ADVANCES IN APPLIED NONLINEAR TIME SERIES MODELING

Dissertation
an der Fakultät für Mathematik, Informatik und Statistik
der Ludwig–Maximilians–Universität
München

vorgelegt von

Muhammad Yousaf Khan

aus Pakistan

München, 2015

Erstgutachter: Prof. Stefan Mittnik, Ph.D.
Zweitgutachter: Prof. Dr. Martin Missong
Tag der mündlichen Prüfung: 17-06-2015

Summary

Time series modeling and forecasting are of vital importance in many real world applications. Recently nonlinear time series models have gained much attention, due to the fact that linear time series models faced various limitations in many empirical applications. In this thesis, a large variety of standard and extended linear and nonlinear time series models is considered in order to compare their out-of-sample forecasting performance. We examined the out-of-sample forecast accuracy of linear Autoregressive (AR), Heterogeneous Autoregressive (HAR), Autoregressive Conditional Duration (ACD), Threshold Autoregressive (TAR), Self-Exciting Threshold Autoregressive (SETAR), Logistic Smooth Transition Autoregressive (LSTAR), Additive Autoregressive (AAR) and Artificial Neural Network (ANN) models and also the extended Heterogeneous Threshold Autoregressive (HTAR) or Heterogeneous Self-Exciting Threshold Autoregressive (HSETAR) model for financial, economic and seismic time series. We also extended the previous studies by using Vector Autoregressive (VAR) and Threshold Vector Autoregressive (TVAR) models and compared their forecasting accuracy with linear models for the above mentioned time series.

Unlike previous studies that typically consider the threshold models specifications by using internal threshold variable, we specified the threshold models with external transition variables and compared their out-of-sample forecasting performance with the linear benchmark HAR and AR models by using the financial, economic and seismic time series. According to our knowledge, this is the first study of its kind that extends the usage of linear and nonlinear time series models in the field of seismology by utilizing the seismic data from the Hindu Kush region of Pakistan.

The question addressed in this study is whether nonlinear models produce 1 through 4 step-ahead forecasts that improve upon linear models. The answer is that linear model mostly yields more accurate forecasts than nonlinear ones for financial, economic and seismic time series. Furthermore, while modeling and forecasting the financial (DJIA, FTSE100, DAX and Nikkei), economic (the USA GDP growth

rate) and seismic (earthquake magnitudes, consecutive elapsed times and consecutive distances between earthquakes occurred in the Hindu Kush region of Pakistan) time series, it appears that using various external threshold variables in threshold models improve their out-of-sample forecasting performance. The results of this study suggest that constructing the nonlinear models with external threshold variables has a positive effect on their forecasting accuracy. Similarly for seismic time series, in some cases, TVAR and VAR models provide improved forecasts over benchmark linear AR model. The findings of this study could somehow bridge the analytical gap between statistics and seismology through the potential use of linear and nonlinear time series models.

Secondly, we extended the linear Heterogeneous Autoregressive (HAR) model in a nonlinear framework, namely Heterogeneous Threshold Autoregressive (HTAR) model, to model and forecast a time series that contains simultaneously nonlinear and long-range dependence phenomena. The model has successfully been applied to financial data (DJIA, FTSE100, DAX and Nikkei) and the results show that HTAR model has improved 1-step-ahead forecasting performance than linear HAR model by utilizing the financial data of DJIA. For DJIA, the combination of the forecasts from HTAR and linear HAR models are improved over those obtained from the benchmark HAR model.

Furthermore, we conducted a simulated study to assess the performance of HAR and HSETAR models in the presence of spurious long-memory type phenomena contains by a time series. The main purpose of this study is to answer the question, for a time series, whether the HAR and HSETAR models have an ability to detect spurious long-memory type phenomena. The simulation results show that HAR model is completely unable to discriminate between true and spurious long-memory type phenomena. However the extended HSETAR model is capable of detecting spurious long-memory type phenomena. This study provides an evidence that it is better to use HSETAR model, when it is suspected that the underlying time series contains some spurious long-memory type phenomena.

To sum up, this thesis is a vital tool for researchers who have to choose the best forecasting model from a large variety of models discussed in this thesis for modeling and forecasting the economic, financial, and mainly seismic time series.

Zusammenfassung

Modellierung und Vorhersage von Zeitreihen sind in zahlreichen realen Anwendungen von großer Bedeutung. Nicht-lineare Modelle der Zeitreihenanalyse haben dabei zunehmend an Aufmerksamkeit gewonnen, nachdem die Schwächen linearer Modelle in zahlreichen empirischen Anwendungen evident wurden. Demgemäß wird in dieser Arbeit eine umfassende Auswahl sowohl gängiger als auch fortgeschrittener linearer und nicht-linearer Modelle der Zeitreihenanalyse betrachten, um deren out-of-sample Vorhersagegüte zu vergleichen. Konkret wurden dabei die Vorhersagegüte von Linear Autoregressive (AR), Heterogeneous Autoregressive (HAR), Autoregressive Conditional Duration (ACD), Threshold Autoregressive (TAR), Self-Exciting Threshold Autoregressive (SETAR), Logistic Smooth Transition Autoregressive (LSTAR), Additive Autoregressive (AAR) und Artificial Neural Network (ANN) Modellen, Heterogeneous Threshold Autoregressive (HTAR) or Heterogeneous Self-Exciting Threshold Autoregressive (HSETAR) bzw. das HSETAR Modell anhand von Finanz- und Wirtschaftsdaten, sowie seismologischer Daten eingehender untersucht. Darüber hinaus wurde anhand der Daten zusätzlich auch die Vorhersagegüte von Vector Autoregressive (VAR) und Threshold Vector Autoregressive (TVAR) Modellen betrachtet und mit den Ergebnissen gängiger linearer Modelle verglichen.

Im Gegensatz zu früheren Forschungsarbeiten, die Threshold Modelle meist mit intrinsischem Threshold spezifizieren, werden in dieser Arbeit Thresholds mit externen Übergangsvariablen modelliert. Die out-of-sample Vorhersagegüte wurde mit HAR und AR Modellen anhand von Finanz- und Wirtschaftsdaten, sowie von seismologischen Daten verglichen. Soviel uns bekannt ist, stellt diese Arbeit die erste Studie dieser Art dar, die lineare und nicht-lineare Modelle der Zeitreihenanalyse auf seismologische Daten in der Hindukusch Region von Pakistan anwendet.

Die Frage, der in dieser Studie nachgegangen wird, ist ob die Anwendung nonlinearer Modelle 1 bis 4 Schritte voraus zuverlässigere Vorhersagen treffen kann als die Vorhersage unter Anwendung linearer Modelle. Die Antwort ist, dass lineare Modelle genauere Vorhersagen für Zeitreihen von Finanz- und Wirtschaftsdaten sowie von seismologischen Daten erzielen, als nicht-lineare Modelle. Zudem scheint es, dass bei der Modellierung

und den Prognosen der Zeitreihen von Finanzdaten (DIJA, FTSE100, DAX und Nikkei), Wirtschaftsdaten (USA BIP Wachstumsrate) und seismologischen Daten (Erdbebenstärken, nachfolgend verstrichene Zeitdauern und Entfernungen zwischen Erdbeben in der pakistanischen Hindukuschregion) eine Verbesserung der out-of-sample Vorhersage erzielt, wenn man verschiedene externe Threshold-Variablen in Threshold-Modellen verwendet. Die Ergebnisse dieser Studie lassen einen positiven Effekt auf die Vorhersagegenauigkeit vermuten, wenn man nonlineare Modelle mit externen Threshold-Variablen erstellt. Auch für seismische Zeitreihen liefern in einigen Fällen TVAR und VAR Modelle bessere Vorhersagen als übliche lineare AR Modelle. Die Resultate dieser Studie könnten die analytische Kluft zwischen Statistik und Seismologie überbrücken, indem man lineare und nonlineare Zeitreihenmodelle anwendet.

Zudem erweiterten wir das lineare Heterogeneous Autoregressive (HAR) Modell auf ein nonlineares Modell, das so genannte Heterogeneous Threshold Autoregressive (HTAR) Modell, um eine Zeitreihe zu modellieren und vorherzusagen, die nonlineare und weitreichende Abhängigkeitsphänomene enthält. Das Modell wurde erfolgreich auf Finanzdaten (DJIA, FTSE100, DAX und Nikkei) angewendet, und die Ergebnisse zeigen, dass das HTAR Modell unter Nutzung der Finanzdaten von DJIA die Ein-Schritt-Vorhersagegüte gegenüber dem HAR Modell verbessert. Für DIJA ist die Kombination der Vorhersagen mittels HTAR und linearen HAR Modellen besser als die des HAR Modells alleine.

Des weiteren führten wir eine Simulation durch, um die Leistung von HAR und HSETAR Modellen zu beurteilen, wenn „spurious long-memory type“ Phänomene in einer Zeitreihe vorhanden sind. Das Hauptziel dieser Arbeit ist es, die Frage zu beantworten, ob HAR und HSETAR Modelle in der Lage sind, in einer Zeitreihe „spurious long-memory type“ Phänomene zu entdecken. Die Ergebnisse der Simulation zeigen, dass das HAR Modell es nicht schafft, zwischen „true“ und „spurious long-memory type“ Phänomenen zu unterscheiden, wohingegen das erweiterte HSETAR Modell dazu in der Lage ist. Diese Arbeit liefert einen Beweis dafür, dass es besser ist, das HSETAR Modell anzuwenden, wenn man vermutet, dass die zugrunde liegende Zeitreihe einige „spurious long-memory type“ Phänomene enthält.

Zusammenfassend ist zu sagen, dass diese Thesis ein wichtiges Instrument für Wissenschaftler bildet, die vor der Wahl stehen, aus einer Vielzahl von Modellen (die in dieser Arbeit diskutiert werden), das Modell mit der besten Vorhersagegüte zu wählen, um Finanz- und Wirtschaftsdaten sowie hauptsächlich seismologische Daten zu modellieren.

Acknowledgments

First and foremost, I thank Allah (subha na wa taala) for endowing me with health, patience, and knowledge to complete this work.

I acknowledge, with deep gratitude and appreciation, the inspiration, encouragement, valuable time, and guidance provided by Prof. Stefan Mittnik. I am very thankful for his encouraging and valuable comments at each and every step of my research.

I am very grateful to my parents and my wife, my son Ali, and my daughter Ayesha, for their emotional and moral support throughout my academic career and for their love, patience, encouragement, and prayers. I dedicate this work to my kind parents and to my loving wife.

I also acknowledge the sincere and valuable support of my colleagues at the University, especially, Naeem Ahmed Qureshi, Andreas Fuest and Serkan Yener. Special thanks go to Martina Brunner, secretary of my adviser, who helped me a lot in every matter. I would also like to say thanks to my friend, Firdos Khan for his support on many occasions.

I am particularly grateful to Matthieu Stigler, who helped me number of times to carry out the computational work in R language.

Muhammad Yousaf Khan, Munich

Contents

List of Figures

List of Tables

Mathematical Notations

Short form	Long form
IID or iid or $i.i.d$	Independent and Identically Distributed
e_t or ϵ_t	White Noise Innovation Process
n	Sample size of a given time series
I_t	Indicator Function used in Threshold Models
h	Forecasting Horizon
y_t	A random variable, representing a time series
$E(.)$	Expectation of a random variable
$M_.$	Earthquake Magnitude
RV_t	Time series of Realized Volatility
M_t	Time Series of Earthquake Magnitude
D_t	Time Series of Consecutive Distances between Earthquakes
E_t	Time Series of Consecutive Elapsed Times between Earthquakes
MC_t	Time Series of Monthly Earthquake Counts
X_t or Y_t	A vector of a specified dimension.

Abbreviations

Short form	Long form explanation
AAR	Additive Autoregressive Model
ACD	Autoregressive Conditional Duration Model
ACF	Autocorrelation Function
ADF	Augmented Dickey-Fuller test of stationarity
AIC	Akaike Information Criterion for Model Selection
ANN	Artificial Neural Network Model
AR	Autoregressive Model
BIC	Bayesian Information Criterion for Model Selection
BS	Bootstrapped
CLS	Conditional Least Square method of estimation
DGPs	Data Generating Process
HAR	Heterogeneous Autoregressive Model
HSETAR	Heterogeneous Self-Exciting Threshold Autoregressive Model
HTAR	Heterogeneous Threshold Autoregressive Model
vs	versus
ln	Natural Logarithm
LS	Least Square method of estimation
LSTAR	Logistic Smooth Transition Autoregressive Model
LSTAR-Ext.	Logistic Smooth Transition Autoregressive Model with external threshold variable
MC	Monte Carlo
MDM	Modified Diebold-Mariano test for predictive accuracy
MSFE	Mean Square Forecasting Error
NLS	Nonlinear Least Square method of estimation
PACF	Partial Autocorrelation Function
RV	Realized Volatility
SETAR	Self-Exciting Threshold Autoregressive Model
Std.	Standard Deviation
TAR	Threshold Autoregressive Model
TAR-Ext.	Threshold Autoregressive Model with external threshold variable
TVAR	Threshold Vector Autoregressive Model
VAR	Vector Autoregressive Model

Chapter 1

Introduction

1.1 Why Nonlinear Time Series Models?

In the field of time series analysis, there are many types of linear time series models that can be used to model and predict a given time series. There are broad classes of linear time series models used in modeling dynamics of a process: the Autoregressive (AR) models, the Integrated (I) models, and the Moving average (MA) models. These three types of linear models depend linearly on previous observations.

Although these models are still in use academic and applied research, mostly it has been found that simple linear time series models usually leave certain aspects of time series particularly economic and financial data unexplained. These findings have encouraged the use of nonlinear time series models, which are able to parsimoniously capture the nonlinear dynamics of a time series.

It is a well known fact that economic and financial systems posses through both structural and behavioral changes. Therefore to model such time series, it seems natural to allow for the existence of different states of the world or regimes and to allow the dynamics to be different in different regimes. According to Franses and van Dijk (2000), 'state-dependent dynamic behavior' of a time series means that certain properties of the time series such as its mean, variance and/or autocorrelation, are different in different regimes. They further show that the means and autocorrelations of returns and squared returns on stock market indices vary during the week. Hence, it can be said that each day of the week constitutes different regimes, which thus encourages the use of nonlinear time series models.

1.2 Overview of the Time Series Models

The ability to predict future and make informed decisions has always been one of mankind's ambitions. This thesis introduces some popular linear and nonlinear time series models and their extensions and expands their application to model and predict the time series from different fields of research, particularly from seismology. In the following paragraphs, the time series models considered in current thesis and their applications are briefly discussed.

Different studies have detected the nonlinearity in the time series of stock prices and exchange rates by various statistical tests (see Hinich and Patterson, 1985; Scheinkmann and LeBaron, 1989; Hsieh, 1989, 1991; Crato and de Lima, 1994; Brooks, 1996, among others). However, explicit modeling of this nonlinearity has received very little attention so far. Further information on nonlinear time series models can be found in articles and books by Tong (1990), Granger and Teräsvirta (1993), and Franses and van Dijk (2000).

LeBaron (1992) shows that the autocorrelations of stock returns are related to the level of volatility of these returns. In particular, autocorrelations tend to be larger and smaller respectively during the periods of low and high volatility. The periods of low and high volatility can be interpreted as distinct regimes. Another example is given by Kräger and Kugler (1993), who argue that exchange rates might show regime-switching behavior, in particular under a system of managed floating such as occurred in the 1980s when an attempt was made to to stabilize the exchange rate of the US dollar.

In recent years, several nonlinear time series models have gained much attention, and it is thought that these models are able to parsimoniously capture the nonlinear dynamics of a given time series. In this thesis, we focus our attention to Threshold Autoregressive (TAR), Self-Exciting Threshold Autoregressive (SETAR), Logistic Smooth Transition Autoregressive (LSTAR), Additive Autoregressive (AAR), and Artificial Neural Network (ANN) models. We also consider the linear Autoregressive (AR), Heterogeneous Autoregressive (HAR), and Autoregressive Conditional Duration (ACD) models, as well as the extended Heterogeneous Threshold Autoregressive (HTAR) and Heterogeneous Self-Exciting Threshold Autoregressive (HSETAR) models. We further assume that in threshold models (TAR and LSTAR) the regimes can be characterized by the lag values of an observable internal or external threshold variables.

The Threshold Autoregressive (TAR) models, initially proposed by Tong (1978) and Tong and Lim (1980), and discussed extensively in Tong (1990), are the extension of Autoregressive models developed in order to allow for a higher degree of flexibility in model parameters through a regime switching behavior. The TAR and SETAR models assumes that the regimes are determined by a specific lag value of the threshold variable. The models consists of different Autoregressive (AR) parts, each for a different regime. However, when the regime switching behavior observed in a time series is not discrete but rather smooth and continuous, the regime switching or transition function can be replaced with continuous transition function. Most popular transition functions include exponential function and first and second-order logistic functions. When the transition function is replaced with logistic function in TAR model, the resultant model is called Logistic Smooth Transition Autoregressive (LSTAR) model. LSTAR models are typically applied to time series data as an extension of Autoregressive models, in order to allow for higher degree of flexibility in model parameters through a smooth transition. Besides a number of other applications, Franses and van Dijk (2000) applied the SETAR and LSTAR models to the absolute values of weekly percentage returns on the Tokyo stock index. Various threshold models have been successfully applied to US GDP/GNP by Beaudry and Koop (1993), Potter (1995), and Pesaran and Potter (1997). The nonlinear behavior of seismic activities in a seismically active region in North China has been studied by Yang *et al.* (1995) by means of TAR models. The LSTAR model was originally suggested by Chan and Tong (1986) and subsequently developed by Teräsvirta (1994). The models were successfully applied to a wide range of industrial production series by Teräsvirta and Anderson (1992).

It has been shown that ANN model is able to approximate almost any nonlinear function arbitrarily close. Hence, when applied to a time series which is characterized by truly nonlinear dynamic relationships, the ANN model will detect these and provide a superior fit compared to linear time series models. The often quoted drawback of an ANN model is, it is difficult and hard to interpret its parameters. For this reason, ANN is mostly used for pattern recognition and forecasting. ANN models have been widely used for modeling and forecasting stock prices (Gencay, 1996; Haefke and Helmenstein, 1996a) and exchange rates (Kuan and Liu, 1995; Franses and van Griensven, 1998).

Nonlinear Autoregressive model, i.e., Additive Autoregressive (AAR) model, is nonlinear nonparametric model, which utilizes the cubic regression splines consisting of lagged time series values. Chen and Tsay (1993) applied the AAR model to river flow data of Iceland and argued that it has improved out-of-sample forecasting gain over linear AR model. Moustra *et al.* (2011) evaluated the performance of ANN model in predicting earthquakes occurring in the region of Greece by incorporating different types of input data.

Heterogeneous Autoregressive model, introduced by Corsi (2009), is a simple cascades long-memory model used to capture a few stylized facts observed in the high-frequency financial data. This model takes into account the volatility components defined over different time periods. The HAR model employs a few predictor terms including the past daily Realized Volatility (RV) averaged over different horizons (typically a day, a week, and a month), and is capable of producing slow-decay patterns in autocorrelations exhibited by many RV series. Its simple estimation and improved volatility forecasting performance prompted its use in several econometric studies, e.g., Andersen *et al.* (2007) on stock, exchange rate, and bond price volatility forecasting, and Forsberg and Ghysels (2007) and Martens *et al.* (2009) on volatility forecasting.

Vector Autoregressive (VAR) is a time series model used to capture the linear dependence among multiple time series. VAR model generalizes the univariate Autoregressive (AR) model by allowing for more than one evolving variable. However, conventional Vector Autoregressive (VAR) models are unable to capture the nonlinear dynamics such as regime switching and asymmetric responses to shocks. To capture regime dependencies and asymmetric responses to shocks, we apply a specified nonlinear Threshold Vector Autoregressive (TVAR) model for seismic data that is a multivariate extension of the univariate Threshold Autoregressive (TAR) model proposed by Tong (1978, 1983). A TVAR model is a relatively simple way to capture nonlinearities such as regime switching and asymmetry to shock responses. Aleem and Lahiani (2013) estimated exchange rate pass-through in Mexico by applying a Threshold Vector Autoregressive (TVAR) model. Afonso *et al.* (2011) applied a threshold VAR analysis to study whether the effects of fiscal policy on economic activity differ depending on financial market conditions.

Due to inherent flexibility of the nonlinear time series models, the possibility of getting improved in-sample fit to any time series data set is very high in comparison to linear mode. For this reason, in this thesis, we are much interested in the out-of-sample forecasting comparison of linear and nonlinear models through real data applications. Linear univariate Autoregressive (AR) and Vector Autoregressive (VAR) models are the simplest type of models and are used by many researchers to forecast the economic and financial time series (see, among others, Teräsvirta *et al.* 2005; Andersen *et al.* 2003; Preve *et al.* 2009).

The explicit objective of Autoregressive Conditional Duration (ACD) model developed by Engle and Russell (1998) is to model irregularly spaced financial transactions data. The modeling of duration is important in the sense that it may signal the arrival of new information concerning the underlying asset. A cluster of short durations corresponds to active trading and, hence, an indication of the existence of new information. Applications of ACD models to trade durations have been reported in numerous papers (see, among others, Engle and Russell, 1998; Jasiak, 1998; Engle, 2000; Manganelli, 2005; Bauwens, 2006).

Most of the computations in this thesis were carried out using the statistical software R (R Development Core Team, 2012).

1.3 Motivation and Brief Results

After researching current and past literature, we arrived at the conclusion that most of the time series models are applied to econometric data, and it is hard to find a study that involves their application to other areas of research, particularly seismology. Secondly, according to our knowledge there are no studies to date that utilize the external threshold variables in threshold models (TAR and LSTAR models) and evaluate its affect on the forecasting performance of threshold models. Thirdly, HAR is a simple long-memory time series model to capture the long-memory property of time series; however, it fails to simultaneously model the long-memory and asymmetric effects in a given time series and is unable to distinguish between true and spurious long-memory as shown in Chapter 4.

Economic theory suggests that number of important time series variables should exhibit nonlinear behavior. Therefore, to model and predict such type of nonlinear dynamics, it is necessary to use nonlinear models. The main purpose of this thesis is to overcome the above shortcomings by carrying the following analysis.

Firstly, we assess the dependence of threshold models (TAR and LSTAR) on external threshold variables regarding out-of-sample forecasting performance and compare these models with linear and threshold models with internal threshold variables. For this purpose, a detailed analysis of the forecasting performance of univariate, AR, HAR, HSETAR, TAR, SETAR, LSTAR, ANN, AAR, and VAR time series models for financial and economic data is conducted by including external threshold variables in Threshold Autoregressive models. Our main goal is to establish whether simple linear models still perform well or whether they should be replaced with the new sophisticated nonlinear models with a combination of internal and external threshold variables. To achieve this goal, we perform modeling and out-of-sample forecasting study by utilizing the financial (Dow Jones Industrial Average (DJIA), London stock exchange (FTSE100), German stock exchange (DAX), Japan stock exchange (Nikkei), and economic (US GDP growth rate with potential external threshold variables) time series. The results support our hypothesis that using the external threshold variables in threshold models has improved forecasting performance over the threshold models with internal threshold variables by utilizing the realized volatility and the GDP growth rate data. This study identifies significant external threshold variables that could be used in threshold models to obtain improved out-of-sample forecasts from these models.

Secondly, we extended the linear Heterogeneous Autoregressive (HAR) model to a nonlinear framework, namely Heterogeneous Threshold Autoregressive (HTAR) and Heterogeneous Self-Exciting Threshold Autoregressive (HSETAR) models, to simultaneously model and forecast time series that contain both the nonlinear and long-range dependence phenomena. The HTAR model is able to simultaneously describe the long-memory and asymmetric property of a given time series. We applied the HTAR model to the realized volatility data of Dow Jones Industrial Average (DJIA), London stock exchange (FTSE100), German stock exchange (DAX), and Japan stock exchange (Nikkei), and compare its out-of-sample forecasting performance with the linear HAR model. The out-of-sample forecasting results indicate that for DJIA, the HTAR model outperforms linear HAR model for 1-step-ahead forecast ($h = 1$). Furthermore, when combined with linear HAR model, the forecasts from (HTAR) model are more viable and flexible for purposes of forecasting realized volatility data of DJIA. HTAR model surpasses almost all nonlinear models in terms of one-step-ahead out-of-sample forecasting comparison by utilizing the financial data of DJIA, which is another significant contribution of this study. We also apply the HTAR model to a simulated time series containing spurious long-memory type phenomena and observed that it correctly detects this phenomena in all simulated Data Generating Process (DGP).

Thirdly, we carried out a study to assess the performance of HAR and HSETAR models in the presence of spurious long-memory type phenomena in a time series. It is known that the HAR model achieve the purpose of modeling the long-memory behavior of time series in a very simple and parsimonious way. In this study, we simulated time series through different Data Generating Processes (DGPS) that contain spurious long-memory type phenomena, whereafter HAR and HSETAR models were applied to these simulated time series to assess their performance. The results show that the HAR model is unable to distinguish between true and spurious long-memory type phenomena in all simulated time series across all sample sizes. However, HSE-

TAR model parsimoniously detects the spurious long-memory type phenomena in all simulated DGPs. Therefore, it is suggested to use HAR model with care whenever a time series possesses a structural break and consequently spurious long-memory type phenomena. In such circumstances it is suggested to use the HSETAR model to avoid misleading results.

Lastly, we introduced extensions of linear AR, ACD, and VAR and nonlinear TAR, SETAR, LSTAR, ANN, AAR, and TVAR models in the field of seismology and present time series analysis through linear and nonlinear time series models by utilizing the seismic data of the Hindu Kush region of Pakistan. According to our knowledge, this is the first study of its kind to be conducted in the field of seismology, by extending the linear and nonlinear time series models for modeling and forecasting the seismic time series (earthquake magnitudes, consecutive elapsed times and consecutive distances between successive earthquakes, and monthly earthquake counts). We first modeled the logarithmic seismic time series of earthquake magnitudes, consecutive elapsed times, consecutive distances between successive earthquakes, and monthly earthquake counts by using the linear and nonlinear models and then compared 1 through 4 step-ahead out-of-sample forecasting performance of nonlinear and VAR models with benchmark AR model. We also incorporated the external threshold variables in TAR and LSTAR models and achieved improved out-of-sample forecasting gain in threshold models over simple linear AR model for the time series of consecutive elapsed times. TVAR and VAR models also provided improved out-of-sample forecasts over benchmark AR model for the time series of consecutive distances. In second part of this study, we utilized the raw seismic data of consecutive elapsed times and consecutive distances and compared the out-of-sample forecasting performance of nonlinear and ACD models with benchmark AR model. Except for a few cases, the results favored the linear AR model when forecasting the raw seismic time series of consecutive elapsed times and consecutive distances.

This study is particularly important for Pakistan because it identifies potential time series models that could be used to model and predict seismic time series of the Hindu Kush region of Pakistan. This study also facilitates earthquake engineer and earthquake prediction analyst. The results of this study could also be utilized in constructing the response spectra, performing seismic risk and hazard analysis and these results are helpful to know the interdependence structure of seismic data.

1.4 Structure of the Thesis

This thesis is concerned with advances in applied nonlinear time series modeling, particularly focusing on the out-of-sample forecasting comparison of standard and extended nonlinear models with linear ones. It is a well known fact that many economic and financial systems go through both structural and behavioral changes (Zivot and Wang, 2006), and such theories have given rise to the use of nonlinear econometric models.

In Chapter 2, the linear and nonlinear time series models considered in this thesis are discussed. The extended Heterogeneous Threshold Autoregressive (HTAR) model and the estimation of threshold models, the models we are mostly interested in, are also discussed in this chapter. In addition, the multivariate Threshold Vector Autoregressive (TVAR) model is introduced in this chapter. To test the nonlinear property of time series, various nonlinearity tests were elaborated. Also, the forecasting procedure of nonlinear models is discussed. At the end of this chapter, the Modified Diebold-Mariano (MDM) test is illustrated in order to compare the forecasting accuracy of various competing models.

Chapter 3 focuses on the out-of-sample forecasting comparison of linear and nonlinear models. Financial (DJIA, FTSE100, DAX, and Nikkei) and economic (USA GDP growth) data were used to extend the threshold models with external threshold variables.

The third chapter of this thesis is concerned about the out-of-sample forecasting comparison of linear and nonlinear models by extending the threshold models with external threshold variables by utilizing the financial (DJIA, FTSE100, DAX, and Nikkei) and economic (USA GDP growth) data.

The performance of HAR and HSETAR models in the presence of structural break and consequently the spurious long-memory type phenomena in time series is assessed in Chapter 4. Different Data Generating Processes (DGPs) are used for simulation purposes that contain spurious long-memory type phenomena whereafter HAR and HSETAR models are applied to these DGPs.

In Chapter 5, an attempt is made to bridge the analytical gap between seismology and statistics by extending the application of many potential linear and nonlinear time series models in the field of seismology with the aim of understanding and predicting the seismic time series of Hindu Kush region of Pakistan. The conclusions of each study is given at the end of each chapter.

1.5 Definitions of Key Terms

The definitions of key terms used in this dissertation were taken from a study by Morley (2013). The seismological definitions were taken from the Glossary in Seismology[1].

Realized volatility

Volatility is a measure for variation and commonly represented through the square root of the second moment of a distribution. Realized Volatility (RV) is the square root of the sum of squared returns. More commonly, the realized variance is computed as the sum of squared intraday returns for a particular day. The realized variance is useful because it provides a relatively accurate measure of volatility which is useful for many purposes, including volatility forecasting and forecast evaluation (Andersen and Bollerslev, 1998).

Long memory time series

Long-range dependency is property of a time series, where the autocorrelation function of time series decay very slowly or decays more slowly than an exponential decay.

Forecast horizon

The future period of time for which a forecast is generated and it is commonly denoted by h.

Persistence

In time series modeling context the persistence is defined as the sum of the autoregressive coefficients.

[1] http://www.imd.gov.in/section/seismo/static/Glossary.pdf

Nonlinear time series in macroeconomics

A field of study in economics pertaining to the use of statistical analysis of data in order to make inferences about nonlinearities in the nature of aggregate phenomena in the economy.

Linear models

Refers to a class of models for which the dependence between two random variables can be completely described by a fixed correlation parameter.

Nonlinear models

Refers to the class of models for which the dependence between two random variables has a more general functional form than a linear equation and/or can change over time.

Structural change

A change in the model describing a time series, with no expected reversal of the change.

Time reversibility

The ability to substitute $-t$ and t in the equations of motion for a process without changing the process.

Aftershock

An earthquake that follows a large magnitude earthquake called, main shock and originates in or around the rupture zone of the main shock. Generally, major earthquakes are followed by a number of aftershocks, which show a decreasing trend in magnitude and frequency with time.

Arrival / Arrival Time

Arrival is the appearance of a wave, representing seismic energy, on a seismic record. The time at which a particular wave / phase arrives at a station or detector is called arrival time.

Body wave

Waves, which propagate through the interior of a body. For the Earth, there are two types of seismic body waves: (1) Compressional or longitudinal (P wave) and (2) Shear or Transverse (S wave).

Earthquake

Earthquakes are the manifestations of sudden release of strain energy accumulated in the rocks over extensive periods of time in the upper part of the Earth. Earthquakes are classified as, Slight ($M < 5.0$), Moderate ($5.0 < M < 6.9$) and Great ($M > 7.0$) depending upon the magnitude on Richter's scale. An earthquake having a magnitude, ($M < 2.0$) is termed as microearthquake.

Earthquake prediction

A statement, in advance of the event, of the time, location and magnitude of a future earthquake.

Epicenter

It is the point on the surface of the Earth, vertically above the place of origin (Hypocenter or Focus) of an earthquake. This point is expressed by its geographical coordinates in terms of latitude and longitude.

Fault

A fracture or fracture zone (a weak plane) in the Earth's crust or upper mantle, along which the two sides have been displaced relative to one another. Faults are caused by earthquakes and earthquakes are likely to recur on pre-existing faults, where stresses are accumulated.

Fault slip

The relative displacement of points on opposite sides of a fault, measured on the fault surface.

Focus (Hypocentre) / Focal Depth

A point inside the Earth, where the rupture of the rocks takes place during an earthquake and seismic waves begin to radiate. Its position is usually determined from arrival times of seismic waves recorded by seismographs. Focal depth is the vertical distance between the Hypocenter (Focus) and Epicenter.

Foreshock

A relatively small tremor (or an earthquake) that commonly precedes a relatively large magnitude earthquake (called the main shock), by seconds to weeks or months and originates in or near the rupture zone of the main shock.

Latitude

The location of a point north or south of the equator. Latitude is shown on a map or globe as east-west lines parallel to the equator.

Longitude

The location of a point east or west of the prime meridian. Longitude is shown on a map or globe as north-south line left and right of the prime meridian, which passes through Greenwich, England.

Magnitude

A measure of the strength of an earthquake or strain energy released by it, as determined by seismographic observations. The amplitude on a seismogram, the magnitude and the energy released are related through a log-linear relationship, which was originally defined by Charles Richter in 1935. An increase of one unit of magnitude (for example, from 4.6 to 5.6) represents a 10-fold increase in wave amplitude on a seismogram or approximately a 30-fold increase in the energy released.

Seismology

The word Seismology is derived from the Greek word ‘Seismos’ meaning earthquake and ‘Logos’ meaning science. Thus, it is the science of Earthquakes and related phenomena.

Seismic waves

They are the waves of energy caused by the sudden breaking of rock within the earth or by an explosion. They carry the released energy and travel through the earth and are recorded on seismographs. There are many types of seismic waves, for example., body waves, surface waves, coda waves, etc.

Plates and plate tectonics

The crust and upper mantle of the earth are made up of about a dozen large plates and several smaller ones that are constantly moving. The movements are very slow and only a few centimeters per year. Where the plates rub against one another, strain builds up, especially at the edges. When the strength of the rock is exceeded, the earth's crust may break and suddenly shift by several meters, causing an earthquake.

P-wave

P-waves are the fastest body waves and arrive at a station before the arrival of the S-waves, or secondary waves. P-waves are also called as Primary, longitudinal, irrotational, push, pressure, dilatational, compressional, or push-pull type wave. The P waves carry energy through the Earth as longitudinal waves, leading to the movement of particles in the same direction as the direction of propagation of the wave. P waves can travel through solid rock and fluids and are generally felt by humans as a bump.

Rayleigh wave

A type of surface wave having a retrograde elliptical motion of the particle, as the wave travels through the Earth's surface. These are the slowest, but often the largest and most destructive, of the wave types caused by an earthquake. They are usually felt as a rolling or rocking motion and move the ground up and down and side-to-side in the same direction that the wave moves. They are named after Lord Rayleigh, the English physicist, who predicted their existence in 1885. They are similar to the waves caused when a stone is dropped into a pond.

S wave

S-waves are the type of body waves, which move slowly in comparison to P waves (other type of body waves), but are usually bigger (in an earthquake). S-waves are also called as Shear, secondary, rotational, tangential, equivoluminal, distortional, transverse, pull or shake waves. S waves carry energy through the Earth as transverse waves, leading to the movement of particles in a direction perpendicular to the direction of propagation of the wave. S waves cannot travel through the outer core because these waves cannot exist in fluids, such as air, water or molten rock.

Source parameters (of an earthquake)

Source parameters of an earthquake include origin time, epicenter, focal depth, magnitude, focal mechanism and moment tensor for a point source model. They include fault geometry, rupture velocity, stress drop, slip distribution, etc. for a finite fault model. The parameters specified for an earthquake source depend on the assumed earthquake model.

Surface waves

Waves, which propagate along the surface of a body or along a subsurface interface. For the Earth, there are two common types of seismic surface waves: Rayleigh waves and Love waves.

Chapter 2

Linear and Nonlinear Time Series Models

2.1 Introduction

A growing body of regime switching models has been developed over the past two decades to capture the nonlinearities in different types of time series. It is mostly assumed in empirical econometric modeling that the relationship between variables are linear. In recent past years, arguments have been presented based on irregularities observed in economic and financial time series; therefore, nonlinear specification may be a more realistic representation of data generation processes. In finance, stock returns represents high and low volatility regimes. Exchange rate, gross domestic product, and many other time series may exhibit nonlinear phenomena. The stylized facts of financial returns are well known, and a range of models evolved from the pioneer Autoregressive Conditional Heteroskedasticity (ARCH) model (Engle, 1982) and later Generalized Autoregressive Conditional Heteroskedasticity (GARCH, Bollerslev (1986)) model and its different extensions up to the stochastic volatility and long-memory Heterogeneous Autoregressive (HAR) model (Corsi, 2009) are used to model these characteristics.

An enormous range of linear and nonlinear time series models that can be potentially useful for modeling and forecasting different kinds of time series exists in the literature. Time series analysts often face problems when choosing a model that is the most appropriate for to study the underlying phenomenon as there are only a very few studies about pros and cons of the standard time series models. A time series analyst usually has two main goals when analyzing a time series: first, to identify the underlying phenomenon; second, to model the phenomenon for description and forecasting purposes. A number of linear and nonlinear time series models designed for description and forecasting purposes can be found in the literature. The main problem encountered, when forecasting time series is the selection of the best model that can provide an accurate forecast. The application of nonlinear time series models in forecasting and testing the implications of economic theories has grown steadily recently.

"A natural approach to modeling economic time series with nonlinear models seems to be to define different states of the world or regimes, and to allow for the possibility that the dynamic behavior of economic variables depends on the regime that occurs at any given point in time (see Priestley, 1980, 1988)"(Frances and van Dijk, 2000). It means that certain properties of time series, like its mean, variance, and autocorrelation, vary over time. Lebaron (1992) showed that the autocorrelation of stock returns depends on the level of volatility. He argues that autocorrelation tends to be larger during a period of low volatility and smaller during a period of high volatility. The periods of high and low volatility can be interpreted as the regime governing process. Regime switching models were first introduced in the seminal work of Tong (1978) and Tong and Lim (1980), and elaborated later by Tong (1990).

In the last few years, a number of time series models have been extensively used to capture the regime switching behavior generated by different stochastic process. Rothman (1998) provided evidence that nonlinear models perform better than their linear counterparts when forecasting the unemployment rate in the USA. Peppenger and Goering (1998) and Chappell *et al.* (1996) showed that nonlinear models forecast better than linear models by using exchange rate data. Stock and Watson (1999) addressed different issues related to compare the forecasts from linear and two nonlinear models for 215 U.S. monthly macroeconomic time series. Teräsvirta and Anderson(1992) forecasted the industrial production with smooth transition autoregressive models. Sarantis (1999) forecasted real exchange rates using linear and nonlinear time series models and found that both classes of models show the same performance. Granger and Andersen (1978) were the first economists who introduced the univariate bilinear nonlinear time series model. Potter (1999) elaborately discusses the nonlinear time series models, their estimation procedures, and testing for nonlinearity.

In Section 2.2 different types of linear and nonlinear time series models are introduced. Their estimation procedures and lag length selection procedures are discussed in Sections 2.3 and 2.4. Section 2.3 also includes a discussion on the extended HTAR model, Section 2.5 contains an outline of the nonlinearity testing methodology, and the evaluation and forecasting methods are further elaborated in Sections 2.6 and

2.7. The Modified Diebold-Mariano (MDM) (Harvey *et al.*, 1997) test to compare the forecasting accuracy of time series models is discussed in Section 2.7.

2.2 Time Series Models

As the current thesis is concerned with the modeling and forecasting of a time series through various standard and extended time series models. So we first discussed all these models and issues related to them. As we are mostly interested in the forecasting comparison of linear and regime switching models, so we only discussed their estimation procedures in detailed and issues related to them.

2.2.1 Linear Autoregressive (AR) Model

A number of linear time series models have been described in literature, but here we only consider the linear Autoregressive (AR) model as an alternative benchmark to the nonlinear autoregressive models. The autoregressive model specifies that the output (dependent) variable depends linearly on its own previous values.

Consider the Autoregressive $AR(p)$ model of order p:

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \ldots + \alpha_p y_{t-p} + e_t \qquad (2.2.1)$$

In the above equation, the current values of y_t depend on its own past values with parameters $(\alpha_0, \alpha_1, \alpha_2, ..., \alpha_p)$, which are estimated by using ordinary least square (OLS) method. The innovation term e_t, is independent and identically distributed (IID) random variable, with mean zero and constant variance. In model (2.2.1) p is the lag order of the process determined by using model selection criteria such as Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). It is further assumed that the time series y_t is stationary.

2.2.2 Vector Autoregressive (VAR) Model

Vector Autoregressive (VAR) is a time series model used to capture the linear dynamics among multiple time series. VAR models generalize the univariate Autoregressive (AR) model by allowing for more than one evolving variable. VAR requires prior knowledge about variables which can be hypothesized to affect each other intertemporally. A VAR model describes the evolution of a set of k endogenous variables over the same time period as a linear function of only their past values. A VAR model with lag length p, denoted as VAR(p) is defined as:

$$y_t = c + A_1 y_{t-1} + A_2 y_{t-2} + \ldots + A_p y_{t-p} + e_t \tag{2.2.2}$$

where y_t is $k \times 1$ vector of endogenous variables and e_t is IID distributed white noise process term of the same dimension. The coefficient matrices $A_1, ..., A_p$ are of dimension $k \times k$ and estimated through OLS. The lag length in VAR(p) model can be selected through using information criteria such as AIC or BIC.

2.2.3 Heterogeneous Autoregressive (HAR) Model

Heterogeneous Autoregressive (HAR) model, introduced by Corsi (2009), is a simple cascades long-memory model used to capture a few stylized facts observed in the high-frequency financial data. This model takes care of the volatility components defined over different time period. Due to the fact that it is simple autoregressive type model in the realized volatility by considering its different components over different time horizons, thus named it Heterogeneous Autoregressive (HAR) realized volatility model. Among other characteristics, volatility distribution has fat tails, which means that extreme values are observed with relatively high probability. Also, volatility has long memory, therefore a value of 'long ago', for example, 20 days ago, may still have an impact on the current or future values. According to Corsi (2009) the observed financial data fluctuated in size of the price changes at all time scales, while standard GARCH and stochastic volatility short-memory models show white noise behavior once aggregated over longer time periods (no scaling behavior). Furthermore Corsi (2009) argues that the agents with different time horizons perceive and react to different volatility components namely, daily, weekly, and monthly. For a time series y_t, the resultant HAR model will have the following form:

$$y_t = c + \beta^{(d)} y_{t-1}^{(d)} + \beta^{(w)} y_{t-1}^{(w)} + \beta^{(m)} y_{t-1}^{(m)} + e_t \tag{2.2.3}$$

where $y_{t-1}^{(d)}$, $y_{t-1}^{(w)}$ and $y_{t-1}^{(m)}$ are respectively the daily, weekly, and monthly observed realized volatilities. The error term e_t in HAR model is assumed to be IID distributed white noise process. Equation (2.2.3) can be seen as a simple autoregressive three-factor realized volatility model, where the three factors are observed volatilities at different time frequencies. Standard ordinary least squares (OLS) method of estimation can be used to estimate the parameters in HAR model, yielding consistent and normally distributed estimates. 'naïve' method of recursive forecasting could be used to generate forecasts from HAR model.

2.2.4 Autoregressive Conditional Duration (ACD) Model

The Autoregressive Conditional Duration (ACD) model was originated by Engle and Russell (1998). In this model the durations between events (trades, quotes, price changes, etc.) are the quantities being modeled. Such type of duration data is irregularly spaced. In ACD models, conditional expected durations are modeled in a fashion similar to the way conditional variances are modeled using ARCH and GARCH models of Engle (1982) and Bollerslev (1986). ACD model has attain considerable interest among researchers representing different fields of science. Duration is commonly defined as the time interval between consecutive events, the time interval between two transitions of stock, or the consecutive elapsed times between seismic events. The modeling of duration between two consecutive events is important by providing some information about new arrivals. In finance, a cluster of short durations corresponds to active trading and hence provides an indication of the existence of new information. Similarly long durations between consecutive seismic events indicates a signal of another large magnitude seismic event at the same place.

Since duration is necessarily non-negative, the ACD model has also been used to model time series that consist of positive observations. An example is the daily range of the log price of an asset. Tsay (2009) argued that for a given asset, longer durations indicate lack of trading activities, which in turn signify a period of no new information. On the other hand, arrival of new information often results in heavy trading and, hence, leads to shorter durations. The dynamic behavior of durations thus contains useful information about market activities. Engle and Russell (1998) used an idea similar to that of the Generalized Autoregressive Conditional Heteroskedastic (GARCH) model to propose the Autoregressive Conditional Duration (ACD) model and show that the model can successfully describe the evolution of time durations for (heavily traded) stocks.

Let t_i be the time, measured with respect to some origin, of the ith event of interest, with t_0 being the starting time. The ith duration is defined as:

$$x_i = t_i - t_{i-1}, \qquad i = 1, 2, ...$$

We assume all $x_i > 0$ by ignoring the zero durations. Let $\psi_i = E(x_i \mid F_{i-1})$ be the conditional expectation of the adjusted duration between the $(i-1)th$ and ith trades, where F_{i-1} is the information set available at the $(i-1)th$ trade. Thus, the ψ_i is the expected adjusted duration given F_{i-1}. The ACD model has the following form:

$$x_i = \psi_i \epsilon_i \tag{2.2.4}$$

where ϵ_i is a sequence of Independent and Identically Distributed (IID) non-negative random variables with $E(\epsilon_i) = 1$. According to Engle and Russel (1998), ϵ_i follows a standard exponential or a standardized Weibull distribution and ψ_i assumes the following form:

$$\psi_i = \alpha_0 + \sum_{j=1}^{p} \alpha_j x_{i-j} + \sum_{v=1}^{q} \beta_v \psi_{i-v} \tag{2.2.5}$$

where p and q are non-negative integers and α_j and β_v are constant coefficients. Since x_i is positive, it is common to assume that $\alpha_0 > 0$, $\alpha_j \geq 0$ and $\beta_v \geq 0$ for $j \in \{1, 2, ..., p\}$ and $v \in \{1, 2, ..., q\}$. To fulfill the stationarity assumptions, the roots of the characteristic polynomials of (2.2.5) should lie outside the unit circle. When the distribution of ϵ_i is exponential, the resulting model is called Exponential ACD model, i.e., $EACD(p, q)$ model. Similarly, if ϵ_i follows a Weibull distribution, the model is called Weibull ACD, i.e., $WACD(p, q)$ model.

