decomposed scales are monofractal or multifractal. If the classification is monofractal, it is considered to be independent in time and space. And if it is multifractal, it has spatial and temporal variation in scale invariant structure. MFDFA results of the chaotic maps are displayed

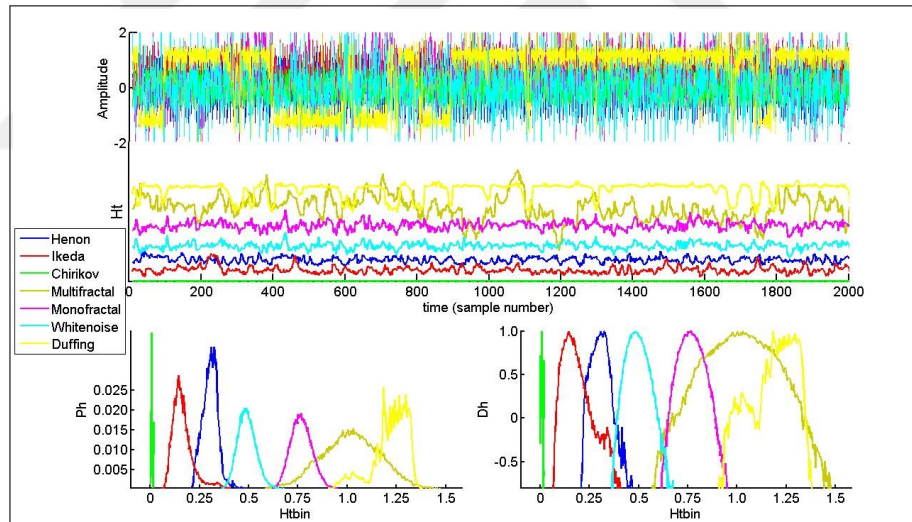


Figure 4.11: MFDFA results of chaotic maps. (Henon = 0.3077, ikeda = 0.1436, chirikov = 0.0067, multifractal = 1.0304, monofractal = 0.7642, white noise = 0.4787)

in Fig. 4.11 with white noise, monofractal and multifractal series for comparison. At the top, local Hurst exponents (H_t) are shown for each series. Large fluctuations indicated multifractal characteristics like the original Duffing series represented in yellow color. The histogram of these fluctuations (Ph), the probability distribution of H_t is given at the bottom left corner. The power law of exponent of Hurst value is obtained by the calculation of

the mean value of local Hurst exponents. The multifractal spectrum (D_h) that is acquired by log transformation of the normalized probability distribution is also located at the right bottom corner therein width of the arc indicated the multifractal structure. So, as seen, Duffing series is identified as multifractal since having wider arc than the white noise and monofractal series.

With the information gathered by the examination of the multifractal structure of the original Duffing series, next, the multifractal structures of the decomposed scales of the noisy Duffing series is investigated and the results in Fig. 4.12 is obtained. According to MFDFA results, scale number six and scale number eight can be described as multifractal. The subplot of local Hurst exponents display that they are recognized with showing larger spikes than the other two scales. What's more, subplots of the probability distributions of local Hurst exponents and the multifractal spectrum illustrates that these two scales have much wider arcs.

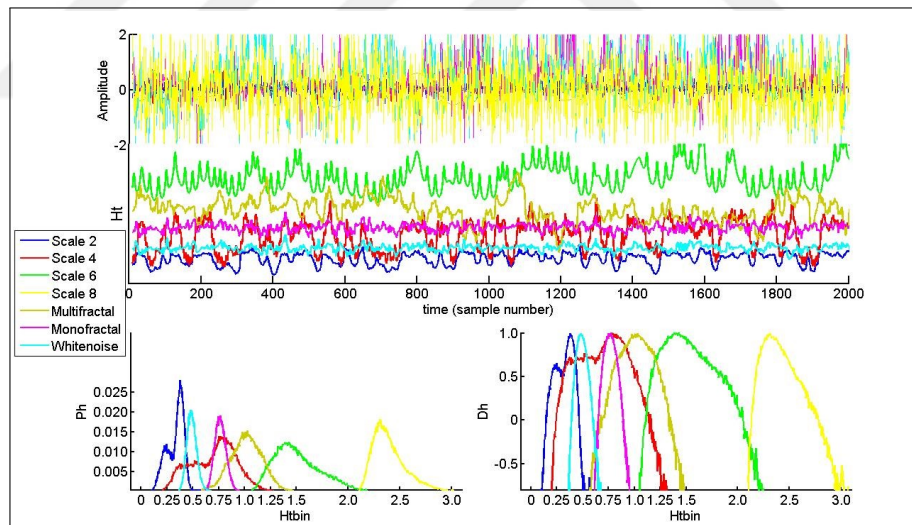


Figure 4.12: The noisy Duffing series with standard deviation of 0.01 and the 8000 number of data points. (Scale 2 = 0.3783, scale 4 = 0.7531, scale 6 = 1.4271, scale 8 = 2.3105, multifractal = 1.0304, monofractal = 0.7642, white noise = 0.4787)

4.2.1.3. *The Noisy Duffing Map and Its Regime Switching Properties*

As MFDFA displays, particularly, the noisy Duffing map carry interesting multifractal quality among the chaotic maps analysed. Additionally, it has some type of regular looking oscillating movement in other words, regime switching characteristics. Variations occur in periods rather than in short instances. A state carries on for a while till a shift takes place and again reversion comes next. This phenomenon particularly has often been detected and studied in financial markets like equity market, income market, foreign exchange market and so on.

Especially, one particular process, namely telegraph process is considered to be a good fit for a regime switching behavior in this research. Actually, telegraph process was introduced for the purpose of treating the turbulent diffusion (Taylor, 1922). It characterized the location of a particle that only can travel in positive direction or negative direction on the real line oscillating with constant velocity v . So, telegraph process only gets the value of -1 or 1 at time t . The number of moves in the period matches the Poisson distribution.

$$X(t) = x_0 + V(0) \int_0^t (-1)^{N(s)} ds \quad (4.22)$$

where $t > 0$ and $V(0)$ is the initial velocity. The fundamental principle of telegraph's equation is the probability of retaining the direction and the probability of reversal of the direction at a certain time period (Kac, 1974).

The noisy Duffing equation evidently displays well-structured regime switching characteristics as previously shown in Fig. 4.9.a. The noisy Duffing series oscillate between positive and negative values at various durations along the axis. Therefore, it can be suggested that the telegraph process and regime switching model possess identical characteristics. The phenomenon of swift variations between positive and negative values and retaining the state corresponds to the regime shift and continuity of the regime switching systems.

For the estimation of regime switching systems, we propose telegraph process. As identified,

the noisy Duffing series possesses ideal setting with its regime switching properties together with multifractal properties to validate both telegraph process estimation with another method searched for more effective prediction of multifractal systems that is greatly challenging.

4.2.2. Modelling with Telegraph Process and Fractional Gaussian Noise

A new model we propose can be outlined as follows.

Modelling outline

- Forming telegraph process representation of data series under investigation
- Telegraph process estimation $[\lambda]$
- Detachment of the noisy section from the telegraph process representation
- fGn estimation $[\mu, \sigma, h]$ of the noisy part

Forecast outline

- Random Generation of estimated telegraph processes
- Calculation of the mean path of random telegraph processes
- Generation of the estimated fGn $[\mu, \sigma, h]$
- Merging of the mean path of the telegraph process with the estimated fGn.

These steps are explained in a great detail in subsequent sections with the application on the noisy Duffing series as well as real world financial data series.

4.2.2.1. The Noisy Duffing Series Estimation by Telegraph Process

Because of the limited availability of the real world data series and for the purpose of making comparisons of the model with it, the length of the noisy Duffing data series set to 600. The computations are executed by Mathematica, Python and R software tools.

As mentioned in its definition before, telegraph process oscillates between the definite values of 1 and -1 . It is regulated by a homogenous Poisson process with the rate of λ . So, the noisy Duffing series path is required a straightforward adjustment to be reconstructed into a new frame that holds the values of either 1 or -1 . According to the Heaviside step function in Eq. 4.23, by replacing every positive value with the value of 1 and every negative value with the value of -1 , intended telegraph process representation of the series are produced.

$$\Theta(y) = \begin{cases} y = 1 & y \geq 0 \\ y = -1 & y < 0 \end{cases} \quad (4.23)$$

where y denotes the value of each data. Fig. 4.13 shows the original noisy Duffing series

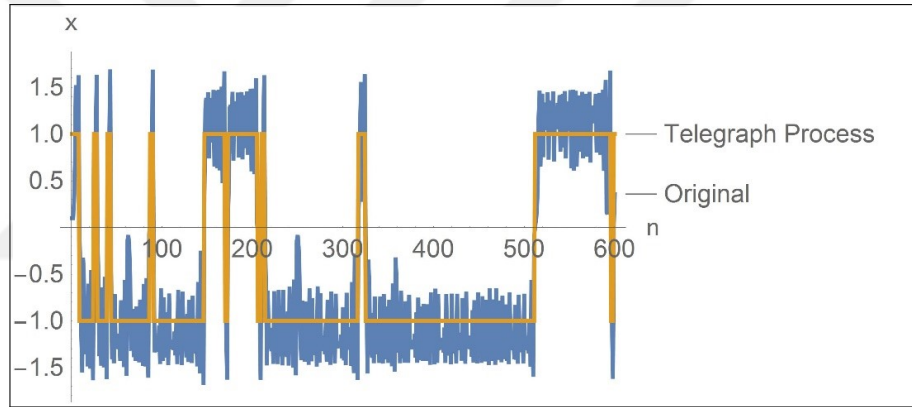
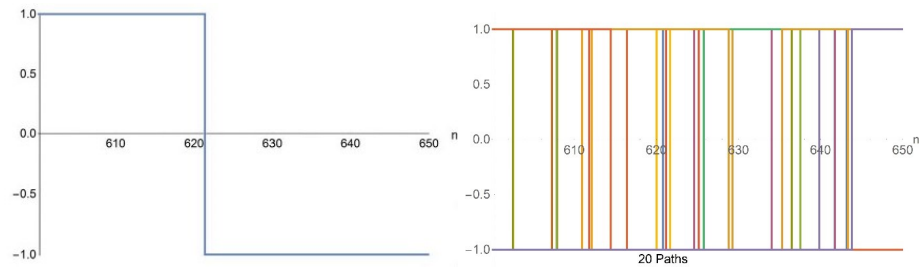


Figure 4.13: The noisy Duffing series versus the modified series as to telegraph process

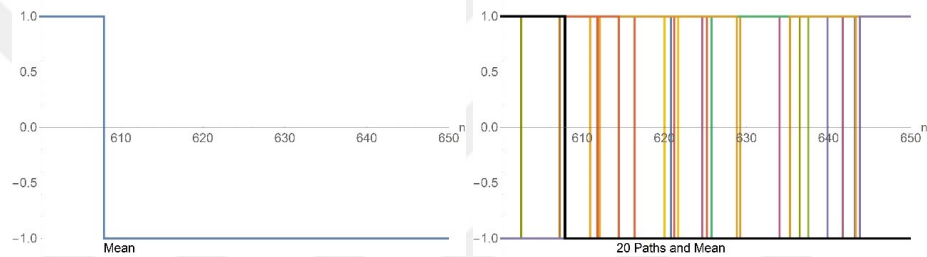
with the modified version as to telegraph process together for neat comparison. It is clearly seen that only the persistent intervals and the instances where the shifts in path takes place are extracted from the original Duffing series to achieve the modification for telegraph process estimation. The computation of the estimation yields the λ value that is the only parameter of telegraph process as 0.02881 with which the prediction can be conducted.

For the prediction of following time steps, random processes of the estimated telegraph process is generated and possible persistent time periods and shifts in direction is obtained. One of these random processes for the 50-forward step is given in Fig. 4.14.a. In addition, all the paths of randomly generated telegraph process for the same duration is presented in



a. Single path of randomly generated telegraph process for next 50 steps. b. 20 number of paths of randomly generated telegraph process for next 50 steps

Figure 4.14: Telegraph process prediction for the noisy Duffing series



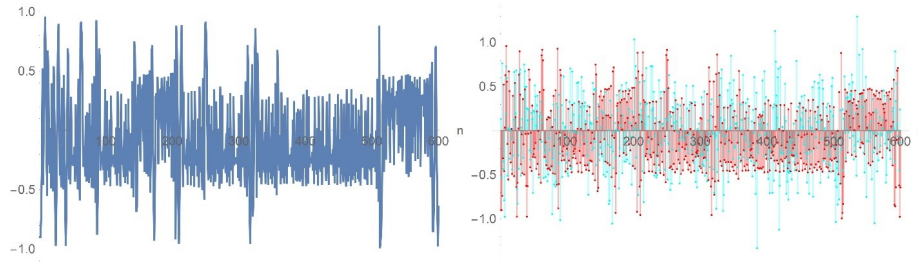
a. The mean of 20 number of paths of estimated telegraph process. b. The mean (in thick black color) and estimated paths together

Figure 4.15: The average value of all estimated telegraph process paths for the noisy Duffing series.

Fig. 4.14.b. The average of these randomly generated processes provides the most likely path which is shown in Fig. 4.15.

4.2.2.2. Fractional Gaussian Noise Process Estimation

With the telegraph process estimation method, how the noisy Duffing series or in more general, any regime switching series possibly shift in terms of the duration of the persistent intervals and where shift occurs become predictable. However, erratic changes in short time intervals around persistent intervals still stay unpredictable. So, as having a different structure from the regime switching series, these noise-like parts need to be investigated independently in order to be able to finalize the estimation and prediction of the noisy



a. The original data versus the noise left b. Original noisy series (in red color) after deduction of telegraph process of the versus estimated fGn series (in light blue). noisy Duffing data.

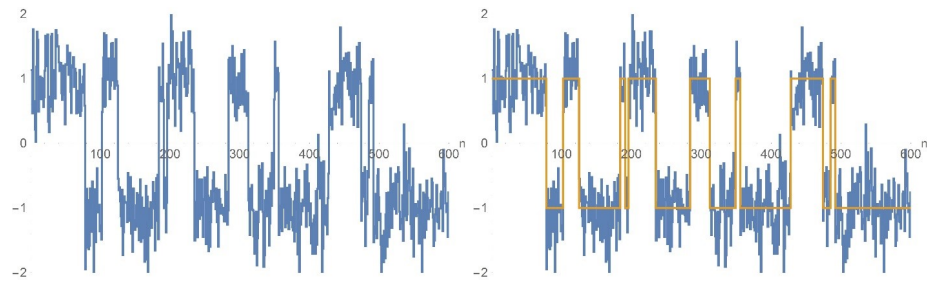
Figure 4.16: The noisy section of the noisy Duffing series and fGn estimation.

Duffing series. For this purpose, noise-like section can be detached from the original Duffing series by subtracting estimated telegraph process from the original Duffing series as shown in Fig 4.16. This section primarily comprises the deterministic chaotic character. So, the most efficient process that would describe the series is considered to be fGn. Because it is a stochastic fractal model which carries strong statistical characteristics, it appears to be yield relevant models for a broad range of empirical observations in variety of fields.

As referred in introduction, fBm is the only self-similar continuous-time Gaussian process with stationary increment and fGn is its discrete-time increment. They simulate the statistical characteristics of self-similarity and long-range dependence. fGn produces a model characterized by Hurst exponent $H \in (0, 1)$, a parameter to measure the Hurst phenomenon (Hurst, 1951) and variance σ^2 . As $H < 0.5$ indicates anti-persistent or negatively auto-correlated behavior, $H > 0.5$ suggests $1/f$, persistent or long memory behavior and $H = 0.5$ implies the conventional Gaussian white noise. fGn has the mean μ and the covariance function of;

$$\gamma(t, s) = \frac{\sigma^2 |t - s - 1|^{2H} - 2 |t - s|^{2H} + |t - s + 1|^{2H}}{2} \quad (4.24)$$

The computation of fGn estimation of the noisy section yields the value of 0.39712 for the standard deviation and 0.56075 for Hurst index as the mean value is fixed to zero. With these parameters, a fGn process can be produced and fitted to the noise-like series as in



a. The random estimated telegraph process b. The random estimated telegraph process series combined with estimated fGn series. series combined with the estimated telegraph process shown separately.

Figure 4.17: The estimated series as a combination of telegraph process and fGn for the noisy Duffing series.

Fig. 4.16.b.

Fig. 4.17 shows the estimated series that is the combination of the random series produced according to the estimated telegraph process over the same period and the fGn process fitted. This new series possess the same statistical characteristics with the original series and clearly the identical probability distributions. The probability distribution of the noisy Duffing series is estimated and found as the combination of a number of normal distributions. The top two best estimations are shown in Fig. 4.18.a in a bar style graph with the probability density functions (pdf) in line style over histogram results.

In addition, the probability distribution of the series estimated with telegraph process and fGn process is computed and the result is the combination of the normal distribution and logistic distribution with more weight on the normal distribution. Fig. 4.18.b illustrates the best two probability density functions of the estimations in line style and the probability distributions as histogram for comparison.

