ADVANCED FORECASTING MODEL ON LAND MARKET VALUE BASED ON USA REAL ESTATE MARKET

A Dissertation By

Lei Wang

Master of Science, University Of Southern Mississippi, 2014

Master of Arts, XI'AN University of Finance and Economics, 2012

Bachelor of Arts, Inner Mongolia University of Finance And Economics, 2009.

Submitted To The Department of Mathematics and Statistics
and the faculty of the Graduate School of
Wichita State University
in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

December 2019

[1]Copyright for original cited results are retained by the original authors. By virtue of their appearance in this dissertation, results are free to use with proper attribution

ADVANCED FORECASTING MODEL ON LAND MARKET VALUE BASED ON USA REAL ESTATE MARKET

The following faculty members have examined the final copy of this dissertation for form and content, and recommend that it be accepted in partial fulfillment of the requirement for the degree of Doctor of Philosophy with a major in Applied Mathematics.

Tianshi Lu, Committee Chair

Xiaomi Hu, Committee Member

Lop-Hing Ho, Committee Member

Mahmoud. Edwin Sawan, Committee Member

Gamal Weheba, Committee Member

Accepted for the College of Liberal Arts and Sciences

Andrew Hippisley, Dean

Accepted for the Graduate School

Coleen Pugh, Dean

DEDICATION

I would like to dedicate the dissertation to my advisor, Dr.Lu, who showed uncommon patience and provided unprecedented support.

ACKNOWLEDGEMENTS

I would like to thank the members of my committee, my friends and my family for the support and encouragement needed to finish my dissertation.

I am sincerely grateful to my advisor, Dr.Lu, for his patience, guidance and instruction. I would like to give special thanks Dr.Edwin.sawan for the support and mentoring he provided over the course of many years. I would like to thank Dr. Longhofer and Dr. Jenchi.Cheng for his penetrating questions and insights during our conversation and Dr.Gamal Weheba for his suggestions. I also want to thank Dr.Ziqi Sun and Jason Clemens for being a special voice during my research; one of practicality, common sense and joy.

My friends and family always supported my further study and I appreciate all of their corroboration and help. Dr. Tianshi Lu, Dr. Lop-Hing Ho, Dr. Xiaomi Hu and my friend Sean Hennessy were constant sources of encouragement. Finally thanks to Mark Arrasmith, William Ingle and Mr and Mrs Joan Phillips for helping me keep my balance. You all helped me keep in mind how much fun statistics and research can be.

ABSTRACT

This research presents a time series estimation and prediction methods with the use of classic and advanced forecasting tools. Our discussion about different time series models is supported by giving the experimental forecast results, performed on several macroeconomic variables. Also, the main section deal with the experience of using such data in econometric analysis. Besides, the implementation of SAS and R software improve the parameter estimation and forecasting accuracy.

The objective in providing crucial statistical techniques is to enable government and investors to make informed decisions regarding real estate. Most importantly, we obtain how to add value to business and apply skills set real estate in a real world environment. Eventually, the summary of various existing forecasting models can provide information to develop an appropriate forecasting model which describes the inherent feature of the series.

ABBREVIATIONS

ADF	Augmented Dickey-Fuller Test. Eq.(2.52) on Page 50
AIC	Akaike information criterion. An estimator of the relative quality of statistical models for a given set of data. Eq.(2.32) on Page 36
AICc	The small-sample corrected Akaike Information Criterion. Eq.(2.33) on Page 36
ARFIMA	Autoregressive Fractionally Integrated Moving Average. Eq.(2.19) on Page 27
ANN	Artificial Neural Networks. Equation.(3.7) on Page 67
AR-GARCH	The combination of ARCH and GARCH model. Eq.(2.48) on Page 42
ARIMA	Autoregressive integrated moving average. Eq. (2.15) on page 24
ARCH	Auto-regressive Conditional Heteroscedasticity. Eq.(2.40-2.42) on Page 40
BVAR	Bayesian Vector Autoregressive model. Eq.(3.4) on Page 62
CCI	Construction Cost Index and Expense incurred by a contractor. Page 7
CPI	A measure of the weighted average of prices of consumer goods and services. Page 7
d.g.p	Data Generating Process. Page 50
DW	Durbin Watson Test. Page 43
ECM	Error-correction models. Table2.11 on Page 54.
FPE	Final Prediction Error. Eq.(2.34) on Page 36

ABBREVIATIONS (continued)

ETS (M,A,N)	Exponential Smoothing Methods with multiplicative and additive errors. Eq.(2.7-2.10) on Page 20
GARCH	Generalized autoregressive conditional heteroscedasticity. [Eq. (2.47) on page 41]
GDP	The total value of goods and services within a nation over a period. Page 6
HQ	Hannan-Quinn criterion. Eq.(2.34) on Page 36
I(d)	The dth difference. Page. 53
IR	The percentage increase in the price of goods and services. Page 6
LCCI	Log transformation of construction cost index dataset. Eq.(2.2) on Page 18
LCPI	Log transformation of consumer cost index dataset. Eq.(2.2) on Page 18
LGDP	Log transformation of GDP dataset. Eq.(2.2) on Page 18
LIR	Log transformation of inflation rate dataset. Eq.(2.2) on Page 18
LM	Lagrange multiplier. Page. 43
LMV	Aggregate market value of residential land. Eq. (1.1) on Page.5
LLMV	Log transformation of land market value dataset. Eq.(2.2) on Page 18
LPP	Log transformation of population dataset. Eq.(2.2) on Page 18
LPMI	Log transformation of purchase manager index dataset. Eq.(2.2) on Page 18
LUR	Log transformation of unemployment rate dataset. Eq.(2.2) on Page 18.

ABBREVIATIONS (continued)

MCMC	Markov Chain Monta Carlo. Page 28
MAE	Mean absolute Error. Eq.(4.2) on Page 79
MAPE	Mean absolute percentage error. Eq.(4.3) on Page 79
ME	Mean Error. Eq.(4.1) on Page 79
MPE	Mean Percentage Error. Eq.(4.4) on Page 79
MS	Markov-switching Model. Page 4
OFHEO	Office of Federal Housing Enterprise and Oversight. Page.4
OLS	Ordinary Least Square Page.39
PMI	An indicator of the economic health of the manufacturing sector. Page 8
PP	human beings in general or considered collectively. Page 7
RBF	Radius Basis Function. Eq.(3.19) on Page 73
RMSE	Root Mean Squared Error. Eq.(4.5) on Page 79
SBC	Schwarz Bayesian Information Criterion. Eq.(4.6) on Page 79
SC	Schwarz Criterion Eq.(2.35) on Page 36
SVR	Support Vector Regression. Eq.(3.14-3.17) on Page 73
SVM	Support Vector Machine. Page 71
UR	A measure of the prevalence of unemployment. Page 7
VAR (p)	Vector Auto-regression with P order. Eq.(2.23) on Page 29
VECM	Vector Error Correction Model. Eq.(2.55) on Page 55

LIST OF TABLES

LIST OF FIGURES

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

Chapter 1

1 Introduction

1.1 Background

Forecasting is an important problem that spans many fields including business and industry, government, economics, environmental sciences, medicine, social science, politics, and finance. A lot of researchers have been studying time series forecasting for approximately one century in order to get better forecasts for the future. To achieve best forecast accuracy level, various time series forecasting approaches have been proposed in the literature. After 1980s, more sophisticated algorithms could be improved since properties of computers were enhanced. Therefore, new time series forecasting approaches such as ARIMA, ARIFMA and artificial neural networks, fuzzy regression, fuzzy inference systems, and fuzzy time series could be proposed. In the applications, these approaches have proved their success in forecasting real life time series. In addition, hybrid forecasting methods which combine these new approaches have also been improved to obtain more accurate forecasts. In recent years, these advanced time series forecasting methods have been used to forecast real life time series and satisfactory results have also been obtained. The aim of this special issue is to collect and summarize advanced forecasting methods and apply these advanced forecasting model to land market value.

The rise and fall of house prices and housing-related activity such as construction and finance over the first decade of this century was by some measures the most dramatic bubble and bust in recorded history for the United States. Prior to the decline in home prices, there was some controversy over whether the sharp rise in housing values, which lasted until 2007, was actually a formal bubble (an increase in asset prices beyond what could reasonably be justified by market fundamentals). As noted, housing is more important for the economy than at any point in recent memory, so better housing forecasts will be useful.

U.S. commercial banks have a greater proportion of real assets on their balance sheets

than at any time in over 30 years. Better forecasts can help banks, as well as their regulators develop better capital requirements. Many institutional investors hold large amounts of mortgage-backed securities. Prepayment risk is a major concern, and Foster and Van Order find volatility to be a determinant of such risk. Better forecasts would help such institutions better manage such risk. Thus improved forecasts promise important benefits for many parties exposed to housing and mortgage activity.

Davis and Heathcote [30] and Davis and Palumbo [31] estimate price indexes for residential land, while Davis (2009) estimates indexes for both residential and commercial land. These indexes, however, are not based on transaction prices. Instead, Davis and his coauthors infer land prices as a residual from data on real estate prices and construction costs, combined with the assumption that the prices of existing structures always equal their replacement cost. This identifying assumption likely is reasonable over long spans of time, but it may not be valid over shorter periods.

As Case and Schiller [50] point out that the housing market is not an auction market. Prices do not fall to clear the market quickly, as one observes in most asset markets. Selling a home requires agreement between buyer and seller. It is a stylized fact about the housing market that bidask spreads widen when demand drops, and the number of transactions drops sharply. This must mean that sellers resist cutting prices.

Thus prices should exhibit nonlinear patterns of adjustment, and forecasting models which take account of such nonlinearity may do better than those which impose symmetric adjustment to rising and falling prices. Forecasts of housing prices, land market value and construction cost may be more important than at any time for the U.S. economy. As noted, consumption is tied more closely to home prices in the U.S. than in other countries. Housing, both directly and indirectly, is responsible for two-fifths of the jobs created in the U.S. since 2001. Home builders would clearly like to be able to forecast prices to avoid unwanted build-ups in inventories. Banks would certainly like to forecast home prices for optimal portfolio management. Despite the growth of secondary mortgage markets, real estate represents

33.5% of the U.S. banking industries as of the 2006, the highest level since 1973 according to Federal Reserve officials.

1.2 Literature Review

There is an existing literature on home price forecasting using linear techniques. Case and Schiller [49] was the canonical paper on forecasting house prices. The authors were investigating the informational efficiency of the housing market in the United States, and found that prices did not follow a random walk. Price changes are thus persistent and therefore forecastable. Gu [53] confirms that different housing markets display varying autocorrelation patterns, and shows that in the case of California, excess returns are in principle possible using a simple trading rule. Gillen et al. [52] find varying patterns of spatial autocorrelation in different housing markets in greater Philadelphia.

Zhou [55] uses linear cointegration and error-correction models (ECM) to forecast prices for the U.S. Barot and Takala [50] use cointegration analysis for house prices in Finland and Sweden.

Muellbauer and Murphy [56] discuss boom and bust cycles in the UK housing market, and then go on to explicitly mention non-linearities in price changes, which they believe arise from transactions costs. However, the authors build a structural, rather than a forecasting model, the goal of which is to estimate elasticities and other parameters, rather than generate predictions about future price changes.

Thus the Crawford and Fratantoni [18] paper is a major contribution. While the testing of hypotheses and estimation of structural parameters, as was the focus in Muellbauer and Murphy, is important, it is also obviously essential to generate accurate forecasts of future prices, given the role home prices now play for both the financial sector and the economy as a whole. And Crawford and Fratantoni is the first study to highlight the importance of the boombust nature of house prices, and to utilize a particular non-linear model in an attempt to capture this cycle and improve forecasts.

The authors employ quarterly OFHEO house price index data from five states California, Florida, Massachusetts, Ohio and Texas over the period 1979 to 2001. They develop three different types of forecasting models for each state. The first two types are well-known-ARMA and GARCH models. While these types of time series models are very useful for producing forecasts, the authors note that housing prices in many markets exhibit boombust cycles. Thus prices may behave differently depending on what state housing conditions are in. Accordingly, the authors employ a MS model to capture this asymmetry.

The authors thus estimate ARMA, GARCH and MS models for all five states, with the number of parameters and lag orders determined by goodness of fit, or minimum RMSE, within the sample. Initially, results seem to be favorable for the nonlinear MS model, as it has far better fit than either alternative in all of the states. When out-of-sample forecasts are performed, however, results are very different. Three different forecast horizons 2 years, 5 years and 10 years are used for the five states, for a total of 15 different forecasts. Of the 15 forecasts, the MS model has the minimum RMSE in only three.

Thus the nonlinear MS model does not appear to improve forecast accuracy. However, since the CrawfordFratantoni paper was published, other research has revealed that the MS model is particularly ill-suited for forecasting. The authors examination of the literature points to a very relevant finding for the CrawfordFratantoni paper-MS models usually display superior in-sample fit but worst out-of-sample forecasting performance relative to linear models. Thus some nonlinear model may do a better job of capturing the boombust nature of the housing market than standard linear estimation, but the MS approach may not be the particular specification needed. Also, Many studies show that forecasting accuracy can be improved from univariate autoregression by adding national economic variables as driving factors (inputs), such as transfer functions.

Intuitively, a larger model such as the vector autoregression (VAR) that allows intertemporal interdependence. among all variables should perform better than transfer functions and univariate regressions. However, VAR suffers from overparameterization and a modified

VAR (Bayesian VAR) imposing some prior restrictions on parameters sometimes perform better [2, 44].

1.3 Data and Predictor variables

The Land Market Value, defined as the total value of land and quantity data are derived from data on housing values, is an important factor in the estimation of structure costs using price indexes for housing and construction costs. Also, the land market value is mathematically defined as the land rental value, minus land taxes, divided by a capitalization rate.

$$Land\ market\ value = \frac{Land\ rental\ value - land\ taxes}{capitalization\ rate} \tag{1.1}$$

Where Land Rental Value is the annual fee individuals are willing to pay for the exclusive right to use a land site for a period of time. This may include a speculative opportunity cost. Land Taxes is the portion of the land rental value that is claimed for the community. Capitalization Rate is a market determined rate of return that would attract individuals to invest in the use of land, considering all of the risks and benefits which could be realized.

Not only is land rent a potentially important source of public revenue, the tax on land is a means of limiting excessive speculation in land prices. This would ensure that the equal opportunity to be productive would be available to all citizens. With limited money to invest, people could invest in productive equipment and wages, rather than in high land prices which produce no additional tangible wealth.

The formula indicates how simple it would be to translate market value to rental value or vice versa, depending upon the policy of any nation. In the United States and most other countries, land values are estimated and assessed. Land taxes, however, are a portion of land rent.

In the Figure (1.1), the graph shows that the changes of eight selected macroeconomic

variables[2] from 1982 to 2015. Also, the dissertation will fit the different forecasting models with respect to these variables.

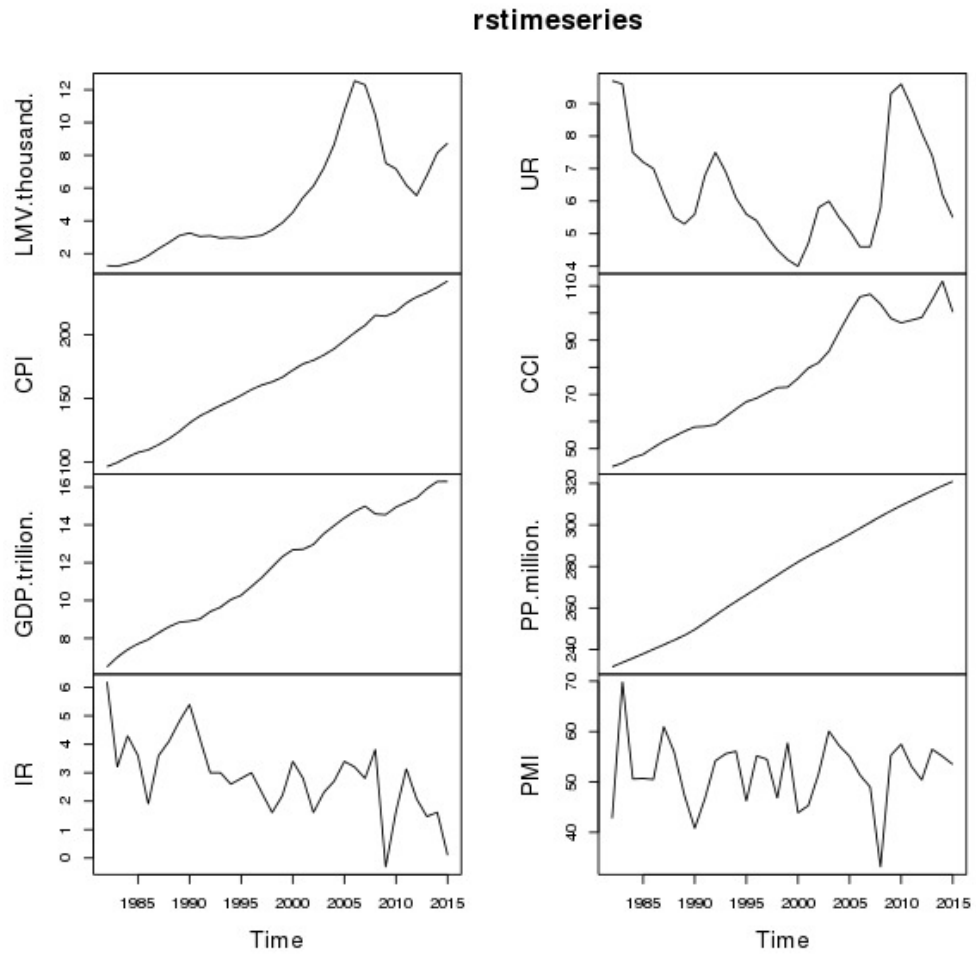


Figure 1.1: Time Series graphing for selected variables

LMV

From the Figure 1.1, we may find that the American land market value shows stable increase from 1982 to 2004,but from 2005 this number increased dramatically and peaked in 2006, 12550 million. In fact, the economic crisis started in 2006 in USA, the economics crisis led to the increased interest.Hence, the LMV rose rapidly. The economy of USA experienced the great recession during this period. Until 2013, the situation recovered and this number rose to 8737.11 million in 2015.

[2]Source: U.S. Bureau of Labor Statistics https://en.wikipedia.org/wiki/Main-Page.

CPI

The CPI is used to adjust income. When the CPI increased rapidly, wages have to increase eventually. The BLS uses the CPI to adjust wages, retirement benefits, tax brackets, and other important economic indicators. Especially, in 2009, the growth of CPI is -0.4, because of the crises influence,the CPI decreased from 215.3 to 214.5.

UR

In the US, the unemployment rate started to increase in the summer 2007, from a level of 4.6% of the labor force (June 2007). It doubled in less than two years and reached a peak of 9.6% in 2009. Since then, it has been very gradually diminishing and was recorded at 8.9% in 2011. The unemployment rate, which rose from 5.5 percent in 2004, is at its highest level since September 2004.

Population

When population growth is strong and long-lasting this encourages development to occur and new housing is build to cope with the increase in demand. To understand this better, we can find it in the graph that the population has a general rise during this period. In 2000, the Population in USA is approximately 280 million, by 2015 this number has reached approximately 321 million in 2015. During this period, the population has increased 14.6%.

As the case of forecasting, the prediction of population fits very well of the actual value. The population of USA is projected to increase to 5.78 in 2018.

CCI

The Figure 1 shows the national construction cost index increase from 2010 to 2015, from low 96.4 to just 100.37. Our research shows a 5.5 percent increase in the national average in construction cost from that January 2014 and December 2014.

PMI

The usefulness of the PMI as an early signal of changes in manufacturing output and GDP. Therefore, an index 50.0 means that the variable is unchanged, a number over 50.0 indicates an improvement while anything below 50.0 suggests a decline. The further away from 50.0 the index is, the stronger the change over the month. For example PMI of 53.5 points in 2015 to a stronger increase in a variable than 50.4 in 2012. In this forecasting model, the trend of PMI will remain around 54 in the next decades.

Time series forecasting uses only information on the variable to be forecast, and makes no attempt to discover the factors which affect its behavior. Therefore it will extrapolate trend and seasonal patterns, but it ignores all other information such as marketing initiatives, competitor activity, changes in economic conditions, and so on. Time series models used for forecasting include ARIMA models, exponential smoothing , structural models and some advanced models.

Time series data are useful when you are forecasting something that is changing over time (e.g., stock prices, sales figures, profits, etc.). such as Daily IBM stock prices, Monthly rainfall, Quarterly sales results for Amazon, Annual Google profits. Anything that is observed sequentially over time is a time series. In this dissertation, regular intervals of time (e.g., hourly, daily, weekly, monthly, quarterly, annually) will be considered as observed time series dataset. Irregularly spaced time series can also occur, but are beyond the scope of this dissertation.

Predictor variables can also be used in time series forecasting. For example, suppose we wish to forecast the LMV in USA from 1982-2015. A model with predictor variables might be of the form

$$LMV = f(GDP, CPI, CCI, UR, IR, PP, PMI)$$

The relationship is not exact—there will always be changes in land market value that

cannot be accounted for by the predictor variables. The error term on the right allows for random variation and the effects of relevant variables that are not included in the model. We call this an explanatory model because it helps explain what causes the variation in land market value.

Because the land market value data form a time series, we could also use a time series model for forecasting. In this case, a suitable time series forecasting equation is of the form

$$LMV_t = f(LMV_{t-1}, LMV_{t-2}, LMV_{t-3}, \cdots, error)$$

where t is the current year, $t+1$ is the next year, $t-1$ is the last year, $t-2$ is two years age and so on. Here, prediction of the future is based on past values of a variable, but not on external variables which may affect the system. Again, the "error" term on the right allows for random variation and the effects of relevant variables that are not included in the model. There is also a third type of model which combines the features of the above two models. For example, it might be given by

$$LMV_{t+1} = f(LMV_t, CCI, UR, PMI, \cdots, error)$$

These types of mixed models have been given various names in different disciplines. They are known as dynamic regression models, panel data models, longitudinal models, transfer function models, and linear system models (assuming f is linear).

An explanatory model is very useful because it incorporates information about other variables, rather than only historical values of the variable to be forecast. However, there are several reasons a forecaster might select a time series model rather than an explanatory model. Firstly, the system may not be understood, and even if it was understood it may be extremely difficult to measure the relationships that are assumed to govern its behavior. Secondly, it is necessary to know or forecast the various predictors in order to be able to forecast the variable of interest, and this may be too difficult. Thirdly, the main concern

may be only to predict what will happen, not to know why it happens. Finally, the time series model may give more accurate forecasts than an explanatory or mixed model.

The model to be used in forecasting depends on the resources and data available, the accuracy of the competing models, and how the forecasting model is to be used.

1.4 Statistical Forecasting Perspective

The thing we are trying to forecast is unknown (or we wouldn't be forecasting it), and so we can think of it as a random variable. For example, the total sales for next month could take a range of possible values, and until we add up the actual sales at the end of the month we don't know what the value will be. So, until we know the sales for next month, it is a random quantity.