The parameters of ACD model can be estimated by using the conditional maximum likelihood method. Forecasts from an ACD model can be obtained using a procedures similar to that of a Autoregressive Moving Average model of order p and q, i.e., $ARMA(p, q)$ model. For example, for $EACD(1, 1)$ model the forecasting formula will take the following form for h-step-ahead forecasting horizon:

$$x_i(h) = \frac{\alpha_0[1 - (\alpha_1 + \beta_1)^{h-1}]}{(1 - \alpha_1 - \beta_1)} + (\alpha_1 + \beta_1)^{h-1} x_i(1)$$

where $x_i(1)$ is the 1-step-ahead forecast of an ACD model, i.e., 1-step-ahead conditional expected duration (ψ_{i+1}). If the fitted ACD model is adequate, then the standardized residuals $\hat{\epsilon}_i = \frac{x_i}{\hat{\psi}_i}$ should behave as an IID sequence of random variables with the assumed distribution. The Autocorrelation Function (ACF) can be drawn to visualize the dependence structure of $\hat{\epsilon}_i$.

2.2.5 Heterogeneous Threshold Autoregressive (HTAR) Model

Heterogeneous Autoregressive (HAR) model, introduced by Corsi (2009), is a simple cascades long-memory model used to capture few stylized facts observed in the high-frequency financial data. This model considers the volatility components defined over different time periods. It has attracted much interest among researchers ever since its introduction and has found many applications in modeling the financial data. Initially, this model served the purpose of modeling and predicting realized volatility data but it can be used for any type of data possessing fat tails and long-memory properties.

A number of studies documenting asymmetric effects in volatility can be found in the literature. Goetzmann *et al.* (2001) found evidence of asymmetry with respect to sign effects in volatility as far back as 1857 for the NYSE. They reported that unexpected negative shocks in monthly returns of the NYSE from 1857 to 1925 increased volatility almost twice as much as do equivalent positive shocks in returns of a similar magnitude. Similar results were also reported by Schwert (1990). McAleer and Medeiros (2008b) argued that most volatility models have been unable to simultaneously describe both nonlinear effects and long-memory and developed multiple regime smoothed transition heterogeneous autoregressive model. They applied this model to several Dow Jones Industrial Average index stocks using transaction level data from the Trades and Quotes database. They further argued that neglected changes in levels or persistence induce estimated high persistence, which has often been called "spurious" high persistence. It is difficult to find realized volatility models that simultaneously consider the effects of nonlinearity and long-memory in the recent literature. By taking inspiration from Threshold Autoregressive (TAR) model, Li and Li (1996) proposed the Double Threshold ARCH (DTARCH) model. Liu *et al.* (1997) generalized the model and proposed the Double Threshold GARCH (DTGARCH) process to model both the conditional mean and conditional variance as threshold processes. Baillie *et al.* (1996) proposed the Fractionally Integrated GARCH (FIGARCH) model as a viable alternative to model long-range dependence in volatility. Thus these studies leads us to extend the HAR model to an asymmetric long-memory volatility model.

In this study, we propose a simple cascade model that merges long-memory and nonlinearities. The new specification is a regime generalization of the Heterogeneous Autoregression (HAR) model suggested by Corsi (2009). The new model is called Heterogeneous Threshold Autoregressive (HTAR) model and it considers asymmetric and long-memory effect in time series. HTAR model constitutes a HAR model in each regime. When in HTAR model the regime switching variable is the lag values of the dependent variable itself, the resultant model is called Heterogeneous Self-Exciting Threshold Autoregressive (HSETAR) model. The HTAR model is threshold autoregressive type model with some restrictions on regressors. This model is globally nonlinear and piecewise linear. As HTAR model consists of an restricted AR(22) type model in each regime, so it has the same properties of stationarity in each regime as an AR model, i.e., the characteristic roots of the AR polynomial should lie outside the unit circle. According to our knowledge, this is the first study that extends the HAR model to HTAR model to simultaneously describe the long-range dependence and asymmetries in time series dynamics.

Denoting realized volatility by a time series y_t, the resultant two regime Heterogeneous Threshold Autoregressive (HTAR) model will have the following form:

$$\begin{aligned} y_t &= \left(c_1 + \alpha^{(d)} y_{t-1}^{(d)} + \alpha^{(w)} y_{t-1}^{(w)} + \alpha^{(m)} y_{t-1}^{(m)}\right)(y_{t-d} \leq r) \\ &+ \left(c_2 + \beta^{(d)} y_{t-1}^{(d)} + \beta^{(w)} y_{t-1}^{(w)} + \beta^{(m)} y_{t-1}^{(m)}\right)(y_{t-d} > r) + e_t \end{aligned} \tag{2.2.6}$$

where $\alpha^{(d)}$, $\alpha^{(w)}$ and $\alpha^{(m)}$ are the parameters in low regime corresponding to daily, weekly and monthly observed aggregated volatility components; similarly, $\beta^{(d)}$, $\beta^{(w)}$, and $\beta^{(m)}$ are the parameters of HTAR model in high regime corresponding to daily, weekly and monthly observed aggregated volatility components. The parameters c_1 and c_2 are the constant terms in lower and upper regime respectively, d is the threshold delay parameter, and r is the threshold value corresponding to the threshold variable y_{t-d}.

The realized volatilities observed over different time horizons is defined as:

$$y_{t-1}^{(w)} = \frac{(y_{t-1} + y_{t-2} + y_{t-3} + y_{t-4} + y_{t-5})}{5}$$

$$y_{t-1}^{(m)} = \frac{(y_{t-1} + y_{t-2} + ... + y_{t-22})}{22}$$

where $y_{t-1}^{(d)}$, $y_{t-1}^{(w)}$, and $y_{t-1}^{(m)}$ are respectively the daily, weekly, and monthly observed realized volatilities. We assumed homoskedastic variance in a two regimes HTAR model and can be heteroskedastic. The HTAR model can easily be extended for more than two regimes. A key step in specifying a HTAR model for a given time series is the identification of the threshold variable (y_{t-d}), the threshold delay parameter (d), and the threshold value (r). The choice of d and y_{t-d} in HTAR setup is relatively simpler than for Threshold Autoregressive (TAR) model. According to the HTAR model specifications, the possible threshold variable could be $y_{t-1}^{(d)}$, $y_{t-1}^{(w)}$, and $y_{t-1}^{(m)}$ or any lagged time series up to maximum 22 lags. Thus, HTAR model is useful in a situation where the underlying time series possesses regime-switching behavior with heterogeneous components. HTAR model allow a researcher to model a time series in a more parsimonious and flexible way that simultaneously contain long-memory, fat tails, and asymmetry properties.

Estimation procedure of the HTAR model is similar to estimating a Threshold Autoregressive (TAR) model, discussed in this chapter using the grid search algorithm of Chan (1993). For given values of threshold delay and thresholds, the HTAR model is linear, and standard OLS can be used to estimate its parameters in each regime. The first and last 15% observations of the threshold variable are excluded from grid search to ensure an adequate number of observations on each side of the threshold. One can estimate the HTAR model with more than two regimes for different potential threshold variables and delay parameters and choose the combination that yields the smallest residual sum of squares (RSS) or AIC. Forecasts from the HTAR model can be generated by using the bootstrapped method, already discussed in this chapter, to generate the forecasts from threshold models.

2.2.6 Threshold Autoregressive (TAR) Model

It is well known that various economic and other time series variables exhibit nonlinear behavior. In this section, we discuss linear and nonlinear time series models, their estimation procedure, and other related issues. TAR model is suited for a time series that posses to regime switching behavior. An example of nonlinear time series includes Gross Domestic Product (GDP), where economic expansion and recession exhibit nonlinear behavior; similarly, unemployment rate is likely to rise more rapidly in recession and then decline slowly in expansion in the long run. Threshold Autoregressive (TAR) models are perhaps the simplest generalization of linear autoregressions. The Threshold Autoregressive (TAR) model was initially proposed by Tong (1978) and Tong and Lim (1980), and later discussed extensively in Tong (1990). The general form of $m-$regime TAR(p) model can be defined as follows by defining a set of $m+1$ thresholds, c_0, $c_1, \ldots, c_m$, such that $-\infty = c_0 < c_1 < \ldots < c_{k-1} < c_m = \infty$

$$y_t = \phi_{0,j} + \phi_{1,j} y_{t-1} + \phi_{2,j} y_{t-2} + \ldots + \phi_{p,j} y_{t-p} + e_t \qquad {\scriptstyle if\, c_{j-1} < y_{t-d} < c_j}$$

for $j = 1, 2, ..., m$, where, y_{t-d} is threshold variable with threshold delay parameter d. The transition variable can also be an exogenous variable, i.e., z_{t-d}.

Threshold Autoregressive (TAR) model captures the variation in nonlinear time series data by dividing it in different regimes and then modeling each regime using a linear autoregressive model. A number of regime switching models differ with respect to the regime switching variable z_t that evolves over time. When using TAR model, we assumed that the regime can be determined by a known threshold observable variable relative to the threshold value c. When the threshold variable z_t is taken to be lagged values of the time series y_t itself, i.e., $z_t = y_{t-d}$ for a certain positive integer d, the regime is determined by the time series itself and the resulting model is called Self-Exciting Threshold Autoregressive (SETAR) model.

Consider the following univariate two regime SETAR model for a stationary time series y_t:

$$y_t = \left[\alpha_0 + \sum_{i=1}^{p_1} \alpha_i y_{t-i}\right] I_t + (1 - I_t)\left[\beta_0 + \sum_{j=1}^{p_2} \beta_j y_{t-j}\right] + e_t \qquad (2.2.7)$$

where I_t is an indicator function defined as:

$$I_t = \begin{cases} 1 & \text{if } z_{t-d} \leq c \\ 0 & \text{if } z_{t-d} > c \end{cases}$$

In model (2.2.7), e_t is the innovation term is assumed to be independent and identically distributed (IID) with mean zero and constant variance. The autoregressive order of the model (2.2.7) is p_1 and p_2 in lower and upper regimes respectively, y_{t-d} is threshold variable by assuming that $z_{t-d} = y_{t-d}$,and d and c are the delay and threshold parameters respectively. Furthermore, the degree of persistence is $\sum_{i=1}^{p_1} \alpha_i$ when $y_{t-d} \leq c$, and $\sum_{j=1}^{p_2} \beta_j$ when $y_{t-d} > c$. A natural approach to modeling economic time series with nonlinear models seems to be to define different states of the world or regimes and to allow for the possibility that the dynamic behavior of economic variables depends on the regime that occurs at any given point in time (Frances and van Dijk, 2000). Regime switching means that the entire time series y_t is nonlinear, but we divide it in two linear parts that depend on the delay (d) and threshold (c)

parameters respectively. The regime dependent model (2.2.7) means that the mean, variance, and autocorrelation of time series varies in different regimes. To ensure the stationarity in TAR model, the characteristic roots of the characteristic polynomial of an AR polynomial in each regime should lie outside the unit circle. To achieve stationarity in TAR model, the characteristic roots of the characteristic polynomial should lie outside the unit circle.

2.2.7 Smooth Transition Autoregressive (STAR) Model

For some time series processes the assumption of a sharp threshold is not reasonable; rather, the shift in the process may be smooth. In such a situation, we applied the Smooth Transition Autoregressive (STAR) model that allows the autoregressive parameters to change slowly. The STAR model was originally suggested by Chan and Tong (1986) and subsequently developed by Teräsvirta and Anderson (1992) and Teräsvirta (1994). The replacement of the indicator function I_t in model (2.2.7) by a continuous function $\theta = G(y_{t-d}; \gamma, c)$, bounded between 0 to 1 provides a gradual transition between different regimes. The resultant model is called the Smooth Transition Autoregressive (STAR) model and is given by:

$$y_t = \left[\alpha_0 + \sum_{i=1}^{p_1} \alpha_i y_{t-i}\right](1-\theta) + \theta\left[\beta_0 + \sum_{j=1}^{p_2} \beta_j y_{t-j}\right] + e_t \tag{2.2.8}$$

A popular choice for the so called transition function $\theta = G(y_{t-d}; \gamma,\, c)$ is the logistic function and defined as:

$$\theta = G(y_{t-d}; \gamma, c) = \frac{1}{1 + exp(-\gamma(y_{t-d} - c))} \tag{2.2.9}$$

The resultant model is then called a Logistic Smooth Transition Autoregressive (LSTAR) model. The threshold variable y_{t-d} could be an external threshold variable, e.g., z_{t-d}. The parameter c in (2.2.9) can be interpreted as the threshold between the two regimes, in the sense that the logistic function changes monotonically from 0 to 1 as y_{t-d} increases, while $\theta = G(c;\, \gamma,\, c) = 0.5$. The smoothness parameter γ determines the smoothness of the change in the value of the logistic function, and thus the transition from one regime to the other. As $\gamma \to \infty$, and $y_{t-d} > c$, then $\theta = 1$, when $y_{t-d} \leq c$, $\theta = 0$, hence model (2.2.8) becomes a TAR (p) model. When $\gamma \to 0$ in (2.2.9), the equation (2.2.8) becomes an AR (p) model. In the limit, as $\gamma \to 0$ or ∞, the LSTAR model becomes an AR(p) model since the value of transition function is constant. For intermediate values of γ, the degree of autoregressive decay depends on the value of y_{t-d}. As $y_{t-d}\rightarrow-\infty$, $\theta \rightarrow 0$, so that the behavior of y_t is given by $\alpha_0 + \sum_{i=1}^{p_1} \alpha_i y_{t-i}$. Similarly, as $y_{t-1}\rightarrow+\infty$, $\theta \rightarrow 1$, so that the behavior of y_t is given by $(\beta_0 + \sum_{j=1}^{p_2} \beta_j y_{t-j})$. Thus, the intercept and autoregressive coefficients smoothly change between these two extremes as the value of y_{t-d} changes. The above two regimes LSTAR model can easily be extended to m regimes LSTAR model analogue to m regimes TAR model discussed previously. To assure the stationarity in STAR model, the characteristic roots of the characteristic polynomial of the AR polynomial in each regime should lie out side the unit circle.

For example, suppose that the time series y_t represents the growth rate of an output variable, and if the threshold variable is the growth rate in the previous period, y_{t-1}, and if $c = 0$, the model distinguishes between periods of positive and negative growth, that is, between expansions and contractions.

2.2.8 Threshold Vector Autoregressive (TVAR) Model

The Threshold Vector Autoregressive (TVAR) model is a multivariate extension of the univariate Threshold Autoregressive (TAR) model proposed by Tong (1978, 1983). A TVAR model parsimoniously capture the nonlinearities such as regime switching and asymmetry to shock responses. Threshold models work by splitting the time series endogenously into different regimes and in each regime the time series is described by a linear model. The stationary TVAR model is defined as follows:

$$Y_t = \sum_{i=1}^{q}(\mu_i + \sum_{j=1}^{p} \phi_{ij} Y_{t-j} + \Gamma_i X_t + \epsilon_{it}) I(c_{i-1} < s_{t-d} \leq c_i) \tag{2.2.10}$$

where, Y_t and ϵ_{it}, $i = 1, ..., p$, are stochastic $m \times 1$ vectors, and μ_i is an $m \times 1$ vector of intercepts in ith regime. Furthermore, ϕ_{ij} are $m \times m$ coefficient matrices, $i = 1, ..., q$, and $j = 1, ..., p$, Γ_i are $m \times n$ coefficient matrices,while c_0 and c_q are the threshold values respectively in first and qth regime against the threshold variable s_{t-d}, with threshold delay parameter d. The indicator function $I(.)$ is equal to 1 if the process is in ith regime otherwise it is equal to zero. In TVAR model, X_t is an $n \times 1$ vector of exogenous variables. The errors ϵ_{it} are serially uncorrelated with mean 0 and positive definite covariance matrix Σ_i, $i = 1, ..., q$. The above TVAR model has q regimes with lag length p in each one. For each equation in TVAR model, the threshold variable s_{t-d} may be lag values of target variable itself or lag values of external threshold variable as suggested by Tena (2009). The threshold delay parameter d is mostly set to one in TVAR model. A systematic modeling strategy for TVAR models with applications can be found in Tsay (1998). The Estimation of TVAR model can be carried out through the grid search algorithmic (see Section 2.4, estimation of TAR model) for the threshold value and then using the OLS method to estimate a separate VAR model in each regime. The lag length of the endogenous and exogenous variables, p, is determined by the usual information criteria (BIC or AIC). Forecasts from TVAR model can be generated by using the Monte Carlo (MC) or Bootstrapped (BS) method as already discussed for the univariate TAR model. To achieve stationarity in TVAR model, the characteristic roots of the characteristic polynomial of VAR model in each regime should lie outside the unit circle.

2.2.9 Artificial Neural Network (ANN) Model

In recent years, neural networks have found their way into financial analysis and other sciences for the purpose of modeling and forecasting. The Artificial Neural Network (ANN) model is simply a nonlinear modeling procedure except that the parameters interpretation is difficult to interpret than other types of nonlinear models. The structure of an ANN model somewhat resembles the human body, and the model is therefore called an Artificial Neural Network (ANN). The ANN has achieved great attention among researchers due to the fact that these models are able to approximate almost any nonlinear function arbitrarily close. Hence,when applied to a time series that truly represent a nonlinear dynamic relationships, the ANN will detect these and provide superior in-sample fit compared to linear time series models without any need of parametric specification of nonlinear time series model. An estimated ANN model does not provide any information on which type of parametric time series models might be suitable to describe the detected nonlinear dynamics. The mostly reported drawback of AANs model is the parameters in the estimated model are difficult, if not impossible, to interpret. For this reason and also because it is usually difficult to meaningfully describe the parameter values, ANNs often are considered as 'black box' models and constructed mainly for the purpose of pattern recognition and forecasting. However, the superior in-sample fit that can be achieved provides no guarantee that an ANN model performs well in out-of-sample forecasting. A simple ANN model takes the following form:

$$y_t = \phi_0 + \tilde{x}_t'\phi + \sum_{j=1}^{q} \beta_j G(x_t'\gamma_j) + e_t$$

More elaborately,

$$y_t = \phi_0 + \phi_1 y_{t-1} + ... + \phi_p y_{t-p} + \sum_{j=1}^{q} \beta_j G(\gamma_{0,j} + \gamma_{1,j} y_{t-1} + ... + \gamma_{p,j} y_{t-p}) + e_t \quad (2.2.11)$$

where $x_t' = (y_{t-1}, y_{t-2}, ..., y_{t-p})$. In equation (2.2.11) the term $\tilde{x}_t'\phi$ can be thought of as representing the linear part of the relationship between y_t and $\tilde{x}_t'$, while the logistic components $G(x_t'\gamma_j)$ measure the amount of nonlinearity that is present. Equation (2.2.11) is referred to as $ANN(k,\, q)$ model to identify the network with k input variables (input layer) $y_{t-1}, y_{t-2}, ..., y_{t-p}$, q logistic components (hidden layer) $G(x_t'\gamma_j)$ and one output variable (output layer or output activation function) y_t. The coefficients $\beta_1, \beta_2, ..., \beta_q$ are the connection weights, connecting the nodes $G(x_t'\gamma_j)$ together into a model. The name 'hidden layer' arises from the fact that it is not directly observed. By increasing the number of logistic components in equation (2.2.11), we can capture any type of nonlinear relationships that might exist between the input and output variables. In the hidden layer, the linear combinations $\tilde{x}_t'\gamma_j$ are formed and transformed into a value between 0 and 1 by the activation functions $G(.)$. Finally, these are multiplied by weights β_j to produce the output y_t. This type of ANN is usually referred to as the single hidden layer feedforward network model, because it contains only one hidden layer, and the information flows only in one directions, from inputs to outputs. The estimation procedure of the ANN model is briefly discussed in the next section.

2.2.10 Additive Autoregressive (AAR) Model

Additive autoregressive model is a nonparametric time series model. Generalized Additive Models (GAMs) received more attention following the work of Hastie and Tibshirani (Generalized Additive Models, 1990). These models assume that the mean of the dependent variable depends on an additive predictor through a nonlinear link function. Therefore making this model more flexible than a linear model. Consider a nonparametric AAR model of the following form:

$$y_t = \mu + \sum_{i=1}^{p} s_i(y_{t-i}) + e_t \quad (2.2.12)$$

where s_i are smooth functions of the lagged time series values represented by penalized cubic regression splines. Spline functions are piecewise polynomials, with the polynomial pieces joining in the so called knots. They are estimated, along with their degree of smoothing by using the Generalized Cross Validation (GCV) score of the whole model. AAR models are estimated by the usual Iteratively Re-weighted Least Squares (IRLS) scheme, except that the least squares problem at each iterate is replaced by a penalized least squares problem, in which the set of smoothing parameters must be estimated alongside the other model parameters: the smoothing parameters are chosen by GCV score (Wood, 2001). The information criteria, AIC or BIC can be used for the lag lag length selection in AAR model. The forecasting procedure for an AAR model is similar as for SETAR or LSTAR models (discussed in the next section). A detailed discussion and estimation procedures of AAR model can be found in Hastie and Tibshirani (Generalized Additive Models, 1990).

2.3 Lag Orders Identification

The important question that arises when estimating nonlinear autoregressive model is the determination of lag orders p_1 and p_2 in general two regimes model, like considering the SETAR and LSTAR models. Frances and van Dijk (2000) pointed out that nonlinear time series may have zero autocorrelation at all lags. In such case, the chances of underestimating a nonlinear time series model are very high; furthermore the autocorrelation structures of a simple nonlinear time series model are quite complicated and can be captured only by a higher-order autoregressive model. They argued that usual model selection criteria, such as AIC and BIC, are not suitable in the same way as used in linear models selection due to the unequal observations in different regimes. Tong (1990) defined an alternative AIC for a 2-regime SETAR model as the sum of the AICs for the AR models in two regimes, that is:

$$AIC(p_1,\ p_2) = n_1 ln(\hat{\sigma}_1^2) + n_2 ln(\hat{\sigma}_2^2) + 2(p_1+1) + 2(p_2+1) \tag{2.3.1}$$

where n_j, $j = 1, 2$, is the number of observations in the jth regime, and $\hat{\sigma}_j^2$, for $j = 1, 2$, is the variance of the residuals in the jth regime. Even if we assume equal variances across regimes, the estimates $\hat{\sigma}_1^2$ and $\hat{\sigma}_2^2$ can be different. The BIC for a SETAR model is defined in a similar way as:

$$BIC(p_1,\ p_2) = n_1 ln(\hat{\sigma}_1^2) + n_2 ln(\hat{\sigma}_2^2) + (p_1+1) ln(n_1) + (p_2+1) ln(n_2) \tag{2.3.2}$$

BIC takes into account the number of observations in each regime. Lag orders in the selected regimes are those for which the information criterion is minimized. Setting $n_2 = p_2 + 1 = 0$ in equations (2.3.1) and (2.3.2) yields the AIC and BIC criteria for the lag order selection in a linear $AR(p)$ model.

2.4 Estimation of Nonlinear Time Series Models

As we are mostly interested in regime switching models, here we only discuss the estimation procedures of TAR, STAR or LSTAR, and ANN models. To make the estimation simple, the discussion is limited to only two regime threshold models, and it is assumed that the lag orders in both regimes are equal and can easily be extended to more than two regimes and for different lag length. For detailed discussions, see Tong (1990) and Hansen (1997, 2000) for the SETAR model and Teräsvirta (1994, 1998) for the STAR model.

2.4.1 Estimation of TAR Models

The estimation procedure of the nonlinear time series models is not easy in comparison to linear models. The parameters of interest in a two regime SETAR model are the lag coefficients, i.e., $\sum_{i=1}^{p_1} \alpha_i$ and $\sum_{j=1}^{p_2} \beta_j$, threshold value γ, threshold delay d, and residual variance $\hat{\sigma}^2$. They can be conveniently estimated by sequential conditional least square. We assumed that the threshold parameter $d = 1$, and the number of lags in both regimes is equal, i.e, $p_1 = p_2$, while estimating the SETAR model. Under the auxiliary assumption that e_t is IID and most probably follows a $N(0, \sigma^2)$, Least Square (LS) is equivalent to the maximum likelihood estimation. To represent the SETAR model in an alternative way for estimation purposes, let

$$y_t = \left[\alpha + \sum_{i=1}^{p_1} \alpha_i y_{t-i}\right] I(z_{t-d} \leq \gamma) + I(z_{t-d} > \gamma) \left[\beta_0 + \sum_{j=1}^{p_2} \beta_j y_{t-j}\right] + e_t \qquad (2.4.1)$$

$$x_t = (1 \ \ y_{t-1} \ \ \ldots y_{t-p})'$$

$$x_t(\gamma) = (x_t' I(z_{t-d} \leq \gamma) \quad x_t' I(z_{t-d} > \gamma))'$$

so above model can be written as:

$$y_t = x_t' \alpha I(z_{t-d} \leq \gamma) + x_t' \beta I(z_{t-d} > \gamma) + e_t \qquad (2.4.2)$$

where $\alpha = (\alpha_0 \, \alpha_1 \ \ldots \alpha_{p_1})'$ and $\beta = (\beta_0 \, \beta_1 \ldots \beta_{p_2})'$ and $\psi = (\alpha' \, \beta')'$

and

$$y_t = x_t(\gamma)'\psi + e_t \tag{2.4.3}$$

The Hansen (1997) sequential Conditional Least Square (CLS) method is used to estimate these parameters.

The CLS estimate of ψ is:

$$\hat{\psi}(\gamma) = \left(\sum_{t=1}^{n} x_t(\gamma)x_t(\gamma)'\right)^{-1}\left(\sum_{t=1}^{n} x_t(\gamma)y_t\right) \tag{2.4.4}$$

where the residuals of the model is $\hat{e_t} = y_t$-$x_t(\gamma)'\hat{\psi}(\gamma)$ and residuals variance is:

$$\hat{\sigma_n}^2(\gamma) = \frac{1}{n}\sum_{t=1}^{n} \hat{\sigma_t}^2(\gamma) \tag{2.4.5}$$

The CLS estimate of γ is the one that produces the lowest value of (2.4.6):

$$\hat{\gamma} = \underset{\gamma\epsilon C}{argmin}\, \hat{\sigma}^2(\gamma) \tag{2.4.6}$$

where C denotes the set of all possible threshold values. The final estimates of the autoregressive parameters are given by $\hat{\psi} = \hat{\psi}(\hat{\gamma})$, and the estimated residual variance is $\hat{\sigma}^2 = \hat{\sigma}^2(\hat{\gamma})$.

The threshold value should be chosen in such a way,that each regime contains enough observations to get precise estimates of the autoregressive coefficients in each regime. A popular choice for C requires that each regime contains at least pre-specified fraction π_0 of the observations, that is:

$$C = \{\gamma \mid y([\pi_0(n-1)]) \leq c \leq y([(1-\pi_0)(n-1)])\} \tag{2.4.7}$$

where $y(0)$, $y(1), ..., y(n-1)$ denote the arranged threshold variable z_{t-d} or y_{t-d} values in ascending order and n denotes the sample size. A popular choice of π_0 is 0.15.

2.4.2 Choosing the Threshold Variable and Delay Parameter

While estimating a TAR model, we assumed that the threshold variable z_{t-d} is known and equal to the lag values of time series itself, i.e., y_{t-d}. In practice, the suitable threshold variable is often unknown and we have to determine this variable. In such case, we estimate the different TAR models with varying threshold variables and choose one that provides best fit to the observed data. Chen (1995) describes number of methods for selecting the threshold variable. In context of SETAR model for a given time series y_t, the threshold variable is taken as its own lag value y_{t-d} for some positive integer d called delay parameter providing that $d\epsilon\{1, 2, \ldots, d^*\}$, where $d^*$is the upper bound. The estimated optimal value of d is chosen in such a way that it provides the minimum residuals variance. Thus, we augmented equation (2.4.6) with d to search for the optimal value of threshold delay (d) and hence equation (2.4.6) becomes:

$$(\hat{\gamma}, \hat{d}) = \underset{\gamma\epsilon C,\, d\epsilon D}{argmin} \hat{\sigma}^2(\gamma,\ d) \tag{2.4.8}$$

where $D = \{1, 2, ..., d^*\}$. Thus, we have nd^*regression to search for a possible threshold value and a threshold delay. In equation (2.4.8), $\hat{\sigma}^2(\gamma,\ d)$, is the estimated residual variance, now depends on d and γ as well. Since the parameter space for d is discrete, the estimate $\hat{d}$ is super-consistent.

2.4.3 Estimation of STAR models

We considered a STAR model where the lag order might vary in two regimes. Consider the following STAR model:

$$y_t = \left[\alpha_0 + \sum_{i=1}^{p_1} \alpha_i y_{t-i}\right](1-\theta) + \theta\left[\beta_0 + \sum_{j=1}^{p_2} \beta_j y_{t-j}\right] + e_t \tag{2.4.9}$$

Estimation of the STAR model is carried out by using the Nonlinear Least Square (NLS) algorithm and the parameters $\psi = (\alpha',\ \beta, \gamma,\ c)'$ can be estimated as:

$$\hat{\psi} = \underset{\psi}{argmin} Q_n(\psi) = \underset{\psi}{argmin} \sum_{t=1}^{n} [y_t - F(x_t; \psi)]^2 \tag{2.4.10}$$

where $F(x_t; \psi)$ is the deterministic part (skeleton) of the model, that is:

$$F(x_t;\psi) = \alpha' x_t(1 - G(y_{t-1};\gamma, c)) + \beta' x_t G(y_{t-1};\gamma, c)$$

By imposing the additional assumption that the error term e_t is normally distributed, NLS estimates are equal to maximum likelihood estimates. Otherwise, the NLS estimates are the quasi-maximum likelihood estimates. Under some regularity conditions, the NLS estimates are consistent and asymptotically normal, that is:

$$\sqrt{T}(\hat{\psi} - \psi_0) \to N(0,\ C) \tag{2.4.11}$$

where ψ_0 denotes the true parameter values and C is the covariance matrix of $\hat{\psi}$. The estimation can be performed using any nonlinear optimization procedure. The starting values play important role in optimization algorithm, concentrating the sum of squares function and the estimate of the smoothness parameter γ in the transition function.

Choosing good starting values provides the best estimates when the optimization algorithm is applied. By fixing the value of the threshold parameter (c) and the smoothness parameter (γ), the STAR model becomes linear in the autoregressive parameters α and β similar to SETAR model. The conditional estimates $\psi = (\alpha'\ \beta')'$ that depend on the values of c and γ can be obtained by OLS as:

$$\hat{\psi}(\gamma,\ c) = \left(\sum_{t=1}^{n} x_t(\gamma,\ c) x_t(\gamma,\ c)'\right)^{-1} \left(\sum_{t=1}^{n} x_t(\gamma,\ c) y_t\right) \tag{2.4.12}$$

where $x_t(\gamma,\ c) = (x_t'(1 - G(y_{t-1};\gamma, c))_t,\ \ x_t' G(y_{t-1};\gamma, c))'$ and the notation $\psi(\gamma,\ c)$ represent the conditionally dependent estimates of ψ on the values of γ and c. The corresponding residuals can be calculated $\hat{e_t} = y_t - \hat{\psi}(\gamma,\ c)' x_t(\gamma,\ c)$ with associated variance $\hat{\sigma_n}^2(\gamma, c) = \frac{1}{n}\sum_{t=1}^{n} \hat{e_t}^2(\gamma, c)$. A convenient method to obtain sensible starting values for the nonlinear optimization algorithm then is to perform a two-dimensional grid search over γ and c and select those parameter estimates which render the smallest estimate for the residual variance $\hat{\sigma_n}^2(\gamma,\ c)$ (Frances and van Dijk, 2000).

2.4.4 Concentrating the Sum of Squares Function

Leybourne *et al.* (1998) suggests another method to ease the estimation problem and is concentrates on the sum of squares function. For a given values of γ and c, STAR model is linear in autoregressive parameters, and the sum of the square function $Q_n(\psi)$

can be concentrated with respect to α and β as:

$$Q_n(\gamma,\, c) = \sum_{t=1}^{n}(y_t - \psi(\gamma,\, c)' x_t(\gamma,\, c))^2 \tag{2.4.13}$$

By minimizing the sum of squares in the above equation with respect to parameters γ and c and this will considerable reduces the dimensionality of NLS estimation problem.

Regarding the precise estimate of the smoothness parameter γ, we used many observations in the immediate neighborhood of the threshold c. Because for large values of γ, the shape of the logistic function does not significantly change.

2.4.5 Estimation of ANN models

The parameter estimates in the $ANN(p,\, q)$ model can be estimated by minimizing the residual sum of squares function:

$$Q_n(\theta) = \sum_{t=1}^{n}[y_t - F(x_t;\theta)]^2 \tag{2.4.14}$$

where

$F(x_t;\theta) = \tilde{x}_t'\phi + \sum_{j=1}^{q}\beta_j G(x_t'\gamma_j)$

the parameter vector θ consists of $k+1+q(k+2)$ parameters in ϕ, $\beta_1, \beta_2, ..., \beta_q$, $\gamma_1, \gamma_2, ..., \gamma_q$. Any conventional nonlinear least squares algorithm can be applied to obtain the estimate $\hat{\theta}_n$. As it is a well-known fact that the ANN model approximates the true data generating process fairly well, the chance of its misspecification is greater. The theory of least squares estimation for misspecified models is by now well-developed (Pötscher and Prucha, 1997) and can be applied to obtain the properties of the nonlinear least squares estimator of θ in (2.4.14), denoted by $\hat{\theta}_n$. Under general conditions $\hat{\theta}_n$ converges to θ^*, defined as:

$$\theta^* = \underset{\theta}{argmin} E([y_t - F(x_t;\theta)]^2)$$

as the sample size n increases without bound. Furthermore, the normalized estimator $\sqrt{n}(\hat{\theta}_n$-$\theta^*)$ converges to a multivariate normal distribution with mean zero and a covariance matrix that can be estimated consistently (Frances and van Dijk, 2000).

To avoid the local minima problems in (2.4.14), it is better to estimate the ANN several times using different starting values by using the simplex algorithm to achieve the global minimum value of $Q_n(\theta)$.

Model (2.2.11) contains $p+1+q(p+2)$ parameters to be estimated. Implementing an $ANN(p,\ q)$ model requires several decisions to be made:

i. Choosing the activation function $G(\,\cdot\,)$.

ii. Choosing the number of hidden units q.

iii. Choosing the number of lags p to use as input variables.

Very often, the choice of the activation function is not considered to be a decision problem. The logistic function is used almost invariably, although other choices, such as the hyperbolic tangent function G(z) = tanh(z), are sometimes applied as well. There are various different strategies one can follow to decide upon the appropriate values of q and p in ANN model. The first is to estimate all possible ANN models with $p \in \{1, 2, ..., p^*\}$ and $q \in \{0, 1, 2, ..., q^*\}$, for certain pre-set values of p^* and q^*, and to use a model selection criterion such as the AIC and BIC to choose the appropriate lag length. Forecasting with ANN models is similar to forecasting with SETAR or LSTAR models (Frances and van Dijk, 2000).

2.5 Testing Nonlinearity

It is usually recommended to perform a test of linearity against nonlinearity before applying a possibly complex nonlinear model. Pretesting for nonlinearity can help protect us from overfitting the data. Therefore, two nonlinearity tests, i.e., Tsay (1986) and Keenan (1985) tests, are discussed in this section.

Keenan and Tsay tests can be interpreted as as Lagrange multiplier tests for specific nonlinear alternatives. Keenan (1985) test of nonlinearity is similar to Tukey (1949) one degree of freedom for nonadditivity test.

2.5.1 Keenan Test

Keenan's (1985) test is motivated by approximating a nonlinear stationary time series by a second order Volterra expansion (Wiener, 1958):

$$y_t = c + \sum_{i=-\infty}^{\infty} \theta_i \varepsilon_{t-i} + \sum_{i=-\infty,}^{\infty} \sum_{j=-\infty}^{\infty} \theta_{ij} \varepsilon_{t-i} \varepsilon_{t-j} \tag{2.5.1}$$

where $(\epsilon_t, \ -\infty < t < \infty)$ is a sequence of independent and identically distributed zero mean random variables. The process y_t becomes linear if the double sum in (2.5.1) is set to zero. Thus, one can test the linearity of a time series by testing whether the double sum in (2.5.1) is zero or not. Suppose we have an observed time series $y_1, ..., y_n$, then Keenan test can be implemented as follows:

i. Regress y_t on $y_{t-1,}...., y_{t-m}$, including an intercept term, where m is the selected lag length then calculate the fitted values $\hat{y}_t$ and the residuals $\hat{e}_t$, for $t = m+1, ..., n$ where n is the sample size and calculate the residual sum of squares, *i.e.*, $RSS = \sum_{t=m+1}^{n} \hat{e}_t^2$.

ii. Regress $\hat{y}_t^2$ on $y_{t-1}, ..., y_{t-m}$, including an intercept term, and calculate the residuals $\hat{\xi}_t$ for $t = m+1, \ldots, n$.

iii. Regress $\hat{e}_t$ on the residuals $\hat{\xi}_t$ without an intercept for $t = m+1, \ldots, n$, and Keenan's test statistic, denoted by $\hat{F}$, is obtained by multiplying $(n-2m-2)/(n-m-1)$ to the $F - statistic$ for testing that the last regression function is identically zero. Specifically, let

$$\delta = \delta_0 \sqrt{\sum_{t=m+1}^{n} \hat{\xi}_t^{\,2}}$$

where δ_0 is the regression coefficient. The test statics will take the following form:

$$\hat{F} = \delta^2 \frac{(n-2m-2)}{RSS - \delta^2} \tag{2.5.2}$$

Under the null hypothesis, this $\hat{F}$ statistic follows $F - distribution$ with degrees of freedom 1 and $n - 2m - 2$.

2.5.2 Tsay Test

This test is a more powerful extension of the Keenan's approach by considering more general nonlinear alternative hypothesis.

A stationary time series y_t can be written as follows:

$$
\begin{aligned}
y_t &= c + \sum_{i=-\infty}^{\infty} \theta_i \varepsilon_{t-i} + \sum_{i=-\infty,}^{\infty} \sum_{j=-\infty}^{\infty} \theta_{ij} \varepsilon_{t-i} \varepsilon_{t-j} \\
&\quad + \sum_{i=-\infty,}^{\infty} \sum_{j=-\infty,}^{\infty} \sum_{k=-\infty,}^{\infty} \theta_{ijk} \varepsilon_{t-i} \varepsilon_{t-j} \varepsilon_{t-k}
\end{aligned} \tag{2.5.3}
$$

where $(\epsilon_t,\ -\infty < t < \infty)$ is a sequence of independent and identically distributed (IID) zero mean random variables. If any of θ_{ij} or θ_{ijk} are nonzero, then it provides evidence of nonlinearity of a time series y_t. Tsay (1986) test consists of the following steps.

i. Regress y_t on $y_{t-1}, \ldots., y_{t-m}$, by using OLS and obtain the residuals $\hat{e}_t$ for $t = m+1, ..., n$, where m and n respectively represent the pre-specified lag length and sample size. The regression model is denoted by

$$y_t = W_t \phi + e_t$$

where $W_t = (1,\ y_{t-1}, \ldots., y_{t-m})$ and $\phi = (\phi_0,\ \phi_1, ..., \phi_m)^t$.

ii. Regress the vector Z_t on $(1,\ y_{t-1}, \ldots., y_{t-m})$ and obtain the residual vector X_t for $t = m+1, ..., n$. Here, the multivariate regression model takes the following form:

$$Z_t = W_t H + X_t$$

where Z_t is an $M = \frac{1}{2}(m(m+1)$ dimensional vector defined by $Z_t^T = vech(U_t^T U_t)$ with $U_t = (y_t, \ldots., y_{t-m})$ and vech denoting the half-stacking vector.

iii. Regress $\hat{e}_t$ on $\hat{X}_t$ as $\hat{e}_t = X_t \hat{\beta} + \epsilon_t$ and define the following test statistic:

$$\hat{F} = \left\{ \frac{((\sum_{t=m+1}^{t} \hat{X}_t \hat{e}_t)(\sum_{t=m+1}^{t} \hat{X}_t^{\ T} \hat{X}_t)^{-1}(\sum_{t=m+1}^{t} \hat{X}_t \hat{e}_t))/M}{(\sum_{t=m+1}^{t} \hat{e}_t^2/(n-m-M-1))} \right\} \tag{2.5.4}$$

Follow the $F-$distribution with degrees of freedom $\frac{1}{2}(m(m+1)$ and $n - \frac{1}{2}m(m+3)-1$.

2.5.3 Testing for Threshold Type Nonlinearity

Before applying TAR model to time series data, it is important to test the significance of TAR model relative to linear autoregressive (AR) model. Consider the null hypothesis that the lag coefficients in two regime SETAR models are equal, i.e., $H_0 : \alpha_i = \beta_j$ for all $i = 1, 2, \ldots, p_1$ and $j = 1, 2, ..., p_2$. As under the null hypothesis the threshold value not identified so the usual F-statistic will not work. In the following paragraph the methodology provided by Hansen (1996) is presented for testing the null hypothesis of linearity.

Davies (1977, 1987) and Andrews-Ploberger (1994) showed that the following F-statistic has near optimal power against different alternative hypothesis by assuming that the errors are independent and identically (IID) distributed. The F-statistic is as follow:

$$F_n = n\left(\frac{\tilde{\sigma}_n^2 - \hat{\sigma}_n^2}{\hat{\sigma}_n^2}\right) \tag{2.5.5}$$

In (2.5.5), $\hat{\sigma}_n^2$ is the residual variance under alternative hypothesis and $\tilde{\sigma}_n^2$ is the residual variance under the null hypothesis and is defined as:

$$\tilde{\sigma}_n^2 = \frac{1}{n}\sum_{t=1}^{n}(y_t - x_t'\tilde{\alpha})^2$$

and

$$\tilde{\alpha} = \left(\sum_{t=1}^{n} x_t x_t'\right)^{-1}\left(\sum_{t=1}^{n} x_t y_t\right)$$

is the OLS estimate of α under the null hypothesis. Since F_n is monotonic function of $\hat{\sigma}_n^2$, it is easy to see that:

$$F_n = \underset{\gamma\epsilon C}{sup} F_n(\gamma)$$

where

$$F_n(\gamma) = n\left(\frac{\tilde{\sigma}_n^2 - \hat{\sigma}_n^2(\gamma)}{\hat{\sigma}_n^2(\gamma)}\right)$$

is the point wise F-statistic under the null hypothesis $H_0 : \alpha_i = \beta_j$, when the threshold value γ is known.

Since γ is unidentified, the test statistic F_n does not follow the chi-square distribution and its value can not be compared to the critical value found in F Table. Hansen (1996a) provides the following bootstrapping method to approximate the asymptotic distribution of F_n.

Let u_t^*, $t = 1, 2, ..., n$, be IID N(0, 1) be normally distributed random numbers, and assume that $y_t^* = u_t^*$. Regress y_t^* on x_t by using all observations and obtain the residual variance $\tilde{\sigma}_n^{*2}$ on $x_t(\gamma)$ to obtain the residual variance $\hat{\sigma}_n^{*2}(\gamma)$. Use these two variances to form the test statistic:

$$F_n^*(\gamma) = n\left(\frac{\tilde{\sigma}_n^{*2} - \hat{\sigma}_n^{*2}(\gamma)}{\hat{\sigma}_n^{*2}(\gamma)}\right) \tag{2.5.6}$$

and therefore

$$F_n^*(\gamma) = \underset{\gamma\epsilon C}{sup} F_n^*(\gamma)$$

Repeat this process for several thousand times to obtain the distribution of F^* to approximate the distribution of F. The bootstrap approximation to the asymptotic p-value of the test is formed by counting the percentage of bootstrap samples for which F_n^* exceeds the observed F_n. Thus, in case of more than two regimes, Hansen (1996a) test statistic tests the null hypothesis of linearity against two possible alternatives of two or three regimes TAR model.

2.6 Diagnostics of Time Series Models

Different diagnostic techniques suitable for evaluating time series models are available. The first is the residual analysis of the fitted model. The residual plot shows any outliers and provides evidence of the period during which the model does not fit the data well. The Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) of the fitted residuals can be used as an evaluation technique to test the independence assumption of the fitted residuals. Heteroskedastic variance can be checked by plotting the sample ACF of the squared standardized or absolute residuals. The information criteria like AIC, BIC, and QQ plot can be used as diagnostic tools to assess the in-sample goodness of fit of linear and nonlinear time series models. Nonlinearity tests discussed in this chapter can also be used as evaluation tools for testing the nonlinearity of time series before using nonlinear time series models. The Hansen (1996a) test can also be utilized an evaluation tool as by using null hypothesis of linearity. If the null hypothesis is rejected, then one could also test the remaining nonlinearity to choose between two or three regime TAR models. In this case, the null hypothesis consists of one regime TAR model against the alternative of a two regime TAR model. Mean Square Forecasting Error (MSFE) can also be used as an evaluation criterion to select the best model among different potential candidates.

2.7 Forecasting

After selecting a parsimonious time series model, the next step is to forecast the future observations of a time series based on the fitted best time series model. These forecasts are then used in policy making decisions. Nonlinear time series models serves as adequate description of the underlying dynamic pattern present in a time series. In many cases, however, we need future values of a time series based on the observed pattern, therefore we employ the model for forecasting time series. Furthermore, out-of-sample forecasting can also be used as an evaluation tool to evaluate the regime switching models, in particular, comparing the forecasts from nonlinear models with those from the linear benchmark models. There are different procedures to forecast the linear and nonlinear time series models and then to compare these forecasts. Generally, if a model describes the features of a time series better with an in-sample fit, there is no guarantee that it also renders better forecasts (Frances and van Dijk, 2000).

