

**MODELING AND FITTING OF THE WIND DATA
USING DIFFERENT TIME SERIES MODELS AND
INVESTIGATING THE RELATED APPLICATIONS
OF FITTED DATA. URLA AND RISØ CASES.**

**A Thesis Submitted to
the Graduate School of Engineering and Sciences of
İzmir Institute of Technology
in Partial Fulfillment of the Requirements for the Degree of
MASTER OF SCIENCE
in Energy Engineering**

**by
Nurseda YILDIRIM**

**December 2014
İZMİR**

We approve the thesis of **Nursedda YILDIRIM**

Examining Committee Members:

Assist. Prof. Dr. Hasan Engin DURAN
Department of City and Regional Planning, Izmir Institute of Technology

Assoc. Prof. Dr. Evrim TURGUTLU
Department of Economics, Faculty of Business, Dokuz Eylul University

Prof.Dr.Aydođan ÖZDAMAR
Department of Mechanical Engineering, Ege University

16 December 2014

Assist. Prof. Dr. Hasan Engin DURAN
Supervisor
Department of City and Regional Planning
Izmir Institute of Technology

Assist. Prof. Dr. Ferhat BİNGÖL
Co-Supervisor
Department of Mechanical Engineering
Izmir Institute of Technology

Prof. Dr. Gülden GÖKÇEN AKKURT
Head of the Department of
Energy Engineering

Prof. Dr. Bilge KARAÇALI
Dean of the Graduate School of
Engineering and Sciences

ACKNOWLEDGMENTS

I would like to thank my supervisor Hasan Engin Duran for his guidance and his support. I am indebted to him for the opportunity to study on wind energy and time series. I would like to thank also my co-supervisor Ferhat Bingöl for his support and guidance. I am grateful to EGENDA company for sharing very kindly their dataset with me. I would like to thank my examination committee for their constructive feedback on my thesis. I am indebted to my former lecturer, Turhan Çoban who taught me open source programming. I am grateful for the supports of the head of energy engineering department Gülden Gökçen Akkurt during my master studies.

I appreciate the Nordtank data which is given to me by Kurt Hansen. I want to also thank to head of ADM meteorological station, he kindly shared with me ADM data.

Many thanks also go to my friends and colleagues and the faculty and staff for converting my time at IZTECH into a great experience.

ABSTRACT

MODELING AND FITTING OF THE WIND DATA USING DIFFERENT TIME SERIES MODELS AND INVESTIGATING THE RELATED APPLICATIONS OF FITTED DATA. URLA AND RISØ CASES.

This thesis is prepared as an outcome of Energy Engineering Master of Science program at IZTECH. Main purpose of this study is to investigate the possible ways of estimating the evolution of wind speed in Turkey, which is useful in predicting the wind power generation.

Wind Energy has recently been recognized as one of the most promising renewable energy sources in the world. Despite its high potential, one major problem is that it is an intermittent energy source which follows, in general, statistically a quite noisy evolution with large variability and difficulty in forecasting. Standard time series models have been employed to forecast the wind speed in the literature (such as ARIMA, ARMA). The majority of these, however, are based on a univariate modelling. This is likely to create a significant loss in forecast accuracy as the important dynamics of wind such as ambient temperature, absolute pressure, wind direction and humidity are ignored.

So, aim of the present study is to incorporate these factors in a multivariate VAR setting and estimate the wind speed in 4 different locations around Urla City (nearby Izmir-Turkey) by employing hourly data between June-2000 and October-2001. To provide a benchmark, I also compare estimations from VAR with the predictions from ARIMA and SARIMA models. The results indicate two important conclusions. First, it has been shown that all models provide an accurate estimate of wind speed. Second, multivariate VAR and SARIMA is clearly shown to outperform the ARIMA model by improving the wind speed predictions and producing less forecast errors.

Thus, these models are demonstrated to be helpful in estimating the wind power generation as well.

ÖZET

RÜZGAR VERİLERİNİN ÇEŞİTLİ ZAMAN SERİSİ YÖNTEMLERİYLE MODELLENMESİ ÜRETİLEN VERİNİN UYGULAMA ALANLARININ İNCELENMESİ. URLA VE RISØ ÖRNEKLERİ

Bu çalışma İzmir Yüksek Teknoloji Enstitüsü, Enerji Mühendisliği Yüksek Lisans programının bitirme tezidir.

Tezin ana amacı rüzgar hızını modellemede ve tahminlemede kullanılan olası istatistiksel yöntemleri araştırmak ve bu yöntemlerin sonuçlarını kıyaslayarak veri temizleme ve eksik datayı tahminleme işlemleri için yöntem önerisi yapmaktır. Halihazırda literatürde varolan ve rüzgar güç santralleri için oldukça önemli bir yere sahip olan istatistiksel yöntemler taranmıştır. Ancak farklı meteorolojik değişkenleri hesaba katabilecek çok değişkenli modellerle ilgili çalışmaların kısıtlı olduğu tespit edilmiştir.

Bu çalışmada VAR yöntemiyle rüzgar hızını tahminlerken basınç, sıcaklık, rüzgar yönü ve nem kayıtları da girdi olarak kullanılmıştır. SARIMA yöntemi mevsimsel etkileri yakalayabilmek için kullanılmış, çok değişkenli VAR ve ARIMA yöntemleri dört farklı data seti için kurulmuş ve tahminlenen değerler doğruluk ölçme araçlarıyla kıyaslanmıştır. Özetle çoklu değişkenin kullanıldığı VAR ve mevsimselliği tahminlemelerde kullanılan SARIMA modeli ARIMA modelinden daha az hata vermiştir. Dolayısıyla, rüzgar enerjisi tahmin ve modellemelerinde bir çok meteorolojik değişkeni ve mevsimselliği göz önüne almak tahmin verimliliğini arttırmaktır.

TABLE OF CONTENTS

LIST OF FIGURES	x
LIST OF TABLES	xii
LIST OF SYMBOLS	xiii
LIST OF ABBREVIATIONS	xiv
CHAPTER 1. INTRODUCTION	1
1.1. Thesis' Aim and Objectives	1
1.2. Organization of Thesis	4
CHAPTER 2. WIND ENERGY MARKETS IN WORLD AND IN TURKEY	6
CHAPTER 3. THEORITICAL BACKGROUND	13
3.1. The Wind	13
3.1.1. Atmosphere and Troposphere	13
3.1.1.1. Atmospheric Forces	14
3.1.1.2. Atmospheric Balances	16
3.1.1.3. Gradient Balance	17
3.1.1.4. Guldberg-Mohn Balance	18
3.2. The Atmospheric Boundary Layer	18
3.2.1. Wind Profile	19
3.3. Anemometer and Measurement Tools	20
CHAPTER 4. METHODOLOGY	23
4.1. Statistical Processes That Impacts ARIMA & VAR	23
4.1.1. STATIONARITY & STABILITY	23
4.1.1.1. Unit Root Test	24
4.2. Univariate Time Series Models	26
4.2.0.2. AR(p) Model Form	26

4.2.0.3.	Backshift Operators & Characteristic Polynomials	27
4.2.0.4.	Inversion and Moving Averages(MAs)	28
4.2.0.5.	ARMA Model	29
4.2.0.6.	ARIMA Model	29
4.2.1.	SARIMA Model	31
4.3.	Sims Methodology(VAR)	34
4.3.1.	Terminology, Notation and General Assumptions	34
4.3.1.1.	General Definition	34
4.3.1.2.	Matrix Notation	35
4.4.	Accuracy Measurement Tools	36
CHAPTER 5. THE WIND SPEED FORECASTING LITERATURE REVIEW		37
CHAPTER 6. DATA		42
6.1.	Location of Study: IZMIR Bay	42
6.2.	Data	42
CHAPTER 7. MODEL SPECIFICATIONS AND RESULTS		48
7.1.	Model Specifications	48
7.2.	Estimation Results	49
7.3.	Accuracy Comparision	62
7.4.	Impulse Response Function	66
7.5.	Robustness Check for model residuals and White Noise Property .	68
7.6.	Forecasted Values	72
CHAPTER 8. ENGINEERING APPLICATION		74
8.1.	Data	74
8.2.	Annual Energy Production (AEP)	75
8.2.1.	Weibull Distribution	76
8.3.	WAsP	77
8.4.	Results	78
CHAPTER 9. CONCLUSION		81

REFERENCES	83
APPENDICES	
APPENDIX A. MODEL PARAMETERS	86
APPENDIX B. NORDTANK TURBINE POWER CURVE	88
APPENDIX C. CODES	89

LIST OF FIGURES

<u>Figure</u>	<u>Page</u>
Figure 2.1. EU Installed Wind Energy in 2013	9
Figure 2.2. Turkey Energy Mix in 2011	10
Figure 2.3. Turkey Wind Energy Installed Capacity between 2008 and 2011	11
Figure 2.4. Turkey Wind Turbine Map	12
Figure 3.1. Troposphere and height change from equator to pole.	15
Figure 4.1. ARIMA 24 hours forecasts	31
Figure 4.2. SARIMA 24 hours forecasts	33
Figure 6.1. Wind Speed Map of Turkey	44
Figure 6.2. Run plot for wind direction	44
Figure 6.3. Data set locations	45
Figure 6.4. Time series plot for Wind Speed ADM	45
Figure 6.5. Time series plot for Wind Speed URLA RES	46
Figure 6.6. Time series plot for Wind Speed CFSR	46
Figure 6.7. Time series plot for Wind Speed MERRA	47
Figure 7.1. ADM CASE Run plot of Actual and Fitted values (1)	50
Figure 7.2. ADM CASE Run plot of Actual and Fitted values (2)	51
Figure 7.3. ADM CASE Run plot of Actual and Fitted values (3)	51
Figure 7.4. ADM CASE Run plot of Actual and Fitted values (4)	52
Figure 7.5. URLA CASE Run plot of Actual and Fitted values (1)	53
Figure 7.6. URLA CASE Run plot of Actual and Fitted values (2)	53
Figure 7.7. URLA CASE Run plot of Actual and Fitted values (3)	54
Figure 7.8. URLA CASE Run plot of Actual and Fitted values (4)	55
Figure 7.9. CFSR CASE Run plot of Actual and Fitted values (1)	56
Figure 7.10. CFSR CASE Run plot of Actual and Fitted values (2)	56
Figure 7.11. CFSR CASE Run plot of Actual and Fitted values (3)	57
Figure 7.12. CFSR CASE Run plot of Actual and Fitted values (4)	58
Figure 7.13. MERRA CASE Run plot of Actual and Fitted values (1)	59
Figure 7.14. MERRA CASE Run plot of Actual and Fitted values (2)	59
Figure 7.15. MERRA CASE Run plot of Actual and Fitted values (3)	60
Figure 7.16. MERRA CASE Run plot of Actual and Fitted values (4)	61

Figure 7.17. ADM CASE IRF Plots	66
Figure 7.18. CFSR CASE IRF Plots	67
Figure 7.19. MERRA CASE IRF Plots	67
Figure 7.20. URLA RES CASE IRF Plots	68
Figure 7.21. Residuals ADM	69
Figure 7.22. Residuals CFSR	70
Figure 7.23. Residuals MERRA	70
Figure 7.24. Residuals Urla RES	71
Figure 7.25. Forecasted Values versus Real Data of 01 July 2000	72
Figure 7.26. Forecasted Values versus Real Data of 01 October 2000	73
Figure 8.1. Nordtank Data Station	75
Figure 8.2. Roughness Effect and Wind Rose	78
Figure 8.3. Wind Atlas	78

LIST OF TABLES

<u>Table</u>	<u>Page</u>
Table 2.1. Wind Power Potential in World According to several authors	7
Table 2.2. Global Installed Wind Power Capacity (MW) Regional Distribution	8
Table 2.3. Turkey Renewable Energy Situation	10
Table 2.4. Turkey Wind Energy Situation in the end of 2013	10
Table 3.1. Force Affect table	17
Table 6.1. Model Progress Data Sets Range 06,2000-10,2001	43
Table 7.1. Comparison of Model Forecasting Performances for 17 months means ...	62
Table 7.2. Comparison of Model Forecasting Performances for ADM	63
Table 7.3. Comparison of Model Forecasting Performances for URLA	64
Table 7.4. Comparison of Model Forecasting Performances for CFSR	65
Table 7.5. Comparison of Model Forecasting Performances for MERRA	65
Table 7.6. Means of Residuals (four digits)	69
Table 8.1. Nordtank Data Set	75
Table 8.2. Measured Power Outputs	79
Table 8.3. Comparison for only 2004 (GWh)	79
Table A.1. Model Parameters for ADM	86
Table A.2. Model Parameters for CFSR	86
Table A.3. Model Parameters for MERRA	87
Table A.4. Model Parameters for URLA RES	87
Table B.1. Power Curve Data of Turbine	88

LIST OF SYMBOLS

Mtoe	million tons of oil
MWh	megawatt hours
kWh	kilowatt hours
WS	wind speed
f	Coriolis parameter
w	Earth's rotation rate
r	radius
w	Earth's rotation rate
z_0	parameter of roughness
ρ	air density
τ	the shear stress
K	von Karman's constant
d	displacement height
t	time index
p	lag parameter
ϵ_t	residual term
E	expected value
$var()$	variance
$cov()$	covariance
μ	mean
n	observation number

LIST OF ABBREVIATIONS

WWEA	World Wind Energy Association
met mast	Meteorological Mast
ADM	Adnan Menderes Airport
MERRA	Modern ERA Retrospective-Analysis For Research and Applications
CFSR	Climate Forecast System Reanalysis
SCADA	Supervisory control and data acquisition system
MCP	Measure Correlate Prediction
PGF H	Pressure Gradient Force Horizontal
CF	Coriolis Force
CeF	Centrifugal Force
FF	Frictional Force
GF	Gradient Force
PGF V	Pressure Gradient Force Vertical
IEC	International Electrotechnical Commission
MEASNET	Network of European Measuring Institutes
RMSE	Root Mean Square Error
ARMA	Autoregressive Moving Average
ANN	Artificial Neural Networks
MAE	Mean absolute error
MRE	Mean relative error
SARIMA	Seasonal autoregressive integrated moving average
MAPE	Mean absolute percentage error
VAR	Vector autoregression
WAsP	Wind atlas analysis and application program
OWC	Observed wind climate
RWC	Regional wind climate

CHAPTER 1

INTRODUCTION

Main goal and the objectives are listed in below section. Organization of thesis is given in second part. Due to increase understandability of study and rough road map which is designed according to contents of chapters and sections.

1.1. Thesis' Aim and Objectives

Wind Energy has recently been recognized as one of the most promising renewable energy sources in the world (Erdem and Shi, 2011). Its increasing trend during the last few decades has undoubtedly confirmed this potential. Such that the World installed capacity has risen from 7,480 MW in 1997 to 74,112 MW in 2006 and to 318,529 MW in 2013 (WWEA's annual report, 2012 and Key statistics of World Wind Energy Report, 2014). In other words, the total capacity has grown more than 44 times over the last 15 years and estimated to reach about 700,000 MW in 2020 (WWEA).

Despite its high potential, one major problem is that wind energy is an intermittent source which follows, in general, a quite noisy evolution with large variability and difficulty in forecasting (Erdem and Shi, 2011). Thus, the power generated from turbines can hardly be predicted (Erdem and Shi, 2011). Forecasting the wind speed is, therefore, critical to predict the amount of power generation (Erdem and Shi, 2011). A small change in the direction and magnitude of the wind might, in fact, create large deviations in power production (Erdem and Shi, 2011).

With regard to the literature on wind speed forecasting, there exist several methodologies. One may partition these approaches into three groups. The first one incorporates physical factors, such as terrain, mountains and obstacles, in wind speed forecasting (Lei et al., 2009), (Giebel et al., 2011), (Erdem and Shi, 2011). Generally, these methods have been used for large-scale places like regions or countries and they are unlikely to fit to the local sites (Lei et al., 2009), (Giebel et al., 2011), (Erdem and Shi, 2011). Second group consists of Artificial Neural Network (ANN) methods which are adopted by a strand of scholars in order to model the wind speed (Li and Shi, 2010) (Carolin Mabel and Fernan-

dez, 2008) (Erdem and Shi, 2011). The main advantage of ANN is that it does not require an assumption of a certain form of regression function, thus, represents a more general approach in that sense. However, the drawback is the weakness of forecast accuracy compared to the other conventional methodologies. (Cadenas and Rivera, 2007), (Gomes and Castro, 2012).

The third group, perhaps the most influential, includes conventional statistical time series methods which rely mostly on ARMA (Autoregressive Moving Average) and ARIMA (Autoregressive Integrated Moving Average) based models (Cadenas and Rivera, 2007), (Erdem and Shi, 2011). ARIMA has, initially been introduced by Box-Jenkins (Erdem and Shi, 2011) and useful in incorporating the persistence, variance and stationarity (order of integration) of wind in predictions (Erdem and Shi, 2011). In addition, more recently developed SARIMA model takes into account the seasonality component. It is argued to outperform the former models and provide more accurate results (Erdem and Shi, 2011). Some examples of these studies are Bivona et al. (2011) who implement wind speed forecasting in Italy (Cammarata Mazzarrone) between 1st Jan 2003 to 31 st December 2006 by employing SARIMA and ANN, Cadenos and Rivera (2007) who predict wind speed in Oaxaca (Mexico) between June 1994- May 2000 and conclude that SARIMA outperforms the ANN based approaches, Pedro et al (2012) who perform similar analysis in Iberia and concludes that ARMA based model produces more accurate results.

The present thesis aims at contributing to literature by departing from other studies in three ways. First, the vast majority of existing studies is based on a uni-variate modelling. In other words, wind speed is estimated solely using its former values. However, this is likely to create a significant loss in forecast accuracy as the important dynamics of wind such as ambient temperature, absolute pressure; wind direction and humidity are ignored (Erdem and Shi, 2011). Two exceptional studies are Zhang et al (2002) and Erdem and Shi (2011) who incorporate wind direction and temperature (in Zhang et al. (2002) in their forecasts and obtained efficient results. Besides these studies, literature considering these dynamics is rather scarce. I intend to take them into account in a Vector Autoregression (VAR) model which is a widely used methodology in interdependent multivariate settings, mostly used in economics and social sciences (Sims, 1980). Specifically, VAR methodology is used to estimate directly the co-variates of endogenous variables by using time lags of all variables. The dynamics of wind speed, such as wind direction, pressure, humidity, etc., are included in VAR model so to capture their possible effect on wind

speed predictions.

Second novelty of our study comes from the fact that we compare the estimations from 4 different locations. It therefore provides an advantage of cross-locational comparison of model predictions and their robustness. Similarly, we compare our results from three methodologies, one is multivariate VAR and the others are uni-variate ARIMA and SARIMA.

Lastly, to the best of our knowledge, this thesis is one of the first attempts which try to model the wind speed in Turkey. In fact, wind is seen as one of the most potential energy sources in Turkey since there exist many persistently windy locations. Indeed, wind energy exhibits an increasing trend in recent years. Such that installed capacity has risen from 363,7 MW in 2008 to 791,6 MW in 2009, to 1320,2 in 2010 and to 1728,2 in 2011. (Turkish Republic Energy Market Regulatory Authority, 2012).

With regard to the locations of study, we analyze the wind speed from 4 different stations. The first one is called URLA-RES which is a met mast located in Urla (sub-province of Izmir-Turkey) and within the campus area of Izmir Institute of Technology. Secondly, we use data measured at Izmir Adnan Menderes Airport (ADM) which is about 50 km away from Urla. The third dataset is called MERRA (MODERN ERA RETROSPECTIVE-ANALYSIS FOR RESEARCH AND APPLICATIONS). Originally, the data is provided by Global Modeling and Assimilation Office of NASA/Goddard Space Flight Center. The wind speed measurement is implemented from a location on the Aegean Sea which is nearby Urla. Specifically, its coordinates are latitude; 38.5 and longitude; 26.6. The fourth data is called CFSR (Climate Forecast System Reanalysis). This data is provided and processed by EMD International A/S. The location of measurement is again nearby places of Urla (Latitude: 38.2, Longitude: 26.3). All the data we use have an hourly frequency and cover the period from June-2000 to October-2001.

Overall, the main objectives of the study can be summarized in a following way:

- i To be able to understand the state of art and summarize the mostly accepted methodologies in the literature
- ii To employ multivariate methodologies which takes into account meteorological variables in the estimation process.
- iii To compare different methodologies' accuracy power using several criteria such as Root mean square error, Mean absolute error, Mean absolute square error tools.

- iv To contribute to the literature with an engineering application by comparing the raw, filled and fitted data in terms of their resulting annual energy production values.

1.2. Organization of Thesis

Structure of the thesis can be summarized as follows. General pictures of wind energy industry in World and in Turkey are presented in chapter 2 herein reader can gain an overview the scales of numbers, statistics of this market. In chapter 3, I provide the theoretical explanations of general concepts such as wind and occurrence of wind, meteorological characteristics and mathematical/physical explanations of wind and measurement tools.

Chapter 4 has been devoted to provide a brief account of the time series models. Therefore, basic terminology and standard procedures of estimations will be summarized in this part. Specifically, I will focus on unit root and stationarity concepts and, then, explain the three models; VAR, ARIMA, SARIMA.

After this point, in Chapter 5, I summarize the related literature by reviewing the empirical published papers and document their main methodologies and results

In Chapter 6, I explain the dataset that I employ, the variables, their measurements and locations. Herein mainly two types of data family will be used. On the one hand, the met mast data from four different stations nearby Urla-Izmir will be used which is given in chapter 6 whereas, SCADA engineering application dataset from Denmark will be introduced in chapter 8.

Chapter 7 has been devoted to reporting the obtained results and comparison of wind speed predictions of 3 different methodologies and 4 locations. Optimization of models; ARIMA(p,d,q), SARIMA(p,d,q)(P,D,Q) and VAR(p) and procedure is also explained. Herein fitted values obtained from estimation of each models and for each month has been displayed together with actual wind speed data to have a preliminary idea on the accuracy of estimations. Residuals have also been plotted to check whether they satisfy the white noise property. Cross-comparison of models in estimation accuracy is evaluated using different criteria. IRFs (Impulse-Response functions) are provided to illustrate better the dynamics of wind speed.

In chapter 8, I pursue the engineering application. After the determination of the most powerful method, I will use this methodology to fill missing data for data cleaning

procedure. Herein annual energy productions of raw, produced and filtered data sets will be used as comparison tool. WAsP program will be employed for the calculation of annual energy production values.

Finally, in chapter 9, I will conclude my study by providing several final remarks.