The outcome of the investigation of the original Duffing series modified with the white noise seems to verify the efficiency of the method proposed. Being able to regenerate the noisy Duffing series by the estimation of telegraph process and fGn indicates the capability of the prediction of the series similarly with a tolerable error. So, the prediction can basically

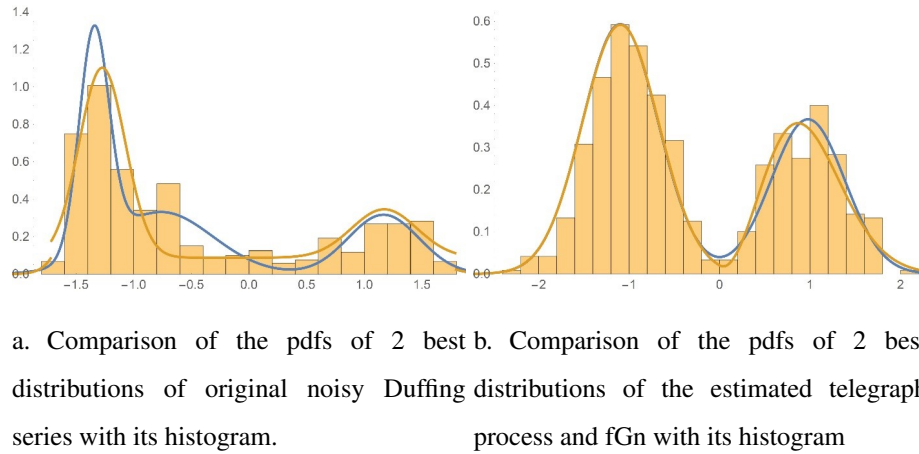
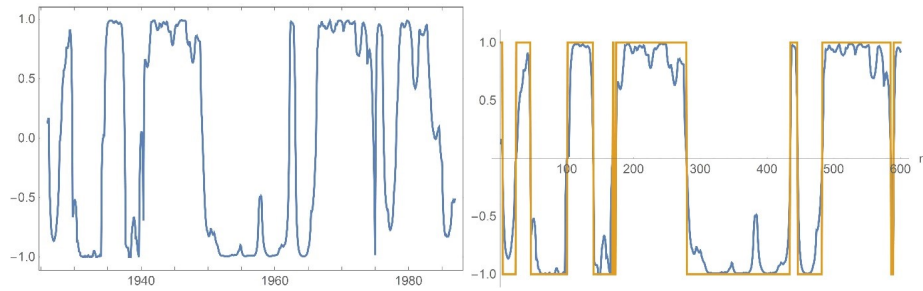


Figure 4.18: The probability distributions of the original Duffing series and the series estimated with telegraph process and fGn.

be made by applying the same method for subsequent steps right away. All these results support that telegraph process and fGn estimation can be applied to describe, estimate and predict even the complex deterministic chaotic systems like the noisy Duffing series. So, it can easily and competently be used for real-world time series with the similar noisy, chaotic and regime switching characteristics in practice. For the purpose of validating its possible practical usages, in subsequent section, it is demonstrated for the estimation and prediction of a real-world time series.

4.2.3. The Demonstration on the Real-World Time Series

The Markov switching model, alias the regime switching model is a popular model used for analyzing the economic and financial time series in different regimes. It can capture dynamic patterns by shifting between multiple structures. It is formed by joining two or more dynamic models. Plainly, Markovian state variable generates random and periodic variations of model structures and the persistence of each regime is determined by its transition probabilities which is the our focus for the utilization of the method proposed. For this purpose, as a real-world time series, the monthly stock returns are used on which Kim et al. (Kim and Nelson, 1998) exhibits good fitting of Markov switching variance model. The smoothed probabilities of a medium variance regime computed for monthly excess stock returns comprises our data set which ranges from January 1926 to December



a. Smoothed probabilities of the medium variance regime for excess stock returns. b. Smoothed probabilities of excess stock returns with telegraph process representation.

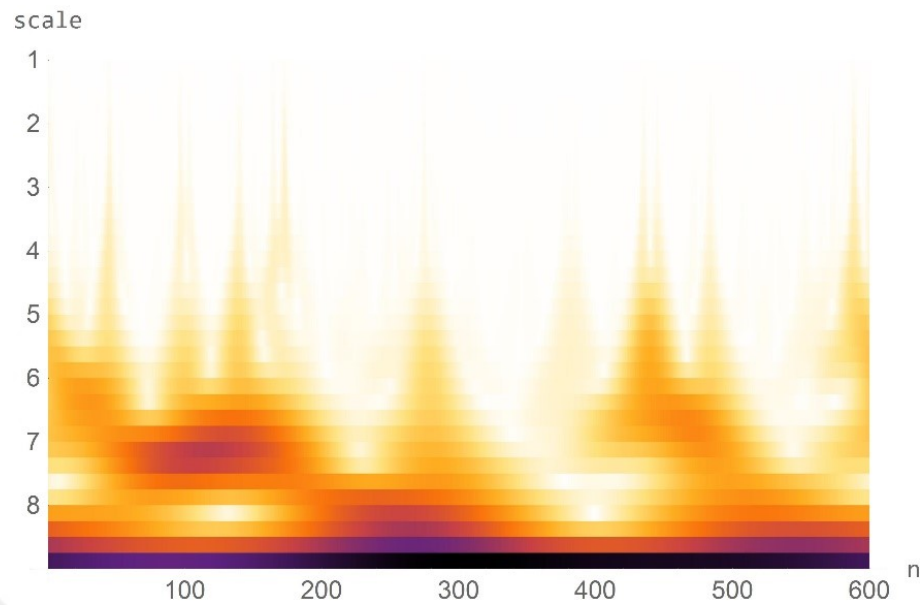
Figure 4.19: Smoothed probabilities of the medium variance regime for excess stock returns with its telegraph process representation.

1975 with the total number of 600 observations and is illustrated on Fig. 4.19.a.

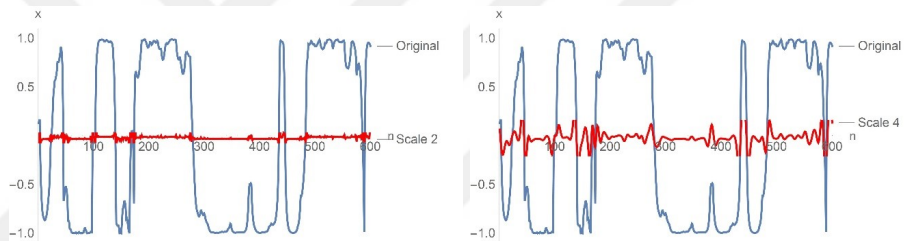
4.2.3.1. *Wavelet Scalogram and Multifractality of the Smoothed Probabilities for Excess Returns*

Fig. 4.20.a displays the wavelet scalogram of the smoothed probabilities for excess returns with the energy levels for the scales from 1 to 8 through all the steps up till 600. Subsequently, by inverse wavelet transform method, the scales from 2 to 8 are withdrawn and for comparisons, they are shown in Fig. 4.20 together with the original series. As seen, the scale 8 has the largest energy.

As presented in Fig. 4.21, the relatively large fluctuations in local Hurst exponent plot at the top and wide arcs in the probability of distribution H_t on the left and multifractal spectrum on the right signifies the multifractal structure for all the scales. Scale 2 appears to be multifractal in contrast to the monofractal determination of scale 2 of the noisy Duffing series in previous section in which chaotic dynamics are retained.

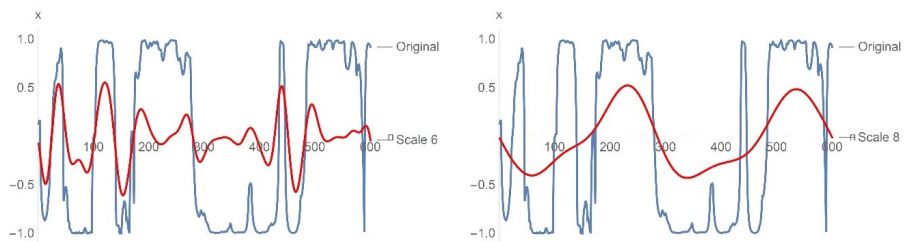


a. The wavelet scalogram result of the smoothed probabilities of excess returns



b. Scale 2 decomposed vs. the smoothed probabilities of excess returns

c. Scale 4 decomposed vs. the smoothed probabilities of excess returns



d. Scale 6 decomposed vs. the smoothed probabilities of excess returns

e. Scale 8 decomposed vs. the smoothed probabilities of excess returns

Figure 4.20: Wavelet scalogram and decomposed scales of the smoothed probabilities of excess returns.

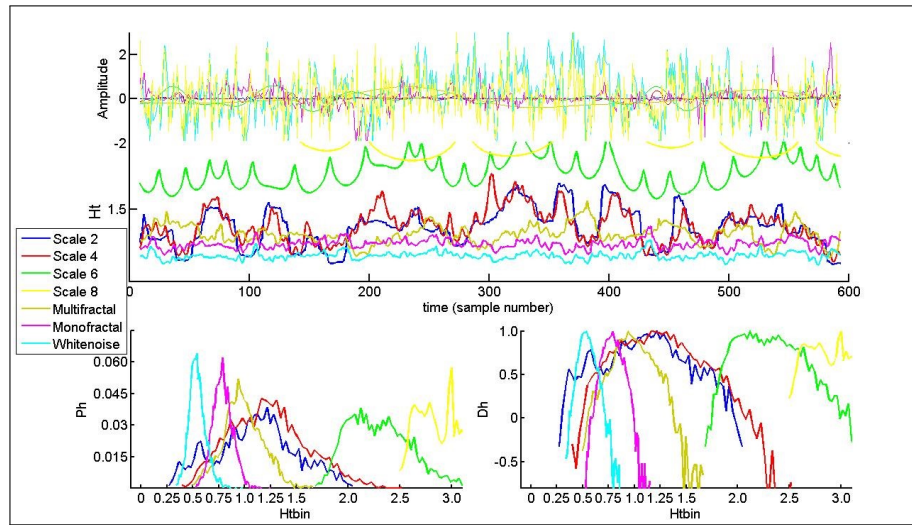


Figure 4.21: Multifractality results of smoothed probabilities for stock returns. (scale 2 = 1.2170, scale 4 = 1.1635, scale 6 = 2.1206, scale 8 = 3.0002, multifractal = 1.0304, monofractal = 0.7642, white noise = 0.4787)

4.2.3.2. Telegraph Process and fGn Estimation

Telegraph process representation of the smoothed probabilities of excess returns are given in Fig. 4.19.b with the original data series. Telegraph process estimation yields the lambda value as 0.02218 as to which random telegraph processes are generated and the average path is computed as shown in Fig. 4.22 and in Fig. 4.23. For the second part of the estimation, volatile section drawn out from the original series after telegraph process section is estimated with fGn process and the mean value of 0, the standard deviation value of 0.3 and Hurst

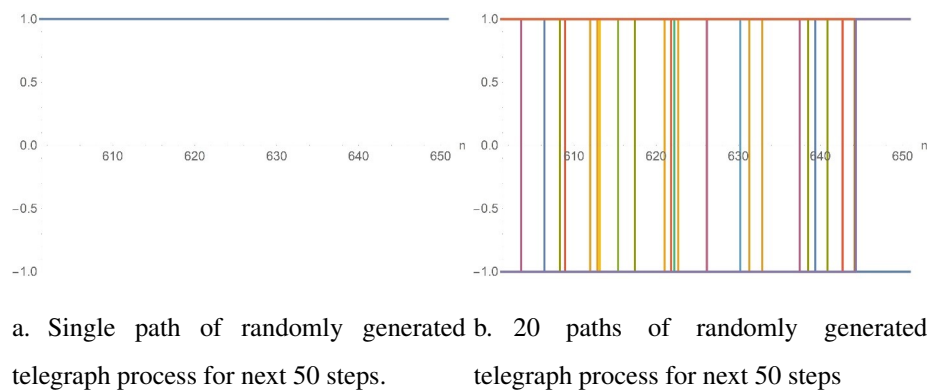


Figure 4.22: Telegraph process prediction for the smoothed probabilities of excess returns.

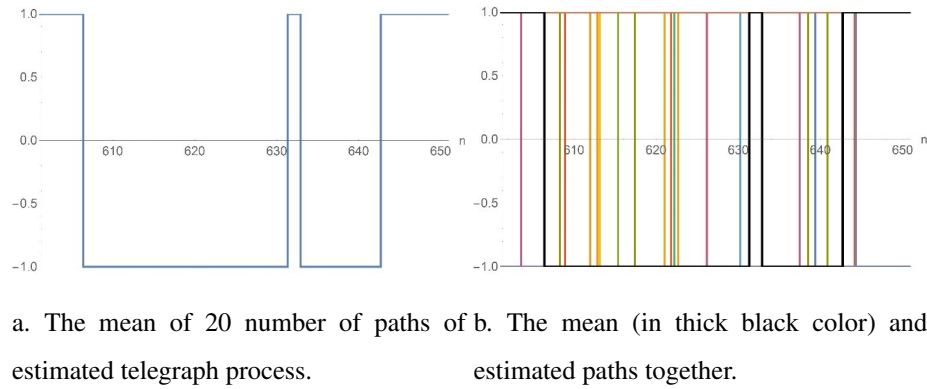


Figure 4.23: The average value of all estimated telegraph process paths for the smoothed probabilities of excess returns.

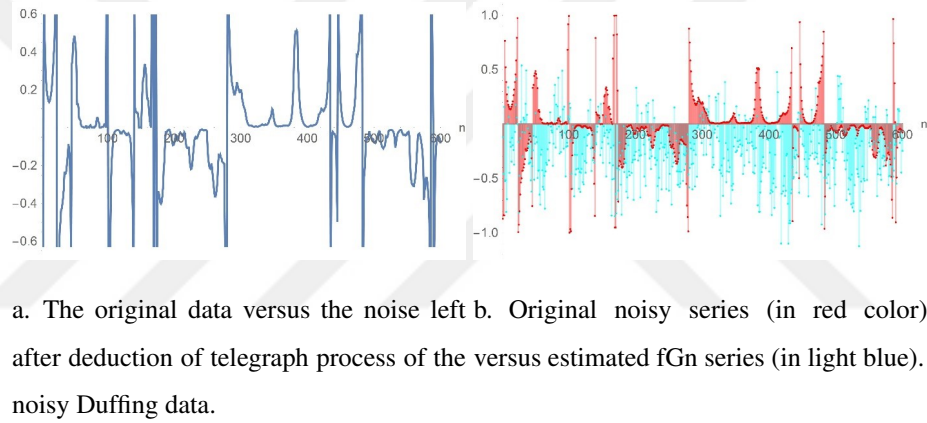
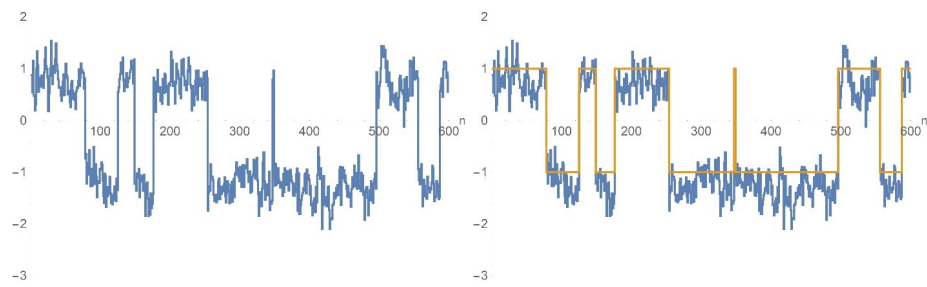


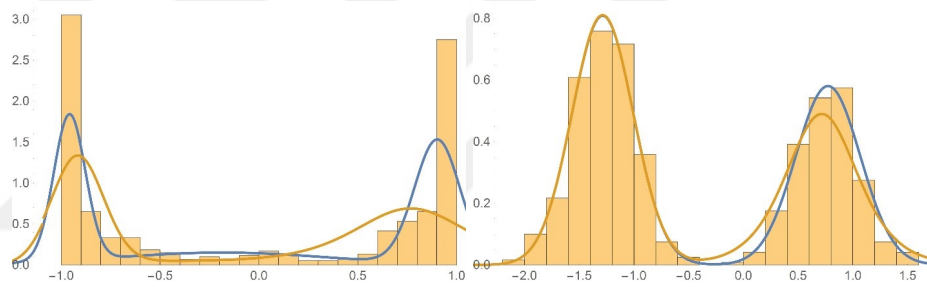
Figure 4.24: The noisy section of the smoothed probabilities of excess returns and fGn estimation.

value of 0.8015 are obtained. These values generated fGn series shown in Fig. 4.24. Both telegraph process estimation and fGn estimation together provide complete solution as in Fig. 4.25. Eventually, as presented in Fig 4.26 the inspection of the probability distributions show that they are the mixture of a number of normal distributions as observed in previous section with the noisy Duffing series. Fig. 4.27 shows that two distributions after estimation with the combination of telegraph process and fGn possess identical distributions. It reveals that how close the estimation of the noisy Duffing series and the estimation of regime switching smoothed probabilities for excess returns. It may suggest that the chaotic noisy Duffing series can be a appropriate model for the real-world series with regime switching characteristics.



a. The random estimated telegraph process b. The random estimated telegraph process series combined with estimated fGn series. series shown separately in yellow color combined with estimated fGn series.

Figure 4.25: The estimated series as a combination of telegraph process and fGn for the smoothed probabilities of excess returns.



a. Comparison of the pdfs of 2 best b. Comparison of the pdfs of 2 best distributions of original smoothed distributions of the estimated telegraph probabilities of excess returns with its process and fGn with its histogram histogram.