Computing point forecasts from nonlinear time series models is not an easy task and it requires a lot of computations. Consider that y_t is a nonlinear autoregressive model with lag length one:

$$y_t = F(y_{t-1}; \psi) + e_t \tag{2.7.1}$$

where $F(y_{t-1}; \psi)$ is a nonlinear function. By using the least square criterion, the conditional expectation is considered as the optimal point forecasts of future values of the time series. The optimal h-step ahead forecast of y_{t+h} at time t is given by:

$$\hat{y}_{t+h|h} = E[y_{t+h} \mid \Omega_t] \tag{2.7.2}$$

whee Ω_t denotes the history of a time series up to time t. By assuming errors are independent and identically distributed and using equation (2.7.1), the optimal 1-step-ahead forecast can easily be obtained as:

$$\hat{y}_{t+1|t} = E[y_{t+1} \mid \Omega_t] = F(y_t; \psi) \tag{2.7.3}$$

which is equivalent to 1-step-ahead forecast from a linear time series model. The calculations become more complicated when we forecast nonlinear time series models for more than 1-step-ahead forecasts. For example the optimal 2-step-ahead forecast follows from (2.7.1) and (2.7.2) as:

$$\hat{y}_{t+2|t} = E[y_{t+2} \mid \Omega_t] = E[F(y_{t+1}; \psi) \mid \Omega_t] \tag{2.7.4}$$

In general, the linear conditional expectation operator E cannot be interchanged with the nonlinear operator F, that is:

$$E[F(y_{t+1}; \psi) \mid \Omega_t] \neq F(E[y_{t+1} \mid \Omega_t]; \psi) \tag{2.7.5}$$

The relation between 1 and 2-step-ahead forecast is given by:

$$\begin{aligned} \hat{y}_{t+2|t} &= E[F(F(y_t; \psi) + e_{t+1}; \psi) \mid \Omega_t] \\ \hat{y}_{t+2|t} &= E[F(\hat{y}_{t+1/t} + e_{t+1}; \psi) \mid \Omega_t] \end{aligned} \tag{2.7.6}$$

Thus, recursive relationship between forecasts at different horizons could be used to obtain multi-step-ahead forecasts. Thus, the 2-step-ahead forecast is as follows:

$$\hat{y}^{(n)}_{t+2|t} = F(\hat{y}_{t+1|t}; \psi) \tag{2.7.7}$$

This approach of forecasting is called ‘naïve’ or ‘skeleton’ method.

Brown and Mariano (1989) show that this ‘naïve’ approach renders biased forecasts. An alternative to cure this problem is to use the Monte Carlo and bootstrap methods for multi-step-ahead forecasts to approximate the conditional expectation (2.7.6). The 2-step-ahead Monte Carlo forecast is given by:

$$\hat{y}^{(mc)}_{t+2|t} = \frac{1}{k} \sum_{i=1}^{k} F(\hat{y}_{t+1/t} + e_i; \psi) \tag{2.7.8}$$

where k is some large number and e_i are random numbers drawn from the presumed distribution of e_{t+1}.

The bootstrap forecast is similar to (2.7.8) except that the residuals are re-sampled from the empirical residuals $\hat{e}_i$ from the estimated model. The bootstrap 2-step-ahead forecast is given by:

$$\hat{y}^{(b)}_{t+2|t} = \frac{1}{k}\sum_{i=1}^{k} F(\hat{y}_{t+1/t} + \hat{e}_i; \psi) \tag{2.7.9}$$

The advantage of bootstrap over the Monte Carlo method is that we did not impose any distributional assumption on e_{t+1}. Lin and Granger (1994) and Clements and Smith (1997) forecasted SETAR and STAR models by using various methods of forecasting and found that Monte Carlo and bootstrap methods perform well relative to other methods to obtain multi-step-ahead forecasts.

The only difference between 1-step-ahead and multi-step-ahead forecasts is that the first one is computed using all the original data and the latter one is computed using not only the original data but also the forecasted values of the previous $(h-1)$ period for h-step-ahead forecast. In 1-step-ahead forecasting, we re-estimated the model every time by adding one more period data into the sample, then estimated the model again and make one-period-ahead prediction on the basis of this new estimated model.

The evaluation criterion used in this study to compare the linear and non linear forecasts is the Mean Square Forecasting Error (MSFE), which is defined as:

$$MSFE = \frac{\sum_{t=1}^{h}(y_t - \hat{y}_t)^2}{h} \tag{2.7.10}$$

where h is the period of forecasting horizon and $\hat{y}_t$ is the forecasted value of a realization y_t at time period t.

2.8 The Modified Diebold-Mariano (MDM) Test

Diebold and Mariano (1995) introduced a new parametric test which by virtue of its robustness to all the error properties has been shown to be superior to its predecessors in a simulation study. The test takes a very general specification, performing a test on the sample mean of a loss differential function, which can be arbitrarily defined, e.g., as the difference between the two mean square forecast errors. The Diebold-Mariano test is straightforward to implement and has very attractive robustness properties compared to its rivals (Harvey *et al.*, 1997).

The Diebold-Mariano test statistic divides the loss differential sample mean by an estimate of its standard error found to be biased in finite samples. The Modified Diebold-Mariano (MDM) test embodies two amendments: first, a finite sample correction to the variance estimator in the test statistic to obtain unbiased estimate of standard error, and second, the use of critical t-distribution values rather than a standard normal distribution. A comprehensive simulation study shows that the Modified Diebold-Mariano (MDM) test exhibits significant and substantial improvements to the original test (Harvey *et al.*, 1997).

Suppose two competing forecasts are made from the estimated model 1 and model 2 with errors e_{1t} and e_{2t} respectively for $t = 1, 2, ..., n$. Denoting the loss function from these forecast errors in period t by $g(e_{1t})$ and $g(e_{2t})$ respectively then the corresponding differential loss function can be constructed as $d_t = g(e_{1t}) - g(e_{2t})$. Now our aim is to test the null hypothesis of equal forecast accuracy, $H_0 : E(d_t) = \bar{d} = 0$. The Diebold-Mariano test statistic is then defined as:

$$S_1 = \frac{\bar{d}}{\sqrt{\hat{V}(\bar{d})}} \tag{2.8.1}$$

The test statistic S_1 follows the standard normal distribution, where

$\bar{d} = \frac{1}{n}\sum_{t=1}^{n} d_t$, $\hat{V}(\bar{d}) = \frac{1}{n}\left[\hat{\gamma}_0 + 2\sum_{i=1}^{n-1}\hat{\gamma}_i\right]$ and $\hat{\gamma}_i = \frac{1}{n}\sum_{i=1}^{n-1}(d_t - \bar{d})(d_{t-i} - \bar{d})$, and where $\hat{\gamma}_i$ denote ith autocovariance of the d_t series.

A general optimal h-steps-ahead forecast error will now be a function of future white noise terms forming an $MA(h-1)$ type process, and this renders zero autocovariances for all lags greater than $h-1$. Harvey (1997) argued that practically this result may not hold, but would be expected for reasonably well-conceived forecasts. By using this result to the variance estimate of the loss differential sample mean, the following estimator for an h-steps-ahead forecast is obtained:

$$\hat{V}(\bar{d}) = \frac{1}{n}\left[\hat{\gamma}_0 + 2\sum_{i=1}^{h-1}\hat{\gamma}_i\right]$$

The problem with this Diebold-Mariano test statistic is that its empirical size significantly exceeds its nominal size in cases of moderate and small sample sizes. As a result, one should use the Modified Diebold-Mariano (MDM) test statistic corrected for small sample bias in original variance estimate, which is given as:

$$MDM = S_1\sqrt{\frac{(n+1-2h+(\frac{h}{n})(h-1))}{n}} \tag{2.8.2}$$

Compare the sample value of MDM test statistic to a tabulated t-value with $n-1$ degrees of freedom.

Chapter 3

Nonlinear Time Series Models with External Threshold Variables for Forecasting Economic and Financial Time Series

A detailed analysis of the forecasting performance of standard and extended linear and nonlinear time series models for financial and economic data was conducted by including external threshold variables in threshold models. Our main goal was to establish whether simple linear models still perform well or whether they should be replaced with the new sophisticated nonlinear models for modeling and forecasting financial and economic time series data. Of course, multivariate time series models could also be considered but this would increase the uncertainty in the model specification and estimation, mainly in the nonlinear time series models. Banerjee and Marcellino (2005) argue that univariate time series models are often more robust than their multivariate counterparts.

Dinghai (2010) argued that recent empirical research work shows that financial asset returns and volatilities exhibit asymmetric behavior in different regimes. Li and Lam (1995) observed significantly asymmetric movements of the conditional mean structure corresponding to the rise and fall of the previous day market. Liu and Maheu (2008) found strong empirical evidence of asymmetry in the volatility regimes. To handle these asymmetric effects, So, Li, and Lam (2002) extended the standard Stochastic Volatility (SV) model into a threshold framework, in which the latent volatility dynamic is determined by the sign of the lagged return. Andersen and Bollerslev (1997) introduced the concept of realized volatility, and there has been considerable interest in its distribution, modeling, and forecasting. Goldman *et al.* (2013) analyzed the dynamics of realized variance and bipower variation of daily stock returns of 30 companies in the Dow Jones index using a threshold fractionally integrated autoregressive and moving average models. Heterogeneous Autoregressive (HAR) model, introduced by Corsi (2009), is also used to capture a few stylized facts observed in the high-frequency realized volatility data.

In this study, a large variety of models are considered, including linear Autoregressive (AR) and Heterogeneous Autoregressive (HAR) models, four types of nonlinear time series models, namely Heterogeneous Self-Exciting Threshold Autoregressive (HSETAR), Self-Exciting Threshold Autoregressive (SETAR), Logistic Smooth Transition Autoregressive (LSTAR), and Artificial Neural Network (ANN) models, and one nonparametric Additive Autoregressive (AAR) model. SETAR model allows the model parameters to change according to the threshold variable, while LSTAR model is useful for time series that change their behavior through a smooth transition variable. The ANN model, which is hard to interpret from an economic point of view, are able to approximate any nonlinear behavior very well, (see studies by Hornik, Stinchcombe and White (1989), Swanson and White (1997)). Many empirical regularities on RV have been well-documented in recent literature and a detailed review has been provided by McAleer and Medeiros (2008a). We are also interested in the forecasting comparison of linear Heterogeneous Autoregressive (HAR) model with its extended version, i.e., Heterogeneous Self-Exciting Threshold Autoregressive (HSETAR) model. Several studies shows the inability of regime switching models to generate improved out-of-sample forecasts over linear models. One possible reconciliation is offered by Dacco and Satchell (1999), who suggest that regime switching models may not provide improved forecasts, owing to the difficulty of forecasting the regime that the series will be in. Thus, any in-sample gain to nonlinear model in a given regime will be lost if the model does not correctly forecast the regime.

In the first part of this chapter, regime switching models with external threshold variable are introduced by using the daily realized variance data from the USA, the UK, Germany, and Japan. Nonlinear time series models with internal and external threshold variables are estimated, and 1 through 4 step-ahead forecasts are obtained by using the bootstrapped method. These models are then compared with a benchmark, the Heterogeneous Autoregressive (HAR) long memory model of Corsi (2009). Secondly, to ascertain the dependence of the threshold models on some external threshold variable, the forecasting accuracy of the threshold time series models with internal and external threshold variables are compared with the linear benchmark HAR model. Furthermore, the threshold models with internal threshold variables are compared with the same model containing external threshold variables. The forecasts from extended HSETAR model are also compared with benchmark HAR model.

In the second part of this chapter, the forecasting accuracy of nonlinear time series models are compared with the benchmark linear AR model by utilizing the real GDP growth rate data of the USA. Different potential macroeconomic external threshold variables are used in TAR and LSTAR models, and their forecasting performance is compared with benchmark AR model and with threshold models with internal threshold variables. By extending this analysis, we also compare the pairwise forecasting performance of threshold models with internal and external threshold variables.

3.1 Financial Data Analysis and Results

First, the time series models were applied to the daily Realized Volatility (RV) data of Dow Jones Industrial Average (DJIA), London stock exchange (FTSE100), German stock exchange (DAX), and Japan stock exchange (Nikkei) covering the time period from January 2006 to October 2012 with more than 1600 observations. In this study, the realized volatility time series is defined as the logarithm of the square root of the realized variance of the respective stock market, i.e., $y_t = Ln(\sqrt{RV_t})$. In this study, the FTSE100 is referred to as FTSE.

We applied the TAR or SETAR, LSTAR, ANN, AAR, VAR, HSETAR, and HAR models to the daily realized volatility data of the USA, the UK, Germany, and Japan. All variables are stationary according to Augmented Dickey Fuller (ADF) test. Expanding window technique of model estimation was adopted here to obtain recursively 1 through 4 step-ahead forecasts, i.e., $h = 1, 2, 3, 4$. We re-estimated each of the above mentioned models by adding one more observation to the estimation period, whereafter the 1 through 4 step-ahead point forecasts were recursively generated by using these estimated models. Each model was specified only once, and these same model specifications were used throughout the expanding window technique. According to minimum Bayesian Information Criteria (BIC), for all realized volatility time series, lag length five and two respectively were used in lower and upper regimes of TAR and LSTAR models, and the threshold delay parameter was set to one in these models to ease the computational burden. Similarly, VAR model of lag length 3 was estimated for all four realized volatility time series. Furthermore, to obtain reliable parameter estimates in each regime, we imposed a condition on threshold models that each regime must contain at least 15 percent of the total observations. In a similar way, the lag length value and the number of hidden units (size) were set to three and two respectively in ANN model, while in AAR model lag length five was used. We used first 500 observations as an estimation period under the expanding window technique (approximately two years of data), then 1 through 4 step-ahead forecasts were generated from each of the time series models considered in this study. This process continued under the expanding window techniques until the final estimation period contained $(n - max(h))$, observations. We left the last four observations from the estimation period in our last estimation window to compare the actual and the forecasted values. Thus, the first and last estimation window respectively contained observations for the time period up to January 2008 and October, 2012. Naïve and Bootstrapped methods of forecasting were respectively used to predict the HAR and nonlinear models. The bootstrapped replication was set to 1000 when forecasting the nonlinear models. To get out-of-sample forecasts from nonlinear models, the forecasted values were averaged over 1000 bootstrapped replications. After getting the forecasting errors from each of the used models, the values of Mean Square Forecasting Error (MSFE) for 1 through 4 step-ahead forecasts were calculated.

This study differs from the previous studies in the sense that it used four types of nonlinear time series models (SETAR, LSTAR, ANN, and AAR) and compared their forecasting performance with a linear benchmark Heterogeneous Autoregressive (HAR) model. Secondly, we estimated SETAR and LSTAR models by using external threshold variables. Henceforth referred to these models as TAR-Ext and LSTAR-Ext and compared their forecasting performance with HAR, SETAR and LSTAR-Int models. In third instance, according to the current state of knowledge, this is the first study of its kind that utilizes the HSETAR model and compared their forecasting performance with HAR model. We also compared the forecasting performance of HAR and VAR models. The SETAR and LSTAR-Int models were referred to as the threshold models, where simply the lag values of dependent variable (y_t) were used as an internal threshold variable. Thus, this study considers a large variety of standard and extended models to forecast the realized volatility data of four countries. The end goal is to choose the model with improved forecasting performance.

The out-of-sample forecasting performance of SETAR and LSTAR models were assessed by assuming the following specifications of external threshold variables in these models (the corresponding names of the model are mentioned in brackets)[1].

i. Using the first lag values of internal threshold variable (SETAR or LSTAR-Int).

ii. Using the first lag values of external threshold variable (TAR-Ext-Simple or

[1]Many other specifications of the external threshold variables (e.g., square root transformation and first difference of external threshold variables, 2, 3 and, 4 period moving average) were also used but they did not improve the out-of-sample forecasting performance of threshold models over the linear model, so these results are not reported here.

LSTAR-Ext-Simple).

iii. Using the first lag values of 5 days moving average of internal threshold variable (SETAR-MA or LSTAR-Int-MA).

iv. Using the first lag values of 5 days moving average of external threshold variable (TAR-Ext-MA or LSTAR-Ext-MA).

As an illustration, assume that the realized volatility data of DJIA is the dependent variable (y_t) in all models considered in this study. Then, according to the above four cases, the respective threshold variables will be, using the first lag values of DJIA, using the first lag values of FTSE, DAX, or Nikkei, using the first lag values of five period moving average of DJIA, and using the first lag five period moving average values of FTSE, DAX, or Nikkei.

To visualize the nonlinear behavior, we plotted the raw and logarithmic realized volatility data from the USA, the UK, Germany, and Japan in Figure 3.1.1. Asymmetric behavior is apparent in this figure for all realized volatility time series, i.e., DJIA, FTSE, DAX, and Nikkei, suggesting the use of threshold models. We observed multiple spikes in realized volatility data for all four stocks with slight difference in their magnitude. The magnitude of spikes in realized volatility data of Nikkei is relatively larger and steeper than for other markets. For all four stocks, we observed a higher magnitude spikes around 2008-2009 (financial crises) and one smaller spike around 2011. The realized volatility data of DJIA are relatively smoother than for other stock markets. Regime-switching models allow us to account for such asymmetric behavior.

Before testing the nonlinear property of realized volatility data, the ACF and PACF plots were used to identify the autoregressive order. These plots are shown in Figure 3.1.2. The ACF function decays very slowly and remains significant even after 30 lags, while the PACF function is significant for the first few lags. This suggests that an autoregressive type model with order ranging from 2 to 6 may be considered as a starting point for modeling and forecasting the logarithms of realized volatility.

Table 3.1 reports the summary statistics, i.e., mean, variance, skewness, Excess kurtosis, minimum, and maximum of the realized volatility data for four countries. The values of skewness and excess-kurtosis indicate that all countries realized volatility time series are approaching towards normality by considering the logarithmic transformation of the square root of realized volatility data, i.e., $y_t = ln(\sqrt{RV_t})$. The smallest and largest variances are respectively observed for Nikkei and DJIA.

Figure 3.1.1: Realized volatility time series

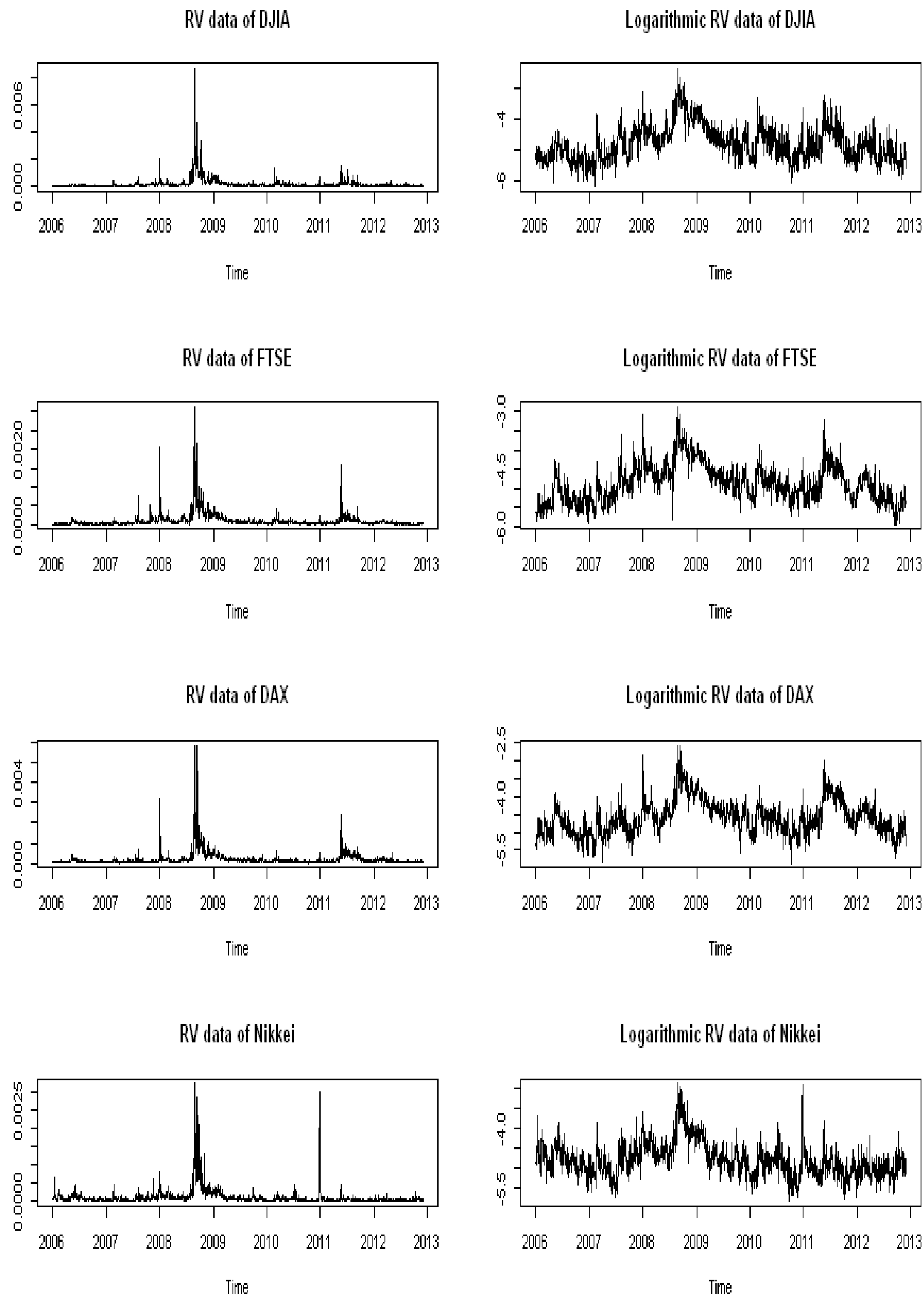


Figure 3.1.2: ACFs and PACFs of the realized volatility time series

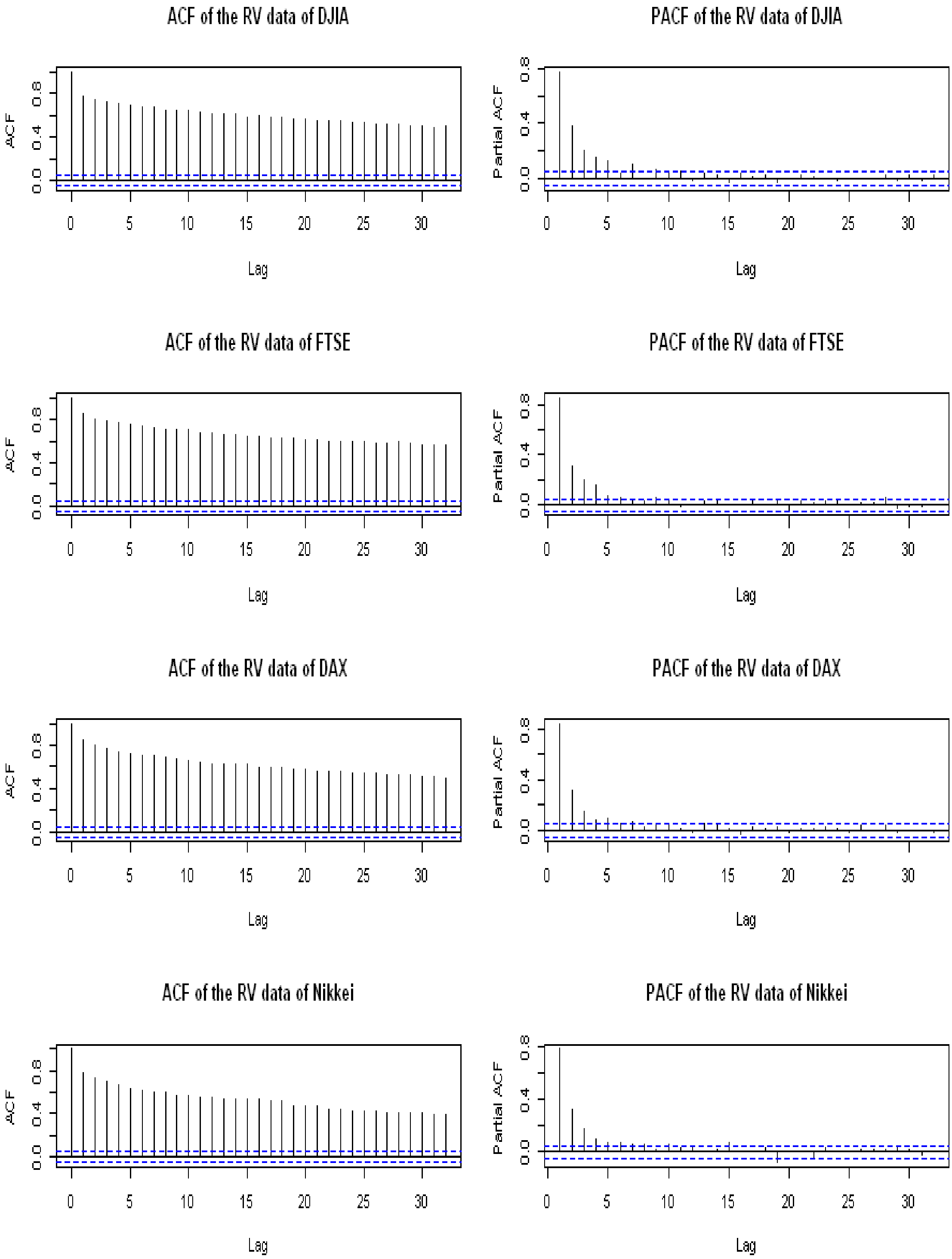


Table 3.1: Summary statistics of realized volatility time series

Variables	Specifications	Mean	Variance	Minimum	Maximum	Skewness	Ex-kurtosis
DJIA	RV	0.00016	0.00000	0.00000	0.00900	10.41500	175.28300
	$\sqrt{RV}$	0.00996	0.00006	0.00200	0.09300	3.28900	17.90700
	$ln(\sqrt{RV})$	-4.79307	0.32394	-6.17000	-2.37700	0.68400	0.427000
FTSE	RV	0.00010	0.00000	0.00000	0.00300	7.79000	90.25400
	$\sqrt{RV}$	0.00868	0.00003	0.00300	0.05600	2.76100	13.11800
	$ln(\sqrt{RV})$	-4.87847	0.23993	-5.94100	-2.88500	0.56200	0.259000
DAX	RV	0.00018	0.00000	0.00000	0.00600	8.78600	112.69500
	$\sqrt{RV}$	0.01139	0.00005	0.00300	0.07700	3.10900	16.60400
	$ln(\sqrt{RV})$	-4.60475	0.23409	-5.86300	-2.56800	0.58900	0.613000
Nikkei	RV	0.00012	0.00000	0.00000	0.00300	7.74900	78.57000
	$\sqrt{RV}$	0.00936	0.00003	0.00300	0.05700	3.49400	18.57600
	$ln(\sqrt{RV})$	-4.77884	0.18570	-5.80700	-2.86800	0.90600	1.582000

All realized volatility time series, i.e., $y_t = ln(\sqrt{RV_t})$ are tested for nonlinearity. The results of Keenan, Tsay, and Hansen tests for testing linearity are presented in Table 3.2. The delay parameter was set to one in Hansen test. All p-values of the linearity tests reported in Table 3.2 are smaller than 0.05, except for the Hansen test which provides a p-value smaller than 0.10 for DAX. Hence, according to the nonlinearity tests, almost all realized volatility time series are nonlinear. In this study an attempt was made to capture these nonlinear dynamics through nonlinear time series models and thereby to achieve improved out-of-sample forecasting gain for these models in comparison to linear models.

Our main focus in this chapter is to compare the forecasting performance of linear and nonlinear models. Secondly, an attempt is made to compare the forecasting performance of threshold models by using the internal and external threshold variables in these models. HAR model is considered to be the benchmark model for all types of nonlinear models, and SETAR and LSTAR-Int models are considered as benchmark models for TAR-Ext and LSTAR-Ext models. To compare the forecasting accuracy of different models, we assessed the significance of the observed differences in the values of Mean Square Forecasting Error (MSFE) by using the pairwise Modified Diebold-Mariano (MDM) test of Harvey, Leybourne, and Newbold (1997). We constructed two null hypotheses: under the first null hypothesis, we assumed that the Mean Square Forecasting Error (MSFE) of the linear HAR model is smaller than MSFE of the corresponding nonlinear model, whereas in the second one, we assumed that the MSFE of the SETAR or LSTAR-Int models is smaller than MSFE of TAR-Ext or LSTAR-Ext models with proper alternatives, i.e., opposite of the null hypothesis. This implies that under the null hypothesis any p-values equal or smaller than 0.05 or 0.10 indicate the rejection of the null hypothesis and thereby an improved forecasting gain to nonlinear models relative to the benchmark model. We compared the observed p-values of MDM test with pre-assumed significance levels of 0.05 and 0.10. By using

Table 3.2: P-values of the linearity tests

Variables	Keenan test	Tsay test	Hansen test
DJIA	0.000	0.041	0.007
FTSE	0.000	0.046	0.047
DAX	0.000	0.005	0.070
Nikkei	0.000	0.001	0.023

Note: Under the null hypothesis in linearity tests, we assume that the time series is linear.

BIC, we set the weekly aggregated volatility as the threshold variable when modeling and forecasting the realized volatility data of DJIA through HSETAR model. We also combined the HAR and HSETAR forecasts by taking their average values and comparing them with benchmark HAR forecasts.

The Mean Square Forecasting Error (MSFE) values of linear and nonlinear models are listed in Tables 3.3 and 3.4. These tables show that almost all MSFE values from all models are similar in magnitude. However, in a few cases, slight differences in the values of MSFE have been observed.

Tables 3.5 to 3.7, contain the Modified Diebold-Mariano (MDM) test results to pair-wise compare the forecasting accuracy of linear and various nonlinear models. The entries presented in these Tables are the p-values of the MDM test, for forecasting horizon 1 through 4. Table 3.5 contains the p-values of the MDM test to compare the forecasting accuracy of linear benchmark HAR model with nonlinear threshold models. Table 3.5 shows that HAR model clearly renders the most accurate forecasts in comparison to threshold models by considering DJIA and DAX as dependent variables (y_t) in all time series models and by assuming different specifications of the FTSE, DAX, and Nikkei as an external threshold variables in threshold models. While modeling the realized volatility data of DJIA, we observed that for forecasting horizon $h = 1$, forecasts from TAR-Ext and LSTAR-Ext models beat the HAR forecasts by using the first lag realized volatility values of FTSE as an external threshold variable in threshold models. Further taking DJIA as the dependent variable (y_t) in all time series models, TAR-Ext and LSTAR-Ext models provide superior forecasts than HAR model for $h = 1$, by using the first lag five period moving average realized volatility values of Nikkei and DAX as an external threshold variables. Nikkei may randomly attributed the nonlinear behavior in DJIA.

Table 3.5, we observed that in most cases HAR model provides superior forecasts than threshold models with a different combination of external threshold variables by considering the realized volatility data of DAX, Nikkei, and FTSE as dependent variables (y_t) in all time series models. Table 3.5 shows that by consid-

Table 3.3: Point forecast evaluation of threshold models with internal and external threshold variables for realized volatility time series: MSFE values

Dependent variables	Models	External Threshold variables															
		DJIA				FTSE				DAX				Nikkei			
		h=1	h=2	h=3	h=4	h=1	h=2	h=3	h=4	h=1	h=2	h=3	h=4	h=1	h=2	h=3	h=4
DJIA	TAR-Ext-Simple	–	–	–	–	109	117	128	136	109	117	129	136	111	121	133	141
	TAR-Ext-MA	–	–	–	–	110	117	129	135	110	119	130	138	110	118	130	138
	LSTAR-Ext-Simple	–	–	–	–	109	116	128	135	111	120	133	143	111	121	133	141
	LSTAR-Ext-MA	–	–	–	–	108	114	124	128	111	120	131	140	110	119	130	138
FTSE	TAR-Ext-Simple	52	63	69	75	–	–	–	–	53	65	71	76	54	67	72	78
	TAR-Ext-MA	53	65	71	77	–	–	–	–	53	66	72	77	53	65	71	76
	LSTAR-Ext-Simple	53	64	71	77	–	–	–	–	55	66	73	81	54	67	73	78
	LSTAR-Ext-MA	52	63	69	73	–	–	–	–	55	67	72	76	54	65	71	77
DAX	TAR-Ext-Simple	52	64	75	83	51	63	73	81	–	–	–	–	54	68	79	88
	TAR-Ext-MA	54	67	78	86	54	66	77	84	–	–	–	–	54	68	80	89
	LSTAR-Ext-Simple	52	64	74	82	51	64	75	84	–	–	–	–	54	67	79	88
	LSTAR-Ext-MA	53	65	75	83	53	65	75	83	–	–	–	–	54	68	79	89
Nikkei	TAR-Ext-Simple	62	73	83	90	63	74	83	90	63	75	85	92	–	–	–	–
	TAR-Ext-MA	63	76	87	94	64	77	87	94	64	76	86	94	–	–	–	–
	LSTAR-Ext-Simple	63	75	85	97	62	73	83	89	63	75	84	93	–	–	–	–
	LSTAR-Ext-MA	64	79	89	97	64	76	86	93	64	77	87	97	–	–	–	–

Note: All entries are multiplied by 1000.

Table 3.4: Point forecast evaluation of HAR and nonlinear models with internal threshold variable for realized volatility time series: MSFE values

Models	Dependent variables															
	DJIA				FTSE				DAX				Nikkei			
	h=1	h=2	h=3	h=4	h=1	h=2	h=3	h=4	h=1	h=2	h=3	h=4	h=1	h=2	h=3	h=4
HAR	112	115	118	118	53	59	61	61	54	63	68	69	64	71	74	74
SETAR-MA	111	119	129	136	54	65	71	76	54	68	78	86	64	76	86	94
LSTAR-Int-MA	107	114	123	126	54	65	70	74	54	66	77	84	64	77	87	97
SETAR-Simple	111	120	131	138	55	67	73	78	55	68	80	89	64	77	88	96
LSTAR-Int-Simple	112	123	135	142	55	66	74	79	54	68	80	89	65	79	90	98
ANN	200	203	215	228	118	126	136	141	106	188	201	110	143	184	241	262
AAR	110	120	133	141	54	66	73	79	55	69	80	90	65	78	89	97
VAR	110	121	137	143	55	68	76	84	55	70	83	94	63	77	88	96
HSETAR	110	119	131	139	56	71	85	99	55	75	92	106	65	80	92	101
HAR+HSETAR	111	116	124	126	53	64	71	75	54	65	74	79	64	74	82	86

Note: All entries are multiplied by 1000.

Table 3.5: Point forecast evaluation of HAR and nonlinear models for realized volatility time series: P-values of the MDM test

Dependent variables	Models Comparison	External Threshold variables															
		DJIA				FTSE				DAX				Nikkei			
		h=1	h=2	h=3	h=4	h=1	h=2	h=3	h=4	h=1	h=2	h=3	h=4	h=1	h=2	h=3	h=4
DJIA	HAR vs TAR-Ext-Simple	–	–	–	–	0.033	0.832	1.000	1.000	0.360	0.999	1.000	1.000	0.018	0.873	1.000	1.000
	HAR vs TAR-Ext-MA	–	–	–	–	0.082	0.874	1.000	1.000	0.067	0.950	1.000	1.000	0.100	0.968	1.000	1.000
	HAR vs LSTAR-Ext-Simple	–	–	–	–	0.015	0.700	1.000	1.000	0.256	0.999	1.000	1.000	0.297	0.981	1.000	1.000
	HAR vs LSTAR-Ext-MA	–	–	–	–	0.082	0.874	1.000	1.000	0.067	0.950	1.000	1.000	0.100	0.968	1.000	1.000
FTSE	HAR vs TAR-Ext-Simple	0.230	0.998	1.000	1.000	–	–	–	–	0.875	1.000	1.000	1.000	0.980	1.000	1.000	1.000
	HAR vs TAR-Ext-MA	0.908	1.000	1.000	1.000	–	–	–	–	0.882	1.000	1.000	1.000	0.932	1.000	1.000	1.000
	HAR vs LSTAR-Ext-Simple	0.718	0.998	1.000	0.998	–	–	–	–	0.954	1.000	1.000	0.999	0.973	1.000	1.000	1.000
	HAR vs LSTAR-Ext-MA	0.908	1.000	1.000	1.000	–	–	–	–	0.882	1.000	1.000	1.000	0.932	1.000	1.000	1.000
DAX	HAR vs TAR-Ext-Simple	0.002	0.735	1.000	1.000	0.000	0.398	0.999	1.000	–	–	–	–	0.468	0.999	1.000	1.000
	HAR vs TAR-Ext-MA	0.437	0.997	1.000	1.000	0.250	0.993	1.000	1.000	–	–	–	–	0.621	1.000	1.000	1.000
	HAR vs LSTAR-Ext-Simple	0.002	0.660	0.998	1.000	0.000	0.676	0.998	0.999	–	–	–	–	0.445	0.996	1.000	1.000
	HAR vs LSTAR-Ext-MA	0.437	0.997	1.000	1.000	0.250	0.993	1.000	1.000	–	–	–	–	0.621	1.000	1.000	1.000
Nikkei	HAR vs TAR-Ext-Simple	0.076	0.944	0.996	0.999	0.143	0.987	0.999	0.999	0.259	0.994	1.000	1.000	–	–	–	–
	HAR vs TAR-Ext-MA	0.368	0.999	1.000	1.000	0.686	1.000	1.000	1.000	0.507	0.999	1.000	1.000	–	–	–	–
	HAR vs LSTAR-Ext-Simple	0.274	0.935	0.988	0.976	0.047	0.926	0.998	0.999	0.309	0.980	0.999	1.000	–	–	–	–
	HAR vs LSTAR-Ext-MA	0.368	0.999	1.000	1.000	0.686	1.000	1.000	1.000	0.507	0.999	1.000	1.000	–	–	–	–

Note: Under the null hypothesis in pairwise MDM test, we assume that HAR model has significantly smaller MSFE than the corresponding MSFE of the nonlinear model.

ering DAX as dependent variable (y_t), TAR-Ext and LSTAR-Ext forecasts beat the forecasts from an HAR model for $h = 1$ by taking the first lag realized volatility values of FTSE and DJIA as an external threshold variable. To sum up, HAR model mostly outperforms the nonlinear time series models, mainly for longer forecasting horizons.

In the next step, the benchmark SETAR and LSTAR-Int models (threshold models with internal threshold variables) were compared with TAR-Ext and LSTAR-Ext (threshold models with external threshold variables) models. In Table 3.6, the p-values of MDM test are presented in order to compare the forecasting accuracy of benchmark SETAR and LSTAR-Int models with TAR-Ext and LSTAR-Ext models. The results shown in Table 3.6 suggest that SETAR and LSTAR-Int forecasts are consistently beaten by the TAR-Ext and LSTAR-Ext forecasts for almost all forecasting horizons when modeling the realized volatility data of DJIA and by considering the first lag and five periods lag moving average values of the realized volatility data of FTSE as an external threshold variable. Similarly, TAR-Ext and LSTAR-Ext models provide more accurate forecasts for $h = 1$, 2, 3 over SETAR and LSTAR-Int forecasts by considering the five period lag moving average realized volatility values of Nikkei as an external threshold variable. TAR-Ext and LSTAR-Ext also beat the simple SETAR and LSTAR-Int forecasts for $h = 1$, 2, 3 by respectively considering the first lag and lag five period moving average of the realized volatility data of DAX.

Table 3.6 shows that by taking the realized volatility data of FTSE as a dependent variable (y_t) in linear and nonlinear time series models, TAR-Ext and LSTAR-Ext models consistently yield better forecasts for almost all forecasting horizons over SETAR and LSTAR-Int models by considering the first lag and lag five period moving average values of the realized volatility data of DJIA. Furthermore, the p-values of MDM test reported in Table 3.6 show that for almost all forecasting horizons, there is some improved forecasting gains to TAR-Ext and LSTAR-Ext models over simple SETAR and LSTAR-Int models by considering the first lag and lag five period moving average realized volatilities values of DAX and Nikkei.

Moreover, results in Table 3.6 show that there is an improved forecasting gain to TAR-Ext and LSTAR-Ext forecasts in comparison to simple SETAR and LSTAR-Int forecasts for almost all forecasting horizons by considering the realized volatility data of DAX as a dependent variable (y_t) in all time series models and taking the first lag and lag five period moving average realized volatilities values of FTSE and DJIA as an external threshold variable. This clearly conveys an empirical message that stock markets of the UK and the USA are the main driving forces of the regime switching behavior in DAX.

Table 3.6: Point forecast evaluation of threshold models with internal and external threshold variables: P-values of the MDM test

Dependent variables	Models Comparison	External Threshold variables															
		DJIA				FTSE				DAX				Nikkei			
		h=1	h=2	h=3	h=4	h=1	h=2	h=3	h=4	h=1	h=2	h=3	h=4	h=1	h=2	h=3	h=4
DJIA	SETAR vs TAR-Ext-simple	–	–	–	–	0.052	0	0.003	0.078	0.034	0.014	0.049	0.169	0.651	0.749	0.952	0.982
	SETAR vs TAR-Ext-MA	–	–	–	–	0.172	0.002	0.031	0.025	0.250	0.091	0.277	0.500	0.083	0.040	0.223	0.460
	LSTAR-Int vs LSTAR-Ext-simple	–	–	–	–	0.003	0.000	0.000	0.005	0.238	0.044	0.179	0.625	0.166	0.040	0.120	0.353
	LSTAR-Int vs LSTAR-Ext-MA	–	–	–	–	0.037	0.000	0.002	0.002	0.063	0.006	0.037	0.116	0.014	0.003	0.034	0.110
FTSE	SETAR vs TAR-Ext-simple	0.000	0.000	0.001	0.006	–	–	–	–	0.012	0.031	0.018	0.028	0.104	0.657	0.164	0.268
	SETAR vs TAR-Ext-MA	0.001	0.017	0.023	0.065	–	–	–	–	0.007	0.152	0.057	0.125	0.005	0.057	0.040	0.080
	LSTAR-Int vs LSTAR-Ext-simple	0.198	0.011	0.089	0.344	–	–	–	–	0.462	0.409	0.398	0.689	0.206	0.812	0.366	0.383
	LSTAR-Int vs LSTAR-Ext-MA	0.006	0.049	0.023	0.060	–	–	–	–	0.035	0.233	0.052	0.096	0.024	0.116	0.049	0.073
DAX	SETAR vs TAR-Ext-simple	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	–	–	–	–	0.243	0.348	0.191	0.372
	SETAR vs TAR-Ext-MA	0.171	0.219	0.024	0.028	0.061	0.025	0.002	0.000	–	–	–	–	0.375	0.574	0.311	0.546
	LSTAR-Int vs LSTAR-Ext-simple	0.000	0.002	0.000	0.001	0.000	0.008	0.015	0.091	–	–	–	–	0.331	0.399	0.363	0.296
	LSTAR-Int vs LSTAR-Ext-MA	0.290	0.379	0.053	0.026	0.128	0.067	0.006	0.000	–	–	–	–	0.520	0.734	0.415	0.535
Nikkei	SETAR vs TAR-Ext-simple	0.082	0.003	0.001	0.005	0.100	0.020	0.002	0.001	0.228	0.069	0.063	0.063	–	–	–	–
	SETAR vs TAR-Ext-MA	0.322	0.188	0.128	0.074	0.703	0.447	0.081	0.081	0.450	0.225	0.078	0.145	–	–	–	–
	LSTAR-Int vs LSTAR-Ext-simple	0.080	0.032	0.143	0.456	0.004	0.003	0.004	0.001	0.085	0.049	0.079	0.065	–	–	–	–
	LSTAR-Int vs LSTAR-Ext-MA	0.075	0.072	0.064	0.055	0.179	0.107	0.040	0.075	0.102	0.085	0.053	0.088	–	–	–	–

Note: Under the null hypothesis in pairwise MDM test, we assume that threshold model with internal threshold variable has significantly smaller MSFE than the MSFE of the threshold model with external threshold variable.

While considering the realized volatility data of Nikkei as a dependent variable (y_t) in all time series models, we observed that for almost all forecasting horizons SETAR and LSTAR-Int forecasts are beaten by TAR-Ext and LSTAR-Ext forecasts by considering the first lag and five periods lag moving average values of the realized volatility data of FTSE, DAX, and DJIA. Thus, stock markets of the USA, the UK, and DAX have great impact on Nikkei when comparing the forecasting accuracy of threshold models with internal and external threshold variables.

The null hypothesis in MDM test is that the Mean Square Forecasting Error (MSFE) of the HAR model is equal or smaller than the corresponding nonlinear model and alternative is the opposite of the null hypothesis. Thus, any p-value that is less than 0.05 or 0.10 provides a rejection of the null hypothesis and thus an improved forecasting gain of the nonlinear models in comparison to HAR model. To compare the HAR and nonlinear models with internal threshold variables, the MDM test is applied to the observed forecasting errors from the respective model; p-values of the test are presented in Table 3.7. This table shows that almost all p-values of the MDM test for 1 through 4 step-ahead forecasting horizons are greater than 0.10 for all realized volatility time series considered in this study. For DJIA only for $h = 1$, HSETAR model has improved forecasting gain in comparison to HAR model at 10% significance level. The interesting result emerging from this study is that for realized volatility time series of DJIA, the combination of 1-step-ahead forecasts of HAR and HSETAR models (taking the average of HAR and HSETAR forecasts) improves the forecasting gain in comparison to the simple HAR model. In most cases, HAR forecasts beat the forecasts from VAR and all nonlinear models for 1 through 4 step-ahead forecasting horizons. Hence, HAR model could be considered as the best forecasting model among all potential candidates to forecast the realized volatility data from the USA, the UK, Germany, and Japan. However, this study conveys a message that the 1-step-ahead forecast combined from HAR and HSETAR models for DJIA is much better than the HAR forecast alone. Among nonlinear models out-of-sample forecasting performance, AAR model performs comparatively better than ANN model.

The correlation matrix of the realized volatility data is shown in Table 3.8. High pairwise correlation is observed for all realized volatility time series. Table 3.9 provides the estimated parameter values of the HAR model, their standard deviations, and p-values by utilizing the whole sample realized volatility data of four countries. The BIC values are reported in the last row of the table. P-values of the estimated coefficients show that all the estimated parameters of HAR model are significant at 1% significance level. Similarly, summary statistics of the estimated threshold models by utilizing the whole sample of the four countries realized volatility time series are

Table 3.7: Point forecast evaluation of HAR and nonlinear models: P-values of the MDM test

Models Comparison	Dependent variables															
	DJIA				FTSE				DAX				Nikkei			
	h=1	h=2	h=3	h=4	h=1	h=2	h=3	h=4	h=1	h=2	h=3	h=4	h=1	h=2	h=3	h=4
HAR vs SETAR-MA	0.338	0.959	1.000	1.000	0.961	1.000	1.000	1.000	0.619	1.000	1.000	1.000	0.507	0.999	1.000	1.000
HAR vs LSTAR-Int-MA	0.013	0.348	0.911	0.946	0.961	1.000	1.000	1.000	0.505	0.977	1.000	1.000	0.679	1.000	1.000	1.000
HAR vs SETAR-Simple	0.276	0.996	1.000	1.000	1.000	1.000	1.000	1.000	0.706	1.000	1.000	1.000	0.539	0.999	1.000	1.000
HAR vs LSTAR-Int-Simple	0.517	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.594	1.000	1.000	1.000	0.848	0.998	1.000	1.000
HAR vs ANN	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.932	0.888	0.906	1.000	0.896	0.885	0.879	0.890
HAR vs AAR	0.120	0.996	1.000	1.000	0.988	1.000	1.000	1.000	0.894	1.000	1.000	1.000	0.949	1.000	1.000	1.000
HAR vs VAR	0.163	0.977	1.000	1.000	0.974	1.000	1.000	1.000	0.821	1.000	1.000	1.000	0.267	0.996	1.000	1.000
HAR vs HSETAR	0.100	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.982	0.999	0.999	0.999
HAR vs (HAR+HSETAR)	0.010	0.890	1.000	1.000	0.770	1.000	1.000	1.000	0.248	0.998	1.000	1.000	0.732	0.999	1.000	1.000

Note: Under the null hypothesis in pairwise MDM test, we assume that HAR model has significantly smaller MSFE than the corresponding MSFE of the nonlinear model.