CHAPTER 2

WIND ENERGY MARKETS IN WORLD AND IN TURKEY

The current chapter is devoted to summarising stylized key facts on wind energy markets in world and in Turkey.

As known, in today's world, there is an undeniable fact that civilization established heavily relies on energy dependent activities like manufacturing, production, transportation, heating, cooling, air conditioning, etc.. Energy, however, is not infinite. A renewable system is, therefore, required.

Conventional energy sources, such as gas, coal, etc., include quite important drawbacks. In detail, they represent a clear threat to sustainability and environment. A list of potential problems is documented below:

- i Green house gasses emission problem and in parallel global warming
- ii Unequal distribution of conventional energy sources, price risks, security problems due to transportation
- iii Limitation in this sources, they are neither sustainable nor renewable.

One possible solution could be through the Wind Energy. Although wind is known as renewable and sustainable energy source there is a little misunderstanding. Wind energy is also limited but compared to conventional sources, it will not be ran out of or perish. In reality, none of the energy types will be perished and they will be transforming each other and burning process. In the most of applications, this process is irreversible. So when you burn a coal there will be heat energy and if you use this energy to boil water you can access the vapor with pressure and the movement of the vapor you can turn propellers. This means one has converted the heat energy into a kinetic energy and from this kinetic energy, electricity can be produced. But here energy conversion procedure depends on the consumption of coal. After some point the energy sources (i.e. coal, petroleum) can not meet the demand as long as they are not renewable .

The global available energy in the wind was calculated by several authors. According to one of the pessimist estimation, its potential is about 1 TW in global scale (Castro et al., 2011). In another words about 8.4% of today's energy demand can be

obtained from the wind for the pessimist estimation. Let's make this calculation to get familiar with units and scales.(Agency, 2014) in 2012 world total energy consumption is 8979 Mtoe.

1 toe = 11.63 MWh

8979 Mtoe = 1.0442577×10^{14} kWh

in 1 year 8760 hours

1.0442577×10^{14} kWh/8760 h=11920750000 kW

1TW=1 000 000 000 kW

As a result 11920750000 kW/1 000 000 000= 11.92075 TW is the total consumption and 1 TW approximately equals to $\approx 8.4\%$ of total consumption.(Castro et al., 2011) also listed other estimations in his study. In the table 2.1 this general summary can be found.

Table 2.1. Wind Power Potential in World According to several authors

Authors	Technical power (TW)	Economic/sustainable power (TW)
Capps and Zender (2010)	39 (offshore)	
DeVries et al. (2007)		4.5(in 2050)
EEA (2009)	8.6 (Europe)	3.5(Europe 2030)
Elliott et al. (2004) and Musial (2005)	1 (EEUU)	
Greenblatt (2005)	70.4 (global)	
Greenpeace (2008) and Greenpeace (2010)		1 in (2050) 1.2 (in 2050)
Hoogwijk et al. (2004)	11	2,4 6 (economic)
Lu et al. (2009)	78(onshore), > 7(offshore)	
Miller et al., (2010)	17– 38 (onshore, geographical potential)	
Schindler et al. (2007)		6.9(sustainable)
WEC (1994)	55.2 (global)	
Wijk and Coelingh (1993)	2.3(onshore,OCDE)	
Zerta et al. (2008)		6.9(sustainable)

In more optimistic estimation,The Global Wind Energy Council and Greenpeace International Global Wind Energy announced that wind power could reach 2000 GW by 2030 which equals to 17-19% of Global electricity demand, this means also over 2 million new jobs and reducing CO² emissions by more than 3 billion tones per year in 2014. Herein it stated also 21000 jobs for a year are created by the investment of every

billion Euro in offshore wind.(Global Wind Energy Council (2014)).

Here, I need to explain some terms more for the sake of clarity. So, our terminology starts with energy mix which is the term to explain a country's primary energy sources used to produce electricity. Installed capacity is a term used for explaining the maximum technological capacity that can be produced by the maximum power in optimum conditions, like wind blowing in the range of 11 m/s and 18 m/s.

With regard to the current situation in world, global installed capacity of wind power and its regional distribution are presented in Table 2.2.

Table 2.2. Global Installed Wind Power Capacity (MW) Regional Distribution
(Source : Global Wind Energy Council.,2014)

	end of 2012	new 2013	end of 2013
AFRICA & MIDDLE EAST	1,165	90	1,255
ASIA	97,715	18,216	115,927
EUROPE	109,817	12,031	121,474
NORTH AMERICA	67,748	3,063	70,811
PACIFIC REGION	3,219	655	3,874
World total	283,194	35,289	318,105

By the end of 2013, total world capacity is 318,105 MW. The region with the largest capacity is Asia (near to 50 % of global capacity) followed by Europe, North America, Pacific, Africa and Middle East.

Regarding the ranking at the country level, the world leader country is China followed respectively by US and Germany. Turkey is 16th in the ranking.

In terms of production techniques of wind energy, 3.6 MW offshore turbine can produce in average the sufficient amount of electricity for 3312 EU households (European Wind Energy Association). On the other hand 2-5 to 3 MW onshore turbine can produce 6 million kWh in a year which is enough to meet 1500 EU households' electricity demands.(European Wind Energy Association).

Figure 2.1 illustrates instead the installed wing energy within EU. The total installed capacity is 117,289 MW while candidate countries have far little capacity (i.e. 2,956 MW), reaching to an overall capacity of 121,474. Regarding the individual countries, Germany, Spain, UK and Italy are the leading countries with respective capacities of 33,230 MW, 22,959 MW, 10,531 MW and 8,551.

Table 2.3. Turkey Renewable Energy Situation(Source: EPDK.,2012))

TYPE	POTENTIAL	IN OPERATION
Hydro	45,000 MW	17359.3 MW
Wind	48,000 MW	1792.7 MW
Solar	300 TWh/year	—
Geothermal	600 MW	114.2 MW
Biomass	17 MTEP	117.4 MW

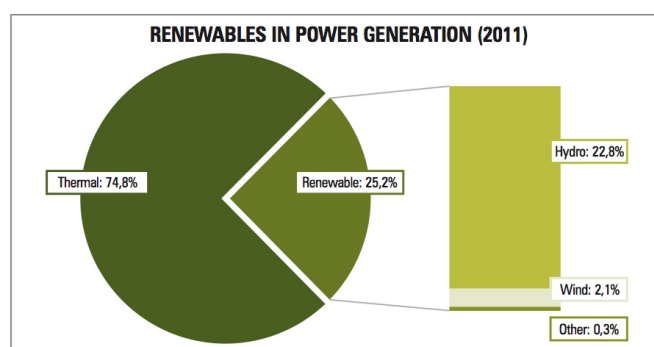


Figure 2.2. Turkey Energy Mix in 2011(Source: EPDK., 2012)

Figure 2.2 and Table 2.4 provide more detailed information on the recent developments of wind energy sector in Turkey. In Figure 2.2, we can observe that wind installed capacity shows a clear tendency to increase over the last years. Its capacity has risen from about 363 MW in 2008 to 1728 MW in 2011. Moreover, its relative share in total energy capacity rises from 0.87 % in 2008 to 3.25 % in 2011.

Table 2.4. Turkey Wind Energy Situation in the end of 2013 (MW)
(Source: TUREB., 2013)

Installed Capacity	Under Construction	Licensed Wind Power Plants
2958.45	980.90	5776

In 2012, Turkey has 2.312 GW wind energy power plants and 646 MW is constructed after that. Finally, by the end of 2013 this value has reached to 2.959 GW. The total number of Licensed Wind Power Plants is 5776.

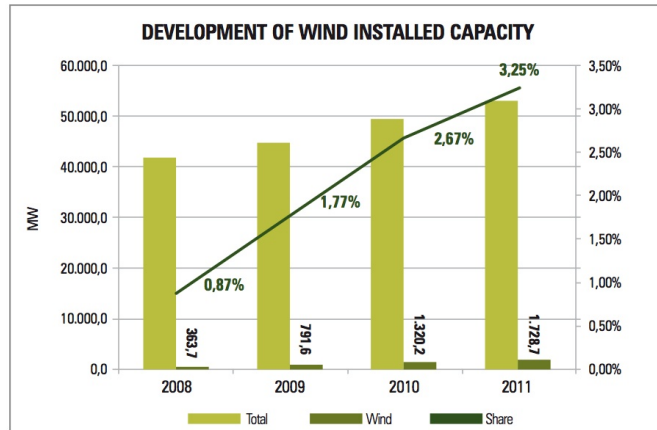


Figure 2.3. Turkey Wind Energy Installed Capacity between 2008 and 2011
(Source: EPDK., 2012)

Finally, a last figure in this chapter regards the geographical distribution of wind power plants in Turkey. In figure 2.4, it is obvious that Turkey’s wind energy applications are denser in the west part of the country. They are mostly concentrated on the Western and Aegean Sea shore since the wind speed in those locations is quite strong and persistently blowing. Another reason for such an agglomeration is the capital structure. Big investors are available around Istanbul and Marmara which directs the investment to accumulate around these zones.



Figure 2.4. Turkey Wind Turbine Map

CHAPTER 3

THEORETICAL BACKGROUND

As a physical definition, wind is the air motion driven by the uneven heating and cooling of the earth's surface. It occurs at Troposphere layer. It has been explained within boundary layer theory and by different balance situations of pressure gradient force, Coriolis force, frictional force, gravitational force, centrifugal force and thermal gradient.

The horizontal movement of the air, parallel to the earth's surface is measured by direction and magnitude of wind. Occurrence of wind is highly uncertain in time and space, which has a complex periodicity and non-linearity.

Thus, the wind speed is closely related to other meteorological factors such as pressure and temperature, and it can also be easily influenced by obstacle and terrain.

An overview of related concepts are provided in this chapter.

3.1. The Wind

Movement of air is a natural phenomena which is caused by changing at balance of atmospheric forces mostly due to nonequivalent warming and cooling of earth.

We will discuss this issue within different aspects in this chapter. In order to give a general overview to reader, first part is reserved for explanation of theoretical basement. Herein we will start with the answers of two crucial question, "where" and "when" to define wind as a natural phenomena with the equivalent of physical and mathematical models. Main concepts are listed below;

- i Where? Atmosphere via Troposphere, layers
- ii Which forces? Definition of atmospheric force, Pressure, Gravity, Coriolis, Thermal.
- iii When? The time of balance between forces

3.1.1. Atmosphere and Troposphere

According to Robert Henson Atmosphere is a life sustaining soup which has the main ingredients oxygen, ozone and water vapour. The ninety percent of atmosphere lies within 32 km of the planet's surface. Comparatively to earth's radius (13 000 km) atmosphere is a thin film covering the surface of earth and "weather layer" of atmosphere where we can find our "wind", is even thinner. Above that there are several atmospheric layers. Each of them has different characteristics like chemical combination, temperature profile, cloud types etc.

The Ground layer is the location where our wind is hidden, Troposphere. It is the bottom of atmosphere extending from sea level up to around 10-16 km high, depending on the season and the location figure 3.1 (poles, equator etc.). As a typical limit of troposphere its height can reach up to a maximum at the place where the air is warmest and at the time where the air is warmest.(Henson, 2011)

What keeps the troposphere moving are contrasts between warm and cold, moist and dry (Henson, 2011). Generally troposphere defined with sub-layers; surface layer, ekman layer, free atmosphere.

3.1.1.1. Atmospheric Forces

Until here we already defined "the place". Right now we will start to define "parameters" which can structure the theory basement. Force is a physical parameter that defined as multiplication of acceleration and mass. The fundamental forces are sorted below; (Ackerman et al., 2013)

- i) Pressure Gradient Force (The PGF always pushes from higher pressure towards lower pressure. Its magnitude is equal to the change in pressure over some distance, divided by the air density.)

$$PGF = \frac{-1}{\rho} * \frac{\text{change in pressure}}{\text{distance}} \quad (3.1)$$

THE TROPOSPHERE

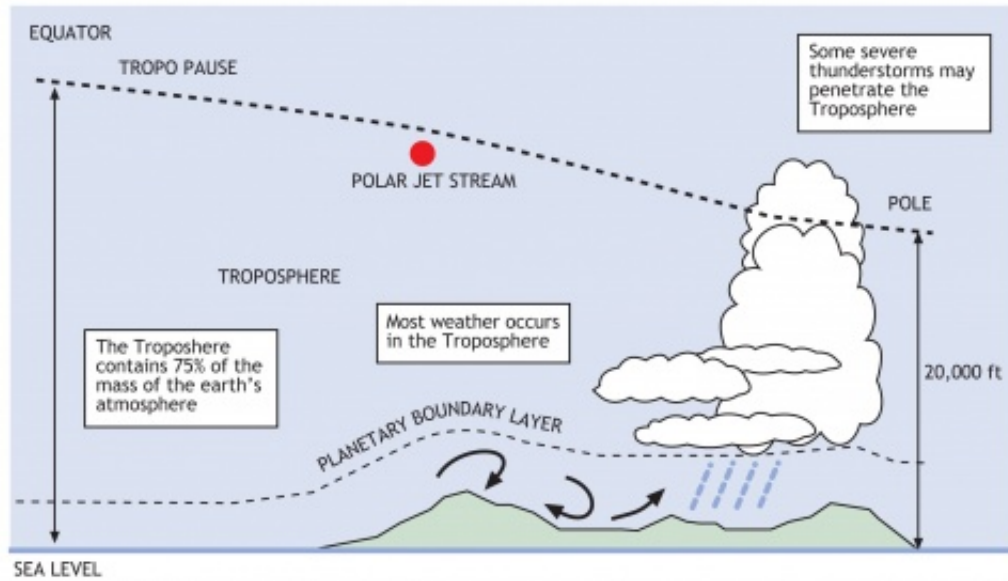


Figure 3.1. Troposphere and height change from equator to pole.
(Source: sky.aero.,2014)

- ii) Coriolis Force (The rotation of the Earth produces this force which is felt in the frame of reference of the spinning planet) The magnitude of the Coriolis force (CF) per unit mass is proportional to the distance from the equator and the speed of the wind. This can be written mathematically as:

$$CF = \pm f * WS \quad (3.2)$$

In equation 3.2 WS is the wind speed and f is the Coriolis parameter, defined as:

$$f = 2\omega \sin \varphi \quad (3.3)$$

w: Earth's rotation rate

$\sin \varphi$: sin of latitude

- iii) Centrifugal Force (When an object changes its direction of motion it is accelerating even if its speed does not change. This acceleration is called centripetal acceleration (centripetal means “pushed toward a center”).(Ackerman et al., 2013))

$$CENTF = \frac{WS^2}{r} \quad (3.4)$$

R is the radius of curvature part.

- iv) Frictional(Viscous)Force (The frictional force in the atmosphere is, for our current purposes, caused by the flow of wind over the roughness of the Earth’s surface.) Mathematically, the frictional force (FF) per unit mass can be written very simplistically as:(Ackerman et al., 2013)

$$FF = -z_0 * WS \quad (3.5)$$

in equation 3.5 z_0 is a parameter describing the roughness of the Earth’s surface and WS is the wind speed. The minus sign emphasizes that friction opposes the the wind regardless of its direction.

- v) Gravitational Force (Gravitational force is defined with the perpendicular direction through ground and with magnitude of the mass times the gravitational acceleration, 9.8 meters per square second.)

3.1.1.2. Atmospheric Balances

In this subsection I will introduce atmospheric balances of troposphere for the definition of time when the wind occurs. In table 3.1 rows represent the balance types and columns represent the each force which have a place in that specific balance situation. By

the way balance means equilibrium of different forces in reverse directions with the same magnitude so in the balance time there will be no motion however in nature attitude of forces towards the balance is weaker than the attitude of forces towards the motion. The wind occurs during to the motion of forces to reach balance.

Table 3.1. Force Affect table

Forces versus Balances	PGF H	CF	CeF	FF	GF	PGF V
Hydrostatic Balance ($PG + GF = 0$)					X	X
Geostrophic Balance ($PG + CF = 0$)	X	X				
Gradient Balance	X	X	X			
Guldberg-Mohn Balance	X	X			X	

PGF H: Pressure Gradient Force Horizontal

CF: Coriolis Force

CeF: Centrifugal Force

FF: Frictional Force

GF: Gradient Force

PGF V: Pressure Gradient Force Vertical

3.1.1.3. Gradient Balance

Geostrophic balance requires that the wind blow in a straight line. However, this is rarely the case. More often, upper-level winds undulate in curvy patterns. This requires the inclusion of a third force, the centrifugal force. The three-way balance of horizontal pressure gradient, Coriolis force and centrifugal force is called gradient balance, and the wind that results from this balance is called the gradient wind. The gradient wind is an excellent approximation to the actual wind observed above the Earth's surface, especially at middle latitudes and in the jet stream.(Ackerman et al., 2013)

Gradient balance illustrates that high and low pressure areas differ not only in the direction but also the strength of the wind. Remember that the centrifugal force always pushes toward the outside of curves and is proportional to the square of the wind speed. Therefore, in the case of a curved low-pressure area, the centrifugal force pushes in the same direction as the Coriolis force. Since both of these forces are related to wind speed,

the wind does not need to be as strong in this case as in the case of uncurved flow in order to achieve balance with the PGF. The wind is therefore slower in this situation than in a purely geostrophic case, and is called subgeostrophic. (Ackerman et al., 2013)

Conversely, in the case of a curved high-pressure area the centrifugal force teams up with the PGF, both pushing outward from the high. The Coriolis force attempts to balance this double-team of forces. To accomplish this, the wind speed to which the Coriolis force is proportional-must be considerably higher than in the purely geostrophic case. This is called supergeostrophic flow.

Therefore, for high- and low-pressure areas that have the same spacing of isobars or isoheights, winds around a high will be stronger than winds around a low because of gradient balance.

3.1.1.4. Guldberg-Mohn Balance

Near the Earth's surface, the frictional force plays a pivotal role in wind. The adjustment to balance in these situations is profoundly affected by the ability of friction to slow the wind and therefore weaken the Coriolis force. This three-way balance of horizontal PGF, Coriolis force, and frictional force is sometimes called Guldberg-Mohn balance. (Ackerman et al., 2013)

The creation of a PGF creates a wind from higher toward lower pressure. However, near the surface this wind triggers not only the Coriolis force, but also the frictional force. Friction opposes the wind, so the wind slows down which in turn makes the Coriolis force weaker. As the situation evolves, the PGF wins the tug-of-war. This is because of the balance of the three forces. Without the influence of the Coriolis force, the wind would blow straight into lower pressure. With the Coriolis force at full strength opposing the PGF, the wind would blow parallel to the isobars as in geostrophic balance. But in this case, friction weakens the wind and the Coriolis force, and the air parcel crosses the isobars toward lower pressure. (Ackerman et al., 2013)

3.2. The Atmospheric Boundary Layer

Ludwig Prandtl defined the concept of boundary layer in 1904 in terms of the layer where a moving fluid comes across with a solid boundary. (Tani, 1977). In this concept atmosphere divides mainly two region according to importance of viscosity.

The Boundary Layer can be defined as: “that part of the troposphere that is directly influenced by the presence of the earth’s surface and responds to surface forcing with a time scale of about an hour or less.”(Stull, 1988)

The lowest one tenth part of the atmospheric boundary layer, is called the surface layer. In this layer the turbulent fluxes remain more or less (within 10%) constant with height, and the Coriolis effects are negligible.

3.2.1. Wind Profile

The concept of boundary layer gave us the opportunity to formulate vertical wind profile.

In the surface layer turbulence is generated by wind shear near the ground; the no-slip condition requires that velocity at the ground is zero. This is expressed in terms of Reynolds stress, and can be defined a velocity scale WS^* .(frictional velocity)

$$WS^* = \frac{\tau}{\rho_0} \quad (3.6)$$

Where is τ the shear stress in an arbitrary layer of fluid and is the density of the fluid.

$$WS(z) = \frac{WS^*}{K} \ln \frac{z}{z_0} \quad (3.7)$$

In equation 3.7, K represents von Karman’s constant $WS(z)$ is the wind speed at height z , z_0 is the roughness length. The roughness of a terrain is included to formulation with the parameter of the roughness length (z_0), in another words terrain roughness is

the measurement of hindrance which is “seen by” the surface wind. (Crockford and Hui, 2007a)

For the forestry region this formulation turns to equation 3.8.

$$WS(z) = \frac{WS^*}{K} \ln \frac{z - d}{z_0} \quad (3.8)$$

The wind profile is essentially shifted upwards by d (tree height), zero plane displacement height. In general its value is based on the estimation which is vary between 0.6 and 0.8 times the height of the canopy (forestry(tree) or city(building)) (Stull, 1988).

3.3. Anemometer and Measurement Tools

Development of wind industry depends on improvement of measurement tools. This is a crucial issue in the improvement of quality, frequency and accessibility of measured data. Wind industry and measurement technology of atmospheric variables affect each others directly.

Basically three main phases require measurements in wind power plants. First one is screening phase which it is focused on description of wind climate in a specific area. Second one is referred as to a development phase where input can be collected, measured, predicted for energy calculations, investment decisions, turbulence and wake analysis, wind shear analysis, extreme wind analysis, choice of turbine type (IEC class 1, class 2, A, B..), site calibration. Finally last phase calls as operational domain. We can define it here as a produced power measurements.

IEC standards (International Electrotechnical Commission) and MEASNET (International Network for Harmonized and Recognized Measurements in Wind Energy) regulations are widely accepted in wind industry. General rules based on requirement of measurements at least 12 complete and consecutive months at wind park area. Data specifications require that it must be recorded within 10-min averages, preferably with a sampling rate of 1 Hz or faster .

If we start to investigate techniques firstly mast measurements should be explained with different construction types and advantages and disadvantages. Met mast tower could be tubular mast or lattice mast.

Tubular masts provide benefits such as having low cost measurements. It is easier to erect the mast and concrete foundation can be avoided comparatively lattice mast. But maintenance issues are really challenging in this type since it is difficult to change the sensors and mast must be taken down. During the winter, this type of mast can not hold heavy ice loads. Furthermore, there is a height limit, i.e. like smaller than 80 m(short masts).