Because next month is relatively close, we usually have a good idea what the likely sales values could be. On the other hand, if we are forecasting the sales for the same month next year, the possible values it could take are much more variable. In most forecasting situations, the variation associated with the thing we are forecasting will shrink as the event approaches. In other words, the further ahead we forecast, the more uncertain we are.

When we obtain a forecast, we are estimating the middle of the range of possible values the random variable could take. Very often, a forecast is accompanied by a prediction interval giving a range of values the random variable could take with relatively high probability. For example, a 95% prediction interval contains a range of values which should include the actual future value with probability 95%.

Generally, we will need to distinguish between a forecast or predicted value of y_t that was made at some previous time period, say, $t-\tau$, and a fitted value of y_t that has resulted from estimating the parameters in a time series model to historical data. Note that τ is the forecast lead time. The forecast made at time period $t-\tau$ is denoted by $\hat{y}_{t-\tau}$. There is a lot of interest in the $lead-1$ forecast, which is the forecast of the observation in period t, y_t , made one period prior, $\hat{y}_{t-1}$. We will denote the fitted value of y_t by $\hat{y}_t$.

We will also be interested in analyzing forecast errors. The forecast error that results from a forecast of y_t that was made at time period $t - \tau$ is the $lead - \tau$ forecast error

$$e_t(\tau) = y_t - \hat{y}_t(t - \tau) \tag{1.2}$$

For example, the lead - 1 forecast error is

$$e_t(1) = y_t - \hat{y}_t(t - 1) \tag{1.3}$$

The difference between the observation y_t and the value obtained by fitting a time series model to the data, or a fitted value $\hat{y}_t$ defined earlier, is called a residual, and is denoted by

$$e_t = y_t - \hat{y}_t \tag{1.4}$$

The reason for this careful distinction between forecast errors and residuals is that models usually fit historical data better than they forecast. That is, the residuals from a model-fitting process will almost always be smaller than the forecast errors that are experienced when that model is used to forecast future observations.

1.5 General Strategy to time series modeling and forecasting

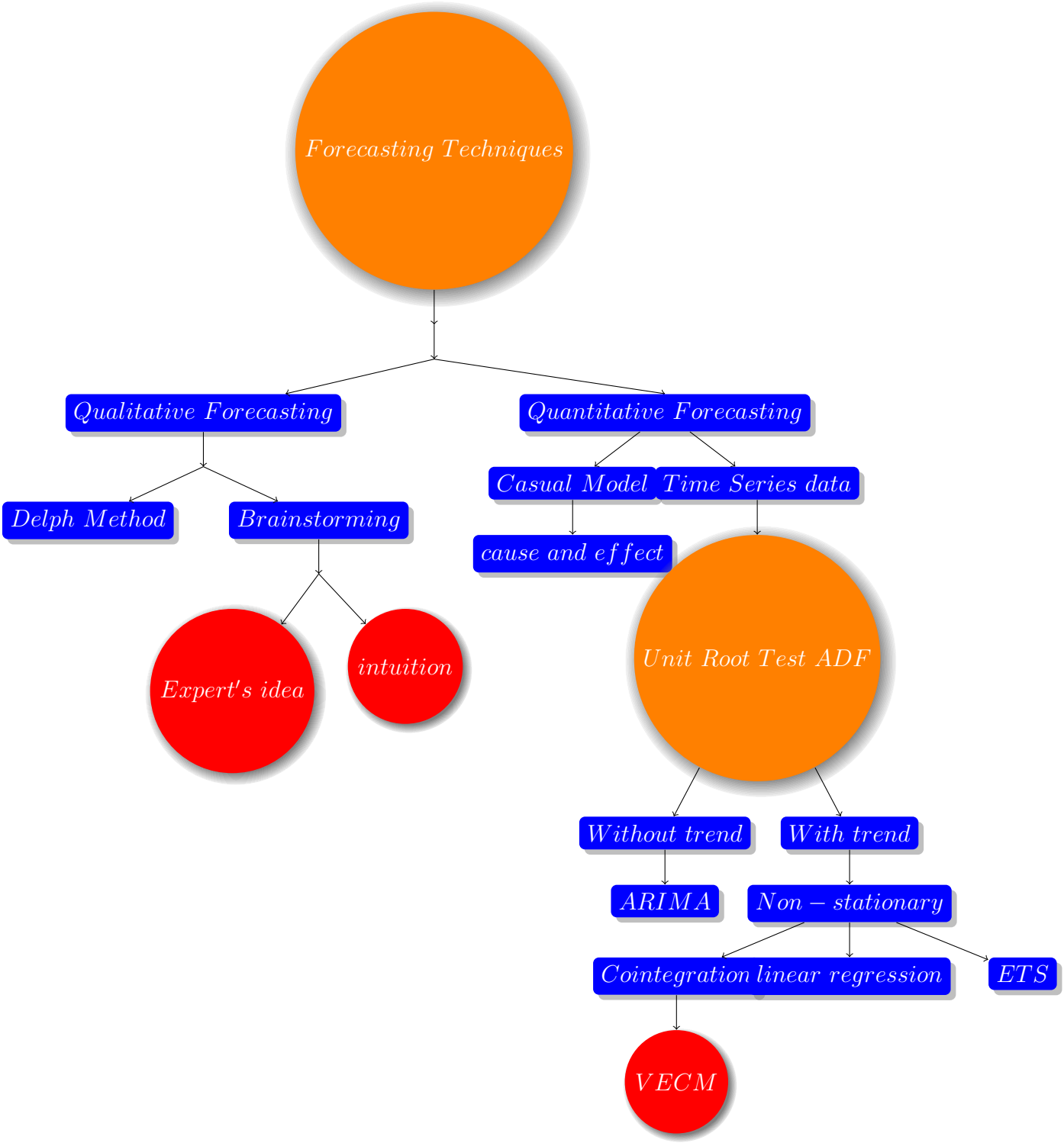
The forecasting techniques can be broadly categorized as consisting of qualitative and quantitative methods.

Qualitative forecasting techniques are mainly exploratory research. Qualitative forecasts are often used in providing the insights into the problems. However, although some data analysis may be performed, the expectations are based on the mathematical studies within the field of biological mathematics, physical and chemical mathematics and others. Perhaps the Delphi Method is the most formal and widely known qualitative forecasting method. This technique was developed by the RAND Corporation. It employs a panel of experts

who are assumed to be knowledgeable about the problem. The panel members are physically separated to avoid their deliberations being impacted either by social pressures or by a single dominant individual.

Quantitative forecasting techniques is used to quantify the problem by way of generating numerical data or data that can be transformed into usable statistics. Quantitative methods are much more structured and reliable results than Qualitative methods. There are several types of forecasting models in general use. In subsequent chapters, we will discuss all several different types of quantitative forecasting models, as Figure 1 shown.

However, although some data analysis may be performed, the expectations are based on the mathematical studies within the field of biological mathematics, physical and chemical mathematics and others. Perhaps the most formal and widely known qualitative forecasting technique is the Delphi Method. This technique was developed by the RAND Corporation [34]. It employs a panel of experts who are assumed to be knowledgeable about the problem. The panel members are physically separated to avoid their deliberations being impacted either by social pressures or by a single dominant individual. Each panel member responds to a questionnaire containing a series of questions and returns the information to a coordinator. Following the first questionnaire, subsequent questions are submitted to the panelists along with information about the opinions of the panel as a group. This allows panelists to review their predictions relative to the opinions of the entire group. After several rounds, it is hoped that the opinions of the panelists converge to a consensus, although achieving a consensus is not required and justified differences of opinion can be included in the outcome.


Forecasting Techniques
Qualitative Forecasting
Quantitative Forecasting
Delph Method
Brainstorming
Casual Model
Time Series data
cause and effect
Expert's idea
intuition
Unit Root Test ADF
Without trend
With trend
ARIMA
Non − stationary
Cointegration linear regression
ETS
VECM

In summary, Often in forecasting, a key step is knowing when something can be forecast accurately, and when forecasts will be no better than tossing a coin. Good forecasts capture the genuine patterns and relationships which exist in the historical data, but do not replicate past events that will not occur again. In this section, our topic is how to tell the difference between a random fluctuation in the past data that should be ignored, and a genuine pattern that should be modeled and extrapolated.

The objective of this dissertation is to present a comprehensive discussion about several popular approaches for time series forecasting, such as the stochastic, neural networks and SVM approaches. This dissertation contains four chapters, which are organized as follows: Chapter 1 gives an introduction to the basic concepts of time series modeling, together with some associated ideas such as stationary, parsimony, etc. Chapter 2 is designed to discuss about the various stochastic time series models. These include the Box-Jenkins or ARIMA models, the generalized ARIMA models as well as some non-linear models such as ARCH, GARCH, etc. In Chapter 3 we have described the application of Bayesian VAR model in time series forecasting, together with two recently developed models, artificial neural networks and SVM. Also, it presents a discussion about the SVR concepts and its usefulness in time series forecasting problems. In Chapter 4, we have introduced about ten important forecast performance measures, often used in literature, together with their salient features. Chapter 4 presents our experimental forecasting results in terms of five performance measures, obtained on six real time series datasets, together with the associated forecast diagrams. After completion of these four chapters, we have given a brief conclusion of our work as well as the prospective future aim in this field.

Chapter 2

2 Empirical Findings

2.1 Simple Model

This chapter aims to provide important information of real estate market in USA and potential problems and opportunities or buyer and seller. As housing is a form of wealth, the purchase of a home represents an important investment, and it is normally a hot topic for the scholars and investors. Because of scarcity, the fluctuation of land market value will have a great influence of the net worth of business and household. In this regard, Davis and Heathcote [30] estimate that wings in residential land prices accounted for most of the variation in house prices over 1975-2006 for the United States as a whole. Davis and Palumbo [31] reach the same conclusion for a large set of metropolitan areas over a somewhat shorter sample period, as do Botic, Longhofer and Redfeam [29] in their detailed analysis of home price changes within a single metropolitan area (Wichita, Kansas).

In addition, the land is an important component of wealth. Also, it is a source of variation in real estate prices, and as collateral for loans, only a handful of studies have calculated land price indexes for the nation as a whole or for a broad set of cities. Davis and Heathcote [30] and Davis and Palumbo [31] estimate price indexes for residential land, while Davis (2009) estimates indexes for both residential and commercial land. Also, Sirmans and Slade [37] use transaction prices to calculate national land prices indexes. The data were collected on the basis of past and present real estate transactions, and develop processes which guide future investment by demonstrating the true future value of the investment. To provide students with sufficient understanding and ability to model, analyze and develop forecasts for engineering and business decisions. The emphasis will be on quantitative methods.

After food and medical care, housing is the largest consumer expenditure in the United states. In 1994, personal consumption expenditures on housing were about $2600 per capita, or 14.9 percent of household budgets. Further, the bulk of expenditures in one of the next

highest categories, household operations,are linked to housing. From the investment side, housing is the largest single form of fixed capital investment in the United States, comprising more than $9 trillion, or roughly half of this nations gross fixed private capital. Other than human capital, housing and land are more widely held than any other form of capital. In United States, as in most countries, the market for housing services can be approximated as a competitive market. Housing production activities have few barriers to entry or large economies of scale. Few landlords or developers are large enough to exert significant market power[27].

The Table (2.1) showed the mechanism of housing market.

Table 2.1

The mechanism of housing market

Inputs		**Production**		**Demand**
Land	**P**	Developers	**P**	
Finance	**R**	Builders	**R**	Renters
infrastructure $\Longrightarrow$	**I**	$\Longrightarrow$ Landlords $\Longrightarrow$	**I**	$\Longrightarrow$ Homeowners
labor	**C**	Homeowners	**C**	(Income and Population)
Materials	**E**		**E**	

Housing is the largest asset of most American households, so the housing market profoundly affects the distribution of wealth; housing's location and tenure could well affect the behavior of its occupants. Hence, we work on how housing market works? In a good economic situation when house prices are consistently rising, most consumers can afford what is perceived to be full market value for a given property, because the inherent assumption is that value will continue to rise.

2.1.1 Regression Analysis and Dynamic Regression Models

For regression method, as shown in Figure (2.1), it mainly measures how the land market value, in terms of GDP, CPI, construction cost index and unemployment rate, inflation rate, population and purchasing manage index, enables a relationship with the real estate market

and investment.

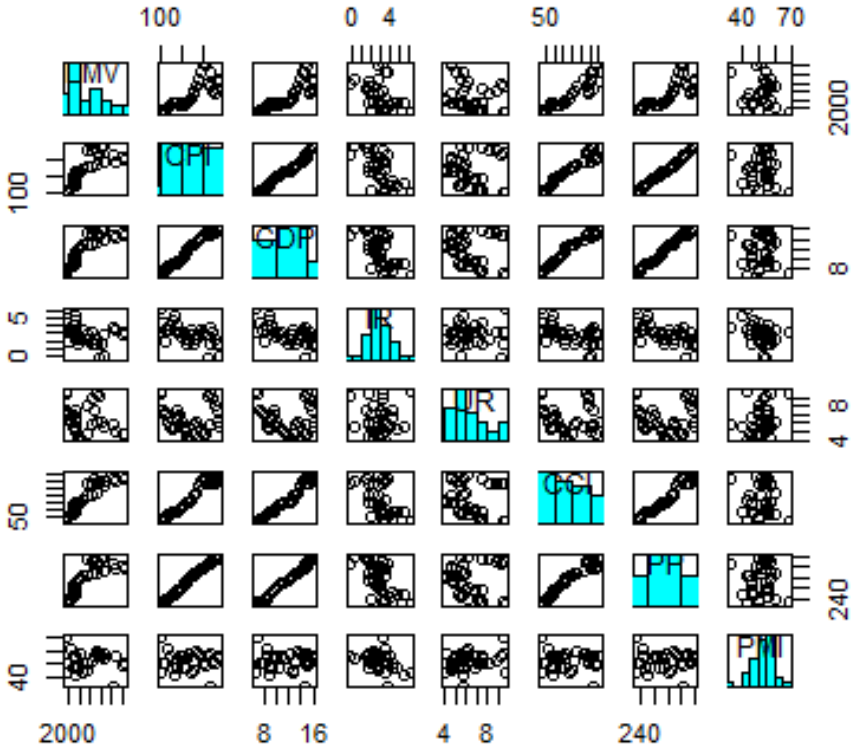


Figure 2.1: LMV

Besides, we explore the quantitative and qualitative relationship among these economic variables at risk scenarios. Firstly, we developed two regression models with raw data and log transformation. After checking all the significance of all explanatory variables and residuals of autocorrelation, we obtain Eq (2.1) and Eq (2.2). From the Eq (2.1), we may simply conclude that the land market value is highly related to IR, UR, CCI and PMI. When UR increased 1 unit and other variables keep unchanged, the land market value will decrease 218.85 million. For Eq (2.2), it involves more explanatory variables than Eq (2.1).

$$LMV = -8400.287 + 574.933IR - 218.85UR + 154.035CCI + 30.048PMI \qquad (2.1)$$

Table 2.2

Regression Table for Eq (2.1)

Model	**Coefficients**	**Standard Error**	**T-value**	**Significance**	**VIF**	**Tolerance**
Constants	-8400.287	2857.725	-2.94	0.006		
IR	574.933	219.673	2.617	0.014	1.89	0.529
UR	-218.850	136.198	-1.607	0.119	1.098	0.911
CCI	154.035	12.589	12.236	0.00	1.635	0.612
PMI	30.048	37.207	0.808	0.426	1.378	0.726

Model	**R-squared**	**Adj R-squared**	**S.E of Est**	**Sample Size**	**F-change**	**Sig**
Eq(1)	0.935	0.875	1220.773	34	50.570	0.00

Considering the multicollinearity within the multiple regressions, we only dropped the variables which are highly related to other variables in the regression, such as GDP and PP in Eq (2.1). Because such variables GDP and PP are capturing the effect of other variables. However,unemployment usually indicate more economic distress and lower production which causes lower demand for economic purchases including land. Hence, it shows the negative relation to the Land Market Value. Because of the big sample size, we keep some insignificant variables in our model, the effects of those variables, PMI and UR, are negligible. Also, we get rid of the effects of multicollinearity for the selected variables in Eq (2.1).

$$\begin{aligned} LLMV = 82.146 - 18.029LPP + 4.2LCCI + 4.45LGDP \\ + 0.505LUR - 0.545LPMI - 0.148LIR \end{aligned} \tag{2.2}$$

The log transformation of the regression model can help stabilize the variance. Hence, we try to fit the log transformed model of Eq (2.2). The transformed model statistical summary shown below:

However, from the table (2.2) and table (2.3), it shows that there are some evidence of autocorrelation in the residuals of Eq (2.1) and Eq (2.2), it indicates there are some information from the data.

Model	Coefficients	Standard Error	T-value	Significance	VIF	Tolerance
Constants	-0.172	1.156	-0.149	0.883		
LIR	0.059	0.053	1.126	0.269	1.338	0.747
LUR	-0.281	0.132	-2.122	0.043	1.12	0.893
LCCI	2.169	0.12	18.088	0.00	1.299	0.770
LPMI	-0.085	0.245	-0.346	0.732	1.153	0.867

Model	R-square	Adj R-squared	S.E of Est	Sample Size	F-value	Sig
Eq(2.2)	0.937	0.928	0.17765	34	107.374	0.00

Hence, we extended the regression method into the general class of dynamic regression models, which simply combined regression models with ARIMA errors. We take into account several formula as the theoretical foundation,

$$y_t = \beta_0 + \beta_1 x_{1,t} + \cdots\cdots\cdots + \beta_k x_{k,t} + n_t \tag{2.3}$$

$$(1 - \phi_1 B - \phi_2 B^2) n_t = e_t \tag{2.4}$$

$$(1 - 1.5024B + 0.7229B^2) n_t = e_t \tag{2.5}$$

$$n_t = 262.8185 + 1.5024 n_{t-1} - 0.7229 n_{t-2} + e_t \tag{2.6}$$

Where n_t denotes the errors from the regression models and e_t denotes the errors from the ARIMA model. Only the ARIMA model errors are assumed to be white noise.

Box.test	X-squared	df	P-value
Box.Ljung	14.0415	20	$0.8284 > 0.05$

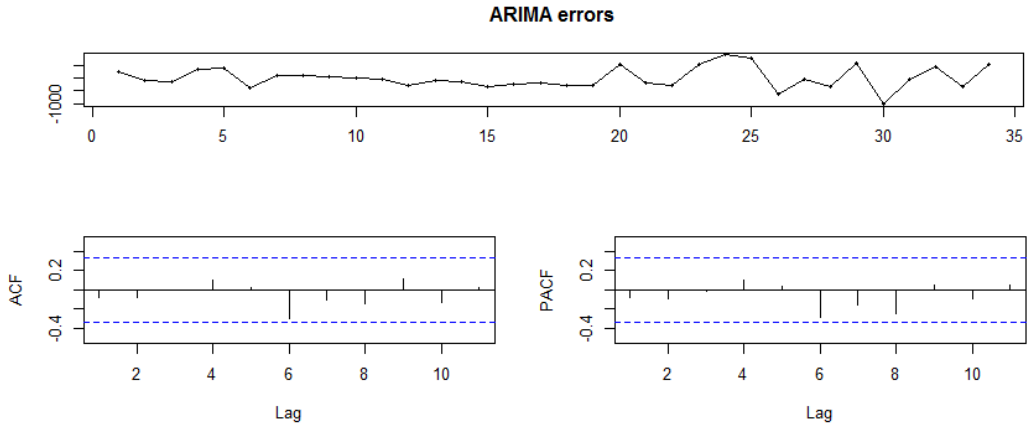


Figure 2.2: Dynamic Regression

From the plots and non-parametric tests of e_t in Figure (2.2), we may conclude that the residuals of $AR(2)$ from the regression model is stationary. Compared to three regression equations, the dynamic method is better than others.

2.1.2 Time Series Decomposition and Smoothing Analysis

For time series decomposition models, we have a couple of options, such as the classical additive decomposition, classical multiplicative decomposition and STL decomposition. The classical decomposition is basic and simple way to forecast the trend. We employ the simple exponential smoothing method, Holt's linear method, exponential smoothing method and additive damped method and multiplicative damped method. Eventually, the exponential smoothing model could have better forecasting on the trend in Figure (2.4). For ETS(M,A,N) model, We take this into innovation by considering multiplicative error equations.

$$y_t = (l_{t-1} + b_{t-1})(1 + \varepsilon_t) \tag{2.7}$$

$$l_t = (l_{t-1} + b_{t-1})(1 + \alpha\varepsilon_t) \tag{2.8}$$

$$b_t = b_{t-1} + \beta(l_{t-1} + b_{t-1})\varepsilon_t \tag{2.9}$$

$$\varepsilon_t = \frac{y_t - (l_{t-1} + b_{t-1})}{l_{t-1} + b_{t-1}} \tag{2.10}$$

Where the $\varepsilon \backsim NID(0, \sigma^2)$, l_t denotes an estimate of the level of the series at time t, b_t denotes an estimate of the trend (slope) of the series at time t, α denotes the smoothing parameter for level. β denotes the smoothing parameter for the trend [15].