Table 3.8: Correlation matrix of realized volatility time series

—	DJIA	Nikkei	FTSE	DAX
DJIA	1.00	0.61	0.86	0.81
Nikkei	0.61	1.00	0.60	0.55
FTSE	0.86	0.60	1.00	0.89
DAX	0.81	0.55	0.89	1.00

Table 3.9: HAR model estimation results for realized volatility time series

Estimates	DJIA			FTSE			DAX			Nikkei		
	Values	Std.	P-values	Values	Std.	P-values	values	Std.	P-values	Values	Std.	P-values
Constant	-0.360	0.078	0.000	-0.284	0.064	0.000	-0.275	0.065	0.000	-0.429	0.082	0.000
$y_{t-1}^{(d)}$	0.264	0.029	0.000	0.386	0.029	0.000	0.436	0.029	0.000	0.379	0.030	0.000
$y_{t-1}^{(w)}$	0.419	0.045	0.000	0.385	0.044	0.000	0.342	0.043	0.000	0.353	0.044	0.000
$y_{t-1}^{(m)}$	0.253	0.038	0.000	0.176	0.034	0.000	0.168	0.034	0.000	0.183	0.036	0.000
BIC	983.952			-86.234			14.570			160.076		

presented in Tables 3.10 and 3.11[2]. For all four realized volatility time series, the observed p-values of the estimated parameters of SETAR and LSTAR-Int models are reported in Tables 3.10 and 3.11. It can be concluded from the observed p-values of the estimated parameters that in both regimes of SETAR and LSTAR-Int models almost all the estimated coefficients are significant at conventional significance level. BIC values almost equally favor both the SETAR and LSTAR-Int models across all four stocks. Consistent results regarding threshold value have been observed for threshold models for all four stocks.

The estimation results of the HSETAR model are provided in Table 3.12. All lag coefficients are significant at 1% significance level. It is important to note that regarding in-sample fitting, HSETAR model dominated over all linear and nonlinear models by providing the smallest values of BIC for all four realized volatility time series.

[2]As we are mostly interested in the forecasting comparison of the HAR and threshold models with internal and external threshold variable, so only for these models we provided the estimation results. Nevertheless, on request estimation results of the other models used in the current study can be provided.

Table 3.10: Threshold models estimation results for realized volatility time series

Models	Estimates	DJIA						FTSE					
		Low regime			High regime			Low regime			High regime		
		Values	Std.	P-values	Values	Std.	P-values	Values	Std.	P-values	Values	Std.	P-values
SETAR	Constant	-0.591	0.126	0.000	-0.419	0.205	0.041	-0.395	0.095	0.000	-0.715	0.219	0.001
	y_{t-1}	0.302	0.027	0.000	0.654	0.047	0.000	0.472	0.028	0.000	0.604	0.050	0.000
	y_{t-2}	0.210	0.033	0.000	0.255	0.071	0.000	0.125	0.032	0.000	0.236	0.070	0.001
	y_{t-3}	0.126	0.029	0.000	—	—	—	0.110	0.030	0.000	—	—	—
	y_{t-4}	0.110	0.029	0.000	—	—	—	0.136	0.031	0.000	—	—	—
	y_{t-5}	0.130	0.028	0.000	—	—	—	0.077	0.028	0.006	—	—	—
	Threshold value	-4.283						-4.374					
	BIC	-3660.800						-4675.412					
LSTAR-Int	Constant	-1.034	0.210	0.000	0.758	0.241	0.002	-1.067	0.299	0.000	0.766	0.309	0.013
	y_{t-1}	0.277	0.032	0.000	0.153	0.045	0.001	0.417	0.041	0.000	0.094	0.049	0.055
	y_{t-2}	0.192	0.045	0.000	-0.005	0.060	0.940	0.073	0.061	0.228	0.048	0.067	0.478
	y_{t-3}	0.101	0.026	0.000	—	—	—	0.107	0.027	0.000	—	—	—
	y_{t-4}	0.096	0.026	0.000	—	—	—	0.116	0.027	0.000	—	—	—
	y_{t-5}	0.127	0.024	0.000	—	—	—	0.083	0.025	0.001	—	—	—
	Threshold value	-4.730						-4692.630					
	BIC	-3668.360						-4692.630					

Table 3.11: Threshold models estimation results for realized volatility time series

Models	Estimates	DAX						Nikkei					
		Low regime			High regime			Low regime			High regime		
		Values	Std.	P-values	Values	Std.	P-values	Values	Std.	P-values	Values	Std.	P-values
SETAR	Constant	-0.583	0.103	0.000	-0.585	0.184	0.002	-0.882	0.178	0.000	-0.431	0.141	0.002
	y_{t-1}	0.468	0.028	0.000	0.661	0.051	0.000	0.346	0.032	0.000	0.683	0.036	0.000
	y_{t-2}	0.172	0.032	0.000	0.196	0.064	0.002	0.201	0.042	0.000	0.230	0.046	0.000
	y_{t-3}	0.094	0.031	0.002	—	—	—	0.145	0.034	0.000	—	—	—
	y_{t-4}	0.055	0.030	0.072	—	—	—	0.051	0.034	0.126	—	—	—
	y_{t-5}	0.087	0.028	0.002	—	—	—	0.075	0.032	0.018	—	—	—
	Threshold value	-4.180						-4.651					
	BIC	-4611.323						-4476.073					
LSTAR-Int	Constant	-1.758	0.294	0.000	1.499	0.304	0.000	-1.142	0.438	0.009	0.738	0.362	0.042
	y_{t-1}	0.423	0.042	0.000	0.107	0.051	0.035	0.316	0.078	0.000	0.447	0.169	0.008
	y_{t-2}	0.010	0.063	0.879	0.189	0.070	0.007	0.226	0.062	0.000	-0.318	0.170	0.062
	y_{t-3}	0.080	0.028	0.004	—	—	—	0.100	0.027	0.000	—	—	—
	y_{t-4}	0.032	0.027	0.237	—	—	—	0.054	0.027	0.043	—	—	—
	y_{t-5}	0.102	0.024	0.000	—	—	—	0.072	0.024	0.003	—	—	—
	Threshold value	-4.843						-4.406					
	BIC	-4624.130						-4484.490					

Table 3.12: HSETAR model estimation results for realized volatility time series

Estimates	DAX						Nikkei					
	Low regime			High regime			Low regime			High regime		
	Values	Std.	P-values	Values	Std.	P-values	Values	Std.	P-values	Values	Std.	P-values
Constant	-0.317	0.213	0.137	-0.003	0.105	0.979	-1.195	0.365	0.001	-0.298	0.092	0.001
$y_{t-1}^{(d)}$	0.396	0.051	0.000	0.532	0.044	0.000	0.293	0.079	0.000	0.387	0.032	0.000
$y_{t-1}^{(w)}$	0.392	0.065	0.000	0.336	0.058	0.000	0.416	0.122	0.001	0.373	0.050	0.000
$y_{t-1}^{(m)}$	0.143	0.057	0.012	0.138	0.042	0.001	0.051	0.113	0.651	0.176	0.040	0.000
BIC	-4648.779						-4521.712					
Threshold value	-4.647						-5.075					
Estimates	DJIA						FTSE					
Constant	-0.238	0.263	0.366	-0.099	0.127	0.438	-0.3294	0.231	0.151	-0.1301	0.108	0.231
$y_{t-1}^{(d)}$	0.322	0.055	0.000	0.212	0.047	0.000	0.2449	0.057	0.000	0.4702	0.041	0.000
$y_{t-1}^{(w)}$	0.398	0.074	0.000	0.612	0.067	0.000	0.5081	0.072	0.000	0.3776	0.058	0.000
$y_{t-1}^{(m)}$	0.227	0.067	0.001	0.160	0.053	0.002	0.1826	0.058	0.001	0.128	0.043	0.003
BIC	-3698.947						-4745.785					
Threshold value	-4.838						-4.999					

3.2 Relationship Between Financial Variables

We observed a high degree relationship between four stocks used in this study, i.e., DJIA, FTSE, DAX, and Nikkei, as already shown in the correlation matrix. The global financial crisis of 2008-2009 affected all stock markets, and this fact is apparent in the figure 3.1.1. We observed co-movement in the realized volatility data for all four stocks. In addition, there is an increasing interdependence among most of the developed and emerging markets after the 2008-2009 financial crises[3].

There is time difference in the opening hours of the four stock markets used in this study, and the four stock markets open in the following order: Nikkei, DAX, FTSE, and DJIA. As a result, the volatility of the stock market that opens earlier may be subsequently affected from the previous day volatility observed in another stock market. For example, the previous day volatility in DAX, FTSE, and DJIA may affect the recent day volatility in Nikkei and could be considered as a regime governing process in it. Similarly, previous day volatility in DJIA may produce variation in present day volatility data of DAX and FTSE, and this fact is shown in the results. Hence, this provides a basis to use one country's realized volatility data as an external threshold variable when modeling and forecasting the realized volatility data of another country.

Kontonikas *et al.* (2012) analyzed the impact of Federal Funds Rate (FFR) surprises on stock returns in the United States from 1989 until 2009by considering recent financial crises. They found that prior to the financial crisis, the cuts in FFR were responsible for the increase in stock prices. Furthermore, they revealed that a structural break occurred during the financial crises, and as a result the stock market prices changed due to FFR shocks. Ghorbel and Trabelsi (2013) used copula techniques to investigate the effects of global financial crisis of 2008-2009 on the dependence between the stock markets of the USA and six other developed countries. Their empirical results showed that there is an increased tendency of market depen-

[3]With the United States being a major investor in many countries and posing a huge political influence on several countries in the world, studies have been done to investigate the causal relationship between the United States and other equity markets. Results indicate that the United States is an important global factor that moves the world markets. Cheung and Mak (1992) examine the causal relationship between the developed markets and Asian emerging markets and find that the United States market is a 'global factor' which leads both the developed and most of the Asian emerging markets. Some studies support the existence of a common trend for world stock markets, while others reject this hypothesis. Arshanapalli and Doukas (1993) analyze the stock markets of the United States, United Kingdom, France, Germany and Japan, using daily closing stock market index time series, in local currency units, covering the period of January 1980 through May 1990, and report an increasing degree of interdependence among world capital markets since the 1987 stock market crash, with Japan's Nikkei Index being the exception (Wong *et al.*, 2004).

dence between US, European, and Brazilian markets during the crisis period and that this increase started around the beginning of 2008. Similarly, they found a very high level of volatility in stock markets around the end of 2008 due to the increased degree of uncertainty during this period.

Thus, the above economic evidence supports the empirical findings of this study, which show that the regime switching behavior in any one country's realized volatility data is governed by the other country's realized volatility data. In this way, improved forecasting gain in threshold models can be achieved by including the external threshold variables.

3.3 Macroeconomic Data Analysis and Results

In the second part of this chapter, out-of-sample forecasting comparison of the nonlinear models with the benchmark linear AR model were taken for the seasonally adjusted macroeconomic data of the USA, namely quarterly real Gross Domestic Product (GDP) (1947 – 2012). Monthly Interest Rates Government Securities and Bonds percent per annum (1954 – 2012), Personal Expenditures (PE) based on monthly (1959 – 2012) and quarterly (1947 – 2012) frequencies, and quarterly real Gross Private Domestic Investment (1947 – 2012) were considered as an external threshold variables in the threshold models[4]. The main purpose of this study is to model and forecast the real GDP growth rate data of the USA using various nonlinear time series models and compare these forecasts with the benchmark AR model. Firstly, an attempt is made to compare the forecasts from the benchmark linear AR model with forecasts obtained from VAR and nonlinear models. Secondly, SETAR and LSTAR models with the first lag values of internal threshold variables are taken as benchmark models and respectively compared with TAR-Ext and LSTAR-Ext models (threshold models with external threshold variables). The purpose of this comparison is to assess whether there is any improvement in the forecasting performance of the threshold models with different external threshold variables when utilizing real GDP growth rate data of the USA. Data on all of the above macroeconomic variables were taken from the Federal Reserve Economic Data (FRED). The percentage of the first difference of the logarithmically transformed real GDP (real GDP growth) data were considered as our target variable in all time series models. All potential external threshold variables

[4]Other economic variables of USA (e.g., all employee, currency circulation, Govt. total spending's, money stock, real exports and real imports) were also used as an external threshold variables in threshold models, but the results are not reported here as these variable are unable to improve their forecasting accuracy.

used in threshold models were logarithmically transformed.

Prediction of the future values of GDP growth rate has recently become the central theme in economic research. Forecasts for real GDP growth rate are either produced by using economics-based theoretical models or by using time series models. Recently, more sophisticated time series models have been able to provide more accurate forecasts for many economic variables. Modeling economic time series by simply using the linear time series models is quite difficult after the changes that have occurred in the policy analysis, especially after World War II. In such situation, nonlinear time series models could have a comparative advantage over linear models. Stock and Watson (2003) conducted a forecasting study to assess the role of financial variables and other macroeconomic time series for forecasting GDP growth rate during the periods 1971-84 and 1985-99. They concluded that nonlinear models have no edge over the linear benchmark AR model. Teräsverta (2005) and White (2005) used a large variety of nonlinear time series models for forecasting different kinds of time series. Stock and Watson (1999) evaluated the relative forecasting performance of a linear and nonlinear models for a large number of macroeconomic variables for the USA and reached the conclusion that although linear models perform best, for some time series nonlinear models can yield substantial gains.

A comparison of the in-sample fitting between linear and nonlinear models could also be conducted but the results would be likely to biased in favor of the latter because of their large number of parameters and flexibility, flexibility (see van Dijk and Frances, 2003). Thus, these comments suggest that it is better to conduct out-of-sample forecasting comparison between linear and nonlinear models for real GDP growth rate data of the USA. In this study, the evaluation of the forecasting performance of the nonlinear time series models by introducing the different external threshold variables in these models is of particular interest. To the best of our knowledge, this is the first time such a comparison is being conducted.

By using BIC, lag length 2 is selected in all time series models used for the prediction purposes of real GDP growth, and the delay parameter is set tot 1 in threshold models. Similarly the number of hidden units (size) were set to one in ANN model. Specification, estimation, and forecasting from the time series models considered in this study, by utilizing the real GDP growth rate data of the USA were carried out recursively by using the expanding window technique of estimation. In this way 32, 53, 61 and 61, 1 through 4 step-ahead point forecasts from all time series model were generated by respectively using the monthly Interest Rates Government Securities and Bonds percent per annum, Personal Expenditures based on monthly and quarterly frequencies and quarterly Real Gross Private Domestic Investment as an external threshold variables in the threshold models. Thus, each time one more observation is added to the estimation period, whereafter 1 through 4 step-ahead forecasts are generated by using the 'naïve' and bootstrap method of forecasting for linear and nonlinear models respectively. To get single out-of-sample forecasts from nonlinear models, the bootstrapped forecasts are averaged over 1000 bootstrapped replications.

According to BIC, VAR model with lag length 4 is estimated by using the quarterly percentage first differences of real GDP, Personal Expenditures (PE), and real Gross Private Domestic Investment. From the VAR model, 61 point forecasts, for each of the forecasting horizon 1 through 4, for real GDP growth rate data were generated. The correlation coefficient between real GDP growth rate with quarterly Personal Expenditures (PE) is 0.60, and it is 0.79 for real GDP growth rate and real Gross Private Domestic Investment. The high correlation between economic variables leads us to estimate VAR model for them.

Figure 3.3.1 shows the plot of the different macroeconomic time series based on monthly and quarterly frequencies. The decline in volatility in real GDP growth rate is obvious after the mid-1980s until the financial crises of 2008-2009. The graph highlights the irregular behavior of the different macroeconomic variables at different time periods. Some higher magnitude spikes are observed in real GDP growth rate and other macroeconomic variables used as an external threshold variable in SETAR and LSTAR models.

Table 3.13, provides the summary statistics of the various specifications of the real GDP growth rate time series (y_t) of USA. These values of skewness and excess kurtosis show that the first difference of the logarithmically transformed real GDP growth rate data is much closer to the normal distribution in comparison to its other transformations considered in this study. For real GDP growth rate data, the P-values and the corresponding lag length used in linearity tests, i.e., Keenan, Tsay, and Hansen, are reported in Table 3.14. From observed p-values of the Tsay and Keenan tests, linearity is rejected respectively at 5 % and 10 % level of significance, while the Hansen test that tests the threshold type nonlinearity accepts the linearity hypothesis for real GDP growth rate data. In Hansen test, lag length 2 is used in AR and SETAR models. Thus, the rejection of the linearity hypothesis by these two tests (Keenan and Tsay) provided an evidence to apply the nonlinear models to model and predict the real GDP growth rate of the USA. The ACFs and PACFs of the real GDP growth rate data are shown in Figure 3.3.2, suggesting that some autoregressive type

Figure 3.3.1: Macroeconomic time series

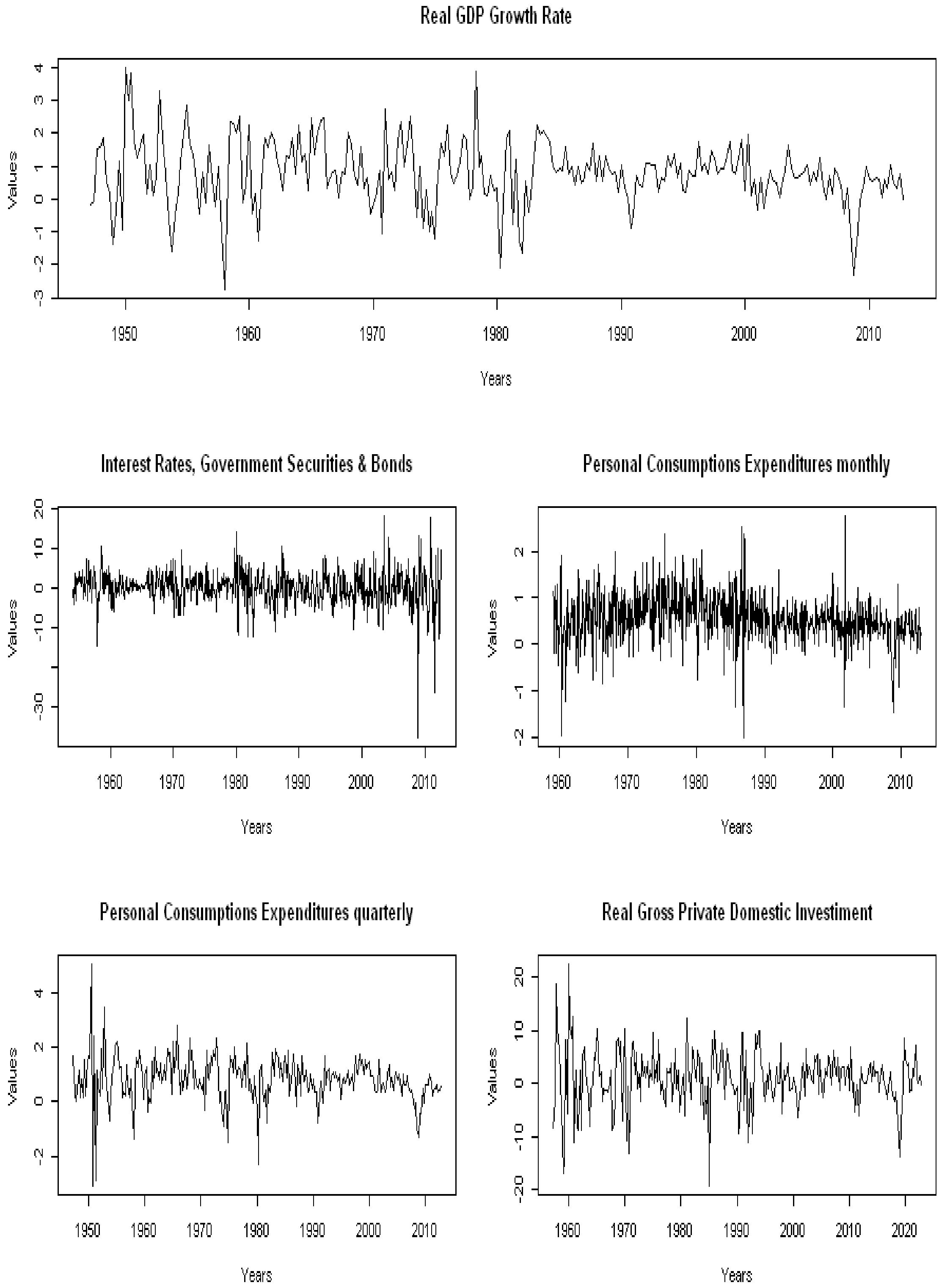
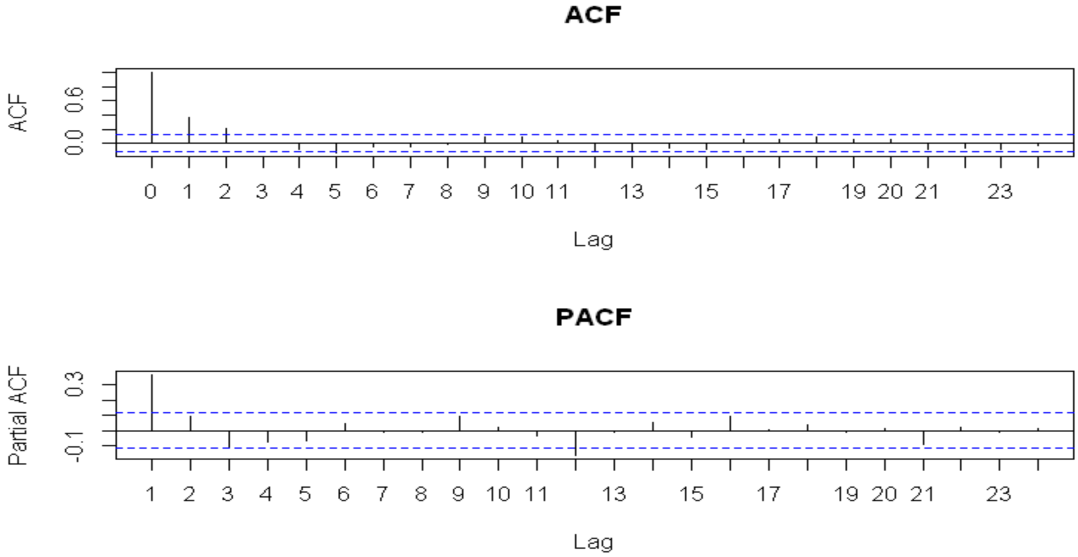


Figure 3.3.2: ACFs and PACFs of real GDP growth rate data



model could be used for modeling and predicting the real GDP growth rate data.

Table 3.13: Summary statistics of the real GDP growth rate data

Real GDP growth rate	Mean	Variance	Minimum	Maximum	Skewness	Ex.-Kurtosis
y_t	6602.408	13859962	1766.500	13652.500	0.480	-1.073
$ln(y_t)$	8.619	0.381	7.477	9.522	-0.187	-1.175
$\triangle(ln(y_t))$	0.776	0.972	-2.745	3.961	-0.096	1.394

Table 3.14: P-values of the linearity tests for real GDP growth rate data

Variable	Keenan test	Tsay test	Hansen test
Real GDP growth	0.100 (4)	0.011 (6)	0.346

Note: Under the null hypothesis in linearity tests, we assume that the time series is linear.

To regard a monthly economic time series as an external threshold variable, the values for the months of February, May, August, and November of a given year corresponding to the four quarters of quarterly real GDP growth rate data were extracted. For example, if the first quarter value of real GDP growth rate is taken, then the value corresponding to the month of February is taken as an external threshold value, and so on. The specifications of the external threshold variables used in this study and the corresponding names given to TAR and LSTAR models are mentioned below. All external threshold variables used in this study are in the form of percentage logarithmic transformation.

i. Using the first lag values of the level series as an external threshold variable (TAR-Ext-lag-(z) or LSTAR-Ext-lag-(z)).

ii. Using the first lag values of the first differenced series as an external threshold variable (TAR-Ext-lag-D(z) or LSTAR-Ext-lag-D(z)).

Table 3.15: Point forecast evaluation of the linear and nonlinear models with internal threshold variables for real GDP growth rate data: MSFE values

Models	h=1	h=2	h=3	h=4
AR	388	471	525	548
AAR	383	469	499	538
SETAR	371	456	520	537
ANN	403	453	509	550
LSTAR	414	487	527	523
VAR	328	455	533	573

Note: All entries are multiplied by 1000.

The Mean Square Forecasting Error (MSFE) values of linear and nonlinear time series models are listed in Tables 3.15 and 3.16. The data in these tables indicate that threshold models with a combination of internal and external threshold variables render smaller MSFE values than linear AR model for almost all forecasting horizons. The significance of these observed differences could be further tested with MDM test. After forecasting the USA real GDP growth rate data through linear and nonlinear time series models by using different external threshold variables in levels and in first differences, MDM test is applied to the forecasting errors, and the results are presented in Tables 3.17 and 3.18.

Table 3.17 reports p-values of the MDM test by pairwise comparing the linear AR as benchmark model with nonlinear time series models for forecasting horizons 1 through 4, i.e., $h = 1, 2, 3, 4$ quarters of the real GDP growth rate data of the USA. The p-values presented in Table 3.17 indicate that TAR-Ext and LSTAR-Ext models provide more accurate forecasts than linear AR model for longer forecasting horizons by using the first differenced lag values of the monthly the USA Interest Rates, Government Securities and Bonds percent per annum as an external threshold variable. By using the USA monthly Personal Expenditures as an external threshold variable in TAR and LSTAR models, little forecasting gain is achieved by the TAR-Ext and LSTAR-Ext models compared to linear AR model for longer forecast horizons. However, the first differenced lag values of the USA Personal Expenditure and Real Gross Private Domestic Investment, both based on quarterly frequencies, are the best external threshold variables as they provide an improved forecasting performance of TAR-Ext and LSTAR-Ext models in comparison to linear AR model. This fact is more obvious for longer forecast horizons. Furthermore when comparing the threshold models with internal and external threshold variables, Table 3.17 shows that using the Real Gross Private Domestic Investment as an external threshold variable in threshold

Table 3.16: Point forecast evaluation of the linear and threshold models with external threshold variables for real GDP growth rate data: MSFE values

Models	External threshold variables (z)															
	Interest rates, Govt.Sec. and bonds				Personal expenditures monthly				Personal expenditures quarterly				Real Gross private domestic investment			
	h=1	h=2	h=3	h=4	h=1	h=2	h=3	h=4	h=1	h=2	h=3	h=4	h=1	h=2	h=3	h=4
AR	461	629	715	749	415	492	564	588	388	471	525	548	388	471	525	548
TAR-Ext-lag-D(z)	453	591	610	639	414	498	535	547	361	417	442	447	366	444	470	467
TAR-Ext-lag-(z)	486	800	919	974	414	493	538	539	366	445	508	525	356	414	450	465
LSTAR-Ext-lag-D(z)	441	548	610	605	426	499	570	566	356	415	433	433	440	495	494	507
LSTAR-Ext-lag-(z)	351	611	780	802	426	496	568	571	384	473	552	566	373	431	486	502

Note: All entries are multiplied by 1000.

Table 3.17: Point forecast evaluation of the linear and threshold models with external threshold variables for real GDP growth rate data: P-values of of the MDM test

Models Comparison	External threshold variables (z)															
	Interest rates, Govt.Sec. and bonds				Personal expenditures monthly				Personal expenditures quarterly				Real Gross private domestic investment			
	h=1	h=2	h=3	h=4	h=1	h=2	h=3	h=4	h=1	h=2	h=3	h=4	h=1	h=2	h=3	h=4
AR vs TAR-Ext-lag-D(z)	0.410	0.230	0.050	0.060	0.470	0.590	0.040	0.020	0.100	0.010	0.010	0.010	0.030	0.040	0.020	0.020
AR vs TAR-Ext-lag-(z)	0.700	0.920	0.900	0.930	0.470	0.520	0.100	0.020	0.280	0.370	0.430	0.390	0.150	0.200	0.170	0.150
AR vs LSTAR-Ext-lag-D(z)	0.270	0.020	0.040	0.050	0.660	0.570	0.550	0.290	0.070	0.020	0.010	0.010	0.900	0.750	0.030	0.040
AR vs LSTAR-Ext-lag-(z)	0.200	0.420	0.740	0.680	0.660	0.550	0.530	0.290	0.450	0.520	0.660	0.640	0.110	0.080	0.110	0.060
SETAR vs TAR-Ext-lag-D(z)	0.380	0.130	0.040	0.040	0.480	0.450	0.200	0.020	0.380	0.180	0.040	0.010	0.000	0.010	0.020	0.020
SETAR vs TAR-Ext-lag-(z)	0.570	0.770	0.830	0.920	0.480	0.390	0.220	0.010	0.440	0.420	0.440	0.440	0.020	0.130	0.190	0.170
LSTAR vs LSTAR-Ext-lag-D(z)	0.170	0.260	0.050	0.060	0.060	0.240	0.540	0.360	0.120	0.150	0.010	0.000	0.540	0.080	0.010	0.060
LSTAR vs LSTAR-Ext-lag-(z)	0.110	0.530	0.850	0.740	0.060	0.240	0.530	0.390	0.240	0.410	0.660	0.780	0.000	0.040	0.070	0.080

Note: Under the null hypothesis in pairwise MDM test, we assume that AR model has significantly smaller MSFE than the corresponding MSFE of the threshold model with external threshold variable.

models has improved forecasting gain over threshold models with internal threshold variables. Thus, it clearly provides an evidence that using the external threshold variable in threshold models leads to a much better forecasting performance than solely using the internal threshold variable in these models.

We then compared the forecasting performance of nonlinear time series models, i.e., SETAR, LSTAR, ANN, AAR, and VAR models, with linear benchmark AR model. Table 3.18 summarizes the results of the MDM test to compare the forecasting accuracy of nonlinear models with linear benchmark AR model for forecasting horizons 1 through 4. The results suggest that linear AR model produced improved forecasts over SETAR, LSTAR, and AAR models for almost all forecasting horizons except for only one case where ANN model beats the linear AR forecasts for $h = 2, 3$. In terms of out-of-sample forecasting, ANN appears to be the best ranked model among nonlinear time series models by using the internal threshold variables. The observed p-values of the MDM test shown in Table 3.18 suggest that AR model has improved forecasting performance over VAR model for 2 through 4 step-ahead forecasts. VAR model provides more accurate forecast than AR model only for $h = 1$. To sum up, the extension of threshold models with external threshold variables has an improved forecasting performance in comparison to AR and nonlinear models with internal threshold variables when modeling and forecasting the real GDP growth rate data of the USA.

Table 3.18: Point forecast evaluation of linear and nonlinear models with internal threshold variable for real GDP growth rate data: P-values of the MDM test

Models Comparison	h=1	h=2	h=3	h=4
AR vs AAR	0.380	0.470	0.160	0.200
AR vs SETAR	0.260	0.350	0.450	0.200
AR vs ANN	0.820	0.070	0.010	0.620
AR vs LSTAR	0.710	0.590	0.560	0.100
AR vs VAR	0.050	0.255	0.550	0.857

Note: Under the null hypothesis in pairwise MDM test, we assume that AR model has significantly smaller MSFE than the corresponding MSFE of the nonlinear model with internal threshold variable.

The estimation results of linear AR and threshold models for the real GDP growth rate data are presented in Tables 3.19 and 3.20. The results in Table 3.19 suggest that the constant and first lag estimated parameter values appear as significant, while the second lag estimated parameter value is insignificant. Table 3.20 shows that approximately all the estimated parameter values are significant for threshold models in both regimes. BIC slightly favors the LSTAR model, while the threshold value for SETAR and LSTAR models are significantly different from each other.

Table 3.19: AR model estimation results for real GDP growth rate data

Estimates	Values	Std.	P-values
Constant	0.450	0.078	0.000
y_{t-1}	0.335	0.062	0.000
y_{t-2}	0.091	0.062	0.143
BIC	-34.012		

Table 3.20: Threshold models estimation results for real GDP growth rate data

Models	Estimates	Low regime			High regime		
		Values	Std.	P-values	values	Std.	P-values
SETAR	Constant	0.289	0.106	0.007	0.551	0.157	0.001
	y_{t-1}	0.338	0.087	0.000	0.341	0.086	0.000
	y_{t-2}	-0.218	0.139	0.119	0.060	0.110	0.586
	Threshold value	0.540					
	BIC	-19.440					
LSTAR	Constant	-0.542	0.295	0.066	0.996	0.313	0.001
	y_{t-1}	0.266	0.123	0.031	0.085	0.141	0.547
	y_{t-2}	-0.802	0.259	0.002	0.897	0.271	0.001
	Threshold value	-0.213					
	BIC	-19.88					

3.4 The Relationship Between Economic Variables

Monthly Interest Rates Government Securities and Bonds percent per annum will affect the USA GDP growth rate in a sense that if the real interest rate is low, the costs of living are also low. This stimulates the economy because home and car loans are more affordable. If people can borrow more, they'll spend more. Low real interest rates also generally weaken the dollar. When the dollar is weak, foreign goods are more expensive, so Americans tend to buy American-made goods. This stimulates the economy even further because high demand for American goods increases employment and wages (Clark, 2008). An increase in interest rates means that consumers, i.e., the households, have to pay more to finance their consumption. Many households buy durable goods on credit, such as cars and white goods. Higher required payments discourage the consumers from buying such goods, which reduces consumption. The same goes for investments, which can be seen as consumption by firms. Higher interest rate for financing of equipment and machinery discourages firms to do investments. In other words, when interest rate increase, investments go down, since it gets more expensive to borrow money and more tempting to save money. Thus, consumption decreases, which leads to decreased demand. This keeps the prices down and inflation decreases (Kudlacek, 2008). Arnold (2013) has suggested an economic relationship between real GDP growth rate and other macroeconomic variables[5]. According to this relationship, low interest rate will enhance the domestic investment, and consumers personal expenditure will also increase. This will increase employment to create a balance between supply and demand, which in turn increases the growth of real GDP.

Thus, the economic theory suggests that the macroeconomic variables used in this study as an external threshold variables in nonlinear time series models will affect the real GDP growth rate of the USA. Hence, the inherent nonlinear dynamics in real GDP growth rate time series of the USA is well explained by the macroeconomic variables, used as an external threshold variables in threshold models. These external threshold variables are the true cause of regime switching behavior in real GDP growth rate time series of the USA.

3.5 Conclusions

In this study, we assessed the forecasting accuracy of two linear models, i.e., HAR and AR, and four types of nonlinear models, i.e., SETAR, LSTAR, AAR, and ANN, by utilizing financial and macroeconomic data. Threshold models were extended by including external threshold variables in these models, and their forecasting accuracy with linear and threshold models with internal threshold variables was compared.

Regarding financial data, almost all realized volatility time series were nonlinear according to Keenan, Tsay, and Hansen tests of linearity. A general conclusion of this study is that HAR model provides more accurate forecasts, mainly for longer forecasting horizons, and in most cases it beats forecasts from nonlinear models. Forecasts provided by HAR model are also superior to those provided by VAR model for almost all forecasting horizons and for all realized volatility time series considered in this study. However, the HSETAR model provides improved 1-step-ahead fore-

[5] The Fed's actions ripple through the economy. Higher interest rate decrease investment and consumption expenditure, increase the foreign exchange price of the dollar, which then decreases net exports a multiplier process then occurs real GDP growth and the inflation rate both slow when the Fed raises the interest rate. The opposite effects occur when the Fed lowers the interest rate. These effects are how the Fed influences the economy. The macroeconomic short run is a period during which some money prices are sticky and real GDP might be below, above, or at potential GDP. If real GDP exceeds potential GDP so there is an inflationary gap, the Fed tightens to avoid inflation. The Fed decreases the quantity of money, which raises the interest rate. The higher interest rate decreases interest sensitive components of aggregate expenditure, such as investment. The decrease in investment leads to a multiplier effect that decreases aggregate demand, thereby lowering the price level and decreasing real GDP so it equals potential GDP. If the Fed eases to avoid a recession, the reverse results occur.

casts over HAR model for DJIA. Similarly, for DJIA, the combined 1-step-ahead forecasts from HAR and HSETAR models has a much more improved forecasting gain in comparison to HAR model. Further, regarding out-of-sample forecasting performance, threshold models with external threshold variables outperform the SETAR and LSTAR-Int models for all realized volatility time series considered in this study. Thus, the inclusion of external threshold variables in threshold models has positive effects on the forecasting accuracy of the nonlinear models, even though it does not beat the HAR forecasts. There is some evidence in favor of threshold models with a combination of external threshold variables, which shows that these models produced improved forecasts in comparison to HAR model for forecasting horizon, $h = 1$ when modeling the realized volatilities data of DJIA and DAX. Among nonlinear models, TAR and LSTAR models with external threshold variables are the best forecasting models for realized volatility data. Thus, these results provided an evidence that the regime switching phenomena observed in one country's stock market is somehow the result of the asymmetric behavior present in other countries stock markets.

Similarly, by using the real GDP growth data, same conclusions as for realized volatility data can be drawn by considering linear AR model as a benchmark in this case. When modeling the real GDP growth data of the USA, nonlinear models with a combination of external threshold variables render improved forecasts in comparison to linear AR model, mainly for longer forecast horizons. Threshold models with external threshold variables provide more accurate forecasts over the same models with internal threshold variables for real GDP growth data of the USA. Among nonlinear models with internal threshold variable, ANN model seems to be the best model for forecasting the real GDP growth data. Only for $h = 1$ VAR model has an improved forecasting gain over linear AR model. In remaining cases, AR model provides superior 2 through 4 step-ahead forecasts over VAR model for real GDP growth rate data.

Hence, the conventional wisdom holds true that linear time series models are robust forecasting devices to predict the realized volatility data with few exceptions. However, for DJIA, the newly extended HSETAR model shows some improved forecasting gain over HAR model. Similarly, by introducing the external threshold variable in threshold models, it seems that there is considerable forecasts improvement in threshold models by modeling the financial (realized volatility) and economic (real GDP growth) data. Thus, this study is an initiative towards the usage of external threshold variables in nonlinear time series models in order to obtain improved out-of-sample forecasts for realized volatility and real GDP growth rate data. Overall, by predicting the real GDP growth rate and realized volatility data, there appears to be virtually some increased forecasting gain to threshold models by incorporating different financial and macroeconomic variables as an external threshold variables over linear benchmark AR model and also over SETAR and LSTAR-Int models for almost all forecast horizons. Thus, as a whole, this study suggests the use of threshold models with combination of external threshold variables to generate improved forecasts from them for financial and economic data.

Chapter 4

True or Spurious long-memory? Performance of The Heterogeneous Autoregressive and Heterogeneous Threshold Autoregressive Models

4.1 Introduction and Motivation

In recent years, research in long-memory time series has gained much attention in econometrics and statistics because of their applicability in many fields of scientific study. Among others, such types of time series are often found in physics, biology, computer sciences, hydrology, and economics. long-memory or long-range dependence implies that in general the autocorrelation of a time series decays hyperbolically rather than exponentially as like for Autoregressive Moving Average (ARMA) process. It provides an evidence that in a time series with long-memory phenomena, observations far away from each other are strongly correlated. Hurst (1951) was among the first to observe the long-range dependence in the water flow of the Nile River at the time of planning of the Aswan Dam. In many circumstances, it has been observed that economic, financial, and other types of data show high persistence in their autocorrelation, and this phenomenon has lead to the use of long-memory models. However, in many circumstances it is not clear whether the observed dependence phenomena are true long-memory in data or rather artifacts of structural breaks or deterministic trends. Granger (1966) discussed extensively the long-range dependence and its consequences in economics. Sibbertsen (2003) pointed out that long-memory in data would have strong consequences. He argued that for out-of-sample forecasting, it is important to know whether the data have real long-range dependence or any artifacts of a deterministic trend.

Mandelbrot and van Ness (1968) introduced the fractional Brownian motion as the first model for a long-memory process. In recent years, the study of structural breaks in financial and economic time series has gained much attention. Structural breaks in a given time series changes the structure of a time series model under consideration. Observing long-memory in a time series does not mean that the true data generating process contains it; rather, this phenomenon could be the result of a structural break into two or more short memory process. Thus, it provides an evidence that a time series may contain spurious long-memory. The spurious long-memory behavior is easily observed in many structural change models that include mean plus noise model of Chen and Tiao (1990), the stochastic permanent break model of Engle and Smith (1999), the sign model of Granger and Terasvirta (1999), and the structural change model of Quandt (1958). Bhattacharya *et al.* (1983) showed that adding a deterministic trend to a short-memory process can cause spurious long-memory. Gil-Alana (2008) pointed out that existence of a structural break may lead to spurious findings of long-memory. Lobato and Savin (1998) argued that in return volatility series the structural break may be responsible for the long-memory. Engle and Smith (1999) investigated the relationship between structural breaks and long-memory in context of unit root process.

In financial markets, stock market volatilities show long-memory phenomena. According to Corsi (2009), the Heterogeneous Autoregressive (HAR) model is able to capture the strong persistence and long-memory properties of time series. A number of studies that utilize the HAR model exist in the literature, including studies by Andersen, Bollerslev and Diebold (2007), Andersen, Bolleerslev and Huang (2011), Bllerslev *et al.* (2009), Corsi *et al.* (2005) and many others. The long-memory property contained by time series may be spurious, and consequently one can obtain misleading results by applying the HAR model to these time series. Liu and Maheu (2008) investigated the evidence of structural breaks in the logarithm of realized volatility by applying the HAR model. They found strong evidence of structural break in realized volatility data of S&P in early 1997 by using their estimation procedure. According to Nelson and Ploser (1982), the random shocks have permanent effects on the long run level of macroeconomic time series, and these fluctuations are not transitory. Perron and Qu (2006) analytically show how a stationary short memory process with level shifts can generate spurious long memory. Pesaran and Timmermann (2005) studied small sample properties of forecasts from autoregressive models under structural breaks. They gave different recommendations depending on the nature of the data considered.

Heterogeneous Threshold Autoregressive (HTAR) model is the extended HAR model in nonlinear framework. It is able to simultaneously model the long-range dependence and the asymmetric property of a time series dynamics. When the threshold variable in HTAR model is the lag values of the target variable itself, the resultant model is called Heterogeneous Self-Exciting Threshold Autoregressive (HSETAR) model. We applied the extended Heterogeneous Self-Exciting Threshold Autoregressive (HSETAR) model (discussed in Chapter 2) to DGP 2 through DGP 7. Thus, we assumed that HSETAR model will model the spurious long-memory time series in a more flexible and parsimonious way and will be able to differentiate it from the true long-memory type phenomena. According to our knowledge, no research studies to date have investigated the performance of HAR and HSETAR models in the presence of superior long-memory type phenomena contained by time series. The current chapter attempting to answer the question whether HAR and HSETAR models have the ability to discriminate between true long-memory and spurious long-memory. All simulated Data Generating Processes (DGP) utilized in this study are short-memory processes, i.e., Autoregressive (AR) or Self-Exciting Threshold Autoregressive (SETAR) models. However, the spurious long-memory type phenomena in these DGPs are produced through a single structural break, which occurs halfway of the total sample size. We limited our current study to a single known structural break by using simulated data from linear AR and nonlinear SETAR models. Single known structural break in Data Generating Processes (DGPs) was produced by allowing shifts in different parameter values of an AR and SETAR models, i.e., shift in intercept, shift in slope, and shift in both these parameters. We estimated the AR(p) model for each DGP containing a structural break, and the optimal lag length was selected by using Akaike Information Criteria (AIC). Afterwards, we compared the lag length, AIC, and minimum absolute characteristic root values obtained from AR(p) and HAR models.

The structure of the current chapter is as follows: in Section 2, the HAR model, its estimation procedure, and some of its properties are discussed; Section 3 elaborates the methodology adopted in this study; Section 4 presents the results and the discussion; Section 5 provides the conclusions.

4.2 Heterogeneous Autoregressive (HAR) Model

Heterogeneous Autoregressive (HAR) model introduced by Corsi (2009), is a simple cascades long-memory model used to capture the few stylized facts observed in the high-frequency financial data. This model takes care of the volatility components defined over different time period. Due to the fact that it is simple autoregressive type model by considering different components of a time series over different time horizons, thus named it Heterogeneous Autoregressive (HAR) realized volatility model. Among other characteristics, volatility distribution has fat tails, which means that extreme values with relatively high probability were observed. Also, volatility has long-memory, so a value of 'long ago', for example, 20 days ago, may still have an impact on the current or future values. According to Corsi (2009) the observed financial data fluctuated in size of the price changes at all time scales, while standard GARCH and stochastic volatility short-memory models show white noise behavior once aggregated over longer time periods (no scaling behavior). Furthermore, Corsi (2009) argues that the agents with different time horizons perceive and react to different volatility components, i.e., daily, weekly, and monthly. The standard definition of the equally spaced return series of the realized volatility over a time period of one day is:

$$RV_t^{(d)} = \sqrt{\sum_{j=0}^{M-1} r_{t-j,\triangle}^2} \tag{4.2.1}$$

where $\triangle = \frac{1d}{M}$ and $r_{t-j,\triangle} = p(t-j,\triangle) - p(t-j-1,\triangle)$ are the continuously compounded $\triangle-$frequency returns that is actually intraday returns at sample interval $\triangle$.