Lattice masts show controversial specifications with tubular masts. In this type, climbing is possible, holding heavy icing loads depend on the design. Tall towers must be designed as lattice masts. If we check disadvantages list here we must face with high investment costs and height so high altitude have influence on the measurements (shadow effect). The desired mast structure and equipment specifications can be listed as follows:

- i) Well documented and tested equipment
- ii) Calibration and mount according to IEC and MEASNET
- iii) Lighting protection
- iv) Sensors at different heights
- v) Two sensors mounted at the top of met mast
- vi) The met mast should be at least 2/3 of hub height
- vii) Usage of suitable sensors for local climate (icing)
- viii) Monitoring the sensors and the met masts continuously
- ix) Ensure a good documentation of the measurement program

The widely accepted tool in the industry is the cup anemometer. But most roughly and simple measurements record from cup anemometer. Cup anemometers locate at turbines and within different altitudes at met mast for collecting information (such as wind speed and wind direction)

Maintenance issue is the main challenge because of high altitude climbing and changing of sensors are generally difficult and at least 1 year measurement duration is interrupted because of icing affects, lighting etc. Loss of data decreases the accuracy of expected annual energy production values so real situation could not be captured directly with met mast measurements.

We have alternative systems of measurement based on remote sensing. Sodar, Lidar, Ultrasonic anemometer are the alternatives to cup anemometers.

Sonic anemometers deviates from cup anemometer with the different measurement methodology based on sound speed.

Light Detection And Ranging so Lidar based on light speed measurement. The most expensive application but it is removable and presents much more wide range for the different atmospheric variables.

Ultrasonic anemometer has already some standards on IEC but it is expensive than cup anemometer but some new application like nacelle montaged ultrasonic anemometers, this technology is the closest one which can be accepted as a standard and take place of cup anemometers.

CHAPTER 4

METHODOLOGY

Herein the main tools, those are summoned due to complete computational and numerical calculations, are listed with details in following sections.

4.1. Statistical Processes That Impacts ARIMA & VAR

There exists some data characteristics and processes in time series literature which have a direct impact on the predictive performance and estimation of AR, ARMA, ARIMA (autoregressive time series analysis) models. These are namely stationarity and seasonality for which the brief definitions are given below.

On the one hand, stationary is a concern about stochastic process which is based on probability distribution that does not tend to change when shifted in time. The main characteristic of stationary variables is the fact that the first and second moment of distributions (mean and variance) do not tend to increase or decrease over time.

Many socio-economic variables follow a non-stationary processes with either varying mean and/or variance. The processes which have a constant mean (i.e. mean reversion) and variance are called stationary processes, if the mean is 0 and variance is constant, the process is referred to as white noise.

On the other hand, seasonality is another concept related to our issue. Many of the variables in nature follow a seasonal pattern. Such that they exhibit a periodic increase and decrease due to a change in the season. For instance, in Izmir, rain fall accelerates during the winter and autumn while it fades out during summer. Likewise, wind speed and direction are likely to have different characteristics over the seasons. That's why, it is a crucial issue to be handled.

4.1.1. STATIONARITY & STABILITY

4.1.1.1. Unit Root Test

A time series is considered to be stationary if the autoregressive component is significantly different from 1. If that is equal to 1, it means the variable has a unit root characteristic and follows a non-stationary process.(Halim and Bisono, 2008)

There are many tests available in the literature that examines whether a variable is stationary or not. Among others the most widely accepted ones are DF (Dickey-Fuller) test, ADF(Augmented Dickey Fuller) test, PP (Philips Perron) test and so on. (Gao et al., 2009)

For this project ADF test is selected as a testing methodology.

In ADF test, the same equation with DF test is accepted;

$$\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \delta_1 \Delta y_{t-1} + \dots + \delta_{p-1} \Delta y_{t-p+1} + \epsilon_t, \quad (4.1)$$

- t is the time index,
- α is an intercept constant called a drift,
- β is the coefficient on a time trend,
- γ is the coefficient presenting process root, i.e. the focus of testing,
- p is the lag order of the first-differences autoregressive process,
- ϵ_t is an independent identically distributes residual term.

(FinMath.,2014)

The focus of testing is whether the coefficient γ equals to zero, what means that the original ADF Time Series process has a unit root; hence, the null hypothesis of $\gamma = 0$ (random walk process) is tested against the alternative hypothesis $\gamma \neq 0$ of stationarity. A time series is a set of time dependent data of some activity (trend,homogeneity, stationarity,white noise etc.), over defined time period. It is the historical record of some

certain activity, with measurements taken at defined regular time domains in the activity and without changing the method of measurement.

Uncertainties of measurement technique and device capabilities like anemometer sensitivity (cup anemometer, Lidar, Sodar etc.) affects method of measurement and time series characteristics.

The following fundamental statistical definitions are reviewed even though it might be familiar to the reader.

The Variance of random variable X with expected value $EX = \mu_x$ is defined as (μ is mean of X series)

$$var(X) = E((X - \mu_x)^2) \quad (4.2)$$

The covariance between random variable Y and Z , with expected values μ_y and μ_z , is defined as

$$cov(Y, Z) = E((Y - \mu_y)(Z - \mu_z)) \quad (4.3)$$

Variance and Covariance are important definitions for defining time series and white noise characteristics. But it is mainly crucial for autoregressive modelling, particularly in parameter (lag) selection part. To be able to decide optimum lagged value and moving average parameter, we need to calculate Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF). Hence, ACF and PACF are depend on Variance and Covariance definitions.

ACF (Autocorrelation Function) is an outcome of the sample autocorrelations, versus h (the time lags). In other terms the ACF represents the randomness degree in a data set. For random data, autocorrelation function should give a result near zero for whole time domain. If non-random, then at least one of the autocorrelations will give a non-zero outcome. (Friendly, 2002) For a stationary process Y_t , the autocovariance between Y_t and Y_{t-j} is

$$\gamma_k = Cov(Y_t - Y_{t-j}) = E[(Y_t - \mu)(Y_{t-j} - \mu)] \quad (4.4)$$

and the autocorrelation function is

$$\rho_k = Corr(Y_t - Y_{t-j}) = \frac{\gamma_k}{\gamma_0} \quad (4.5)$$

$$\gamma_0 = Var(Y_t) \quad (4.6)$$

PACF(Partial Autocorrelation Function) is the correlation between Y_t and Y_{t-j} after their mutual linear dependency on the intervening variables $Y_{t-1}, Y_{t-2}, \dots, Y_{t-j+1}$ has been removed. The conditional correlation

$$Corr(Y_t - Y_{t-j})|Y_{t-1}, Y_{t-2}, \dots, Y_{t-j+1}) = \phi_{kk} \quad (4.7)$$

is usually referred as the partial autocorrelation in time series. One example:

$$\phi_{11} = Corr(Y_t - Y_{t-1}) = \rho_1 \phi_{22} = Corr(Y_t - Y_{t-2}|Y_{t-1}) \quad (4.8)$$

After defining the basic statistical derivations and definitions we will now introduce the autoregressive models in next sub-section. Autoregressive models give us opportunity to have an initial idea on how one can model the wind speed and other variables using time series models.

4.2. Univariate Time Series Models

The methods of estimations in this study rely on autoregressive and moving average procedures, a brief account of which are documented in this part. So, AR (autoregressive), MA (moving average) and a combination of previous ones ARMA, ARIMA models will be introduced below.

4.2.0.2. AR(p) Model Form

Let y be random univariate time dependent variable and $y(t) = y(t_1), y(t_2) \dots y(t_n), y(t)$ series are generated from an AR (p) model

$$y_t = \sum_{j=1}^p \phi_j y_{t-j} + \epsilon_t \quad (4.9)$$

ϵ is error term, model deviation from actual value. $\epsilon_t \sim N(0, \sigma_\epsilon^2)$ meaning that the ϵ_t are identically independently distributed each with a normal distribution having mean equals 0 and constant variance. Properties of errors ϵ are independent of y .

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \epsilon_t \quad (4.10)$$

for $t=1, 2, \dots$ (and conceptually $t=0, -1, \dots$) and with $\epsilon_s \perp \epsilon_t$ for $t \neq s$

In a typical autoregressive model, y_t is regressed on historical values of y . And, the order p represents the length of past periods that will be incorporated into the model. (Gerolimetto, 2010)

AR model is univariate linear prediction model. y_t follows a linear, homogeneous Gaussian process, and has time reversible characteristic. But other accepted common models with non-Gaussian shocks are linear yet but not time reversible. The main characteristic of Gaussian is that the process is completely specified through its first moments (mean, variance, covariance). So, if mean and variance, covariance matrix Σ are known, the process is known.

In AR(p) processes: the ACF function is infinitely extended and PACF is infinitely extended (to order p) for identifying an AR(p) process the PACF plays a very important role.

4.2.0.3. Backshift Operators & Characteristic Polynomials

- Operator B such that $B y_s = y_{s-1}$ and so $B^k y_s = y_{s-k}$ for all k .

$$- y_{t-p} = \sum_{j=1}^p \Phi_j B_j y_t + \epsilon_t$$

$$- \Phi(B)y_t = \epsilon_t$$

- Characteristic Polynomial $\Phi(u) = 1 - \Phi_1 u - \Phi_2 u^2 - \dots - \Phi_p u^p$. This is a polynomial of order p defined on $|u| < 1$.

4.2.0.4. Inversion and Moving Averages(MAs)

Moving average is another important concept that explains the degree of smoothness of the error terms in time series processes. It can be expressed mathematically in this way:

MA(1);

$$y_t = \mu + \epsilon_t + \alpha_1 \epsilon_{t-1} \tag{4.11}$$

μ is the mean of y_t series.

$$y_t = \sum_{j=1}^{\infty} \alpha \epsilon_{t-j} + \epsilon_t \tag{4.12}$$

As in the AR(1) case, iterative substitution of y_{t-1} then y_{t-2} and so on in the right hand side of the model equation represents y_t as a linear function of more distant past y_s and $\epsilon_t, \epsilon_{t-1}, \dots$. Formally, assuming the inversion can be performed,

$$y_t = \alpha_t - \epsilon \alpha_{t-1} - \epsilon_2 \alpha_{t-2} \dots - \epsilon_t \tag{4.13}$$

α : is a weight

$$y_t = \Phi(B)^{-1}\epsilon_t \quad (4.14)$$

Operator B and characteristic polynomial are defined in previous section. (Gerolimetto, 2010)

This gives the explicit representation of y_t as a linear combination of independent innovations, a linear (Gaussian) process. This is also referred to as a Moving Average (MA) representation. If the weights α_j cut-off to zero after some finite lag q, we have a finite MA(q) representation.

In MA(q) processes: the ACF function is infinitely extended (to order q) and PACF is infinitely extended for identifying an MA(q) process the ACF plays a very relevant role.

4.2.0.5. ARMA Model

ARMA(p,q) models are defined by

$$X_t = \underbrace{\Phi_1 X_{t-1} + \Phi_2 X_{t-2} + \dots + \Phi_p X_{t-p}}_{\text{autoregressive part}} - \underbrace{\theta_1 \alpha_{t-1} - \dots - \theta_q \alpha_{t-q}}_{\text{moving average part}} \quad (4.15)$$

In the ARMA equation above, p represents the lag length and the order of autoregressive process and q represents a parameter for a determination of degree of moving average. Supposing p=q=1, then we have ARMA(1,1)

$$X_t = \Phi_1 X_{t-1} - \theta_1 \alpha_{t-1} + \alpha_t \quad (4.16)$$

4.2.0.6. ARIMA Model

ARIMA is a comprehensive model that incorporates autoregressive, moving average and stationarity components. As a de-trending treatment, regular differencing is ap-

plied together with AR and MA. I means “integrated” and it explains the need for a differencing procedure. ARIMA might be seen as improved version of ARMA which considers additional aspects like non stationarity in mean and seasonal behaviours. (Gerolimetto, 2010)

If the sequence is not stationary, the smoothing of time series become essential for procedure. In reality, many atmospheric variable (wind speed, wind direction, temperature, pressure, humidity, thermal radiation etc.) may or may not be stationary in first moment (mean).

In order to model the atmospheric parameters, the source of variation must be removed. Generally this treatment is done by differencing (I) the series.(Shoba and Shobha, Shoba and Shobha)

Suppose X_t is nonstationary in mean, to set up an ARMA model on the series w_t , de-trending process is defined here as the result of the operation of differencing the series d times (in general $d = 1$): $w_t = \delta^d X_t$. (Gerolimetto, 2010)

So, ARIMA models are the generalized version if ARMA models. When $I=0$, ARIMA becomes equivalent to ARMA. If the series are non-stationary, d^{th} order differencing process is as follows:

Introduce orderly differential operator $\delta y = 1 - B y$ for the original non-stationary time series first order difference transform

$$\delta y_t = (1 - B)y_t = y_t - y_{t-1} \quad (4.17)$$

The original time sequence through d-order difference,

$$\delta^d y_t = (1 - B)^d y_t \quad (4.18)$$

Difference series and final ARIMA equation becomes;

$$\Phi(B)\Delta^d y_t = \theta B \alpha_t \quad (4.19)$$

where $\Phi(B)\Delta^d$ is called generalized autoregressive operator and $\Delta^d y_t$ is a result after the differentiation operator cleaned the non-stationarity and now this time series can be modelled with an autoregressive- moving average methodology. (Hui et al., 2011)

Examples:

- ARIMA(1,0,0)=AR(1)
- ARIMA(0,0,1)=MA(1)
- ARMA(1,0,1)= ARMA(1,1)

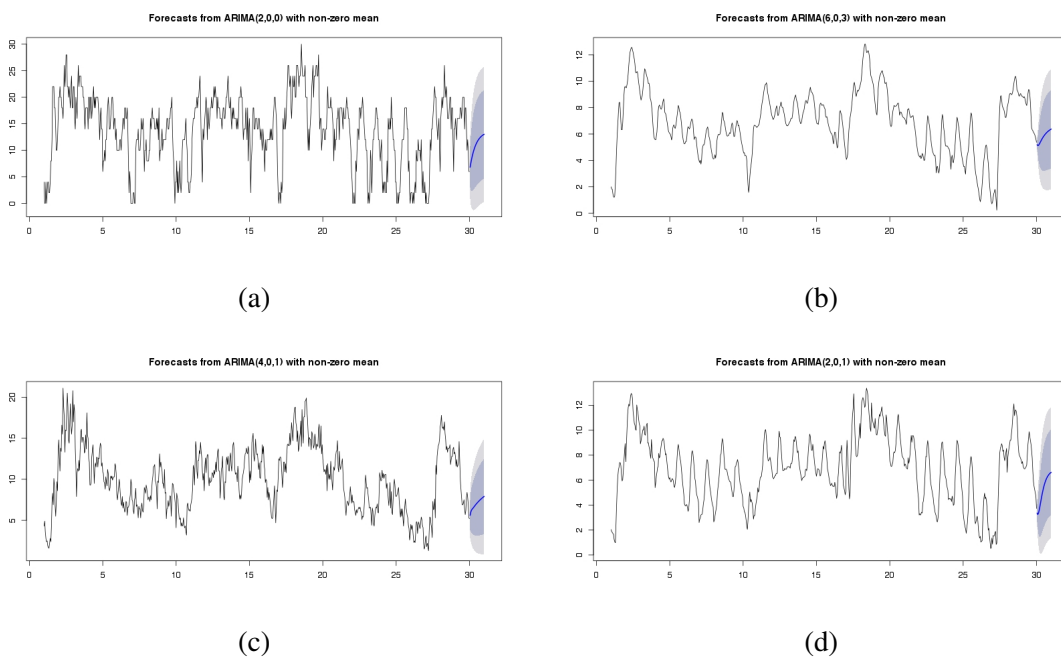


Figure 4.1. a) ADM b) MERRA c) URLA d) CFSR ARIMA 24 hours forecasts for 01-July-2000

Above in Figure 4.1, an initial example of 24 hours forecasts of wind speed for 4 different stations are depicted. All are estimated using an ARIMA model. It is important to mention that these graphs are relevant just as an initial example. We will present our empirical analysis and results at the later stages after explaining our data and models.

4.2.1. SARIMA Model

As we mentioned in previous sub-section ARIMA model has three parametric linear components, AR(p) lagged value, I(d) integration value and MA(q) is the moving average parameter. Adopting the dataset, these parameters are used to estimate the model coefficients. This model can forecast the future value from the previous records with a random error. Inclusion of seasonality component in SARIMA model is expected to rise the forecast accuracy while decreasing the magnitude of errors. This seems plausible as many socio-economic variables as well as wind and meteorological variables are likely to have a seasonal component.

Mostly used seasonal components in SARIMA modelling are listed below.

- i For hourly observations $s=24$ (24 in 1 day)
- ii For daily observations $s=365$ (365 in 1 year)
- iii For quarterly observations; as representatives of autumn,winter,spring,summer $s=4$ (4 in 1 year)
- iv For monthly observations $s=12$ (12 in 1 year)

Now we can define seasonal part of SARIMA model. Non-seasonal version of SARIMA model is a typical ARIMA model. So herein already p,d,q components will be used and additionally P,D,Q values will be added to catch seasonality affect.

SARIMA are ARIMA(p,d,q) models whose residuals α_t are ARIMA(P,D,Q).(Gerolimetto, 2010)

$$\varphi_p(L)\Phi_p B^s \nabla^d \nabla_s^D y_t = \theta_q(B)\Theta_Q(B^s)\alpha_t \quad (4.20)$$

α_t noise component and P seasonal autoregressive component, D seasonal difference component finally Q seasonal moving average parameter ∇_D , Θ , Φ seasonal coefficients.(Ruiz-Aguilar et al., 2014)

The effect of seasonality is very clear in forecasted values

Above in Figure 4.2, an initial example of 24 hours forecasts of wind speed for 4 different stations are depicted. All are estimated using an SARIMA model. It is important

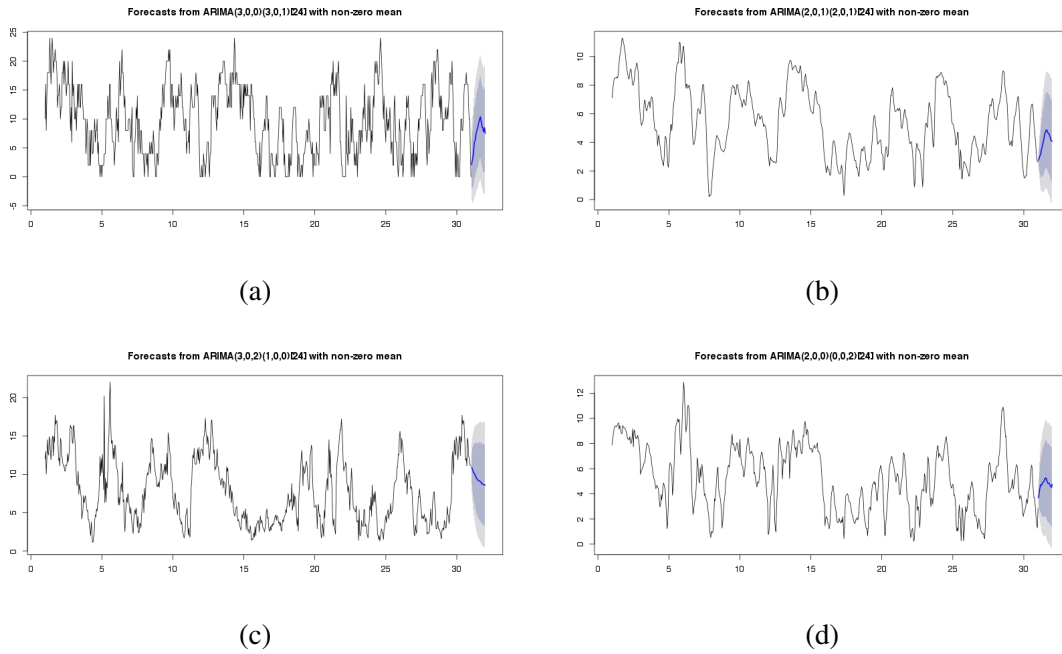


Figure 4.2. a) ADM b) MERRA c) URLA d) CFSR SARIMA 24 hours forecasts for 01-July-2000

to mention that these graphs are relevant just as an initial example. We will present our empirical analysis and results at the later stages after the explaining our data and models.

4.3. Sims Methodology(VAR)

The vector autoregression (VAR) model is widely used and suggested methodology for the modelling multivariate time series. It is an improved version of the univariate autoregressive model that captures dynamic multivariate time series characteristics.

Sims (1980) published his studies on vector autoregressive (VAR) models.(Sims, 1980) His aim was to explain the dynamic structure of the variables. So in this model, every variables affects one another and with time lags. There are two types of variables. First, endogenous variables; which affect other variables and affected by other variables, second, exogenous variables, which affect variables but unaffected from any variable. (Gujarati, 2012)

The differences between ARIMA and SARIMA are below summarised.

- i) In (S)ARIMA models we only the derive the actual value from past values. (univariate case)
- ii) VAR models give us chance to investigate exogenous variables affects
- iii) VAR models give us chance to see dynamic relationships between variables.
- iv) In VAR methodology value is explained by its own history and by the history of other variables as well. (multivariate case)

(Gujarati, 2012)

4.3.1. Terminology, Notation and General Assumptions

4.3.1.1. General Definition

Definition of main equation based on $Y_t = (y_{1t}, y_{2t}, \dots, y_{nt})$ 0. It denotes an (nx1) vector of time series variables. The first order vector autoregressive VAR(1) model with 2 endogenous variables defined as $y_{1,t}, y_{2,t}$

$$y_{1,t} = \alpha_{1,1}y_{1,t-1} + \alpha_{1,2}y_{2,t-1} + \epsilon_{1,t} \quad (4.21)$$

$$y_{2,t} = \alpha_{2,1}y_{1,t-1} + \alpha_{2,2}y_{2,t-1} + \epsilon_{2,t} \quad (4.22)$$

Assumption about the error terms is that expected residuals have a zero mean and not autocorrelated.(Gujarati, 2012)

4.3.1.2. Matrix Notation

$$y_t = A_1 y_{t-1} + \epsilon_t \quad (4.23)$$

$$A_1 = \begin{pmatrix} \alpha_{1,1} & \alpha_{1,2} \\ \alpha_{2,1} & \alpha_{2,2} \end{pmatrix} \quad (4.24)$$

for example VAR(2) model;

$$\begin{pmatrix} y_{1,t} \\ y_{2,t} \end{pmatrix} = \begin{pmatrix} \alpha_{1,1} & \alpha_{1,2} \\ \alpha_{2,1} & \alpha_{2,2} \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{pmatrix} \quad (4.25)$$

VAR coefficients can be calculated with Cramer or Inverse Matrix rule.

In literature (Baum, 2013) different types of VAR models could be find.