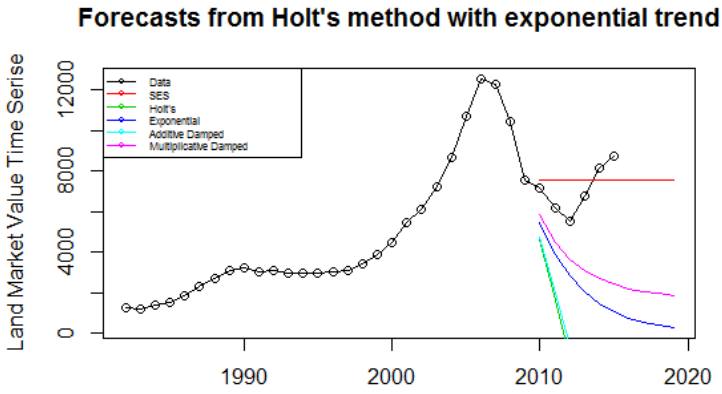


Figure 2.3: Holt's method

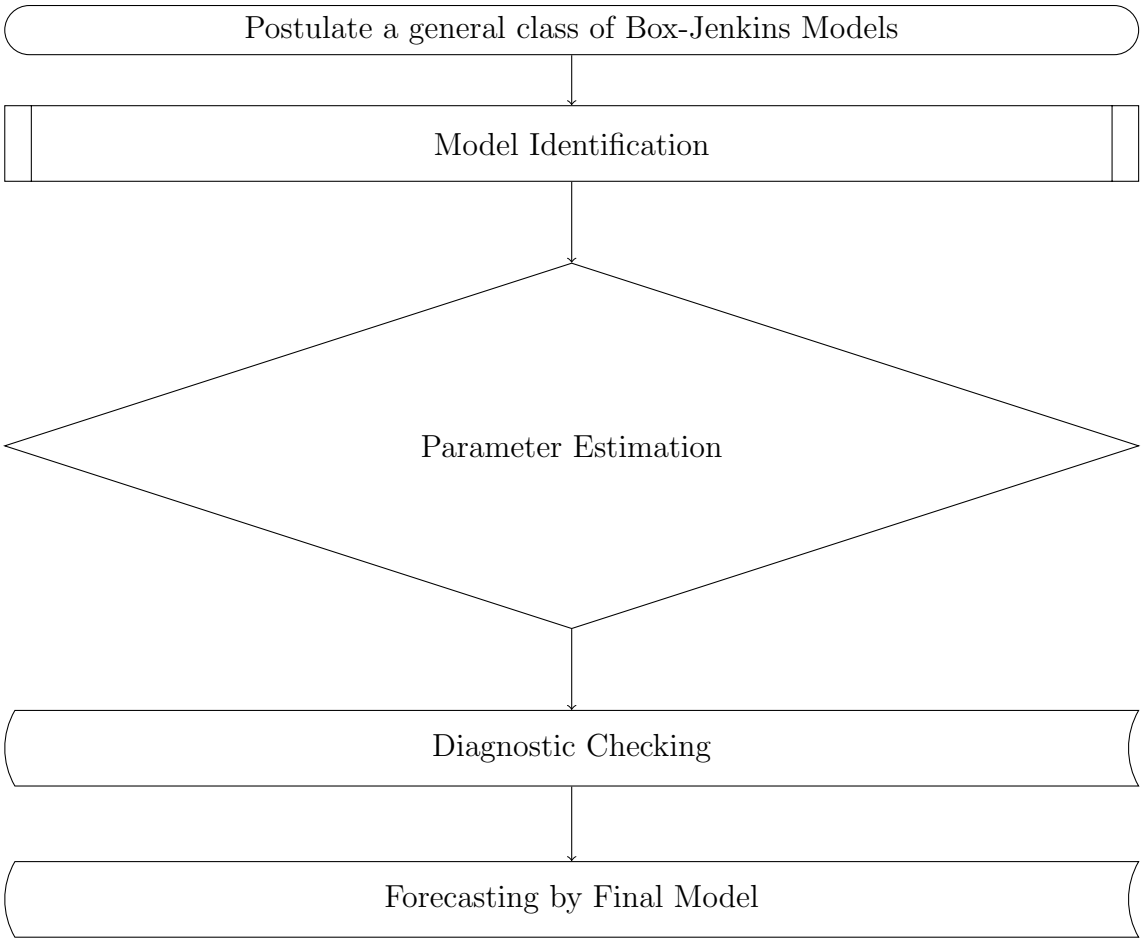
By the method of minimizing the "likelihood". We estimate the smoothing parameters α, β , b and l. In our model, the estimated parameters are $\alpha = 0.9051$, $\beta = 0.9051$, $l = 836.1576$, $b = 388.3622$. The possible values that the smoothing parameters can take is restricted. Traditionally the parameters have been constrained to lie between 0 and 1 so that the equations can be interpreted as weighted averages.For the state space models, we have set $0 < \alpha < \beta < 1$.

2.2 Time Series Analysis

2.2.1 Introduction of ARMA model

In this section, an important parametric family of stationary time series and non-stationary time series was introduce, the auto-regressive moving-average, or ARMA, processes. The family of ARMA processes plays a key role in the modeling of time series data. The linear structure of ARMA processes also leads to a substantial simplification of the general methods for linear prediction.

In the previous chapter, we discussed forecasting techniques that, in general, were based on some variant of exponential smoothing. The general assumption for these models was that any time series data can be represented as the sum of two distinct components: deterministic and stochastic (random). The former is modeled as a function of time whereas for the latter we assumed that some random noise that is added on to the deterministic signal generates the stochastic behavior of the time series. One very important assumption is that the random noise is generated through independent shocks to the process. In practice, however, this assumption is often violated. That is, usually successive observations show serial dependence. Under these circumstances, forecasting methods based on exponential smoothing may be inefficient and sometimes inappropriate. because they do not take advantage of the serial dependence in the observations in the most effective way. To formally incorporate this dependent structure, in this chapter we will explore a general class of models called autoregressive integrated moving average (MA) models or ARIMA models (also known as BoxJenkins models).


Postulate a general class of Box-Jenkins Models
Model Identification
Parameter Estimation
Diagnostic Checking
Forecasting by Final Model

Limited by finite number of available observations, we often construct a finite parametric model to describe a time series process. The $ARMA$ model contains a very broad class of parsimonious time series processes found useful in describing a wide variety of time series. After giving detailed discussions on the characteristics of each process in terms of the autocorrelation and partial autocorrelation functions, we will illustrate the results with examples [10].

Next, we will show the general $ARMA(p,q)$ process. Suppose X_t is an $ARMA(p.q)$ process if X_t is stationary and if for every t,

$$X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p} = Z_t + \eta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q} \tag{2.11}$$

where $Z_t \sim WN(0, \sigma_t)$ and the polynomials $(1 - \phi_1 z - \cdots - \phi_p z^p)$ and $(1 + \theta_1 z + \cdots + \theta_q z^q)$ have no common factors.

The process X_t is said to be an $ARMA(p,q)$ process with mean μ if $X_{x-\mu}$ is an $ARMA(p,q)$ process. It is very convenient to use concise form of (2.11):

$$\phi(B)X_p = \theta(B)Z_q \tag{2.12}$$

Where $p(\cdot)$ and $q(\cdot)$ are the $p-th$ and $q-th$ degree polynomial.

$$\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p \tag{2.13}$$

And

$$\theta(z) = 1 - \theta_1 z - \cdots - \theta_q z^q \tag{2.14}$$

And B is backward shift operator.($B^j X_t = X_{t-j}, B^j Z_t = Z_{t-j}, j = 0, +1, \cdots$). The time series X_t is said to be an autoregressive process of order p or ($AR(p)$) if $\theta Z \equiv 1$, and a

moving-average process of order q or $(MA(q))$ if $\phi Z \equiv 1$.

2.2.2 ARIMA model

The time series processes we have discussed so far are all stationary processes, but many applied time series, particularly those arising from economic and business areas, are non-stationary. With respect to the class of covariance stationary processes, non-stationary time series can occur in many different ways. They could have non-constant means μ_t, time-varying second moments such as non-constant variance σ_t^2, or both of these properties. In this chapter, we will explain the construction of a very useful class of homogeneous non-stationary time series models, the autoregressive integrated moving average models. Some useful differencing and variance stabilizing transformations are introduced to connect the stationary and non-stationary time series models [11].

Many models used in practice are of the simple ARIMA type, which have a long history and were formalized in Box and Jenkins [4]. ARIMA stands for Autoregressive Integrated Moving Average and an $ARIMA(p; d; q)$ model for an observed series $\{y_t\}$, $t = 1 \cdots T$ is a model where the dth difference $z_t = y_t - y_{t-d}$ is taken to induce stationarity of the series. The process $\{z_t\}$ is then modeled as $z_t = \mu + \epsilon_t$ with

$$\epsilon_t = \phi_1 \epsilon_{t-1} + \phi_2 \epsilon_{t-2} + \cdots + \phi_p \epsilon_{t-p} + u_t - \eta_1 u_{t-1} - \cdots - \eta_q u_{t-q} \tag{2.15}$$

or in terms of polynomials in the lag operator L (defined through $L^s x_t = x_{t-s}$):

$$\phi(L)\epsilon_t = \eta(L)u_t \tag{2.16}$$

where u_t is white noise and usually Normally distributed as $u_t \sim N(0; \sigma^2)$. The stationarity and invertibility conditions are simply that the roots of $\phi(L)$ and $\eta(L)$, respectively, are outside the unit circle. An accessible and extensive treatment of the use of Bayesian methods for ARIMA models can be found in Bauwens et al. (1999). The latter chapters also has a

useful discussion of multivariate modelling using Vector Autoregressive (VAR) models and co-integration.

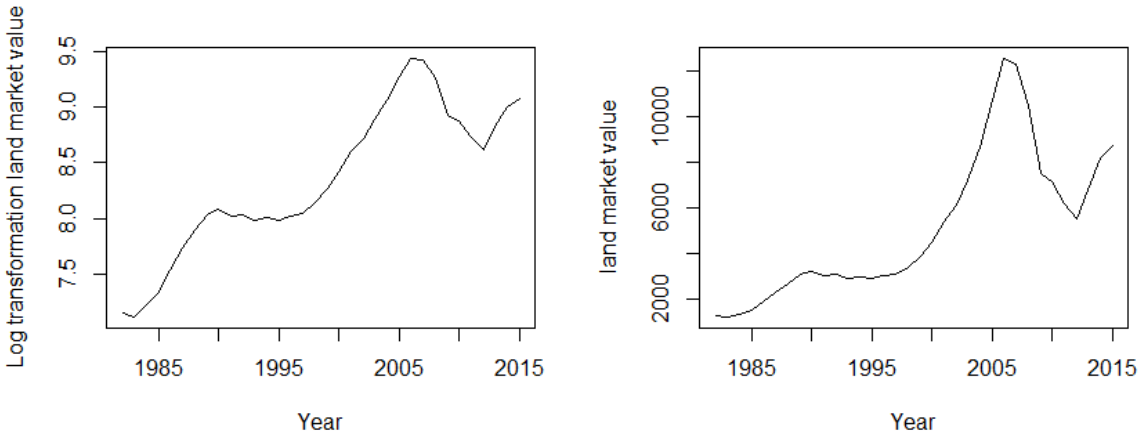


Figure 2.4: tslmv

Obviously, in Figure (2.5), Land Market Value is an increasing time series dataset. The original data and log transformation data were fitted to the ARIMA model. Eventually, the Box-Ljung test of residual met the assumption of the non-parametric, it indicates that autocorrelation come from the white noise. Hence, we selected the Log ARIMA (2,0,1) for forecasting model. [Eq.(2.32) on Page 36]

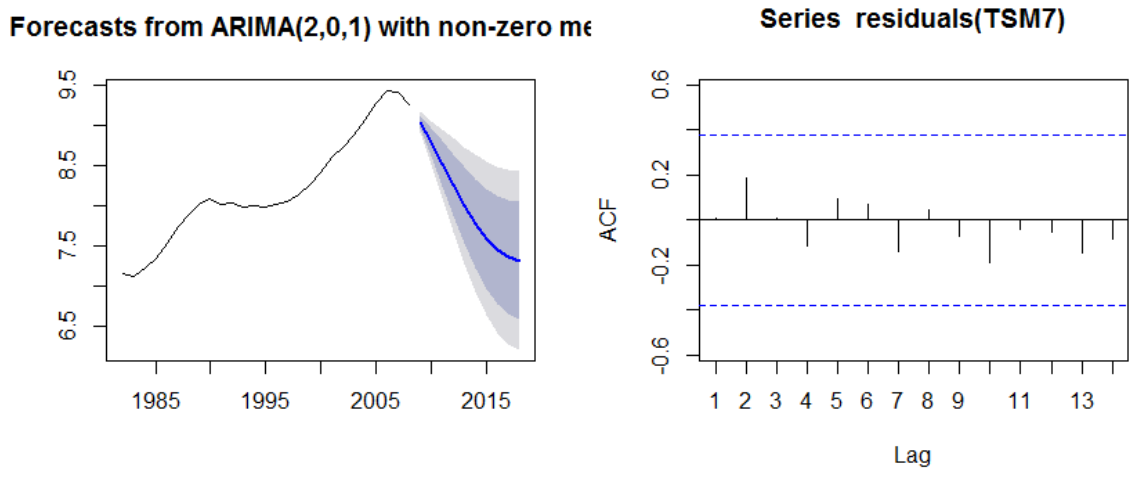


Figure 2.5: TSLMV

In the mature economies LMV illustrate the importance of land as a source of wealth, but in rapidly growing economies land has an even more significant role in determining economic welfare and a host of incentives for the performance of the economy. From the time series graph, we may find that the American land market value shows stable increase from 1982 to 2004, but from 2005 this number increased dramatically and peaked in 2006, 12.55. In fact, the economic crisis started in 2006 in USA, the economics crisis led to the increased interest, hence, the LMV rose rapidly. However, LMV decreased from 12.55 to 5.54 which is the relative lower level in 2012. The economy of USA experienced the great recession during this period. Until 2013, the situation recovered and this number rose to 7.594 in 2015.

2.2.3 Autoregressive Fractionally Integrated Moving Average

The time series processes we have discussed so far are all stationary processes, but many applied time series, particularly those arising from economic and business areas, are non-stationary. With respect to the class of covariance stationary processes, non-stationary time series can occur in many different ways. They could have non-constant means μ_t, time-varying second moments such as non-constant variance σ_t^2, or both of these properties. In this chapter, we will explain the construction of a very useful class of homogeneous non-stationary time series models, the autoregressive integrated moving average models. Some useful differencing and variance stabilizing transformations are introduced to connect the stationary and non-stationary time series models [37].

Many models used in practice are of the simple ARIMA type, which have a long history and were formalized in Box and Jenkins [4]. ARIMA stands for Autoregressive Integrated Moving Average and an $ARIMA(p; d; q)$ model for an observed series $\{y_t\}$, $t = 1 \cdots T$is a model where the dth difference $z_t = y_t - y_{t-d}$is taken to induce stationarity of the series.

The process $\{z_t\}$ is then modeled as $z_t = \mu + \epsilon_t$ with

$$\epsilon_t = \phi_1\epsilon_{t-1} + \phi_2\epsilon_{t-2} + \cdots + \phi_p\epsilon_{t-p} + u_t - \eta_1 u_{t-1} - \cdots - \eta_q u_{t-q} \tag{2.17}$$

or in terms of polynomials in the lag operator L (defined through $L^s x_t = x_{t-s}$):

$$\phi(L)\epsilon_t = \eta(L)u_t \tag{2.18}$$

Where u_t is white noise and usually Normally distributed as $u_t \sim N(0; \sigma^2)$. The stationarity and invertibility conditions are simply that the roots of $\phi(L)$ and $\eta(L)$, respectively, are outside the unit circle. An accessible and extensive treatment of the use of Bayesian methods for ARIMA models can be found in Bauwens et al. (1999). The latter chapters also has a useful discussion of multivariate modelling using Vector Autoregressive (VAR) models and co-integration.

ARIMA models will either display perfect memory (if there are any unit roots) or quite short memory with geometrically decaying autocorrelations (in the case of a stationary ARMA model). ARFIMA models [22] have more flexible memory properties, due to fractional integration which allows for hyperbolic decay. Consider $z_t = \Delta y_t - \mu$, which is modeled by an $ARFIMA(p, \delta, q)$ model as:

$$\phi(L)(1-L)^\delta z_t = \eta(L)u_t \tag{2.19}$$

Where u_t is white noise with $u_t \sim N(0, \sigma^2)$, and $\delta \in (0, -1, 0.5)$,the fractional differencing operator $(1-L)^\delta$ is defined as:

$$(1-L)^\delta = \sum_{j=0}^{\infty} c_j(\delta)L^j. \tag{2.20}$$

where $c_0(\cdot) = 1$ and for $j > 0$:

$$c_j(a) = \prod_{k=1}^{j} \left(1 - \frac{1+a}{k}\right) \tag{2.21}$$

This model takes the entire past of z t into account, and has as a special case the $ARIMA(p, 1, q)$ for y_t (for $\delta = 0$). If $\delta > -1$, z_t is invertible (Odaki, 1993) and for $\delta < 0.5$, we have stationarity of z_t. Thus, we have three regimes:

- $\delta \in (-1, -0.5)$: y_t trend-stationary with long memory.
- $\delta \in (-0.5, 0)$ z_t stationary with intermediate memory.
- $\delta \in (0, 0.5)$ z_t stationary with long memory.

Of particular interest is the Impulse Response Function $I(n)$, which captures the effect of a shock of size one at time t on y_{t+n}, and is given by

$$I(n) = \sum_{i=0}^{n} c_i(-\delta - 1)J(n - i) \tag{2.22}$$

With $J(i)$ the standard $ARMA(p, q)$ impulse responses (i.e. the coefficients of $\phi^{-1}(L)\eta(L)$). Thus, $I(\infty)$ is 0 for $\delta < 0$, $\eta(1)/\phi(1)$ for $\delta = 0$ and ∞ for $\delta > 0$. Koop et al.[1997] analyses the behavior of the impulse response function for real U.S. GNP data using a set of 32 possible models containing both ARMA and ARFIMA models for z_t. They use Bayesian model averaging to conduct predictive inference and inference on the impulse responses, finding about one third of the posterior model probability concentrated on the ARFIMA models [39]. Koop et al. (1997) use importance sampling to conduct inference on the parameters, while MCMC methods are used in Pai and Ravishanker (1996) and Hsu and Breidt (2003).

2.3 VAR models

Forecasts of the future tendency of economic variables such as GDP, inflation rate and unemployment rate, arise many interests from business and government. Also, Modeling the land market at the national level can capture rich dynamic presenting in interdependent economies. In this section, we studied a VAR of Land Market Value and five US macroeconomic variables. We employed the VAR model for forecasting Land Market Value in USA and analyzed annual data on the main macroeconomic variables of interest going back to 1982. Most importantly, we explore the mutual influence between Land market value and selected macroeconomics variables to enable government and investor to make informed decision regarding real estate market [2].

There are a variety of methods available for forecasting economic variables. One common type of forecast is Vector auto-regression modeling for multivariate Time Series approach. This type of forecast is predominantly in economics and financial analysis.

A VAR model is an useful and flexible approach to describe the dynamic behavior of economic activity and financial time series dataset; that is, a vector of time series. In this system, we consider one equation for one variable as dependent variable with constant and lags. Each variable is assumed to influence with each other in the system, which makes direct interpretation of the estimated coefficients very difficult [15]. We write a multi-dimensional $VAR(p)$ as

$$Y_t = C + \Phi_1 \begin{bmatrix} LLMV_{t-1} \\ LCPI_{t-1} \\ LUR_{t-1} \\ LPP_{t-1} \\ LCCI_{t-1} \\ LPMI_{t-1} \end{bmatrix} + \Phi_2 \begin{bmatrix} LLMV_{t-2} \\ LCPI_{t-2} \\ LUR_{t-2} \\ LPP_{t-2} \\ LCCI_{t-2} \\ LPMI_{t-2} \end{bmatrix} + \cdots + \Phi_{t-p} \begin{bmatrix} LLMV_{t-p} \\ LCPI_{t-p} \\ LUR_{t-p} \\ LPP_{t-p} \\ LCCI_{t-p} \\ LPMI_{t-p} \end{bmatrix} + a_t \qquad (2.23)$$

where a_t are white noise process. $E(a_t) = 0$ and

$$E(a_t a_t') = \begin{cases} 0 \quad when\ t = \tau, \\ \Omega \quad when\ t \neq \tau, . \end{cases}$$

In the reduced form, we will include a six variable VAR with one lag in our forecasting model.

$$Y_t = \begin{bmatrix} LLMV_t \\ LCPI_t \\ LUI_t \\ LCCI_t \\ LPP_t \\ LPMI_t \end{bmatrix} \quad \Phi_1 = \begin{bmatrix} \phi_{11} & \phi_{12} & \phi_{13} & \phi_{14} & \phi_{15} & \phi_{16} \\ \phi_{21} & \phi_{22} & \phi_{23} & \phi_{24} & \phi_{25} & \phi_{26} \\ \phi_{31} & \phi_{32} & \phi_{34} & \phi_{34} & \phi_{35} & \phi_{36} \\ \phi_{41} & \phi_{42} & \phi_{43} & \phi_{44} & \phi_{45} & \phi_{46} \\ \phi_{51} & \phi_{52} & \phi_{53} & \phi_{54} & \phi_{55} & \phi_{56} \\ \phi_{61} & \phi_{62} & \phi_{63} & \phi_{64} & \phi_{65} & \phi_{66} \end{bmatrix}$$

Coefficient $\phi_{ii,i}$ indicates the influence of the iith lag of variable Y_i on itself, while coefficient $\phi_{ij,i}$ indicates the influence of the iith lag of variable Y_j on Y_i.

A VAR in levels is known as the series modeled are stationary, we forecast them directly by fitting a VAR to the data . A VAR in differences is known as the series are non-stationary, we firstly take differences to make them stationary, and then we fit a VAR model. In both cases, the models and coefficients are estimated equation by equation using the principle of least squares.

We applied the VAR selection package for forecasting the raw data.The function returns information criteria and final prediction error for sequential increasing the lag order up to a VAR(p)-process. which are based on the same sample size. For each equation, the parameters are estimated by minimizing the sum of squared $e_{i,t}$ values.

Before running R software, the data is firstly adjusted by taking log transformation to stabilize the variance. And then, we set the 80% of the data as training set and the remaining data as the test set [18]. The statistical summary of transformed model shown below:

Table 2.3

VAR ESTIMATION RESULTS FOR VAR (1)

Endogenous variables: LLMV, LUR, LPP,LCPI,LPMI, LCCI
Deterministic variables: const
Sample size: 28
Log Likelihood: 472.079
Roots of the characteristic polynomial:
0.9685 0.9403 0.9403 0.807 0.5849 0.5849

Residuals of VAR object var	**Chi-squared**	**DF**	**P-value**
Portmanteau Test (asymptotic)	333.73	324	0.3429

Next, we can check the stability conditions for VAR Systems by the value of Roots of the characteristic polynomial. Because they are all less than 1 in Table (2.3), we could conclude that the VAR (1) model is stationary.

Firstly, if a VAR (2) model is estimated. The null hypothesis of no autocorrelation is rejected since the p-value of 0.03325 is lower than the significance level of 0.05. Since autocorrelation is an undesirable feature of the model, we moves on to look for another model that does not have autocorrelation. As shown in Table (2.3), we estimates a VAR (1) model, tests for autocorrelation, and finds that the null of no autocorrelation cannot be rejected because the p-value of 0.3429 is greater than the significance level of 0.05. Since there is not enough evidence of presence of autocorrelation, we satisfied and sticks to the VAR (1) model.

A portmanteau test is used for autocorrelation in errors:

H_0: There is no evidence show that there are autocorrelation in residuals for some lag p.

H_1: There are some evidences show that there are autocorrelation in residuals for some lag p.

$$LLMV = 1.198LLMV_{t-1} - 0.341LUR_{t-1} + LCCI_{t-1} + 7.556LPP_{t-1} - 1.496LPMI_{t-1} - 1.886LCPI_{t-1} - 29.122 \quad (2.24)$$

$$LCCI = 1.198LLMV_{t-1} - 0.341LUR_{t-1} + LCCI_{t-1} + 7.556LPP_{t-1} - 1.496LPMI_{t-1} - 1.886LCPI_{t-1} - 29.122 \quad (2.25)$$

When VAR is used for forecasting tool, one has to choose the number of variables (denoted by K) and the number of lags (denoted by p). The number of coefficients to be estimated in a VAR is equal to $K + pK^2$ (or $1 + pK$ per equation). For example, in our VAR model with $K = 6$ variables and $p = 1$ lags, there are 7 coefficients per equation making for a total of 42 coefficients to be estimated. The more coefficients to be estimated the larger the estimation error entering the forecast.