Figure 4.26: The probability distributions of the original smoothed probabilities of excess returns and the series estimated with telegraph process and fGn.

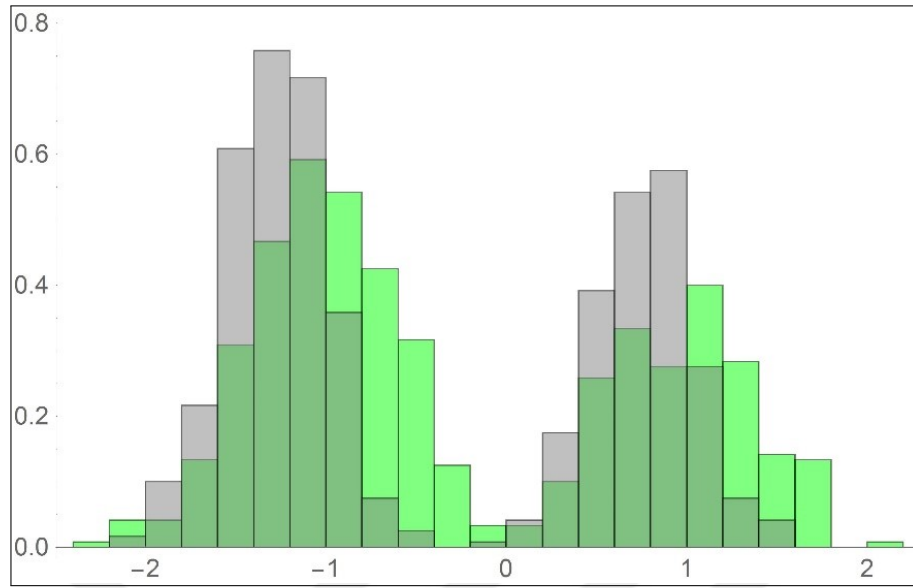
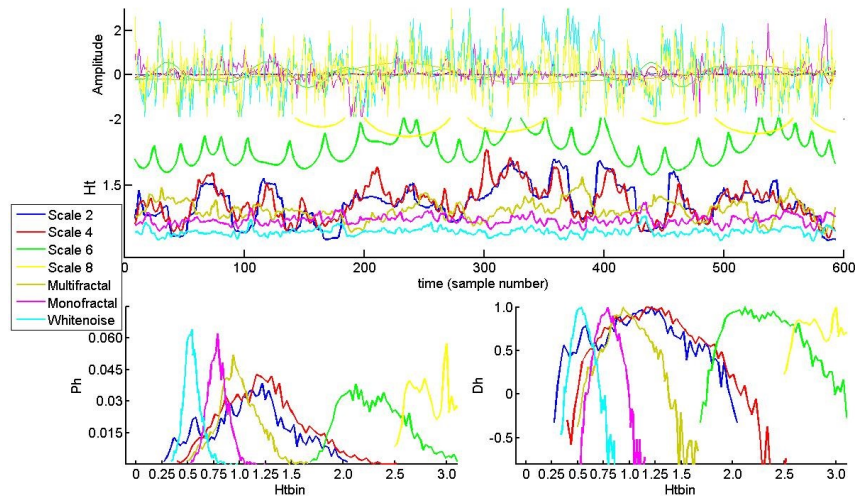


Figure 4.27: Comparison of the results of probability distributions of estimated noisy Duffing series and the smoothing probabilities of excess return series.

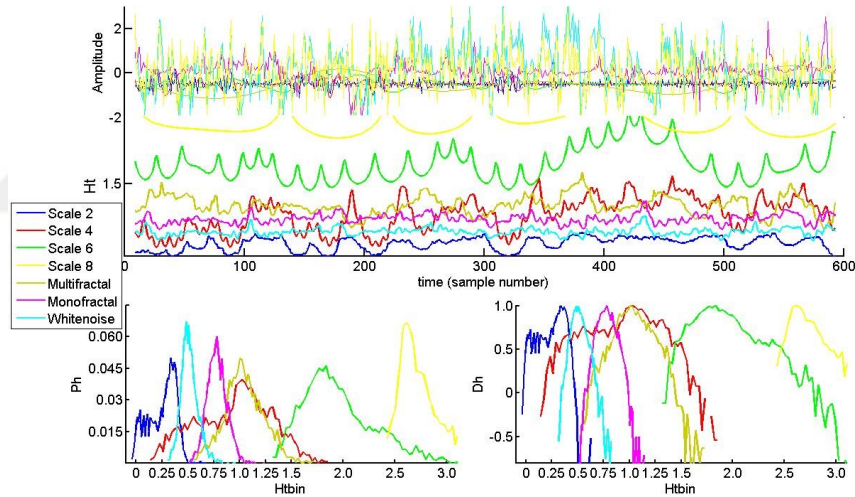
4.2.3.3. The Comparison of the Results of the Analyses of the Noisy Duffing Series and the Smoothed Probabilities for Excess Returns

In this section, to provide visual comparison for better evaluation, the findings so far are presented next to each other starting from the original data series and then continuing with the estimated versions. As seen through analyses so far, starting with MFDFA (Fig. 4.28) the noisy Duffing series and the smoothed probabilities for excess returns show identical regime switching patterns. Besides, estimation with telegraph process (Fig. 4.29) and fGn (Fig. 4.30) firstly individually and then together (Fig. 4.31) yield similar patterns. Probability distributions are also compared with their histograms before and after the estimation with telegraph process and fGn (Fig. 4.32) and supported mutual identical statistical properties.

Until here, the focus has been the introduction and estimation of the proposed method which is the combination of telegraph process and fGn and their analyses in detail. In subsequent section, the forecast practice and performance are explored with demonstrations comprehensively.



a. MF DFA results of the smoothed probabilities for excess returns, the length is 600 (Scale 2 = 1.2170, scale 4 = 1.1635, scale 6 = 2.1206, scale 8 = 3.0002, multifractal = 1.0304, monofractal = 0.7642, white noise = 0.4787).



b. MF DFA results of the noisy Duffing series with 600 number of data points (Scale 2 = 0.3319, scale 4 = 1.0252, scale 6 = 1.8345, scale 8 = 2.6054, multifractal = 1.0304, monofractal = 0.7642, white noise = 0.4787)

Figure 4.28: MF DFA results of the noisy Duffing series and the smoothed probabilities for excess returns

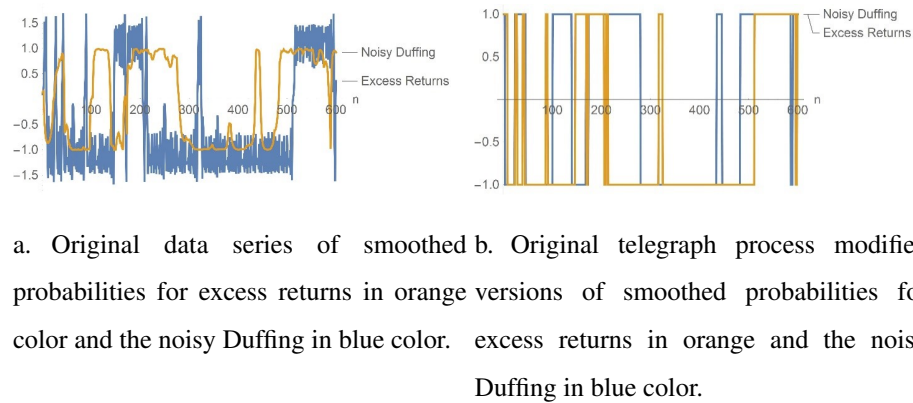


Figure 4.29: Telegraph process representations of the noisy Duffing series and the smoothed probabilities of excess returns with the original series.

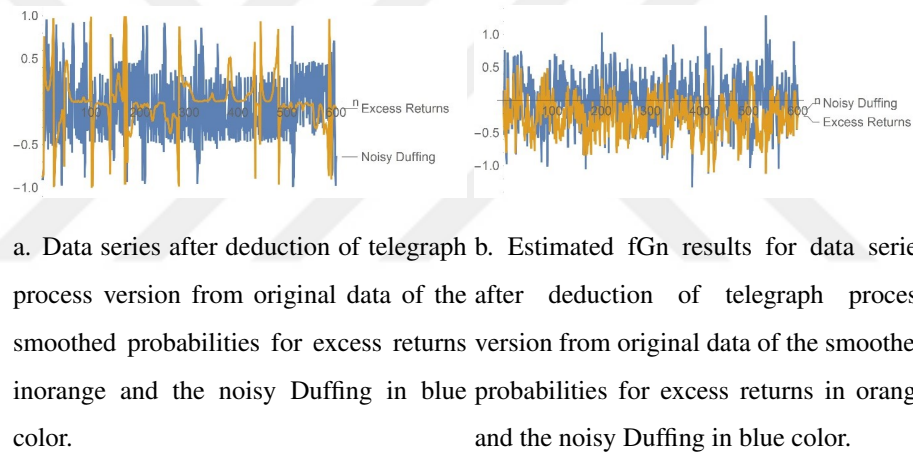
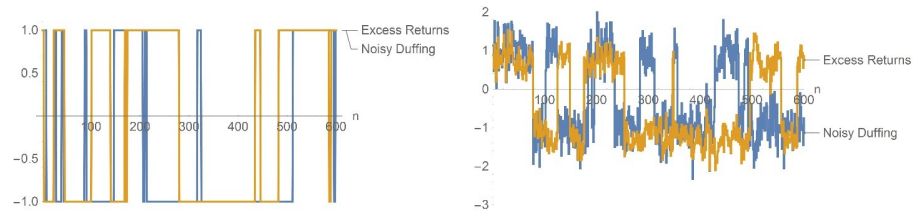
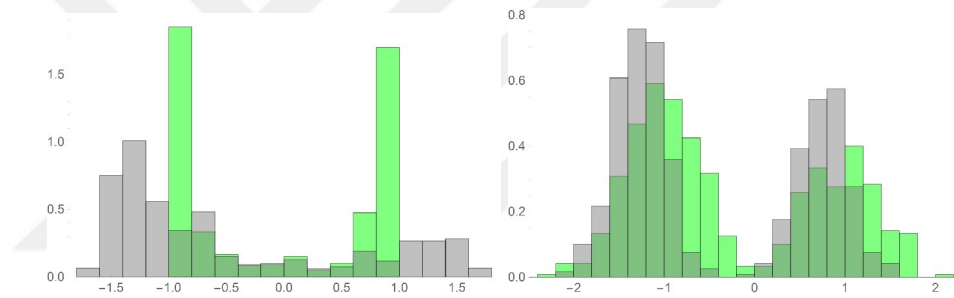


Figure 4.30: The noise-like section with the estimated fGn results for the noisy Duffing series and the smoothed probabilities of excess returns.



a. Estimated telegraph process series of the b. Estimated telegraph process and fGn smoothed probabilities for excess returns series of the smoothed probabilities for in orange color and noisy Duffing in blue excess returns in orange color and noisy Duffing in blue color.

Figure 4.31: Telegraph process and fGn estimations of the noisy Duffing series and the smoothed probabilities of excess returns.



a. Histogram of the smoothed probabilities b. Histogram of the smoothed probabilities for excess returns data series in gray color for excess returns data series estimated versus histogram of the noisy Duffing by telegraph process and fGn in gray data series with data length of 8000 and versus histogram of noisy Duffing data standard deviation of white noise process series estimated by telegraph process and of 0.01 in green. fGn with data length of 600 and standard deviation of white noise process of 0.019 in green.

Figure 4.32: Histogram results of the noisy Duffing series and the smoothed probabilities of excess returns.

4.2.4. Forecast Results

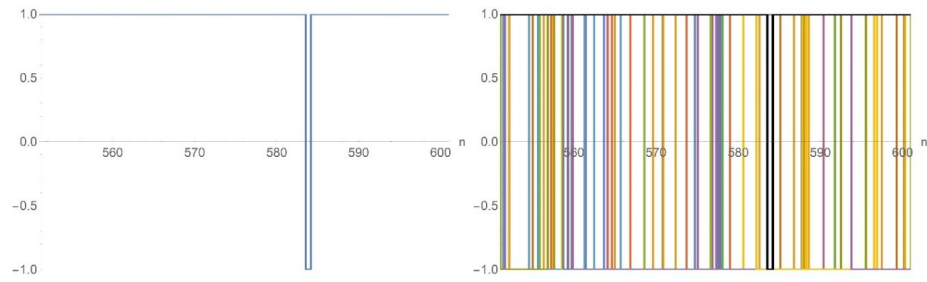
For the purpose of demonstrating the forecast efficiency of the method introduced, as a data set, the first 600 points are extracted. Then, the forecast method is employed for a number of time horizons that are 25, 50 and 100 steps ahead. With the λ value estimated for telegraph process, total number of 50 random processes are generated for each forecast range. The most likely prediction path for telegraph process is computed by averaging all the random paths generated as to fitted telegraph process.

Subsequently, the estimation is made for the noise-like section by fGn and Hurst exponent value is found with the standard deviation. Accordingly, with these parameters fGn is used to predict the related section. Eventually, by combining telegraph process and fGn prediction results, complete forecast for the data series is obtained. This methodology is applied for the forecast of the noisy Duffing series and the smoothed probabilities of excess returns. The accuracy of the forecast results are compared by plotting them with the actual results for clear evaluation.

Firstly, the forecast is demonstrated for the range of 50-step ahead which is fairly long for conventional methods. Despite of such a long forecast horizon, the forecast results are come out quite good as displayed in subsequent section in a progressive way. Next, the method is experimented on much longer forecast range of 100-step ahead and shorter forecast range of 25-step ahead. Once more, the forecast results similarly emerge as satisfactory. Conclusively, all these results achieved for different distances might be considered as an indication of the efficiency of the proposed method and as an alternative to conventional methods.

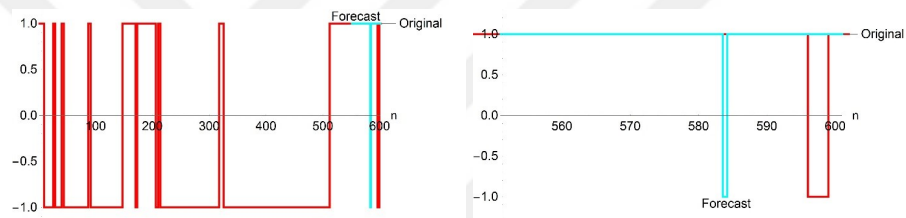
4.2.4.1. The Noisy Duffing Series Forecast Results for the Range of 50-step ahead

The telegraph process estimation for the noisy Duffing series for the range of 50-step ahead between the steps of 550 and 600 is given in Fig. 4.33. The average of 50 random telegraph process paths generated according to the estimated telegraph process of the noisy Duffing series is given in thick black line individually in Fig. 4.33.a and together with all other paths



a. The average path of 50 number of random processes for the noisy Duffing series
b. The average path over all random paths together for the noisy Duffing series

Figure 4.33: The average path of the 50 random processes generated for the noisy Duffing series.



a. Telegraph process forecast (in blue color) of the noisy Duffing series for 50-step ahead through all period.
b. Telegraph process forecast (in blue color) of the noisy Duffing series for 50-step ahead through the forecast period with the original series (in red color).

Figure 4.34: Telegraph process forecast of the noisy Duffing series.

in Fig. 4.33.b for apparent presentation.

Fig. 4.34.a displays the forecast results of telegraph process for the noisy Duffing series in blue color with the original series in red color through all period. Besides, the comparison of forecast results are presented in Fig. 4.34.b through forecast period with the actual path. In Fig. 4.35.a, fGn estimation with the blue line and the original noisy Duffing series with red line are shown for the forecast range of 50-step ahead. Fig. 4.35.b displays the full results comprised of fGn and telegraph process forecasts versus the actual outcome. As seen, the forecast and actual paths mostly move together with a minor deviation and the length of the regime diverges lightly.

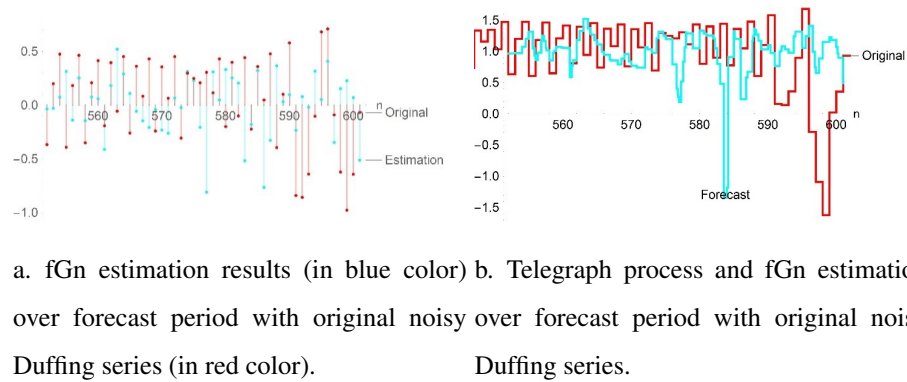


Figure 4.35: Telegraph process and fGn estimation through forecast range with the original noisy Duffing series.