- i) Recursive VARs: based on construction the error term in each regression to be uncorrelated with the error in the preceding equations.
- ii) Reduced VARs: based on assumption of neglecting contemporaneous correlation between the variables
- iii) Structural VARs: on the contrary of reduced VARs, here the basement is to organise the contemporaneous relationships between the variables

- iv) Vector Error Correction Model VECM: special case solution for not covariance stationary variables

4.4. Accuracy Measurement Tools

After having defined the time series models that we would like to use in our analysis, now, a few words on how we can compare their accuracy and estimating performances should be mentioned. In the literature, there are three main criteria for the comparison of models' estimate accuracy. The First one is RMSE (Root Mean Square Error) (Hyndman and Koehler, 2006):

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (\hat{W}S_t - WS_t)^2}{n}}. \quad (4.26)$$

where WS is the actual value of wind speed and $\hat{W}S_t$ is the predicted one. So, it represents the magnitude of residuals per observation. Second criterion is MASE (Mean Absolute Scaled Error): which takes the following form (Hyndman and Koehler, 2006):

$$MASE = \frac{\sum_{t=1}^n |\hat{W}S_t - WS_t|}{\frac{n}{n-1} \sum_{t=2}^n |\hat{W}S_t - WS_t|}. \quad (4.27)$$

Lastly, the third criterion is MAE (Mean Absolute Error) (Hyndman and Koehler, 2006):

$$MAE = \frac{1}{n} \sum_{t=1}^n |\hat{W}S_t - WS_t|. \quad (4.28)$$

which represents the sum of total error terms in absolute values divided into the number of observations.

All three criteria summarize actually the magnitude of forecast errors. Hence, it means that the models with lowest values have lower forecast errors and higher predicting power.

CHAPTER 5

THE WIND SPEED FORECASTING LITERATURE REVIEW

Weather forecasting is a complex and difficult study field. However, estimation of wind characteristics is more challenging than temperature or humidity estimation. The main features of the wind are referred to as its direction and magnitude.

Forecasted wind characteristics are crucial in various fields like structure, transportation, fishing, power production etc. For example to construct a skyscrapers, there is a need for crane rent and crane usage has certain regulations related with weather conditions and wind speed levels.

Indeed, wind speed prediction is the main parameter for wind energy power plants for the control, scheduling, maintenance, transportation of wind turbine parts and blades, erection of turbine tower. As an example wind turbine blades can reach up to 63 meters (Enercon, 2014). So, transportation of this long parts and assembly of blades to nacelle must be done at proper weather conditions which are defined according to specific regulations.

In fact power production issue has high priority among all problems due to intermittency of wind speed. In the field of power systems, a traditional generator is known as dispatchable which means ability to provide power regardless of weather conditions. Wind power production is, instead, famous with non dispatchable characteristics. (Foley et al., 2012)

The literature on wind speed and power forecasting applications, methodologies and their classifications are very well summarized in a study by Soman.(Soman et al., 2010) The Reader can find the details of the classification procedure in time zone; very short term, short term, medium term and long term forecasting applications related with engineering applications like grid security, wind turbine maintenance schedule, electricity market price estimations in power industry.

Short-term wind speed forecasting can be implemented over different time frequencies. Usually, hourly forecasts of expected winds are helpful in dispatching decision-making, daily forecasts of hourly winds are useful for the load scheduling strategy, and weekly forecasts of day-to-day winds greatly facilitate maintenance scheduling.

To handle wind speed prediction, many methods have been developed. Mainly;

- (i) physical models (Numerical Weather Forecasting)
- (ii) spatial correlation models (Meso-Scale Models)
- (iii) conventional statistical models (Box-Jenkins, Sims Methodology etc.)
- (iv) artificial intelligence and new models (Artificial Neural Networks, Fuzzy logic etc.)

More specifically, one may partition these approaches into three groups. The first one incorporates physical factors, such as terrain, mountains and obstacles in wind speed predictions. Generally, these methods have been used for large-scale places like regions or countries and they are unlikely to fit to the local sites. Second group consists of Artificial Neural Network (ANN) methods which are adopted by a strand of scholars in order to model the wind speed. The main advantage of ANN is that it does not require an assumption of a certain form of function, thus, represents a more general approach in that sense. However, the drawback is the weakness of forecast accuracy compared to the other methodologies.

The third group, perhaps the most influential, includes conventional statistical time series methods which rely mostly on ARMA (Autoregressive Moving Average) and ARIMA (Autoregressive Integrated Moving Average) based models. ARIMA has, initially been introduced by Box-Jenkins. It is a useful model in incorporating the persistence, variance and stationarity (order of integration) of wind in predictions. In addition, more recently developed SARIMA model takes into account the seasonality component as well. It is argued to outperform the former models and provide more accurate results. Some examples of these studies are Bivona et al. (2011) who perform a wind speed forecasting in Italy (Camarata Mazzarrone) between 1st Jan 2003 to 31st December 2006 by employing SARIMA and ANN, Cadenos and Rivera (2007) who predict the wind speed in Oaxaca (Mexico) between June 1994- May 2000 and conclude that SARIMA outperforms the ANN, Pedro et al (2012) who perform similar analysis in Iberia and concludes that ARMA based model produces more accurate results.

Soman et al. (2010) have summarized well the statistical forecasting family, physical and hybrid approaches. In their study, a list of methodology starts with the naive and persistence models which are considered as reference methodologies. Indeed the new trend is to use hybrid methodologies in the literature for wind speed forecasting. (Chang, 2014) Main idea is to increase the accuracy of models and decreasing time consumption.

The hybrid systems can be in form of combinations of physical and statistical models or artificial intelligence and statistical models etc. Numerical weather prediction models can take into account the geographical factors and prediction methods are classified from global to local scales.

Foley, Aoife M.(2012) stated that one of the main problems of current numerical weather prediction models is their design environment assumption which is based on onshore site. The outcomes of these models deviate from real data for offshore site errors become more serious. Another problem is the poor availability of recorded data. (Foley et al., 2012)

In current thesis, our selected family is conventional statistical models. Several papers have been published on this issue. We summarize their main findings and methodology below.

Firstly for the statistical models we need to handle with stationarity as a common obligation for Box-Jenkins methodology.(Arwade and Gioffre, 2014) But Sanjay R. Arwade and Massimiliano Gioffre presented a study that is based on statistic tests to control the assumption of stationarity of wind speed. Their key finding is that wind speed stationarity assumption is valid for one week duration and for monthly data it causes more serious prediction errors. However, in the end just the division of two period, roughly winter and summer seasons the wind speed shows consistent characteristics.

Garcia Torres have implemented a comparison of models based on measured wind speed within 10 minutes intervals in Navarre Spain at five different station. Herein authors have adopted two methods; persistence model and ARMA. As an accuracy measurement tool root mean square error has been used. Results are obtained from five different cases and ARMA was found to outperform the persistence methodology. Also the lowest RMSE values of forecasts are corresponded to the highest measured wind speeds. to forecast hourly averaged wind speed.(Torres et al., 2005)

Another study is written by Gomes and Castro which is very valuable study and compares the AutoRegressive Moving Average (ARMA) and Artificial Neural Networks (ANN) methodologies.(Gomes and Castro, 2012). The details of short term power forecasting can be found in this study. Fundamental tools; persistence, ARMA and ANN are employed by the researcher for the comparison cases. MAE , MRE and RMSE accuracy measurement tools approve that ARMA and ANN methodology outperforms reference “persistence” methodology The main result is that ARMA and ANN; Autoregressive models can achieve to reach smaller errors. However, their process period is

longer than ANN models.

ARMA models are powerful accepted methodology but one of the weakness of them is the fact that it simulates seasonal effects especially when the subject is wind, season, day and night time.

Indeed (Cadenas and Rivera, 2007) inspected this effect with the monthly period in Mexico case. Cadenas, Erasmo have selected SARIMA and ADALINE which is the one of the ANN tools to investigate seasonal effects. MAE, MSE and MAPE have been adopted. SARIMA methodology showed lower statistical errors than ADALINE.

Another important study is implemented by Bivona et al. (2011) which is an empirical article. Here SARIMA and ANN methodologies are investigated with R and Java environments. Authors pointed out as a focusing concern creation of a general class of stochastic models for hourly wind speed.

Unknown parameters and requirement of large data sets are the challenging main problems. So, authors employed two considerable data sets for 4 years from Sicily Italy, after that with the derivation of deterministic model they reached to stochastic model. Using SARIMA methodology, they took into consideration diurnal seasonality and as performance criteria they used MAE, MASE, MAPE.

Extension of auto regressive methods to multivariate space is called vector auto regressive (VAR). It is supposed to have a greater forecasting ability as it considers a range of variables rather than just 1 variable. Erdem and Shi (2011) have shown this in comparison to ARMA-linked models.

Another study in which the wind speed estimation using VAR is done by Ewing et al. (2007) In this study, the author used only wind speed measurements which are recorded at different heights. So, estimation of wind speed and turbulence are done with the consideration of vertical wind profile. .

Our way of thinking is consistent with VAR model and its proponent. We argue that using multivariate models may help increasing the forecast accuracy of wind speed. Indeed, there exists many meteorological variables that might play a significant role in affecting wind speed dynamics.

Forecasting accuracy is the main parameter to validate studies with these purpose accuracy of forecasting is given in four section by Jung, Jaesung. (Jung and Broadwater, 2014). According to his study accuracy should be investigated with the terms of accuracy by forecast time horizon and as a rule when the prediction length increases, accuracy of forecast decreases, accuracy by location complexity, terrain complexity is another param-

eter decreases accuracy comparatively flat sites, accuracy by seasonal variability herein the researcher claims winter months have smaller forecasting errors comparatively summer months, accuracy by weather condition in another words weather stability which is evaluated by meteorological risk index (MRI), has a linear relationship with forecasting accuracy when MRI increases error of forecasts increase.

CHAPTER 6

DATA

In this chapter reader can find the location information and collection procedure of taken datasets.

6.1. Location of Study: IZMIR Bay

Izmir is the third biggest city of Turkey. It is famous as an international port city, due to its location in the western part of Anatolia along Aegean sea. Climatology of Izmir is classified as Mediterranean climate. In other words long, hot and dry summers and rainy, mild cool and short winters characterize the climate.

Izmir's important physical features are determined by Gediz river(second biggest river in Anatolia) and Gulf of Aegean sea (between the peninsula of Karaburun and Foca) Region has a quite hilly topography.

Two important circulations are valid in Izmir Bay. First one is horizontal and vertical mixing and advection between mainland (Anatolia) and sea (Aegean). This circulation refers to as wind-driven circulation. The second one is the thermoline-driven or baroclinic circulation.(Sayın, 2003)

Shape of Izmir Bay seems like roughly L letter. Region must be investigated with three main sub-classes, these are Outer, Middle and Inner Bay. Our data set is taken from Gulbahce Bay and Karaburun Peninsula where is located in Outer Bay.

6.2. Data

To begin with the detailed description of data measurement, we provide in table 6.1 main properties of 4 data sources (locations) and variables with their units.

In the first data type, obtained from URLA RES EGENDA Company, the wind speed, direction, temperature, pressure and relative humidity were measured within the campus area of Izmir Institute of Technology (IZTECH) in Urla which is a sub-province

Table 6.1. Model Progress Data Sets Range 06,2000-10,2001

Data Set	Wind Speed	Wind Dir.	Pressure	Temperature	Humidity	Radiation	Disp. Height	Heat Flux
Urla Res	(m/s)	(deg)	(hPa)	(degC)	% RH	(x)	(x)	(x)
CFSR	(m/s)	(deg)	(x)	(degC)	g/kg	W/m ²	(x)	W/m ²
MERRA	(m/s)	(deg)	(hPa)	(degC)	(x)	(x)	(m)	(x)
ADM	(m/s)	(deg)	(hPa)	(degC)	% RH	(x)	(x)	(x)

of Izmir. The met mast is located within the campus area and the data measurement has been implemented at two different heights; 30 meters and 10 meters, specifically sensors which are located at 10 meters measure the direction and 30 meters anemometers (THIES 4) measure the wind speed. This data is recorded in 10 minutes intervals. For the sake of tractability, we create hourly average of wind speeds as input data.

The second dataset is taken from Izmir Adnan Menderes Airport Meteorology Station. Measurements are done at 10 meter height. Relative humidity, pressure, temperature and wind speed measurements (hourly) are used for our research.

For detailed comparison and validation of our analysis, we employed also MERRA and CFSR datasets. MERRA (Modern ERA RETROSPECTIVE-Analysis for Research and Applications), CFSR (Climate Forecast System Reanalysis) are obtained from Wind-Pro program (EMD) . MERRA dataset gave us the opportunity to create a multivariable set which consist of pressure, wind speed, wind direction, temperature and displacement height variables.

On the other hand CFSR dataset includes the following variables; wind speed, wind direction, temperature, cloud cover, solar radiation, humidity and heat flux. Both of them are available as hourly averaged wind speeds.

As mentioned before, our period of analysis runs from 2000 June to 2001 October on an hourly basis. Although, some of the wind speed data is available for a longer period, we lack some other variables and that's why, we resort only to 2000 June-2011 October period.

With regard to the geographical and climate characteristics of Urla, it is located along the Aegean Sea and in a quite mountainous place. It has recently attracted attention and known as a convenient place for wind energy harvesting. To support this visually, we present in Figure 6.1 the map of wind speed in Turkey.

It is seen at a glance that Urla belongs to the orange layer which represents the

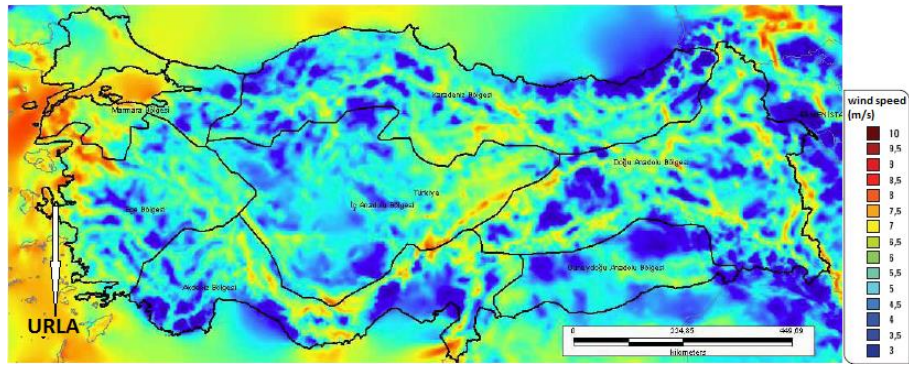


Figure 6.1. Wind Speed Map of Turkey (Source:energy.gov.tr,2014)

second high-potential group for wind energy and has the average wind speed about 7.5 meters/second (measured at 50 meters height). In addition, the campus area of IZTECH is likely to exhibit even higher values as it is located towards the main wind direction (north).

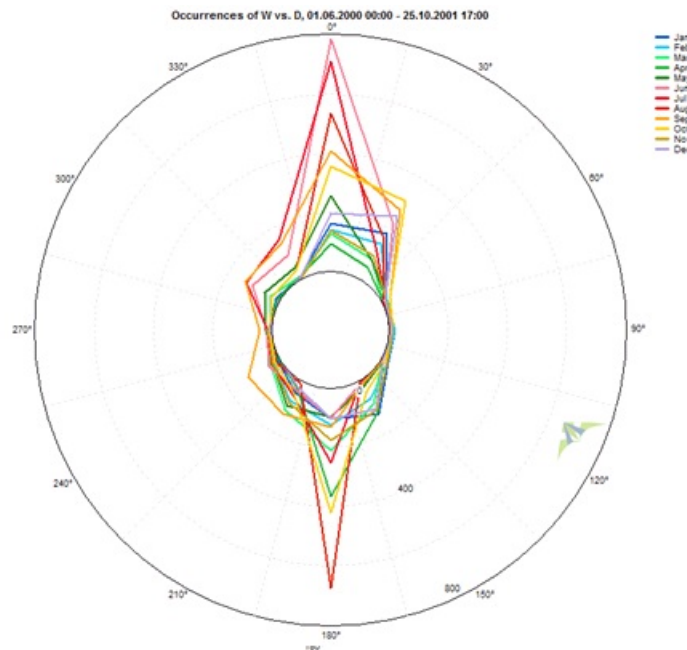


Figure 6.2. Run plot for wind direction from 1 June 2000 to 25 October 2001 (windographer software tool is employed)

To provide more information on the wind characteristics of Urla, we present a related figure 6.2 that captures the main wind direction during different months (illustrated using windographer software) We observe that the two main wind directions are north

and south. The wind blowing from north dominates during the summer times, especially during June, July, August, September and October months. The wind blowing from South is mostly observed during June, May and April. The east-west direction is quite rarely observed.



Figure 6.3. Data set locations (Obtained from WindPro program)

In Figure 6.3 above, the data-set locations are depicted where URLA-RES is just within the IZTECH campus area, CFSR is in the south, MERRA is in the north and ADM is in the east of URLA-RES.

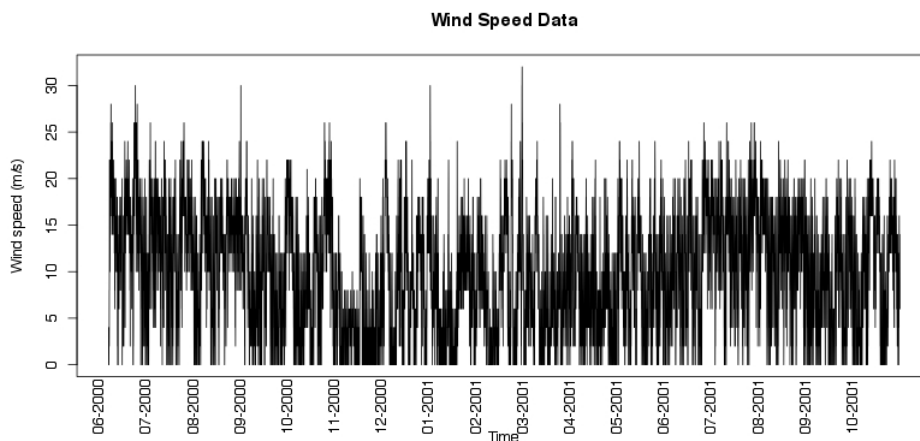


Figure 6.4. Time series plot for Wind Speed ADM

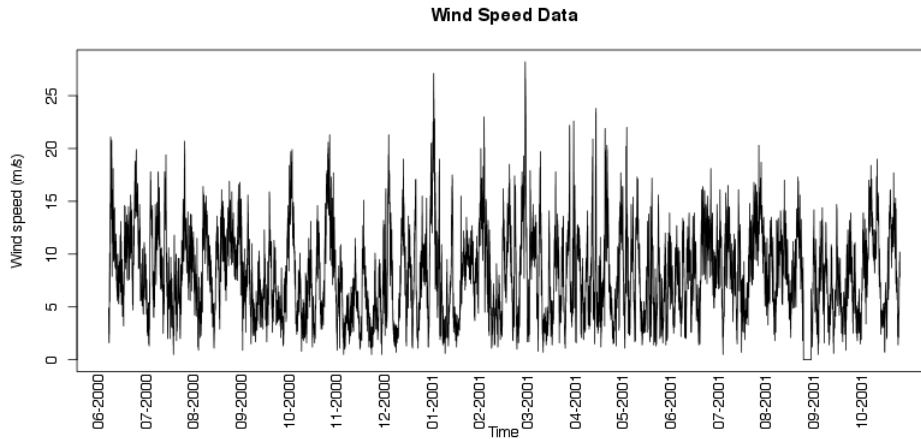


Figure 6.5. Time series plot for Wind Speed URLA RES

We think it is useful to illustrate the time series plots of our variables in order to have an initial idea about them.

To begin with the ADM case, temperature (in Celcius degrees), wind speed (meter-s/second), humidity (% RH) and pressure data (hPA) clearly follow a stationary process with constant mean and variance. Quite similarly, the same variables for URLA-RES dataset also seem to follow a stationary process, exhibiting mean reversion in the level and constant variance.

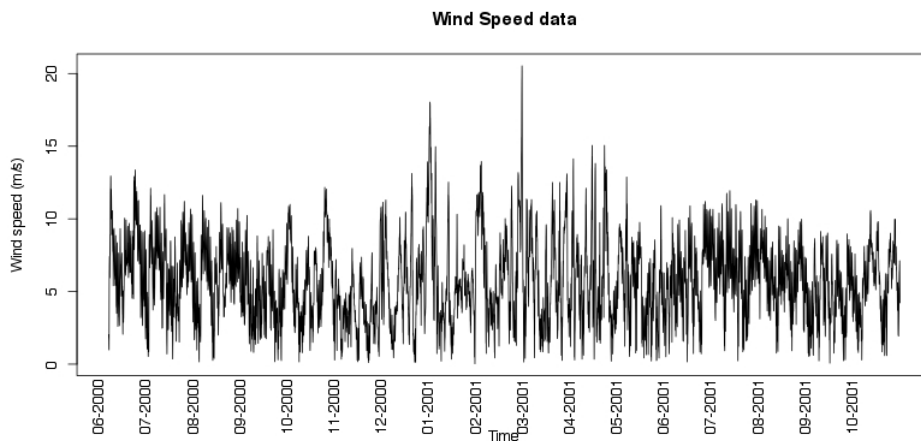


Figure 6.6. Time series plot for Wind Speed CFSR

In CFSR case the available three endogenous variables, temperature, wind speed and humidity, and in MERRA data, wind speed, temperature and displacement height data

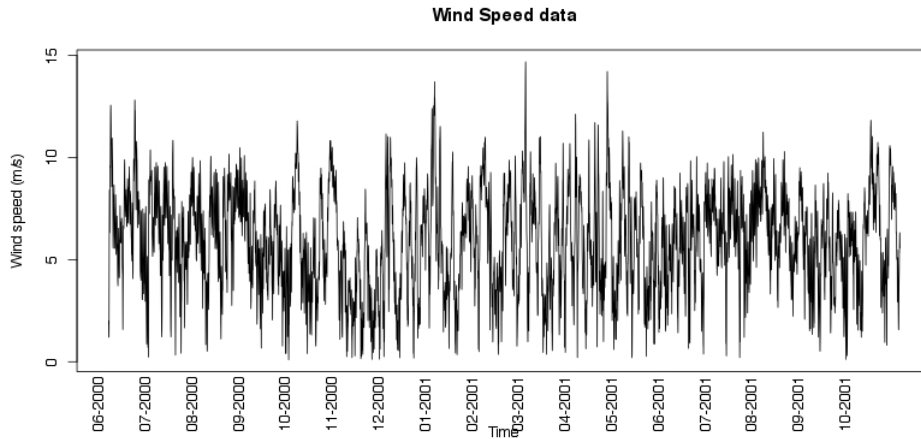


Figure 6.7. Time series plot for Wind Speed MERRA

(displacement height is a variable of special tree heights located around the mat mast and which might affect vertical wind speed profile) seem to exhibit desired stationary evolutions with a constant mean and variance. Hence, all variables have convenient evolution for our analysis.