Table 2.4

VAR forecasting model for LLMV

Model	**Estimate**	**Standard Error**	**T-value**	**Significance**
LLMV.l1	1.1981	0.1402	8.548	4.13e-08 ***
LCPI.l1	-1.8860	0.6243	-3.021	0.006750 **
LPP.l1	7.5556	2.2532	3.353	0.003164 **
LPMI.l1	0.007910	0.001851	4.273	0.000514 ***
LCCI.l1	0.031452	0.007257	4.334	0.000450 ***
LUR.l1	-0.3411	0.1045	-3.265	0.003874 **
constant	-29.1223	8.7337	-3.334	0.003304 **

Table 2.5

VAR forecasting model for LCCI

Model	**Estimate**	**Standard Error**	**T-value**	**Significance**
LLMV.l1	0.077	0.034	2.255	0.035 *
LCPI.l1	-0.15	0.152	-0.989	0.335
LPP.l1	1.418	0.548	2.586	0.018 *
LPMI.l1	0.123	0.027	4.475	0.000232 ***
LCCI.l1	0.480	0.177	2.711	0.0135*
LUR.l1	-0.0047	0.025	-0.183	0.857
constant	-6.053	2.126	-2.848	0.00995 **

Mul R-squared	**Adj R-squared**	**Residual.s.e**	**S. Size**	**F-stat**	**P-value**
0.9969	0.9951	0.04333 on 17 df	29	544.6 on 10	$<$ 2.2e-16

In Table (2.4) and Table (2.5), we list the forecasting model for LLMV and LCCI and ignore the other variables, because we select the variables we interests regarding real estate market. In table 2, we noticed that the p value are extremely small. It indicates that $LLMV$ can be estimated by the $LLMV_{t-1}$, $LCPI_{t-1}$, LPP_{t-1}, $LPMI_{t-1}$, $LCCI_{t-1}$ and LUR_{t-1}. The forecasting model fits very well for LLMV. Our goal is to forecast the tendency of Land market value to provide the information for policy-maker or decision-maker. The construction cost index is highly related to real estate market. In table (2.5), the forecasting performance for LCCI is influenced by the $LLMV_{t-1}$,LPP_{t-1}, $LPMI_{t-1}$ and $LCCI_{t-1}$. So we mainly studied the performance of the two variables in our forecasting model. Also, the performance of forecasting of other variables are not good enough, so we will not discuss them in this section.

2.3.1 Forecasting

VAR model generate the forecasting in a recursive structure. The VAR is a system in which each variable is regressed on a constant and p of its own lags as well as on p lags of each of the other variables in the VAR. To illustrate the process, assume that we have fitted the multi-dimensional VAR (1) described in equations Eq (2.17) for all observations up to

time T. Then the one-step-ahead forecasts are generated by

$$\hat{y}_{1,T+1|T} = \hat{c}_1 + \hat{\phi}_{11,1} y_{1,T} + \hat{\phi}_{12,1} y_{1,T} \tag{2.26}$$

$$\hat{y}_{2,T+1|T} = \hat{c}_1 + \hat{\phi}_{21,1} y_{1,T} + \hat{\phi}_{22,1} y_{2,T} \tag{2.27}$$

$$\hat{y}_{3,T+1|T} = \hat{c}_1 + \hat{\phi}_{31,1} y_{1,T} + \hat{\phi}_{32,1} y_{3,T} \tag{2.28}$$

$$\hat{y}_{4,T+1|T} = \hat{c}_1 + \hat{\phi}_{41,1} y_{1,T} + \hat{\phi}_{42,1} y_{4,T} \tag{2.29}$$

$$\hat{y}_{5,T+1|T} = \hat{c}_1 + \hat{\phi}_{51,1} y_{1,T} + \hat{\phi}_{52,1} y_{5,T} \tag{2.30}$$

$$\hat{y}_{6,T+1|T} = \hat{c}_1 + \hat{\phi}_{61,1} y_{1,T} + \hat{\phi}_{62,1} y_{6,T} \tag{2.31}$$

This is the same form as Eq (2.26) to Eq (2.31) except that the errors have been set to zero and parameters have been replaced with their estimates.

2.3.2 Conclusion

The VAR forecasting model has been widely used in many area of finance in recent years and it increased the understanding of tendency of land market value. Compared to separate univariate models, the VAR models have the advantage over traditional large-scale macroeconomics variables, but are easily interpreted and available. However, VAR models have also been much criticized, but the criticism usually refers to particular applications and interpretations of empirical results, rather than the methodology itself. For example, if a VAR model deal with a risk as the longer the lags, it will estimate the greater the number of parameters and the fewer the degrees of freedom.

As noted, Land is nature's gift to mankind, which enables life to continue and prosper. Because of its uniqueness of fixed supply and immobility, housing and land are more important for the economy that at any point in recent memory, the better housing forecasts results will be useful and necessary for USA real estate market. The tremendous rise in house prices over the decade has been both a national and global phenomenon.

In other hand, how to measure the forecasting accuracy is also our interest and con-

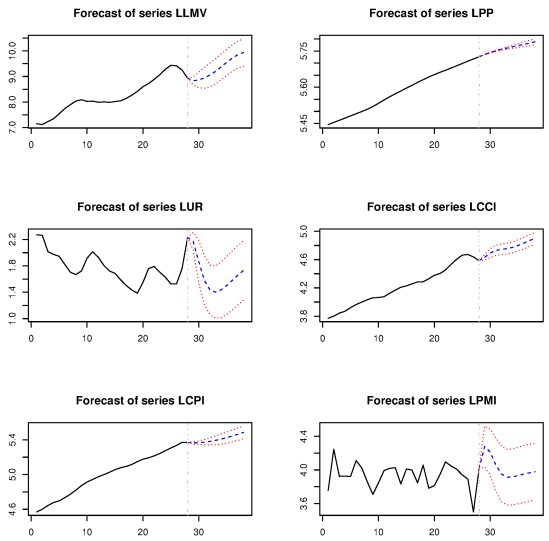


Figure 2.6: VAR forecasting

clusion from this paper. Measuring the forecasting accuracy is an efficient way to select a better model for the vector time series datasets. Forecasting evaluation is relevant to the decision-maker when choosing on a model specification for subsequent use. The preference or loss function of forecast evaluations the accuracy measures are some forms of average error, typically root mean squared error or mean absolute error, but many other possibilities are available.

The VAR package contains the function VARselect for selection of lags p by four different information criteria: AIC, HQ, SC and FPE. We have met the AIC before, AIC also created by Akaike. SC stands for Schwarz Criterion after Gideon Schwarz who proposed it. It is also simply another name for BIC. HQ is the Hannan-Quinn criterion and FPE is the Final Prediction Error criterion. The criterion of AIC is usually used to choose large numbers of

lags. Instead, for VAR models, we prefer to use the AIC and BIC [22].

Based on the same sample size and the scope of different information criterion, and the final prediction error are computed:

$$AIC(p) = ln(|S(p)|) + \frac{2pm^2}{n} \tag{2.32}$$

$$AICc = AIC + \frac{2p*(p*+1)}{n-p*-1} \tag{2.33}$$

$$HQ(p) = ln(|S(p)|) + \frac{2ln(ln(n))pm}{n} \tag{2.34}$$

$$SC(p) = ln(|S(p)|) + \frac{2ln(n)pm}{n} \tag{2.35}$$

$$FPE(p) = \left(\frac{n+p*}{n-p*}\right)^m |S(p)|) \tag{2.36}$$

Where $|S(p)|$ is the residual sum of squares and cross-products. P^* is the total number of parameters in each equation.p assigns the lag order.

```
VARselect(mydatats, lag.max=8, type="const")$selection
AIC(n) HQ(n) SC(n) FPE(n)
  4      4     4      3
```

As you can see, the four criterions are so similar. According to the AIC,HQ,SC the optimal lag number is $p = 4$, whereas the FPE criterion indicates $p = 3$. We estimated for one lag of VAR including a constant and a trend as deterministic regressors and conducted diagnostic tests with respect to the residuals.

2.4 Time-varying volatility models

2.4.1 The heteroscedasticity and homoscedasticity problem

Many economic and financial time series shows the serious heteroscedastic problems. The conditional heteroscedasticity usually depends on the past value. In this case, ARCH and GARCH models are developed to handle changes in conditional variances. Also, the ARCH and GARCH model, some extensions and various applications of the forecasting model, are

widely studied by many scholars and researches. The focus will be restricted on forecasting land market value on 34 years' annual dataset regarding real estate scenario. Finally, I take a short summary of several variety of ARCH and GARCH models.

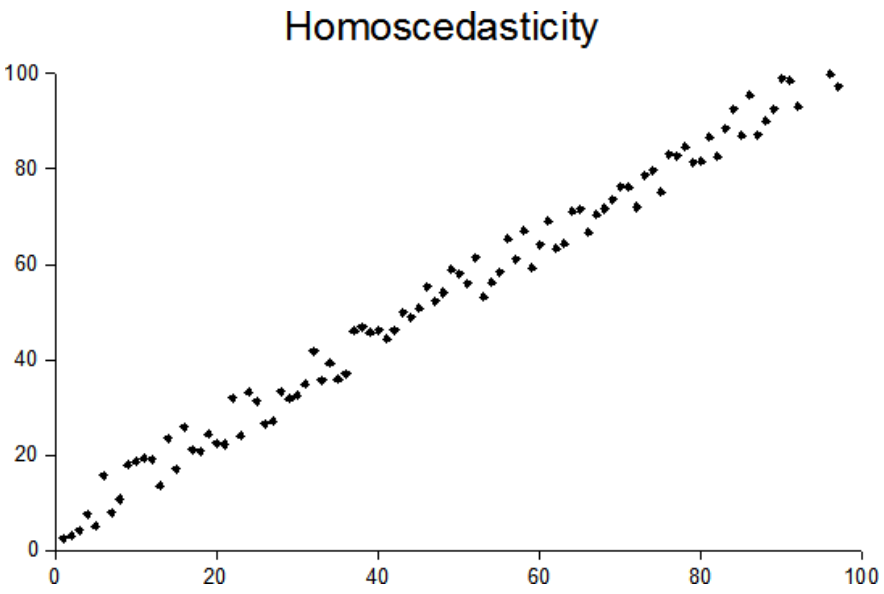


Figure 2.7: Homoscedasticity

The assumption of homoscedasticity (meaning same variance) is central to linear regression models. Homoscedasticity describes a situation in which the error term (that is, the noise or random disturbance in the relationship between the independent variables and the dependent variable) is the same across all values of the independent variables. Heteroscedasticity (the violation of homoscedasticity) is present when the size of the error term differs across values of an independent variable. The impact of violating the assumption of homoscedasticity is a matter of degree, increasing as heteroscedasticity increases.

To improve the accuracy of the forecasting model, we need to stabilized the variance. One approach for dealing with heteroscedasticity is to transform the dependent variable using one of the variance stabilizing transformations. A logarithmic transformation can be applied to highly skewed variables, while count variables can be transformed using a square root transformation.

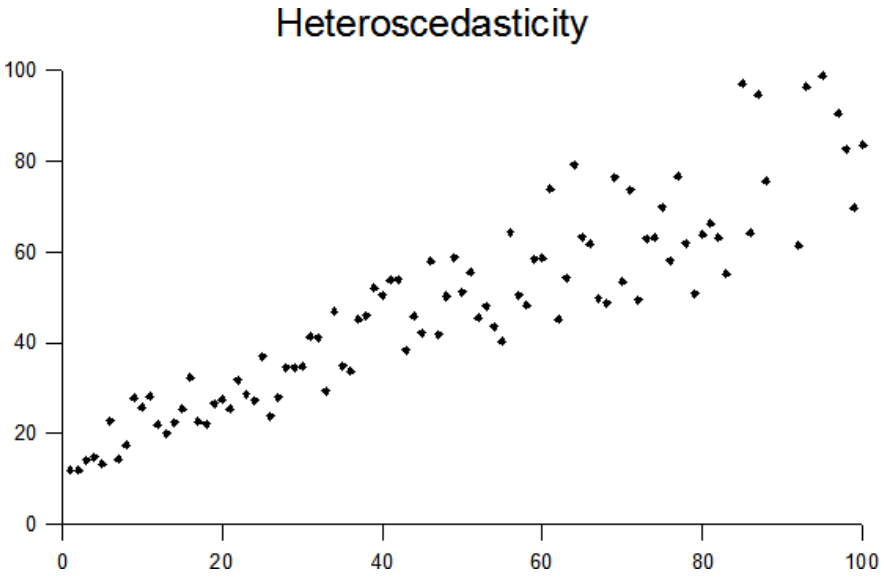


Figure 2.8: Heteroscedasticity

When data exhibits heteroscedasticity, there are two ways to handle that: If the function of heteroscedasticity is known, we transform it to homoscedasticity. If the function of heteroscedasticity is unknown, we try to apply the ARCH model. we can check the residual graphing and the squared residual graphing to detect the existence of heteroscedasticity. Also, we can use the autocorrelation coefficient of the squared residuals series to detect heteroscedasticity.

$$\rho = \sqrt{\frac{Cov(\epsilon_t^2, \epsilon_{t-k}^2)}{Var(\epsilon(_t)^2)}}. \tag{2.37}$$

We need to consider two different cases; When $\rho = 0, k = 1, 2, ...$; It indicates that the function of heteroscedasticity is the stochastic process. For this sake, the dataset has nothing to do with the estimation of heteroscedasticity. This is the most difficult problem. So far, no one can extract the heteroscedasticity information from it. Secondly, for $\rho \neq 0$ and $k = 1, 2, \cdots$. This is the case with which we are mainly concerned. If the ρ is not zero, it indicates the heteroscedastic function have the autocorrelation problems.

There are several approaches to deal with heteroscedasticity. If the error of variance at

different times is known, weighted regression is a good method. If, as is usually the case, the error variance is unknown and must be estimated from the original data, you can model the changing error variance.

In the standard regression and time series models we have covered so far, many diagnostic checks were based on the assumptions that we imposed on the errors: independent, identically distributed with zero mean, and constant variance. Our main concern has mostly been about the independence of the errors. The constant variance assumption is often taken as a given. In many practical cases and particularly in finance, it is fairly common to observe the violation of this assumption. A linear trend model, an exponential smoother, or even an ARIMA model would have failed to capture this phenomenon, as all assume constant variance of the errors. This will in turn result in the underestimation of the standard errors calculated using OLS and will lead to erroneous conclusions. There are different ways of dealing with this situation. For example, if the changes in the variance at certain time intervals are known, weighted regression can be employed. However, it is often the case that these changes are unknown to the analyst. Moreover, it is usually of great value to the analyst to know why, when, and how these changes in the variance occur. Hence, if possible, modeling these changes (i.e., the variance) can be quite beneficial.

Consider, for example, the simple $AR(m)$ model

$$y_t = \delta + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-m} + Z_t \tag{2.38}$$

Where Z_t is the uncorrelated, zero mean noise with changing variance. Please note that we used e_t to distinguish it from our general white noise error ϵ_t. Since we let the variance of e t change in time, one approach is to model Z_t^2 as an $AR(m)$ model as

$$Z_t = \xi_0 + \xi_1 z_{t-1}^2 + \xi_2 z_{t-2}^2 + \cdots + \xi_m z_{t-m}^2 + a_t \tag{2.39}$$

where a_t is a white noise sequence with zero mean and constant variance σ_a^2. In this no-

tation Z_t is said to follow an autoregressive conditional heteroskedastic process of order m, ARCH(m).

To check for a need for an ARCH model, once the ARIMA or regression model is fitted, not only the standard residual analysis and diagnostics checks have to be performed but also some serial dependence checks for Z_t^2 should be made.

In 1982, Engle introduced the ARCH and explicitly recognizes this type of temporal dependence. Moreover, Crawford and Fratantoni [25] applied a Markov-switching model to U.S. home price, and compare the performance with ARIMA and GARCH models. In 1991, Nelson put forward E-Garch (Exponential Garch) and Glosten, later on Johnathan and Runkle [53], introduced GJR-GARCH model. Besides, William Miles [21] evaluate the Forecasting Performance of Linear and Non-linear Models of House Prices. To correlate changes in volatility with changes in returns, Engle, Lilien, and Robins [52] proposed the GARCH-M model [33,34].

The ARCH model assumes that the changes in variance is a function of the realizations of squares of preceding errors.

To model a time series ϵ_t using an ARCH process, we assume that:

$$\sqrt{h_t} Z_t = \epsilon_t \tag{2.40}$$

$$Var(\epsilon_t |\ past\ \ data) = h_t. \tag{2.41}$$

$$\frac{\epsilon_t}{\sqrt{h_t}} \sim IID(0,1). \tag{2.42}$$

Where h_t is the conditional variance.

Especially, the order of the model needs to be specified for each of the parametric models before fitting the model to the in-sample data. The drawback of the ARCH model is that it usually required quite a high order to accurately be able to model the conditional variance.

The advantage of the ARCH model is the ability to model volatility clustering can be seen in the definition of the conditional variance where it is evident that a large Z_{t-i}^2 will

give rise to a large h_t^2.

$$h_t = \zeta_0 + \zeta_1 Z_{t-1}^2 + \zeta_2 Z_{t-2} + \cdots + \zeta_p Z_{t-p} \tag{2.43}$$

Hence the conditional variance of Z_t is

$$\begin{aligned} Var(Z_t|Z_{t-1}, \cdots) &= E(Z_t^2|Z_{t-1}^2, \cdots) && (2.44)\\ &= h(t) && (2.45)\\ &= \zeta_0 + \zeta_1 Z_{t-1}^2 + \zeta_2 Z_{t-2} + \cdots + \zeta_p Z_{t-p} && (2.46) \end{aligned}$$

We can also argue that the current conditional variance should also depend on the previous conditional variances as

$$h_t = \zeta_0 + \varsigma_1 h_{t-1} + \varsigma_2 h_{t-2} + \cdots + \varsigma_p h_{t-p} + \zeta_1 Z_{t-1}^2 + \zeta_2 Z_{t-2}^2 + \cdots + \zeta_q Z_{t-q} \tag{2.47}$$

In this notation, the error term h_t is said to follow a generalized autoregressive conditional heteroskedastic process of orders p and q, $GARCH(p, q)$, proposed by Bollerslev (1986). The equation (2.42) for conditional variance resembles an ARMA model. However, it should be noted that the model in Equation (2.17) is not a proper ARMA model, as this would have required a white noise error term with a constant variance for the MA part. But none of the terms on the right-hand side of the equation possess this property.

The properties of the $GARCH(p, q)$ model is quite similar to that of the ARCH but requires far less parameters to adequately model the volatility process. It is a model with additional lag terms. Such higher order models are often useful when a long span of data is used, like several decades of daily data or a year of hourly data. With additional lags, such models allow both fast and slow decay of information.

Although the error term, ϵ_t, can be autocorrelated in the regression model, it should be stationary. A non-stationary error structure could produce a spurious regression where

a significant regression can be achieved for a totally unrelated series as shown in Granger and Newbold (1986), and Phillips (1986). In such a case, one should properly difference the series before estimating the regression [18].

2.4.2 Modeling AR-GARCH

In 1985, Bollerslov proposed the GARCH model. The GARCH model might perform better in cases where theory suggests that the data-generating process produces true autoregressive conditional heteroscedasticity. This is the case in some economic theories of asset returns, and GARCH type models are often used for analysis of applied econometrics.
To fit the $AR(m) - GARCH(p,q)$ model, we consider the following formula:

$$\begin{cases} x_t &= f(t, x_{t-1}, x_{t-2}, \cdots) + \epsilon_t \\ \epsilon_t &= \sum_{k=1}^{m} \beta_k \epsilon_{t-k} + v_t \\ v_t &= \sqrt{h_t} e_t \\ h_t &= \omega + \sum_{i=1}^{p} \eta_i h_{t-i} + \sum_{j=1}^{q} \lambda_j v_{t-j}^2 \end{cases} \tag{2.48}$$

Where $f(t, x_{t-1}, x_{t-2}, \cdots)$ is the regression function of x_t; $e_t \sim N(0.1)$.

GARCH model is a generalization of ARCH model. It is similar to the ARCH model but with the addition of lagged conditional variances, h_{t-j}^2, as well as the lagged squared returns Z_{t-i}^2. The addition of the lagged conditional variances avoids the need for adding many lagged squared returns as was the case for the ARCH model to be able to appropriately model the volatility. In GARCH (p,q), when p=0, it is ARCH(q) model.

To fit a GARCH model for ϵ_t, we must follow a portmanteau Q test is used for autocorrelation in errors:

H_0: There is no evidence show that there are autocorrelation in residuals for some lag p.

H_1: There are some evidence show that there are autocorrelation in residuals for some lag p.

Table 2.6

Portmanteau Q and LM tests for ARCH Disturbances

Order	**Q**	$Pr > Q$	**LM**	$Pr > LM$
1	20.5339	$< .0001$	18.9236	$< .0001$
2	23.5834	$< .0001$	24.8639	$< .0001$
3	23.6226	$< .0001$	26.0909	$< .0001$
4	23.7134	$< .0001$	26.4050	$< .0001$
5	24.2642	0.0002	26.4050	$< .0001$
6	26.5167	0.0002	26.4050	0.0002
7	28.0117	0.0002	26.7401	0.0004
8	28.2805	0.0004	27.6699	0.0005
9	28.2825	0.0009	27.7360	0.0011
10	28.5920	0.0015	28.1616	0.0017
11	29.6284	0.0018	28.1646	0.0031
12	31.4359	0.0017	28.1658	0.0052

The table (2.6) indicates that the residuals of Land market value with log transformation present long-term heteroscedastic problems. The DW statistic is a number that tests for autocorrelation in the residuals from a statistical regression analysis. The DW statistic is always between 0 and 4. A value of 2 means that there is no autocorrelation in the sample. Values approaching 0 indicate positive autocorrelation and values toward 4 indicate negative autocorrelation.

An uncorrelated time series can still be serially dependent due to a dynamic conditional variance process. A time series exhibiting conditional heteroscedasticityor autocorrelation in the squared seriesis said to have ARCH effects. Engles ARCH test is a LM test to assess the significance of ARCH effects. LM tests tend to under-reject for small values of α, and over-reject for large values of α.

In table (2.7), we conclude that the DW statistics is 0.1937. It shows that the residuals are positively correlated. Under the α level 0.05, The probability of LM are all less than 0.05. Thus, we need to fit a ARCH model to get rid of auto-correlated problems in the residuals series.