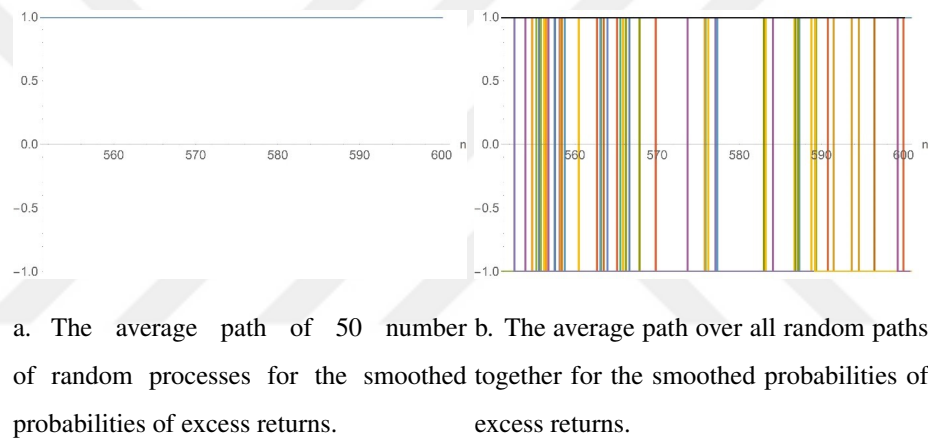
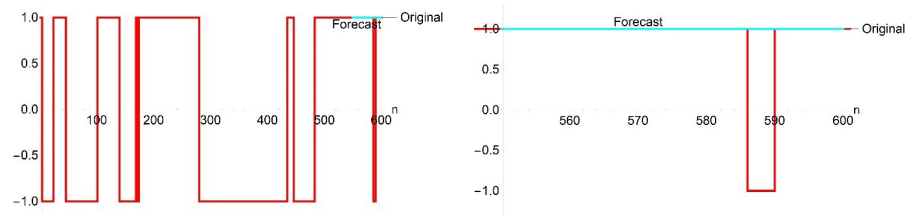


Figure 4.36: The average path of the 50 random processes generated for the smoothed probabilities of excess returns.

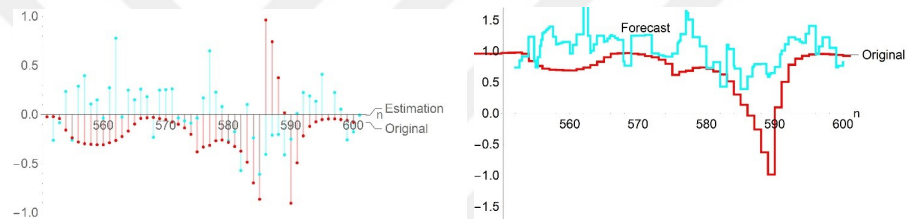
4.2.4.2. The Smoothed Probabilities of Excess Returns Forecast Results for the Range of 50-step ahead

The estimation results of telegraph process for the smoothed probabilities of excess returns over the range of 50-step ahead is given in Fig. 4.36. The average of 50 random generated paths is shown in Fig. 4.36.a individually and then in Fig. 4.36.b with all the other paths together. The forecast results of estimated telegraph process is given in Fig. 4.37 in blue color against the original series presented in red color over the whole time period and then over the forecast range next to each other. These results display close and parallel lines between forecast result and actual result through the most of the forecast range except the



a. Telegraph process forecast (in blue color) of the smoothed probabilities of excess return for 50-step ahead through all period.
b. Telegraph process forecast (in blue color) of the smoothed probabilities of excess return for 50-step ahead through the forecast period with the original series (in red color).

Figure 4.37: Telegraph process forecast of the smoothed probabilities of excess returns.



a. fGn estimation results (in blue color) over forecast period with original data series (in red color).
b. Telegraph process and fGn estimation over forecast period with original data series.

Figure 4.38: Telegraph process and fGn estimation through forecast range with the smoothed probabilities of excess return series.

short-remained regime switch around the point of 590. The fGn estimation of the smoothed probabilities of the smoothed probabilities for excess returns is shown in blue color in Fig. 4.38.a. Then, complete forecast results with the telegraph process and fGn against the original series over the range of 50-step ahead are displayed in Fig. 4.38.b. Although the noise-like section deviated from the actual path, in general the results are appeared to be close.

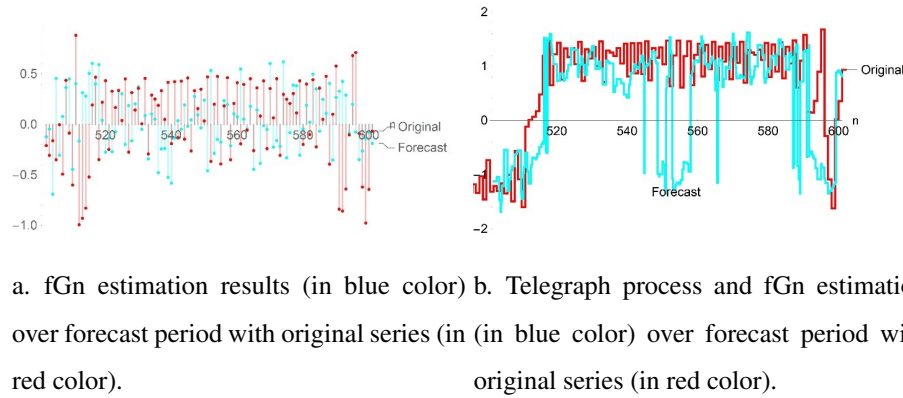


Figure 4.39: Telegraph process and fGn estimation for the forecast range of 100-step ahead for the noisy Duffing series.

4.2.4.3. The Noisy Duffing Series Forecast Results for the Range of 100-step ahead

The fGn forecast results of the noisy Duffing series are given in Fig. 4.39.a in blue color with the actual path in red color through the forecast range. Besides, in Fig. 4.39.b, the complete forecast outcome as a combination of telegraph process estimation and fGn estimation is displayed versus the original outcome. Despite the overestimated regime shifts along the forecast range, considering the quite long forecast horizon, the forecast outcome seems to be meeting with the actual outcome, particularly at the beginning and at the end of the forecast region. In addition, forecast outcome and the actual outcome move together closely at the several section for the volatile section where the fGn estimation performs.

4.2.4.4. The Smoothed Probabilities of Excess Returns Forecast Results for the Range of 100-step ahead

Fig. 4.40.a displays the fGn estimation with the blue line and the smoothed probabilities of excess returns with the red line. The full forecast result as the combination of telegraph process prediction and fGn prediction is illustrated in Fig. 4.40.b. According to these figures, it seems that the volatility is overestimated through the period taking into account the soft course of actual outcome through the forecast range.

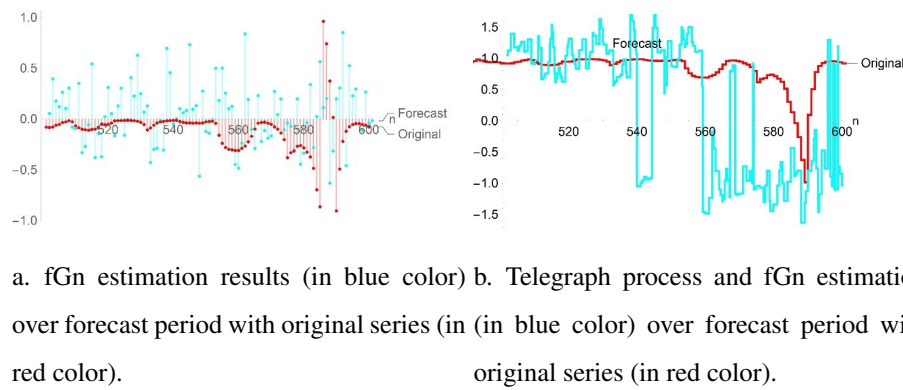


Figure 4.40: Telegraph process and fGn estimation for the forecast range of 100-step ahead for the smoothed probabilities of excess returns.

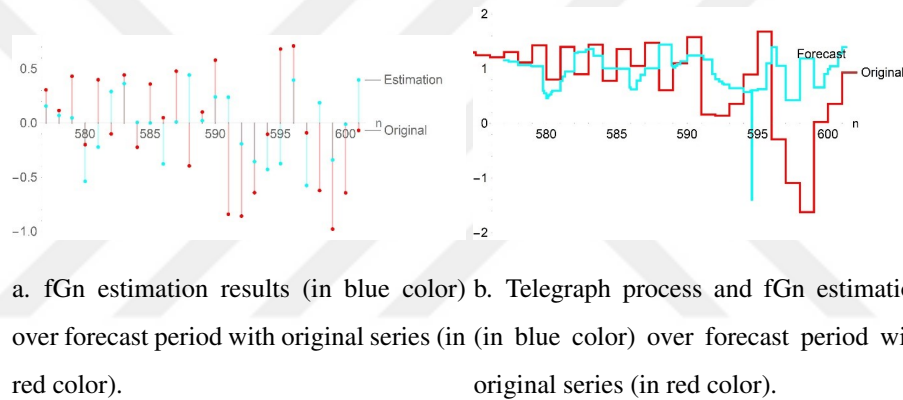


Figure 4.41: Telegraph process and fGn estimation for the forecast range of 25-step ahead for the noisy Duffing series.

4.2.4.5. The Noisy Duffing Series Forecast Results for the Range of 25-step ahead

In Fig. 4.41.a, the fGn estimation of the noisy Duffing map is displayed in blue color against the actual outcome in red color. The complete forecast with the telegraph process and the fGn in blue color are illustrated in Fig. 4.41.b with the original series in red color through the forecast period. Based on these results, the forecast outcome and actual outcome seems to be matching over most of the forecast range with relatively small deviations.

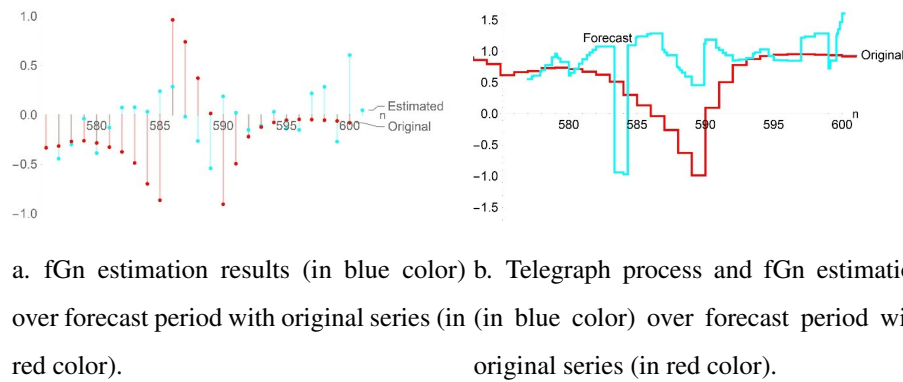


Figure 4.42: Telegraph process and fGn estimation for the forecast range of 25-step ahead for the smoothed probabilities of excess returns.

4.2.4.6. *The Smoothed Probabilities of Excess Returns Forecast Results for the Range of 25-step ahead*

Fig. 4.42.a shows the fGn estimation of the smoothed probabilities of excess returns and the original series. The complete forecast result which is the combination of telegraph process and fGn with blue line and the original series with red line is illustrated in Fig. 4.42.b. These figures indicate that the forecast and the original outcome are matched at the beginning and at the end of the forecast horizon. There are deviations in the middle of the forecast range but the location of the regime shift is captured even though the duration of the regime is not matched.

Conclusively, all these forecast results show that the telegraph process model yields mostly close results with varying characteristics in different forecast periods. Relatively good results are also obtained by fGn model particularly on the noisy Duffing series. On the other hand, the efficiency on the real-world series doesn't seem to be as good. Apparently, the fGn model produces much better estimations as the data possesses the multifractal and long memory features like the noisy Duffing series. If the real-world series resembles to the noisy Duffing series statistically, it is very likely to get much better results. Whatsoever, the constraint in the availability of the real-world financial data, the demonstrations had to be made on the data sets with the limited number of data points. The longer data sets may also help to increase the performance of the modelling and forecast results.

4.2.5. Final Remarks

The noisy Duffing map has interesting chaotic, multifractal and regime switching properties which are also found in a number of real-world systems at some extent including economic and financial systems. The noisy Duffing series can be used as a model for the systems with identical regime switching and multifractal characteristics that can be detected by wavelet analysis and MFDFA.

A new modelling and forecasting method comprised of the telegraph process and the fGn was introduced and demonstrated on the noisy Duffing series and the real-world financial time series. Favorable results suggest that the noisy Duffing map can be used as an alternative model for the real-world regime switching systems with the long-range dependence. In addition, the results of the method for modelling and forecasting showed its beneficial performance over the traditional methods even on the relatively long forecast periods and chaotic series.

These results encourage to try similar methodology on the other chaotic maps to find out if they can be used as a model for any real-world systems. Moreover, it is also inferred that the larger data sets can enhance the efficiency of the results.

This part of the research is also published (Yilmaz and Unal, 2019).

4.3. MULTISCALE CORRELATION DIMENSION METHOD

In this section, a new method of evaluating chaotic characteristics of nonlinear time series. It is named as the multiscale correlation dimension analysis (MSCD) since correlation dimension computation in multiple scale is the fundamental principle of the method.

As introduced in previous sections, correlation dimension was developed by Grassberger and Procaccia (Grassberger and Procaccia, 1983b) and used extensively for the analysis of complexity of time series. It evaluates strange attractors for describing deterministic chaos. Nevertheless, correlation dimension computation only function over single scale for the

analysis of the chaotic characteristics of time series.

Costa et al. (Costa et al., 2005) in 2005 introduced a multiscale complexity method for measuring the complexity of biological signals on multiple scales. It is also suggested that biological systems are required to function over multiple spatial and temporal scaled as a result of their complexity. Their method is comprised of the coarse-graining algorithm and the sample entropy as a complexity measure.

With a resembling approach, by enhancing the correlation dimension algorithm, we intend to measure the information supplied by physical and financial dynamical systems over multiple scales. So, we develop the multiscale correlation dimension (MSCD) method by utilizing mean filter, Gaussian filter and correlation dimension algorithm. We apply it on stochastic time series, chaotic time series and financial time series. Firstly, the various alternative filters are investigated for their suitability and for the explanation of why certain ones are preferred. These filters are Wiener filter, mean shift filter, bilateral filter, total variation filter, standard deviation filter, max filter, harmonic mean filter and gradient filter.

Fractal dimension (D) and Hurst exponent (H) has a relationship of $D = 2 - H$ (Mandelbrot, 1983). As D is a measure related to the chaotic nature of time series, H is a value that ranges between 0 and 1 and is associated with the long-memory dependence of time series. We examine this relationship between H and D in monoscale, by the assist of MSCD method in multiscale. These algorithms also are experimented on the stochastic time series, the chaotic time series and the financial time series.

An interesting and steady pattern is observed converging throughout the upper scales which can be expressed in a functional term as $D = 1.5 - H$. However, the same computations on the chaotic data series yield varying convergence outcomes for a number of chaotic map data series. As a result of these unique outcomes between the stochastic time series and the chaotic time series, this approach provides an identification tool for the time series, a new procedure for characterizing for the chaotic time series depending on the link between H and D in multiscale.

In next section, as a part of the MSCD method, the scaling filter, how the preference is made is explained by presenting the experiment outcomes of a number of filters. Subsequently, the mean filter or the coarse-graining algorithm and Gaussian filter as a preferred scaling filters and correlation dimension algorithm are introduced. Then, the MSCD method is applied and examined on stochastic financial time series and chaotic time series in individual subsections. After that, a new method is exhibited and an interesting links revealed between D and H in multiscale over the chaotic time series and stochastic time series are presented.

4.3.1. The Details of the Method

The Multiscale Correlation Dimension method is comprised of two successive steps. These are the filtering operation for scaling that is coarse graining procedure by choice and correlation dimension procedure. Coarse-graining procedure seems to be the most preferred method for forming multiple scales (Costa et al., 2005; Azami and Escudero, 2016; Shaobo et al., 2016; Zhao et al., 2018). Nevertheless, there are various low pass filter as the coarse-graining procedure especially employed for image processing. Therefore, the questions of that how other low pass filters perform and which ones would be the most suitable for generating multiple scales arise.

To find an answer to this question, numerous possibly alternative filtes are experimented on all the stochastic time series, the chaotic time series and financial time series used in this research. In addition to mean filter or coarse-graining filter, the filters found relevant to experiment are listed as Gaussian filter, Wiener filter, mean shift filter, bilateral filter, total variation filter, standard deviation filter, max filter, harmonic mean filter and gradient filter.

4.3.1.1. *Scaling Filters*

a) Coarse-graining Procedure (the Mean Filter)

Coarse-graining procedure or mean filter is a moving average process. It is applied to time series in an order by employing a scale factor τ with a low-pass filter. In this research, overlapping window is utilized to minimize data loss to a single one in each step. With the

help of windowing, in each step, data series are altered. The information in each interval which is linked to other intervals can be grabbed by each window.

For one-dimensional time series of $x_1, x_2, \dots, x_N$, coarse graining procedure is defined as:

$$y_{n,\tau} = \frac{1}{\tau} \sum_{i=0}^{\tau-1} x_{n+i} \quad (4.25)$$

where, consecutively, n, N, i and τ denote the subscript of coarse-grained data series, the length of data series, the loop key and the scale factor. Eq. 4.25 yields each data point on each scale. Besides, the procedure is presented with an depiction in Fig. 4.43. As scale 1 is the original time series, at each scale, values of subsequent data pairs are averaged to get the value of each point of next scale. As a result of this downsampling, the length of data is shrinked on each scale. But, it keeps the loss of data at minimum by decreasing the length of the series only by one that assists to deliver more information through scales, in comparison to other coarse-graining procedures that spend more number of data points.