To have a better understanding of the wind speed characteristics of our data set, one may compare the wind speed values obtained from 3 data-set with the meteorological mast of Urla-Res. As a measure of closeness, lower values of RMSE indicate more similar values of wind speed. So, having calculated it, MERRA seems to exhibit the most similar wind speed data (with a RMSE score of 4,7743) followed by ADM and CFSR which takes the respective scores (5,14046 and 6,8842). This information is important to know accuracy power of data sets like MERRA and CFSR which are taken from WindPro program database. In general they are taken as main input data in the calculation of annual energy production estimation of wind energy power plants.

CHAPTER 7

MODEL SPECIFICATIONS AND RESULTS

Now I will introduce main parameters and statistical models which are used during this thesis study.

7.1. Set of Variables, Selection of Model Parameters and AIC

Criteria

With regard to the variables employed in models. IN SARIMA adn ARIMA models, only wind speed has been used. In VAR model, however, the following set of variables are adopted:

$$X_{urlares}=endogenous : P, W, D, H; exogenous : T$$

$$X_{merra}=endogenous : P, W, D; exogenous : T, Hd$$

$$X_{cfsr}=endogenous : F, W, D, H; exogenous : T, R$$

$$X_{adm}=endogenous : P, W, D, H; exogenous : T$$

Where W is the wind speed (meters/second). T is the ambient temperature (in Celcius-degrees), H is the level of humidity, P is the air pressure and D represents the wind direction, specifically the deviation of wind from north, measured in degrees. Hd is the displacement height data as explained before. F denotes the Heat Flux and R is the Solar radiation. As mentioned before, all variables are measured on hourly basis and for the period June 2000- October 2001.

In terms of determination of model parameters, We choose the p,d,q values for ARIMA and SARIMA according to Akaike Information Criterion(AIC). Although there exist other criteria in the literature that are used for model comparison, such as BIC (Bayesian Information Criterion) or HQ (Hannan-Quinn), AIC is known as one of the

most widely accepted one. For this reason, we prefer adopting it. AIC is based on the following formula.(Akaike, 1974)

$$AIC = \ln(\hat{\sigma}_\varepsilon^2) + \frac{2k}{T} \quad (7.1)$$

Where k is the number of parameters estimated and T is the number of observations, $\hat{\sigma}_\varepsilon^2$ is the penalty parameter and the estimated variance. The lower values of AIC indicate greater estimation accuracy for the model. In modelling, we select the parameters that minimize the AIC for each month.

- 1)ARIMA=auto.arima(...seasonal=FALSE,ic=c("aic"))
- 2)SARIMA=auto.arima(...seasonal=TRUE,ic=c("aic"))
- 3)VAR=VAR(...lag.max = 10, ic = c("AIC"))

7.2. Estimation Results

To begin with the results, for all stations, we perform the estimations for each month. So, there are 17 estimations in total as there are 17 months in dataset. (from June-2000 to October-2001). The data is in the form of hourly estimation.

We estimate the three models and present the fitted values versus actual data. Figures from 7.1 to 7.16 present these results for 4 different data-stations and 3 models. Within the graphs, red dots represent the observed data, black-crosses represent the fitted values from ARIMA model, purple crosses represent the fitted values from SARIMA and green-crosses represent the fitted values from VAR model.

Figures from 7.1 to 7.4 illustrate the results for ADM case. For the summer months, all the three models seem to exhibit a fair match with raw data.

However, at the beginning of autumn, the accuracy tends to weaken. Actually September and April are the most problematic months in which sudden and unexpected jumps in wind can be observed.

Figures 7.5 to 7.8 presents the results for URLA-RES case in which the wind was measured at mat mast. For all months, the three models seem to perform well in fitting the real data.

Rest of the graphs from 7.9 to 7.12 represents the fitted values for MERRA (Figures 7.9- 7.12) and CFSR (Figures 7.13-7.16) datasets both of which were obtained from

WindPro program. Although in both cases models provide a quite powerful match with the actual data, CFSR seems to provide slightly better estimates of wind speed.

Overall, an interesting feature appears to emerge from graphs. All models provide a reasonably good fit to the actual data. In other words, the models we run produce low levels of forecast errors as the actual data overlaps well with the fitted data.

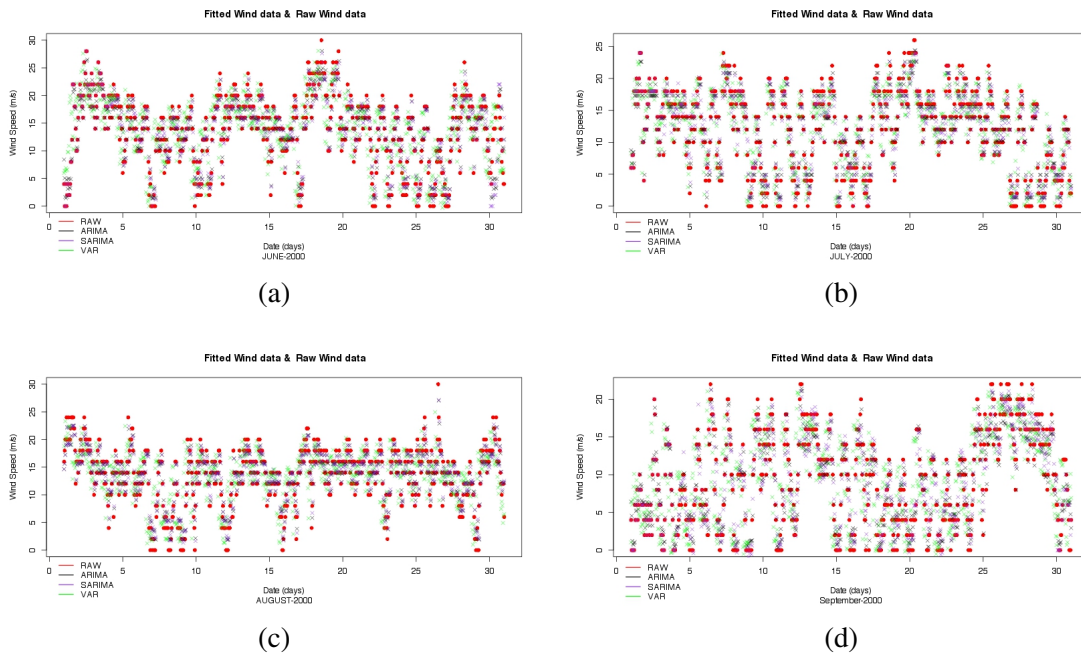
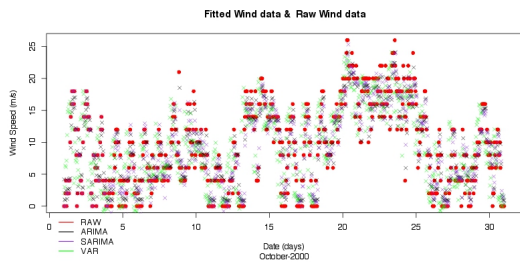
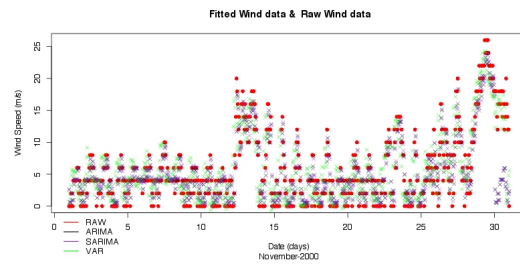


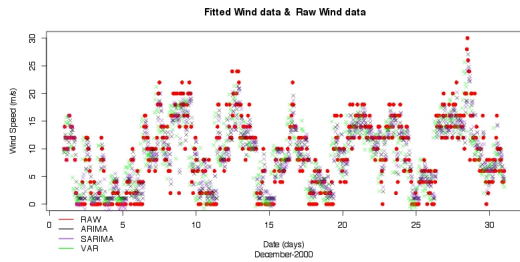
Figure 7.1. ADM CASE Run plot of Actual and Fitted values (1)



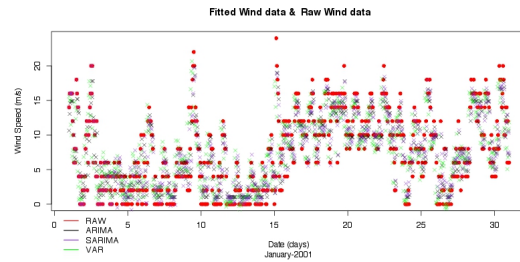
(a)



(b)

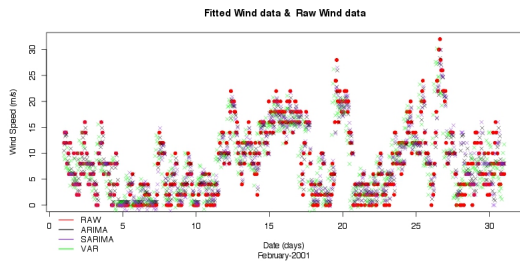


(c)

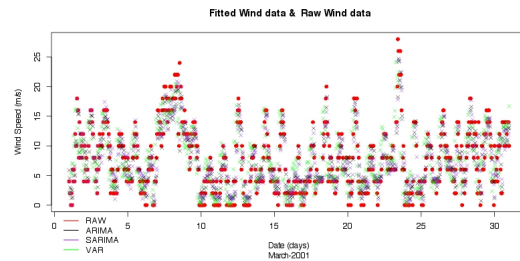


(d)

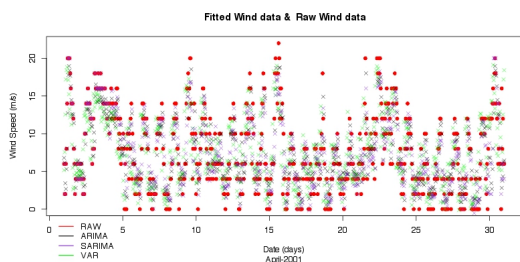
Figure 7.2. ADM CASE Run plot of Actual and Fitted values (2)



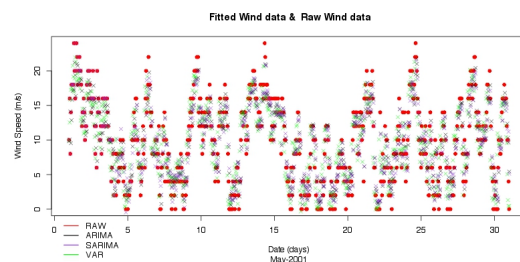
(a)



(b)



(c)



(d)

Figure 7.3. ADM CASE Run plot of Actual and Fitted values (3)

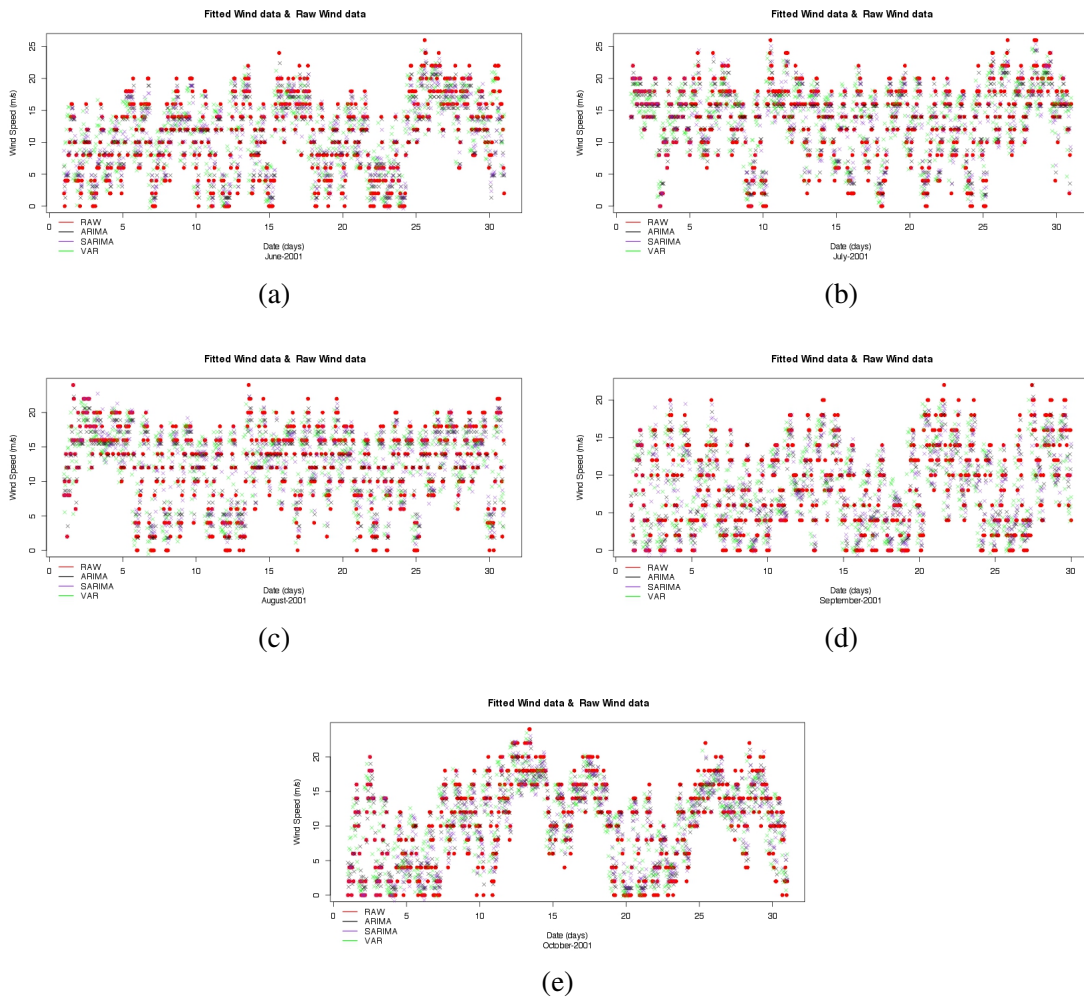


Figure 7.4. ADM CASE Run plot of Actual and Fitted values (4)

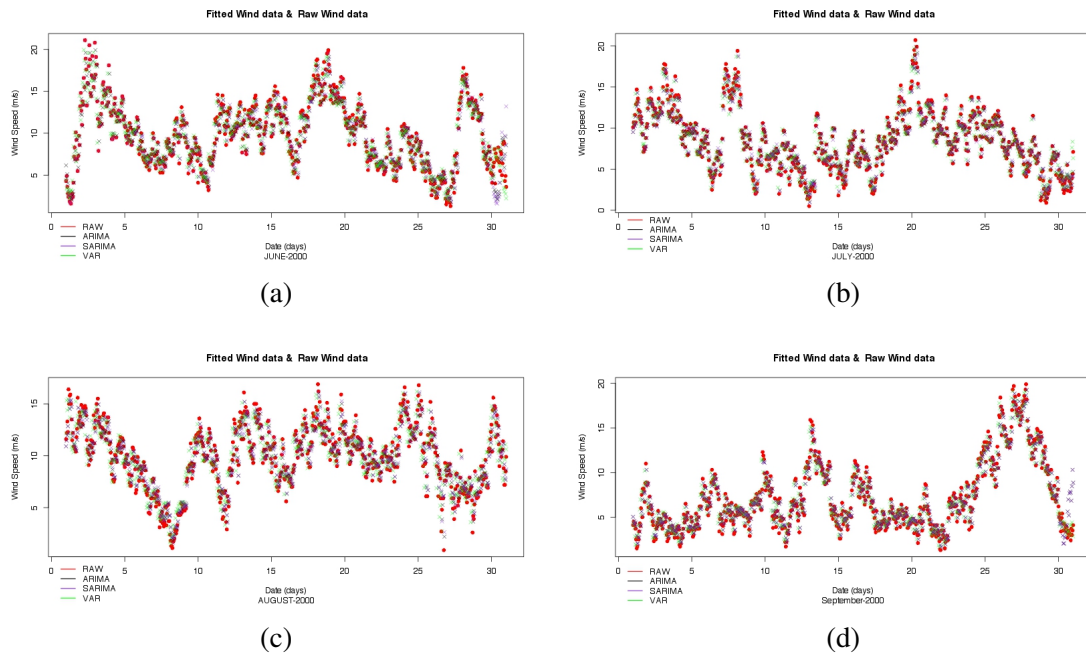


Figure 7.5. URLA CASE Run plot of Actual and Fitted values (1)

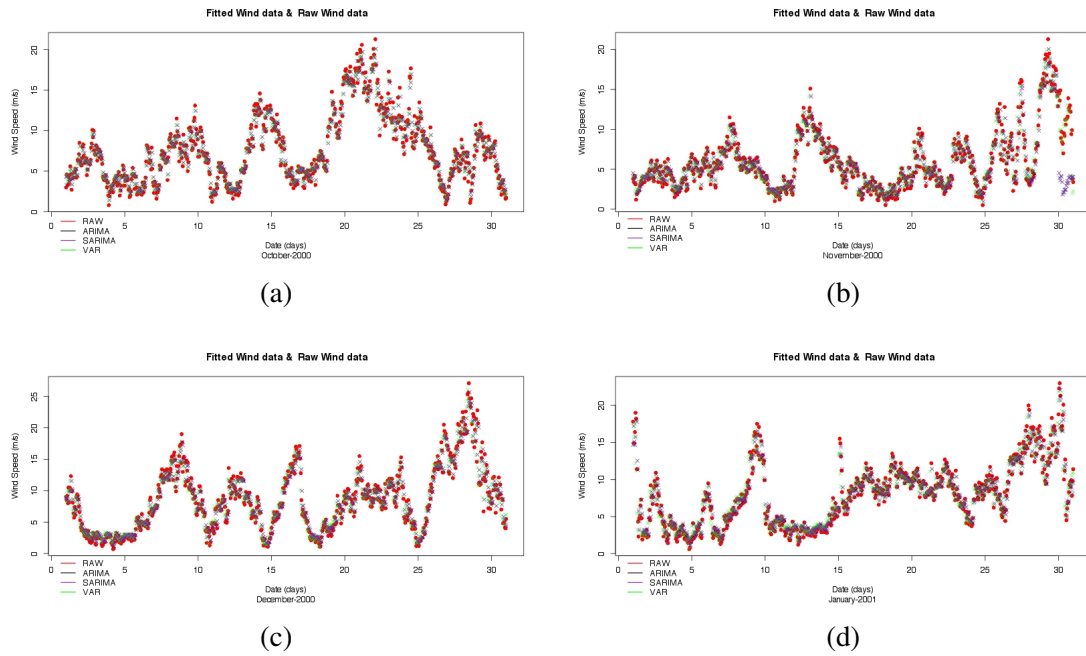


Figure 7.6. URLA CASE Run plot of Actual and Fitted values (2)

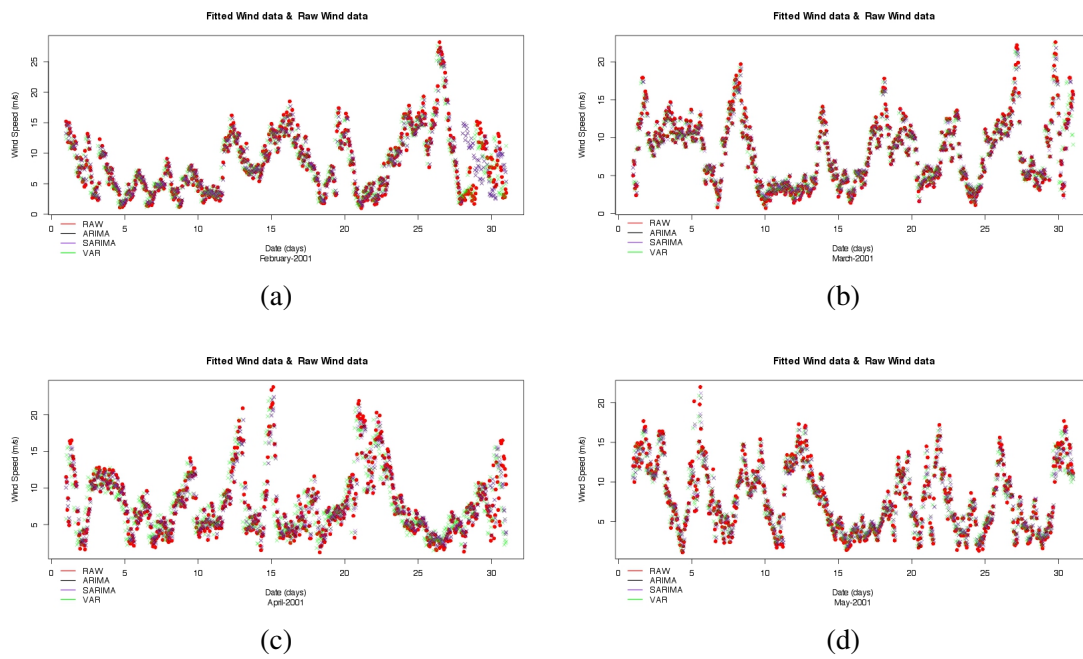


Figure 7.7. URLA CASE Run plot of Actual and Fitted values (3)