Table 2.7

Ordinary Least Squares Estimates

SSE	2.71086564	DFE	32
MSE	0.08471	RMSE	0.29106
SBC	17.5513957	AIC	14.4986747
MAE	0.46842205	AICC	14.8857714
MAPE	2.74993807	Regress R-square	0.8127
Durbin-watson	0.1937	Total-square	0.8127

Specifically, the R^2 statistics indicated that the model as fitted explains 81.27% of the variability in Land Market Value. It shows that the model perfectly fits the data. All the estimates are significant except ARCH0 coefficient, the probability 0.2155 is greater than 0.05. It meets the assumption that the GARCH residual function $\frac{\epsilon_t}{\sqrt{h_t}}$ follow normal distribution. We combine the two results and fit the final $AR(2) - GARCH(1,1)$ model, as shown below:

$$\begin{cases} x_t &= 7.1365x_{t-1} + u_t \\ u_t &= -0.115u_{t-1} + \epsilon_t \\ \epsilon_t &= \sqrt{h_t}a_t, a_t \sim N(0, 0.01948) \\ h_1 &= 0.002816 + 1.0663\epsilon_{t-1}^2 - 0.0206\epsilon_{t-2}^2 + 0.0193h_{t-1} \end{cases}$$

Table 2.8

AR-Garch Estimates

Variable	**DF**	**ESTIMATES**	**Error**	**T-value**	**P-value**
Intercept	1	7.1365	0.1456	49.02	$< .0001$
t	1	0.1150	0.10215	5.34	$< .0001$
AR1	1	-0.9937	0.0370	-26.85	$< .0001$
ARCH0	1	0.002816	0.002273	1.24	0.2155
ARCH1	1	1.0663	0.5857	1.82	0.0686
ARCH2	1	-0.0206	0.005284	-3.89	$< .0001$
GARCH1	1	0.0193	0.005638	3.42	0.0006

Table 2.9

Parameter Estimates

SSE	0.66245257	Observations	34
MSE	0.01948	Uncond Var	
Log Likelihood	27.077195	Total R-Square	0.9542
SBC	-32.99622	Normality Test	5.8737
AIC	-42.15439	Pr > ChiSq	0.0531
AICC	-39.04328	RMSE	0.13957077
MAE	0.1009846	MAPE	1.19139757
HQC	-39.031194		

The investigation of conditional variance models has been one of the main areas of study in time series analysis of financial markets. Towards these ends, the GARCH model and its variations have been applied to many risk and volatility studies. We use above simple examples to illustrate the procedures [22].

In table (2.8), the parameter estimates table includes rows for the GARCH parameters. ARCH0 represents the estimate for the constant ω, ARCH1 represents η_1 , and GARCH1 represents λ_1. All the parameters are significant except ARCH0. But, in table (2.9), the total R-Squared value is around 95%. The normality test is not significant (p = 0.053), which is consistent with the hypothesis that the residuals from the GARCH model, are normally distributed. The MSE and MAE are extremely small, we consider we fit the model successfully.

In 1982, Engle introduced the auto-regression conditional heteroskedastic and explicitly recognizes this type of temporal dependence. The impact of violating the assumption of homoscedasticity is a matter of degree, increasing as heteroscedasticity increases. In econometrics, the well-known ARCH and GARCH models are the basic and important tools used to dealing with time series heteroscedasticity. In this Chapter, we gives the motivation behind the simplest ARCH and GARCH model and illustrates its usefulness in forecasting land market value in U.S. This section described the ARCH and GARCH models and will illustrate how they can be estimated with PROC AUTOREG in the SAS software.

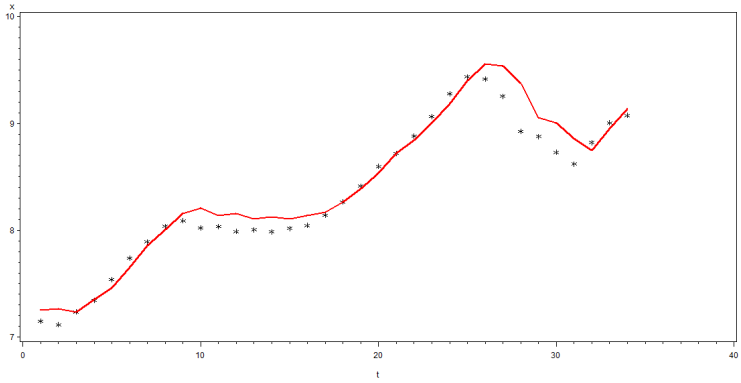


Figure 2.9: SAS FORECASTING GRAPHING

2.5 Introduction of VECM

This section aims to investigate the short and long run equilibrium between the $LLMV$ and $LCCI$ of America economy during $1982 - 2015$. Unit root test, co-integration test and finally error correction model are the econometric tools to establish the relationship between macroeconomic variables. In addition to this ordinary least square method is used to find out the LLMV and spurious regression. The findings reveal that the variables are non-stationary at their level and they become stationary in their first difference. There are two co-integration equations showing the long run relationship between LLMV and LCCI. There is short and long run equilibrium as indicated by the statistically significant coefficient of LCCI and error correction term.

Modeling the long-run when the variables are non-stationary is a new and expanding area of econometrics.It is new in that while it is possible to find antecedents in the literature dating back to, for example, the seminal work of Sargan (1964) on early forms of the error-correction model, it was really only in 1986 that co-integration became a familiar term in the literature. It is also rapidly expanding area, as witnessed by the number of articles that have been published since the mid-1980s. There have been, and continue to be, major new developments. However, co-integration analysis may have reached a plateau, and initial concentration on the theoretical side has now probably been superseded by attempts to apply the theory [50].

2.5.1 Spurious regression and Co-integration

Trends in the data can lead to spurious correlations that imply relationships between the variables in a regression equation, when all that is present are correlated time trends. The time trend in a trend-stationary variable can either be removed by regressing the variable on time or nullified by including a deterministic time trend as one of the regressors in the model. In such circumstances, the standard regression model is operating with stationary series which have constant means and finite variances, and thus statistical inferences are valid. Generally time series data are non-stationary if used to run regression may produce spurious regression which is not desirable. Saying the same thing again, regression of a non-stationary time series on another non-stationary time series may cause a spurious regression. Running regression on the non-stationary series at their level would generally be produced spurious regression. Such as the following Figure (2.11) shows the two non-stationary series LLMV and LCCI may be co-integrated,

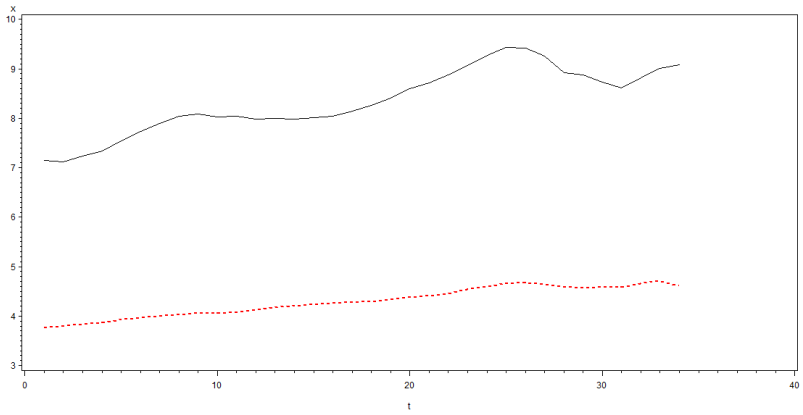


Figure 2.10: Co-integration for LLMV and LCCI

The economic interpretation of co-integration is that if two (or more) series are linked to form an equilibrium relationship spanning the long-run, then even though the series themselves may contain stochastic trends they will nevertheless move closely together over time and the difference between them will be stable. Thus the concepts of co-integration

mimics the existence of a long-run equilibrium to which an economic system converges over time, and u_t defined above can be interpreted as the disequilibrium error(i.e., the distance that the system is away from the equilibrium at time t). Co-integration is also linked very closely to the use of short-run error-correction models, thus providing a useful and meaningful link between the long- and short-run approach to econometric modeling [6].

Regressing a non-stationary variable on a deterministic trend generally does not yield a stationary variable. Thus, using standard regression techniques with non-stationary data can lead to the problem of spurious regressions involving invalid inferences based on t test and F test. For instance, consider the following equations:

$$y_t = y_{t-1} + u_t \tag{2.49}$$

$$x_t = x_{t-1} + v_t \tag{2.50}$$

This is, both x and y are uncorrelated non-stationary variables such that when the following regression model is estimated:

$$y_t = \beta_0 + \beta_1 x_t + \epsilon_t \tag{2.51}$$

It should generally be possible to accept the null $H_0 : \beta_1 = 0$. However, because of the non-stationary nature of the data, implying that ϵ_t is also non-stationary, any tendency for both time series to be growing leads to correlation which is picked up by the regression model, even though each is growing for very different reasons and at rates which are uncorrelated (i.e., δ_1 converges in probability to zero in the regression ($\Delta y_t = \delta_0 + \delta_1 \Delta x_t + \eta_t$)). Thus, correlation between non-stationary series does not imply the kind of causal relationship that might be inferred from stationary series.

The problem of spurious correlation, resulting in a non-zero estimate of β_1, is compounded by the fact that t statistics and F-statistics do not have the standard distributions generated by stationary series; with non-stationary series, there is a tendency to reject the null in both

cases, and this tendency in fact increases with the sample size. In a Monte Carlo experiment reported in Banerjee et al.(1993, pp.73-5), equation (2.42) was estimated 10,000, with x and y as defined in (2.40, 2.41), resulting in an estimated mean value for β_1 of -0.012, and an associated standard error of 0.006(Given a sample size of T=100), thus rejecting the null that $E(\beta_1) = 0$. Based on the 10,000 replications, the probability of rejecting the null of no association at the conventional significance level of 0.05 was found to be 0.753. (i.e., in 75.3 per cent of the regressions, values of $|t| > 1.96$ were obtained). This was due to the fact that the mean t-statistic obtained from the experiment was -0.12 instead of zero, with an associated standard deviation of 7.3. The non-standard distribution of the t-statistic accounts for the very high rejection rate of the null.

In summary, there is often a problem of falsely concluding that a relationship exists between two unrelated non-stationary series. This problem generally increases with the sample size and it can't be solved by attempting to de-trend the underlying series as would be possible with trend-stationary data. This leads to the question of when it is possible to infer a causal long-run relationships between non-stationary time series, based on estimating a standard regression such as (2.42).

2.5.2 Unit root

During the last decade applied economists attempting to estimate time series econometric models have been aware of certain difficulties that arise when unit roots are present in the data. To ignore this fact, and to proceed to estimate a regression model containing non-stationary variables at best ignores important information about the underlying processes generating the data, and at worst leads to nonsensical results. For this reason, it is incumbent on the applied researcher to test for the presence of unit root and if they are present to use appropriate modeling procedures. Detrending is not appropriate and simply differencing the data to remove non-stationary trend is only part of the answer. While the use of of differenced variables will avoid the spurious regression problems, it will also remove any

long-run information but to ensure that it reflects the co-movement of variables due to the underlying equilibrating tendencies of economic forces, rather than those due to common,but unrelated , time trends in the data.

Generally, time series data contains unit root meaning that these series are not stationary. ADF test (1979), generally popular method, is being applied to test the unit root under the hypothesis series has unit root. Akaike criterion has been followed to lag selection. The model to check the unit root is:

$$\Delta LLMV_t = \lambda_0 + \lambda_1 LLMV_{t-1} + \lambda_2 T + \sum_{i=1}^{n} \phi_i \lambda LLMV_{t-i} + \epsilon. \tag{2.52}$$

Where Δ is the difference operator X is the natural logarithm of the series. T is a trend variable. λ and ϕ are parameters to be estimated and $\epsilon$is the error term.

When discussing stationary and non-stationary time series, the need to test for the presence of unit roots in order to avoid the problem of spurious regression was stressed. If a variable contains a unit root then it is non-stationary and unless it combines with other non-stationary series to form a stationary co-integration relationship , then regressions involving the series can falsely imply the existence of a meaningful economic relationship.

In principle, it is very important to test the order of integration of each variable in a model, to establish whether it is non-stationary series. Also, as will be seen, testing whether a linear combination of variables co-integrate to form a stationary, equilibrium relationship.

- It is necessary to take account of the possibility that the underlying d.g.p. may inter alia, include a time trend.
- The d.g.p may be more complicated than a simple AR(1) process, and indeed may involve MA terms.
- It is known that when dealing with finite samples the standard tests for unit roots are biased towards accepting the null hypothesis of non-stationary when the true d.g.p is in fact stationary but close to having unit root.

Table 2.10

ADF unit root test

Type	**Lags**	ρ	$Pr < \rho$	τ	$Pr < \tau$	**F**	$Pr > F$
Zero Mean	0	-4.6496	0.1302	-1.62	0.0978		
	1	-25.1613	$< .0001$	-3.29	0.0017		
	2	296.4800	0.9999	-4.75	$< .0001$		
Single Mean	0	-4.5745	0.4526	-1.57	0.4864	1.37	0.728
	1	-24.8294	0.0006	-3.24	0.0262	5.64	0.0293
	2	297.5548	0.9999	-4.77	0.0006	11.72	0.0010
Trend	0	-4.6015	0.8320	-1.53	0.7976	1.19	0.9338
	1	-28.9145	0.0012	-3.30	0.0844	5.45	0.1295
	2	72.7197	0.9999	-5.30	0.0008	14.41	0.0010

In Dickey-Fuller tests, the second column in table (2.10) specifies three types of models, which are zero mean, single mean, or trend. The third column (ρ) and the fifth column (τ) are the test statistics for unit root testing. Other columns are their -values. You can see that both series have unit roots.

In summary, this section has shown that while in principle it is necessary to test for the presence of unit roots in order to avoid the problem of spurious regression, this is by no means a simple exercise. An approach testing strategy is based on the augmented Dickey-Fuller test with a generous lag structure which allows for both constant and trend terms, and then follows the sequential testing strategy suggested by Perron (1988). Clearly, the most important problem faced when applying unit root tests is their probable poor size and power properties (i.e., the tendency to over-reject the null when it is true and under-reject the null when it is false, respectively). This problem occurs because of the near equivalence of non-stationary and stationary processes in finite samples which makes it difficult to distinguish between trend-stationary and difference-stationary processes[8]. It is not really possible to make such definitive statements as "real GNP is not stationary "; Rather, unit root tests are more useful for indicating whether the finite sample data used exhibits stationary or non-stationary attributes.

2.5.3 Johansen Co-integration Test

Testing for co-integration using a single equation is problematic. If there are $n > 2$ variables in the model, and if $n - 1$ of these are not weakly exogenous, the single equation approach can be misleading, particularly if more than one co-integration relationship is present. If single equation methods are to be used, it would seem that the unrestricted dynamic modeling approach is more likely to produce unbiased estimates of the long-run relationship, with appropriate t-statistic and F-statistics. The test of co-integration associated with this approach is also more powerful against alternatives, such as the usual Engle-Granger static model. However, given that the number of co-integration vectors is unknown, and give the need to allow all variables to be potentially endogenous. There seems little advantages in starting from the single equation model. Rather, the multivariate VAR approach developed by Johansen (1998) is the more obvious place to begin testing for co-integration.

The Johansen technique is fast becoming an essential tool for applied economists wishing to estimate time series models. The implication that non-stationary variables can lead to spurious regressions unless at least one co-integration vector is present means that some form of testing for co-integration is almost mandatory. The uses of Engle-Granger procedure is giving way to the determination of co-integration relationship exists. The major problem facing those wishing to use Johansen's (1988) technique is that it is only just becoming available in a user-friendly fashion, following the release of PcFiml and Cats.

Johansen co-integration test procedure consists of estimating a vector autoregressive (VAR) models which includes difference as well as the levels of the non-stationary variables. The equation for Johansen co-integration test is given by:

$$\Delta LLMV_t = \gamma_1 \Delta LLMV_{t-1} + \gamma_{k-1} \Delta LLMV_{t-k+1} + \cdots + \pi LLMV_{t-k} + \epsilon_t \tag{2.53}$$

Where ϵ_t is Gaussian random variable and π and γ are matrices of parameters estimated using OLS. And $\pi = \alpha \times \beta'$ where α and β are $r \times k$ matrices.
The procedure of implement Johansen technique is followed by these steps:

- testing the order of integration of each variable that enters the multivariate model;
- setting the appropriate lag-length of the VAR model (in order to ensure Gaussian error terms in the VECM) and determining whether the system should be conditioned on any predetermined $I(0)$ variables (including dummies to take account of possible policy interventions);
- testing for reduced rank , including the issue of testing whether the system should be treated as an $I(2)$ rather than an $I(1)$ system;
- identifying whether there are trends in the data and therefore whether deterministic variables (a constant and trend) should enter the co-integration space or not;
- testing for weak exogeneity (which leads to the modelling of a partial system with exogenous variables);
- testing for linear hypotheses on co-integration relations;
- testing for unique co-integration vectors;
- joint tests involving restrictions on α and β.

Each of these will be considered in turn and examples provided. However, it is necessary first to briefly outline the Johansen model and the method of reduced rand regression used to estimate it.

Johansen co-integration test procedure consists of estimating a VAR models which includes difference as well as the levels of the non-stationary variables. In the cointegration rank test, the last two columns in table (2.11) explain the drift in the model or process.

Since the NOINT option is specified, the model is

$$\Delta\ LLMV_t = \pi LLMV_{t-1} + \Phi\Delta\ LLMV_{t-1} + \epsilon_t \tag{2.54}$$

The column Drift In ECM means there is no separate drift in the error correction model, and the column Drift In Process means the process has a constant drift before differencing.

Table 2.11

Co-integration Rank Test

$H_0 : R = r$	$H_1 : R > r$	**EV**	**Trace**	5% **CV**	**Drift in ECM**	**Drift in pro**
0	1	0.3111	14.6164	12.21	NOINT	Constant
1	1	0.0679	2.3190	4.14		

H_0 is the null hypothesis, and H_1 is the alternative hypothesis. The first row tests $r = 0$ against $r = 1$; the second row tests $r = 1$ against $r > 1$. The Trace test statistics in the fourth column are computed by $-T\sum_{i=r+1}^{k} log(1 - \lambda_i)$ where T is the available number of observations and λ_i is the eigenvalue in the third column. By default, the critical values at 5% significance level are used for testing. You can compare the test statistics and critical values in each row. There is one cointegrated process in this example since the Trace statistic for testing $r = 0$ against $r > 0$ is greater than the critical value, but the Trace statistic for testing $r = 1$ against $r > 1$ is not greater than the critical value.

However, there are two drawbacks of the Johansen method. One is that it takes a little getting used to interpreting the results and formulating hypotheses in this setting. In the VAR system all variables are treated symmetrically, as opposed to the standard univariate models that usually have a clear interpretation in terms of exogenous and endogenous variables. The other drawback of the VAR system is one has to model all the variables at the same time, which will be a problem if the relation for some variable is flawed. This may give bias in the whole system and one may have been better of conditioning on that variable. Further, the multidimensional VAR model uses many degrees of freedom [8].

2.5.4 VECM model specification

A vector error correction model (VECM) is a theoretically-driven approach useful for estimating both short-term and long-term effects of one time series on another.
The $VECM(P)$ form with the cointegration rank $r(\leq k)$ is written as

$$\Delta\ y_t = \delta + \pi y_{t-1} + \Phi\Delta\ y_{t-1} + \epsilon_t \tag{2.55}$$

Where ϵ_t is Gaussian random variable and are matrices of parameters estimated using OLS. The component π produces different linear combinations of levels of the time series X_t as such the matrix y_t contains information about the long run properties of the system describe by the model. For instance, if the rank of the matrix π is 0, then no series of the variables can be expressed as a linear combination of the remaining series. This indicates that there does not exists a long run relationship among the series of the VAR model as a test of co-integration a rank of 0 means integration is rejected. On the other hand, if the rank of the coefficient matrix π is 1, or greater than 1 then there exists 1 or more co-integrating vectors. This indicates a long run relationship or that the series exhibits significant evidence or behaving as a co-integrated system.

Table 2.12

Model Parameter Estimates

Equation	**Parameter**	**S.E**	**Error**	**t-value**	$Pr > \|t\|$	**Variable**
$LLMV_t$	$AR(111)$	-0.21347	0.07000			$LLMV_{t-1}$
	$AR(112)$	0.41092	0.13475			$LCCI_{t-1}$
	$AR(211)$	0.61141	0.17196	3.56	0.0015	$D(LLMV_{t-1})$
	$AR(212)$	1.22096	0.60819	2.01	0.0556	$D(LCCI_{t-1})$
$LCCI_t$	$AR(121)$	-0.01568	0.01988			$LLMV_{t-1}$
	$AR(122)$	0.03018	0.03827			$LCCI_{t-1}$
	$AR(221)$	0.01805	0.0488	0.37	0.7148	$D(LLMV_{t-1})$
	$AR(222)$	0.81207	0.17274	4.70	0.0001	$D(LCCI_{t-1})$

The values and t-values corresponding to the parameters $AR1$ are missing since the parameters $AR(1)$ have non-Gaussian distributions. The parameter $AR2$ corresponds to the elements in the differenced lagged AR coefficient matrix. The "D" prefixed to a variable name in Table (2.12) implies differencing.

The following statements fit a VECM(2) form to the simulated data. From the result in Table (2.12) , the time series are cointegrated with rank=1. You specify the ECM= option with the RANK=1 option. For normalizing the value of the cointegrated vector, you specify the normalized variable with the NORMALIZE= option. The PRINT=(IARR) option provides the VAR(2) representation. The VARMAX procedure output is shown in Figure through Table (2.12) .

This process can be given the following $VECM(2)$ representation with the co-integration rank one:

$$\Delta y_t = \begin{bmatrix} -0.21347 & 0.41092 \\ -0.01568 & 0.03018 \end{bmatrix} y_t + \begin{bmatrix} 0.61141 & 1.22096 \\ 0.01805 & 0.81207 \end{bmatrix} \Delta y_{t-1} + \epsilon_t \tag{2.56}$$

The ECM= option produces the estimates of the long-run parameter, β, and the adjustment coefficient,α. In table (2.12), "1" indicates the first column of the and matrices. Since the cointegration rank is 1 in the bivariate system, and α and β are two-dimensional vectors. The estimated cointegrating vector is 1 . Therefore, the long-run relationship between x_t and y_t and is $x_t = 1.925y_t$. The first element of $\hat{\beta}$ is 1 since it is specified as the normalized variable.

The estimated value of $AR(211)$ is 0.61142 . It is individually significant at 5% level (Table 2.13). This coefficient represents the short run coefficient and represent the short run equilibrium. It tells about the rate at which the previous period disequilibrium of the system is being corrected. The value of $AR(211)$ is 0.61142 meaning that system corrects its previous period disequilibrium at a speed of 61.142% between variables LLMV and LCCI.

$LLMV_{t-1}$ is one period lag error correction term or residual. It guides the variables

(LLMV and LCCI) of the system to restore back to equilibrium or it corrects disequilibrium. To happen this, the sign of this should be negative and significant. Parameter $AR(111)$ represents its coefficient. It tells about the rate at which it corrects the previous period disequilibrium of the system if it is negative and significant. The coefficient of $AR(111)$ is negative (-0.21347, Table 2.13) and is significant at 1% level meaning that system corrects its previous period disequilibrium at a speed of 21.35% annually. It implies that the model identified the sizable speed of adjustment by 21.35% of disequilibrium correction yearly for reaching long run equilibrium steady state position.