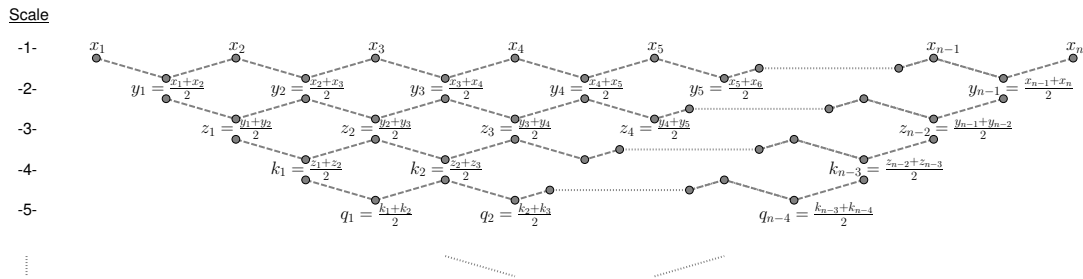


Figure 4.43: Multiscale overlapping coarse-graining algorithm

b) Gaussian Filter

Gaussian filter particularly has got many applications in image processing. Plainly, the average of weighting values replace the intensity value of the pixel and its neighbor pixels. The Gaussian filter employs convolution of required Gaussian function g . This function

which is governed by the variance σ^2 is defined as;

$$g(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-\left(\frac{x^2+y^2}{2\sigma^2}\right)} \quad (4.26)$$

Eq. 4.26 is utilized to estimate the coefficients for a Gaussian template and then convolved (Nixon and Aguado, 2012). The Gaussian filter can be applied in one or more dimensions. The advantage of the Gaussian filter compared to direct averaging is the increased efficiency as a result of maintaining more features.

c) Wiener Filter

The Wiener filter is simply a linear estimation of a signal and particularly, beneficial while working with noisy signals. So, it finds broad-range applications for linear prediction, signal restoration and system identification. For an original signal x and an additive noise n , the Wiener filter is defined as;

$$W(u, v) = \frac{H^*(u, v) S_{xx}(u, v)}{|H(u, v)|^2 S_{xx}(u, v) + S_{nn}(u, v)} \quad (4.27)$$

where u, v are the location parameters of frequency. $H(u, v)$ is blurring or degradation filter and H^* is its conjugate. $S_{xx}(u, v)$ denotes the power spectrum of the original signal which is computed by the Fourier transform of the signal autocorrelation. $S_{nn}(u, v)$ is the power spectrum of the additive noise which is acquired by the Fourier transform of the noise autocorrelation (Vaseghi, 2008).

d) Mean Shift Filter

Mean shift is mainly related to data clustering. It has many applications in the areas of computer vision, smoothing, segmentation and tracking. As Fukunaga's introduction, the mean shift vector described as;

$$m(x) = x_{t+1} - x_t = \frac{\sum_{i=1}^n g\left(\frac{x-x_i}{h}\right) x_i}{\sum_{i=1}^n g\left(\frac{x-x_i}{h}\right)} - x \quad (4.28)$$

where $g(x) = -K(x)$. K is a kernel (for example Gaussian kernel that is the most popular one) used in order to calculate the probability density. The algorithm works by iteratively calculating the mean of a window around a data point as moving the center of the window till the convergence. The mean shift vector is computed until the convergence for each point x_i as to a chosen search window. The algorithm looks for a local maximum of density of a distribution (Fukunaga and Hostetler, 1975; Comaniciu and Meer, 2002).

d) Bilateral Filter

Bilateral filter functions as a smoothing filter by replacing each point with the nonlinear combination of the neighbouring values. It has applications in denoising, optical-flow estimation, texture editing and so on. As I represents an image and p, q denotes some pixel locations, Bilateral filter is defined as:

$$BF[I]_p = \frac{1}{W_p} \sum_{q \in S} G_{\sigma_s}(\|p - q\|) G_{\sigma_r}(I_p - I_q) I_q \quad (4.29)$$

The normalization factor W_p in Eq. 4.29 is;

$$W_p = \sum_{q \in S} G_{\sigma_s}(\|p - q\|) G_{\sigma_r}(I_p - I_q) \quad (4.30)$$

and the two-dimensional Gaussian kernel $G_\sigma(x)$ in Eq. 4.30 is;

$$G_\sigma(x) = \frac{1}{2\pi\sigma^2} \exp\left(\frac{-x^2}{2\sigma^2}\right) \quad (4.31)$$

So, as to Eq. 4.31, G_{σ_s} is the Gaussian kernel linked with position that diminishes the effects of far points and G_{σ_r} is the Gaussian related to value which decreases the effects of points q with an intensity value different from I_p (Paris et al., 2007; Tomasi and Manduchi, 1998).

e) Total Variation Filter

The total variation algorithm was developed for image denoising and reconstruction (Rudin

et al., 1992). For a signal x with an additive noise n that is monitored in the form of $y = x + n$, in order to estimate x , total variation filtering technique calculates the amount of shifts between signal values and is defined as the minimization of subsequent formula:

$$J(x) = \|y - x\|_2^2 + \lambda \|Dx\|_1 \quad (4.32)$$

the L1 norm matrix $\|Dx\|_1$ provided in Eq. 4.32 can also be declared as;

$$\sum_{n=2}^N |x(i) - x(i-1)| \quad (4.33)$$

for $1 \leq i \leq N$ for N -point signal $x(i)$ in Eq. 4.33. And λ represents the regularization parameter(Selesnick and Bayram, 2010).

f) Standard Deviation Filter

In the standard deviation filter, the standard deviation of the data points in a specific range neighbourhood are employed. After that these are returned to the location of every data point. The standard deviation formula is provided as:

$$SD = \sqrt{\frac{\sum (x_i - \bar{x})^2}{r * c - 1}} \quad (4.34)$$

where x_i represents the value of specific pixel and $\bar{x}$ is the mean of the pixel values in the filter range. Moreover, r and c are the size of the filter in rows and columns consecutively(Research, 2020b).

g) Max and Min Filter

Max filter and min filter are categorized among nonlinear filters. They filter an image by refining only the minimum or maximum of all pixels in a local region $R_{u,v}$ of an image. Each pixel in an image is assigned a new value equal to the maximum or minimum value in

a neighborhood around itself. This process can be summarized as,

$$I'(u, v) \leftarrow \min\{I(u + i, v + j) \mid (i, j) \in R\}, \quad (4.35)$$

$$I'(u, v) \leftarrow \max\{I(u + i, v + j) \mid (i, j) \in R\}, \quad (4.36)$$

where R is the filter region, I and I' represent the original image and the filtered image, u and v denote the position parameters. Algorithms replace every value in time series by the maximum or minimum in a determined range (Burger and Burge, 2016).

h) Harmonic Mean Filter

Harmonic mean filter, basically is a different form of mean filter. It is very effective for the elimination of Gaussian noise as well as retaining edge features. For two dimensional space, the harmonic mean filter is described as;

$$HM(I) = \frac{mn}{\sum_{(i,j) \in W} \frac{1}{I(x+i, y+j)}} \quad (4.37)$$

where x and y represent the coordinates over the image I , i and j denote the coordinates in a window W with the size of m, n that are the length of each dimension (Marques, 2011; Research, 2020a). The algorithm operates as replacing every value by the harmonic mean value in a determined range.

i) Gradient Filter

The gradient of the function I at location (u, v) is given as a vector:

$$\nabla I(u, v) = \begin{pmatrix} I_x(u, v) \\ I_y(u, v) \end{pmatrix} = \begin{pmatrix} \frac{\partial I}{\partial x}(u, v) \\ \frac{\partial I}{\partial y}(u, v) \end{pmatrix} \quad (4.38)$$

In essence, the partial derivatives of horizontal and vertical lines construct the gradient

function. There are the horizontal and vertical gradient filters respond to sudden variations in horizontal axes and the vertical axis consecutively (Burger and Burge, 2016).

Gradient filter based on the vector provided in Eq. 4.38 compute the magnitude of the gradient of an image that is the rate of surge and defined as;

$$|\nabla I| = \sqrt{\left(\frac{\partial I}{\partial x}(u, v)\right)^2 + \left(\frac{\partial I}{\partial y}(u, v)\right)^2} \quad (4.39)$$

Considering the magnitude doesn't change when the image is rotated or oriented in a different location, it is particularly utilized in edge detection.

Findings on Scaling Filters

For the application of these low pass filters, built-in functions of Mathematica v11.0.1.0 are employed and the results are illustrated in a way to help for comparisons. Considering roughly each pair outcome is approximately similar, merely figures for FARIMA time series are presented in Fig. 4.44 with ten subfigures to prevent repetition and to summarize findings. Each subfigure belongs to each filter and exhibits correlation dimension value on y axis and scale number up to 12 on x axis.

As observed, mean filter, Gaussian filter, Wiener filter, mean shift filter, bilateral filter and total variation filter yields identical patterns with different values in each scale. It is not quite definite to conclude the most effective scaling filter due to very close patterns. Nevertheless, these filters are applied to upcoming analysis at each step and outcomes showed us the consistency of the mean filter and Gaussian filter relative to others. As a consequence, they are preferred to be used for the analysis as a scaling filter.

4.3.1.2. Correlation Dimension

Although there are various methods available for distinguishing between chaotic and stochastic systems, the correlation dimension method is perhaps the most fundamental one.

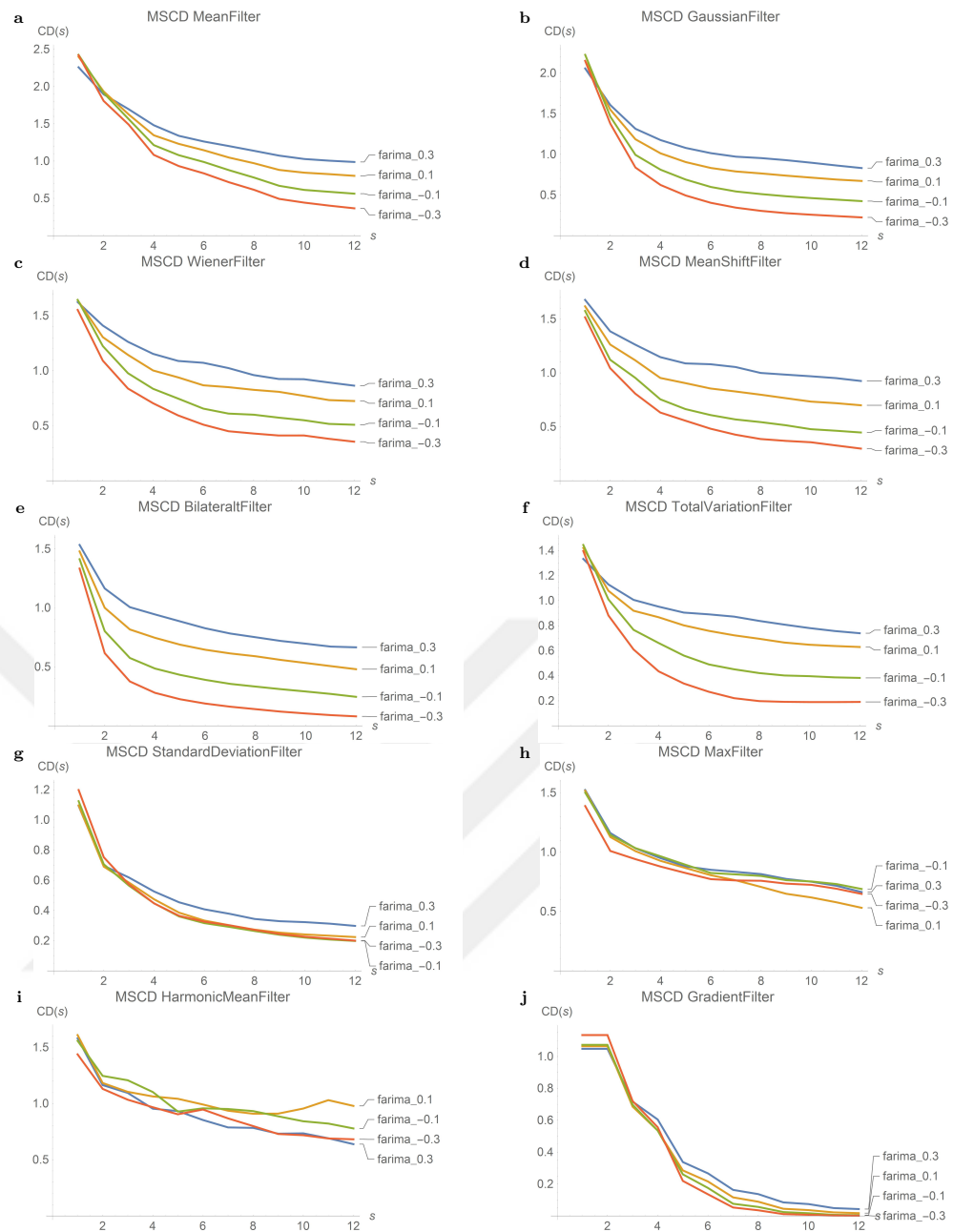


Figure 4.44: MSCD results of All Filters

The correlation dimension was developed by Grassberger and Procaccia (Grassberger and Procaccia, 1983a). Correlation dimension is a measure for the strangeness of the attractor. It consists of information about formation and final appearance of the attractor (Sprott and Sprott, 2003). The correlation summation at scale ε for a given embedding dimension is described as;

$$C_m(N, \varepsilon) = \frac{2}{N(N-1)} \sum_{j=1}^N \sum_{i=j+1}^N \Theta(\varepsilon - \|X_i - X_j\|) \quad (4.40)$$

where $\Theta(\cdot)$ is the Heavyside function;

$$\Theta(x) = \begin{cases} 0, & x \leq 0 \\ 1, & \text{otherwise} \end{cases} \quad (4.41)$$

and $N = T - (m-1)\tau$ is the number of embedding points in an m -dimensional space, T is the length of time series.

If $C_m(N, \varepsilon)$ scales like;

$$C_m(N, \varepsilon) = \varepsilon^D \quad (4.42)$$

Then, D is called the correlation dimension of the time series or equivalently the slop of the log-log plot of $C_m(N, \varepsilon)$ versus ε .

$$C_m(N, \varepsilon) = \frac{2}{N^2} \sum_{\substack{i,j \\ (1 \leq i < j < N)}} \Theta(\varepsilon - \sqrt{(X_{i+j} - X_i)^2 + \dots + (X_{i+j+(m-1)l} - X_{i+(m-1)l})^2}) \quad (4.43)$$

Correlation dimension calculation is made for each time series scaled by coarse-graining procedure in MSCD method. All correlation dimension outcomes for each scale are gathered in a plot on which y axis shows correlation dimension values and x axis shows the scale factor. In subsequent sections, MSCD method is demonstrated on chaotic time series and stochastic financial time series.

4.3.2. Applications

4.3.2.1. Chaotic Systems

A chaotic system is a complex dynamic nonlinear deterministic system that is unpredictable in the long run due to its sensitivity to variation of initial conditions. For that reason, even the smallest error in one state grows rapidly and transferred to the subsequent states. There are number of algorithms like correlation dimension to measure the chaotic properties of the systems in monoscale. In this section, MSCD is used to inspect the feature of chaotic structure of chaotic time series in multiscale. MSCD is applied to the three chaotic time series generated from two dimensional chaotic discrete maps of Duffing map, Henon map and Ikeda map.

a) Duffing Map

Duffing map is governed by two equations given below.

$$x_{n+1} = y_n \tag{4.44}$$

$$y_{n+1} = -bx_n + ay_n - y_n^3$$

So, Duffing chaotic time series are obtained by assigning the values of 2.75 and 0.2 to a and b parameters (Wiggins, 2003).

b) Henon Map

Henon map was introduced by Michel Henon in 1976 as;

$$x_{n+1} = 1 - ax_n^2 + y_n \tag{4.45}$$

$$y_{n+1} = bx_n$$

where a , b values are given 1.4 and 0.3 to generate chaotic Henon time series for the calculation (Hénon, 1976).

c) Ikeda Map

The Ikeda map is defined by these three equations (Ikeda, 1979).

$$\begin{aligned}x_{n+1} &= 1 + u(x_n \cos(z) - y_n \sin(z)) \\y_{n+1} &= u(x_n \sin(z) + y_n \cos(z)) \\z &= 0.4 - \left(\frac{6}{1 + x_n^2 + y_n^2} \right)\end{aligned}\tag{4.46}$$

where u parameter is set to 0.918 to generate chaotic time series.

Fig. 4.45 outlines the outcomes of MSCD computations on these three chaotic time series created by employing Eq. 4.45, Eq. 4.44 and Eq. 4.46 with the total number of 1250 data points. It is observed that correlation dimension values drop up to the scale factor nine and then converges to a constant value for Duffing apparently and Henon to some extent as Ikeda map seems to be diverging from them.