Denoting realized volatility by a time series y_t, the resultant HAR model has the following form:

$$y_t = c + \beta^{(d)} y_{t-1}^{(d)} + \beta^{(w)} y_{t-1}^{(w)} + \beta^{(m)} y_{t-1}^{(m)} + e_t \tag{4.2.2}$$

and the realized volatilities observed over different time horizons are defined as:

$$y_{t-1}^{(w)} = \frac{(y_{t-1} + y_{t-2} + y_{t-3} + y_{t-4} + y_{t-5})}{5} \tag{4.2.3}$$

$$y_{t-1}^{(m)} = \frac{(y_{t-1} + y_{t-2} + ... + y_{t-22})}{22} \tag{4.2.4}$$

where $y_{t-1}^{(d)}$, $y_{t-1}^{(w)}$ and $y_{t-1}^{(m)}$ are respectively the daily, weekly, and monthly observed realized volatilities. Thus, the three volatility components could be explained as the short-term traders with daily or higher trading frequencies, the medium term investors who typically rebalanced their positions weekly, and the long-term agents with a time period of one or more months. The error term in model (4.2.2) is assumed to be IID white noise process. Equation (4.2.2) can be seen as a simple autoregressive three-factor realized volatility model, where the three factors are observed volatilities at different time frequencies.

Standard Ordinary Least Squares (OLS) method of estimation can be used to estimate the parameters in HAR model, yielding consistent and normally distributed estimates. 'naïve' method of recursive forecasting can be used to forecast the HAR model. According to Corsi (2009), HAR model achieves the purpose of reproducing the empirical properties of realized volatility like long-memory, fat tails, and self similarity in a simple and parsimonious way. Thus this model estimates the unobservable latent volatility by means of realized volatility. The heterogeneous market hypothesis, which states that the reaction of the agents with different time horizons perceive, react, and cause different types of volatility components, is the key reason for the birth of HAR model. According to Corsi (2009), volatility at longer time periods has stronger influence on volatility over shorter time intervals than conversely.

4.3 Simulation Methodology

We used Monte Carlo simulations technique to investigate the performance of long memory HAR and HSETAR models, which we applied to the data containing the structural break and consequently spurious long-memory type phenomena as well. The procedures outlined in this section permit the testing for the significance of the long-memory coefficient in HAR model in a relatively general time series process, which allows one-time break in the mean of the series or its rate of growth, or both. The simulated data in this study can be considered as any type of time series containing spurious long-memory type phenomena. We simulated data from short-memory linear Autoregressive model with lag length one, i.e., AR(1) and nonlinear two regimes Self-Exciting Threshold Autoregressive model with threshold delay parameter one, i.e., SETAR(2,1,1). The current study takes into account the existence of three kinds of known structural breaks in a stationary linear AR(1) and nonlinear SETAR(2,1,1) time series models, i.e., a shift in level or intercept, a shift in slope, and a simultaneous shift in the intercept and slope. The single known structural break occurred at halfway of the total sample size[1]. Later on, we provided structural breaks of different magnitude in each DGPs with varying sample sizes, to assess its influence on the performance of HAR and HSETAR models. The main purpose of this study is to investigate the effect of long-memory coefficients in estimated HAR and HSETAR models when applied to a time series containing spurious long-memory

[1]The results remain same for the case where the structural break occurs at the first or last quarter of the total sample size for a given DGP.

type phenomena produced by a structural break.

Consider an AR(1) model for a stationary time series variable y_t:

$$y_t = \alpha_1 + \beta_1 y_{t-1} + \epsilon_t \tag{4.3.1}$$

The coefficients α_1 and β_1 respectively represent the intercept and slope parameters of an AR(1) model. The univariate two regime SETAR (2,1,1) model for a stationary time series y_t with the lag length and the threshold delay values equal to one can be presented as follows:

$$y_t = [\alpha_1 + \beta_1 y_{t-1}]\, I_t + (1 - I_t)\, [\alpha_2 + \beta_2 y_{t-1}] + \epsilon_t \tag{4.3.2}$$

where I_t is an indicator function defined accordingly by the threshold variable (y_{t-1}) relative to the threshold value (c):

$$I_t = \begin{cases} 1 & \text{if } y_{t-1} \leq c \\ 0 & \text{if } y_{t-1} > c \end{cases}$$

The α_1 and β_1 are the coefficients in regime 1, and α_2 and β_2, are the coefficients in regime 2 of the SETAR model shown in (4.3.2). The error term ϵ_t in both models is assumed to be IID and follow a standard normal distribution. Perron (1989) considered three kinds of single-known structural breaks, i.e., a shift in the level or intercept of a series, a shift in the slope, and a shift in the intercept and slope of a series simultaneously. He argued that 1929 crash and the 1973 oil price shocks have persistent effects.

By taking the work of Perron (1989) into consideration, the three types of single known structural breaks taking place in AR(1) and SETAR (2, 1, 1) models are considered in the current study. The different Data Generating Processes (DGPs) can be specified as fallows:

- No shift (DGP 1)

$$y_t = \alpha_1 + \beta_1 y_{t-1} + \epsilon_t \;\; for \;\; t = 1, 2, ..., n \tag{4.3.3}$$

- Intercept shift (DGP 2)

$$y_t = \begin{cases} \alpha_1 + \beta_1 y_{t-1} + \epsilon_t & t = 1, 2, ..., k \\ \alpha_2 + \beta_1 y_{t-1} + \epsilon_t & t = k+1, ..., n \end{cases} \tag{4.3.4}$$

- Slope shift (DGP 3)

$$y_t = \begin{cases} \alpha_1 + \beta_1 y_{t-1} + \epsilon_t & t = 1, 2, ..., k \\ \alpha_1 + \beta_2 y_{t-1} + \epsilon_t & t = k+1, ..., n \end{cases} \tag{4.3.5}$$

- Simultaneous shift in intercept and slope (DGP 4)

$$y_t = \begin{cases} \alpha_1 + \beta_1 y_{t-1} + \epsilon_t & t = 1, 2, ..., k \\ \alpha_2 + \beta_2 y_{t-1} + \epsilon_t & t = k+1, ..., n \end{cases} \tag{4.3.6}$$

We defined the following types of structural breaks in a similar way, where the process shifts from an AR(1) model to SETAR(2,1,1) model:

- Intercept shift (DGP 5)

$$y_t = \begin{cases} \alpha_1 + \beta_1 y_{t-1} + \epsilon_t & t = 1, 2, ..., k \\ (\alpha_1 + \beta_1 y_{t-1}) I_t + (1 - I_t)(\alpha_2 + \beta_1 y_{t-1}) + \epsilon_t & t = k+1, ..., n \end{cases} \tag{4.3.7}$$

- Slope shift (DGP 6)

$$y_t = \begin{cases} \alpha_1 + \beta_1 y_{t-1} + \epsilon_t & t = 1, 2, ..., k \\ (\alpha_1 + \beta_1 y_{t-1}) I_t + (1 - I_t)(\alpha_1 + \beta_2 y_{t-1}) + \epsilon_t & t = k+1, ..., n \end{cases} \tag{4.3.8}$$

- Simultaneous shift in intercept and slope (DGP 7)

$$y_t = \begin{cases} \alpha_1 + \beta_1 y_{t-1} + \epsilon_t & t = 1, 2, ..., k \\ (\alpha_1 + \beta_1 y_{t-1}) I_t + (1 - I_t)(\alpha_2 + \beta_2 y_{t-1}) + \epsilon_t & t = k+1, ..., n \end{cases} \tag{4.3.9}$$

In DGPs (5 to 7), the indicator function I_t can be defined as follows:

$$I_t = \begin{cases} 1 & \text{if } y_{t-1} \leq c \\ 0 & \text{if } y_{t-1} > c \end{cases}$$

where y_{t-1} is the threshold variable and c is the threshold value.

In all above mentioned DGPs, we assumed that the single known structural break has taken place at halfway ($k = \frac{n}{2}$) of the total sample of size n. The effect of the magnitude of structural break on the performance of HAR and HSETAR models was evaluated by allowing structural breaks of different magnitudes in DGPs (2 to 3 and 5 to 6) in the form of intercept and slope shift. We generated data from DGPs (1 to 7) with a sample size n = 500, 1000, 2000, and 5000 to evaluate its effects on the performance of HAR and HSETAR models. Table 4.1 lists the specifications of all DGPs used for simulations purposes. DGP 1 is a simple AR(1) model without any structural change, while DGP 2 to DGP 7 are time series models in which different parameter values shift due to single structural break. In DGP 1, no structural break has taken place, and it could be considered as a benchmark DGP. Similarly, DGP 4 is an AR model, where both the intercept and slope parameters have shifted from (α_1, β_1) to (α_2, β_2) at the time period $k + 1$.

We first estimated two separate AR(1) models by respectively using the logarithmically transformed realized volatility time series of Down Jones Industrial Average (DJIA) for pre-crises (2006-2007) and crises (2008-2009) periods. To make our simulations empirically realistic, these parameter values were used in all DGPs. The structural breaks that we allow in all DGPs were even smaller in magnitude than the ones observed in AR(1) model parameters estimated for logarithmically transformed realized volatility time series of DJIA for pre-crises (2006 - 2007) and crisis (2008 - 2009) period. Thus, we assessed the HAR and HSETAR model performance in more challenging conditions. In Table 4.1, DGP 1 is an AR model without any structural breaks, while for DGP 2 to DGP 7 at least one parameter changes halfway through the total sample size and thus provide a single known structural break and consequently spurious long-memory. For example, in DGP 2 the intercept coefficient (α_1) in AR(1) model changes from $\alpha_1 = -3.60$ to $\alpha_2 = -2.60$. Similarly, in DGP 7, the AR(1) process shifts to SETAR(2,1,1) process halfway through the total sample size, and the intercept and slope coefficients change from (α_1, β_1) to (α_2, β_2), i.e., (-3.60, 0.65) to (-2.60, 0.85). In DGP 7, the coefficients of AR and SETAR models in regime 1 are the same but they change in the second regime. Therefore, DGP 7 provides low level of long memory type phenomena, which is shown by estimated HAR and HSETAR models. In DGPs 5 to 7, the threshold delay parameter and the threshold value are

Table 4.1: Parameter values used for simulation

Process	DGPs	Types of structural break	Parameters shift			
			From	To	From	To
			α_1	α_2	β_1	β_2
AR(1)	DGP 1	No Shift	-3.60	-3.60	0.65	0.65
	DGP 2	Intercept shift	-3.60	-2.60	0.65	0.65
	DGP 3	Slope shift	-3.60	-3.60	0.65	0.85
	DGP 4	Intercept and slope shift	-3.60	-2.60	0.65	0.85
AR(1) to SETAR(2,1,1)	DGP 5	Intercept shift	-3.60	-2.60	0.65	0.65
	DGP 6	Slope shift	-3.60	-3.60	0.65	0.85
	DGP 7	Intercept and slope shift	-3.60	-2.60	0.65	0.85

respectively set to 1 and -10 in SETAR(2,1,1) model. Further, in all DGPs the error term (ϵ_t) is assumed to follow a standard normal distribution. The simulated realizations with different sample sizes and 500 replications are generated from each of the DGPs listed in Table 4.1. The first 200 and 600 simulated realizations respectively from AR and SETAR models are discarded to obtain stable realizations. In DGP 5 to DGP 7, we imposed a condition that each regime of SETAR model must contains at least 15 percent of the total observations.

4.4 Results and Discussion

We estimated the HAR and two regime HSETAR models for each draw under each DGP with 500 replications. This study can be considered as a miss-specification type test of HAR and HSETAR models in the presence of spurious long-memory type phenomena contained by time series. Although Goerg (2010) argued that the structural break that occurs over time is responsible for producing long-memory in time series,in fact it does not. In this study, we considered a single known structural break that occurs halfway of the total sample size in a stationary time series and producing spurious long-memory type phenomena in it. According to the Heterogeneous Autoregressive (HAR) model specifications, the long-memory property of a time series can be captured by a coefficient corresponding to time series aggregated over one month, i.e., $\beta^{(m)}$. For a given time series, if an estimated value of $\beta^{(m)}$ appears as the least insignificant, then it provides evidence that HAR model shows long-memory in its results, which is in fact spurious. Secondly, we applied the extended Heterogeneous Self-Exciting Threshold Autoregressive (HSETAR) model to each realization of all DGPs listed in Table 4.1. According to the HSETAR model specifications, in each regime of HSETAR model, one can capture the long memory type phenomena in a time series by a coefficient corresponding to time series aggregated over one month, i.e., $\alpha^{(m)}$ or $\beta^{(m)}$.

The data presented in Table 4.2 shows the percentage of the total time series out of 500 replications for which the p-values (P) of the coefficient $\beta^{(m)}$ in HAR model are less than or equal to our presumed significance levels, i.e., 0.01, 0.05, and 0.10, and thus turned out significant for different DGPs and sample sizes. It can be seen from Table 4.2 that for most of the time series across all DGPs and sample sizes the estimated values of $\beta^{(m)}$ appeared as least insignificant. However, for DGP 1, for almost all sample sizes the $\beta^{(m)}$ values are insignificant. This fact is true because DGP 1 does not contain any structural breaks and thus no spurious long-memory, as shown by the HAR model. In DGP 2, where the structural break occurs when the intercept shifts from -3.60 to -2.60, for $n = 500$, approximately 94% of the time series out of 500 replications have p-values of $\beta^{(m)}$ smaller or equal to 0.01, 98% time series have p-values of $\beta^{(m)}$ smaller or equal to 0.05, and 99% time series have p-values of $\beta^{(m)}$ smaller or equal to 0.10. Similarly, for DGP 6, the structural break occurs when the slope parameter shifts from 0.65 to 0.85, for $n = 2000$, 90% of the time series out of 500 replications have p-values of $\beta^{(m)}$ smaller or equal to 0.01. Hence these results provide the evidence that the estimated HAR model show long-memory, which is in fact spurious. For DGPs 5 to 7, the observed long-memory is somehow weaker than those observed for DGPs 2 to 4. The possible reason for weaker long-memory observed in DGPs 5 to 7, can be the equality of the coefficients in the AR(1) model and in the first regime of SETAR(2,1,1) model. Similarly, it can be further seen from Table 4.2 that the percentage of the time series for which the p-values of $\beta^{(m)}$ in the estimated HAR model are smaller or equal to 0.01, 0.05, and 0.10, and thus tuned out as significant, is greater for DGP 5 than that observed for DGP 6.

Table 4.2 shows that with an increasing sample size the percentage of time series for which the estimated $\beta^{(m)}$ values turn out as significant also increases. Thus increasing sample size produces strong spurious long-memory in simulated time series across all DGPs, and it is well captured by an estimated HAR model. Hence, positive relationship is observed between spurious long-memory type phenomena in time series with increasing sample size. Furthermore, it can be seen that each DGP with sample of size $n = 5000$, provides the highest percentage of time series for which the long memory coefficient in the estimated HAR model is significantly different from zero.

We provided different magnitudes of single known structural break in DGPs 2 to 3 and DGPs 5 to 6 and simulated 2000 realizations with 500 replications from each of these DGPs to evaluate its effect on the performance of HAR model. After estimating the HAR model for each realization of these DGPs, significance of the estimated values of $\beta^{(m)}$ was assessed through its p-values. Table 4.3 shows that for all DGPs, as the magnitude of the structural breaks increases, the percentage of the

Table 4.2: Significance of the $\beta^{(m)}$ in HAR model

DGPs	Sample Size (n)	$P \leq 0.01$	$P \leq 0.05$	$P \leq 0.10$
DGP 1	500	1	6	12
	1000	1	5	10
	2000	1	5	12
	5000	1	4	9
DGP 2	500	94	98	99
	1000	100	100	100
	2000	100	100	100
	5000	100	100	100
DGP 3	500	25	53	71
	1000	91	98	99
	2000	100	100	100
	5000	100	100	100
DGP 4	500	58	81	89
	1000	97	99	99
	2000	100	100	100
	5000	100	100	100
DGP 5	500	67	83	89
	1000	95	98	99
	2000	99	100	100
	5000	100	100	100
DGP 6	500	24	45	58
	1000	53	77	84
	2000	90	97	98
	5000	95	99	100
DGP 7	500	8	21	33
	1000	23	41	53
	2000	53	72	80
	5000	91	98	99

Note: Table entries represent the percentage time series out of 500 replications for which the coefficient corresponding to a time series aggregated over one month in estimated HAR model appears significantly different from zero at significance levels 0.01, 0.05, and 0.10.

time series for which the p-values of $\beta^{(m)}$ are less than or equal to 0.01, 0.05, and 0.10 also increases. For example, in DGP 2, when the intercept shifts by magnitude of 0.25 units, i.e., from -2.0 to -2.25, approximately 31%, 53%, and 66% of the time series out of 500 replications have p-values of $\beta^{(m)}$ respectively smaller or equal to 0.01, 0.05, and 0.10. However, this percentage goes up by at least 48% when the magnitude of the structural break increases from 0.25 to 0.30. A slight increase in the magnitude of slope parameter in DGP 3 provides a significant increase in the percentage of time series for which the estimated values of $\beta^{(m)}$ appeared as significant. For example, in DGP 3, when the slope parameter shifts by a magnitude of 0.05 units, i.e., from 0.30 to 0.35, approximately 10%, 26%, and 37% of the time series out of 500 replications have p-values of $\beta^{(m)}$ respectively smaller or equal to 0.01, 0.05, and 0.10. But, when the slope parameter increases by a magnitude of 0.10 units, approximately 95%, 99%, and 99% of the time series out of 500 replications have p-values of $\beta^{(m)}$ respectively smaller or equal to 0.01, 0.05, and 0.10. Thus, a slight increase in the magnitude of the slope parameter provides a strong long-memory phenomena in time series captured by the estimated HAR model. An interesting point noted in DGPs 2 to 6 is that for 100% of the time series, out of 500 replications have p-values smaller than 0.01 when the intercept and slope parameters are respectively shifted by a magnitude of 0.50 and 0.20. Thus, it can be concluded from he results presented in Table 4.3 that as the magnitude of the structural break increases, the percentage of the time series for which the $\beta^{(m)}$ values appeared as significant in estimated HAR model also increases. Hence, the above results supported our assertion that for any magnitude structural break provided in simulated DGPs, HAR model fails to distinguish between true and spurious long-memory in it.

The estimation results of the two regimes HSETAR model for DGP 2 to DGP 7 are presented in Table 4.4. The data presented in Table 4.4 show the percentage of total time series out of 500 replications for which the p-values (P) of $\alpha^{(m)}$ and $\beta^{(m)}$ in the estimated HSETAR model are less than or equal to our conventionally presumed significance levels. The first period lag values of time series aggregated over one month, i.e., $y_{t-1}^{(m)}$, are used as an threshold variable in the HSETAR model. Table 4.4 shows that in a two regimes HSETAR model, for a low percentage of time series out of 500 replications, the coefficients $\alpha^{(m)}$ and $\beta^{(m)}$ appear significantly different from zero across all DGPs. This percentage tends to decrease with an increasing sample size. For example, in DGP 2 with a sample of size 500, 6%, 12%, and 21% of the time series out of 500 replications have p-values of $\alpha^{(m)}$ less than or equal to 1%, 5%, and 10% significance levels; these values in the high regime of HSETAR model are 9%, 15%, and 23%. These values indicate that for a high percentage of time series, the long memory coefficients in the two regimes HSETAR model are least significant.

Hence, this study signifies the ability of HSETAR model to detect the spurious long-memory type phenomena in time series produced by a single known structural break. Thus, the HSETAR model parsimoniously differentiates between true and spurious long-memory type phenomena in all simulated DGPs. The results of the current study show that the utilization of the first lag time series averaged over the last 22 values as a threshold variable in the HSETAR model enables it to differentiate between true and spurious long-memory type phenomena.

Table 4.3: Structural break of different magnitudes and significance of the $\beta^{(m)}$ in HAR model

DGPs	Parameters shift				P-values of $\beta^{(m)}$		
	α_1	α_2	β_1	β_2	$P \leq 0.01$	$P \leq 0.05$	$P \leq 0.10$
DGP 2	-2.00	-2.25	0.30	0.30	31	53	66
	-2.00	-2.30	0.30	0.30	60	76	84
	-2.00	-2.35	0.30	0.30	84	94	97
	-2.00	-2.50	0.30	0.30	100	100	100
DGP 3	-2.50	-2.50	0.30	0.35	10	26	37
	-2.50	-2.50	0.30	0.40	95	99	99
	-2.50	-2.50	0.30	0.45	99	100	100
	-2.50	-2.50	0.30	0.50	100	100	100
DGP 5	-2.00	-2.25	0.30	0.30	33	56	67
	-2.00	-2.30	0.30	0.30	64	83	88
	-2.00	-2.35	0.30	0.30	84	94	96
	-2.00	-2.50	0.30	0.30	100	100	100
DGP 6	-2.50	-2.50	0.30	0.35	10	35	36
	-2.50	-2.50	0.30	0.40	95	99	99
	-2.50	-2.50	0.30	0.45	98	100	100
	-2.50	-2.50	0.30	0.50	99	100	100

Notes: The results presented in this table show the affect of structural break of different magnitudes in simulated DGPs on the performance of HAR model to detect the spurious long-memory type phenomena.

Table entries represent the percentage time series out of 500 replications for which the coefficient corresponding to a time series aggregated over one month in estimated HAR model appears significantly different from zero at significance levels 0.01, 0.05, and 0.10.

We also estimated the AR(p) and HAR model for DGP 2 to 7 that contains one known structural break. We compared the median AIC and the median minimum absolute values of the characteristic roots of the HAR and AR models. The optimal lag length in AR model was selected through AIC. Table 4.5 presents the median values of AIC and the median minimum absolute values of the characteristic roots of the HAR and AR models for each DGP across 500 replications. In a similar way, the values of the median lag length of the AR(p) model are also given in this table. After estimating the AR(p) model for each of these DGPs, we could conclude from Table

Table 4.4: Significance of the α^m and β^m in HSETAR model

DGPs	Sample Size (n)	Low regime (α^m)			High regime (β^m)		
		$P \leq 0.01$	$P \leq 0.05$	$P \leq 0.10$	$P \leq 0.01$	$P \leq 0.05$	$P \leq 0.10$
DGP 2	500	6	12	21	9	15	23
	1000	4	7	12	3	8	14
	2000	1	7	14	2	8	14
	5000	2	8	15	4	10	15
DGP 3	500	5	13	21	3	10	14
	1000	9	16	23	1	7	13
	2000	6	12	16	2	6	12
	5000	1	4	8	1	6	14
DGP 4	500	9	20	19	6	16	20
	1000	6	11	18	3	7	14
	2000	1	7	11	1	7	12
	5000	1	6	10	3	6	11
DGP 5	500	9	19	25	7	17	28
	1000	7	12	18	11	19	26
	2000	4	9	16	8	15	21
	5000	4	10	18	9	19	27
DGP 6	500	5	14	22	5	13	22
	1000	4	11	18	4	11	19
	2000	4	12	18	3	11	18
	5000	6	18	27	8	19	28
DGP 7	500	5	14	24	7	17	25
	1000	6	17	22	6	15	22
	2000	4	12	20	8	20	28
	5000	2	11	17	14	17	28

Note: Table entries represent the percentage time series out of 500 replications for which the coefficients corresponding to a time series aggregated over one month in low and high regimes of the estimated HSETAR model appears significantly different from zero at significance levels 0.01, 0.05, and 0.10.

Table 4.5: AR(p) and HAR models in-sample fitting comparison

DGPs	Sample Size	AR(p) model			HAR model	
		Median			Median	
		p	AIC	Min(abs(roots)))	AIC	Min(abs(roots)))
DGP 2	500	4	-6420.56	1.331	-4819.86	1.064
	1000	5	-12757.40	1.262	-9760.43	1.065
	2000	8	-25409.80	1.163	-19603.10	1.061
	5000	16	-63031.40	1.081	-48783.50	1.062
DGP 3	500	9	-9198.79	1.134	-7632.86	1.061
	1000	11	-18402.00	1.111	-15417.90	1.058
	2000	14	-36771.20	1.092	-30979.00	1.057
	5000	19	-91623.40	1.061	-77443.90	1.057
DGP 4	500	11	-9210.17	1.133	-7630.65	1.062
	1000	10	-18394.20	1.124	-15401.50	1.059
	2000	13	-36664.10	1.095	-30870.80	1.059
	5000	19	-91649.80	1.062	-77409.80	1.058
DGP 5	500	4	-6455.10	1.272	-4865.94	1.069
	1000	5	-12845.20	1.243	-9844.79	1.064
	2000	8	-25620.70	1.152	-19815.40	1.062
	5000	15	-63375.40	1.094	-49144.80	1.064
DGP 6	500	4	-9172.42	1.415	-7537.12	1.117
	1000	5	-18368.40	1.403	-15310.80	1.114
	2000	6	-36588.20	1.352	-30697.60	1.103
	5000	10	-91216.90	1.201	-76917.40	1.099
DGP 7	500	4	-9319.88	1.454	-7674.65	1.132
	1000	5	-18632.20	1.434	-15567.30	1.131
	2000	5	-37171.30	1.414	-31257.90	1.122
	5000	8	-92474.60	1.272	-78073.40	1.112

Note: AR(p) and HAR models are estimated for each DGP with different sample sizes and then median values of the lag length (p), AIC, and absolute characteristic roots are calculated across 500 replications and results are reported in above table.

4.5 that the optimal lag length of the AR model selected through AIC is higher than the lag length used in the data generating process, and it tends to increase with an increasing sample size. The high lag length values in the estimated AR model show long-memory in DGP 2 to DGP 7.

According to the observed median values of AIC for each sample size and DGP, the AR model performs better than the HAR model. Across all DGPs, the median values of AIC for both models tend to decrease with an increasing sample size. Therefore, regarding in-sample fitting, the AR model dominates the HAR model across all DGPs and sample sizes. After estimating the HAR and AR models for DGP 2 to DGP 7 for sample of size 2000 across 500 replications, we reported the median minimum absolute characteristic root values of the AR and HAR models in Table 4.5. It shows that the median minimum absolute characteristic root values for the HAR model are smaller than the AR model for almost all DGPs. Thus, it shows more persistence in parameter estimates of the HAR model. In all DGPs, as the sample size increases, the median value of the minimum absolute characteristic roots of the AR model tends towards one, providing evidence of an increased persistence in the AR model parameters with an increasing sample size. For the HAR model, the median of the minimum absolute characteristic root values remain similar with an increasing sample size for all DGPs.

The histograms of the lag lengths of the AR model, selected through AIC for different DGPs for $n = 2000$ are shown in Figure 4.4.1. It shows that the estimated lag length values for the AR model are greater than one, and for DGP 2 and DGP 3 the true lag length values are approximately 22. Thus, the true model for DGP 2 and DGP 3 is AR(22) and for DGP 4 to 7, the lag length values selected through AIC in estimated AR(p) model are also quite high, providing evidence of the presence of long-memory type phenomena in these DGPs. The box plots of the minimum absolute characteristic root values for the AR and HAR models for all DGPs and a sample of size 2000 are shown in Figure 4.4.2. This figure suggests that the median values of minimum absolute characteristic roots of the HAR model are slightly smaller than for the AR model in all DGPs. All minimum absolute characteristic roots values of the estimated HAR and AR models are greater than one, providing evidence of the stationarity of these models. The box plots of AIC for the AR and HAR models for different DGPs and a sample of size 2000 are shown in Figure 4.4.3. This figure also shows the minimum values of AIC are observed for an AR model across all DGPs providing evidence of an improved in-sample fitting gain over the HAR model.

Based on these results, we come to a conclusions that for simulated DGPs 2 to 7, the estimated HAR model shows strong persistence in its parameters and poor in-sample fitting performance in comparison to the AR model.

The current results provide the evidence that the HAR model shows long-memory in its estimation results, which is in fact spurious, and does not have the ability to distinguish between true and spurious long-memory. The estimation results of the HAR model shows strong long-memory in its parameter values for almost all simulated DGPs, which is in fact spurious. Hence, the HAR model is unable to differentiate between true and spurious long memory in any given time series and consequently may yield misleading results, which can cause serious distortions in policy implementation.

To conclude, the HAR model may provides unreliable parameter estimates and consequently poor in-sample-fitting and worst out-of-sample forecasting performance for time series containing spurious long-memory type phenomena. Overall, this

Figure 4.4.1: Histograms of the lag length, $AR(p)$ model for $n = 2000$

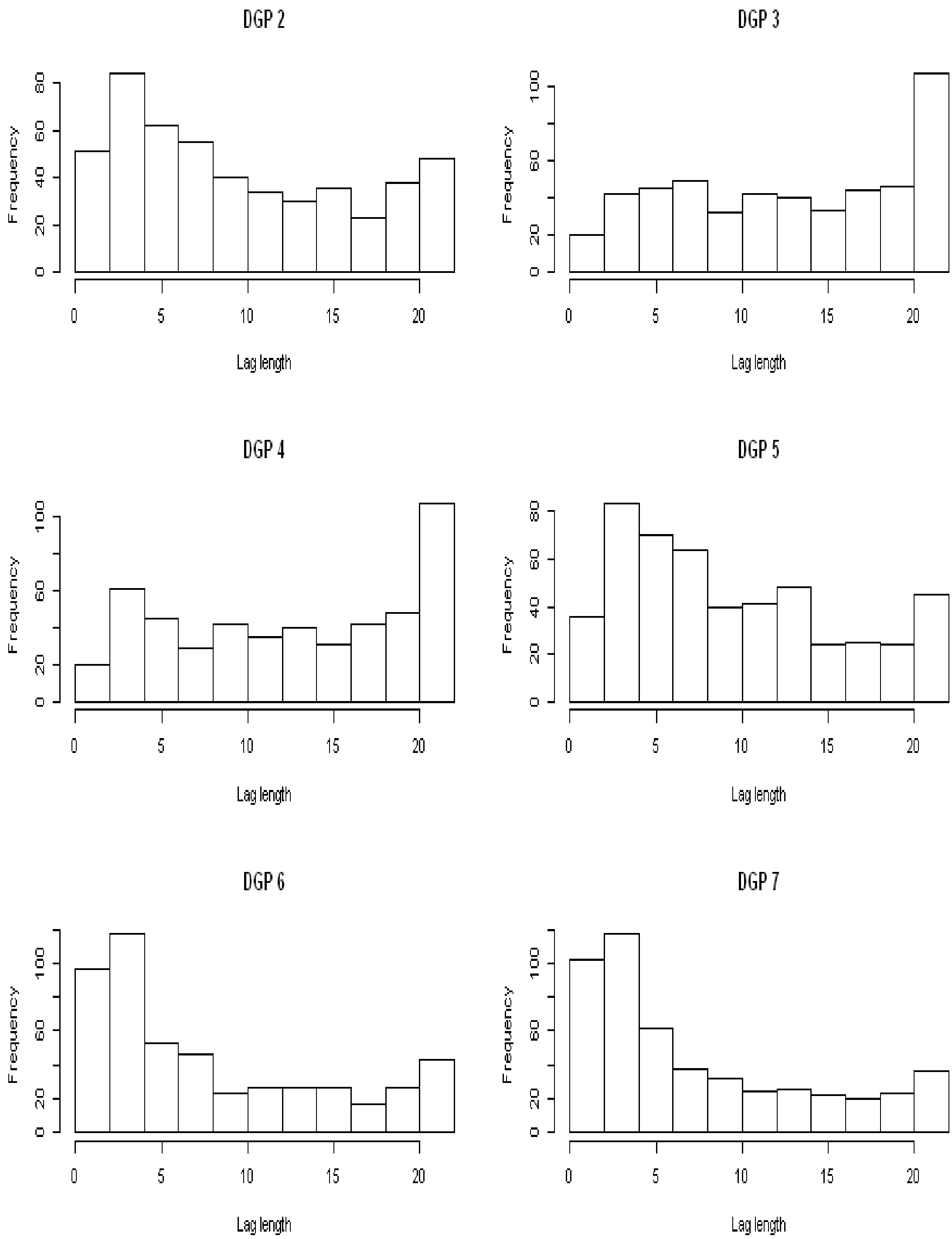


Figure 4.4.2: Box plots of the minimum absolute characteristic root values of the HAR and $AR(p)$ models for $n = 2000$

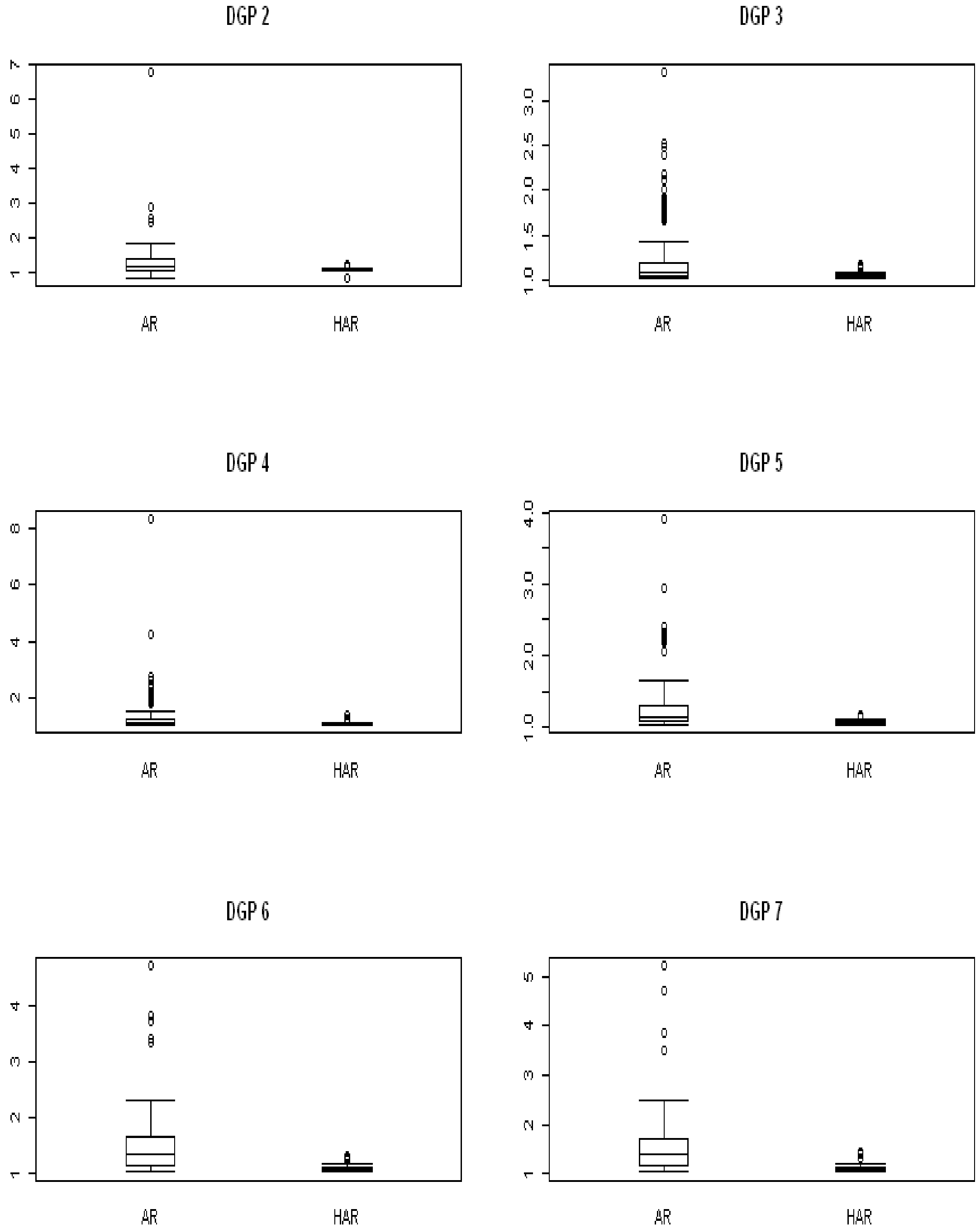
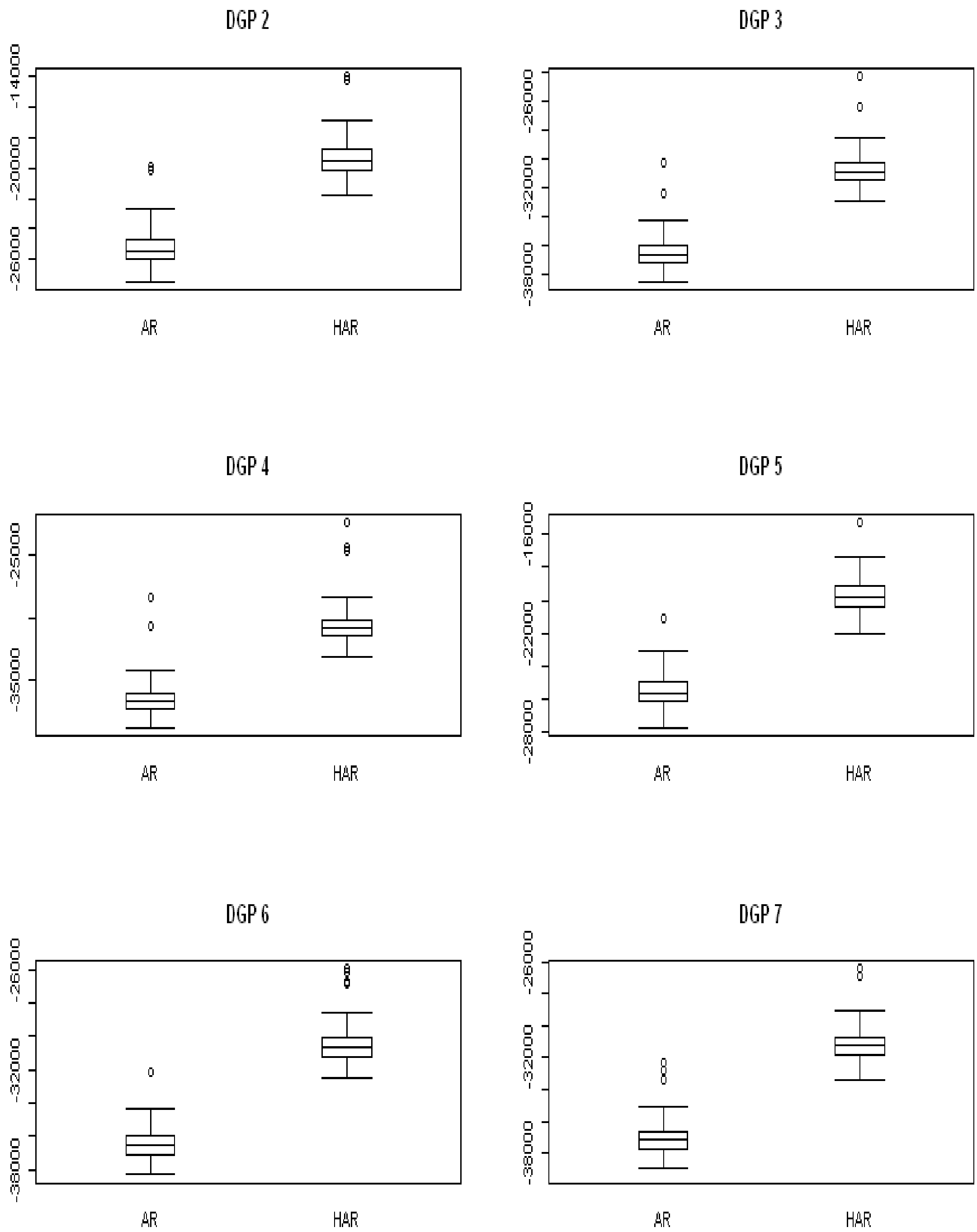


Figure 4.4.3: Box plots of the AIC for HAR and $AR(p)$ models for $n = 2000$



study provides the evidence that the HAR model can be easily confused with spurious long-memory, which can lead an investigator to draw unrealistic results. However, the HSETAR model parsimoniously detects the spurious long-memory type phenomena in a given time series. It is therefore suggested that under such circumstances it is better to use the HSETAR model to avoid misleading results.

4.5 Conclusions

It is known that the HAR model serves the purpose of modeling the long-memory behavior of time series in a very simple and parsimonious way. In this study, the performance of the HAR and HSETAR models was assessed in the presence of spurious long-memory type phenomena contained by a time series. The main focus in this study was to assess whether the HAR and HSETAR models have the ability to distinguish between true and spurious long-memory generated by a structural break. We observed that for all simulated DGPs, except DGP 1 over 500 replications, the estimated coefficient, $\beta^{(m)}$ in the HAR model appeared to be significantly different from zero. It provides the evidence that the HAR model shows long-memory in its estimation results, which is in fact spurious. Hence, the HAR model is unable to differentiate between true and spurious long-memory type phenomena contain by time series.

The results show that over 500 replications across all DGPs, the percentage of time series for which the estimated values of $\beta^{(m)}$ in the HAR model appeared significantly different from zero increases as the sample size increases. The results provide the evidence that with an increasing sample size or with an increasing magnitude structural break, the spurious long-memory captured by the HAR model also increases for all simulated DGPs containing a structural break.

We also applied the two regimes HSETAR model to those simulated DGPs that contain spurious long-memory type phenomena. The results indicate that for a high percentage of time series, the long-memory coefficients in the estimated two regimes HSETAR model are not significantly different from zero. Thus, the HSETAR model parsimoniously differentiates between true and spurious long-memory for all simulated DGPs. The current results clearly convey a message that if a time series contains spurious long memory type phenomena, the HSETAR model is able to parsimoniously detect it.

We further show that regarding in-sample fitting the AR model dominates the HAR model for all simulated DGPs. The median minimum absolute characteristic root values of the AR and HAR models are close to one. This provided an evidence of high persistence in the parameter estimates of these models across all DGPs and sample sizes.

The analytical findings contained in this study show that neglecting the spurious long-memory contained by a time series may lead the investigator to conclude that the series under investigation has also significant long-memory through the application of the HAR model, but in fact it does not. However, the HSETAR model correctly detects the spurious long-memory type phenomena contained by time series. Therefore, in such circumstances, it is suggested to use the nonlinear HSETAR model rather than the HAR model to avoid unrealistic results.

Chapter 5

Linear and Nonlinear Time Series Modeling and Forecasting of Seismic Data

5.1 Introduction and Motivation

The application of statistical methods and models to seismic data in order to assess the seismic hazard and risk in a particular area is of great importance. The recent occurrence of earthquakes has shown that most of the buildings are erected in and around the epicentral areas, therefore putting more emphasis on earthquake research. It has been assumed until now by the scientific communities that earthquakes are almost impossible to predict. In fact, scientists cannot predict the exact location, magnitude, and time of an earthquake occurring in a particular area. However, several earthquake scientists are conducting research on earthquake precursors and forecasting studies.

Statistical models and methods play a very useful role in seismology and involve the estimation of seismic hazard, assessment of earthquake prediction schemes, and quantification of uncertainties in the estimates of earthquake locations or magnitudes. However, there is still much room left for the application of statistics in seismology. More emphasis on research in this field is needed to answers the questions such as: which assumptions may be reasonable, which statistical models or methods are more useful under conditions appropriate to seismological problems. It has been observed that seismologists describe the earthquakes by using deterministic and physical models. However, applied statisticians have been increasingly exploring the use of stochastic models for modeling earthquake behavior and forecasting. It has been observed in previous studies that in the field of earthquake seismology little attention has been given to the utilization of the time series models.

Before carrying out this study, some know-how about seismology was needed. From the Greek seismos means earthquakes and logos means science, so seismology means the scientific study of Earth and earthquakes. Seismic data contain the earthquake occurrence time, its date, locations, magnitudes, and depth. Seismic waves, i.e., body and surface waves, produced during an earthquake are the energy caused by the sudden breaking of the rock within the Earth or an explosion. After recording these seismic waves, it is possible to calculate the above mentioned earthquake parameters by using different techniques. The main purpose of the seismic analysis is to understand the internal structure of the Earth, particularly the aspects related to the earthquake phenomena. Furthermore, an earthquake seismologist is also interested in ways of mitigating the earthquake effects. For earthquake mitigation, seismic hazard analysis is a very important field of research.

Applications of statistical models and methods in earthquake science with the goal of improving our knowledge of how the Earth works is called Statistical Seismology. Earthquake forecasts are very important in terms of estimating hazard and managing emergency system (D'Amico (2012)). According to a book edited by D'Amico (2012) interoccurrence times between successive earthquakes are not independent random variables, but have "long-term memory". Vere-Jones (2010) discussed the importance of statistics in seismology as "The role of statistical modeling ideas in seismology has increased to the stage where more serious attention should be given to the possibility of incorporating some serious statistical courses in the undergraduate and postgraduate statistical programmes. Better later than earlier, I think, and with the emphasis on statistical modeling, not on cook-book recipes". Similarly, Vere-Jones (2006) emphasis the used of statistics in the field of seismology as "Seismology in particular has presented statistics with an array of challenging and fruitful problems: epicenter location and velocity modeling; the interpretation of seismograms and a host of related problems in time-series analysis; the analysis of building response and its uses in earthquake resistant design and the assessment of risk from seismic and other geophysical events". Ogata (1988, 1998) introduced the Epidemic-Type-Aftershock-Sequence (ETAS) model based on Gutenberg Richter law and modified Omori law. David Harte's "Statistical Seismology Library" (Harte, 1998) provided the statistical and graphical R routines for modeling earthquake catalogs. Michael and Wiemer (2010) argued that the earthquakes are not independent; rather, they interact and cluster with respect to space and time. Amei *et al.* (2012) fitted the ARIMA model to empirical recurrence rates (ERRs) of earthquake counts that occurred all around the world. They argued that besides the use of ARIMA model for long-term earthquake prediction, it also serves as a linking bridge between point processes and the classical time series. In this study, they predict 12 large earthquakes in next six year worldwide by using the best fitted model. Martienz *et al.* (2005) investigated the elapsed times and distances between consecutive seismic events that occurred in the Southern Iberian Peninsula from 1985 to 2000. They argued that beside satisfactory performance of power law model for earthquake elapsed times and distances, in some cases a fit with a Weibull distribution for elapsed times performs better. Wang *et al.* (2014) used the Self-Exciting Threshold Poisson Autoregression model for the major earthquake counts worldwide. Kanaori (1997) showed positive relationship between earthquake magnitudes and elapsed times between earthquakes.