Table 2.13

Infinite Order AR Representation

Lag	**Variable**	Y_1	Y_2
1	x	1.39794	1.63188
	y	0.00237	1.84225
2	x	-0.61141	-1.22096
	y	-0.01805	-0.81207
3	x	0	0
	y	0	0

The PRINT=(IARR) option in the previous SAS statements prints the reparameterized coefficient estimates. For the LAGMAX=3 in the SAS statements, the coefficient matrix of lag 3 is zero.

One final point that is worth emphasizing, and which should be fairly obvious from the above overview of this section, is that an applied economist should really begin his or her analysis by using a multivariate framework and not by using single equation approach. The exception will obviously be when only two variables are involved. The main reason for taking a system approach from the outset is that to do otherwise restricts the practitioner to considering only one interested in one vector, it is probable that he will not obtain consistent and efficient estimates without allowing for the possibility of other co-integration vectors [35].

Chapter 3

3 Advanced Time Series forecasting model

3.1 Bayesian Time Series forecasting model

The analysis of time series data presents some of the most difficult analytical challenges: you typically have the least amount of data to work with, while needing to inform some of the most important decisions. For example, time series analysis is frequently used to do demand forecasting for corporate planning, which requires an understanding of seasonality and trend, as well as quantifying the impact of known business drivers. But herein lies the problem: you rarely have sufficient historical data to estimate these components with good precision. And, to make matters worse, validation is more difficult for time series models than it is for classifiers and your audience may not be comfortable with the embedded uncertainty.

Bayesian forecasting is a natural product of a Bayesian approach to inference. The Bayesian approach in general requires explicit formulation of a model, and conditioning on known quantities, in order to draw inferences about unknown ones. In Bayesian forecasting, one simply takes a subset of the unknown quantities to be future values of some variables of interest. This chapter presents the principles of Bayesian forecasting, and describes recent advances in computational capabilities for applying them that have dramatically expanded the scope of applicability of the Bayesian approach [11]. It describes historical developments and the analytic compromises that were necessary prior to recent developments, the application of the new procedures in a variety of examples, and reports on two long-term Bayesian forecasting exercises.

In many forecasting problems there is little or no historical information available at the time initial forecasts are required. Consequently, the initial forecasts must be based on subjective considerations. As information becomes available, this subjective information can be modified in light of actual data. An example of this is forecasting demand for seasonal

clothing, which, because of style obsolescence, has a relatively short life. In this industry a common practice is to, at the start of the season, make a forecast of total sales for a product during the season and then as the season progresses the original forecast can be modified taking into account actual sales.

Bayesian methods can be useful in problems of this general type. The original subjective estimates of the forecast are translated into subjective estimates of the forecasting model parameters. Then Bayesian methods are used to update these parameter estimates when information in the form of time series data becomes available. This section gives a brief overview of the Bayesian approach to parameter estimation and demonstrates the methodology with a simple time series model. The method of parameter estimation makes use of the Bayes theorem.

3.1.1 Bayesian Priors, Posteriors, and Estimators

Let y be a random variable with probability density function that is characterized by an unknown parameter θ. We write this density as $f(y|\theta)$ to show that the distribution depends on θ. Assume that θ is a random variable with probability distribution $\pi(\theta)$ which is called the prior distribution for θ. Usually, the recent data are much more useful for forecasting and the own lags are more useful for that specific variable prediction. If the number of parameters increased, the model has to be more restrictive.

$$\pi(\theta|y) \propto f(y|\theta)\pi(\theta) \tag{3.1}$$

These prior views are tested in the training out-of-sample forecasting exercise. Hyperparameters which govern prior views are selected so that they produce most accurate forecasts in the training sample. The prior distribution summarizes the subjective information that we have about θ, and the treatment of θ as a random variable is the major difference between Bayesian and classical methods of estimation. If we are relatively confident about the

value of θ we should choose prior distribution with a small variance and if we are relatively uncertain about the value of θ we should choose prior distribution with a large variance.

If $Y_1, Y_2, Y_3, \cdots, Y_n$ denote the random variables associated with a sample of size n, we used the notation $L(y_1, y_2, \cdots, y_n|\theta)$ to denote the likelihood of the sample. In the discrete case, this function is defined to be the joint probability $P(Y_1 = y_1, Y_2 = y_2, \cdots, Y_n = y_n)$, and in the continuous case, it is the joint density of $Y_1, Y_2, Y_3, \cdots, Y_n$ evaluated at $y_1, y_2, \cdots, y_n$. The parameter θ is included among the arguments of $L(y_1, y_2, \cdots, y_n|\theta)$ denote that this function depends explicitly on the value of some parameter θ. In the Bayesian approach, the unknown parameter θ is viewed to be a random variable with a probability distribution, called the prior distribution of θ. This prior distribution is specified before any data are collected and provides a theoretical description of information about θ that was available before any data were obtained. In our initial discussion, we will assume that the parameter θ has a continuous distribution with density $g(\theta)$ that has no unknown parameters [36]. Using the likelihood of the data and the prior on θ, it follows that the joint likelihood of $Y_1, Y_2, \cdots, Y_n, \theta$ is

$$f(y_1, y_2, \cdots, y_n, \theta) = L(y_1, y_2, \cdots, y_n|\theta) \times g(\theta)$$

And that the marginal density or mass function of $Y_1, Y_2, Y_3, \cdots, Y_n$ is

$$m(y_1, y_2, \cdots, y_n) = \int_{-\infty}^{+\infty} L(y_1, y_2, \cdots, y_n|\theta) \times g(\theta) d\theta$$

Finally, the posterior density of $\theta|y_1, y_2, \cdots, y_n$ is

$$g*(\theta|y_1, y_2, \cdots, y_n) = \frac{L(y_1, y_2, \cdots, y_n|\theta) \times g(\theta)}{\int_{+\infty}^{-\infty} L(y_1, y_2, \cdots, y_n|\theta) \times g(\theta) d\theta}$$

The posterior density summarizes all of the pertinent information about the parameter θ by making use of the information contained in the prior for θ and the information in the

data.

3.1.2 Bayesian VAR model

The Bayesian approach was introduced to a reevaluation of the VAR approach based on the Bayesian principles. Thus, the VAR approach was characterized by several deficiencies, especially due to the over-parameterization problems. The Bayesian approach proposes a solution to this problem due to the fact that it does not ponder too much any of the parameters of the model. However the emphasis falls on the use of prior distributions for the parameters, the prior distributions being a key factor in the BVAR approach. Therefore, the advantages and disadvantages of VAR model and BVAR were listed in table 3.1.

Table 3.1

Comparison of Bayesian VAR and VAR models

	advantages	**disadvantages**
Bayesian VAR model	reducing root mean square imposing some prior restrictions on parameters percent error Easier and more accurate assessment of uncertainty.	
VAR model	Fairer assumptions about data interaction of different related variables in forecasting macroeconomic variable	over-parameterization the loss of degrees of freedom which exponentially decrease for the number of lags included over-fitting phenomenon

A particular type of Bayesian VAR (BVAR) imposing the Minnesota prior or Littermans prior have been used in many regional studies recently. Some of the important studies using Bayesian VAR of Littermans type are for Minnesota [2], New York state [44], Texas [47], ROMANIAN GDP [9], South African [20] and Indiana Reginal economy [29].

In this section, we specify a BVAR model with Littermans prior. Also, we first develop a VAR model for the Land Market Value in USA and then modify the model using Littermans

prior to develop a Bayesian VAR model. Next, we compare the out of sample forecasts from VAR with our modified VAR model.

When developing the BVAR model, Litterman has made some assumptions on the unrestricted VAR model given by the following equations [60].

$$Y_t = \mu + \Pi_1 y_{t-1} + \Pi_2 y_{t-2} + \cdots + \Pi_p y_{t-p} + \epsilon_t \tag{3.2}$$

$$y_t - y_{t-1} = c + \epsilon_t \tag{3.3}$$

As for example, writing the nth equation in a BVAR model

$$\begin{aligned} Y_{i,t} &= c_i + \Phi^1_{i1}\ y_{1,t-1} + \Phi^1_{i2}\ y_{2,t-1} + \Phi^1_{in}\ y_{n,t-1} \cdots + \Phi^2_{i1}\ y_{1,t-2} \\ &\quad + \Phi^2_{i2}\ y_{1,t-2} + \cdots + \Phi^2_{in}\ y_{n,t-2} + \cdots + \Phi^p_{n1}\ y_{1,t-p} + \Phi^p_{i2}\ y_{n,t-p} + \Phi^p_{in} y_{n,t-p} + \epsilon_{n,t}. \end{aligned} \tag{3.4}$$

The Φ^s_{ij} gives the coefficient relating y_{it} to $y_{j,t-s}$ Litterman (1980) assumed that $\Phi^1_{ii} = 1$ and all the other $\Phi^{(s)}_{ij} = 0$. These 0 and 1 values characterize the mean of the prior distribution for the coefficients. Moreover, Litterman (1980) assumed that

$$\Phi^{(1)}_{ii} \sim N\big(1, \gamma^2\big) \qquad \Phi^{(s)}_{ij} \sim N\left(0, \frac{\gamma^2}{s^2}\right) \qquad \Phi^{(s)}_{ij} \sim N\big(0, [S(i,j,l)]^2\big) \tag{3.5}$$

Although each equation $i = 1, 2, ..., n$ of the VAR is estimated separately, the same value γ is used for each i. Smaller values of γ mean greater confidence in the prior information.

According to Litterman (1986), the prior for the variance is the only other prior to be set, the standard error on the coefficient estimate of lag l of variable j in equation i is denoted by a standard deviation of the form $S(i, j, l)$:

$$S(i,j,l) = \frac{[\gamma g(l) f(i,j)](\sigma_i)}{\sigma_j} \tag{3.6}$$

And

$$f(i,j) = \begin{cases} 1 \;\; when\ i = j, \\ \\ w_{ij} \;\; when\ i \neq j, . \end{cases}$$

where $\frac{\sigma_i}{\sigma_j}$ is correction for the scale for the series i compared with j and $0 < \gamma < 1$. In Eq (26), the model requires choosing specific values for $g(l)$(the lag decay) and γ, the tightness parameter and the standard deviation on the first own lag, will improve forecasting performance. Thus, the parameter $g(l)$ measures the tightness on lag l with respect to lag 1, and is assumed to have a harmonic shape with a decay factor of λ [24].

Litterman [2] a found that tight priors around zero on coefficients of other variables provide better forecast. Lai and Roy [34] recommended a value of $\lambda = 0.7$ in concert with $\gamma = 0.9$. Kinal and Ratner [44] used $\lambda = 0.40$ and $\gamma = 0.90$.

In this section, we used the Litterman prior as described in the previous paragraph to forecast the macroeconomic variables of the Land Market Value in USA. This paper extends the existing literature by developing a Bayesian VAR forecasting model for USA economy [44].

Next, we fit the BVAR (2) for LLMV with prior λ is 1 and prior θ is 0.9. The variance estimate for the innovation is approximately 0.00957.

Table 3.2

Model Parameter Estimates for BVAR(2)

Equation	**Parameter**	**Estimates**	**S.E**	**t-value**	$Pr > \|t\|$	**Variable**
$LLMV_t$	$AR(1\ 1\ 1)$	1.60022	0.12774	12.53	0.0001	$LLMV_{t-1}$
	$AR(2\ 1\ 1)$	-0.59738	-0.59738	-4.65	0.0001	$LLMV_{t-2}$

The macroeconomic modeling did not remain indifferent to these developments which actually coincided with an internal need for redefining itself. The macroeconomics concept, as Schorfheide [11] pointed out, was basically associated with the structural equation models

due to Cowles Commission between 1950 and 1970. Following Lucas critique [45], who showed that the structural parameters are not inelastic with respect to changes in economic policy, which led to the first major revolution through the introduction of VAR models, due to Sims [6].

The Table (3.2) shows that how we improve the accuracy and select the best model for land market value. Because when we increase the p with different parameters, the lagged variables are not significant and the RMSE is bigger than BVAR (2).

Bayesian analysis requires explicit specification of the prior distribution to be used in the analysis. Since, both the prior distribution and sample data are required to obtain the posterior distribution, the choice of the prior distribution depends on the knowledge and experience of the researcher. Some theoretical assumptions are also effective in the choice of the priors. There are a lot of different views about including the prior information in the analysis. The main idea in BVAR models is that the model parameters are random variables. The mechanism of this idea is representing the prior information for all the unknown quantities through a prior distribution and combining them with the objective infor- mation coming from observations to obtain the posterior distributions. Posterior distributions are obtained by the application of Bayes Theorem. In general, the choice of the prior distribution depends on the structure of the available information.

Usually, the choice of a particular econometric model is not pre-specified by theory and many competing models can be entertained. Comparing models can be done formally in a Bayesian framework through so-called posterior odds, which is the product of the prior odds and the Bayes factor. The Bayes factor between any two models is the ratio of the likelihoods integrated out with the corresponding prior and summarizes how the data favor one model over another. Given a set of possible models, this immediately leads to posterior model probabilities. Rather than choosing a single model, a natural way to deal with model uncertainty is to use the posterior model probabilities to average out the inference.

3.2 Artificial Neural Networks

In the previous Chapter we have discussed the important stochastic methods for time series modeling and forecasting. Artificial neural networks (ANNs) approach has been suggested as an alternative technique to time series forecasting and it gained immense popularity in last few years. The basic objective of ANNs was to construct a model for mimicking the intelligence of human brain into machine []. Similar to the work of a human brain, ANNs try to recognize regularities and patterns in the input data, learn from experience and then provide generalized results based on their known previous knowledge. Although the development of ANNs was mainly biologically motivated, but afterwards they have been applied in many different areas, especially for forecasting and classification purposes [30]. Below we shall mention the salient features of ANNs, which make them quite favorite for time series analysis and forecasting.

First, ANNs are data-driven and self-adaptive in nature [66]. There is no need to specify a particular model form or to make any a priori assumption about the statistical distribution of the data; the desired model is adaptively formed based on the features presented from the data. This approach is quite useful for many practical situations, where no theoretical guidance is available for an appropriate data generation process.

Second, ANNs are inherently non-linear, which makes them more practical and accurate in modeling complex data patterns, as opposed to various traditional linear approaches, such as ARIMA methods [5, 8, 20]. There are many instances, which suggest that ANNs made quite better analysis and forecasting than various linear models.

Finally, as suggested by Hornik and Stinchcombe [22], ANNs are universal functional approximators. They have shown that a network can approximate any continuous function to any desired accuracy. ANNs use parallel processing of the information from the data to approximate a large class of functions with a high degree of accuracy. Further, they can deal with situation, where the input data are erroneous, incomplete or fuzzy [20].

A neural network can be thought of as a network of neurons which are organised in layers.

The predictors (or inputs) form the bottom layer, and the forecasts (or outputs) form the top layer. There may also be intermediate layers containing hidden neurons.

The simplest networks contain no hidden layers and are equivalent to linear regressions. Figure 3.1 shows the neural network version of a linear regression with four predictors. The coefficients attached to these predictors are called weights. The forecasts are obtained by a linear combination of the inputs. The weights are selected in the neural network framework using a learning algorithm that minimises a cost function such as the MSE. Of course, in this simple example, we can use linear regression which is a much more efficient method of training the model. Once we add an intermediate layer with hidden neurons, the neural network becomes non-linear. A simple example is shown in Figure 3.1.

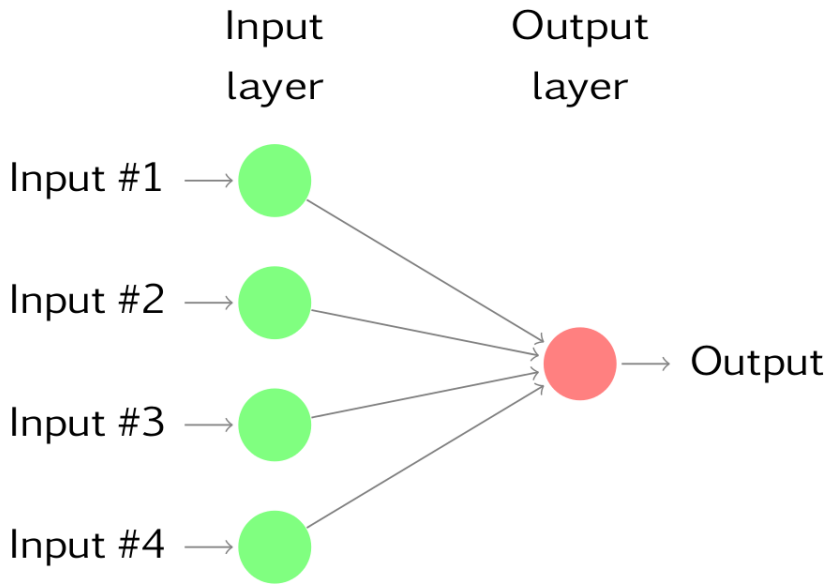


Figure 3.1: ANN 1:A simple neural network equivalent to a linear regression

This is known as a multilayer feed-forward network, where each layer of nodes receives inputs from the previous layers. The outputs of the nodes in one layer are inputs to the next layer. The inputs to each node are combined using a weighted linear combination. The result is then modified by a nonlinear function before being output. For example, the inputs into hidden neuron j in Figure 3.2 are combined linearly to give the input for the next layer.

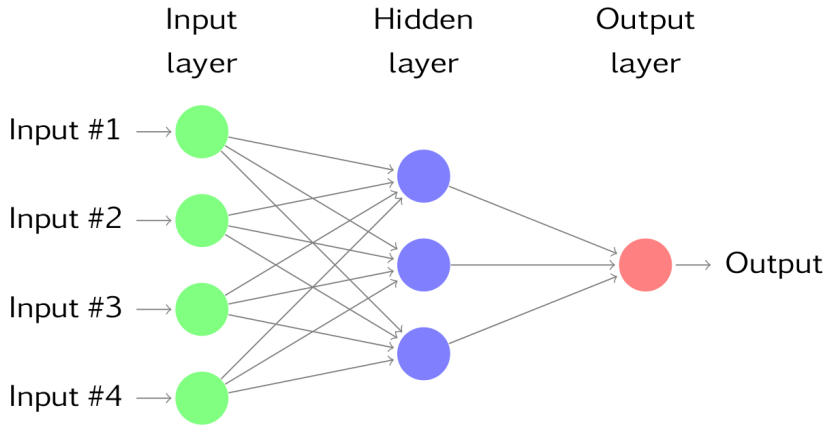


Figure 3.2: ANN2: A neural network with four inputs and one hidden layer with three hidden neurons.

This tends to reduce the effect of extreme input values, thus making the network somewhat robust to utilizers.

$$z_j = b_j + \sum_{i=1}^{4} w_{i,j} x_i. \tag{3.7}$$

In the hidden layer, this is then modified using a nonlinear function such as a sigmoid,

$$s(z) = \frac{1}{1 + e^{-z}} \tag{3.8}$$

The parameters b_1, b_2, b_3 and $w_{1,1}, \cdots, w_{4,3}$ are "learned" from the data. The values of the weights are often restricted to prevent them from becoming too large. The parameter that restricts the weights is known as the decay parameter, and is often set to be equal to 0.1.

The weights take random values to begin with, and these are then updated using the observed data. Consequently, there is an element of randomness in the predictions produced by a neural network. Therefore, the network is usually trained several times using different random starting points, and the results are averaged.

The number of hidden layers, and the number of nodes in each hidden layer, must be

specified in advance. The number of nodes in hidden layer should be between the number of inputs and outputs. An alternative way to calculate the nodes in the hidder layer is two thirds of inputs plus output functions. We will consider how these can be chosen using cross-validation later in this chapter [30].

3.2.1 Neural network autoregression

With time series data, lagged values of the time series can be used as inputs to a neural network, just as we used lagged values in a linear autoregression model. We call this a neural network autoregression or NNAR model.

In this section, we only consider feed-forward networks with one hidden layer, and we use the notation $NNAR(p, k)$ to indicate there are p lagged inputs and k nodes in the hidden layer. For example, a NNAR(9,5) model is a neural network with the last nine observations $(y_{t1}, y_{t2}, \cdots, y_{t9})$ used as inputs for forecasting the output y_t, and with five neurons in the hidden layer. $ANNAR(p, 0)$ model is equivalent to an $ARIMA(p, 0, 0)$ model, but without the restrictions on the parameters to ensure stationarity.

The $nnetar()$ function fits an $NNAR(p, P, k)m$ model. If the values of p and P are not specified, they are selected automatically. For non-seasonal time series, the default is the optimal number of lags (according to the AIC) for a linear $AR(p)$ model. When it comes to forecasting, the network is applied iteratively. This process proceeds until we have computed all the required forecasts.

3.2.2 Prediction Interval

Unlike most of the methods considered in this book, neural networks are not based on a well-defined stochastic model, and so it is not straightforward to derive prediction intervals for the resultant forecasts. However, we can still compute prediction intervals using simulation where future sample paths are generated using bootstrapped residuals.

The neural network fitted to the sunspot data can be written as

$$y_t = f(y_{t-1}) + \epsilon_t \tag{3.9}$$

where $y_{t1} = (y_{t1}, y_{t2}, \cdots, y_{t10})$ is a vector containing lagged values of the series, and f is a neural network with 6 hidden nodes in a single layer. The error series $\{\epsilon_t\}$ is assumed to be homoscedastic (and possibly also normally distributed).

We can simulate future sample paths of this model iteratively, by randomly generating a value for $\{\epsilon_t\}$, either from a normal distribution, or by resampling from the historical values. So if ϵ^*_{T+1} is a random draw from the distribution of errors at time $T+1$, then

$$y^*_{T+1} = f(y_T) + \epsilon^*_{T+1} \tag{3.10}$$

is one possible draw from the forecast distribution for y_{T+1}. Setting $Y^*_{T+1} = (y_{T+1}, y_T, \cdots, y_{T-6})$, we can then repeat the process to get

$$y^*_{T+2} = f(y^*_{T+1}) + \epsilon^*_{T+2}. \tag{3.11}$$

In this way, we can iteratively simulate a future sample path. By repeatedly simulating sample paths, we build up knowledge of the distribution for all future values based on the fitted neural network. A major advantage of neural networks is their ability to provide flexible nonlinear mapping between inputs and outputs. They can capture the nonlinear characteristics of time series well.

Here is a simulation of 9 possible future sample paths for the $LLMV$ data. Each sample path covers the next 30 years after the observed data.

```
sim <- ts(matrix(0, nrow=30L, ncol=9L),
  start=end(mydata10)[1L]+1L)
for(i in seq(9))
```

```
  sim[,i] <- simulate(fit, nsim=30L)
autoplot(mydata10) + autolayer(sim)
```

A desired network model should produce reasonably small error not only on within sample (training) data but also on out of sample (test) data [20]. Due to this reason immense care is required while choosing the number of input and hidden neurons. However, it is a difficult task as there is no theoretical guidance available for the selection of these parameters and often experiments, such as cross-validation are conducted for this purpose [30].

Another major problem is that an inadequate or large number of network parameters may lead to the overtraining of data [66]. Overtraining produces spuriously good within-sample fit, which does not generate better forecasts. To penalize the addition of extra parameters some model comparison criteria, such as AIC and BIC can be used [39].