4.3.2.2. Stochastic Processes

For the purpose of demonstrating MSCD method on stochastic processes, we employ white noise and fGn time series. As white noise is essentially a random noise, fGn is a self-similar stochastic process with long-range dependence. These properties make them convenient for the examination and comparison of the algorithm with other specific time series.

Fractional Gaussian noise is a Gaussian noise with mean function μ and covariance function

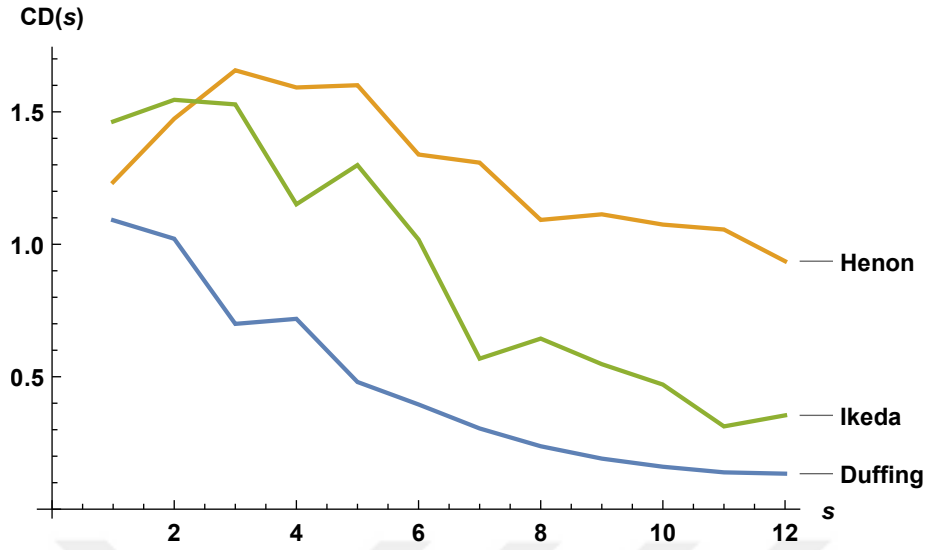


Figure 4.45: Multiscale correlation dimension results of chaotic maps

is defined as;

$$\frac{\sigma^2(|t-s-1|^{2H} - 2|t-s|^{2H} + |t-s+1|^{2H})}{2} \quad (4.47)$$

where σ is volatility and $H \in (0, 1)$.

As a data set, we use white noise with three different generated fGn processes with different Hurst values of 0.35, 0.5 and 0.75. Fig. 4.46 show the results of that when white noise MSCD curve decreases at each scale, all fGn values stay consistent. In comparison with chaotic time series, more regular and smooth lines through consecutive scales of the stochastic time series are observed.

4.3.2.3. Financial Time Series Models

a) FARIMA

Autoregressive fractionally integrated moving average model (FARIMA), is a modification of autoregressive process (AR) and moving average process (MA) models allowing fractional differencing.

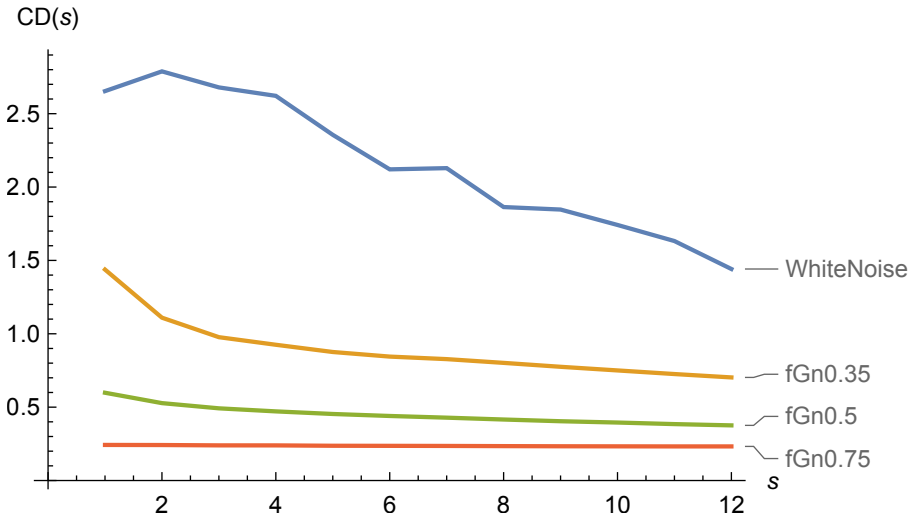


Figure 4.46: Multiscale correlation dimension results of stochastic processes

MA(q) model is defined as;

$$X_t = \mu + \sum_{i=1}^q \theta_i \epsilon_{t-i} + \epsilon_t \quad (4.48)$$

And AR(p) is written as;

$$X_t = \mu + \sum_{i=1}^p \phi_i X_{t-i} + \epsilon_t \quad (4.49)$$

where $\mu = E[y_t]$, θ and ϕ are the parameters of the MA model and AR model consecutively. ϵ is the white noise process with the properties of $E(\epsilon_t) = 0$ and $var(\epsilon_t) = \sigma^2$ (Brooks, 2008).

ARIMA(p, d, q) as a combination of AR(p) described in Eq. 4.49 and MA(q) presented in Eq. 4.48 models is defined as follows (Box et al., 2015);

$$(1 - \sum_{i=1}^p \phi_i B^i)(1 - B)^d(X_t - \mu) = (1 + \sum_{i=1}^q \theta_i B^i)\epsilon_t \quad (4.50)$$

where $(1 - B)^d$ is the difference operator as d gets integer values. B represents the backshift operator. It works in the notation as $B^i X_t = X_{t-i}$.

Therefore, ARIMA models presented in Eq. 4.50 are particularly useful when it comes to short-range dependence (Sheng et al., 2011). Besides, as d is allowed to take fractional values, it is implied that the model gets better at capturing long-range dependence (Granger and Joyeux, 1980). Then, as denoting the autoregressive order, the difference coefficient and the moving average order with p , d and q , in general form FARIMA(p, d, q) process can be described as;

$$\phi(B)(1 - B)^d X_t = \theta(B)\epsilon_t \quad (4.51)$$

where $d \in (-0.5, 0.5)$, $\phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$, $\theta(B) = 1 + \theta_1 B + \dots + \theta_q B^q$ and $(1 - B)^d$ is the fractional difference operator. The MSCD results of FARIMA is presented in Fig. 4.47.

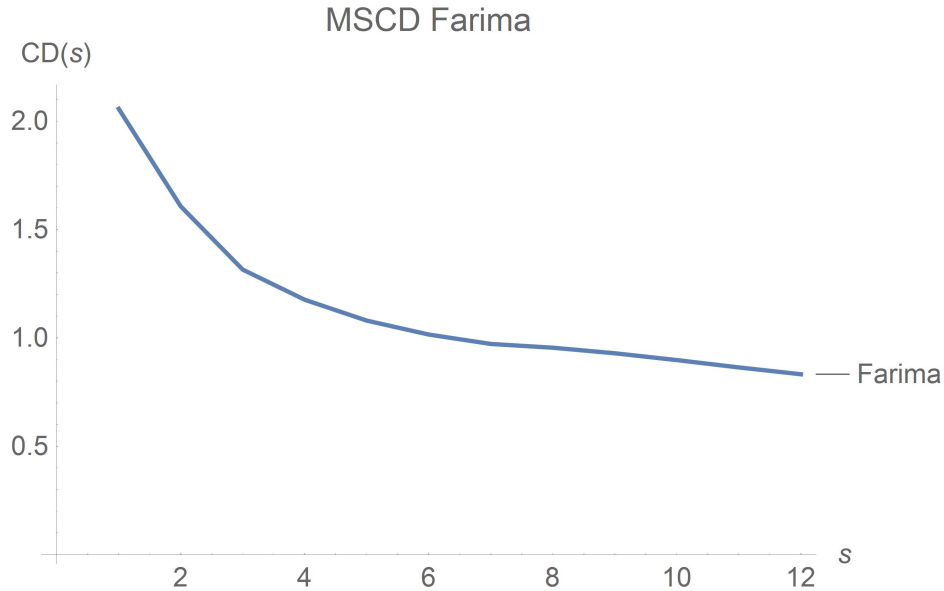


Figure 4.47: Multiscale Correlation Dimension Results of FARIMA

b) FIGARCH

The MSCD method is demonstrated on a specific financial time series of fractionally integrated generalized autoregressive conditional heteroskedastic (FIGARCH) process whose chaotic properties were examined thoroughly in previous sections.

FIGARCH model in a similar sense to FARIMA, is the modified version of GARCH model that provides more persistence on the conditional variance to be able to estimate long memory of conditional volatility. FIGARCH(p, d, q) is obtained by replacing the first difference operator with fractional differencing operator in GARCH(p, q) and described as;

$$\phi(L)(1-L)^d u_t^2 = \omega + [1 - \beta(L)]\epsilon_t \quad (4.52)$$

where $0 < d < 1$ (Baillie et al., 1996).

Although, FIGARCH appeared to possess fractal characteristics, examinations of chaotic features have concluded with the rejection of carrying deterministic chaotic properties. Correlation dimension calculations in a single scale were provided ambiguous outcomes by converging unevenly at some dimension points over 15. As a result of these findings, MSCD method is applied for checking it as an enhancement of single dimension correlation dimension analysis and validate former results on deterministic chaotic structure of FIGARCH. Fig. 4.48 presents the MSCD results of FIGARCH. The declining correlation dimension values seems to be supporting the prior findings of that FIGARCH can't be settled as deterministic chaotic process. Moreover, MSCD results indicates the hints of generating more obvious picture than its single scale equivalent on the FIGARCH process.

FIGARCH and FARIMA lines follow a similar pattern as in MSCD results of white noise and fGn to some extent examined previously. These similar patterns despite appearing in different values might imply that it is a characteristics of stochastic time series in multiscale.

These patterns are very different from the ones explored with the chaotic time series. In comparison with the chaotic time series, more regular and smooth lines through consecutive

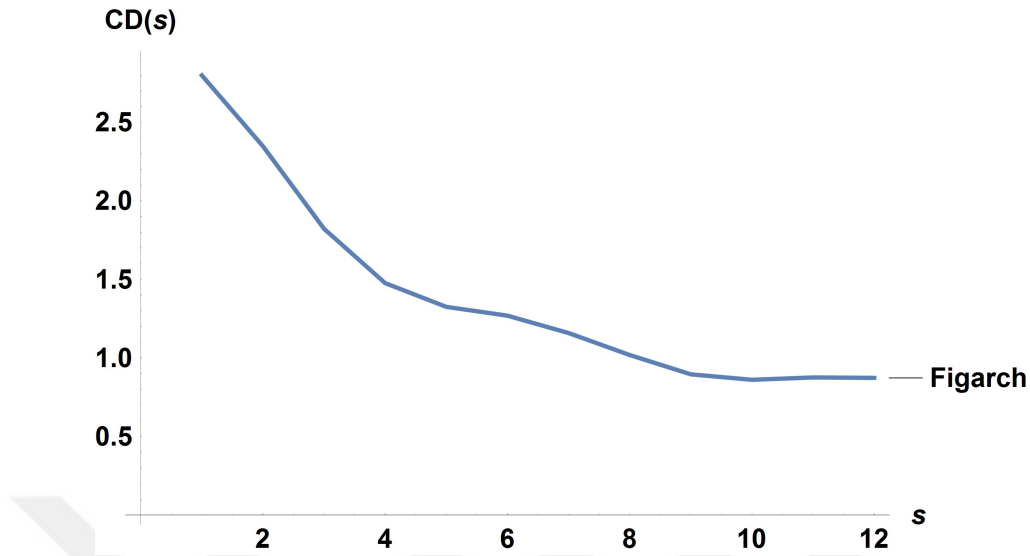


Figure 4.48: Multiscale correlation dimension results of FIGARCH

scales in the stochastic time series figure take the place of irregular and jagged lines in the chaotic time series.

4.3.3. The Relationship between Correlation Dimension and Hurst Exponent in Multiscale

The relationship between correlation dimension (D) and Hurst exponent (H) in multiscale is investigated in this section. Throughout previous sections, MSCD method has been introduced and demonstrated on time series with different properties. By MSCD method, correlation dimension has been extended to multiple scales and its distinguishing quality has been discovered. Similarly, Hurst value computations are utilized on multiple scales numerically, in order to investigate if the relationship between H and D also holds in multiple scales.

The relationship between Hurst value and fractal dimension was introduced by Mandelbrot

(Mandelbrot, 1983) as;

$$D = 2 - H \quad (4.53)$$

where, H represents the Hurst value and D represents the fractal dimension which is computed by correlation dimension algorithm. The Hurst value characterizes the level of long memory dependence and varies between 0 and 1. The system is independently distributed and Brownian motion, in case that $H = 0.5$. Besides, if H value is different from 0.5, then system carries memory of prior points and is not identified as independent any more. If $H < 0.5$, it is characterized as a short-term memory or anti-persistent system and if $H > 0.5$ as long-term memory or persistent system. Therefore, theoretically D and H values are associated to each other as in Eq. 4.59.

The relationship between D and H in multiscale is investigated and demonstrated on chaotic data series, stochastic data series and financial data series. There observed a considerable distinctive properties between the chaotic data series and stochastic data series. In addition, long-range dependence processes of fGn and fBm which are related to self-similar processes and described by the Hurst exponent are included to this analysis. Consequently, the time series data set for the analysis are comprised of white noise, fGn($H = 0.5$), fBm($H = 0.5$), FARIMA, FIGARCH processes and chaotic maps of Duffing map, Henon map and Ikeda map. As in previous section, several randomly generated processes with the length of 1250 data points are scaled and then the values of H and D computed for each scale.

As correlation dimension algorithm is used for the calculation of D as in the previous section, for the calculation of H value, Mandelbrot's method is applied (Mandelbrot, 1969) that is defined as;

$$R/S_t = ct^H \quad (4.54)$$

The formula is linearized by taking logarithms of both sides as;

$$\log(R/S_t) = \log(c) + H\log(t) \quad (4.55)$$

where R/S_t is,

$$R/S_t = [1/M \sum_{m=1}^M (R_{I_m}/\sigma_{I_m})] = ct^H \quad (4.56)$$

$$R_{I_m} = \max(X_{k,m}) - \min(X_{k,m}) \quad (4.57)$$

$$X_{k,m} = \sum_{k=1}^n (N_{k,m} - \mu_{I_m}) \quad (4.58)$$

where, σ , μ , I_m , N , $N_{k,m}$, c , t and k consecutively, represent the standard deviation, the mean of M sub-periods for $m = 1, 2, \dots, M$, each of the M subperiods, the number of points in time series, each element of a given time series, a constant, $t = N/M$ and $k = 1, 2, \dots, t$ (Raimundo and Okamoto Jr, 2018).

Time series of N observations are divided to t length of M subperiods. Standard deviation and mean value of each subperiods are computed. Then by calculation of variation R_{I_m} given in Eq. 4.57 dependent on $X_{k,m}$ computed as in Eq. 4.58 yields to estimation of the mean R/S_t in Eq. 4.56 by R_{I_m}/σ_{I_m} of all subperiods. Eventually, H value can be computed by a linear regression of Eq. 4.55 as a solution of relationship given in Eq. 4.54. This algorithm is employed to get H values with the correlation dimension algorithm for D at each scale that is acquired by Gaussian filter. Total number of scales are determined as 12.

Before the presentation of the results of the analysis of the H and D relationships in multiscale, the MSCD results of each time series are given in Fig. 4.49 with separate subplots. Multiple random samples are created for each stochastic process and their mean path is computed and then correlation dimension values are computed and illustrated as D-

Scale plot. The computations of the chaotic map are provided afterwards in the same setup. Each plot acquired for each time series are exhibited on a single figure in order to give effective visual review. Subplots are arranged starting from stochastic time series through financial time series to chaotic time series. In subplots, when y axis represents D and H values, x axis represents the scale number. MSCD method was already introduced and the results of those processes were given and reviewed in the previous section.

The outcome of the analysis of the relationship between H and D in multiple scales is displayed in Fig. 4.50 in which the subplots from a to h shows the changes of H and D values with $D + H$ for all time series separately. According to the outcomes, the Hurst exponent is related to the correlation dimension in multiscale for the stochastic time series. It is seen that there is a general order and convergence about the value of 1.5 as the sum of H and D values. This relation seems to be persistent for all stochastic processes but the chaotic maps under examination. So, the relationship between H and D can be formulized for the stochastic processes through upper scales as;

$$D = 1.5 - H \quad (4.59)$$

Contrary to the stochastic data series, the chaotic map series doesn't follow the same pattern. They converge around different values. While, in Duffing's case, the convergence value is 1.25, for Henon it happens at 1.1 and 0.75 for Ikeda map data series. These such variant observations show that how distinguishing properties the stochastic time series and the chaotic time series possess in terms of H , D and the sum of H and D in multiple scales.

Algorithms are run one by one for stochastic data series and the mean of all samples with Gaussian filter as well as the mean filter. Both filter yields closely steady results even though Gaussian filter slightly outperforms coarse-graining procedure. So, we present the mean value of the samples with Gaussian filter in the figures.