Yang *et al.* (1995) studied the nonlinear behavior of seismic activities in Northern China by using threshold autoregressive and exponential autoregressive models. They applied both of these models to the earthquake magnitudes sequence and reached to a conclusion that the seismicity of different areas possesses different active levels. Morales-Esteban *et al.* (2010) proposed pattern recognition based on K-means algorithm to forecast earthquakes with magnitude greater or equal to 4.5 by utilizing Spanish seismic temporal data. They arrived at the conclusion that the proposed technique performs well, especially when the earthquake uncertainty in its occurrences is taken into account. Juan and Guzman-Vargas (2013) presented the visibility graph method by using the earthquake magnitude data from Italy, Southern California, and Mexico. Feng *et al.* (1997) carried out a dynamic modeling for earthquake magnitude series form China and Japan region for prediction purposes. They argued that the underlying model performs well in terms of prediction. Tepei (1998) estimated earthquake intensity around year 2000 in Guangdong area by using the threshold autoregressive model. Ling-Yi and De-Fu (1985) established the Self-Exciting Threshold Autoregressive model, i.e., SETAR (2, 4, 3) for the time series of the maximal earthquake magnitude occurring in China since 1901. Kahraman *et al.* (2012) modeled the earthquake magnitude sequence through SETAR model for the purpose of out-of-sample prediction. Panakkat and Adeli (2007) used neural network models for earthquake magnitude prediction by incorporating multiple seismic indicators in the model. Ho (2008) modeled the empirical recurrence rate time series for volcanism through ARIMA models for Russia and predicted volcanic activity.

Earthquake prediction has a vital importance for the safety of human beings. Earthquakes are nonlinear and complicated dynamic processes and cannot be completely described by using deterministic models. Therefore some stochastic model is needed to characterize the earthquakes phenomena. Studies that extend the application of nonlinear time series models in the field of seismology are hard to find in the literature. This study is the first of its kind, and we hope that this study will enable us to understand the underlying seismic processes and provide a way of utilizing the time series models in earthquake seismology. Similarly, this study is an attempt to bridge the analytical gap between seismology and statistics by using many potential linear and nonlinear time series models to understand and predict the seismic time series of Hindu Kush region of Pakistan. We also extended the Threshold Vector Autoregressive (TVAR) model by characterizing the nonlinear dynamics of seismic data. It has been observed that TVAR has mostly been used in the econometric literature (see, for example, Sims and Zha (2006), and Hubrich and Tetlow (2012)). The time series models considered in current study are linear Autoregressive (AR), Vector Autoregressive (VAR), and five nonlinear time series models, i.e., Threshold Autoregressive (TAR) or Self-Exciting Threshold Autoregressive (SETAR), Logistic Smooth Transition Autoregressive (LSTAR), Artificial Neural Network (ANN), Threshold Vector Autoregressive (TVAR) models, and one nonparametric Additive Autoregressive (AAR) model. By utilizing the seismic data from the Hindu Kush region of Pakistan for a given time series model, a very important and basic question in this study is: how adequately does the model describe the data (observations or simulations) to which it applies? Which time series model is the most plausible among the competing models and could be utilized for forecasting purposes. This study is an attempt to answer the above stated questions.

Autoregressive Conditional Duration (ACD) models, discussed in Chapter 2, are used to model the irregularly spaced financial duration data. In literature, the ACD model has mostly been used to model and forecasts the econometric data. For example, Tsay (2009) applied the ACD model to the adjusted transaction durations of the IBM stock from November 1 to November 7, 1990 and to the daily range of the log price of Apple stocks from January 4, 1999 to November 20, 2007; he arrived at the conclusion that the ACD model parsimoniously captured the dynamics of these time series.

In this study, we also extended the applications of the ACD model to model and forecast the seismic data from the Hindu Kush region of Pakistan. As we know that the basic ACD model works only with positive quantity, we also modeled and forecasted raw seismic time series to achieve this assumption besides modeling and forecasting the logarithmically transformed seismic data. We estimated and forecasted the ACD, VAR, and nonlinear AAR, SETAR, ANN, and LSTAR models and compared their 1 through 4 step-ahead out-of-sample forecasting performance with the benchmark AR model. We chose the forecasting horizons from 1 through 4 to assess the short-term horizon ($h = 1$, 2) and the long-term horizon ($h = 3$, 4) forecasting performance of these models. In the second section of this chapter, we present the modeling and forecasting results of raw seismic data.

The structure of the current chapter is as follows. In Section 5.2, we discuss the dynamics of earthquakes. In Section 5.3, we discuss the Hindu Kush region and its geological structure. In Section 5.4, we elaborate the seismic data, earthquake magnitude conversion, calculation of the time series of consecutive distances and elapsed times, and modeling and forecasting methodology of the time series models considered in this study. Section 5.5 contains the modeling and forecasting results and a discussion on the time series models in the context of logarithmically transformed and raw seismic data. Section 5.6 contains the conclusions.

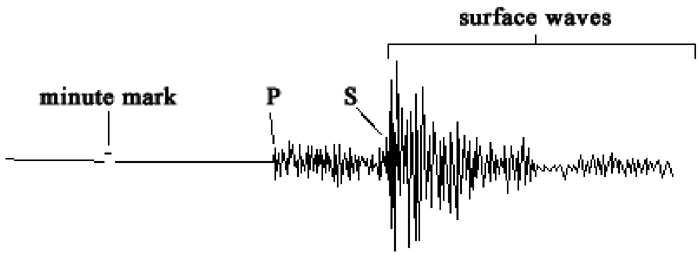
5.2 Earthquakes and Seismic Waves

An earthquake is the result of a sudden release of energy in the Earth's crust that creates different types of seismic waves. Earthquakes are usually caused when rock underground suddenly breaks along a fault. This sudden release of energy causes the seismic waves that make the ground shake. When two blocks of rock or two plates are rubbing against each other, they stick a little. They don't just slide smoothly; the rocks catch on each other. The rocks are still pushing against each other, but not moving. After a while, the rocks break because of all the pressure that's built up. When the rocks break, the earthquake occurs. During the earthquake and afterward, the plates or blocks of rock start moving, and they continue to move until they get stuck again. The spot underground where the rock breaks is called the focus of the earthquake. The place right above the focus (on top of the ground) is called the epicenter of the earthquake.

Seismology is the scientific study of earthquakes and seismic waves that travel through and around the earth. A seismologist is a scientist who studies phenomena related to earthquakes. Seismic waves are the waves of energy caused by the sudden breaking of rock within the earth or an explosion. They seismic waves are recorded on seismographs. When an earthquake occurs, two main types of seismic waves are produces: body waves and surface waves. Body waves traveling through the interior of the earth, are of two types: P waves and S waves (Figure 5.2.1).

The first kind of body wave is the P wave or primary wave. This is the fastest kind of seismic wave, and, consequently, the first to arrive at a seismic station. The P wave can move through solid rock and fluids, like water or the liquid layers of the

Figure 5.2.1: A typical seismogram



Source: UPSeis, an educational site for budding seismologists.

earth. It pushes and pulls the rock it moves through just like sound waves push and pull the air. P waves are also known as compressional waves, because of the pushing and pulling they do. Subjected to a P wave, particles move in the same direction that the wave is moving in, which is the direction that the energy is traveling in. The second type of body wave is the S wave or secondary wave, which is the second wave you feel in an earthquake. An S wave is slower than a P wave and can only move through solid rock, not through any liquid medium. It is this property of S waves that led seismologists to conclude that the Earth's outer core is a liquid. S waves move rock particles up and down, or side-to-side, perpendicular to the direction that the wave is traveling in (the direction of wave propagation).

Surface waves traveling only through the crust, are of a lower frequency than body waves, and are easily distinguished on a seismogram as a result. Though they arrive after body waves, it is surface waves that are almost entirely responsible for the damage and destruction associated with earthquakes. This damage and the strength of the surface waves are reduced in deeper earthquakes. The first kind of surface wave is called a Love wave, named after A.E.H. Love, a British mathematician who worked out the mathematical model for this kind of wave in 1911. It's the fastest surface wave and moves the ground from side-to-side. The other kind of surface wave is the Rayleigh wave, named for John William Strutt, Lord Rayleigh, who mathematically predicted the existence of this kind of wave in 1885. A Rayleigh wave rolls along the ground just like a wave rolls across a lake or an ocean. Because it rolls, it moves the ground up and down, and side-to-side in the same direction that the wave is moving. Most of the shaking felt from an earthquake is due to the Rayleigh wave, which can be much larger than the other waves[1].

[1] Source: The text contained by this section is taken from UPSeis, an educational website for budding seismologists, http://www.geo.mtu.edu/UPSeis/waves.htm.

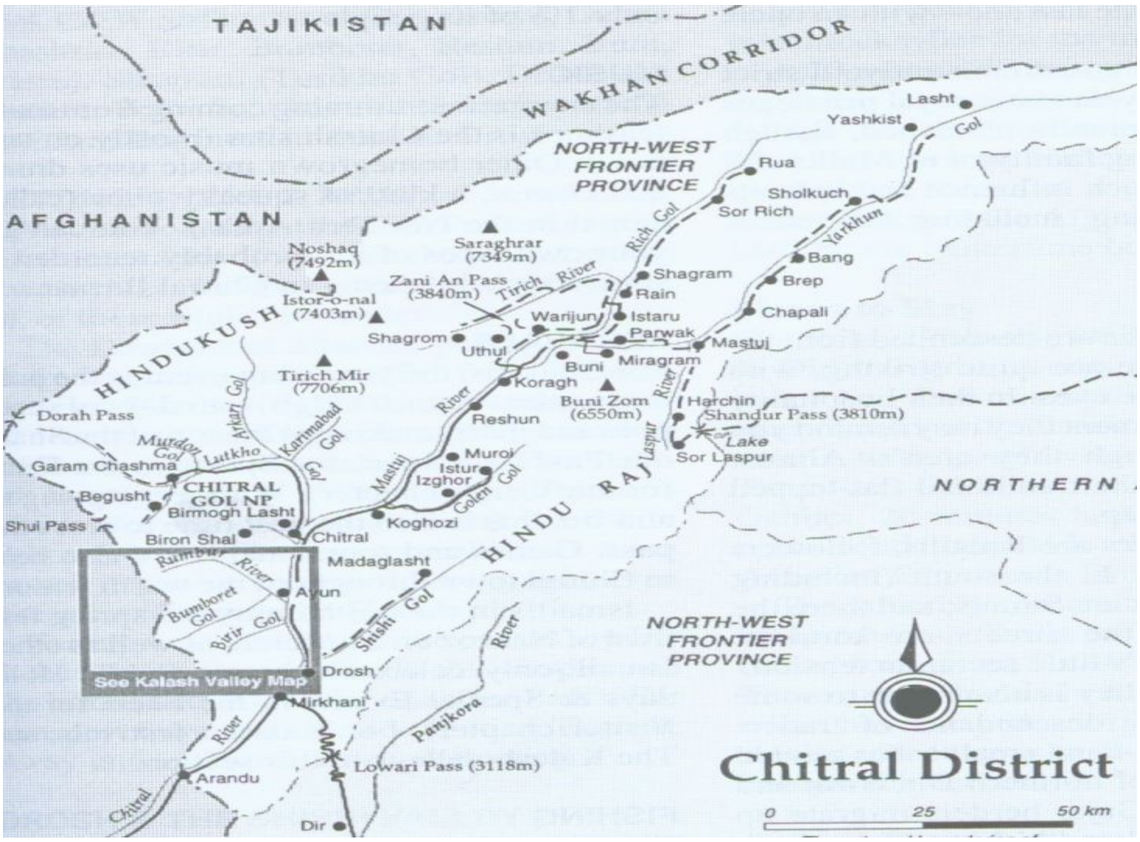
5.3 The Hindu Kush Region

The Hindu Kush is a mountain system of Central Asia, extending 400 miles south-west from the Pamir Knot. These mountains roughly form the boundary between Pakistan and Afghanistan. It is the westernmost extension of the Pamir Mountains, the Karakoram Range, and is a sub-range of the Himalayas. In the east, the mountains are round and wide, and rise only to 18,000 feet, low by central Asian standards. The mountains of the Hindu Kush system diminish in height as they stretch westward: toward the middle, near Kabul, they extend from 4,500 to 6,000 meters, in the west, they attain heights of 3,500 to 4,000 meters. The average altitude of the Hindu Kush is 4,500 meters. The Hindu Kush system stretches about 966 kilometers (km) laterally, and its median north-south measurement is about 240 km. Only about 600 km of the Hindu Kush system is called the Hindu Kush Mountains. Most of the high summits rise from subsidiary ridges south of the main crest. As the mountains stretch further west and south, they gradually get smaller and spread out as dusty barren hills into the central Afghanistan. Compared with most other areas in Central Asia, many of the high glaciated peaks are conveniently accessible, and the weather, though hot in the summer and snowy in the winter, is generally predictable and stable. The Hindu Kush region has provided a new area of scientific investigation for a number of international geoscientists, especially during the last decade, due to its peculiar structural position and in the context of development of the plate tectonics hypothesis.

Nowroozi (1971) discussed the broad features of the Hindu Kush seismic zone. The concentrated zone of the seismic activity shows an east-west trend along the great circle latitude of $36.22°N$. The zone extends from $70°E$ to $71.5°E$ meridian, and the focal depths exceed 200 km. In the vicinity of $71.5°E$, the general $N45°E$, $S45°W$ trend of the Pamir appears; this zone is broader than the east-west trend of Hindu Kush and extends approximately from $36°N$, $70°E$ to $38°N$, $73°E$. The focal depths are generally shallower than 200 km and usually fall between 100 and 200 km. At the intersection of the two zones, the seismic activity scatters and is more complex, and the focal depths usually exceed 200 km. Nowroozi (1972) made some detailed profiles across the Hindu Kush seismic zone. The east-west alignment of deeper earthquakes is sharply defined. The zone has a length of about 120 km, a width of 25 to 30 km, and a majority of the earthquakes occur within 175 to 200 km in a vertical slab. These earthquakes are assumed to be in a slab within the mantle, similar to those beneath the island arcs. Presently, however, this slab is overlain by continental crust, and it is difficult to assess whether the slab was originally made of continental or oceanic lithosphere. In the island-arc structure, the earthquakes are in a slab that is made of oceanic lithosphere. Thus, it is suggested that the Hindu Kush deep seismic zones are within pieces of oceanic lithosphere which are probably remnants of the Tethys Ocean floor.

In the Hindu Kush region intermediate earthquakes occur to a maximum depth of 300 km, but are mostly in the range from 50 - 200 km and are mainly distributed within a zone 200 km deep dipping northward. However, in the Pamir zone it inclines southward forming a 'V' distribution. The 'V' form distribution focal depth may be related to a bent plate structure caused by the collision of the Indian subcontinent with the Asian continent. The continuing north-south horizontal compressive force between the two continents causes earthquakes to take place frequently in this region. Because of its complex seismic characteristics as well as typical tectonic locations, this region probably deserves more attention than it has received so far. Map of the Hindu Kush region of Pakistan is shown in Figure 5.3.1. The Hindu Kush range is mostly located in Khyber Pakhtunkhwa (its old name was North-West Frontier Province (NWFP)) of Pakistan and along the border with the Nuristan and Badakhshan provinces of Afghanistan.

Figure 5.3.1: Map of the Hindu Kush region of Pakistan



source: http://www.caingram.info/Pakistan/htm/Hindu_kush_map.htm

5.4 Methodology

In this study, we applied the linear and nonlinear time series models[2] to the earthquake data from the Hindu Kush region of Pakistan. This study is an attempt to bridge the analytical gap between statistics and seismology. From historical studies it emerges that mostly financial data have been used to compare the forecasting performance of linear and nonlinear time series models. This study, however, utilizes seismic data that includes earthquake magnitude, consecutive earthquake elapsed times, consecutive earthquake distances, and monthly earthquake counts time series. Earthquakes are complex phenomena and nonlinear in nature, therefore we used four types of nonlinear time series models, i.e., Threshold Autoregressive (TAR), Logistic Smooth Transition Autoregressive (LSTAR), Artificial Neural Network (ANN), and Threshold Vector Autoregressive (TVAR) models, and one nonparametric nonlinear Additive Autoregressive (AAR) model, and compared their out-of-sample forecasting performance with the linear benchmark Autoregressive (AR) model. TAR model with internal threshold variable is referred to as Self-Exciting Threshold Autoregressive (SETAR) model. We also compared the Vector Autoregressive (VAR) forecasts with those from the univariate AR model. Secondly, we used these models to model and forecast the monthly number of the earthquakes in the Hindu Kush region of Pakistan. The raw earthquake counts data does not follow the normal distribution; however, the logarithmic monthly earthquake counts follow normal distribution (shown below in histogram) and thus encourages us to model this time series by using the linear and nonlinear time series models.

The TAR model allows the model parameters to change according to the threshold variable, while the LSTAR models are useful for a time series that changes its behavior through a smooth transition variable. For linear and nonlinear models, 1 through 4 step-ahead forecasts were obtained by respectively using the naïve and Bootstrapped method of forecasting. To obtain a single out-of-sample forecasted value from nonlinear time series models, the forecasted values were averaged over 1000 bootstrapped replications. To compare the out-of-sample forecasts from linear and nonlinear models, Mean Square Forecasting Error (MSFE) and pairwise Modified Diebold-Mariano (MDM) tests were used. All these models, their procedures of estimation and forecasting, and MDM test are elaborately discussed in Chapter 2. Expanding window technique of model estimation was adopted to obtain recursively 1 through 4 step-ahead forecasts, i.e., $h = 1, 2, 3, 4$. We re-estimated each of the above models by adding one more observation to the estimation period and then the

[2]All these models are elaborately discussed in chapter two.

first four step-ahead forecasts were recursively generated by using these estimated models. Each model is specified only once, and the same model specifications are used throughout the expanding window technique of model estimation. The first estimation window used 60% of the total observations of the four seismic time series (i.e., earthquake magnitudes, consecutive elapsed times, consecutive distances, and monthly earthquake counts) from the Hindu Kush region of Pakistan. Following that, 1 through 4 step-ahead forecasts were generated from each of the models used in this study. In the next step, in each iteration one more observation were added to the estimation period under the expanding window scheme, and 1 through 4 step-ahead forecasts were again generated. This process was continued until our last estimation window utilized $(n - max(h))$ observations, where n and h respectively represent the total sample size of a time series and forecasting horizon. Each time series model was specified only once by using Bayesian Information Criterion (BIC), and then the same specifications of the model were used throughout the expanding window scheme.

5.4.1 Data

In our study, the data set consists of 1692 earthquakes that occurred in the Hindu Kush region from January 1975 to April 2013 with moment, body, and surface wave magnitude ranging from 4.5 to 7.3. To get homogenous earthquake magnitudes, all body and surface waves magnitudes respectively denoted by (m_b) and (M_s) were converted into moment magnitudes (M_w) by using the empirical relation provided by Scordilis (2006). After converting the (m_b) and (M_s) to (M_w), (M_w) ranged from 4.8 to 7.3. These earthquakes have focal depth less than 300 km in a region, which extends from longitudes $66^\circ E$ to $73^\circ E$ and latitudes $34.5^\circ N$ to $37^\circ N$. The data used in the present study were taken from the website of the United States Geological Survey (USGS)[3]. The USGS data set contains the occurrence time of the earthquakes, its magnitudes, latitudes, longitudes, and depth. The consecutive distances time series is constructed by using the formula provided in the next section. Similarly, the time series of consecutive elapsed times were generated by subtracting the previous earthquake occurrence time from the recent one. The time series of consecutive elapsed times and distances between earthquakes were respectively represented in days and kilometers. Since earthquake magnitudes calculations are based on the logarithmic scale, all time series used for modeling and predicting purposes in this study are logarithmically transformed except for the earthquake magnitudes. Similarly, logarithmically transformed time series were used as an external threshold variable in the TAR and LSTAR models. Threshold models with external threshold variables are

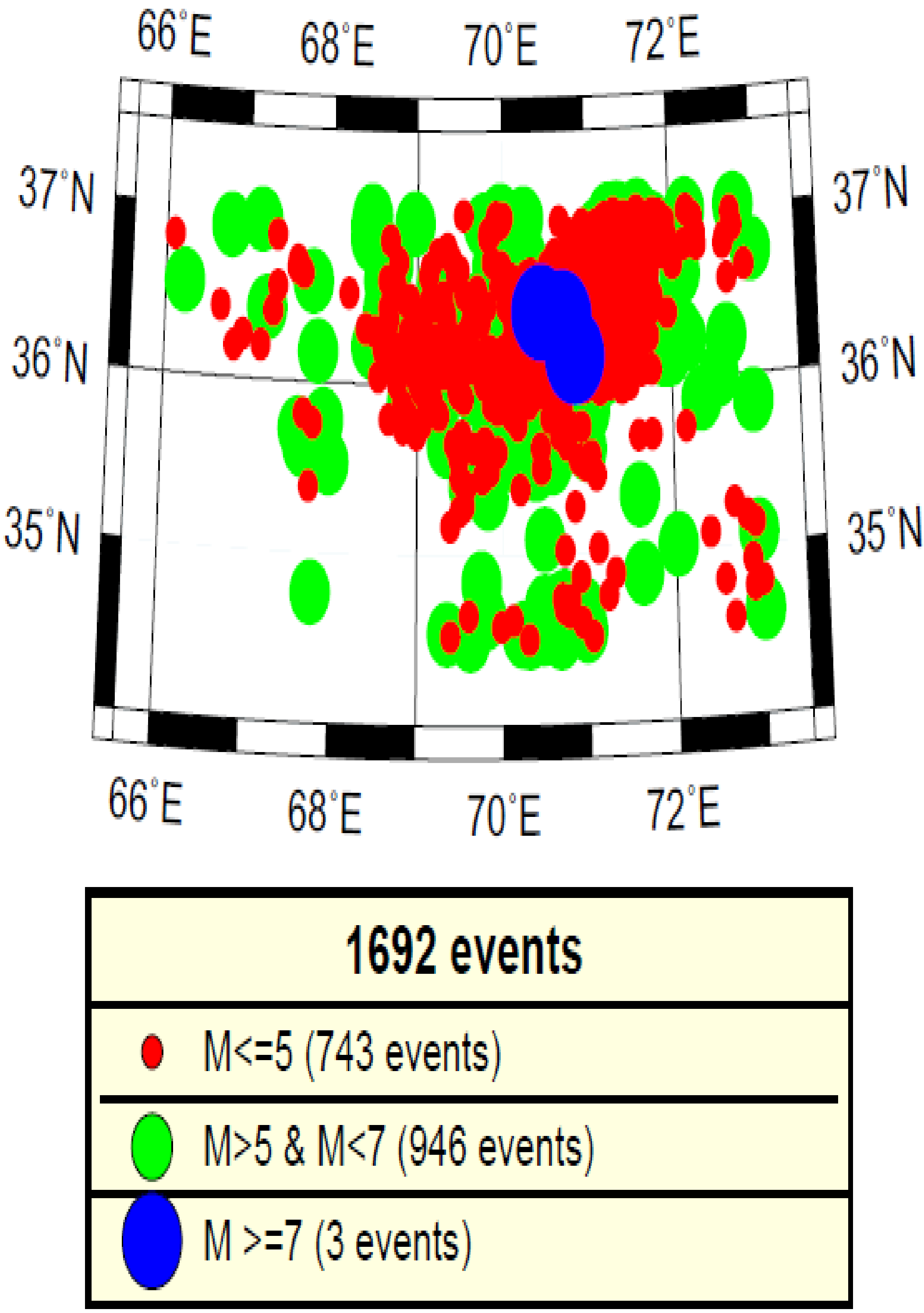
[3]http://earthquake.usgs.gov/earthquakes/search/

referred as TAR-Ext( ) and LSTAR-Ext( ). The time series of monthly earthquake counts were constructed by counting the monthly number of earthquake occurrences in the study period. All aftershocks from the data set are eliminated in order to avoid misleading results (Ferraes, 1968 and 1973). The map of the seismic activity observed in the study area is shown in Figure 5.4.1.

5.4.2 Magnitude Conversion and Constructing the Consecutive Distances Time Series

There are several magnitude scales widely used to measure the size of an earthquake, each one based on measuring a specific type of seismic waves in a specified frequency range with a particular instrument. Richter or local magnitude (M_L) was the first widely used instrumental magnitude scale to be applied in the USA (Richter, 1935). The different magnitude scales that have been used to measure the size of an earthquake are body wave magnitudes (m_b), local magnitudes (M_L), and surface wave magnitudes (M_S). Surface and body wave magnitude calculations are based on surface and body waves produced by an earthquake. Scordilis (2006) pointed out that the main problem of M_L, m_b, and M_S scales are that they do not behave uniformly for all magnitude ranges. Another problem is that the M_L, M_s, and m_b scales exhibit saturation effects at different levels for large earthquakes. The M_s and m_b respectively saturate for earthquakes with $M_s > 8.0$ and $m_b > 6.5$. Thus, both the body and surface wave magnitudes underestimate the true earthquake magnitudes. Consequently, the body and surface wave magnitude scale progressively underestimates the actual energy released. These limitations led Kanamori (1977) and Hanks and Kanamori (1979) to propose a new magnitude scale, namely moment magnitude M_W. The moment magnitude scale is the most recently developed scale and is fundamentally different from the earlier scales. Further, Scordilis (2006) argues that M_w does not saturate, since it is directly proportional to the logarithm of seismic moment, resulting in a uniform behavior for all magnitude ranges. Scordilis (2006) provides an empirical study by converting different magnitude scales to the most reliable and useful scale of magnitude, the moment magnitude (M_w). The data set that we extracted from the USGS website contains body wave magnitudes (m_b), surface-wave magnitude (M_s), and moment magnitude (M_w *or* M). To get the homogeneity in magnitude time series, we converted the m_b and M_s into M_w. The conversion of m_b and M_s into M_w was carried out by using the empirical global magnitude converting relationship provided by the Scordilis (2006). All earthquake magnitudes considered in the current study are moment magnitudes (M_w).

Figure 5.4.1: Seismic activity in the Hindu Kush region from 1975 to 2013



The moment magnitude is calculated from the seismic moment using the relation of Hanks and Kanamori (1979):

$$M_w = \frac{2}{3}log_{10}(M_0) - 6.0$$

where M_0 is the seismic moment defined as:

$$M_0 = DA\mu$$

where D is the average displacement over the entire fault surface, A is the area of the fault surface, and μ is the average shear rigidity of the faulted rocks. The value of D is estimated from the observed surface displacements or from displacements on the fault plane reconstructed from instrumental or geodetic modeling. A is derived from the length multiplied by the estimated depth of the ruptured fault plane, as revealed by surface rupture, aftershock patterns, or geodetic data. The method thus assumes that the rupture area is rectangular.

The following empirical relationship has been provided by Scordilis (2006) for magnitude conversion:

$$M_w = 0.67M_s + 2.07 \quad for \;\; 3.0 \leq M_s \leq 6.1$$

and

$$M_w = 0.99M_s + 0.08 \quad for \;\; 6.2 \leq M_s \leq 8.2$$

The above empirical relationship between M_w and M_s holds for shallow earthquakes, i.e., earthquakes having focal depth less than 70 km.

Similarly, the empirical relationship between M_w and m_b, provided by Scordilis (2006) is given as follows:

$$M_w = 0.85m_b + 1.03 \quad for \;\; 3.5 \leq m_b \leq 6.2$$

The relationship between earthquake magnitude and the corresponding energy released can be described as: an earthquake that measures 5.0 on the Richter scale has a shaking amplitude 10 times larger than one that measures 4.0, and corresponds to a 31.6 times larger release of energy. Vassiliou and Kanamori (1982) provide the magnitude energy relationship. According to them, the energy release of an earthquake, which closely correlates to its destructive power, scales with the 3/2 power of the shaking amplitude. Thus, a difference in magnitude of 1.0 is equivalent to a factor of 31.6 ($= (10^{1.0})^{(3/2)}$) in the energy released, a difference in magnitude of 2.0 is equivalent to a factor of 1000 ($= (10^{2.0})^{(3/2)}$) in the energy released. Similarly a difference in magnitude of 0.1 is equivalent to the factor of 1.4 ($= (10^{0.1})^{(3/2)}$).

The consecutive elapsed times time series is calculated by subtracting the previous earthquake occurrence time from the current one. The time series of consecutive elapsed times between successive earthquakes are represented in days. Similarly, the time series of consecutive distances between successive earthquake is generated by using the haversine formula (great circle distance formula) (Sinnott, 1984):

$$Consecutive\ Distances\ (D_t)\ km = Radius\ of\ the\ earth * 2 * \lambda$$

where

$$\lambda = \arcsin\sqrt{\sin^2\left(\frac{\theta}{2}\right) + \cos(Lat1) * \cos(Lat2) * \left(sin^2\left(\frac{\delta}{2}\right)\right)}$$

and

$$\theta = Lat1 - Lat2,\ \ \delta = Long1 - Long2$$

In the above formula, the Lat1 and Long1 represents the latitude and longitude in degrees of the earthquake that occurs first and second with respect to time. Lat2 and Long2 represents the latitudes and longitudes in degrees of the second successive earthquake. We used 6371 km as the Earth's radius. In mathematics and cartography, a Great Circle Distance is the shortest path between two points on the surface of a sphere (and we assumed that the Earth is a perfect sphere, even though it is not in reality). To calculate the Great Circle Distance between points, we first calculated the spherical central angle between the two points (λ) and then multiplied that angle (in radians) by the radius of the Earth. As Microsoft Excel works with radians, to calculate the consecutive distances, we first converted the latitudes and longitudes in radians.

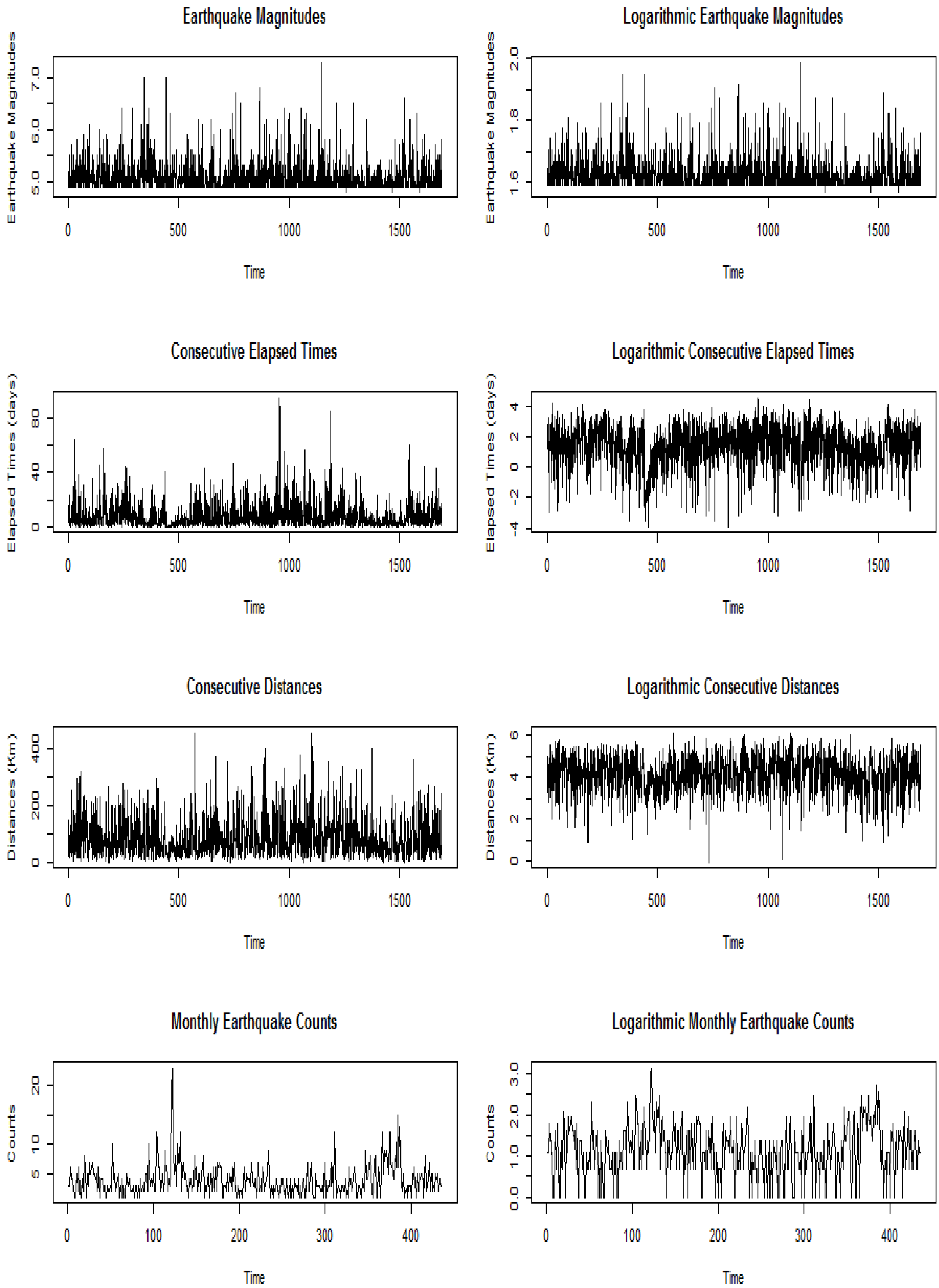
5.5 Results and Discussion

5.5.1 Analysis of Logarithmically Transformed Seismic Data

The visual display of the seismic data, i.e., earthquake magnitudes, consecutive elapsed times, consecutive distances, and monthly earthquake counts are respectively shown in first, second, third, and fourth panel of Figure 5.5.1. The first and second column of Figure 5.5.1 respectively shows the raw and logarithmically transformed seismic data. We observed asymmetric behavior of all four seismic time series shown in Figure 5.5.1. This figure suggests that time series of consecutive distances between earthquakes appeared to be the most asymmetric time series. For earthquake magnitudes time series, similar persistent is observed for the whole time period. The earthquake magnitudes time series shows that a larger earthquake is followed by smaller ones. Similarly for consecutive elapsed times and consecutive distances time series, the second half of the total sample is more persistent than the first one. The last panel of Figure 5.5.1 represents the monthly earthquake counts from the Hindu Kush region of Pakistan. We observed that for smaller magnitude earthquakes the consecutive elapsed times and consecutive distances are also short. For monthly earthquake counts time series, a high magnitude spike is observed between sample point 121 and 123, i.e., the corresponding time period is February, 1986. The mean earthquake magnitude, consecutive elapsed times, consecutive distances, and monthly earthquake counts are 5.10, 8.25 days, 90.92 km and 3.88, respectively. The right panel of Figure 5.5.1 shows the logarithmically transformed seismic time series and shows a fair amount of persistence as observed in the raw data.

The Autocorrelation Function (ACF) and the Partial Autocorrelation Function (PACF) of the four seismic time series are shown in Figure 5.5.2. From the observed ACFs plotted in Figure 5.5.2, consecutive elapsed times and monthly earthquake counts provide significant ACF even at higher lags, while for consecutive distances time series the ACF is the only significant one at first lag. The ACF for earthquake magnitudes time series is insignificant at all lags proving no long-memory in series. However, as ACF measures the linear correlation between a time series and its lag values and does not account for nonlinear relationship, therefore we modeled this time series by using the time series models. The significant ACF of the consecutive elapsed times, consecutive distances, and monthly earthquake counts time series led us to utilize time series techniques for modeling and forecasting purposes. For the time series of consecutive elapsed times, most of the PACF values are significant at higher lags, while only a few significant PACF values are observed for consecutive

Figure 5.5.1: Raw and logarithmically transformed seismic data



earthquake distances and monthly earthquake counts time series.

Table 5.1 shows that the skewness and kurtosis are high for raw seismic time series, so we logarithmically transformed them, thus providing much closer values of these statistics to the normal distribution. This is depicted in the histogram of four seismic time series provided in Figure 5.5.3. For this reason, it seemed reasonable to model the logarithmically transformed seismic time series, except earthquake magnitudes, rather than raw series. Since the earthquake magnitude is a logarithmic measure of seismic energy and already on the logarithmic scale, we modeled and predicted the earthquake magnitudes time series, not the logarithmically transformed ones.

The results of summary statistics for the four seismic time series, i.e., earthquake magnitudes (M_t), consecutive elapsed times (E_t), consecutive distances (D_t), and monthly earthquake counts (MC_t) are provided in Table 5.1. The observed sample size for each of the time series (M_t), (E_t), and (D_t) is 1602, while for (MC_t) time series the sample size is equal to 436. As stated above, the raw seismic data provided high values of variance, skewness, and kurtosis, indicating high dispersion and departure from normality. In this part of study, all seismic time series considered were logarithmically transformed, except for the earthquake magnitudes, to achieve the purpose of normality and low dispersion. After logarithmic transformation, the variances of elapsed times, consecutive distances, and monthly earthquake counts were quite low, thus indicating low dispersion. Skewness and kurtosis for the logarithmically transformed seismic time series were also low and much closer to normality than the raw ones. For earthquake magnitude time series, the minimum and maximum earthquake magnitude values are 4.8 and 7.3, respectively. The minimum and maximum consecutive elapsed times between earthquakes are 0.02 and 94.94 days, respectively. The minimum and maximum consecutive distances between earthquakes are 0.95 and 455.80 kilometers (km), respectively. Similarly, the minimum and maximum number of the monthly earthquake counts are 1 and 23, respectively. The histograms of the raw and logarithmically transformed seismic data are shown in Figure 5.5.3. Raw seismic time series are positively skewed, while the logarithmically transformed seismic data, except the earthquake magnitudes, has much closer shape to follow the normal distribution.

The Augmented Dicky-Fuller (ADF) test was used for testing the stationarity properties of the seismic time series. The linearity property of the seismic time series was tested using Tsay (1986) and Keenan (1985) tests of nonlinearity. The p-values of the stationarity and linearity tests are presented in Table 5.2. The zero p-values pro-

Figure 5.5.2: ACFs and PACFs of seismic data

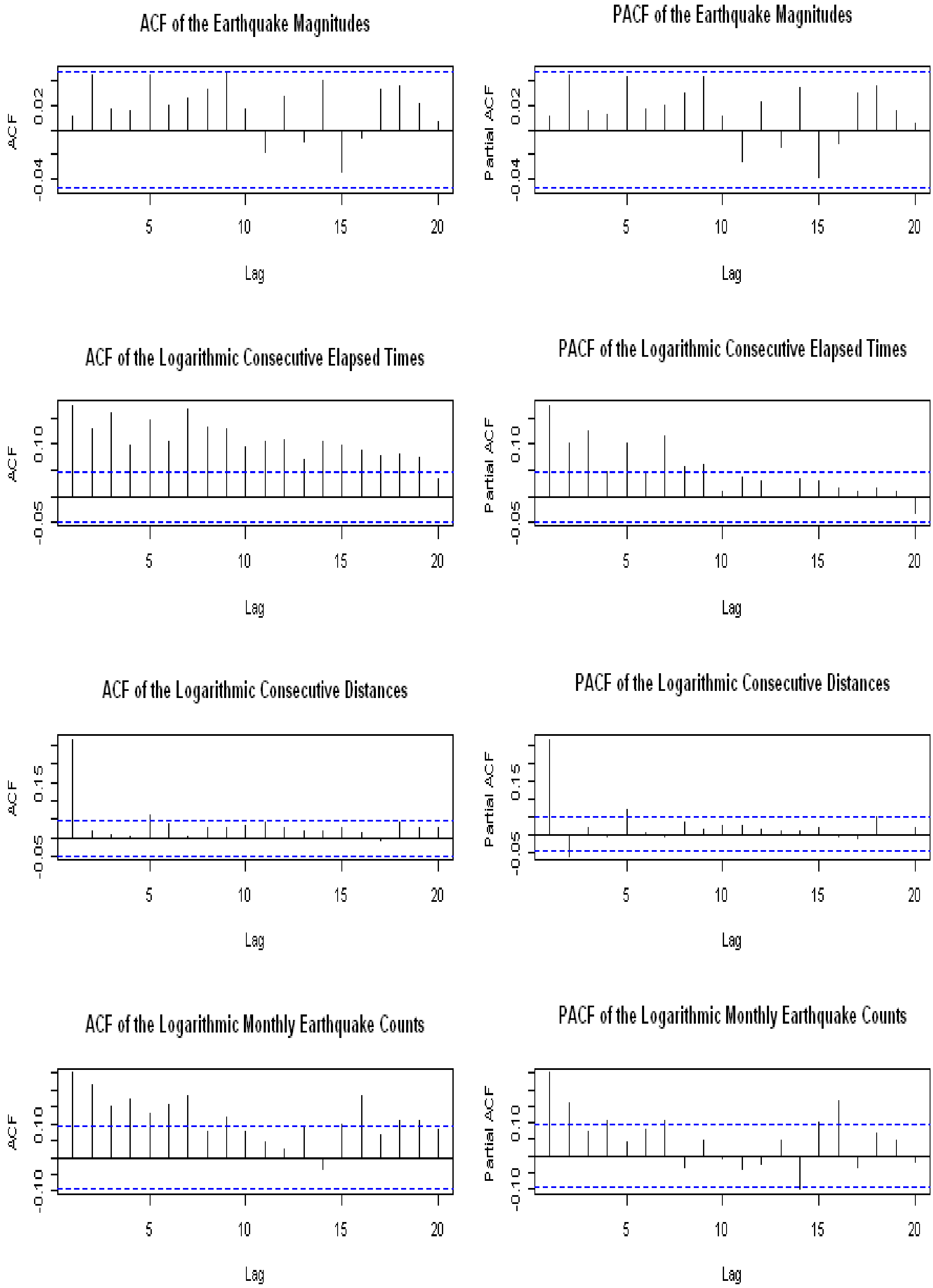


Figure 5.5.3: Histograms of raw and logarithmically transformed seismic data

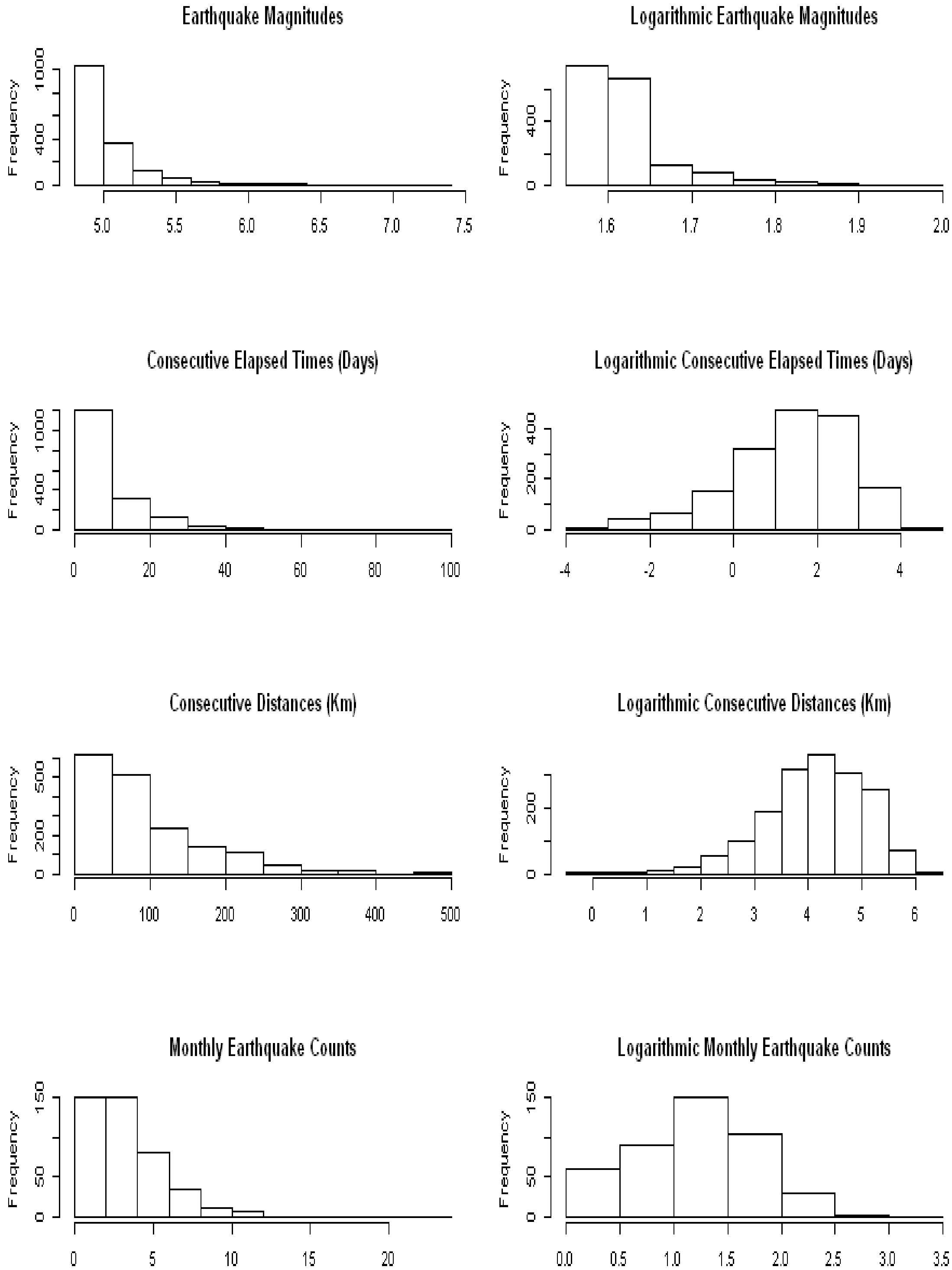


Table 5.1: Summary statistics of seismic data

Variables	Specifications	Minimum	Maximum	Mean	Variance	Skewness	Ex-kurtosis
Earthquake Magnitude (M_t)	M_t	4.800	7.300	5.087	0.084	2.741	9.861
Consecutive Elapsed Times (E_t) (days)	E_t	0.020	94.940	8.248	94.872	2.578	10.974
	$ln(E_t)$	-3.912	4.553	1.369	2.046	-0.775	0.529
Consecutive Distances (D_t) (km)	D_t	0.950	455.800	90.920	5482.133	1.371	1.864
	$ln(D_t)$	-0.051	6.122	4.155	0.836	-0.578	0.486
Monthly Earthquake Counts (MC_t)	(MC_t)	1.000	23.000	3.881	7.108	2.169	8.819
	$ln(MC_t)$	0.000	3.135	1.150	0.428	-0.115	-0.391

Table 5.2: P-values of the stationarity and linearity tests

Variables	ADF test	Keenan test	Tsay test
M_t	0.00	0.00	0.02
$Ln(E_t)$	0.00	0.00	0.00
$Ln(D_t)$	0.00	0.00	0.00
$Ln(MC_t)$	0.00	0.00	0.00

Note: Under the null hypothesis in linearity tests, we assume that the time series is linear.

vided by the ADF test in Table 5.2 indicate that all seismic time series are stationary at level. The zero p-values of the Kennan and Tsay tests reject the null hypothesis of linearity and suggest that all seismic time series are nonlinear at conventional significance level. However, non-rejection of the nonlinearity hypothesis of the seismic time series does not suggest that nonlinear models will provide more accurate out-of-sample forecasts. This could be achieved if the out-of-sample forecasting period contains high degree of nonlinearity needed by the nonlinear models.