In summary we can say that ANNs are amazingly simple though powerful techniques for time series forecasting. The selection of appropriate network parameters is crucial, while using ANN for forecasting purpose. Also a suitable transformation or rescaling of the training data is often necessary to obtain best results.

One of the most important steps in developing a satisfactory forecasting model such as ANN and SVM models is the selection of the input variables. In this study, the input structures having various input variables are determined by setting the input layer nodes equal to the number of the lagged variables from river flow data, $(x_{t-1}, x_{t-2}, ..., x_{t-p})$ where, p is time delay [73].

In this study, a typical feed-forward ANN model is constructed for time series forecasting. The training and testing data were normalized in the range zero to one. From the input layer to the hidden layer, the hyperbolic tangent sigmoid transfer function has been commonly used in time series was applied. From the hidden layer to the output layer, a linear function was employed as the transfer function because the linear function is known to be robust for a continuous output variable.

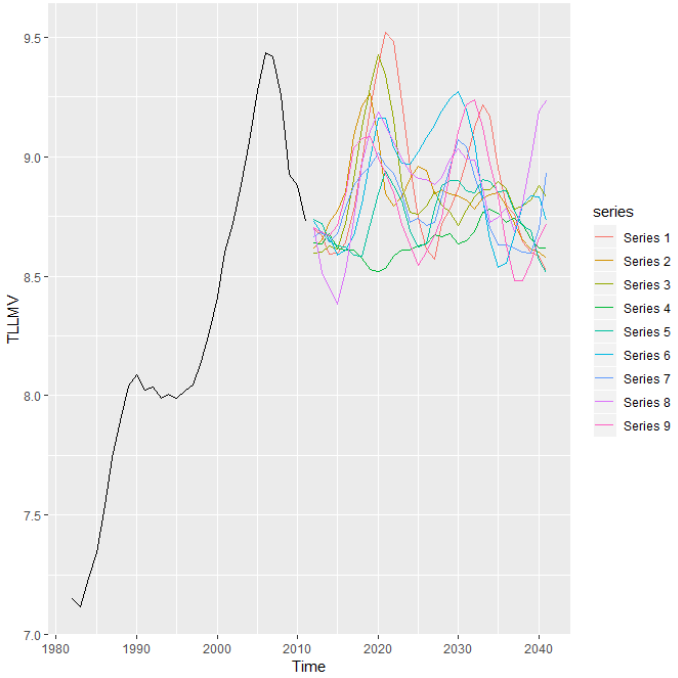


Figure 3.3: ANN3:Plot

3.3 Support Vector Regression

Support vector regression has been introduced to solve regression and prediction problems. It has become a hot topic of intensive study due to its successful application in classification tasks and regression tasks, especially on time series prediction and financial related applications. SVMs are state-of-the-art tools for non-linear input-output knowledge discovery. The support vector regressor (SVR) is for regression and function approximation. The combination of clustering and a state-of-the-art technique has been proposed for knowledge gain and accuracy improvement in a complex pharmacokinetic prediction problem [75]. The power and versatility of the SVR machines have allowed fast and reliable prediction schemes [67].

Support vector regression (SVR) is a powerful technique for predictive data analysis. SVM is also an effective tool in high-dimensional spaces, which is particularly applicable to document classification and sentiment analysis where the dimensionality can be extremely large. The goal of the regression problem is to approximate a function. The application of SVM for time series forecasting is relatively new. The initial results shown in this research

have clearly demonstrated the potential of this approach in predicting time series data. Future works can be centered on investigating the performance of SVM with different kernel functions and optimal hyper parameters of SVM forecasting model, which has the potential to improve the accuracy of the forecast. The research results demonstrate that SVM method is a promising alternative approach for the prediction of the time series data [41].

3.3.1 SVR models

In this section, we introduce ϵ-support vector regression model which are used to forecast the Land market value.

Suppose we are given a training dataset of nnn points of the form (x_i, y_i) where the y_i indicate the class to which the point x_i belongs to. Suppose that $|x_i| = d$ for all i, that is, all the points live in R^d.

A support vector machine is looking for a maximum-margin hyperplane that divides the groups such that the distance from the hyperplane to any point from either group is maximized. One way to think about this hyperplane is the set of points v that satisfy.

$$w \cdot v - b = 0 \qquad \textit{where } w \textit{ is thenormal vector to the hyperplane.}$$

While implementing SVR technique, the user needs to select the appropriate kernel function. The selection of kernel function is a tricky and requires optimization techniques for the best selection. A discussion on kernel selection is outside the scope of discussion for this section. In the constructed SVR model, we used the automated kernel selection provided by R. Radius Basis Function (RBF) kernel is used in the above model. Given a non-linear relation between the variables of interest and difficulty in kernel selection, we would suggest the beginners to use RBF as the default kernel [58]. The kernel function transforms our data from non-linear space to linear space. The kernel trick allows the SVR to find a fit and then

data is mapped to the original space. Now let us represent the constructed SVR model:

$$Y_i = W \cdot x_i + b \tag{3.12}$$

$$Y_i = W \cdot K(x_i, y_i) + b \tag{3.13}$$

Where $K(x_i, y_i)$ denotes the kernel function. The value of the kernel is equal to the inner product of two vectors x_i and y_i in the feature space $\phi(x_i)$ and $\phi(y_i)$, that is, $K(x_i, y_i) = \phi(x_i) \times \phi(y_i)$. The $\phi(x)$ denotes the high dimensional space. The corresponding conclusions are similar with the linear case except that the inner product $K(x_i, y_i)$ is replaced by (x_i, y_i). The coefficient are W and b are estimated by minimizing the regularized function:

$$Minimize \qquad R(w, b, \epsilon^*) = \frac{1}{2}||w||^2 + C\sum_{i=1}^{l}\left(\xi_i + \xi_i^*\right), \tag{3.14}$$

$$S.T(w \cdot x_i) + b - y_i \leq \xi + \xi_i, i = 1, \cdots, l \tag{3.15}$$

$$y_i - (w \cdot x_i) - b \leq \xi + \xi_i^*, i = 1, \cdots, l \tag{3.16}$$

$$\xi_i^* \geq 0, i = 1, \cdots, l \tag{3.17}$$

Where (*) is a shorthand implying both the vector with and without asterisks. $\xi^* = (\xi_1, \xi_1^*, \cdots, \xi_l, \xi_l^*)^T$ is slack variable and $C > 0$ is penalty parameter. For dataset in this dissertation, the calculated value of parameters for W and b is 4.6665 and 0.0646 respectively [70].

SVR technique relies on kernel functions to construct the model. The commonly used kernel functions are:

- Linear $\qquad K(x_i, y_i) = x_i^T y_i \qquad x_i,\, y_i \in R^d$
- Polynomial $\qquad K(x_i, y_i) = \left(x_i^T y_i + r\right)^n \qquad x_i,\, y_i \in R^d$
- Laplacian kernel $\qquad K(x, y) = e^{-\alpha\|x_i - y_i\|} \qquad x_i,\, y_i \in R^d$
- Radial Basis Function $\qquad K(x_i, y_i) = exp\{-\gamma||x_i - y_i||^2\} \;\; \gamma > 0 \; x_i,\, y_i \in R^d$ (3.19)

Here, γ, r and d are kernel parameters. The kernel parameters should be carefully chosen as it implicitly defines the structure of the high dimensional feature space and thus controls the complexity of the final solution [75].

3.3.2 Simulation of SVR

Support Vector Regression (SVR) works on similar principles as Support Vector Machine (SVM) classification. Also, the SVR is the adapted form of SVM when the dependent variable is numerical rather than categorical. A major benefit of using SVR is that it is a non-parametric technique, the output model from SVR does not depend on distributions of the underlying dependent and independent variables. Instead the SVR technique depends on kernel functions. Another advantage of SVR is that it permits for construction of a non-linear model without changing the explanatory variables, helping in better interpretation of the resultant model. The basic idea behind SVR is not to care about the prediction as long as the error (ϵ_i) is less than certain value. This is known as the principle of maximal margin. This idea of maximal margin allows viewing SVR as a convex optimization problem. Now let us fit SVR model on our sample data. R package e1071 is required to call svm function [71].

By default *svm* function in R considers maximum allowed error (ϵ_i) to be 0.1. In order to avoid over-fitting, the *svm* function allows us to penalize the regression through cost function. The SVR technique is flexible in terms of maximum allowed error and penalty cost. This flexibility allows us to vary both these parameters to perform a sensitivity analysis in attempt to come up with a better model. Now we will perform sensitivity analysis, by training a lot of models with different allowable error and cost parameter. This process of searching for the best model is called tuning of SVR model [72].

The Figure 3.4 shows the performance of various models using color coding. Darker regions imply better accuracy. The use of this plot is to determine the possible range where we can narrow down our search to and try further tuning if required. The best model is the

one with lowest MSE. The darker the region the lower the MSE, which means better the model. In our sample data MSE is lowest at epsilon 0 and cost 128. We do not have to do this step manually, R provides us with the best model from the set of trained models.

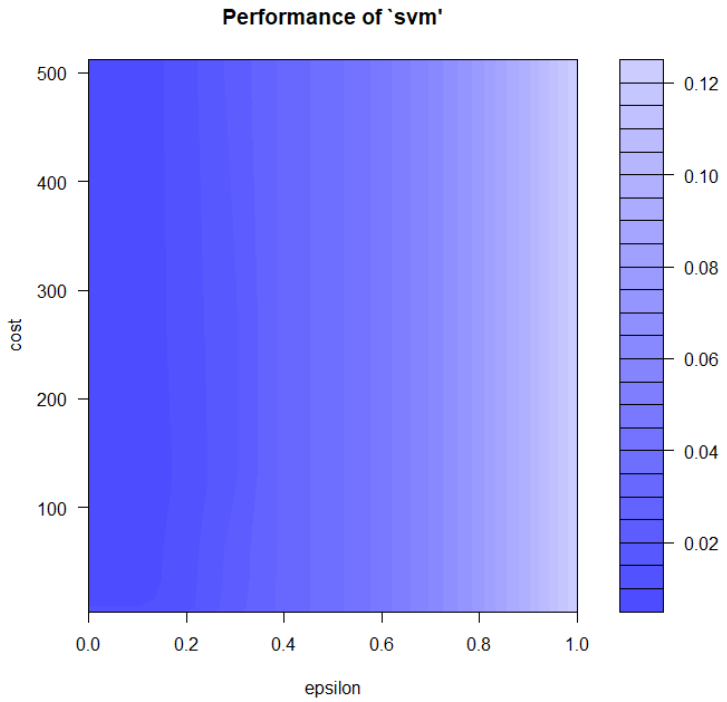


Figure 3.4: Performacne of SVR

Further, we explain the idea of tuning SVR model. Tuning of SVR model can be performed as the technique provides flexibility with respect to maximum error and penalty cost. Tuning the model is extremely important as it optimizes the parameters for best prediction. In the end we compare the performance of SVR and tuned SVR model. The RMSE of SVR is 0.133238, the rmse of tuned SVR is 0.07804601. As expected, the tuned SVR model provides the best prediction [73].

The solution to a support vector regression problem is a function that accepts a data point and returns a continuous value. The support vector regression problem also allows for a zone of insensitivity defined typically by a parameter ϵ_i. The support vector regressor must use a cost function to measure the empirical risk in order to minimize the regression

error. For financial data, due to the embedded noise, one must set a suitable margin in order to obtain a good prediction. The financial data are usually highly noisy and contain outliers. A novel two-phase SVR training algorithm has been proposed to detect outliers and reduce their negative impact. However, SVMs are very sensitive to the choice of the kernel parameters. In situations where the number of features for each object exceeds the number of training data samples, SVMs can perform poorly [74]. This can be seen intuitively as if the high-dimensional feature space is much larger than the samples. Then there are less effective support vectors on which to support the optimal linear hyperplanes, leading to poorer classification performance as new unseen samples are added.

Chapter 4

4 Conclusion

4.1 Forecasting Performance

In this dissertation, the focus is on statistical methodology and forecast macroeconomic variable on time series datasets regarding real estate scenario. The table (4.1) showed all the potential univariate forecasting models. However, the Log transformation reduced the RMSE significantly, which is almost near to 0.1.

Table 4.1

Comparison Table for univariate forecasting model

Model	**ME**	**RMSE**	**MAE**	**MPE**	**MAPE**	**AIC**
Regression	-1.103e-13	1127.403	918.1104	0.6624	21.1810	586.37193
Log Regression	-2.7e-17	0.1661	0.1489	-0.0371	1.7665	-12.84073
Regression with AR (2) errors	-5.2326	590.5238	424.728	-1.1753	7.9002	545.45
ETS (M,A,N)	-91.4374	617.2248	339.1251	-0.8840	408.4312	413.6146
ARIMA (1,1,0)	88.7736	713.7366	477.83	2.6121	7.973	532.83
Log ARIMA (2,0,1)	0.001	0.063	0.057	0.006	0.692	-181.9942
ANN(2,2)	-1.209355e-05	0.0533	0.0454	-0.00464	0.5411	-195.3637

We mainly evaluate forecasting models based on the two performance measures of RMSE and AIC for univariate forecasting model. As was the case with the forecast in Table (4.1), land market value is projected to continue increase in the following years. It shows the stable increase in the future. This number will significantly rise to 7.3 in 2018.

The author measured forecasting performance by RMSE and AIC, as the authors did, and we will also compare performance in terms of MAE, AICC and so on. The MAE criterion is most appropriate when the cost of a forecast error rises proportionally with respect to the absolute size of the error. With RMSE, the cost of the error rises as the square of the error, and so large errors can be weighted far more than proportionally. Whether MAE or RMSE is

most appropriate surely varies according to circumstances and individual institutions, and in any case we will find that the two measures pick the same model in all but several instances.

The advantage of decomposition is that decomposition models do not involve a lot of mathematics or statistics; they are relatively easy to explain to the end user. This is a major advantage because if the end user has an appreciation of how the forecast was developed, he or she may have more confidence in its use for decision making. The disadvantage of decomposition methods is that the hypothesis may be too strong for the epidemic behavior, so that the model may not perform well sometimes. The ARIMA model has obvious advantages in its well-known statistical properties and effective modeling process. It can be easily realized through mainstream statistical software. The model can be used when the seasonal time series are stationary and have no missing data, but the disadvantage of the ARIMA model is that it can only extract linear relationships within the time series data. Therefore, the ARIMA model did not efficiently extract the full relationship hidden in the historical data.

These time series come from different location and have different statistical characteristics. These data utilized to forecast through an application aimed to handle real life time series. From the experimental results comparing the performance of several models, it indicates that ANN significantly outperform other models in term of RMSE and MAE.

In addition, while RMSE and AIC are good relative measures, both depend on the scale of the forecast variable. Moreover, each could hypothetically be quite low, and still contain systematic bias and do a poor job of forecasting average value changes. Measuring the forecasting accuracy is an efficient way to select a better model for prediction. These measures were calculated by using the folllowing Equations. P_t is the predicted value at time t, Z_t is the observed value at time t and N is the number of predictions.

$$ME = \frac{\sum_{i=1}^{N}(P_t - Z_t)}{N} \tag{4.1}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|P_t - Z_t| \tag{4.2}$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}|\frac{P_t - Z_t}{Z_t}| \tag{4.3}$$

$$MPE = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{Z_t - P_t}{Z_t}\right) \times 100\% \tag{4.4}$$

$$RMSE = \sqrt{MSE} = \sqrt{\frac{\sum_{i=1}^{N}(P_t - Z_t)^2}{N}} \tag{4.5}$$

$$SBC = N\ln(MSE) + p\ln(N). \tag{4.6}$$

A given forecasting model may have a systematic positive or negative bias and do a poor job of tracking the actual mean of value changes, and measures such as RMSE and MAE could well miss this defect. We will thus evaluate forecasts based on the four performance measures of RMSE, AIC and MAE.

Table 4.2

Comparison Table for multivariate forecasting model

Model	**VAR (1)**	**BVAR (2)**	**VECM (2)**	**AR(2) -GARCH(1,1)**	**Tune SVR**
RMSE	0.08125	0.00797	0.0059	0.139570	0.0781
AICC	-860.1579	-4.5159	-12.8474	-39.04328	-169.0845
AIC	-860.304	-4.52431	-12.8951	-39.04328	-169.4716
SBC	-805.7328	-4.43271	-12.6122	-32.99622	-166.419

For comparison purpose, the training and the forecast performance of SVM were compared with the ANN model. Table 4.1 shows the performances of the ANN and other linear models during the forecasting periods 34 years. Based on the table 4.2, VECM(2) and BVAR(2) forecasts are closely to actual values. We compare the multivariate forecasting

model on RMSE and AIC criterions in terms of SBC and AICC. It show that the both approaches work well for the data set used.

The limitations of the study should also be acknowledged. The relatively short length of the series may influence the forecasting efficacy of the different methods. Second, we only predicted the land market with the six typical forecasting methods. Each forecasting model has the advantages and disadvantages. What is more, there are some other hypotheses on the long term trend in decomposition methods, such as generalized models which assume a nonlinear function among the time series. Many other models were developed to make up deficiencies of listed models. In this paper, we choose six very typical existing time series methods to make a comparison.

4.2 Outlook

Land market value, both directly and indirectly, is related to the housing market, commercial and residential buildings, construction industry, job-hunting market and home price. Therefore, the improved forecasting promise important benefits for any parties exposed to housing market [38]. The forecasting in real estate market is more important and necessary for the economy of American, because the tendency of Land Market Value would be helpful for government and investor to examine the problem in housing market, make the appropriate policy and regulate the housing market. Thus, a given forecasting model did a good job of tracking the actual value of land market changes. On the other hand, forecasting techniques are widely used in the area of finance and housing market. As rapidly rising housing prices are the hot topic in the growing number of metropolitans around the world. Most importantly, forecasting modeling is ever more significant in predicting the direction of future tendency.

The table (4.3) shows the VECM (2) forecasting model for the general tendency of the land market value for about 10 years. The forecasting results shows the stable increase in the future. This number will rise to 9.12 in 2022.

Table 4.3

Forecasting VECM (2) Model for LLMV

Year	**Forecasting**	**Standard Error**	**95% L B**	**95% U B**
2013	8.61155	0.07584	8.46290	8.76019
2014	8.66936	0.15104	8.37334	8.96539
2015	8.76250	0.22786	8.31590	9.20910
2016	8.86550	0.30301	8.27161	9.45939
2017	8.96023	0.37445	8.22633	9.69413
2018	9.03590	0.44069	8.17217	9.89963
2019	9.08805	0.50070	8.10670	10.06941
2020	9.11704	0.55394	8.03134	10.20275
2021	9.12638	0.60040	7.94961	10.30315
2022	9.12117	0.64054	7.86573	10.37661

Time series analysis and forecasting is an active research area over the last five decades. Thus, various kinds of forecasting models are developed and researchers have relied on statistical techniques to predict time series data. The successful forecasting usually depends on the proper selection of model order, parameters, the number of inputs and neurons in the forecasting model. The accuracy of time series forecasting is fundamental to many decisions processes and hence the research for improving the effectiveness of forecasting models has never been stopped. However, the time series datasets are often nonlinear and irregular.

A variety of forecasting methods often apply to any particular risk scenario. Business and investor use multiple forecasting methods tuned to perform well at different phases of housing market, chosen to best exploit the available historical data and degree of market knowledge. The key is to pick the most effective and flexible forecasting models, blend their best features, and shift between them as needed to keep forecast accuracy at its peak. This research paper delves into the details of ten forecasting methods, including why, when and how they should be used to realize the greatest overall improvements in forecast accuracy.

Over the last 35 years there has been considerable information accumulated about forecasting techniques and how these methods are applied in a wide variety of settings. Conflict-

ing results are very common when performing advanced forecasting competitions between different methods. As forecasting tasks can vary by many dimensions, it is unlikely that one method will be better than all others for all forecasting scenarios. What we require from a forecasting method are consistently sensible forecasts, and these should be frequently evaluated against the task at hand.

REFERENCES

REFERENCES

[1] Davis, M. A and Jonathan H, (2007)
The Price and Quantity of Residential Land in the United States.
The Journal of Monetary Economics 54 (8): 2595-2620.

[2] Davis, M. A and Michael G. P, (2008)
The price of Residential Land in Large US Cities,
Journal of Urban Economics, 63 (1): 352-84.

[3] Ghulam. A , (2015)
Cointegration VAR and VECM and ARIMAX Econometric Approaches for Water Quality Variates,
Journal of Statistical and Econometric Methods, 4 (1): 1-38.

[4] RATNADIP A, (2013)
An Introductory Study on Time Series Modeling and Forecasting ,
LAP Lambert Academic Publishing, ISBN 10: 3659335088.

[5] Wennstrm. A, (2014)
Volatility Forecasting Performance: Evaluation of GARCH type volatility models on Nordic equity indices. Second Edition,
Master of Science Thesis, Department of Mathematics, Royal Institute of Technology (KTH), Stockholm, Sweden.

[6] Barot. B and Takala. K, (1998)
House prices and inflation : a cointegration analysis for Finland and Sweden,
Research Discussion Papers, Bank of Finland.

[7] Debasish B, Srimanta P and Dipak Chandra P, (2007)
Support Vector Regression,
Neural Information Processing Letters and Reviews 11 (10): 203-224.

[8] Litterman R B, (1980)
A Bayesian Procedure for Forecasting with Vector Autoregression,
Ph.D Dissertation., Department of Economics.

[9] Litterman R B, (1984a)
Forecasting and Policy Analysis with Bayesian Vector Autoregression Models,
Federal Reserve Bank of Minneapolis Quarterly Review, 8 (4): 30-41.

[10] Litterman R B, (1986)
Forecasting with Bayesian Vector Autoregressions-Five Years of Experiencem,
Journal of Business and Economic Statistics, American Statistical Association, 4 (1): 25-38.

[11] Tim B, (2006)
A Conditionally Heteroskedastic Time Series Model For Speculative Prices and Rates of Return,
The review of Economics and Statistics, 69 (3): 542-547.

[12] Montgomery, D. C., Jennings, C. L., and Kulahci, M., and Wiley, J., and Sons., (2015)
Introduction to Time Series Analysis and Forecasting,
Springer: Second Edition.

[13] Javier E. C and Wilfredo.P., (2013)
Statistical analysis of autoregressive fractionally integrated moving average models in R,
Computational Statistics 28 (5): 23092331.

[14] Gruben W.C. and Long III, W.T., (1988a)
Forecasting the Texas economy: Application and evaluation of a systematic multivariate time-series model outlook for 1989,
Review, Federal Reserve Bank of Dallas.

[15] Sims C., (1980)
Microeconomics and Reality,
Econometrica, 48 (1): 1-48.

[16] Sims C., (2007)
Bayesian Methods in Applied Econometrics, or, Why Econometrics Should Always and Everywhere Be Bayesian,
Hotelling lecture, Duke University.

[17] Matteo.C., and Alessandro.R, (2003)
Bayesian VARS: A Survey of the Recent literature with an Application to the European Monetary System,
International Monetary Fund.