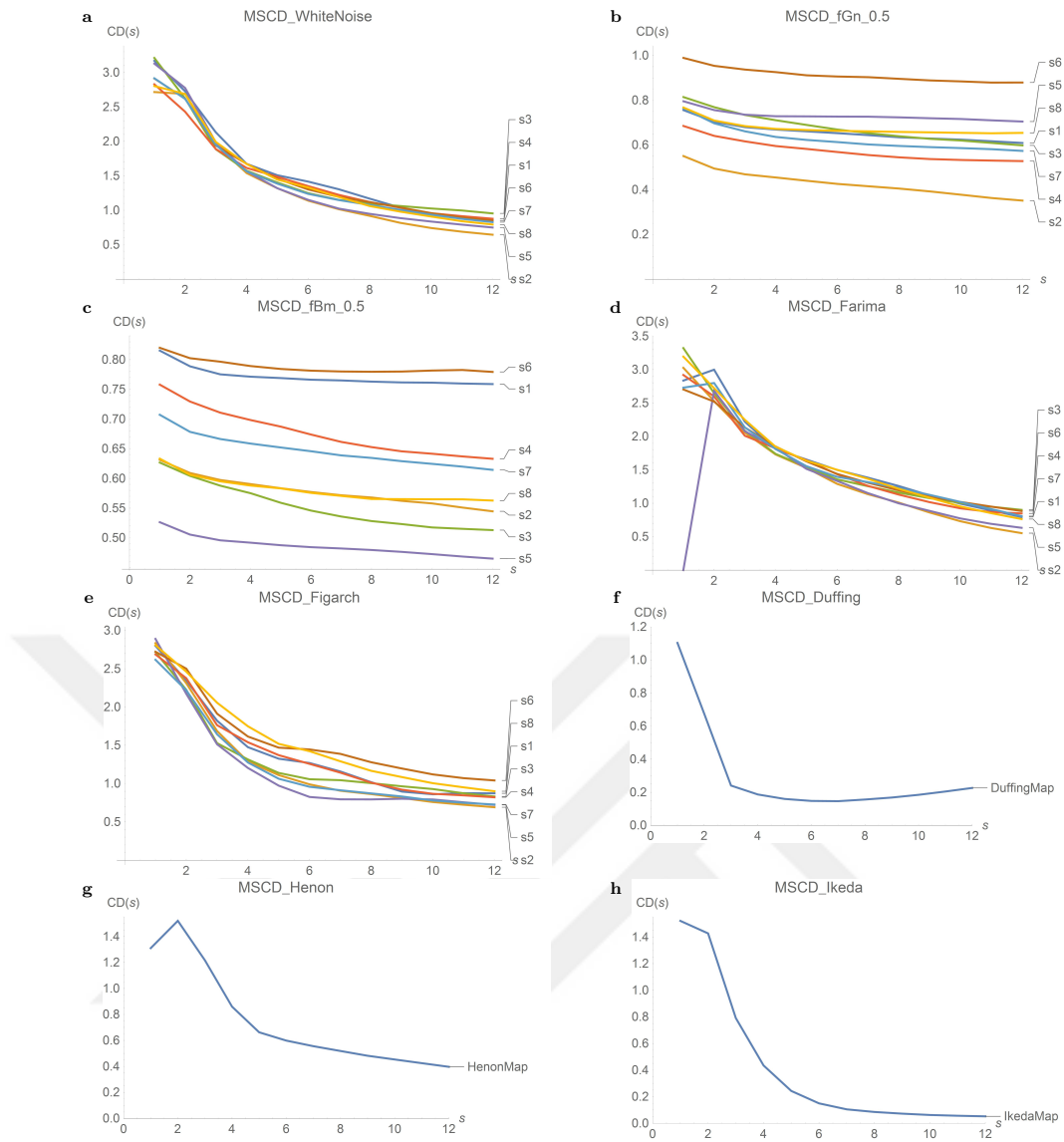


Figure 4.49: MSCD with Gaussian filter results for specific time series.

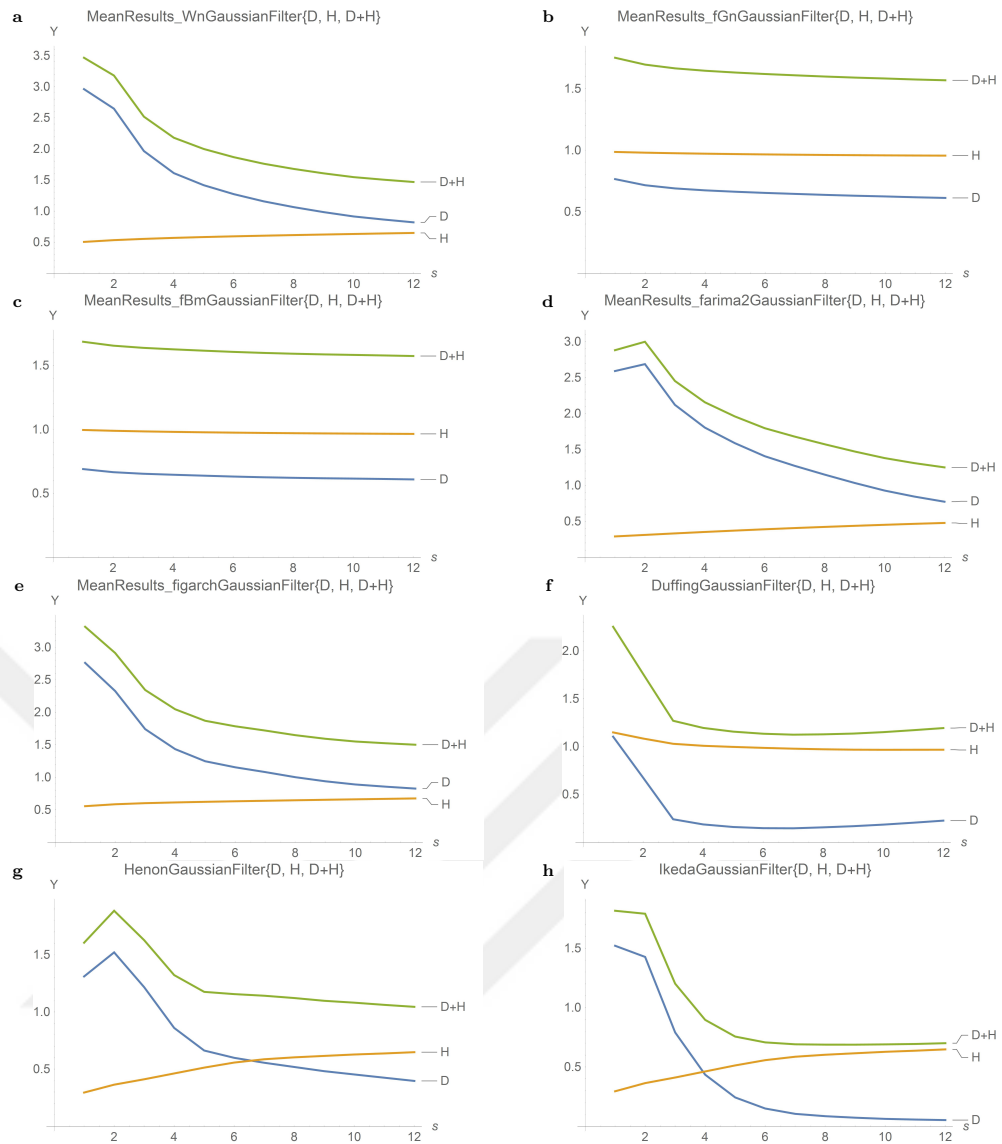


Figure 4.50: The relationship between D and H for specific time series.

4.3.4. Final Remarks

In this research the multiscale correlation dimension (MSCD) is introduced as a new measure for chaotic characteristics which captures multiscale properties for time series by utilizing coarse-graining procedure and Gaussian filter as a scaling tool and correlation dimension algorithm as a measure of chaotic structure.

The choice of the most convenient scaling filter is experimented by applying a number of advanced filtering methods on the stochastic and chaotic time series for which several filters are shown to be potentially effective as an addition to commonly used coarse-graining procedure in multiscale literature.

MSCD method is demonstrated on a number of stochastic time series and chaotic time series. Distinguishing results between these two different group of time series are observed to be supporting the applicability and functionality of the method as an alternative extension of current multiscale measuring methods.

Moreover, how the relationship between fractal dimension(D) and Hurst value(H) holds in multiple scales is examined by utilizing MSCD algorithm and stretching Hurst computation through scales at the same time, on the same time series data set. As a result, inspiring patterns are arised which is generalized as $D = 1.5 - H$. As this relationship is detected to be holding for the stochastic time series, opposite results are obtained for the chaotic time series. The results of these calculations indicate specific patterns for the stochastic time series and chaotic time series distinctively through multiple scales based on the sum value of D and H that suggests the possible implementation of the methods developed to reveal these unique multiscale properties for the categorization of time series.

This part of the research is also published (Yilmaz and Unal, 2020).

5. CONCLUSIONS

The focus of the researches that comprise this thesis is the nonstationary and nonlinear methods and subjects for the analysis of financial time series. A number of new methods, the enhancements and the new application methods for several existing techniques are developed, explored and demonstrated on both simulated and real-world time series.

At the first part of the thesis, wavelet coherence based methods are introduced and investigated in detail. Wavelet coherence is validated as a very powerful method for obtaining valuable information about dynamic correlation and its impact on time series. Its power in the analysis of the interrelationships between multiple time series in time-frequency space is demonstrated in various time series sets.

Firstly, wavelet coherence methods are applied for measuring interconnections between one of the major stock markets, S&P 500 and FTSE and several main international stock markets. As a result of these analysis, clear effects of S&P 500 and FTSE on other international markets are revealed in a great detail by providing information about how interdependencies change by country and in which frequencies and time intervals.

These unique methods subsequently are used to analyze the interrelationships between major international REITs. Once again, comprehensive information about these interrelationships are obtained in time-frequency space with which VARMA models are developed. Wavelet coherence powered VARMA model is demonstrated on these time series pairs at time intervals identified based on the strength of intercorrelations provided by wavelet analysis. As a consequence, very efficient forecast results are able to be acquired with much less errors by such composite model design.

Moreover, the wavelet coherence analysis are expanded with multiple wavelet coherence method (MWC) to allow the analysis of three different time series together. Hence, REITs are investigated with MWC for the comovements in groups of three time series. With the revelation of the interrelationships between these time series, powerful VARMA forecast

results are demonstrated according to time intervals identified.

All these results not only indicate the effectiveness of the wavelet coherence techniques but also the proper modeling of time series for the correctly determined time intervals. These relationships also provide quite valuable information, especially, for portfolio diversification and risk elimination. As a result of such relationships, from investor's perspective, it can be advantageous to diversify financial assets considering time periods and correlation frequencies. It is also important to take into account that shifting interrelations over time suggests that diversifications based on varying time periods could be beneficial.

At the second part of the thesis, the focus is nonlinear time series analysis. Fractionally integrated generalized autoregressive conditional heteroskedasticity (FIGARCH) emerges in modeling of financial time series. FIGARCH is essentially governed by a system of nonlinear stochastic difference equations. The chaoticity properties of FIGARCH stochastic difference equations are investigated by employing correlation dimension and Lyapunov exponents calculations. Numerical simulations of difference equations are used to determine the embedding dimension and embedding delay which then operates as parameters for the computation of the correlation dimension and the Lyapunov exponent. Fractal properties of FIGARCH is recognized by the correlation dimension and positive Lyapunov exponents acquired by Wolf's algorithm indicates possible chaotic qualities. However, application of Kant's algorithm and the method with Jacobian derivative based on the stochastic difference equations give negative Lyapunov exponents for all possible cases.

It is suspected that such conflicting results on Lyapunov exponent might be the issue of the embedding method which leads to the failure of Lyapunov exponent calculation under Takens' embedding theorem. Additionally, the Wolf's algorithm might fail to make use of all the available data due to its focus on one fiducial trajectory causing loss in information. As a result, it is concluded that these results suggest the rejection of FIGARCH possessing deterministic chaotic characteristics but with the hint of having some kind of nonlinear determinism. It could be misleading any observed irregularity and self-similarity to deterministic chaos, especially when the data is modelled by stochastic difference equations.

In subsequent section, a new stochastic modelling and forecasting method for the chaotic noisy Duffing map series and the financial time series with identical characteristics is introduced. The noisy Duffing map shares multifractal and regime switching properties with numerous real-world systems like financial and economic time series. So, the noisy Duffing map becomes a possible model for these systems as an alternative to traditional models. A new modelling tool based on the noisy Duffing map is developed by integrating the telegraph process and the fractional Gaussian process. It is successfully demonstrated on the noisy Duffing map series first and then on the real-world financial time series which is the smoothed probabilities for excess returns for both modelling and forecast.

As a consequence of favorable results even for relatively long forecast horizons and for the time series with chaotic influences, it is proposed as a very promising modelling and forecasting tool for the time series with the characteristics similar to the noisy Duffing series. It is also noticed that the longer data sets are very likely to increase the efficiency of results even further. Moreover, these methodology and results inspire studies on other chaotic maps as a potential modelling tools for real-world systems.

Eventually, a new method as an upgraded version of the correlation dimension for the analysis in multiple scales is proposed. Named as multiscale correlation dimension (MSCD) method, it is demonstrated on numerous stochastic time series and chaotic time series thoroughly. At the beginning, various scaling filters are examined for the purpose of the determination of the most efficient method for the MSCD, as well as a contribution to multiscale analysis literature in addition to conventional methods generally accepted. As a result of detailed experiments, Gaussian filter and mean filter are observed and suggested to be the most effective methods for scaling of time series.

The result of MSCD demonstrations support the efficiency and usefulness of the method for the analysis of time series with different characteristics. Based on MSCD, we further study the correlation dimension and Hurst exponent in multiscale with exploring exciting patterns on the same set of time series. We generalize these relationships observed as $D = 1.5 - H$ for the stochastic time series. On the other hand, this relationship is detected not to be holding for the chaotic time series suggesting that these unique patterns allows the

characterizing the chaotic and stochastic time series in a new way.

Even though, linear models mostly based on stationarity assumption govern the research and applications of time series, empirical time series are mainly nonstationary and nonlinear. Linear methods are easier to use and usually yield good enough approximation by describing regular structure in a time series. So, irregular structure of the data is associated to some random external noise. Nonlinear methods are still young and require more knowledge and experience because of the complexity, variety of methods and data sets. It is an extremely large area and quite hard to construct general theories. However, nonlinear methods are able to interpret these underlying irregular dynamical processes that is the essential part of the system. Nonlinear modelling area looks quite promising for improvement, development and exploration in near future due to the need for more efficient modelling tools and with the support of increase in computing power and the availability of large data sets.

REFERENCES

- Abarbanel, H. (2012). *Analysis of observed chaotic data*. Springer Science & Business Media.
- Addison, P. S. (2017). *The illustrated wavelet transform handbook: introductory theory and applications in science, engineering, medicine and finance*. CRC press.
- Aguiar-Conraria, L., Azevedo, N., and Soares, M. J. (2008). Using wavelets to decompose the time–frequency effects of monetary policy. *Physica A: Statistical mechanics and its Applications*, 387(12):2863–2878.
- Aguiar-Conraria, L. and Soares, M. J. (2014). The continuous wavelet transform: moving beyond uni-and bivariate analysis. *Journal of Economic Surveys*, 28(2):344–375.
- Almasri, A. and Shukur, G. (2003). An illustration of the causality relation between government spending and revenue using wavelet analysis on finnish data. *Journal of Applied Statistics*, 30(5):571–584.
- Azami, H. and Escudero, J. (2016). Improved multiscale permutation entropy for biomedical signal analysis: Interpretation and application to electroencephalogram recordings. *Biomedical Signal Processing and Control*, 23:28–41.
- Bachelier, L. (1900). Théorie de la spéculation. In *Annales scientifiques de l'École normale supérieure*, volume 17, pages 21–86.
- Baillie, R. T., Bollerslev, T., and Mikkelsen, H. O. (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 74(1):3–30.
- Barunik, J. and Vacha, L. (2013). Contagion among central and eastern european stock markets during the financial crisis. *arXiv preprint arXiv:1309.0491*.
- Bezruchko, B. P. and Smirnov, D. A. (2010). *Extracting knowledge from time series: An introduction to nonlinear empirical modeling*. Springer Science & Business Media.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31(3):307–327.

- Box, G. E., Jenkins, G. M., Reinsel, G. C., and Ljung, G. M. (2015). *Time series analysis: forecasting and control*. John Wiley & Sons.
- Bremaud, P. (2020). *Probability theory and stochastic processes*. Springer.
- Brooks, C. (2008). *Introductory econometrics for finance*. Cambridge university press.
- Budroni, M. A., Baronchelli, A., and Pastor-Satorras, R. (2017). Scale-free networks emerging from multifractal time series. *Physical Review E*, 95(5):052311.
- Burger, W. and Burge, M. J. (2016). *Digital image processing: an algorithmic introduction using Java*. Springer.
- Cajueiro, D. O. and Tabak, B. M. (2008). Testing for long-range dependence in world stock markets. *Chaos, Solitons & Fractals*, 37(3):918–927.
- Cochard, G.-M. (2019). *Introduction to Stochastic Processes and Simulation*. John Wiley & Sons.
- Comaniciu, D. and Meer, P. (2002). Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 24(5):603–619.
- Conrad, C. and Haag, B. R. (2006). Inequality constraints in the fractionally integrated garch model. *Journal of Financial Econometrics*, 4(3):413–449.
- Costa, M., Goldberger, A. L., and Peng, C.-K. (2005). Multiscale entropy analysis of biological signals. *Physical review E*, 71(2):021906.
- Das, A. and Das, P. (2006). Does composite index of nyse represents chaos in the long time scale? *Applied mathematics and computation*, 174(1):483–489.
- Das, A. and Das, P. (2007). Chaotic analysis of the foreign exchange rates. *Applied mathematics and computation*, 185(1):388–396.
- Daubechies, I. (1988). Orthonormal bases of compactly supported wavelets. *Communications on pure and applied mathematics*, 41(7):909–996.
- Daubechies, I. (1992). *Ten lectures on wavelets*, volume 61. Siam.