The correlation matrix of the seismic data is provided in Table 5.3. Low correlation is observed between all seismic time series. However, it needs to be clarified that the correlation coefficient measures the degree of linear relationship between two variables but it does consider nonlinear relationship. However, the plots of the seismic time series show presence of some nonlinear phenomena in these time series. Prediction of the occurrence time, magnitude, and epicentral location of future earthquakes has been the subject of many studies with different conclusions. Some scientists have concluded that prediction of these variables for a single earthquake cannot be done, while others have suggested several procedures and techniques to predict them (Panakkat, 2007).

Table 5.3: Correlation matrix of seismic variables

	M_t	$Ln(E_t)$	$Ln(D_t)$
M_t	1.000	0.039	-0.030
$Ln(E_t)$	0.039	1.000	0.139
$Ln(D_t)$	-0.030	0.139	1.000

Table 5.4: Lag length (m) specifications in time series models

Models	$Ln(E_t)$	$Ln(D_t)$	M_t	$Ln(MC_t)$
AR	8	2	1	2
AAR	1	1	1	1
SETAR	mL=8, mH=1	mL=1, mH=1	mL=1, mH=1	mL=1, mH=1
NNET	m=3, size=1	m=2, size=1	m=1, size=1	m=3, size=2
LSTAR	mL=8, mH=2	mL=1, mH=1	mL=2, mH=1	mL=2, mH=1
TAR-Ext(magnitude)	mL=8, mH=1	mL=1, mH=1	—	—
LSTAR-Ext(magnitude)	mL=8, mH=2	mL=1, mH=1	—	—
TAR-Ext(distance)	mL=8, mH=1	—	mL=1, mH=1	—
LSTAR-Ext(distance)	mL=8, mH=2	—	mL=2, mH=1	—
TAR-Ext(elapsed times)	—	mL=1, mH=1	mL=1, mH=1	—
LSTAR-Ext(elapsed times)	—	mL=1, mH=1	mL=2, mH=1	—
VAR	(1,1,1)	(1,1,1)	(1,1,1)	—
ACD	(1, 1)	(1, 1)	(1, 1)	(1, 1)
TVAR	(1,1,1)	(1,1,1)	(1,1,1)	(1,1,1)

Note: The symbols mL and mH respectively denoted the lag length in upper and lower regime of threshold models. Size denoted the number of hidden unites in neural network models. The values (1,1,1) represent the lag length of each variable used in a trivariate VAR model.

The main purpose of the current study is to bridge the analytical gap between statistics and seismology by applying the linear and nonlinear time series models to seismic data and compare their forecasting performance. So this study could be considered as an effort to build a bridge between statistics and seismology. We make an advancement in TAR and LSTAR models by including the external threshold variables in these models. TAR and LSTAR models with internal and external threshold variables were estimated and then compared with linear AR benchmark model by using Mean Square Forecasting Error (MSFE) and pairwise Modified Diebold-Mariano (MDM) test of Harvey, Leybourne, and Newbold (1997). For each seismic time series analyzed in this study, we estimated the SETAR and LSTAR models by considering the first lag values of the dependent variable of the model itself as an internal threshold variable. These models were also estimated by using the first lag values of the external threshold variables; these models are referred to as TAR-Ext (threshold autoregressive model with external threshold variable) and LSTAR-Ext (logistic smooth transition autoregressive model with external threshold variable). A comparison between VAR and AR forecasts was also carried out. For monthly earthquake counts time series, forecasts from the AR model were compared with those obtained from nonlinear time series models by using the first lag values of the target variable itself as an internal threshold variable.

As mentioned earlier, the first estimation window consist of first 60% of the total observations, i.e., the first estimation window for (M_t), $Ln(E_t)$, and $Ln(D_t)$ utilized the first 1016 observations, while for $Ln(MC_t)$ the first estimation windows contains the first 262 observations under the expanding window scheme. We then recursively generated the first four-step-ahead forecasts, i.e., $h = 1, 2, 3, 4$ from each model. In this way, we generated 673 point forecasts for each of the forecasting horizons 1 through 4, for earthquake magnitudes, consecutive elapsed times, and for consecutive distances time series. Similarly, for monthly earthquake counts time series, we generated 171 point forecasts for each of the forecasting horizons 1 through 4. All time series considered in this study for modeling and prediction purposes, except the earthquake magnitudes, were logarithmically transformed. The values of optimal lag length selected through BIC for different time series models are listed in Table 5.4. For consecutive distances time series, we estimated the SETAR model by using the first differenced lag values of the target variable as an internal threshold variable, i.e., Δy_{t-1}. The optimal threshold delay parameter is set to one in TAR and LSTAR models with internal and external threshold variables for all four seismic time series. The lag length and the number of hidden units (size) in Artificial Neural Network (ANN) model were set to 1 for earthquake magnitudes. For consecutive distances time series, values of the lag length and number of hidden units (size) were set to 2 and 1 respectively, while these parameters for the consecutive elapsed times series were set to 3 and 1 respectively. In n AAR model, by using BIC, for all four seismic time series the optimal lag length appear as one. Similarly, for the AR model the, optimal lag length values selected through BIC were 8, 2, 1 and 2 respectively for $Ln(E_t)$, $Ln(D_t)$, (M_t), $and\, Ln(MC_t)$, while in SETAR model lag length one is used in both regimes for $Ln(D_t)$, (M_t), $and\, Ln(MC_t)$, i.e., mL=1, mH=1. Similarly, The lag length values were set to 8 and 2 in the lower and upper regime of the LSTAR model, i.e., mL=8, mH=2. The VAR model with lag length 1 was estimated for $Ln(E_t)$, $Ln(D_t)$, $and\, (M_t)$ time series. By utilizing the BIC, TVAR model with lag length one in each regime was estimated for the time series of earthquake magnitudes, logarithmic consecutive elapsed times, and logarithmic consecutive distances. The threshold variable in TVAR model was the first lag values of the target variable (dependent variable).

After estimating the linear and nonlinear time series models and obtaining the 1 through 4 step-ahead forecasts, i.e., $h = 1, 2, 3, 4$, we computed the Mean Square Forecasting Error (MSFE) for each seismic time series. The estimated MSFE values are reported in Table 5.5. For earthquake magnitude time series, a slight difference in the values of MSFE for different models were observed. However, by using the consecutive distances as an external threshold variable, the MSFE values of TAR-Ext(distance) and LSTAR-Ext(distance), i.e., using the consecutive distances time series as an external threshold variable in TAR and LSTAR models, are smaller than AR model respectively for $h = 1, 2$ and $h = 1$.

Table 5.5: Point forecast evaluation for seismic data: MSFE values

Models	Time Series											
	Earthquake Magnitudes				Consecutive Elapsed Times				Consecutive Distances			
	h=1	h=2	h=3	h=4	h=1	h=2	h=3	h=4	h=1	h=2	h=3	h=4
AR	76.68	76.66	77.36	77.62	1836.30	1851.10	1863.40	1873.60	842.40	910.30	910.60	908.70
AAR	76.68	76.56	77.62	77.66	1893.30	1908.40	1921.10	1912.10	830.00	904.90	906.60	911.30
SETAR	77.04	77.09	76.97	77.62	1840.20	1852.80	1859.70	1852.50	823.30	908.30	908.70	908.60
NNET	76.67	76.22	77.43	77.88	1863.80	1901.70	1898.80	1896.40	875.60	918.00	908.90	913.60
LSTAR	76.28	76.06	77.34	77.77	1820.00	1846.30	1855.50	1869.20	837.00	914.90	919.30	909.30
TAR-Ext(elapsed times)	76.40	77.14	77.99	78.13	—	—	—	—	821.60	873.60	879.20	882.10
LSTAR-Ext(elapsed times)	76.29	76.80	77.83	78.30	—	—	—	—	837.00	914.20	914.90	911.10
TAR-Ext(distance)	76.16	76.20	77.63	77.60	1855.00	1867.30	1857.20	1859.40	—	—	—	—
LSTAR-Ext(distance)	75.81	76.25	77.35	77.88	1820.00	1841.40	1859.60	1874.00	—	—	—	—
TAR-Ext(magnitude)	—	—	—	—	1852.70	1873.50	1879.70	1892.60	848.60	912.90	916.10	917.20
LSTAR-Ext(magnitude)	—	—	—	—	1820.00	1838.60	1862.80	1870.00	837.00	907.90	922.10	911.70
VAR	77.01	76.76	77.37	77.62	1899.50	1902.50	1910.50	1912.80	831.50	900.70	903.50	907.50
TVAR	79.60	79.53	80.85	80.06	1855.95	1892.25	1908.18	1922.82	825.28	887.38	889.37	895.96

Note: All entries are multiplied by 1000.

For the time series of consecutive elapsed times, the observed values of MSFE are reported in Table 5.5. It shows that the observed values of MSFE from SETAR and LSTAR models are slightly smaller than from the AR model respectively for 3 through 4 and 1 through 4 step-ahead forecasts. Similarly, using the external threshold variables in threshold models yields a slight decrease in the values of the MSFE. Thus, SETAR and TAR-Ext (consecutive distances) models provide more accurate forecasts than the AR model for $h = 3,\ 4$. Also, the LSTAR model provides more accurate forecasts than the AR model for $h = 1,\ 2,\ 3,\ 4$. The significance of the observed differences in the values of MSFE was tested through MDM test, and the results are provided in Table 5.6.

Table 5.5 shows that for consecutive distances time series the observed values of MSFE from SETAR, TAR-Ext(elapsed times), and VAR models are smaller than MSFE values of the AR model for $h = 1,\ 2,\ 3,\ 4$. Thus, the linear VAR, nonlinear SETAR, and TAR-Ext(elapsed times) models have an improved forecasting gain over benchmark AR model. Similarly, the values of the MSFE from the AAR model are slightly smaller than the AR model for almost all forecasting horizons. The LSTAR model with internal and external threshold variables yields smaller MSFE values for $h = 1$ than the AR model. Hence, using the MSFE values as an out-of-sample evaluation criteria provides an evidence of an improved forecasting gain of the VAR, AAR, TAR-Ext(elapsed times) and SETAR models over the linear AR model for almost all forecasting horizons. Also, the LSTAR model shows improved forecasting accuracy over the AR model for $h = 1$. In the remaining out-of-sample comparison, none of the models listed in Table 5.5 outperform the benchmark linear AR model.

The TVAR model provides smaller magnitude MSFE values for the time series of consecutive distances for one through four step-ahead forecasts. The significance of the observed differences in MSFE values was tested with the MDM test.

When performing the MDM test, we constructed the null hypothesis stating that the Mean Square Forecasting Error (MSFE) of the linear AR model is smaller or equal to the MSFE of the corresponding nonlinear model, i.e., $MSFE(AR) \leqslant MSFE(Nonlinear\,Model)$ and alternative hypothesis is opposite of the null hypothesis. The squared loss function was used in the MDM test. We compared the observed p-values of the MDM test with a pre-assumed significance level 0.05. This implied that under the null hypothesis p-values smaller than 0.05 indicate rejection of the null hypothesis and thereby improved forecasting performance of the nonlinear models relative to the benchmark linear AR model. After securing the forecasting errors from linear and nonlinear models, we applied the MDM test and its p-values are shown in Table 5.6.

Table 5.6: Point forecast evaluation for seismic data: P-values of the MDM test

Models Comparison	Time Series											
	Earthquake Magnitudes				Consecutive Elapsed Times				Consecutive Distances			
	h=1	h=2	h=3	h=4	h=1	h=2	h=3	h=4	h=1	h=2	h=3	h=4
AR vs AAR	0.50	0.26	0.87	0.60	0.95	0.91	0.90	0.84	0.05	0.27	0.20	0.86
AR vs SETAR	0.83	0.97	0.06	0.50	0.53	0.49	0.45	0.20	0.05	0.42	0.26	0.67
AR vs NNET	0.47	0.02	0.63	0.94	0.81	0.94	0.87	0.76	0.98	0.88	0.24	0.99
AR vs LSTAR	0.34	0.28	0.47	0.79	0.21	0.31	0.19	0.27	0.25	0.78	0.92	0.64
AR vs TAR-Ext(elapsed times)	0.22	0.95	0.97	0.91	—	—	—	—	0.01	0.00	0.00	0.00
AR vs LSTAR-Ext(elapsed times)	0.18	0.65	0.91	0.99	—	—	—	—	0.25	0.72	0.77	0.81
AR vs TAR-Ext(distance)	0.09	0.03	0.85	0.46	0.75	0.73	0.42	0.32	—	—	—	—
AR vs LSTAR-Ext(distance)	0.02	0.14	0.47	0.87	0.21	0.15	0.34	0.52	—	—	—	—
AR vs TAR-Ext(magnitude)	—	—	—	—	0.91	0.94	0.86	0.94	0.83	0.63	0.81	0.91
AR vs LSTAR-Ext(magnitude)	—	—	—	—	0.21	0.10	0.47	0.31	0.25	0.35	0.97	0.87
AR vs VAR	0.94	0.85	0.51	0.52	0.96	0.89	0.86	0.85	0.10	0.10	0.02	0.09
AR vs TVAR	0.99	0.99	0.99	0.99	0.63	0.77	0.78	0.79	0.09	0.05	0.07	0.10

Note: Under the null hypothesis in pairwise MDM test, we assume that AR model has significantly smaller MSFE than the corresponding MSFE of the nonlinear model.

The p-values of the MDM test for earthquake magnitude time series reported in Table 5.6, suggest that at 10% significance level there is an improved forecasting gain to the SETAR and NNET models in comparison to the AR model respectively for $h = 3$ and $h = 2$. Similarly, the TAR-Ext(distance) and LSTAR-Ext(distance) models (i.e., using the consecutive distances time series as an external threshold variable in TAR and LSTAR models) showed an improved forecasting gain in comparison to the AR model respectively for $h = 1$, 2 and $h = 1$ at 10% significance level.

Furthermore, Table 5.6 shows that for consecutive elapsed times almost all p-values of the MDM test are greater than 0.05, thus providing the evidence that in out-of-sample forecasting performance none of the nonlinear models outperform the linear AR model. There is no improvement in the forecasting accuracy of the TAR and LSTAR models by incorporating external threshold variables in these models. One possible reason of equal forecasting performance of linear and nonlinear models would be that the forecasting time period may not contain enough nonlinearity needed by the nonlinear models to yield improved out-of-sample forecasts over the linear AR model.

The observed p-values of the MDM test for consecutive distances time series reported in Table 5.6 show that the AAR and SETAR models outperform the linear AR model in out-of-sample forecasting performance for $h = 1$. In addition, the pairwise comparison of the AR and TAR-Ext(elapsed times) models shows that the p-values of the MDM test are smaller than 0.05 for 1 through 4 step-ahead forecasts, thus the TAR-Ext(elapsed times) model provides more accurate forecast over the linear AR model for $h = 1, 2, 3, 4$. In second case, when elapsed times are assumed as an external threshold variable in the TAR model, the TAR-Ext(consecutive elapsed times) model provides more accurate forecasts than the SETAR model for all forecasting horizons.

The comparison of the forecasts from VAR and TVAR models with benchmark AR model (see Table 5.6) shows that for earthquake magnitudes and consecutive elapsed times all p-values provided by the MDM are greater than 0.05. Thus, there is no forecasting gain to VAR and TVAR models relative to the linear AR model. Furthermore to compare the VAR and TVAR models with benchmark AR model for consecutive distances time series, we observed in Table 5.6 that the p-values of the MDM test are smaller or equal to 0.10 for 1 through 4 step-ahead forecasts. Thus, the VAR and TVAR models provide improved 1 through 4 step-ahead forecasts in comparison to the linear AR model at 10 % significance level[4].

[4]TVAR model with various potential external threshold variables is also estimated and forecasted for seismic data, however it does not improved it out-of-sample forecasting performance over linear

Table 5.7: Point forecast evaluation for monthly earthquake counts: MSFE values

Models	Forecasting horizons			
	h=1	h=2	h=3	h=4
AR	441.50	459.30	454.40	457.80
AAR	423.40	470.50	464.10	468.20
SETAR	422.50	436.60	453.80	455.10
NNET	460.90	521.20	476.90	471.90
LSTAR	434.70	453.20	456.80	459.20

Note: All entries are multiplied by 1000.

Table 5.8: Point forecast evaluation for monthly earthquake counts: P-values of the MDM test

Models Comparison	Forecasting horizons			
	h=1	h=2	h=3	h=4
AR vs AAR	0.08	0.71	0.85	0.96
AR vs SETAR	0.01	0.01	0.48	0.29
AR vs NNET	0.87	0.93	0.86	0.94
AR vs LSTAR	0.18	0.12	0.61	0.71

Note: Under the null hypothesis in pairwise MDM test, we assume that AR model has significantly smaller MSFE than the corresponding MSFE of the nonlinear model.

The MSFE values of the linear and nonlinear models for monthly earthquake counts time series are presented in Table 5.7. These results show that the AAR model has a smaller value of MSFE in comparison to the AR model for $h = 1$. Similarly, the observed MSFE values of SETAR and LSTAR models are smaller than those of the AR model respectively for $h = 1, 2, 3, 4$ and $h = 1, 2$. Thus, according to MSFE criterions, the SETAR and LSTAR models provide more accurate forecasts in comparison to the linear AR model respectively for $h = 1, 2, 3, 4$ and $h = 1, 2$.

The p-values of the MDM test in a pairwise model comparison for monthly earthquake counts are provided in Table 5.8. According to this table, in pairwise comparison of the AR and SETAR models, p-values corresponding to $h = 1, 2$ are smaller than 0.05, thus the SETAR model provides more accurate forecasts in comparison to the linear benchmark AR model for $h = 1, 2$.

In summary, we reached the conclusion that the SETAR and AR models provide more accurate forecasts than the benchmark linear AR model for $h = 1$ for consecutive distances time series. Secondly, for the same time series, the TAR-Ext(elapsed times), i.e., using the elapsed times as an external threshold variable in the TAR model, provides more accurate forecasts than the linear AR model for $h = 1, 2, 3, 4$. The VAR model provides improved forecasts in comparison to the AR

AR model, so we did not report these results.

model for longer forecasting horizons for consecutive distances time series.

Thus, for time series of earthquake magnitudes and consecutive elapsed times, the AR model turns out as the best forecasting candidate, while for consecutive distance time series the VAR, TAR(elapsed times), and AAR appeared to be the best forecasting models.

The results show that the time series of consecutive elapsed times attributed nonlinearity in consecutive distances. This make sense: if an area has experienced a high magnitude earthquake, then the probability of the next large earthquake will be high with high elapsed time. Thus, for the occurrence of another high magnitude earthquake near to this place, the consecutive distances will be high, and vice versa for small magnitude earthquakes. For example, if an earthquake of magnitude 5 occurs in a particular area, then it is expected that the consecutive distance will increase for the occurrence of another earthquake of magnitude 5 or greater. As most of the stored energy in the Earth's crust will be released in the first earthquake, it will take time to develop potential amount of energy at the same fault, and thus for a new earthquake to occur. In other words, since we know that the convectional forces push the tectonic plates continuously (because there is a source of heat in the Earth), they are continuously being under high stress. When elapsed time between successive earthquakes increases, the stress also increases, and a huge amount of energy is accumulated in the rocks. As a result, large magnitude earthquakes occur after longer elapsed time. When an earthquake occurs on some fault at a particular location, it stimulates another earthquake on the same fault at a new location or on some another fault nearby the previous fault. The earthquake can be of any magnitude, depending on the accumulated energy.

The cross correlation plots of the seismic time series depicted in Figure 5.5.4 shows that the time series of consecutive elapsed times and consecutive distances show a fair amount of correlation at various lags. Thus, it provides further support for the estimation of these time series in the VAR framework. The left and right panels of Figure 5.5.5 respectively show the ACFs of the residuals of the estimated AR and SETAR models (the best forecasting models among various potential candidates) by utilizing the full sample of seismic time series. We did not find any dependence structures in residuals shown in Figure 5.5.5. Therefore, the residuals from the estimated AR and SETAR models are white noise.

The observed and one step-ahead forecasted values of the AR and SETAR models are plotted respectively in Figures 5.5.6 and 5.5.7. We only provided the forecasted values of the best forecasting models (AR and SETAR) among all potential candidates. The observed and forecasting values for consecutive elapsed times, consecutive distances, and monthly earthquake counts are quite close in both models; however, for earthquake magnitude time series, the observed and forecasted values deviate slightly more from each other. The forecasted values of consecutive elapsed times from the AR model (the best forecasted model for elapsed times) are much closer to the observed values in comparison to the forecasted values from the SETAR model. In these figures, we also plotted the forecasted earthquake magnitudes and logarithmic consecutive distances time series by using the TAR(distance) and TAR(elapsed times) models, i.e., using the consecutive distances and consecutive elapsed times as an external threshold variables in TAR models, while respectively predicting the earthquake magnitudes and consecutive distances time series.

As we are mostly interested to compare the forecasting performance of the threshold autoregressive and linear AR models, we only provide the estimation results for these models[5]. The estimation results of the VAR model is also provided to understand the relationship between seismic variables. The parameter estimates of the linear AR, VAR, two regimes SETAR, and LSTAR models along with their standard deviations and p-values are presented in Tables 5.9 to 5.13. By using BIC, we estimated the AR(8), AR(2), and AR(1) models respectively for consecutive elapsed times, consecutive distances, and earthquake magnitudes time series. The p-values of parameter estimates smaller than 0.05 listed in Table 5.9 show that for consecutive elapsed times and consecutive distances time series, almost all coefficients of the linear AR model are significant at 5 % significance level. For earthquake magnitude time series, the first lag coefficient is insignificant. Thus, the significant coefficients provided by the estimated AR model for two seismic time series provide the evidence of a high persistence in them. The estimation results of the VAR model provided in Table 5.10 shows that for consecutive elapsed times and consecutive distances time series most of the estimated values of the first lag coefficients are significant at conventional significance level, while for earthquake magnitudes time series all lags coefficients are insignificant. The estimated parameter values of the SETAR and LSTAR models along with their standard deviations and p-values are reported in Tables 5.11 to 5.12 for all three seismic time series. BIC was used to estimate the SETAR model with optimal threshold value and lag length. The first differenced first lag values of consecutive distances are used as an internal threshold variable in the estimation of the SETAR model for the time series of consecutive distances. From the observed p-values of the estimated parameters of the SETAR and LSTAR models, for consecutive elapsed times and consecutive distances time series almost all parameter estimates

[5]Parameter estimates their standard errors and p-values of the VAR and AAR models can be provided on request

Table 5.9: Estimation results of linear AR model

Coefficients	Consecutive Elapsed Times			Consecutive Distances			Earthquake Magnitudes		
	Values	Std.	P-values	Values	Std.	P-values	Values	Std.	P-values
Constant	0.607	0.072	0.000	3.221	0.124	0.000	5.028	0.124	0.000
y_{t-1}	0.117	0.024	0.000	0.285	0.024	0.000	0.012	0.024	0.634
y_{t-2}	0.054	0.024	0.027	-0.060	0.024	0.014	—	—	—
y_{t-3}	0.096	0.024	0.000	—	—	—	—	—	—
y_{t-4}	0.014	0.024	0.552	—	—	—	—	—	—
y_{t-5}	0.084	0.024	0.001	—	—	—	—	—	—
y_{t-6}	0.028	0.024	0.257	—	—	—	—	—	—
y_{t-7}	0.109	0.024	0.000	—	—	—	—	—	—
y_{t-8}	0.055	0.024	0.023	—	—	—	—	—	—
BIC	1113.510			-415.470			-4170.859		

Table 5.10: Estimation results of VAR model

Coefficients	Consecutive Elapsed Times			Consecutive Distances			Earthquake Magnitudes		
	Values	Std.	P-values	Values	Std.	P-values	Values	Std.	P-values
Constant	2.290	0.627	0.000	4.293	0.073	0.000	5.043	0.024	0.000
M_{t-1}	-0.234	0.118	0.048	-0.250	0.015	0.000	0.012	0.005	0.625
E_{t-1}	0.174	0.024	0.000	0.068	0.024	0.000	-0.003	0.008	0.531
D_{t-1}	0.008	0.038	0.840	0.251	0.389	0.000	-0.003	0.129	0.714
BIC	10970.451								

are significant at 5 % significance level in both regimes. For earthquake magnitudes, the estimated coefficients of the SETAR and LSTAR models are insignificant in both regimes at 5 % significance level.

For monthly earthquake counts time series, the parameter estimates, their standard deviations, and p-values are provided in Table 5.13. Almost all estimated parameters of AR, SETAR, and LSTAR models are significant at 5 % significant level. The estimation results of TVAR model is provided in Table 5.14, and it shows that for the time series of consecutive distances and consecutive elapsed times, most of the estimated first lag coefficients are significant at 5% significance level.

Regarding in-sample model comparison through BIC, the best fitted model is the AR model, followed by the LSTAR and SETAR models for consecutive elapsed times time series. For consecutive distances time series, SETAR model appeared to be the best in-sample fitted model by providing the lowest value of BIC, followed by the AR and LSTAR models. Similarly, for earthquake magnitudes time series the AR model possesses the best in-sample fitting, followed by the SETAR and LSTAR models. For monthly earthquake counts time series, the AR model provides the lowest value of BIC and thus shows improved in-sample fitting in comparison to SETAR and LSTAR models.

Figure 5.5.4: Cross correlation plots of seismic data, whole sample

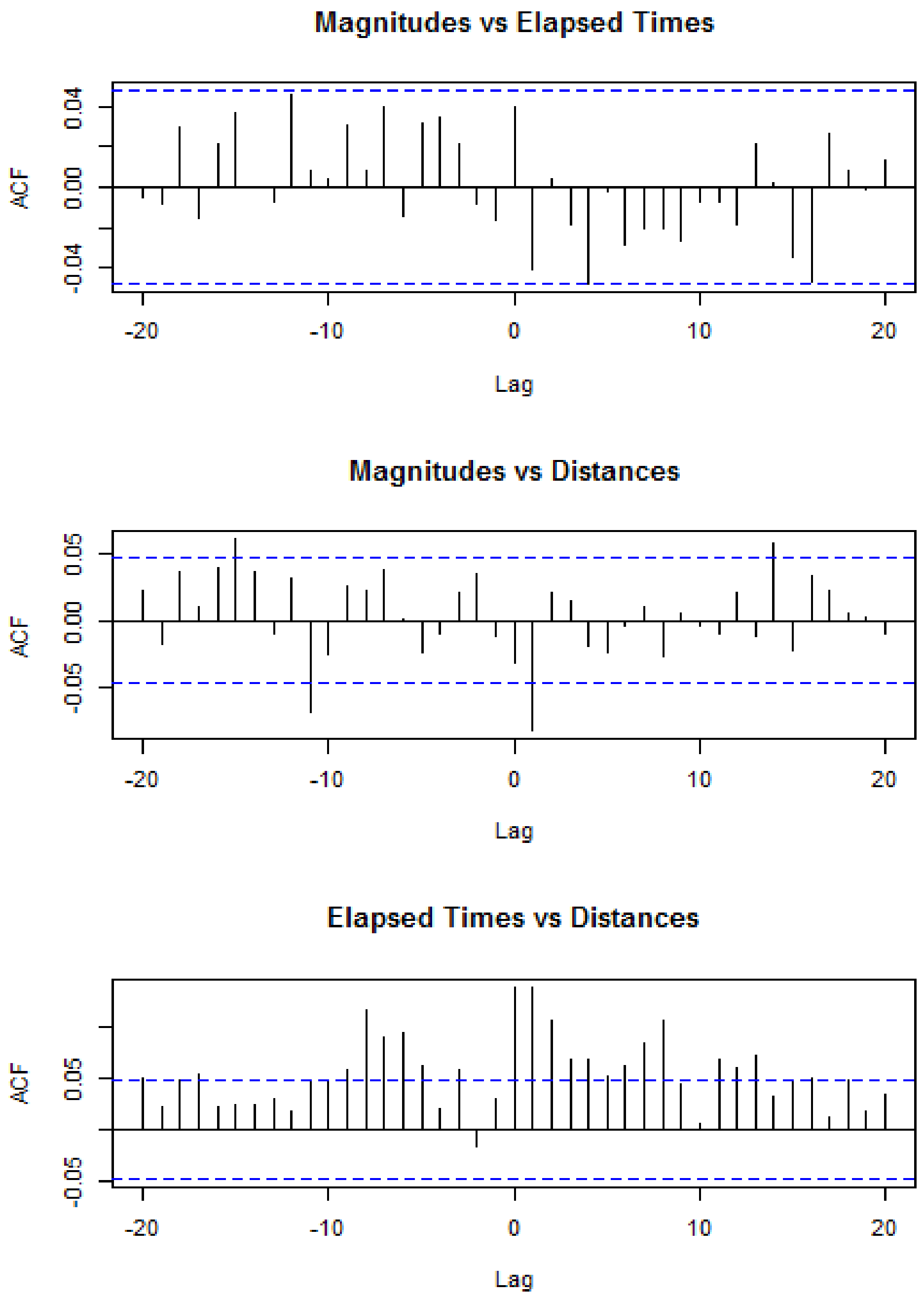


Figure 5.5.5: ACFs of the residuals from the AR(left panel) and SETAR (right panel) models

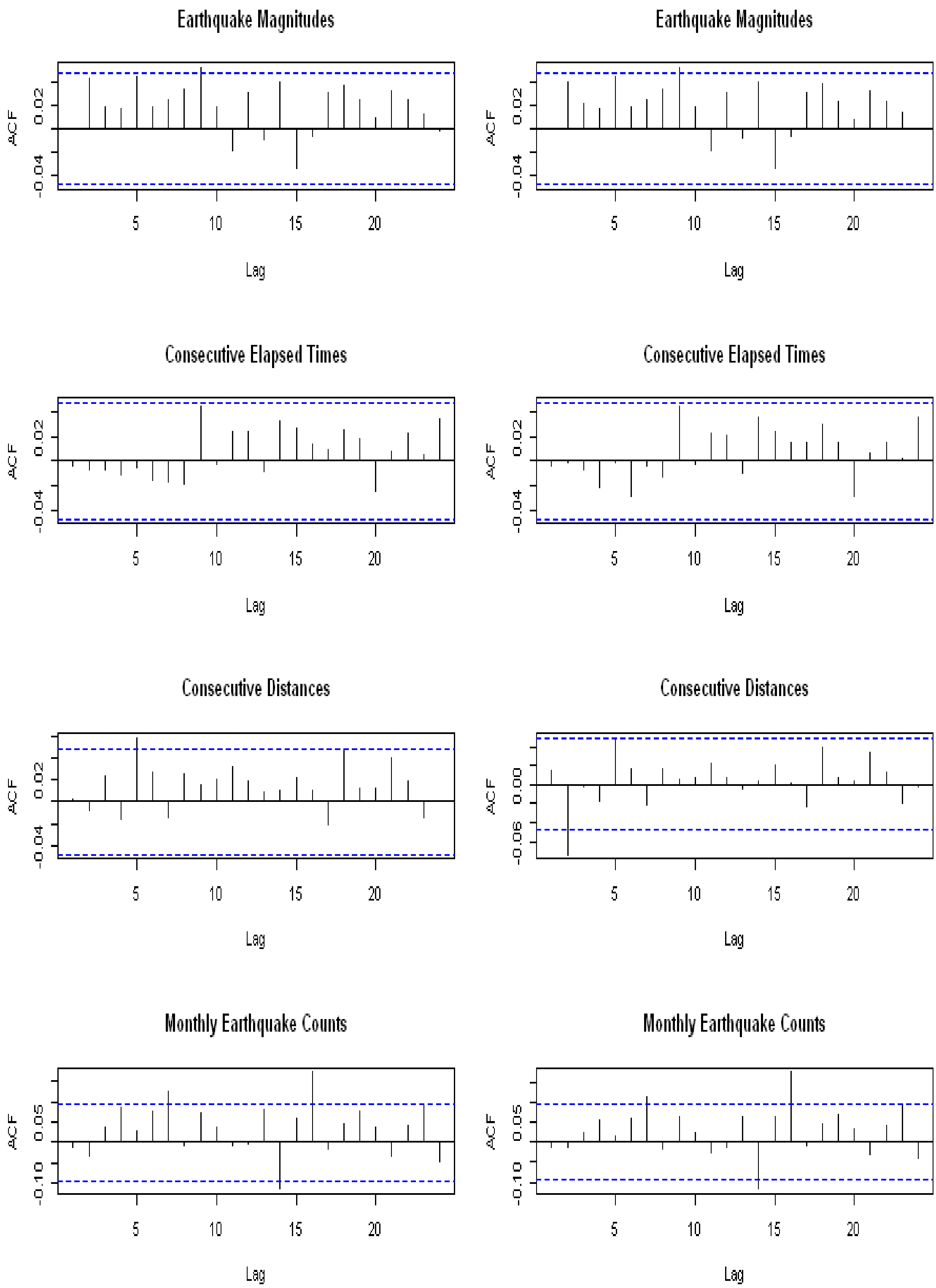


Figure 5.5.6: Observed and 1 step-ahead forecasted values of AR model

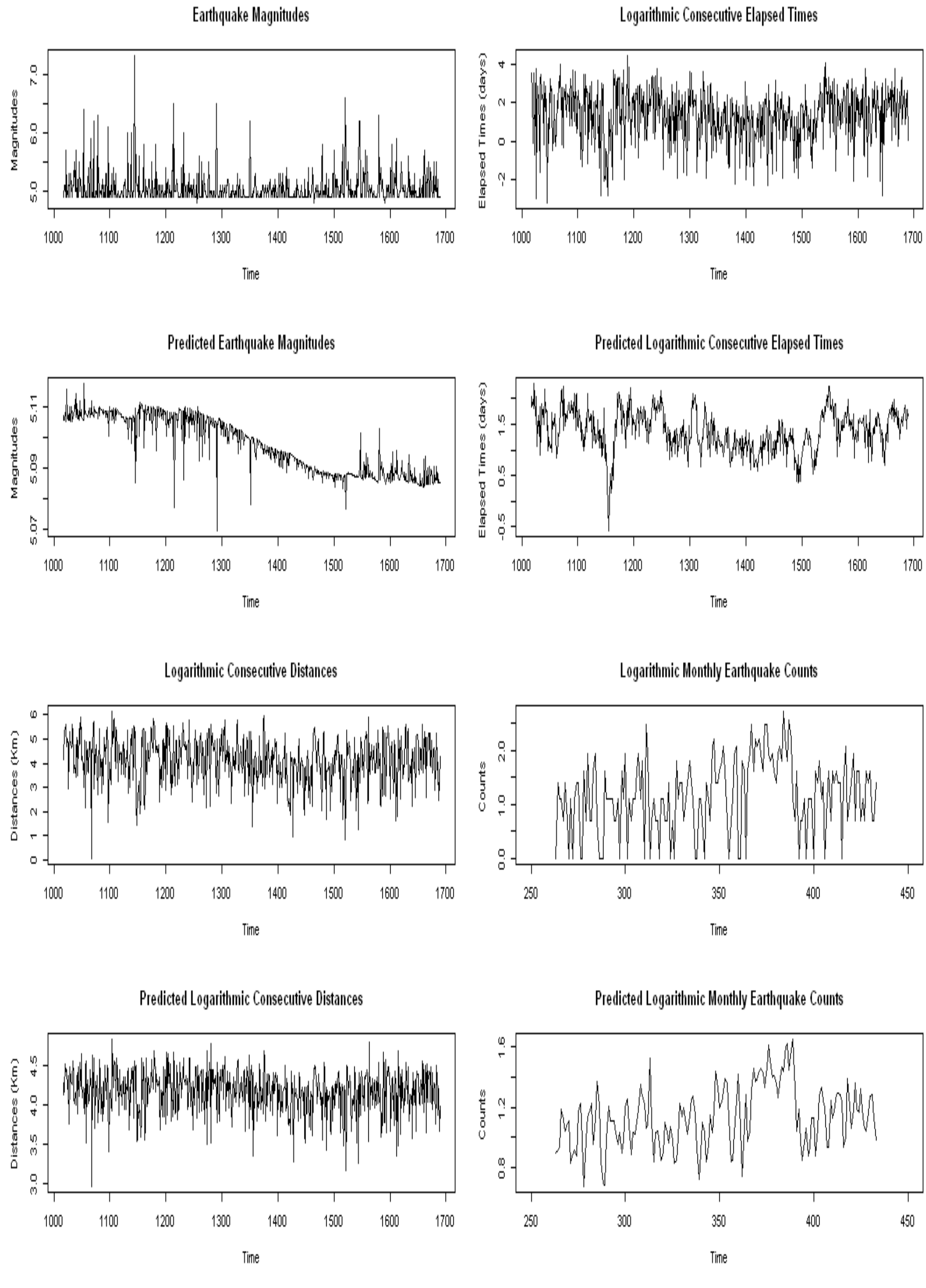


Figure 5.5.7: Observed and 1 step-ahead forecasted values of SETAR model

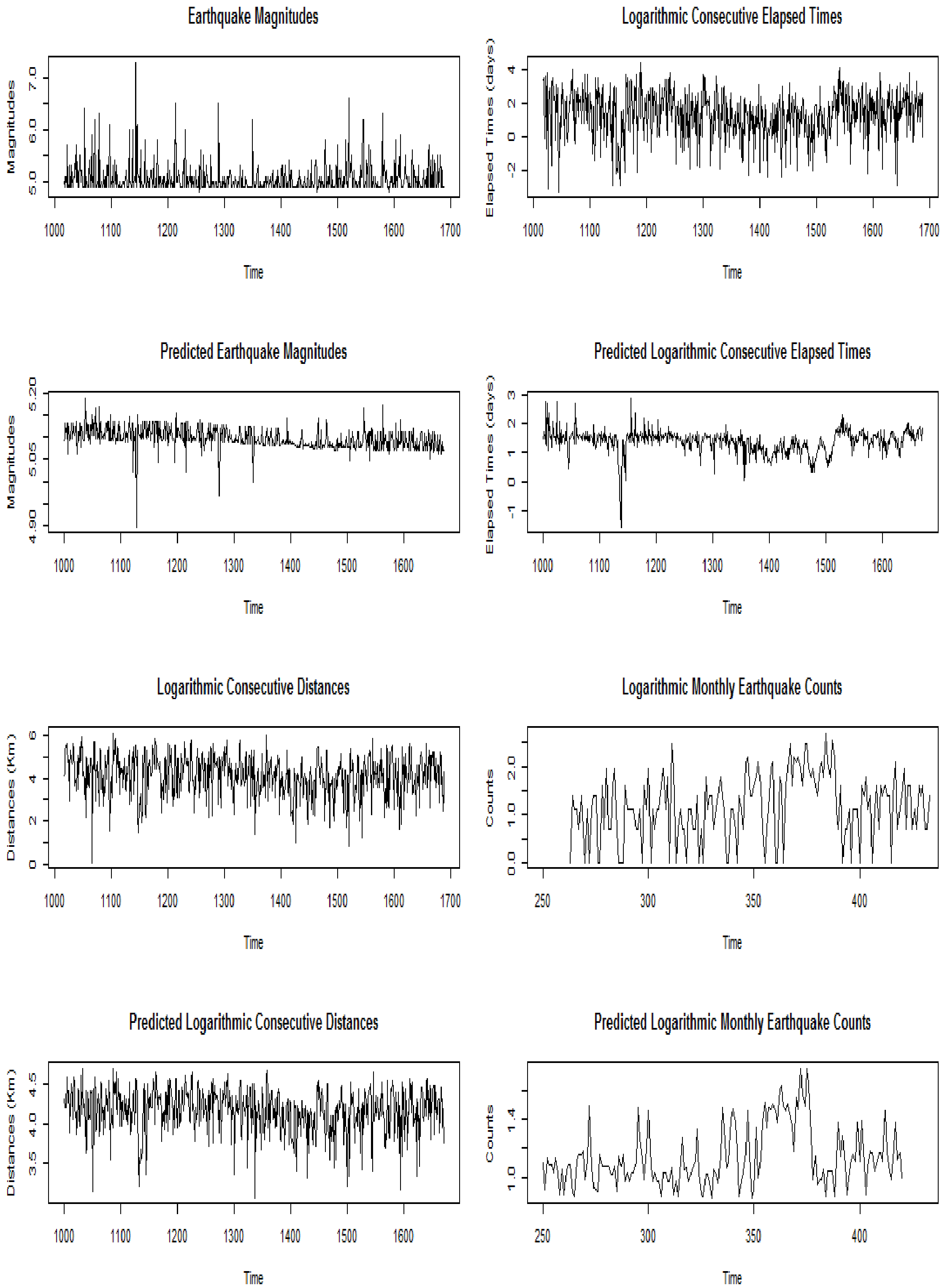


Table 5.11: Estimation results of SETAR model

Coefficients	Consecutive Elapsed Times						Consecutive Distances						Earthquake Magnitudes					
	Low regime			High regime			Low regime			High regime			Low regime			High regime		
	Values	Std.	P-values	Values	Std.	P-values	Values	Std.	P-values	Values	Std.	P-values	Values	Std.	P-values	Values	Std.	P-values
Constant	0.553	0.074	0.000	1.386	0.134	0.000	2.708	0.124	0.000	3.534	0.177	0.000	5.232	0.188	0.000	4.874	0.165	0.000
y_{t-1}	0.113	0.026	0.000	0.119	0.061	0.052	0.360	0.030	0.000	0.145	0.039	0.000	-0.029	0.037	0.439	0.042	0.032	0.193
y_{t-2}	0.049	0.029	0.094	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
y_{t-3}	0.110	0.027	0.000	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
y_{t-4}	0.027	0.026	0.300	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
y_{t-5}	0.092	0.026	0.000	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
y_{t-6}	0.047	0.027	0.079	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
y_{t-7}	0.107	0.026	0.000	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
y_{t-8}	0.055	0.027	0.039	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
BIC	1129.940						-417.334						-4151.091					
Threshold value	2.75						4.46						4.9					

Table 5.12: Estimation results of LSTAR model

Coefficients	Consecutive Elapsed Times						Consecutive Distances						Earthquake Magnitudes					
	Low regime			High regime			Low regime			High regime			Low regime			High regime		
	Values	Std.	P-values	Values	Std.	P-values	Values	Std.	P-values	Values	Std.	P-values	Values	Std.	P-values	Values	Std.	P-values
Constant	0.561	0.186	0.003	0.093	0.204	0.650	2.728	0.121	0.000	0.771	0.222	0.001	6.441	4.234	0.245	-4.272	3.521	0.225
y_{t-1}	0.365	0.059	0.000	-0.292	0.064	0.000	0.356	0.030	0.000	-0.206	0.051	0.000	-0.284	0.642	0.423	0.899	0.721	0.221
y_{t-2}	0.088	0.108	0.416	-0.004	0.113	0.974	—	—	—	—	—	—	-0.212	0.532	0.352	—	—	—
y_{t-3}	0.089	0.024	0.000	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
y_{t-4}	0.010	0.024	0.690	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
y_{t-5}	0.081	0.024	0.001	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
y_{t-6}	0.020	0.024	0.409	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
y_{t-7}	0.102	0.024	0.000	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
y_{t-8}	0.053	0.024	0.027	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
Gamma	100						94.923						0.022					
Threshold value	-0.232						4.528						5.40					
BIC	1128.10						-408.466						-4151.209					

Table 5.13: Estimation results of AR, SETAR and LSTAR models for monthly earthquake counts

Coefficients	AR			SETAR						LSTAR					
				Low regime			High regime			Low regime			High regime		
	Values	Std.	P-values	Values	Std.	P-values	Values	Std.	P-values	Values	Std.	P-values	Values	Std.	P-values
Constant	0.721	0.070	0.000	0.900	0.071	0.000	0.950	0.121	0.000	0.881	0.090	0.001	0.024	0.151	0.900
y_{t-1}	0.212	0.050	0.000	0.131	0.060	0.030	0.311	0.080	0.000	0.131	0.062	0.020	0.171	0.102	0.080
y_{t-2}	0.161	0.050	0.000	—	—	—	—	—	—	0.031	0.070	0.630	—	—	—
Threshold value	—			1.386						1.50					
BIC	-392.100			-387.741						-375.823					

Table 5.14: Estimation results of TVAR model

Coefficients	Consecutive Elapsed Times						Consecutive Distances						Earthquake Magnitudes					
	Low regime			High regime			Low regime			High regime			Low regime			High regime		
	Values	Std.	P-values	Values	Std.	P-values	Values	Std.	P-values	Values	Std.	P-values	Values	Std.	P-values	Values	Std.	P-values
Constant	-0.288	1.412	0.838	3.003	0.696	0.000	4.294	0.626	0.000	4.121	0.536	0.000	5.092	0.169	0.000	5.609	1.565	0.001
M_{t-1}	0.146	0.268	0.585	-0.310	0.130	0.018	-0.138	0.114	0.225	-0.377	0.096	0.000	0.001	0.033	0.969	-0.061	0.245	0.806
E_{t-1}	0.309	0.068	0.000	0.191	0.048	0.000	0.106	0.023	0.000	0.043	0.020	0.033	-0.002	0.008	0.760	0.010	0.061	0.876
D_{t-1}	0.205	0.070	0.004	-0.085	0.045	0.050	0.059	0.056	0.287	0.431	0.054	0.000	-0.002	0.005	0.728	-0.075	0.051	0.149
BIC	1191.861						-430.742						-4141.462					
Threshold value	0.66						3.98						6.0					

5.5.2 Analysis of Raw Seismic Data

In this section, raw seismic data are analyzed by applying linear and nonlinear time series models. The main purpose for analyzing raw seismic data is to compare the forecasting performance of benchmark AR model with SETAR, AAR, ANN, LSTAR, bivariate VAR, and Autoregressive Conditional Duration (ACD) models[6]. We were particularly interested to compare the forecasting performance of the ACD model with the linear AR and nonlinear models. Since the ACD model is working with time series containing positive observations, we analyzed the raw seismic data to ensure this condition. We analyzed the consecutive elapsed times and consecutive distances time series as these time series respectively represent the time and distance duration between consecutive earthquakes and therefore contain positive observations.