[18] Petre. C, (2010)
Forecasting Romanian GDP using a BVAR model,
Journal for Economic Forecasting, 13 (4): 76-87.

[19] Robertson.J.C and Tallman E.W, (First Quarter 1999)
Vector Autoregressions: Forecasting and Reality,
Federal Resene Bank of Atlanta & Economic Review. 84 (1): 4-18.

[20] Kamal.R.D, (2014)
Estimation of Short and Long Run Equilibrium Coefficients in Error Correction Model: An Empirical Evidence from Nepal,
International Journal of Econometrics and Financial Management, 2 (6): 214-219.

[21] Schorfheide F., (2008)
Bayesian Methods in Macroeconometrics,
The New Palgrave Dictionary of Economics, 402-406.

[22] Box G. and Jenkins G, (1970)
Time Series Analysis: Forecasting and Control,
Holden-Day, San Francisco.

[23] Anthony Y G, (2002)
The Predictability of House Prices,
Journal of Real Estate Research, 24 (3): 213-233.

[24] Richard, H, (1995)
Using Co-integration Analysis in Econometric Modelling,
Harvester Wheatsheaf, Prentice Hall, ISBN 0-13-355892-4.

[25] Faraway, J. J., (2014)
Linear Models with R,
CRC Press, Taylor adn Francis Group.

[26] Muellbauer.J and Anthony M, (1997)
Booms and Busts in the UK Housing Market,
The Economic Journal, 107 (445): 1701-1727.

[27] Faming L, (2005)
Bayesian neural networks for nonlinear time series forecasting,
Statistics and Computing, 15 (1): 1329.

[28] William, M, (2008)
Boom-bust Cycles and the Forecasting Performance of Linear and Non-linear Models of House Prices,
Journal of Real Estate Finance and Economics, 36 (3): 249-264.

[29] Brockwell P.J and Davis R.A, (2002)
Introduction to Time Series and Forecasting, Second Edition,
Springer, New York.

[30] Engle. R, (2001)
An Introduction to the Use of ARCH/GARCH models in Applied Econometrics,
New York University No. FIN-01-030.

[31] Crawford G W and Fratantoni M.C, (2003)
Assessing the Forecasting Performance of Regime-Switching, ARIMA and GARCH Models of House Prices,
Real Estate Economics, 31 (2): 223-243.

[32] William, W, (2006)
Time Series Analysis:Univariate and Multivariate Methods. Second Edition,
Pearson Addison Wesley.

[33] Bostic, R. W., Stanley D. L and Christian L.R, (2007)
Land Leverage: Decomposing Home Price Dynamics,
Real Estate Economics, 35 (2): 183-208.

[34] Richard K.G and Stephen M (2008).
A primer on U.S. Housing Markets and Housing Policy,
The Urban Institute Press, ISBN 0-87766-702-0.

[35] Joseph B. N, Stephen D. O and Michael R. M, (2010)
Commercial and Residential Land Price Across the United States,
Federal Reserve Board.

[36] Choon-Shan L and Roy. A, (2004)
Accuracy of Bayesian VAR in forecasting the economy of Indiana,
Proceedings of the Midwest Business Economics Association, 67-71.

[37] Hyndman, R. J and Athanasopoulos G, (2014)
Forecasting: Principles and Practice,
OTexts, 2nd edition, ISBN 978-0-9875071-0-5.

[38] James D.H, (1994)
State-space models,
Elsevier Science. University of California, San Diego.

[39] Xu chen J, Man jin C, Yu hong X, Fu qiang Q, Ying jie T, (2014)
Support Vector Regression and Time series Analysis for the forecasting of Bayannur's Total Water Requirement,
Procedia Computer Science, 31: 523-531.

[40] Sirmans, C.F and Barrett A. S, (2009)
National Transactional-based Land Price Indices,
Journal of Real Estate Finance and Economics, 45 (4): 829-845.

[41] Zou H.F., Xia G.P., Yang F.T. and Wang H.Y, (2007)
An investigation and comparison of artificial neural network and time series models for chinese food grain price forecasting,
Neurocomputing, 70 (16-18): 2913-2923.

[42] William M, (2008)
Bubbles, Busts, Breaks and Segmentation,
The Journal of Housing Research, 23 (1): 57-72.

[43] William M, (2008)
Bubbles, Busts and Breaks in UK Housing,
International Real Estate Review, 18 (4): 455-471.

[44] Dennis D. W., William M., and Richard L.S, (2008)
Mathematical Statistics with applications, 7th Edition,
Thomson Higher Education.

[45] Gupta, R. and Sichei, M. M, (2006)
A BVAR model for the South African economy,
South African Journal of Economics, 74 (3): 391-409.

[46] James H.S., and Mark W. W, (2012)
Dynamic Factor Models,
Oxford Handbook of Economic Forecasting.

[47] Ricky H, (2012)
Predictive Analytics: NeuralNet, Bayesian, SVM, KNN,
Pragmatic Programming Techniques.

[48] Kinal T. and Ratner J, (1986)
A VAR Forecasting Model of a Regional Economy: Its Construction and Comparative Accuracy,
International Regional Science Review, 10 (2): 113-126.

[49] Robert E and Lucas Jr, (2008)
Econometric poeicy evaluation: A CRITIQUE,
Public Policy, 1 (1): 19-46.

[50] Lesage, J. P, (1990)
A comparison of the forecasting ability of ECM and VAR Models,
The Review of Economics and Statistics, 72 (4): 664-671.

[51] Peter P, (1986)
Understanding Spurious Regression in Econometrics,
The Journal of Econometrics, 33 (3): 311-340.

[52] Case K and Shiller R, (1989)
The Efficiency of the Market for Single-Family Homes,
American Economic Review, 79 (1): 125-137.

[53] Case K and Shiller R, (2003)
Is There a Bubble in the Housing Market?,
Brookings Papers on Economic Activity, 2: 299-342.

[54] Gillen K., Thibodeau T. and Wachter S, (2001)
Anisotropic autocorrelation in house prices,
Journal of Real Estate Finance and Economics, 23 (1): 5-30.

[55] Lippmann R, (1987)
An introduction to computing with neural nets,
IEEE ASSP Mag., 4 (2): 4-22.

[56] Samsudin R, Shabri A and Saad P, (2010)
A Comparison of Time Series Forecasting using Support Vector Machine and Artificial Neural Network Model,
Journal of Applied Sciences, 10 (11): 950-958.

[57] Engle R, Lilien M. and Robins R. P, (1987)
Estimating Time Varying Risk Premia in the Term Structure: The Arch-M Model,
Econometrica 55 (2): 391-407.

[58] Glosten L. R., Jagannathan, R., and Runkle D. E, (1993)
On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks,
The Journal of Finance, 48 (5): 1779-1801.

[59] Bukola Titilayo O, (2006)
Support Vector Regression for Non-Stationary Time Series,
University of Tennessee - Knoxville, Master Theses.

[60] Kang. S, (1991)
An investigation of the use of feedforward neural network for forecasting,
Ph.D. Thesis, Kent State University.

[61] Volkan S and Gul E, (2009)
Usage of different prior distributions in Bayesian vector autoregressive models,
Journal of Mathematics and Statistics, 38 (1) : 85-93.

[62] Wong F. S, (1991)
Time series forecasting using backpropagation neural networks,
Neurocomputing, 2 (4): 147-159.

[63] Andrew M, (2001)
Tutorial slides: Support Vector Machines,
http://www.cs.cmu.edu/awm.

[64] SAS institute Inc.
SAS/ETS14.1 Users Guide The VARMAX Procedure,
SAS Institute Inc., SAS Campus Drive, Cary, NC.

[65] Mark, S, (1980)
Bayesian Time Series Analysis,
The New Palgrave Dictionary of Economics, 2nd edition, 1-9.

[66] Chaitanya S, (2017)
Tutorial Support Vector Machines in R,
https://www.r-bloggers.com/machine-learning-using-support-vector-machines/.

[67] James L, (2018)
Tutorial Support Vector Machines in R,
https://www.datacamp.com/community/tutorials/support-vector-machines-r.

[68] Chaitanya S, (2017)
Building Regression Models in R using Support Vector Regression,
https://www.kdnuggets.com/2017/03/building-regression-models-support-vector-regression.html.

[69] Nichols N.V and Pope J.C, (2013)
The value of residential land and structures during the great housing boom and bust,
Land Economics, 89 (1): 1-29.

[70] Calomiris C W, Longhofer S.D and Miles W.R, (2013)
The ForeclosureHouse Price Nexus: A Panel VAR Model for U.S. States, 1981-2009,
Real Estate Economics, 41 (4): 709-746.

[71] Crawford G W and Fratantoni M C, (2003)
Assessing the Forecasting Performance of Regime-Switching, ARIMA and GARCH Models of House Prices,
Real Estate Economics, 31 (2): 223-243.

[72] Wang L, (2017)
Prediction on Land Market Value Based on the Real Estate Market in USA,
Journal of Mathematics and Statistics, 13 (2): 143-151.

[73] Wang L, (2018)
A VAR Model for Forecasting Land Market Value in USA ,
Journal of Mathematics and Statistics, 14 (1): 1-6.

[74] Wang L, (2018)
Advanced Multivariate Time Series Forecasting Models,
Journal of Mathematics and Statistics, 14 (1): 253-260.

[75] Zhong Guo. Z, (1997)
Forecasting Sales and Price for Existing Single-Family Homes: A VAR Model with Error Correction,
Journal of Real Estate Research, 14 (2): 155-168.

[76] Zhang, G.P., Patuwo G.E. and Hu.M.Y, (2001)
A simulation study of artificial neural networks for nonlinear time-series forecasting,
Computers and Operations Research, 28 (4): 381-396.

APPENDIX

```
panel.hist <- function(x,...){
usr <- par("usr")
on.exit(par(usr))
par(usr=c(usr[1:2],0,1))
h $<-$ hist(x,plot=FALSE,breaks="FD")
breaks <- h$breaks
nB <- length(breaks)
y <- h$counts
y <- y/max(y)
rect(breaks[-nB],0,breaks[-1],y,col="cyan",...)}
pairs(LMV.forecasting2[,(2:9)],diag.panel=panel.hist)

fit <- Arima(LMV.forecasting3[,1], xreg=LMV.forecasting3[,2:8],
order=c(2,0,0))
tsdisplay(arima.errors(fit), main="ARIMA errors")

TSLMV1 <- window(TSLMV,start=1982,end=2008)
fit1 <- ses(TSLMV1)
fit2 <- holt(TSLMV1)
fit3 <- holt(TSLMV1,exponential=TRUE)
fit4 <- holt(TSLMV1,damped=TRUE)
fit5 <- holt(TSLMV1,exponential=TRUE,damped=TRUE)
plot(fit2$model$state)
plot(fit4$model$state)
plot(fit2$model$state)
flwd=1, plot.conf=FALSE)
lines(window(TSLMV,start=2015),type="o")
lines(fit1$mean,col=2)
```

```
lines(fit2$mean, col=3)
lines(fit4$mean, col=5)
lines(fit5$mean, col=6)
legend("topleft", lty=1, pch=1, col=1:6,
c("Data","SES","Holt's","Exponential",
"Additive Damped","Multiplicative Damped"),cex=0.75)

fit0 <- ets(TSLMV1)
summary(fit0)
plot(forecast(fit0,h=8),
ylab="Lank Market Value (millions)")
fit1 <- lm(LMV ~ IR+UR+CCI+PMI, data=LMV.forecasting3)
summary(fit1)
accuracy(fit1)

fit2 <- lm(LLMV ~ LIR+LUR+LCCI+LPMI, data=LMV.forecasting3)
summary(fit2)
accuracy(fit2)
Box.test(residuals(fit), fitdf=5,lag=10,type="Ljung")

TSLMV <- ts(LMV, start=1982,frequency=1)
LTSLMV <- log(TSLMV)
par(mfrow=c(1,2))
plot(LTSLMV, ylab="Log transformation land market value", xlab="Year")
plot(TSLMV, ylab="land market value", xlab="Year")

LTSLMV1 <- window(LTSLMV, start=1982,end=2008)
TSM1 <- arima(TSLMV, order=c(1,1,0))
```

```
Acf(residuals(TSM1))
summary(TSM1)
accuracy(TSM1)
forecast(TSM1)
plot(forecast(TSM1))
TSM7 <- arima(LTSLMV1, order=c(2,0,1))
Acf(residuals(TSM7))
summary(TSM7)
accuracy(TSM7)
forecast(TSM7)
plot(forecast(TSM7))

 VARselect(mydata7, lag.max=8, type="const")$selection
AIC(n) HQ(n) SC(n) FPE(n)
4 4 4 3
 var <- VAR(mydata7, p=2, type="const")
 serial.test(var, lags.pt=10, type="PT.asymptotic")
Portmanteau Test (asymptotic)
data: Residuals of VAR object var
Chi-squared = 333.59, df = 288, p-value = 0.03325
 var <- VAR(mydata7, p=1, type="const")
 serial.test(var, lags.pt=10, type="PT.asymptotic")
 pre <- predict(var)
 plot(pre)
 print(pre)
```

```
data LLMV;
input x@@;
t=_n_;
cards;
7.150607 7.116597 7.235014 7.343717 7.538543 7.739416 7.893120
8.038370 8.088755 8.023008 8.035862 7.988960 8.004953 7.987881
8.017594 8.045787 8.142360 8.265179 8.413873 8.599398 8.719822
8.883061 9.064873 9.278833 9.437262 9.416564 9.255757 8.927688
8.878195 8.729766 8.620392 8.821296 9.006019 9.075335;
;
proc gplot data=LLMV;
plot x*t=1;
symbol1 c=black i=join v=star;
proc autoreg data=LLMV;
model x=t/nlag=5 dwprob archtest;
model x=t/nlag=1 garch=(p=1,q=2);
output out=out p=xp;
proc gplot data=out;
plot x*t=2 xp*t=3/overlay;
symbol2 v=star i=none c=black;
symbol3 v=none i=join c=red w=2;
run;

data llmvlcci;
input x y@@;
t=_n_;
cards;
7.150607 3.770459 7.116597 3.799974 7.235014 3.843744 7.343717
```

```
3.869116 7.538543 3.919991 7.739416 3.964615 7.893120 3.998201
8.038370 4.032469 8.088755 4.060443 8.023008 4.063885 8.035862
4.075841 7.988960 4.123903 8.004953 4.168214 7.987881 4.209160
8.017594 4.228293 8.045787 4.257030 8.142360 4.283587 8.265179
4.286341 8.413873 4.329417 8.599398 4.37827 8.719822 4.403054
8.883061 4.453184 9.064873 4.533674 9.278833 4.60517 9.437262
4.663439 9.416564 4.672829 9.255757 4.637637 8.927688 4.585987
8.878195 4.568506 8.729766 4.578826 8.620392 4.589041 8.821296
4.652054 9.006019 4.716712 9.075335 4.608863
;
proc gplot;
plot x*t=1 y*t=2/overlay;
symbol1 c=black i=join v=none;
symbol2 c=red i=join v=none w=2 l=2;
run;
proc arima data=llmvlcci;
identify var=x stationarity= (adf=1);
identify var=y stationarity= (adf=1);
run;

proc arima;
identify var= y crosscorr=x;
estimate method=ml input=x noint plot;
run;

forecast lead= 0 id=t out=out;
proc arima data=out;
```

```
identify var= residual stationarity= (adf=2);
run;

proc varmax data=llmvlcci;
   model x y / p=2 noint lagmax=3
                ecm=(rank=1 normalize=x)
                print=(iarr estimates);
run;

proc varmax data=llmvlcci;
   model x y / p=2 noint dftest cointtest=(johansen);
run;

proc varmax data=llmvlcci;
      model x y / p=1 print=(parcoef pcorr pcancorr) lagmax=6 ;
   run ;
proc varmax data=llmvlcci;
   model x y / p=1 noint lagmax=3
                print=(estimates diagnose);
   output out=for lead=5;
run;

proc varmax data=llmvlcci;
   model x  / p=2 noint
      prior=(lambda=1 theta=0.9 ivar);
           output out=for lead=10;
run;
```

```
 library(fpp2)
library(e1071)
 fit <- nnetar(TLLMV, lambda=0)
 accuracy(fit)
 sim <- ts(matrix(0, nrow=20, ncol=9), start=end(TLLMV)[1]+1)
   for(i in seq(9))
   sim[,i] <- simulate(fit, nsim=20)
 fcast <- forecast(fit, PI=TRUE, h=20)
 autoplot(fcast)
 autoplot(TLLMV) + autolayer(sim)

 svm <- svm(mydata7$LLMV~LCCI, mydata7)
 predYsvm <- predict(svm, mydata7)
 points(mydata7$LLMV,predYsvm,col="red",pch=16)
W <- t(svm$coefs)%*%svm$SV
  W
          [,1]
[1,] 4.666511
 b<- modelsvm$rho
> b<- svm$rho
>  b
[1] 0.06458899
RMSEsvm <- rmse(predYsvm,LCCI)
```

```
library(e1071)
model <- svm(LLMV ~ LCCI, mydata7)
 predictedY <- predict(model,mydata7)
 points(LCCI,predictedY,col="red",pch=4)
 error <- mydata7$LLMV-predictedY
 svrPredictionRMSE <- rmse(error)
 rmse <- function(error){
 sqrt(mean(error^2))
 }
 error<- model$residuals
 error <- mydata7$LLMV-predictedY
 svrPredictionRMSE <- rmse(error)
  svrPredictionME <- mean(error)
> print(svrPredictionME)
[1] -0.00425145
> svrPredictionMAE <- mean(abs(error))
> print(svrPredictionMAE)
[1] 0.102405
 print(svrPredictionRMSE)
[1] 0.133238
tuneResult <- tune(svm, mydata7$LLMV~LCCI, data=mydata7,
ranges=list(epsilon=seq(0,1,0.1),cost=2^(2:9)))
 print(tuneResult)
 plot(tuneResult)
 summary(tuneResult)
Parameter tuning of   svm   :
- sampling method: 10-fold cross validation
```

```
- best parameters:
 epsilon cost
       0  128

- best performance: 0.006416961

 tuneModel <- tuneResult$best.model
 tuneModelY <- predict(tuneModel,mydata7)
 print(tuneModelY)
 error <- mydata7$LLMV-tuneModelY
 tuneModelRMSE <- rmse(error)
 print(tuneModelRMSE)
[1] 0.07804601

 tune <- tune.svm(LLMV~LCCI, data=mydata9)
 tune <- tune.svm(LLMV~LCCI, data=mydata9,
 gamma=10^(-6:-1),cost=10^(1:4))
 summary(tune)

Parameter tuning of  svm :
- sampling method: 10-fold cross validation
- best parameters:
 gamma cost
   0.1  100
 best performance: 0.02570149
```

```
model <- svm(LLMV~LCCI, data=mydata9, method="C-classification",
kernal="radical", probability=T, gamma=0.1, cost=100)
prediction <- predict(model, mydata9)
 print(prediction)
  error2 <- mydata9$LLMV-prediction
  svrPredictionRMSE2 <- rmse(error2)
 print(svrPredictionRMSE2)
[1] 0.140136

 svrPredictionME2 <- mean(error)
> print(svrPredictionME2)
[1] -0.00425145
> svrPredictionMAE2 <- mean(abs(error2))
> print(svrPredictionMAE2)
[1] 0.1184444

  W <- t(model$coefs)%*%model$SV
 print(W)
 plot(prediction)
          [,1]
[1,] 24.36648
 b<- model$rho
 print(b)
[1] 0.5752832
```

Appendix[a]: 1982-2015 Land Market Value Datasets

Year	LMV [b]	CPI	GDP [c]	IR	UR	CCI	PP	PMI	Year	LMV	CPI	GDP [d]	IR	UR	CCI	PP	PMI
1982	1274.88	96.5	6.49	6.2	9.7	43.4	231.66	42.8	2000	4509.19	172.2	12.68	3.4	4	75.9	282.16	43.9
1983	1232.25	99.6	7	3.2	9.6	44.70	233.79	69.9	2001	5428.39	177.1	12.71	2.8	4.7	79.7	284.97	45.3
1984	1387.16	103.9	7.4	4.3	7.5	46.7	235.82	50.6	2002	6123.09	179.9	12.96	1.6	5.8	81.7	287.63	51.6
1985	1546.45	107.6	7.71	3.6	7.2	47.9	237.92	50.7	2003	7208.82	184.13	13.53	2.3	6	85.9	290.11	60.1
1986	1879.09	109.6	7.94	1.9	7	50.4	240.13	50.5	2004	8646.18	188.9	13.95	2.7	5.5	93.1	292.81	57.2
1987	2297.13	113.6	8.29	3.6	6.2	52.7	242.29	61	2005	10708.93	195.3	14.37	3.4	5.1	100	295.52	55.1
1988	2678.79	118.3	8.61	4.1	5.5	54.5	244.50	56	2006	12547.31	201.6	14.72	3.2	4.6	106	298.38	51.4
1989	3097.56	124.8	8.85	4.8	5.3	56.4	246.82	47.4	2007	12290.28	207.3	14.99	2.8	4.6	107	301.23	49
1990	3257.63	130.7	8.91	5.4	5.6	58	249.62	40.8	2008	10464.64	215.3	14.58	3.82	5.8	103.3	304.09	33.1
1991	3050.34	136.2	9.02	4.2	6.8	58.2	252.98	46.8	2009	7537.82	214.5	14.54	-0.32	9.3	98.10	306.77	55.3
1992	3089.8	140.3	9.41	3	7.5	58.9	256.51	54.2	2010	7173.83	218.1	14.94	1.64	9.6	96.4	309.35	57.5
1993	2948.23	114.5	9.65	3	6.9	61.8	259.92	55.6	2011	6184.28	224.9	15.19	3.14	8.9	97.4	311.72	53.1
1994	2995.76	148.2	10.05	2.6	6.1	64.6	263.13	56.1	2012	5543.56	229.6	15.43	2.08	8.1	98.4	314.11	50.4
1995	2945.05	152.4	10.28	2.8	5.6	67.3	266.28	46.2	2013	6777.04	233	15.92	1.46	7.4	104.8	316.5	56.5
1996	3033.87	156.9	10.74	3	5.4	68.6	269.39	55.2	2014	8152	237.2	16.29	1.61	6.2	111.8	318.86	55.1
1997	3120.62	160.5	11.21	2.3	4.9	70.6	272.65	54.5	2015	8737.11	242.1	16.3	0.1	5.5	100.37	320.99	53.5
1998	3437.02	163.11	11.77	1.6	4.5	72.5	275.85	46.8									
1999	3886.17	166.6	12.32	2.2	4.2	72.7	279.04	57.8									

[a]The data was based on the 34 years' national data on past and present real estate transaction from 1982 to 2015.

[b]The unit of land market value is million

[c]The unit of GDP is trillion.

[d]http://www.statista.com/statistics/188105/annual-gdp-of-the-united-states-since-1990/

Source: U.S. Bureau of Labor Statistics https://en.wikipedia.org/wiki/Main-Page.