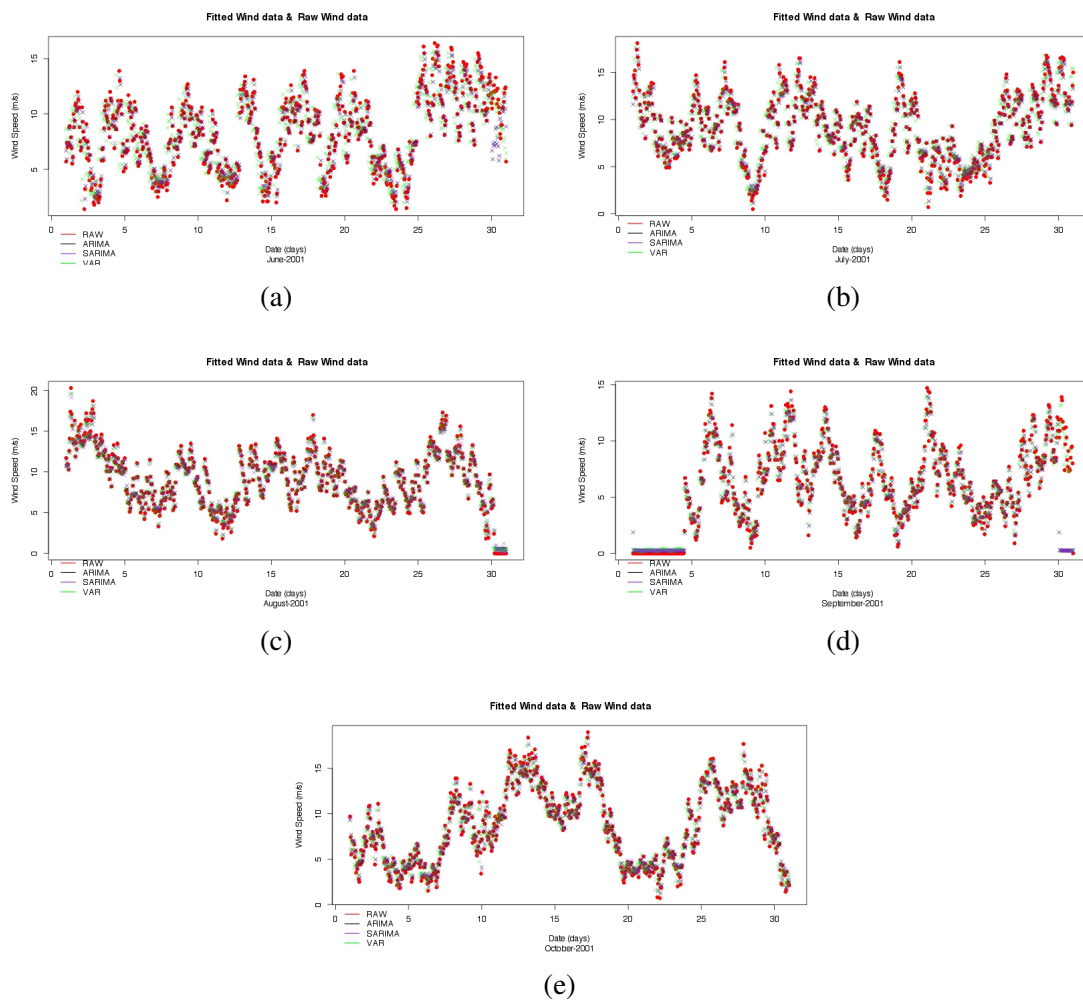


Figure 7.8. URLA CASE Run plot of Actual and Fitted values (4)

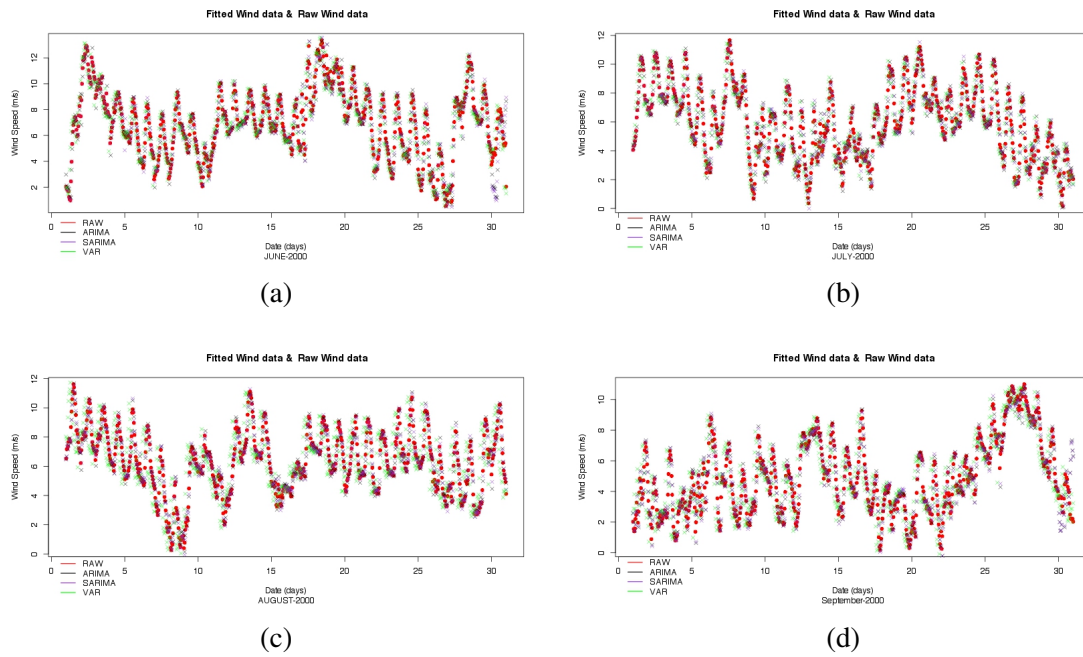


Figure 7.9. CFSR CASE Run plot of Actual and Fitted values (1)

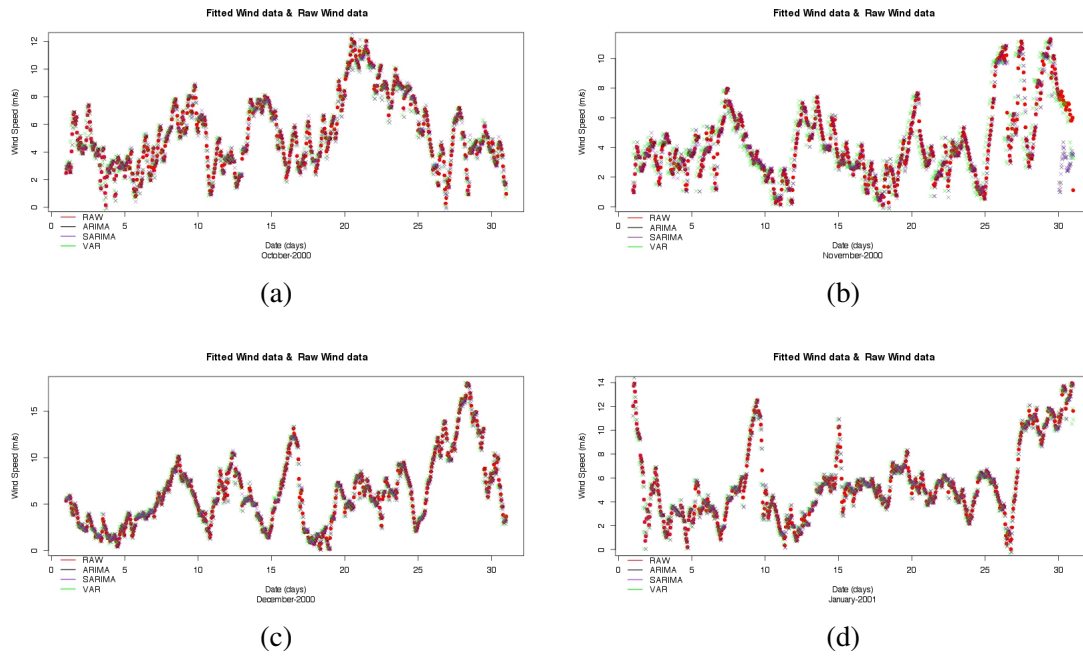


Figure 7.10. CFSR CASE Run plot of Actual and Fitted values (2)

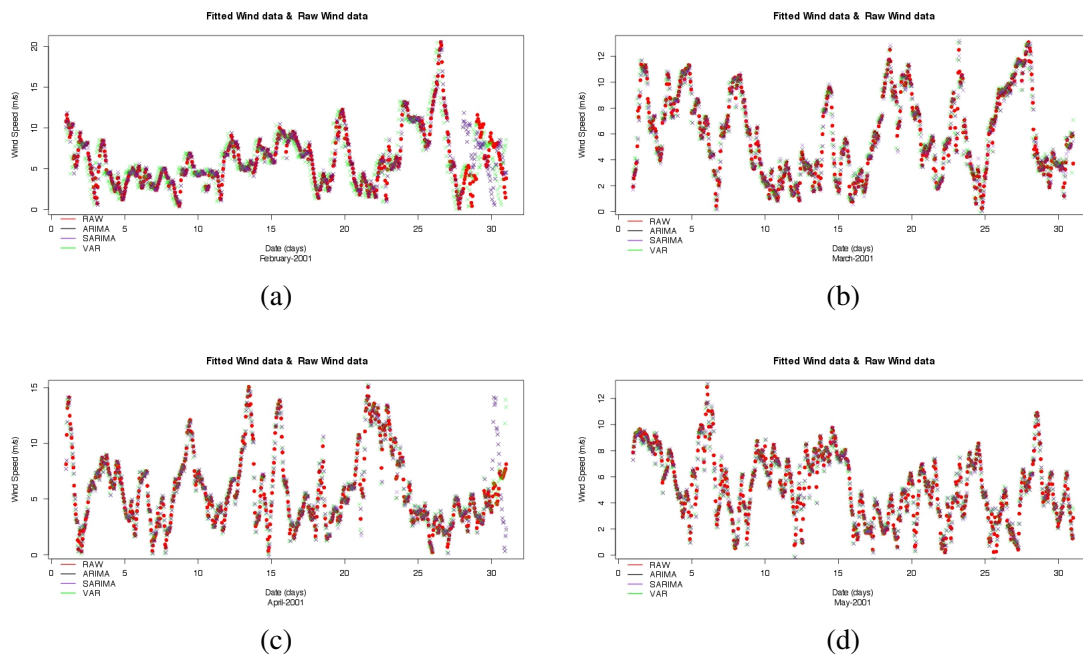


Figure 7.11. CFSR CASE Run plot of Actual and Fitted values (3)

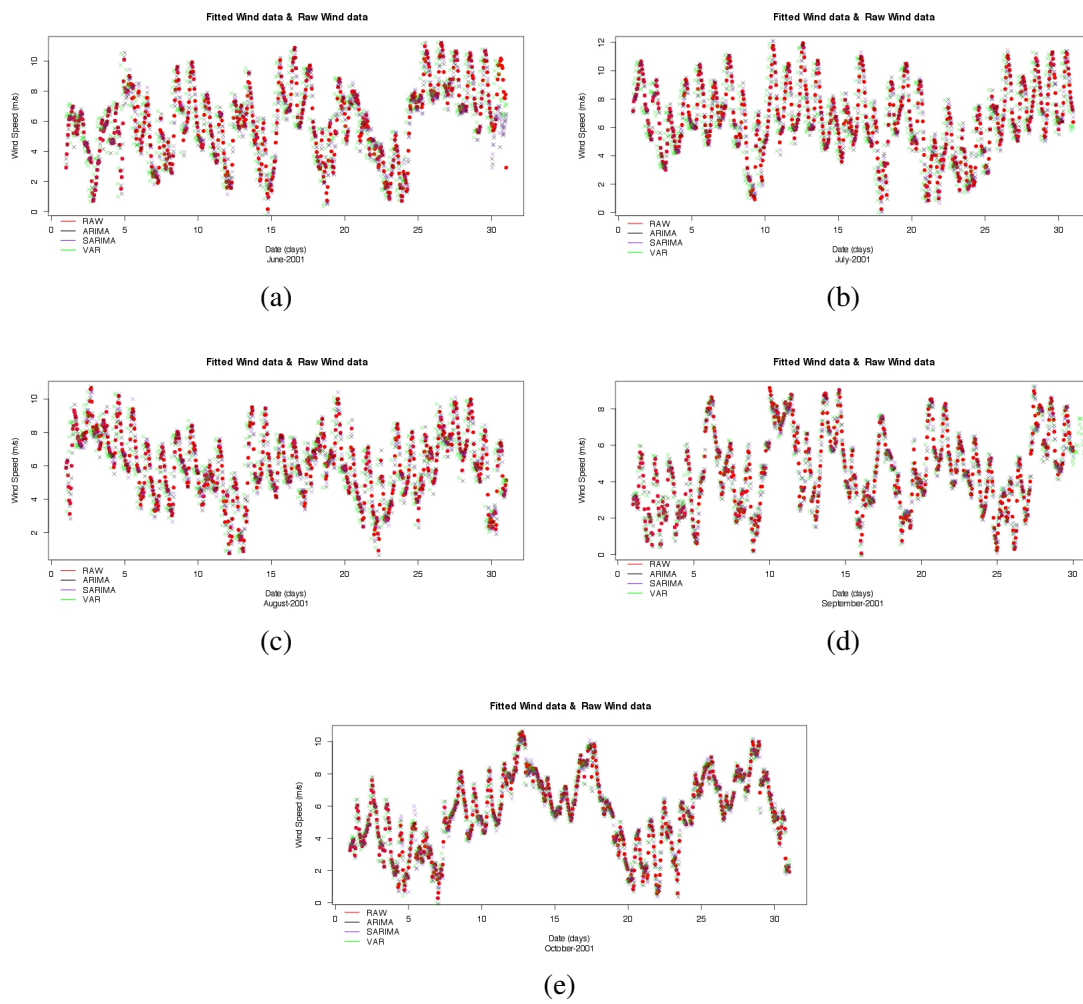


Figure 7.12. CFSR CASE Run plot of Actual and Fitted values (4)

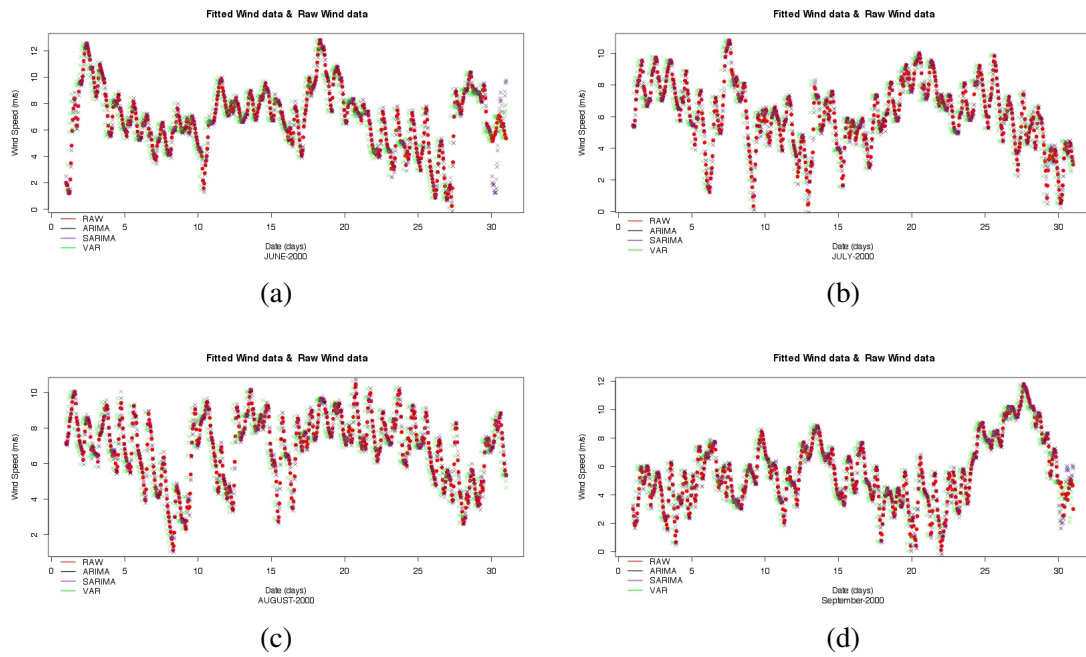


Figure 7.13. MERRA CASE Run plot of Actual and Fitted values (1)

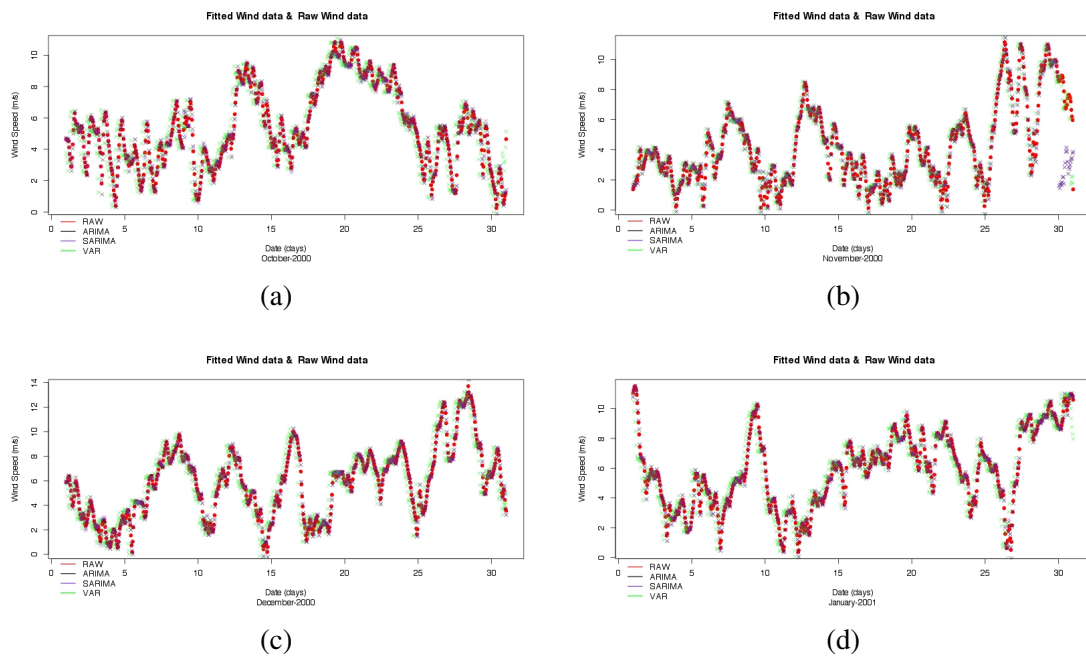


Figure 7.14. MERRA CASE Run plot of Actual and Fitted values (2)

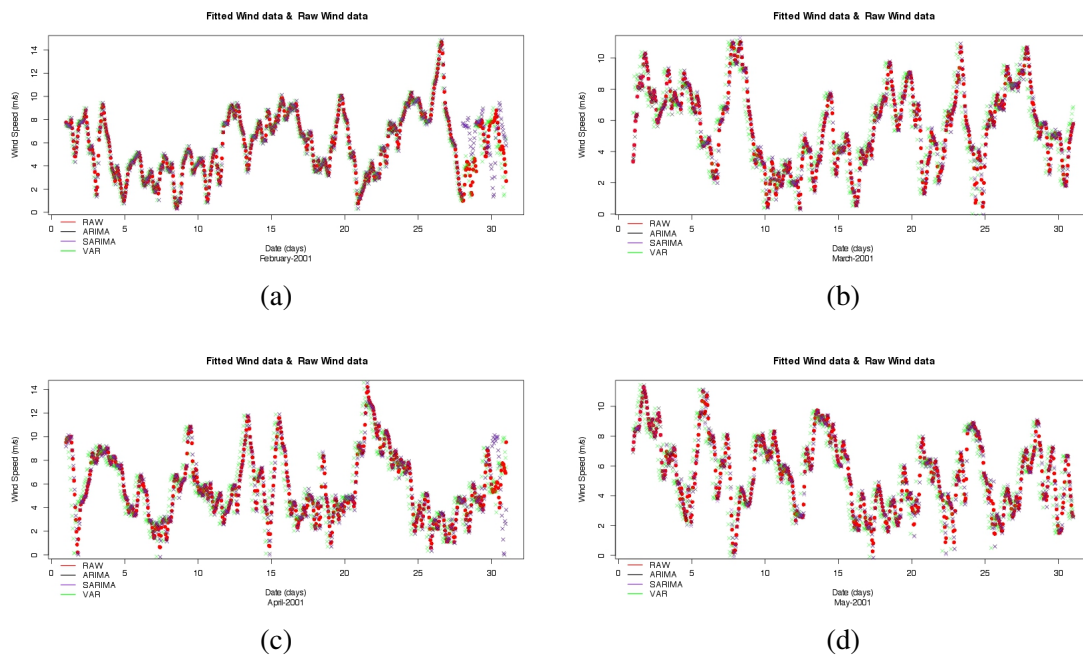


Figure 7.15. MERRA CASE Run plot of Actual and Fitted values (3)

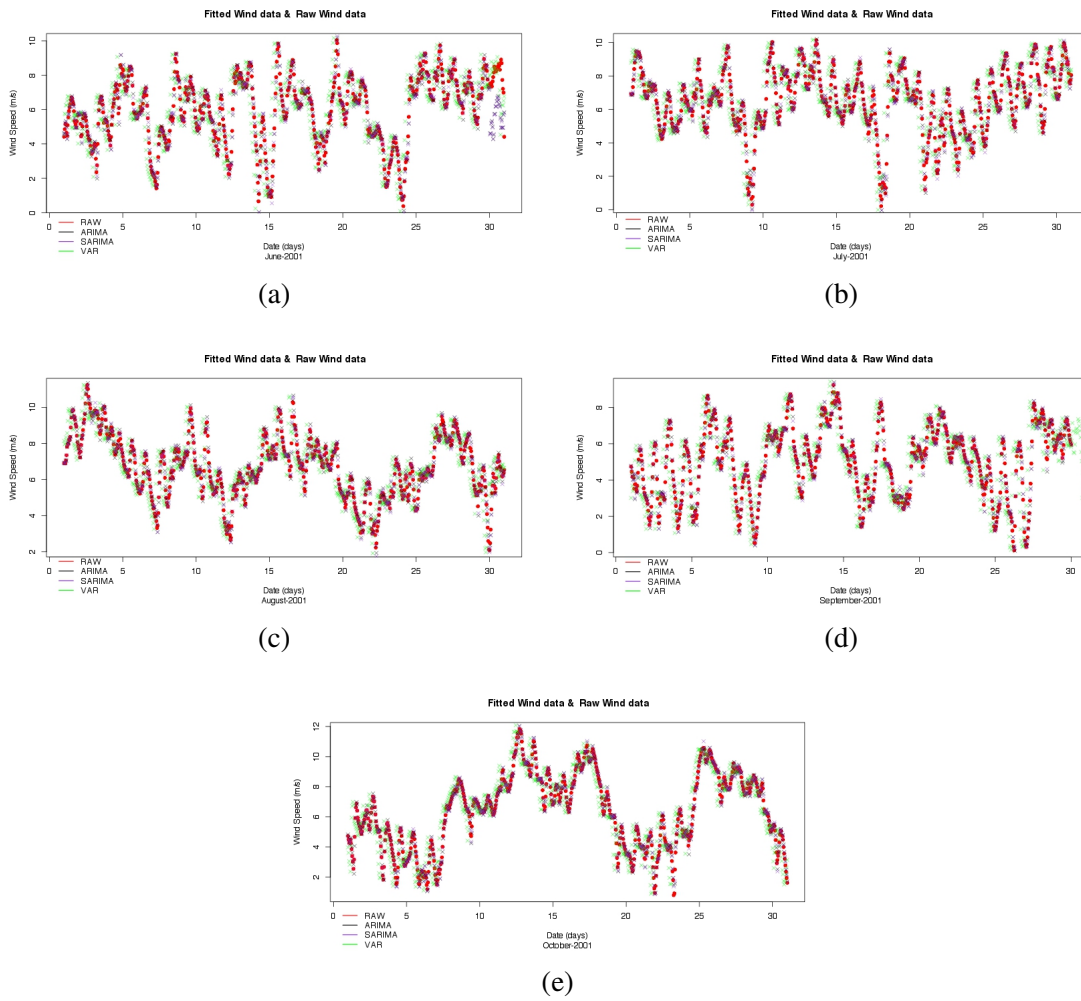


Figure 7.16. MERRA CASE Run plot of Actual and Fitted values (4)

7.3. Accuracy Comparison

In order to provide a detailed and numerical cross comparison between the performances of models, we adopt three most commonly used performance criteria.