- Daubechies, I., Grossmann, A., and Meyer, Y. (1986). Painless nonorthogonal expansions. *Journal of Mathematical Physics*, 27(5):1271–1283.
- Davidson, J. (2004). Moment and memory properties of linear conditional heteroscedasticity models, and a new model. *Journal of Business & Economic Statistics*, 22(1):16–29.
- De Gooijer, J. G. et al. (2017). *Elements of nonlinear time series analysis and forecasting*. Springer.
- Dechert, W. D. and Gençay, R. (2000). Is the largest lyapunov exponent preserved in embedded dynamics? *Physics Letters A*, 276(1-4):59–64.
- Di Crescenzo, A. and Pellerrey, F. (2002). On prices' evolutions based on geometric telegrapher's process. *Applied Stochastic Models in Business and Industry*, 18(2):171–184.
- Douc, R., Moulines, E., and Stoffer, D. (2014). *Nonlinear time series: Theory, methods and applications with R examples*. Chapman and Hall/CRC.
- Dutta, S. (2010). Multifractal properties of ecg patterns of patients suffering from congestive heart failure. *Journal of Statistical Mechanics: Theory and Experiment*, 2010(12):P12021.
- Engle, R. (1982). Arch with estimates of variance of united kingdom inflation. *Econometrica*, 50(4):987–1007.
- Engle, R. F. and Bollerslev, T. (1986). Modelling the persistence of conditional variances. *Econometric reviews*, 5(1):1–50.
- Essex, C., Lookman, T., and Nerenberg, M. (1987). The climate attractor over short timescales. *Nature*, 326(6108):64–66.
- Fallahi, F. (2011). Causal relationship between energy consumption (ec) and gdp: a markov-switching (ms) causality. *Energy*, 36(7):4165–4170.
- Feldman, D. P. (2012). *Chaos and fractals: an elementary introduction*. Oxford University Press.

- Ferrara, L. and Guegan, D. (2000). Forecasting financial times series with generalized long memory processes. In *Advances in quantitative asset management*, pages 319–342. Springer.
- Foong, S. and Kanno, S. (1994). Properties of the telegrapher’s random process with or without a trap. *Stochastic processes and their applications*, 53(1):147–173.
- Franses, P. H., Van Dijk, D., et al. (2000). *Non-linear time series models in empirical finance*. Cambridge University Press.
- Fraser, A. M. and Swinney, H. L. (1986). Independent coordinates for strange attractors from mutual information. *Physical review A*, 33(2):1134.
- Fukunaga, K. and Hostetler, L. (1975). The estimation of the gradient of a density function, with applications in pattern recognition. *IEEE Transactions on information theory*, 21(1):32–40.
- Ghahramani, S. (2018). *Fundamentals of probability: with stochastic processes*. CRC Press.
- Goupillaud, P., Grossmann, A., and Morlet, J. (1984). Cycle-octave and related transforms in seismic signal analysis. *Geoexploration*, 23(1):85–102.
- Graham, M., Kiviahho, J., and Nikkinen, J. (2012). Integration of 22 emerging stock markets: A three-dimensional analysis. *Global Finance Journal*, 23(1):34–47.
- Granger, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: journal of the Econometric Society*, pages 424–438.
- Granger, C. W. and Joyeux, R. (1980). An introduction to long-memory time series models and fractional differencing. *Journal of time series analysis*, 1(1):15–29.
- Grassberger, P. and Procaccia, I. (1983a). Characterization of strange attractors. *Physical review letters*, 50(5):346.
- Grassberger, P. and Procaccia, I. (1983b). Measuring the strangeness of strange attractors. *Physica D: Nonlinear Phenomena*, 9(1-2):189–208.

- Grinsted, A., Moore, J. C., and Jevrejeva, S. (2004). Application of the cross wavelet transform and wavelet coherence to geophysical time series. *Non-Linear Processes in Geophysics*, 11:561–566.
- Grossmann, A. and Morlet, J. (1984). Decomposition of hardy functions into square integrable wavelets of constant shape. *SIAM journal on mathematical analysis*, 15(4):723–736.
- Günay, S. (2015). Chaotic structure of the bric countries and turkey's stock market. *International Journal of Economics and Financial Issues*, 5(2).
- Hafner, C. (2013). *Nonlinear time series analysis with applications to foreign exchange rate volatility*. Springer Science & Business Media.
- 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.
- Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica: Journal of the Econometric Society*, pages 357–384.
- Hamilton, J. D. and Raj, B. (2002). New directions in business cycle research and financial analysis. *Empirical Economics*, 27(2):149–162.
- Hegger, R., Kantz, H., and Schreiber, T. (1999). Practical implementation of nonlinear time series methods: The tisean package. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 9(2):413–435.
- Hénon, M. (1976). A two-dimensional mapping with a strange attractor. In *The Theory of Chaotic Attractors*, pages 94–102. Springer.
- Higuchi, T. (1988). Approach to an irregular time series on the basis of the fractal theory. *Physica D: Nonlinear Phenomena*, 31(2):277–283.
- Holmes, E. E., Lewis, M. A., Banks, J., and Veit, R. (1994). Partial differential equations in ecology: spatial interactions and population dynamics. *Ecology*, 75(1):17–29.

- Hurst, H. E. (1951). Long-term storage capacity of reservoirs. *Trans. Amer. Soc. Civil Eng.*, 116:770–799.
- Ihlen, E. A. F. E. (2012). Introduction to multifractal detrended fluctuation analysis in matlab. *Frontiers in physiology*, 3:141.
- Ikedda, K. (1979). Multiple-valued stationary state and its instability of the trasmitted light by a ring cavity system. *Optics communications*, 30(2):257–261.
- Janczura, J. and Weron, R. (2012). Efficient estimation of markov regime-switching models: An application to electricity spot prices. *AStA Advances in Statistical Analysis*, 96(3):385–407.
- Kac, M. (1974). A stochastic model related to the telegrapher's equation. *The Rocky Mountain Journal of Mathematics*, 4(3):497–509.
- Kantz, H. (1994). A robust method to estimate the maximal lyapunov exponent of a time series. *Physics letters A*, 185(1):77–87.
- Kantz, H. and Schreiber, T. (2004). *Nonlinear time series analysis*, volume 7. Cambridge university press.
- Kapitaniak, T. (2012). *Chaos for engineers: theory, applications, and control*. Springer Science & Business Media.
- Karatas, C., Unal, G., and Yilmaz, A. (2017). Co-movement and forecasting analysis of major real estate markets by wavelet coherence and multiple wavelet coherence. *Chinese Journal of Urban and Environmental Studies*, 5(02):1750010.
- Kim, C.-J. and Nelson, C. R. (1998). Business cycle turning points, a new coincident index, and tests of duration dependence based on a dynamic factor model with regime switching. *Review of Economics and Statistics*, 80(2):188–201.
- Kolesnik, A. D. and Ratanov, N. (2013). *Telegraph processes and option pricing*, volume 204. Springer.

- Lau, K.-M. and Weng, H. (1995). Climate signal detection using wavelet transform: How to make a time series sing. *Bulletin of the American meteorological society*, 76(12):2391–2402.
- Liu, K., Chen, Y., and Zhang, X. (2017). An evaluation of arfima (autoregressive fractional integral moving average) programs. *axioms*, 6(2):16.
- Lorenz, E. N. (1963). Deterministic nonperiodic flow. *Journal of the atmospheric sciences*, 20(2):130–141.
- Lowen, S. B. and Teich, M. C. (2005). *Fractal-based point processes*, volume 366. John Wiley & Sons.
- Lütkepohl, H. (2006). Forecasting with varma models. *Handbook of economic forecasting*, 1:287–325.
- Mallat, S. G. (1989). Multiresolution approximations and wavelet orthonormal bases of l^2 (r). *Transactions of the American mathematical society*, 315(1):69–87.
- Mallick, K. and Marcq, P. (2004). Noise-induced reentrant transition of the stochastic duffing oscillator. *The European Physical Journal B-Condensed Matter and Complex Systems*, 38(1):99–102.
- Mandelbrot, B. (1969). Long-run linearity, locally gaussian process, h-spectra and infinite variances. *International Economic Review*, 10(1):82–111.
- Mandelbrot, B. B. (1977). Fractals. *Form, chance and dimension*.
- Mandelbrot, B. B. (1983). *The fractal geometry of nature*, volume 173. WH freeman New York.
- Mandelbrot, B. B. and Van Ness, J. W. (1968). Fractional brownian motions, fractional noises and applications. *SIAM review*, 10(4):422–437.
- Mandelbrot, B. B. and Wallis, J. R. (1969). Computer experiments with fractional gaussian noises. *Water resources research*, 5(1):228–267.
- Manneville, P. (2010). *Instabilities, chaos and turbulence*, volume 1. World Scientific.

- Marques, O. (2011). *Practical image and video processing using MATLAB*. John Wiley & Sons.
- Michael, S. (2005). *Applied nonlinear time series analysis: applications in physics, physiology and finance*, volume 52. World Scientific.
- Moeini, A., Ahrari, M., Madarshahi, S. S., et al. (2007). Investigating chaos in tehran stock exchange index. *Iranian economic review*, 18:103–120.
- Nixon, M. and Aguado, A. S. (2012). *Feature extraction and image processing for computer vision*. Academic Press.
- Oxmetrics (2006). Time series modelling. <http://www.doornik.com>.
- Paris, S., Kornprobst, P., Tumblin, J., and Durand, F. (2007). A gentle introduction to bilateral filtering and its applications. In *ACM SIGGRAPH 2007 courses*, page 1. ACM.
- Peitgen, H.-O., Jürgens, H., and Saupe, D. (2006). *Chaos and fractals: new frontiers of science*. Springer Science & Business Media.
- Perkins, E. (2017). Effects of noise on the frequency response of the monostable duffing oscillator. *Physics Letters A*, 381(11):1009–1013.
- Raimundo, M. S. and Okamoto Jr, J. (2018). Application of hurst exponent (h) and the r/s analysis in the classification of forex securities. *Int. J. Model. Optim*, 8:116–124.
- Ratanov, N. (2007). A jump telegraph model for option pricing. *Quantitative Finance*, 7(5):575–583.
- Research, W. (2020a). Harmonic mean filter. <https://reference.wolfram.com/language/ref/HarmonicMeanFilter.html>. Accessed: 2019-09-11.
- Research, W. (2020b). Standard deviation filter. <https://reference.wolfram.com/language/ref/StandardDeviationFilter.html>. Accessed: 2019-09-11.
- Rosenstein, M. T., Collins, J. J., and De Luca, C. J. (1993). A practical method for calculating largest lyapunov exponents from small data sets. *Physica D: Nonlinear Phenomena*, 65(1-2):117–134.

- Rua, A. and Nunes, L. C. (2009). International comovement of stock market returns: A wavelet analysis. *Journal of Empirical Finance*, 16(4):632–639.
- Rudin, L. I., Osher, S., and Fatemi, E. (1992). Nonlinear total variation based noise removal algorithms. *Physica D: nonlinear phenomena*, 60(1-4):259–268.
- Rybski, D., Bunde, A., Havlin, S., Kantelhardt, J. W., and Koscielny-Bunde, E. (2011). Detrended fluctuation studies of long-term persistence and multifractality of precipitation and river runoff records. In *In Extremis*, pages 216–248. Springer.
- Schreiber, T. (1999). Interdisciplinary application of nonlinear time series methods. *Physics reports*, 308(1):1–64.
- Schroeder, M. (2009). *Fractals, chaos, power laws: Minutes from an infinite paradise*. Courier Corporation.
- Selesnick, I. W. and Bayram, I. (2010). Total variation filtering. *White paper*.
- Shang, P., Lu, Y., and Kamae, S. (2008). Detecting long-range correlations of traffic time series with multifractal detrended fluctuation analysis. *Chaos, Solitons & Fractals*, 36(1):82–90.
- Shaobo, H., Kehui, S., and Huihai, W. (2016). Modified multiscale permutation entropy algorithm and its application for multiscroll chaotic systems. *Complexity*, 21(5):52–58.
- Sheng, H., Chen, Y., and Qiu, T. (2011). *Fractional processes and fractional-order signal processing: techniques and applications*. Springer Science & Business Media.
- Sheppard, K. (2006). Mfe toolbox. <https://www.kevinshppard.com/code/matlab/mfe-toolbox/>.
- Sivakumar, B. (2004). Chaos theory in geophysics: past, present and future. *Chaos, Solitons & Fractals*, 19(2):441–462.
- Sørbye, S. H. and Rue, H. (2018). Fractional gaussian noise: Prior specification and model comparison. *Environmetrics*, 29(5-6):e2457.
- Sprott, J. C. and Sprott, J. C. (2003). *Chaos and time-series analysis*, volume 69. Oxford University press.

- Stadje, W. and Zacks, S. (2004). Telegraph processes with random velocities. *Journal of Applied Probability*, 41(3):665–678.
- Strogatz, S. H. (2018). *Nonlinear dynamics and chaos with student solutions manual: With applications to physics, biology, chemistry, and engineering*. CRC press.
- Takens, F. (1981). Detecting strange attractors in turbulence. In *Dynamical systems and turbulence, Warwick 1980*, pages 366–381. Springer.
- Tang, W. K. and Liu, Y. (2011). Formation of high-dimensional chaotic maps and their uses in cryptography. In *Chaos-Based Cryptography*, pages 99–136. Springer.
- Tarnopolski, M. (2016). On the relationship between the hurst exponent, the ration of the mean square successive difference to the variance and the number of turning points. *Physica A*, 461:662–673.
- Taylor, G. I. (1922). Diffusion by continuous movements. *Proceedings of the london mathematical society*, 2(1):196–212.
- Telesca, L., Lapenna, V., and Macchiato, M. (2005). Multifractal fluctuations in earthquake-related geoelectrical signals. *New Journal of Physics*, 7(1):214.
- Thompson, J. R. and Wilson, J. R. (2016). Multifractal detrended fluctuation analysis: Practical applications to financial time series. *Mathematics and Computers in Simulation*, 126:63–88.
- Tiwari, A. K. (2012). Decomposing time-frequency relationship between interest rates and share prices in india through wavelets. *MPRA Paper No. 39693*.
- Tomasi, C. and Manduchi, R. (1998). Bilateral filtering for gray and color images. In *Iccv*, volume 98, page 2.
- Torrence, C. and Compo, G. P. (1998). A practical guide to wavelet analysis. *Bulletin of the American Meteorological society*, 79(1):61–78.
- Trippi, R., Chorafas, D. N., and Foreword By-Sekiguchi, M. (1994). *Chaos theory in the financial markets*. McGraw-Hill, Inc.

- Tsay, R. S. and Chen, R. (2018). *Nonlinear time series analysis*, volume 891. wiley.
- Vaseghi, S. V. (2008). *Advanced digital signal processing and noise reduction*. John Wiley & Sons.
- West, B. J. (2012). *Fractal physiology and chaos in medicine*, volume 16. World Scientific.
- Wiggins, S. (2003). *Introduction to applied nonlinear dynamical systems and chaos*, volume 2. Springer Science & Business Media.
- Wolf, A., Swift, J. B., Swinney, H. L., and Vastano, J. A. (1985). Determining lyapunov exponents from a time series. *Physica D: Nonlinear Phenomena*, 16(3):285–317.
- Wong, K.-W. (2009). Image encryption using chaotic maps. In *Intelligent computing based on chaos*, pages 333–354. Springer.
- Yilmaz, A. and Unal, G. (2016a). Chaoticity properties of fractionally integrated generalized autoregressive conditional heteroskedastic processes. *Bulletin of Mathematical Sciences and Applications*, 15:69–82.
- Yilmaz, A. and Unal, G. (2016b). Co-movement analysis of asian stock markets against ftse100 and s&p 500: Wavelet-based approach. *International Journal of Financial Engineering*, 3(04):1650033.
- Yilmaz, A. and Unal, G. (2019). Stochastic duffing equation in modelling of financial time series. *International Journal of Dynamics and Control*, 7(4):1173–1194.
- Yilmaz, A. and Unal, G. (2020). Multiscale correlation dimension method. *International Journal of Modern Physics C (IJMPC)*, 31(03):1–17.
- Yilmaz, A., Unal, G., and Karatasoglu, C. (2016). Wavelet based analysis of major real estate markets. *Journal of Advanced Studies in Finance*, 7(2):107–116.
- Zhang, Y.-J. and Zhang, L. (2015). Interpreting the crude oil price movements: Evidence from the markov regime switching model. *Applied Energy*, 143:96–109.
- Zhao, X., Sun, Y., Li, X., and Shang, P. (2018). Multiscale transfer entropy: Measuring information transfer on multiple time scales. *Communications in Nonlinear Science and Numerical Simulation*, 62:202–212.