The specifications of the linear and nonlinear models considered in this study and selected through BIC remained the same as observed when analyzing the logarithmic seismic data. BIC suggested to estimate the bivariate VAR model with lag length one. The forecasting methodology adopted in this section is the same as the one used in forecasting the logarithmic seismic data in the previous section.

Engle and Russell (1998) proposed an Autoregressive Conditional Duration (ACD) model and showed that this model can successfully describe the evolution of time durations for different stocks. ACD models are mostly used in finance to model and forecast the time intervals between trades. However, in this study we extended the application of ACD model in the field of seismology by utilizing the time series of consecutive elapsed times and consecutive distances between earthquakes. As longer duration between consecutive elapsed times and consecutive distances indicate the absence of earthquake, therefore it is assumed that a fair amount of elastic strain energy is stored in the rocks during this period, which in turn increases the probability of another earthquake in a given region. On the other hand, a small magnitude earthquake often has small elapsed times and consecutive distances in a given region, which in turn reduces the probability of a large magnitude earthquake in this region because most of the energy stored in the rocks is released in the form of various small magnitude earthquakes. Thus, the dynamic behavior of seismic durations contains useful information about the seismic activity.

We estimated a number of ACD (p, q) models by assuming different conditional distribution of the error term (innovation). By considering log-likelihood (LL) criteria, and by assuming the exponential distribution as a conditional distribution of

[6]The ACD model is discussed in chapter 2.

Table 5.15: Point forecast evaluation for raw seismic data: MSFE values

Models	Time Series							
	Consecutive Elapsed Times				Consecutive Distances			
	h=1	h=2	h=3	h=4	h=1	h=2	h=3	h=4
AR	87.27	86.48	86.66	87.28	5033.35	5790.23	5799.24	5788.46
AAR	91.24	90.32	89.69	89.37	5067.21	5797.51	5762.88	5769.47
SETAR	87.62	86.61	87.72	87.20	4941.75	5791.93	5844.24	5798.62
NNET	90.10	90.79	89.76	89.29	5752.30	5868.13	5810.83	5782.11
LSTAR	89.60	87.20	86.97	87.81	4941.65	5795.96	5873.16	5789.25
ACD	86.06	86.12	87.27	88.049	5067.71	5752.98	5752.52	5775.70
VAR	90.24	89.75	89.69	89.69	5034.59	5753.79	5740.23	5768.77

the error term and using the lag length one for observed durations and for conditional expectation of observed duration, the Exponential Autoregressive Conditional Duration (EACD (1, 1)) model turned out to be the best fitting model for both consecutive elapsed times and consecutive distances time series.

The behavior of ACFs and some properties (stationary and linearity) of the consecutive elapsed times and consecutive distances remained the same as observed for its logarithmically transformed version, which has already been discussed in the previous section. We estimated large variety of linear and nonlinear time series models for both time series and generated 1 through 4 step-ahead forecasts from linear and nonlinear models by respectively using the naïve and bootstrapped methods of forecasting. To get a single out-of-sample forecasted value from nonlinear time series models, the forecasted values were averaged over 1000 bootstrapped replications. The Mean Square Forecasting Error (MSFE) values of the respective model for 1 through 4 step-ahead forecasts are shown in Table 5.15. This table shows that for the time series of consecutive elapsed times, for almost all forecasting horizons, the AR model provides low magnitude MSFE values relative to those observed for VAR and nonlinear models. However, the MSFE values of the ACD model for $h = 1, 2$, is slightly smaller in magnitude than those observed for an AR model.

Similarly for the time series of consecutive distances, Table 5.15 shows that the SETAR and LSTAR models provide smaller MSFE values for $h = 1$ than the MSFE value of the AR model. Furthermore, the AAR model also provides smaller MSFE values for $h = 3, 4$ than the MSFE values of the AR model. ACD and VAR models show some improved forecasting performance by rendering smaller MSFE values than those observed for the AR model for $h = 2, 3, 4$. The ANN model does not perform well in terms of out-of-sample forecasting performance for both time series.

To assess whether the observed differences in the values of MSFE for different

Table 5.16: Point forecast evaluation for raw seismic data: P-values of the MDM test

Models Comparison	Time Series							
	Consecutive Elapsed Times				Consecutive Distances			
	h=1	h=2	h=3	h=4	h=1	h=2	h=3	h=4
AR vs AAR	0.98	0.99	0.98	0.91	0.75	0.55	0.14	0.15
AR vs SETAR	0.62	0.54	0.95	0.41	0.22	0.50	0.84	0.74
AR vs NNET	0.96	1.00	0.99	0.90	1.00	0.97	0.74	0.40
AR vs LSTAR	0.93	0.90	0.73	0.84	0.22	0.54	0.94	0.34
AR vs ACD	0.20	0.43	0.68	0.69	0.74	0.20	0.05	0.10
AR vs VAR	0.98	0.99	0.98	0.94	0.62	0.23	0.04	0.10

Note: Under the null hypothesis in pairwise MDM test, we assume that AR model has significantly smaller MSFE than the corresponding MSFE of the nonlinear model.

models for 1 though 4 step-ahead forecasting horizons are significant or not, we used the Modified Diebold-Mariano (MDM) test developed by Harvey, Leybourne, and Newbold (1997), discussed in Chapter 2. The results of pairwise forecasting comparison of the benchmark AR model with the remaining models by utilizing the MDM test are provided in Table 5.16. The entries in Table 5.16 are the p-values of the MDM test under the null hypothesis which states that $MSFE(linear\, AR model) <= MSFE(any\ other\ model)$. Table 5.16 shows that for the time series of consecutive elapsed times, all p-values of the MDM test are greater than our presumed significance level, i.e., 0.05 and 0.10, showing evidence that no model provides improved 1 through 4 step-ahead forecasts in comparison to the benchmark AR model. Hence, for the time series of raw consecutive elapsed times, the simple linear AR model could be used for modeling and forecasting purposes. For the time series of consecutive distances, by pairwise comparing the linear AR model with nonlinear AAR, SETAR, ANN, and LSTAR models for 1 through 4 step-ahead forecasting horizon, the p-values of the MDM test are greater than 0.05 or 0.10, providing evidence of an improved out-of-sample forecasting gain to the AR model in comparison to nonlinear models. An interesting conclusion emerges from this study: in the pairwise comparison of the ACD and VAR models with the benchmark AR model, the p-values of the MDM test are less than or equal to 0.10, for $h = 3$, 4, and it showed improved forecasting performance of ACD and VAR models in comparison to the benchmark AR model. Thus, for long forecasting horizons, i.e., $h = 3$, 4, the ACD and VAR models provided improved forecasts in comparison to the AR model. However, for short-term forecasting horizons, i.e., $h = 1$, 2, AR forecasts dominates over the ACD and VAR forecasts. Hence, the AR is still the best forecasting device for consecutive distances time series for short-term forecasting horizon, while for long-term forecasting horizon the ACD and bivariate VAR model could be used as the best forecasting device.

For raw seismic time series, only the parameter estimates of the first three

Table 5.17: Estimation results of linear AR model for raw seismic data

Coefficients	Consecutive Elapsed Times			Consecutive Distances		
	Values	Std.	P-values	Values	Std.	P-values
Constant	4.294	0.469	0.000	65.997	3.045	0.000
y_{t-1}	0.117	0.024	0.000	0.370	0.024	0.000
y_{t-2}	0.061	0.024	0.012	-0.096	0.024	0.000
y_{t-3}	0.061	0.024	0.013	—	—	—
y_{t-4}	0.053	0.024	0.031	—	—	—
y_{t-5}	0.065	0.024	0.008	—	—	—
y_{t-6}	-0.003	0.024	0.910	—	—	—
y_{t-7}	0.066	0.024	0.007	—	—	—
y_{t-8}	0.060	0.024	0.014	—	—	—
BIC	7662.407			14366.431		

Table 5.18: Estimation results of linear ACD model for raw seismic data

Coefficients	Consecutive Elapsed Times			Consecutive Distances		
	Values	Std.	P-values	Values	Std.	P-values
Constant	0.192	0.071	0.000	63.51	9.500	0.000
α	0.091	0.022	0.000	0.300	0.035	0.000
β	0.892	0.021	0.000	-0.004	0.103	0.960
LL	5168.098			9415.361		

best forecasting models for respective time series are provided. The estimation results of AR, ACD, VAR, and SETAR models are shown in Tables 5.17 to 5.20. These tables show that almost all coefficients in all these models are significant at a conventional significance level, providing an evidence of high persistence in consecutive elapsed times and consecutive distances time series from the Hindu Kush region of Pakistan. Thus, regarding in-sample fitting, these models performed quite well by parsimoniously capturing the dependence structure present in raw seismic data.

Figure 5.5.8 shows the ACF plot of the residuals of the first three best forecasted models for the whole sample period of raw time series of consecutive elapsed times and consecutive distances. The ACFs of the residuals of the best fitted models fail to indicate any serial dependence. Thus, these best models models are adequate for describing and forecasting the dynamic dependence structure of the time series of consecutive elapsed times and consecutive distances.

Table 5.19: Estimation results of VAR model for raw seismic data

Coefficients	Consecutive Elapsed Times			Consecutive Distances		
	Values	Std.	P-values	Values	Std.	P-values
Constant	7.134	0.412	0.000	57.418	2.982	0.000
E_{t-1}	0.156	0.024	0.000	0.377	0.174	0.031
D_{t-1}	-0.002	0.003	0.560	0.334	0.023	0.000
BIC	31641.315					

Table 5.20: Estimation results of SETAR model for raw seismic data

Coefficients	Consecutive Elapsed Times						Consecutive Distances					
	Low regime			High regime			Low regime			High regime		
	Values	Std.	P-values	Values	Std.	P-values	Values	Std.	P-values	Values	Std.	P-values
Constant	3.379	0.538	0.000	8.873	0.767	0.000	46.928	3.148	0.000	78.618	4.877	0.000
y_{t-1}	0.145	0.029	0.000	0.044	0.046	0.332	0.564	0.031	0.000	0.116	0.034	0.001
y_{t-2}	0.162	0.060	0.008	—	—	—	—	—	—	—	—	—
y_{t-3}	0.034	0.028	0.216	—	—	—	—	—	—	—	—	—
y_{t-4}	0.073	0.028	0.009	—	—	—	—	—	—	—	—	—
y_{t-5}	0.070	0.027	0.010	—	—	—	—	—	—	—	—	—
y_{t-6}	0.008	0.028	0.770	—	—	—	—	—	—	—	—	—
y_{t-7}	0.093	0.028	0.001	—	—	—	—	—	—	—	—	—
y_{t-8}	0.068	0.027	0.013	—	—	—	—	—	—	—	—	—
BIC	7673.472						14293.742					
Threshold value	15.680						96.430					

The observed and predicted values of the time series of consecutive elapsed times and consecutive distances are plotted in Figure 5.5.9. We only plotted the predicted values of the first three best forecasting models for respective seismic time series. This Figure shows that the first best forecasted model for the time series of consecutive elapsed times is the AR model, while for consecutive distances time series the first two best forecasted models are the VAR and ACD, and these models show a forecasted pattern similar to observed time series for consecutive distances. In short, for both raw time series of consecutive elapsed times and consecutive distances, there is not much forecasting gain to ACD model in comparison to the benchmark AR model.

Figure 5.5.8: ACFs of the residuals from the best forecasting models for raw seismic data

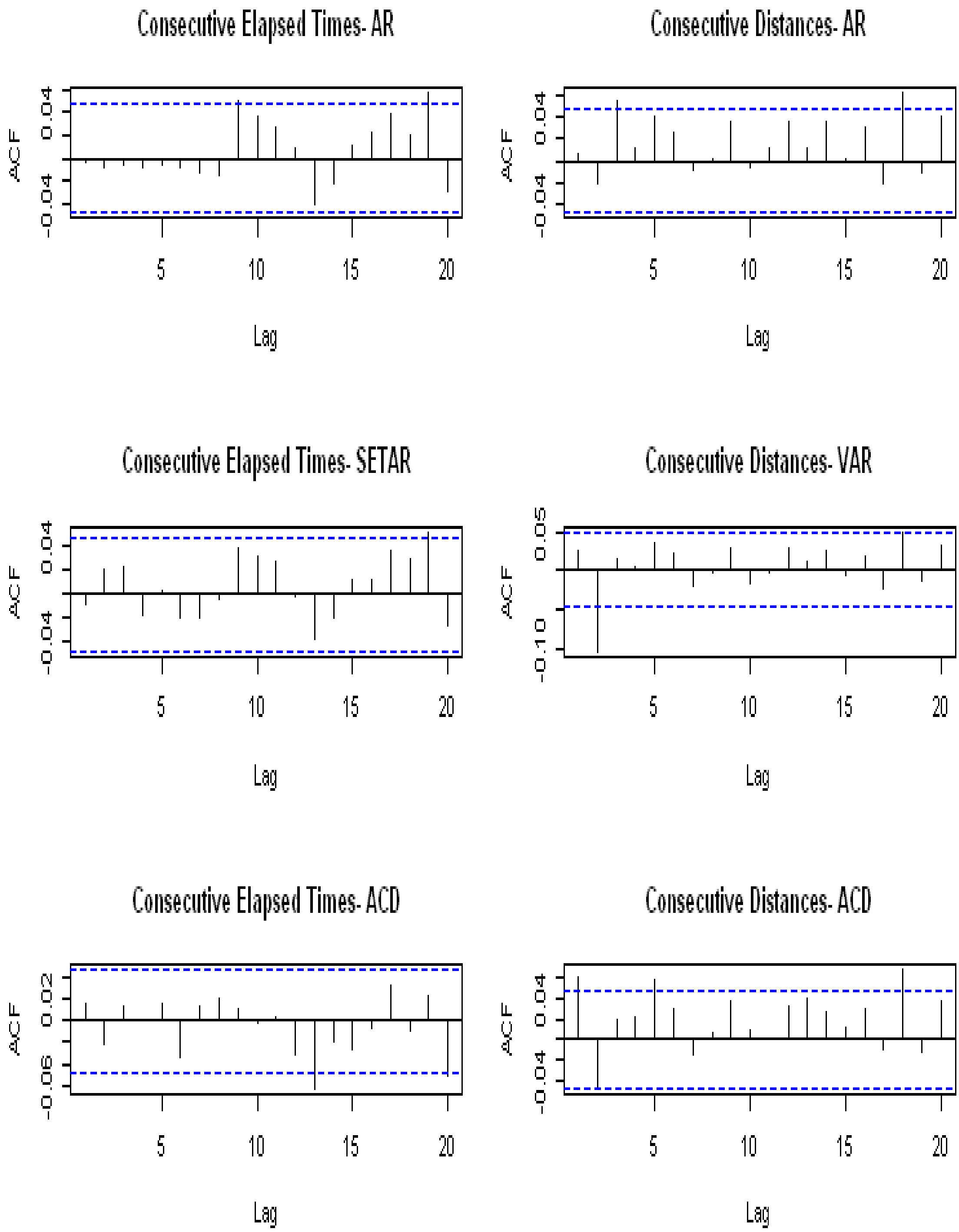
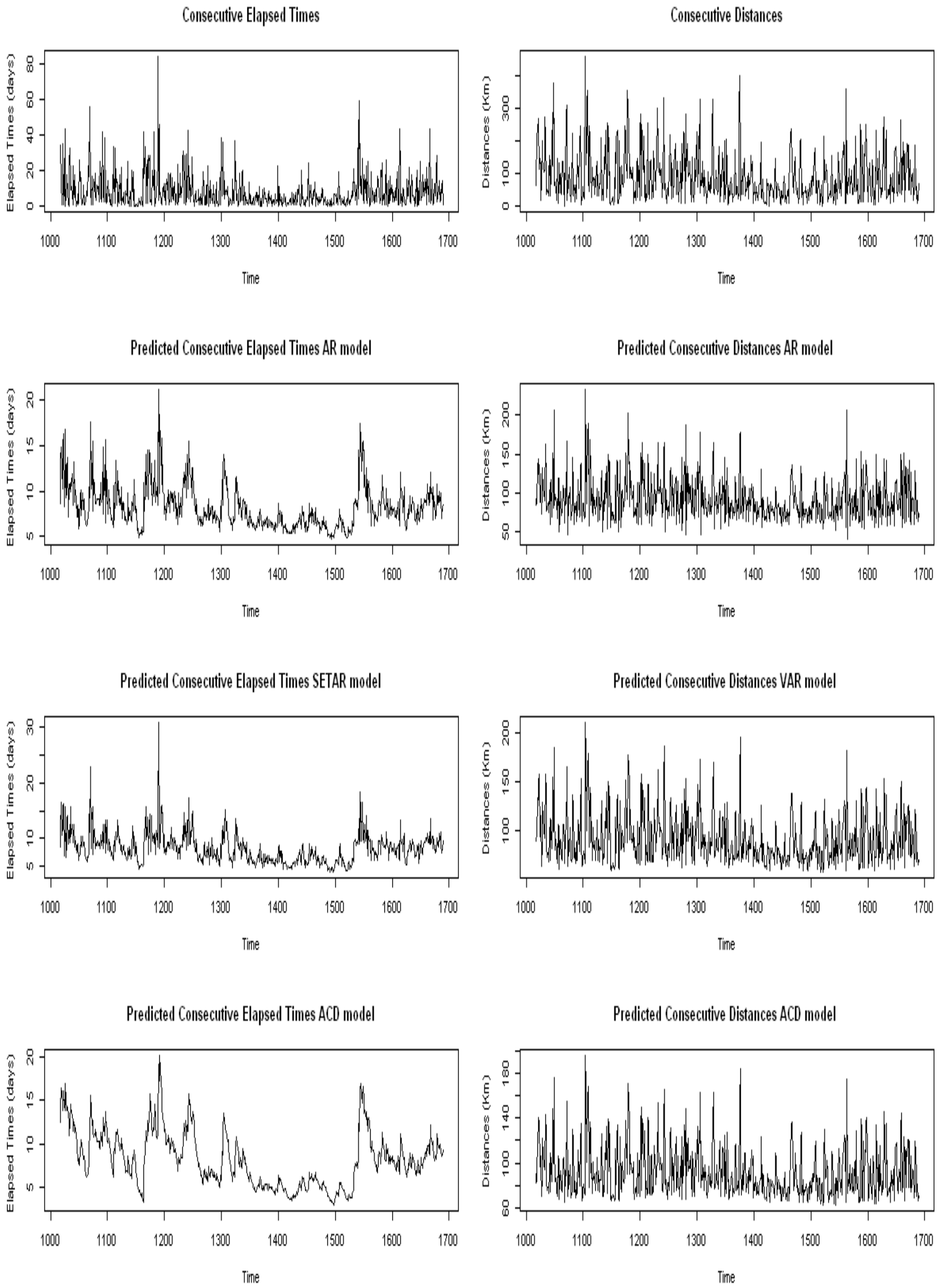


Figure 5.5.9: Observed and 1 step-ahead forecasted values for raw seismic data



5.6 Conclusions

In recent years, research in statistical seismology has been a growing subject all over the world, but still more research in this field is necessary. The current study is an attempt to bridge the research gap between statistics and seismology by describing and forecasting the seismic data from the Hindu Kush region of Pakistan through the application of linear and nonlinear time series models. Four seismic time series from the Hindu Kush region of Pakistan, i.e., earthquake magnitudes, consecutive elapsed times, consecutive distances between earthquakes, and monthly earthquake counts have been modeled through linear and nonlinear time series models. We generated 1 through 4 step-ahead forecasts from linear and nonlinear time series models for raw and logarithmically transformed seismic data and compared the forecasts from nonlinear models with benchmark linear AR model. We also compared the forecasting performance of TAR and LSTAR models with internal and external threshold variables for logarithmically transformed seismic data.

According to ADF test, all variables are stationary at level. Tsay (1986) and Keenan (1985) test of nonlinearity rejected the linearity hypothesis, provided that all four seismic time series are nonlinear. After estimating the linear and nonlinear models with internal and external threshold variables under expanding window scheme, we obtained the 1 through 4 step-ahead forecasts for each of these models.

The MDM test has provided results consistent with those provided by the MSFE criteria for logarithmically transformed seismic time series. According to the p-values provided by the MDM test for earthquake magnitudes time series, TAR and LSTAR models with consecutive distances as an external threshold variable have improved forecasting gain over linear AR model respectively for $h = 1, 2$ and $h = 1$ at 10% significance level. Similarly for consecutive elapsed times, almost all p-values of the MDM test are greater than 0.05, providing evidence that in terms of out-of-sample forecasting performance none of the nonlinear models outperform the linear AR model. For consecutive distances time series, AAR and SETAR models render more accurate 1 step-ahead forecasts than benchmark linear AR model. Comparing the AR and TAR-Ext(elapsed time) models pairwise for consecutive distances time series, it has been observed that the p-values of the MDM test are smaller than 0.05 for 1 through 4 step-ahead forecasts. Thus, TAR-Ext(elapsed times) model provides more accurate forecasts over linear AR model for $h = 1, 2, 3, 4$. Similarly for monthly earthquake counts time series, according to MDM test SETAR model provides more accurate forecasts over linear benchmark AR model for $h = 1, 2$. Only for consecutive distance time series, VAR model provides improved 2 through 4 step-ahead forecasts over linear

AR model at 10% significance level. Furthermore, according to the MSFE criterion and MDM test, TVAR model predicts the time series of consecutive elapsed times fairly well by providing improved out-of-sample forecasting accuracy over benchmark AR model for 1 through 4 step-ahead forecasts.

In the second part of this chapter, we modeled raw seismic data and compared the forecasting performance of linear ACD, VAR, and nonlinear models with a linear benchmark AR model. According to the MSFE values and MDM test, we observed that no model dominates the AR model for the time series of consecutive elapsed times. However, for consecutive distances time series, only for $h = 3$, 4 ACD and VAR models provide improved forecasts over the AR model. On the whole, the AR model appears to be the best forecasting device to model and forecast the raw seismic time series of consecutive elapsed times and consecutive distances between successive earthquakes occurred in the Hindu Kush region of Pakistan. Thus, the use of ACD model does not seem suitable for modeling and predicting the raw seismic time series of elapsed times and distances from the Hindu Kush region of Pakistan.

Hence, a general conclusions can be made that for the Hindu Kush region of Pakistan it is more plausible to use threshold models with a combination of external threshold variables to predict the logarithmically transformed seismic time series of earthquake magnitudes and consecutive distances. For the time series of elapsed times (logarithmically transformed), the linear AR model is the best forecasting device among all potential candidates.

In general, we introduced an extension of time series models in the field of seismology and presented a time series analysis by applying linear and nonlinear time series models. In addition, this study contributes to the field of statistical seismology by analyzing the seismic activity in the Hindu Kush region of Pakistan. The results of this study indicate that when choosing the best forecasting model from a large family of models, the use of external threshold variables in threshold models (mainly for the time series of consecutive distances) may substantially improve the precision of forecasts. This study is particularly important for Pakistan, because it facilitates the work of earthquake engineers and earthquake prediction analysts. The results of this study could be also utilized in seismic hazard and risk analysis.

References

Afonso, A., J. Baxa and M. Slavik (2011). Fiscal Developments And Financial Stress: A Threshold VAR Analysis, Institute of Economic Studies, working paper No.1319.

Aleem, A. and A. Lahiani (2013). A threshold vector autoregression model of exchange rate pass-through in Mexico, Research in International Business and Finance 30, 24-33.

Amei, A., W. Fu and C-H. Ho (2012). Time Series Analysis for Predicting the Occurrences of Large Scale Earthquakes, International Journal of Applied Science and Technology 2 (7), 64-75.

Andersen, G. T., T. Bollerslev, X. F. Diebold and P. Labys (2003). Modeling and forecasting realized volatility, Econometrica 71, 529-626.

Andersen, T. G. and T. Bollerslev (1997). Intraday periodicity and volatility persistence in financial markets, Journal of Empirical Finance 4, 115-158.

———. and ———. (1998). Answering the sceptics: yes standard volatility models do provide accurate forecasts, International Economic Review 39, 885-905.

Andersen, T. G., T. Bollerslev and F. X. Diebold (2007). Roughing It Up: Including Jump Components in the Measurement, Modeling and Forecasting of Return Volatility, Review of Economics and Statistics 89, 701-720.

Andersen, T. G., T. Bollerslev and X. Huang (2011). A reduced form framework for modeling volatility of speculative prices based on realized variation measures, Journal of Econometrics 160, 176-189.

Andrews, D. W. K. and W. Ploberger (1994). Optimal tests when a nuisance parameter is present only under the alternative, Econometrica 62, 1383-1414.

Arnold, R. A. (2013). Book: Economics, California State University, San Marcos.

Arshanapalli, B. and J. Doukas (1993). International Stock Market Linkages: Evidence from the Pre and Post-October 1987 Period, Journal of Banking and Finance 17(1), 193-208.

Baillie, R., T. Bollerslev and H. Mikkelsen (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity, Journal of Econometrics 74, 3-30.

Banerjee, A. and M. Marcellino (2005). Are there any reliable leading indicators for US Inflation and GDP Growth?, International Journal of Forecasting 22, 137-151.

Bauwens, L. (2006). Econometric Analysis of Intra-Daily Activity on Tokyo Stock Exchange, Monetary and Economic Studies 24, 1-23.

Beaudry, P. and G. Koop (1993). Do recessions permanently change output?, Journal of Monetary Economics 31, 149-163.

Bhattacharya, R. N., V. K. Gupta and E. Waymire (1983). The Hurst effect under trends, Journal of Applied Probability 20, 649-662.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity, Journal of Econometrics 31, 307-27.

Bollerslev, T., U. Kretschmer, C. Pigorsch and G. E. Tauchen (2009). A Discrete-Time Model for Daily S&P 500 Returns and Realized Variations: Jumps and Leverage Effects, Journal of Econometrics 150 (2), 151-166.

Brooks, C. (1996). Testing for nonlinearity in daily sterling exchange rates, Applied Financial Economics 6, 307-17.

Brown, B. Y. and R. S. Mariano (1989). Predictors in dynamic nonlinear models: large sample behaviour, Econometric Theory 5, 430-52.

Chan, K. S. (1993). Consistency and Limiting Distribution of the Least Squares Estimator of a Threshold Autoregressive Model, The Annals of Statistics 21, 520-533.

Chan, K. S. and H. Tong (1986). On estimating thresholds in autoregressive models, Journal of Time Series Analysis 7, 179-190.

Chappell, D., J. Padmore, P. Mistry and C. Ellis (1996). A threshold model for the French Franc/Deutschmark exchange rate, Journal of Forecasting 15, 155-164.

Chen, C. and G. Tiao (1990). Random level-shift time series models, arima approximations, and level-shift detection, Journal of Business and Economics Statistics 8, 83-97.

Chen, R. and R. S. Tsay (1993). Nonlinear Additive ARX Models, Journal of the American Statistical Association 88, 955-967.

Cheung, Y. L. and S. C. Mak (1992). The International Transmission of Stock Market Fluctuations between the Developed Markets and the Asian - Pacific Markets, Applied Financial Economics 2, 43-47.

Clark, J. (2008). Why does the Fed change the interest rate?, HowStuffWorks notes.

Clements, M. P. and J. Smith (1997). The performance of alternative forecasting methods for SETAR models, International Journal of Forecasting 13, 463-75.

———. and ———. (1999). A monte carlo study of the forecasting performance of empirical SETAR models, Journal of Applied Econometrics 14, 123-42.

Corsi, F. (2009). A Simple Approximate long-memory Model of Realized Volatility, Journal of Financial Econometrics 7, 174-196.

Corsi, F., U. Kretschmer, S. Mittnik and C. Pigorsch (2005). The Volatility of Realized Volatility, Econometric Reviews 27, 46-78.

Crato, N. and P. J. F. de Lima (1994). Long-memory and nonlinearity: a time series analysis of stock returns and volatilities, Managerial Finance 20, 49-67.

Dacco, R. and S. Satchell (1999). Why do regime-switching models forecast so badly?, Journal of Forecasting 18(1), 1-16.

D'Amico, S. (2012). Earthquake research and analysis Statistical studies, Observations and Planing, Book Published by InTech.

Davies, R. B. (1977). Hypothesis testing when a nuisance parameter is present only under the alternative, Biometrika 64, 247-254.

———. (1987). Hypothesis testing when a nuisance parameter is present only under the alternative, Biometrika 74, 33-43.

Diebold, F. X. and R. S. Mariano (1995). Comparing predictive accuracy, Journal of Business and Economic Statistics 13, 253-63.

Dinghai, Xu. (2010). A Threshold Stochastic Volatility Model with Realized Volatility, Working Papers from University of Waterloo, Department of Economics.

Engle, R. and A. Smith (1999). Stochastic permanent breaks, Review of Economics and Statistics 81, 533-574.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, Econometrica 50, 987-1007.

———. (2000). The Econometrics of Ultra-High Frequency Data, Econometrica 68, 1-22.

Engle, R. F. and J. R. Russell (1998). Autoregressive conditional duration: a new model for irregularly spaced transaction data, Econometrica 66, 1127-1162.

Feng, H. and J. Liu (2003). A SETAR model for Canadian GDP: non-linearities and forecast comparisons, Applied Economics, Taylor and Francis Journals 35 (18), 1957-1964.

Feng, X. T., M. Seto and K. Katsuyama (1997). Neural dynamic modeling on earthquake magnitude series, Geophysical Journal International 128, 547-556.

Ferraes, S. G. (1968). Statistical relation of magnitude to frequency of occurrence for earthquakes in Mexico City, Earth and Planetary Science Letters 3, 449-452.

———. (1973). Earthquake magnitude probabilities and statistical independence for Mexico City earthquakes, Bulletin of Seismological Society of America 63, 1913-1919.

Franses, P. H. and D. van Dijk (2000). Non-linear time series models in empirical finance, Cambridge, New York.

Franses, P. H. and K. van Griensven (1998). Forecasting exchange rates using neural networks for technical trading rules, Studies in Nonlinear Dynamics and Econometrics 2, 109-16.

Gencay, R. (1996). Non-linear prediction of security returns with moving average rules, Journal of Forecasting 15, 165–74.

Georg, M. G. (2010). Time Series Analysis of long-memory versus Structural Breaks: A Time-Varying Memory Approach, Book Verlag Dr. Mueller, 1-120.

Ghorbel, A. and A. Trabelsi (2013). The impact of global financial crisis on the dependence structure of equity markets and on risk management, International Journal of managerial and financial accounting 5, 1-32.

Ghysels, E. and J. Jasiak (1998). GARCH for Irregularly Spaced Financial Data: The ACD-GARCH Model, Studies in Nonlinear Dynamics & Econometrics 2(4), 133-149.

Gil-Alana, L. A. (2008). Fractional integration and structural breaks at unknown periods of time, Journal of Time Series Analysis 29, 163-185.

Goetzmann, W. N., R. G. Ibbotson and L. Peng (2001). A new historical database for the NYSE 1815 to 1925: Performance and predictability, Journal of Financial Markets 4, 1-32.

Goldman, E., J. Nam, H. Tsurumi and J. Wung (2013). Regimes and long-memory in realized volatility, Studies in Nonlinear Dynamics and Econometrics 17, 521-549.

Granger, C. W. J. (1966). The Typical Spectral Shape of an Economic Variable, Econometrica 34, 150-161.

Granger, C. W. J. and A. Andersen (1978). An Introduction to Bilinear Time Series Models, Gottingen, ISBN-13: 978-3525112397.

Granger, C. W. J. and T. Teräsvirta (1993). Modeling Nonlinear Economic Relationships, Oxford: Oxford University Press.

———. and ———. (1999). Simple nonlinear time series model with misleading linear properties, Economics Letters 62, 161-165.

Haefke, C. and C. Helmenstein (1996a). Forecasting Austrian IPOs: an application of linear and neural network error-correction models, Journal of Forecasting 15, 237-51.

Hanks, T. C. and Kanamori, H. (1979). Moment magnitude scale, Journal of Geophysical Research 84, 2348-2350.

Hansen, B. E. (1996). Inference when a nuisance parameter is not identified under the null hypothesis, Econometrica 64, 413-30.

———. (1997). Inference in TAR models, Studies in Nonlinear Dynamics and Econometrics 2, 1-14.

———. (2000). Sample splitting and threshold estimation, Econometrica Härdle,W., 1990. Applied Nonparametric Regression, Cambridge: Cambridge University Press.

Harvey, D. I. (1997). The Evaluation of Economic Forecasts, Ph.D. thesis.

Harvey, D. I., S. Leybourne and P. Newbold (1997). Testing the equality of prediction mean squared errors, International Journal of Forecasting 13, 281-291.

Hastie, T. J., and R. J. Tibshirani (1990). Generalized Additive Models, New York: Chapman and Hall.

Hinich, M. J. and D. M. Patterson (1985). Evidence of nonlinearity in daily stock returns, Journal of Business & Economic Statistics 3, 69-77.

Ho, C-H. (2008). Empirical recurrence rate time series for volcanism: Application to Avachinsky volcano, Russia, Journal of Volcanology and Geothermal Research 173 (1–2), 15-25.

Hornik, K., M. Stinchcombe and H. White (1989). Multilayer feedforward networks are universal approximators, Neural Networks 2, 359-66.

Hsieh, D. A. (1989). Testing for nonlinear dependence in daily foreign exchange rates, Journal of Business 62, 339-69.

———. (1991). Chaos and nonlinear dynamics: application to financial markets, Journal of Finance 46, 1839-77.

Hubrich, K. and R. J. Tetlow (2012). Financial stress and economic dynamics: The transmission of crises, The Federal Reserve Board, Finance and economics discussion series 20, 12-82.

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

Juan, A-S. and L. Guzman-Vargas (2013). Earthquake magnitude time series: scaling behavior of visibility networks, The European Physical Journal B 86 (11).

Kahraman, U., Y. Asar, H. Basbpzkurt, S. Akogul and A. Genc (2012). Statistical Modeling of Seismicity of Van Region by Using SETAR Model, Journal of Selcuk University Natural and Applied Science1 (3), 40-53.

Kanamori, H. (1977). The energy release in great earthquakes, Journal of Geophysical Research 82 (20), 2981-2876.

Keenan, D. M. (1985). A Tukey nonadditivity-type test for time series nonlinearity, Biometrika 72, 39-44.

Kontonikas, A., R. MacDonald and A. Saggu (2012). Stock market reaction to fed funds rate surprises: state dependence and the financial crisis, Journal of Banking and Finance 37(11), 4025-4037.

Kräger, H. and P. Kugler (1993). Nonlinearities in foreign exchange markets: a different perspective, Journal of International Money and Finance 12, 195-208.

Kuan, C. M. and T. Liu (1995). Forecasting exchange rates using feedforward and recurrent neural networks, Journal of Applied Econometrics 10, 347-64.

Kudlacek, S. (2008). The Interest Rate and Related Macro Economic Variables, Notes.

LeBaron, B. (1992). Some relationships between volatility and serial correlations in stock market returns, Journal of Business 65, 199-219.

Leybourne, S., P. Newbold and D. Vougas (1998). Unit roots and smooth transitions, Journal of Time Series Analysis 19, 83-97.

Li, C. W. and W. K. Li (1996). On a double threshold autoregressive heteroskedastic time series model, Journal of Applied Econometrics 11, 253-274.

Li, W. K. and K. Lam (1995). Modeling asymmetry in stock returns by a threshold ARCH model, The Statistician 44, 333-341.

Lin, J. L. and C. W. J. Granger (1994). Forecasting from nonlinear models in practice, Journal of Forecasting 13, 1-9.

Ling-Yi, L. and L. De-Fu (1985). Threshold modeling on the earthquake magnitude series, Chinese Journal of Geophysics.

Liu, C. and J. M. Maheu (2008). Are there structural breaks in realized volatility?, Journal of financial econometrics 6, 326-360.

———. and ———. (2009). Forecasting Realized Volatility: A Bayesian Model Averaging Approach, Journal of Applied Econometrics 24, 709-733.

Liu, J., W. K. Li and C. W. Li (1997). On a threshold autoregression with conditional heteroskedastic variances, Journal of Statistical Planning and Inference 62 (2), 279-300.

Lobato, I. N. and N. E. Savin (1998). Real and spurious long-memory properties of stock market data, Journal of Business and Economic Statistics 16, 261-268.

Mandelbrot, B. and J. W. van Ness (1968). Fractional Brownian motions, fractional noises and applications, SIAM Review 10 (4), 422-437.

Manganelli, S. (2005). Duration, Volume and Volatility Impact of Trades, Journal of Financial Markets 8, 377-399.

Martinez, M. D., X. Lana., A. M. Posadas and L. Pujades (2005). Statistical distribution of elapsed times and distances of seismic events: the case of the Southern Spain seismic catalogue, Nonlinear Processes in Geophysics 12, 235-244.

McAleer, M. and M. C. Medeiros (2008a). Realized Volatility: A Review, Econometric Reviews 27, 10-45.

———. and ———.(2008b). A multiple regime smooth transition Heterogeneous Autoregressive model for long-memory and asymmetries, Journal of Econometrics 147, 104-119.

Michael, A. J. and S. Wiemer (2010). Theme I – Introductory Material CORSSA: the Community Online Resource for Statistical Seismicity Analysis.

Morales-Esteban, A., F. Martínez-Álvarez, A. Troncoso, J. L. Justo and C. Rubio-Escudero (2010). Pattern recognition to forecast seismic time series, Expert Systems with Applications 37, 8333-8342.

Morely, J. (2013). Macroeconomics, Non-linear Time Series in, Encyclopedia of Complexity and Systems Science.

Moustra, M., M. Avraamides, and C. Christodoulou (2011). Artificial neural networks for earthquake prediction using time series magnitude data or Seismic Electric Signals, Expert systems with applications 38, 15032-15039.

Nowroozi, A. A. (1971). Seismotectonics of the Pakistan plateau, eastern Turkey, Caucasus, and Hindu Kush regions, Bulletin of Seismological Society of America 61, 317-341.

———. (1972). Focal mechanisms of earthquake in Persia, Turkey, West Pakistan, and Afghanistan and plate tectonics of the Middle East, Bulletin of Seismological Society of America 62, 823-850.

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

———. (1998). Space-time point-process models for earthquake occurrences, Annals of the Institute of Statistical Mathematics 50(2), 379-402.

Panakkat, A. and H. Adeli (2007). Neural network models for earthquake magnitude prediction using multiple seismicity indicators, International Journal of Neural Systems 17(1), 13-33.

Perron, P. (1989). The great crash, the oil price shock, and the unit root hypothesis, Econometrica 57, 1361-1401.

Perron, P. and Z. Qu (2006). An analytical evaluation of the log-periodogram estimate in the presence of level shifts and its implications for stock returns volatility, Working paper, Boston University.

Pesaran, M. H. and A. Timmerman (2005). Small sample properties of forecasts from autoregressive models under structural breaks, Journal of Econometrics 129, 183-217.

Pesaran, M. H. and S. Potter (1997). A Floor and Ceiling Model of US Output, Journal of Economic Dynamics and Control 21, 661-695.

Pippenger, M. K. and G. E. Goering (1998). Exchange rate forecasting: Results from a threshold autoregressive model, Open Economies Review 9, 157-170.

Pötscher, B. M. and I. V. Prucha (1997). Dynamic Nonlinear Econometric Models Asymptotic Theory, Berlin: Springer-Verlag.

Potter, S. M. (1995). A Nonlinear approach to US GNP, Journal of Applied Econometrics 10, 109-125.

———. (1999). Nonlinear Time Series Modeling: An Introduction, Federal Reserve Bank of New York.

Preve, D., A. Eriksson and J. Yu (2009). Forecasting Realized Volatility Using A Nonnegative Semiparametric Model, Working Papers, Singapore Management University, School of Economics.

Priestley, M. B. (1980). State-dependent models: a general approach to non-linear time series analysis, Journal of Time Series Analysis 1, 47-71.

———. (1988). Nonlinear and Nonstationary Time Series Analysis, London: Academic Press.

Quandt, R. (1958). The estimation of the parameters of a linear regression system obeying two separate regimes, Journal of the American Statistical Association 53, 873-880.

R Development Core Team (2012). R: A language and environment for statistical computing, Vienna Austria: R Foundation for Statistical Computing, ISBN 3-900051-07-0.

Richter, C. F. (1935). An instrumental earthquake magnitude scale, Bulletin of the Seismological Society of America 25 (1-2), 1-32.

Rothman, P. (1998). Forecasting asymmetric unemployment rates, Review of Economics and Statistics 80, 164-168.

Sarantis, N. (1999). Modeling non-linearities in real effective exchange rat es, Journal of International Money and Finance 18 (1), 27-45.

Scheinkmann, J. A. and B. LeBaron (1989). Nonlinear dynamics and stock returns, Journal of Business 62, 311-37.

Schwert, G. W. (1990). Stock volatility and the crash of 87, Review of Financial Studies 3, 77-102.

Scordilis, E. M. (2006). Empirical global relations converting M_s and m_b to moment magnitude, Journal of Seismology 10, 225-236.

Sibbertsen, P. (2003). Log-periodogram estimation of the memory parameter of a long-memory process under trend, Statistics and Probability Letters 61, 261-268.

Sims, C. and T. Zha (2006). Were there regimes shifts in U.S. monetary policy?, American Economic Review 96, 54-81.

Sinnott, R. W. (1984). Virtues of the Haversine, Sky and Telescope 68 (2), 159.

So, M. K. P., W. K. Li and K. Lam (2002). A Threshold Stochastic Volatility Model, Journal of Forecasting 21, 473-500.

Stock, J. H. and M. W. Watson (1999). A comparison of linear and nonlinear univariate models for forecasting macroeconomic time series, in Engle, R. and White, R. (eds), Cointegration, causality, and forecasting: A festschrift in honor of Clive W. J. Granger, Oxford: Oxford University Press, 1-44.

———. and ———. (2003). Forecasting output and inflation: the role of asset prices, Journal of Economic Literature 41(3), 788-829.

Swanson, N. R. and H. White (1997). A model selection approach to real-time macroeconomic forecasting using linear models and artificial neural networks, Review of Economics and Statistics 79, 540-50.

Tena, J. D. and A. R. Tremayne (2009). Modeling monetary transmission in UK manufacturing industry, Economic Modeling 26, 1053-1066.

Tepei, L. (1998). An application of the threshold autoregressive model to estimation of seismicity intensity around 2000 year in Guangdong area, South China Journal of Seismology 18 (3).

Teräsvirta, T. (1994). Specification, Estimation and Evaluation of Smooth Transition Autoregressive Models, Journal of the American Statistical Association 89, 208-218.

———. (1998). Modelling economic relationships with smooth transition regressions, in A. Ullah and D.E.A. Giles (eds.), Handbook of Applied Economic Statistics, New York: Marcel Dekker.

———. (2005). Forecasting with Nonlinear Models, in G. Elliott, C. W. J. Granger and A. Timmermann (eds.), Handbook of Economic Forecasting.

Teräsvirta, T. and H. M. Anderson (1992). Characterizing nonlinearities in business cycles using smooth transition autoregressive models, Journal of Applied Econometrics 7, 119-136.

Teräsvirta, T., D. van Dijk and M. C. Medeiros (2005). Linear models, smooth transition autoregressions, and neural networks for forecasting macroeconomic time series: A re-examination, International Journal of Forecasting 21(4), 755-774.

Tong, H. (1978). On a threshold model, Pattern Recognition and Signal Processing, Amsterdam: Sijthoff & Noordhoff, 101-41.

———. (1983). Threshold Models in Nonlinear Time Series Analysis, New York: Springer Lecture Notes in Statistics.

———. (1990). Non-Linear Time Series: A Dynamical Systems Approach, Oxford: Oxford University Press.

Tong, H. and K. S. Lim (1980). Threshold autoregressions, limit cycles, and cyclical data, Journal of the Royal Statistical Society B 42, 245-92.

Tsay, R. S. (1986). Nonlinearity tests for time series, Biometrika 73, 461–6.

———. (1998). Testing and modeling multivariate threshold models, Journal of the American Statistical Association 93, 1188-1202.

———. (2009). Autoregressive conditional duration models in Palgrave Handbook in Econometrics 2, eds. T. C. Mills and K. Patterson, Chapter 21, 1004-1024.

Tukey, J. (1949). One degree of freedom for non-additivity, Biometrics 5, 232-242.

van Dijk, D. and P. H. Franses (2003). Selecting a nonlinear time series model using weighted tests of equal forecast accuracy, Econometric Institute Report 315, Erasmus University Rotterdam.

van Dijk, D., P. H. Frances., M. P. Clements and J. Smith (2003). On SETAR nonlinearity and forecasting, Journal of Forecasting 22(5), 359-375.

Vassiliou, M. and H. Kanamori (1982). The Energy Release in Earthquakes, Bulletin of the Seismological Society of America 72, 371-387.

Vere-Jones, D. (2006). The development of statistical seismology: A personal experience, Tectonophysics 413(1-2), 5-12.

———. (2010). Introductory Material: How to educate yourself as a statistical seismologist, 1-14.

Wang, C., H. Liu, J-F. Yao, and R. A. Davis (2014). Self-excited Threshold Poisson Autoregression, Journal of the American Statistical Association, published online.

White, H. (2005). Approximate Nonlinear Fore casting Methods, in G. Elliott, C. W. J. Granger and A. Timmermann (eds.), Handbook of Economic Forecasting.

Wiener, N. (1958). Nonlinear Problems in Random Theory, Cambridge, MA: MIT Press.

Wong, W. K., J. Penm, R. D. Terrell and K. Y. C. Lim (2004). The Relationship Between Stock Markets of Major Developed Countries And Asian Emerging Markets, Journal of Applied Mathematics and Decision Sciences 8 (4), 201-218.

Wood, S. N. (2001). mgcv:GAMs and Generalized Ridge Regression for R, R News 1(2), 20-25.

Yang, Li-M., Te-L. Shi and Da-Q. Guo (1995). Mathematical modeling of nonlinear behavior of seismicity, Acta Seismologica Sinica 8 (2), 223-232.

Zivot, E. and J. Wang (2006). Modeling financial time series with S-Plus, second edition, Springer, USA.

Eidesstattliche Versicherung

(Siehe Promotionsordnung vom 12.07.11, § 8, Abs. 2 Pkt. .5.)

Hiermit erkläre ich an Eidesstatt, dass die Dissertation von mir selbstständig, ohne unerlaubte Beihilfe angefertigt ist.

Khan, Muhammad Yousaf
Name, Vorname

München, 17.06.2015 ________________________________
Ort, Datum Unterschrift Doktorand/in