Initially, Table 7.1 summarizes the performance criteria values for all models and all stations. The values are averaged over 17 months and lower values indicate lower forecast errors and greater accuracy.

Table 7.1. Comparison of Model Forecasting Performances for 17 months means

CASES	ARIMA	SARIMA	VAR
ADM	RMSE=2.91	RMSE=2.69	RMSE=2.75
	MAE=2.22	MAE=1.98	MAE=2.11
	MASE=1.03	MASE=0.91	MASE=0.43
URLA RES	RMSE=1.35	RMSE=1.33	RMSE=1.34
	MAE=1	MAE=0.99	MAE=1
	MASE=0.99	MASE=0.97	MASE=0.31
MERRA	RMSE=0.22	RMSE=0.2	RMSE=0.22
	MAE=0.16	MAE=0.14	MAE=0.16
	MASE=0.54	MASE=0.46	MASE=0.09
CFSR	RMSE=0.47	RMSE=0.42	RMSE=0.45
	MAE=0.33	MAE=0.29	MAE=0.31
	MASE=0.81	MASE=0.70	MASE=0.15

As a general result, both VAR and SARIMA models outperform, in almost all cases, the performance of ARIMA. Hence, VAR and SARIMA models are found to be superior in wind speed predictions. This actually means that taking into account other meteorological variables and seasonality increases the fitting performance of estimations. This shows once more the validity of our approach to wind speed estimations. This aspect has not yet been adequately studied in the literature which should, therefore, be considered as our major contribution to literature.

In detail, there are 4 stations x 3 performance criteria. Hence, we have 12 cross method comparison cases in total. Out of 12, VAR outperforms others in 4 cases. SARIMA performs best in 8 cases by producing lowest errors. And, in none of the cases, ARIMA performs best. Hence, one may derive a general ranking in a following way; SARIMA performs better than VAR that performs better than ARIMA.

Table 7.2 summarizes the performance values in ADM case. It presents detailed results for all months. According to MASE criteria, VAR model outperforms in all

months. However, in MAE criteria SARIMA is dominating as a superior model. Likewise, as for the RMSE criterion, SARIMA outperforms others in all most all cases.

It is obvious that ARIMA is the model which has the ability neither to catch seasonality nor multivariate parameters and, therefore, has the poorest performance results.

ADM case shows that in the forecasting of summer months' wind speeds, other meteorological variables like temperature and pressure increases their effects. So, VAR model increases the forecasting accuracy.

Table 7.2. Comparison of Model Forecasting Performances for ADM

Months	RMSE	MAE	MASE	Months	RMSE	MAE	MASE	Months	RMSE	MAE	MASE
1	2.99	2.32	1.02	1	2.74	2.00	0.88	1	2.77	2.17	0.43
2	2.96	2.28	1.05	2	2.77	2.03	0.93	2	2.82	2.18	0.41
3	2.73	2.10	1.02	3	2.71	2.07	1.01	3	2.58	1.97	0.50
4	2.86	2.14	1.05	4	2.47	1.77	0.87	4	2.64	2.01	0.37
5	3.07	2.31	1.06	5	2.77	1.99	0.91	5	2.84	2.15	0.38
6	2.54	1.92	1.01	6	2.51	1.89	1.00	6	2.42	1.83	0.42
7	3.22	2.39	1.02	7	3.12	2.28	0.98	7	3.05	2.27	0.43
8	2.96	2.21	1.04	8	2.64	1.85	0.88	8	2.86	2.14	0.45
9	2.76	2.08	1.04	9	2.44	1.68	0.85	9	2.66	2.03	0.36
10	2.84	2.15	1.01	10	2.63	1.92	0.91	10	2.75	2.08	0.45
11	2.90	2.26	1.00	11	2.69	2.00	0.89	11	2.83	2.21	0.51
12	3.04	2.38	1.00	12	2.81	2.08	0.88	12	2.89	2.24	0.45
13	2.99	2.35	1.03	13	2.86	2.26	0.99	13	2.81	2.21	0.41
14	2.91	2.18	1.05	14	2.55	1.77	0.85	14	2.70	2.04	0.43
15	2.98	2.32	1.03	15	2.70	2.00	0.89	15	2.75	2.14	0.45
16	2.92	2.25	1.01	16	2.78	2.10	0.95	16	2.75	2.08	0.41
17	2.85	2.19	1.04	17	2.60	1.91	0.90	17	2.67	2.07	0.40
Means	2.91	2.22	1.03	Means	2.69	1.98	0.91	Means	2.75	2.11	0.43

(a) ARIMA

(b) SARIMA

(c) VAR

In table 7.3, forecasting performance of models are presented for Urla RES met mast data case. In MASE criterion, the leader is clearly the VAR model among others. This happens to be true consistently for all months.

MAE performance criterion claims that all tree models approximately have the same level of forecasting errors.

RMSE gave us an interesting picture between months; June 2000 and July 2000 are the months in which SARIMA outperforms. August 2000 have same forecast errors between SARIMA and VAR.

As a general result, in RMSE, the leader is ambiguous and varying between SARIMA and VAR.

Table 7.4 summarizes the performance values of models for CFSR case. To start with RMSE, SARIMA clearly outperforms the others, in almost all months. Regarding MAE, once more SARIMA is clearly shown to outperform the other models. However,

Table 7.3. Comparison of Model Forecasting Performances for URLA

Months	RMSE	MAE	MASE	Months	RMSE	MAE	MASE	Months	RMSE	MAE	MASE
1	1.38	1.06	0.99	1	1.14	0.83	0.77	1	1.38	1.07	0.34
2	1.31	1.01	0.99	2	1.29	1.00	0.97	2	1.32	1.02	0.35
3	1.27	0.96	0.97	3	1.25	0.96	0.96	3	1.25	0.95	0.38
4	1.18	0.89	0.97	4	1.18	0.89	0.97	4	1.16	0.89	0.29
5	1.31	0.97	0.97	5	1.31	0.97	0.97	5	1.30	0.96	0.28
6	1.22	0.89	0.98	6	1.21	0.89	0.98	6	1.23	0.89	0.32
7	1.56	1.12	0.98	7	1.56	1.12	0.98	7	1.58	1.12	0.28
8	1.43	0.98	0.98	8	1.42	0.97	0.97	8	1.41	0.97	0.28
9	1.50	1.09	0.99	9	1.49	1.09	0.99	9	1.48	1.08	0.26
10	1.43	1.06	0.99	10	1.42	1.06	0.99	10	1.37	1.03	0.28
11	1.64	1.18	0.98	11	1.65	1.18	0.98	11	1.67	1.18	0.34
12	1.52	1.11	0.99	12	1.53	1.11	0.99	12	1.50	1.08	0.30
13	1.34	1.00	0.98	13	1.33	1.00	0.98	13	1.30	0.98	0.34
14	1.29	1.00	0.99	14	1.29	1.00	0.98	14	1.27	0.99	0.34
15	1.27	0.97	1.00	15	1.25	0.96	0.98	15	1.25	0.95	0.31
16	1.17	0.87	1.01	16	1.16	0.87	1.01	16	1.17	0.88	0.30
17	1.18	0.88	0.98	17	1.18	0.88	0.98	17	1.17	0.87	0.24
Means	1.35	1.00	0.99	Means	1.33	0.99	0.97	Means	1.34	1.00	0.31

(a) ARIMA

(b) SARIMA

(c) VAR

in MASE criterion, VAR shows a very clear and distinguished accuracy performance. In total, SARIMA and VAR clearly outperforms ARIMA which is robust across seasons and performance criteria.

Finally, Table 7.5 documents the performance values of models in MERRA case. The findings are quite similar to the previous cases. In terms of RMSE and MAE, SARIMA is the best model again in fitting the wind speed. And, VAR takes the lead in MASE case which performs best.

Overall, the message conveyed in this part points to the benefits of incorporating meteorological variables and seasonality in wind speed forecasting. Hence, VAR and SARIMA models are shown to perform better as they take into account these additional factors.

Table 7.4. Comparison of Model Forecasting Performances for CFSR

Months	RMSE	MAE	MASE	Months	RMSE	MAE	MASE	Months	RMSE	MAE	MASE
1	0.52	0.36	0.77	1	0.43	0.28	0.59	1	0.48	0.32	0.16
2	0.44	0.33	0.72	2	0.39	0.27	0.60	2	0.40	0.29	0.14
3	0.40	0.30	0.74	3	0.34	0.24	0.61	3	0.35	0.25	0.15
4	0.46	0.33	0.76	4	0.43	0.30	0.70	4	0.43	0.31	0.15
5	0.44	0.31	0.89	5	0.36	0.24	0.68	5	0.43	0.29	0.14
6	0.41	0.27	0.84	6	0.39	0.25	0.79	6	0.40	0.26	0.14
7	0.50	0.33	0.91	7	0.48	0.33	0.90	7	0.50	0.34	0.11
8	0.41	0.28	0.87	8	0.36	0.23	0.71	8	0.41	0.28	0.12
9	0.54	0.36	0.85	9	0.53	0.35	0.83	9	0.52	0.35	0.13
10	0.56	0.38	0.89	10	0.52	0.35	0.82	10	0.56	0.38	0.14
11	0.59	0.40	0.82	11	0.59	0.40	0.81	11	0.57	0.39	0.15
12	0.56	0.40	0.84	12	0.55	0.39	0.82	12	0.55	0.39	0.18
13	0.49	0.35	0.76	13	0.38	0.26	0.56	13	0.45	0.32	0.16
14	0.44	0.32	0.67	14	0.38	0.26	0.55	14	0.39	0.28	0.14
15	0.43	0.32	0.78	15	0.35	0.24	0.59	15	0.37	0.27	0.17
16	0.47	0.34	0.84	16	0.41	0.27	0.65	16	0.45	0.32	0.18
17	0.40	0.28	0.88	17	0.33	0.21	0.68	17	0.38	0.26	0.14
Means	0.47	0.33	0.81	Means	0.42	0.29	0.70	Means	0.45	0.31	0.15

(a) ARIMA

(b) SARIMA

(c) VAR

Table 7.5. Comparison of Model Forecasting Performances for MERRA

Months	RMSE	MAE	MASE	Months	RMSE	MAE	MASE	Months	RMSE	MAE	MASE
1	0.25	0.17	0.56	1	0.19	0.13	0.41	1	0.24	0.17	0.10
2	0.24	0.18	0.56	2	0.20	0.14	0.44	2	0.24	0.17	0.11
3	0.23	0.17	0.55	3	0.18	0.13	0.41	3	0.22	0.16	0.10
4	0.23	0.16	0.52	4	0.21	0.15	0.47	4	0.22	0.16	0.09
5	0.24	0.15	0.54	5	0.21	0.12	0.44	5	0.19	0.14	0.07
6	0.22	0.15	0.52	6	0.21	0.15	0.51	6	0.22	0.15	0.07
7	0.20	0.14	0.53	7	0.18	0.13	0.49	7	0.19	0.14	0.06
8	0.17	0.12	0.50	8	0.16	0.12	0.48	8	0.17	0.12	0.06
9	0.18	0.14	0.47	9	0.17	0.13	0.45	9	0.19	0.14	0.06
10	0.21	0.16	0.54	10	0.19	0.14	0.48	10	0.22	0.16	0.07
11	0.26	0.18	0.51	11	0.25	0.18	0.49	11	0.25	0.18	0.08
12	0.27	0.19	0.58	12	0.26	0.17	0.52	12	0.27	0.19	0.09
13	0.26	0.19	0.59	13	0.23	0.16	0.51	13	0.25	0.18	0.11
14	0.25	0.18	0.56	14	0.20	0.15	0.44	14	0.23	0.17	0.11
15	0.21	0.15	0.59	15	0.17	0.12	0.44	15	0.21	0.15	0.11
16	0.21	0.15	0.49	16	0.19	0.14	0.43	16	0.21	0.15	0.09
17	0.20	0.14	0.53	17	0.16	0.11	0.40	17	0.19	0.14	0.07
Means	0.22	0.16	0.54	Means	0.20	0.14	0.46	Means	0.22	0.16	0.09

(a) ARIMA

(b) SARIMA

(c) VAR

7.4. Impulse Response Function

Finally, it is worthwhile to investigate the dynamics of wind speed more in detail. In order to do this, we calculate the Impulse-Response Functions (IRF) of wind speed from VAR estimations with respect to different variables. In general, IRF summarizes over time the impact of unexpected movements in one variable to another. Figures 7.17,7.20,7.18,7.19 present below the IRF of wind speed to unexpected changes in different variables for 4 datasets.

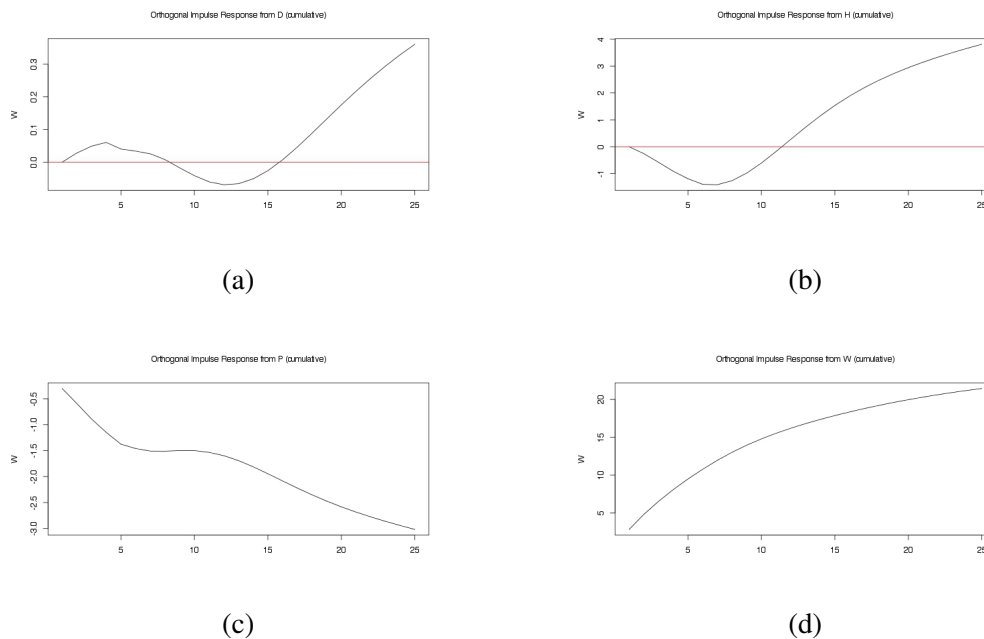
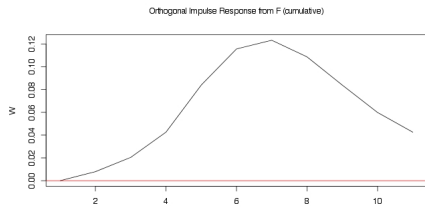
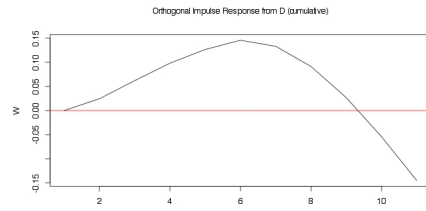


Figure 7.17. ADM CASE IRF Plots

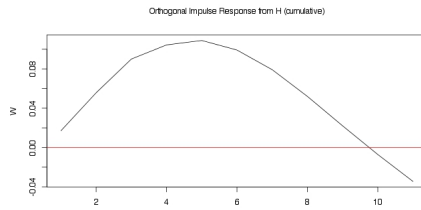
It seems that an unexpected increase in pressure decreases the wind speed persistently in 4 cases. The mechanism is likely to work in a following way; it is well known that pressure is higher above the sea. As air pressure rises on land, it will balance the pressure between sea and land, thus, the wind speed should slowdown. In contrast, if the pressure decreases further on land, then the unbalance of pressure between sea and land would indeed widen that will accelerate the wind speed. This case fits perfectly to our URLA-RES case since the mat mast is on the campus area which just towards the Aegean Sea. Hence, air pressure stands as one of the most important dynamics of wind speed. With regard to humidity, an unexpected increase in humidity accelerates the wind speed.



(a)



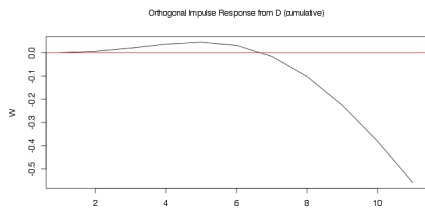
(b)



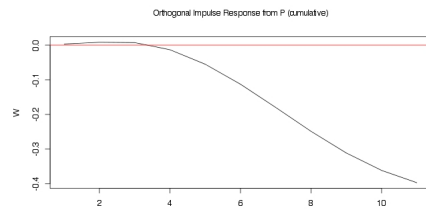
(c)

Figure 7.18. CFSR CASE IRF Plots

Wind direction gives, in contrast, an ambiguous message. Such that In URLA-RES and ADM case, deviation of wind direction from north to south lowers the wind speed in the first few hours but then accelerates it afterwards. Controversially, in CFSR and MERRA case, its impact is positive: change in the direction towards south increases the wind speed in the first few hours then decreases.



(a)



(b)

Figure 7.19. MERRA CASE IRF Plots

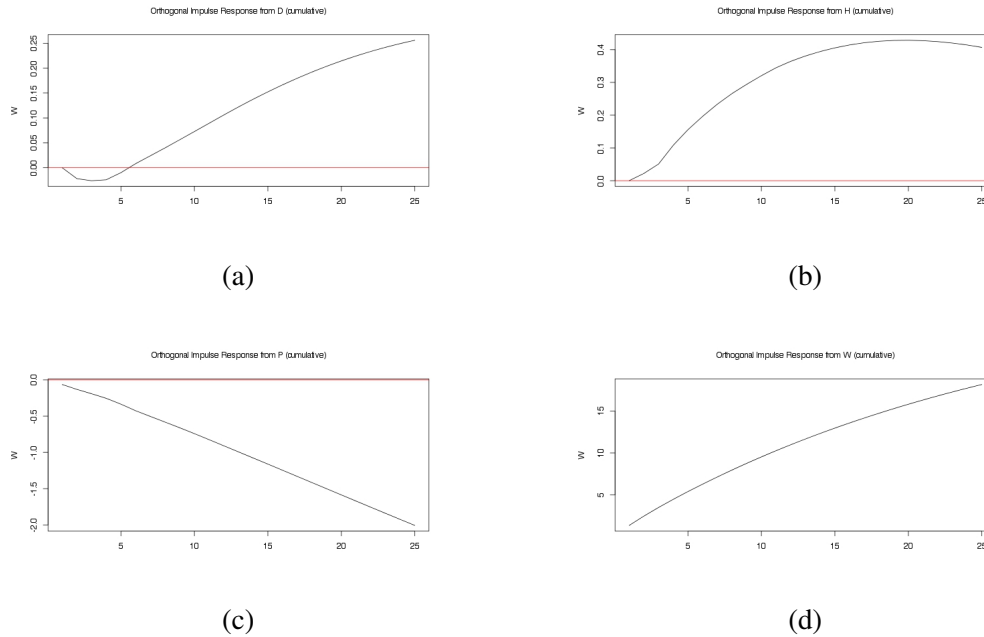


Figure 7.20. URLA RES CASE IRF Plots

7.5. Robustness Check for model residuals and White Noise Property

This sub-section is devoted to checking whether obtained residuals satisfy the White Noise characteristics. As one can remind, White Noise is a special case and desired property of residuals which follow a normal distribution with zero mean and constant variance.

In order to check this, initially we present the average value of residuals for each model in Table 7.6. The error terms in this table are estimated from aggregated 17-months time series models. We observe that VAR model produces errors with mean extremely close to zero. Error terms obtained from SARIMA are positive and from ARIMA are negative. But in any case, they are pretty close to zero which approves the validity of zero mean.

As for the variance, we depict the evolution of residuals for all models and stations in figures from 7.21 to 7.24. In all cases, error terms seem to follow a very clear stationary process, with mean close to zero and variance remaining more or less constant over time. Hence, one may argue that the models we run are technically adequate to satisfy the White Noise property.

Table 7.6. Means of Residuals (four digits)

Model	ADM	CFSR	MERRA	URLA RES
VAR	0	0	0	0
ARIMA	0.0055	0.0009	-0.0019	-0.001
SARIMA	-0.0063	0.002	-0.0005	-0.0025

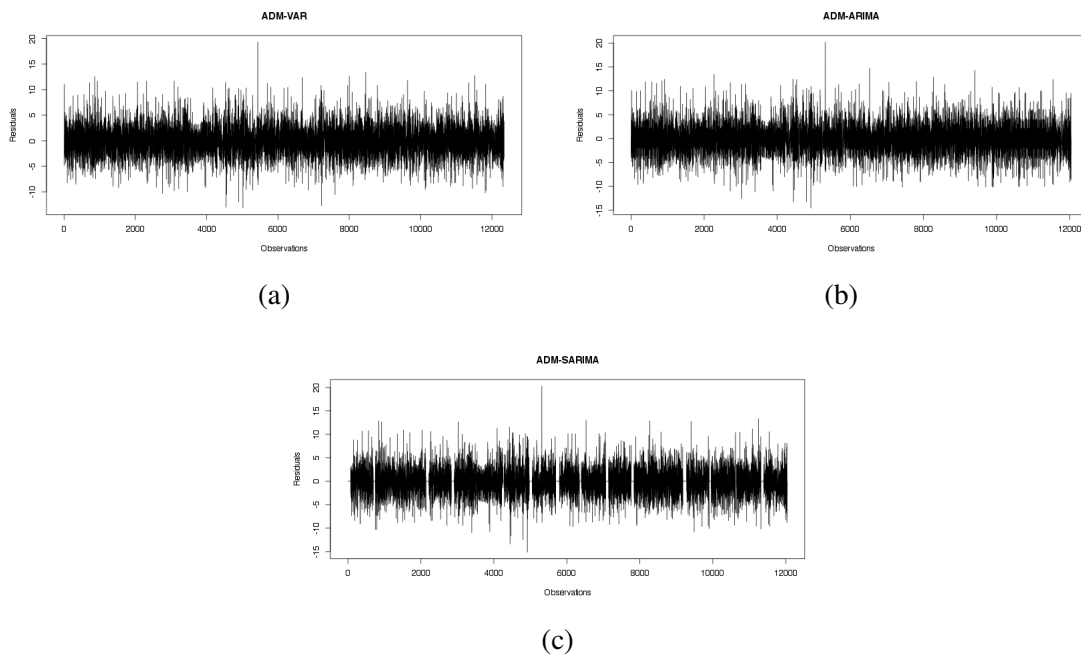


Figure 7.21. ADM CASE residuals for a)VAR model b)ARIMA model c)SARIMA model

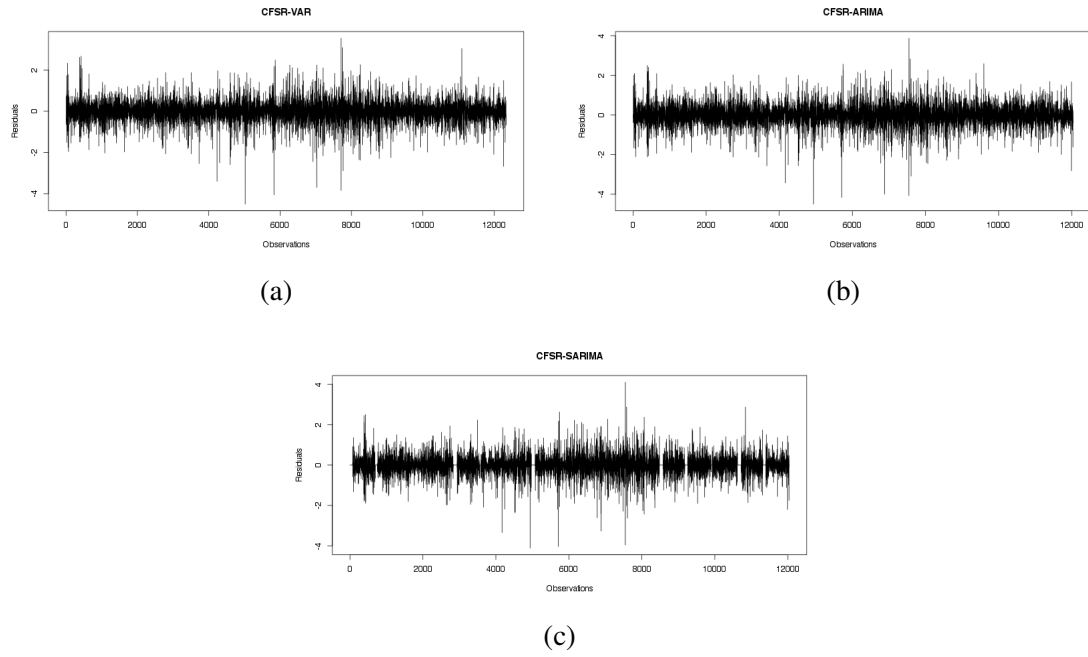


Figure 7.22. CFSR CASE residuals for a)VAR model b)ARIMA model c)SARIMA model

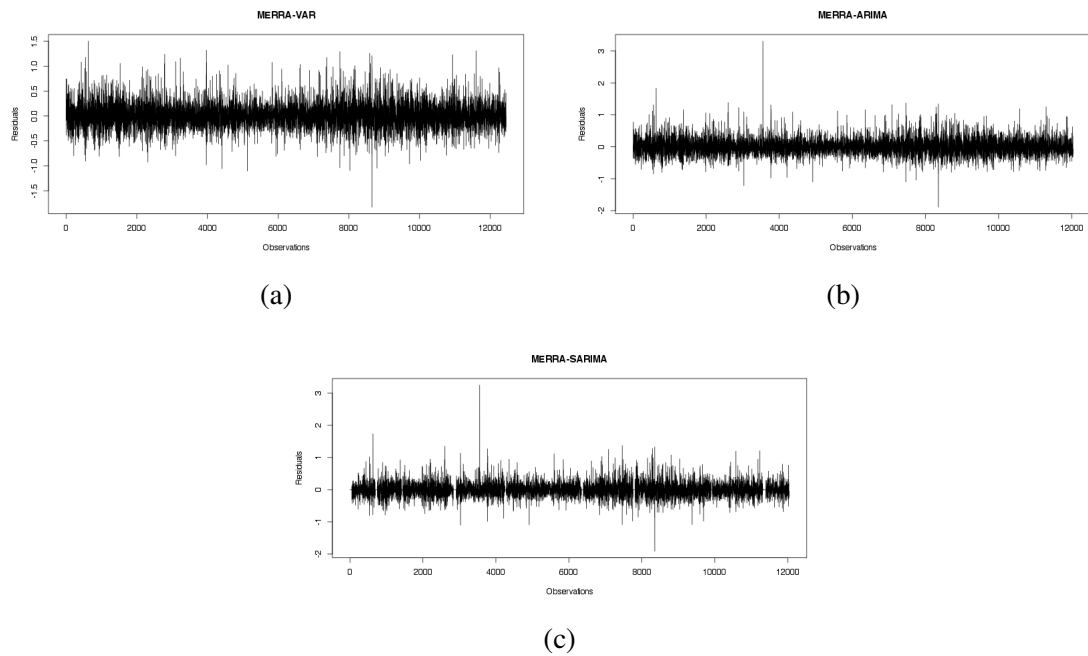


Figure 7.23. MERRA CASE residuals for a)VAR model b)ARIMA model c)SARIMA model

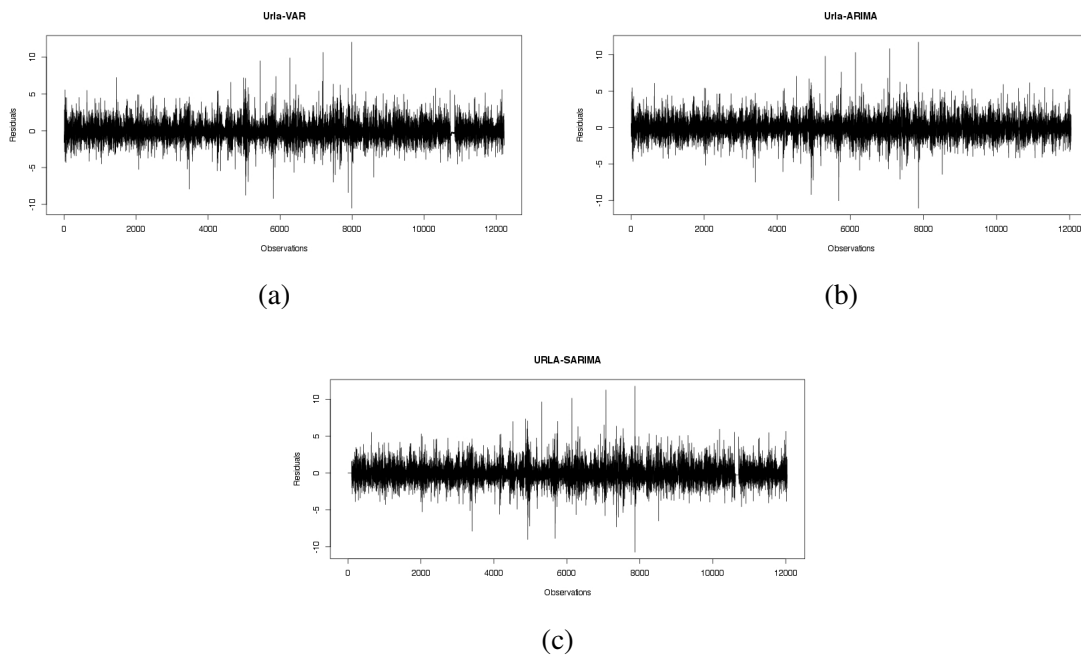


Figure 7.24. URLA RES CASE residuals for a)VAR model b)ARIMA model c)SARIMA model

7.6. Forecasted Values

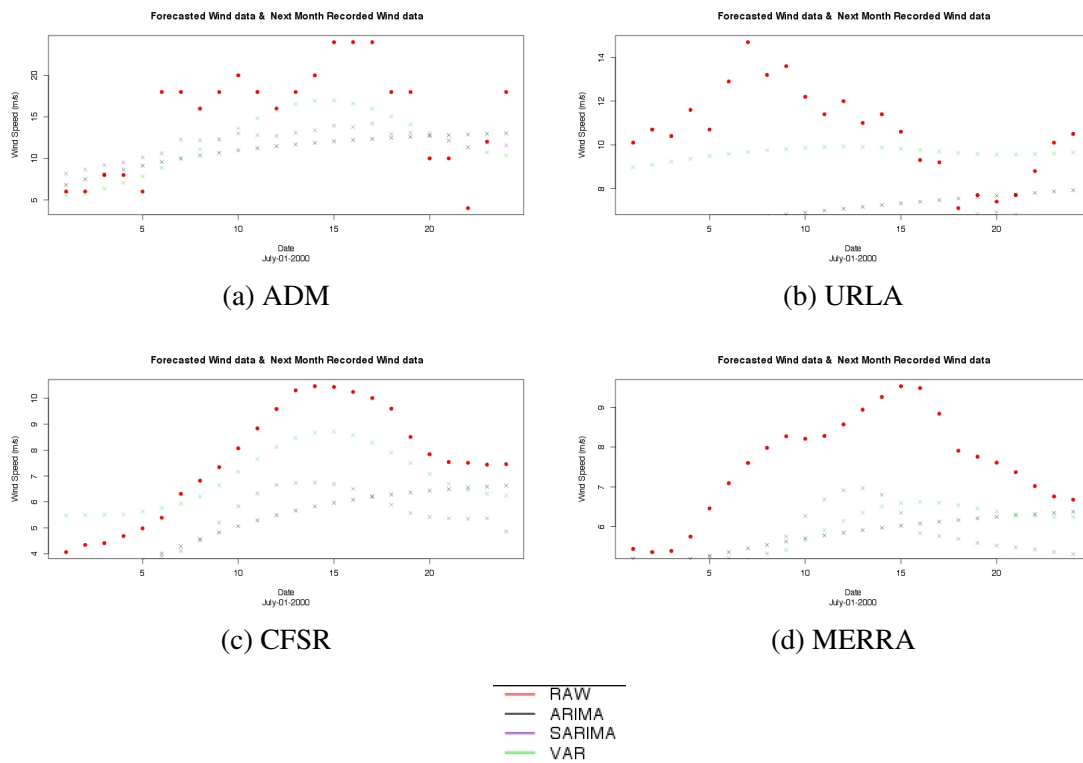


Figure 7.25. Forecasted Values versus Real Data of 01 July 2000

Red dots are the original measured values and green ones VAR forecasted values, whereas black ones ARIMA and purple ones SARIMA future forecasts. Reader should be aware the difference between future forecast and fitted values which were presented in previous sections. In figure 7.25 Herein it can be seen VAR model is superior to SARIMA and ARIMA models. In figure 7.25 and 7.26 it can be seen also forecast accuracy depends site and case.

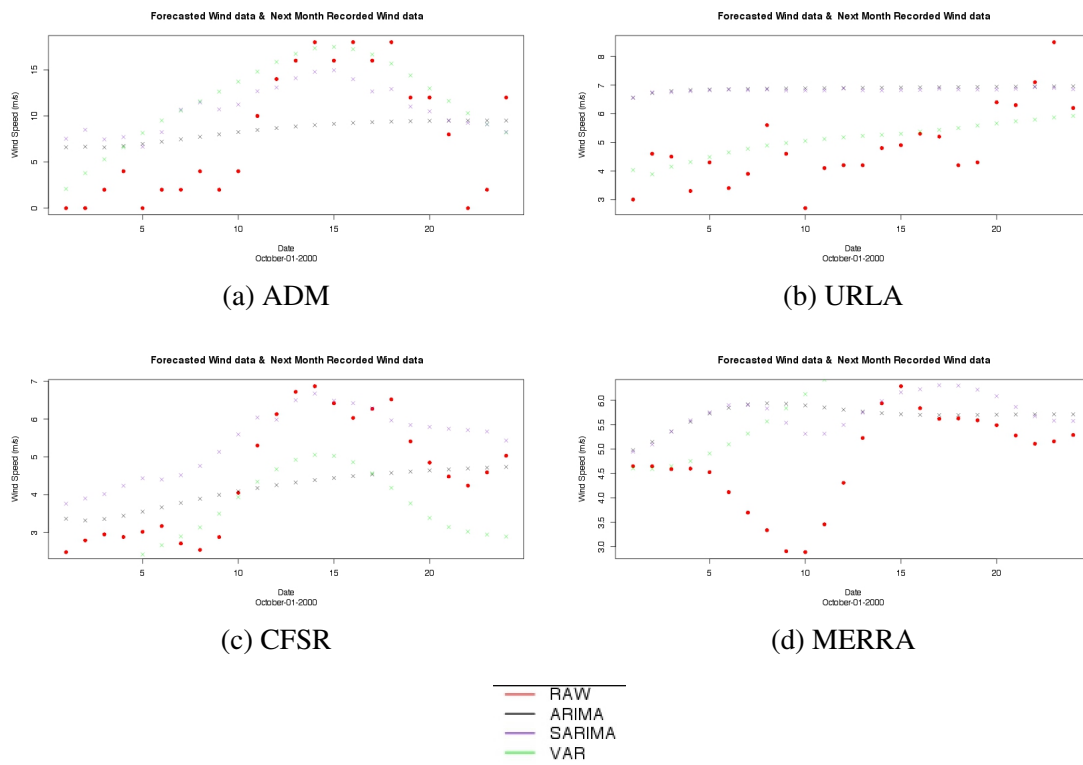


Figure 7.26. Forecasted Values versus Real Data of 01 October 2000

CHAPTER 8

ENGINEERING APPLICATION

In the most of the cases engineers must handle with missing datasets actually regulations state clearly that minimum 1 year in another words four seasons must be captured for wind resource assessment which is obligatory for commercial applications. Actually minimum three years data set is more trustful representative for correct characteristics of terrain. However anemometers can be broken or data loggers could cause problems, also the maintenance issues or natural phenomena like thunder, icing etc. can interrupt both Met mast and SCADA records from turbine.

As a result the recovery rate of measured data could not be 100%. Although full data set is required to determine the characteristics and capacity of terrain, simulations, energy estimations etc. can not be performed with actual measured full data. As we investigated in our previous chapters our time series analysis and methodologies can predict future value or missing value. For now we will see the comparison of raw missing data with recovery rate approximately 50% annual energy production and measured annual energy production of data versus filled data with our time series models.

8.1. Data

Here for the engineering application part Nordtank data set is employed to see affect of selected methodology in filling procedure. The Data is provided by DTU Wind Energy Department, Kurt Hansen. The reader can ask the question why we are not modelling Urla case and change our data set. The answer is in Urla case the project is still on the way with 15 MW wind farm (EGENDA, 2014) so there is no available produced power data. For the comparison and to know accuracy of estimations this information is crucial. Because of that in this part we will change our data set.

DTU and Risø National laboratory erected a test turbine for the mainly three scientific purposes; meteorological data collection, wind turbine operational statistics determination, investigation of structural loads. Our data is taken from these project.

In figure 8.1 mast measurement heights, turbine characteristics heights and loca-

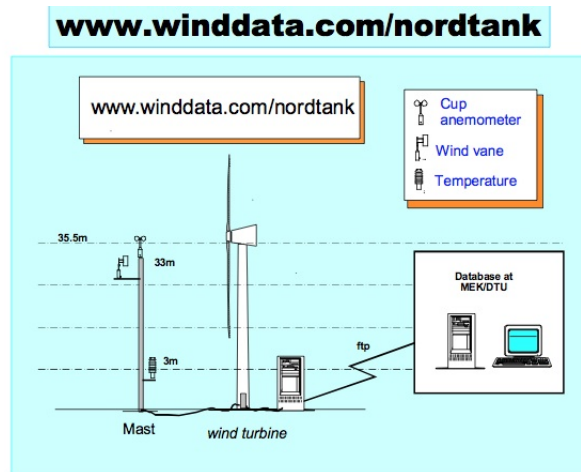


Figure 8.1. Nordtank Data Station (Source:Hansen et al., 2006)

tion of sensors are illustrated for the measurement station. (Hansen et al., 2006)

Location of data set is the Risø Campus, Roskilde Denmark, a country generally known as with impressive wind characteristics and typical example of flat orography which is highly desirable for wind energy applications and Roskilde has pastoral terrain characteristics (Larsen and Hansen, 2004). Three-bladed stall regulated Nordtank, with 36 meters hub height wind turbine is located with the consideration of main leading wind direction which is free undisturbed Westerly wind sector. (Hansen et al., 2006).

In this project wind speed, wind direction, pressure, temperature, grid connected power production records with 10 minutes intervals are employed. Units and details of used data are listed below in table 8.1.

Table 8.1. Nordtank Data Set

Parameter	Unit	Cover Range
Wind Speed	(m/s)	12,2003-04,2008
Wind Direction	(deg)	12,2003-04,2008
Pressure	(hPa)	12,2003-04,2008
Temperature	(degC)	12,2003-04,2008
Power(grid connected)	(kW)	12,2003-2005

8.2. Annual Energy Production (AEP)

Annual energy production calculation actually is based on probability density function of wind speed and power curve characteristics of turbine in one year (approximately 8760 hours).

$$AEP = 8760 \sum_{i=1}^n P(W S_i) R(W S_i) \quad (8.1)$$

$R(W S_i)$ is the distribution of wind speed this distribution is generally represents with Rayleigh distribution and $P(W S_i)$ is the wind turbine's power curve, i represents observations .(Crockford and Hui, 2007b)

In another way to calculate it frequency table method.

$$AEP = \sum_{j=1} f_j P(W S_j) \quad (8.2)$$

(Gerdes, 2005) f_j frequency of wind speed class j , $P(W S_j)$ is the power which produced in wind class j according to turbine manufacturer power curve versus wind speed graph or power curve table. j is the representative of wind classes generally it starts from 1 m/s and stops at 25 m/s, but generally power values are given listed from 3 m/s to 25m/s due to cut in and cut out wind speeds of wind turbine.

8.2.1. Weibull Distribution

$$f(v) = \frac{k}{A} \left(\frac{WS}{A}\right)^{k-1} \exp\left(-\frac{WS}{A}\right) \quad (8.3)$$

WS is the wind speed k is the Weibull shape parameter A is the scale parameter. Rayleigh Distribution is a type of Weibull distribution which is the k parameter is set to 2. These probability density functions actually creates a curve which fits the actual data for our case y axes probability (from zero to one) in another words occurrence probability of

certain wind speed classes and x axes represents wind speed which starts from zero stops generally at 25 m/s. Variation of wind speeds can be seen easily from distribution plot.

8.3. WAsP

WAsP, Wind atlas analysis and application program is one of engineering tool, the widely accepted in wind energy sector which is developed in DTU. It depends on linear model (The logarithmic wind profile (Based on geostrophic balance)) and suitable for flat terrains. Although WAsP is a time efficient program comparatively with non linear solvers and CFD (computational fluid dynamics) based programs, it has some limitations about wake calculations and turbulence, number of wind turbines (high number of wind turbines means at the same time wake occurrences, for example WAsP suggests only a fixed ratio for wakes like %10 reduction in AEP due to wake.) and terrain characteristics (weakness on complex terrain). WAsP module can take into account Stability, Roughness (for example open sea has 0.0002 m roughness length and city has 2 m), Shelter and Orography (mountain, hills or elevation of the terrain).

Wind atlas methodology is known with two main approaches in the one hand observational wind atlas and in the other hand numerical wind atlas.

For now observational wind atlas part is crucial for our application. Indeed until now we dealt with raw data, cleaning procedure, fitting and statistics of data so we are already familiar with the first step, observed wind climatology and the all process can be summarized by below listed steps.

Meteorological mast or SCADA data of wind turbine in another words measurement sensor is placed in space with certain three coordinates. First input is the 3-D coordinates of measurement sensor, and for the second input measurements are required in chapter 3 we explained the data measurement and collection procedure. In third input we have raw data which is recorded in certain time period at end creation of our wind direction and wind speed time series is completed. Finally if we take statistics of these time series we can reach mean, variance, Weibull distribution etc. in terms of observed wind climate (OWC) and then if we add site information such as roughness, elevation to OWC, wind atlas of region can be calculated which equals to regional wind climate (RWC). (WAsP, 2014)

Now if add to wind turbine coordinates and wind farm site descriptions to RWC

we can reach predicted wind climate. Predicted wind climate and power curve combination will give us to annual energy production (AEP) of a wind turbine.

8.4. Results

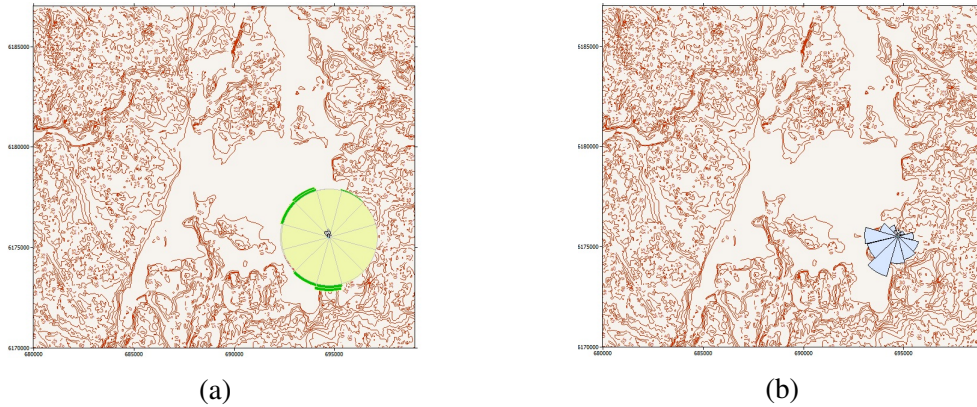


Figure 8.2. Roughness Effect and Wind Rose

In figure 8.2a terrain and test case can be seen. Westerly winds are dominant, West, South-South-West directions are the main leading wind directions in figure 8.2b

		R-class 0 (0.000 m)	R-class 1 (0.030 m)	R-class 2 (0.100 m)	R-class 3 (0.400 m)	R-class 4 (1.500 m)
Height 1 (z = 10 m)	U [m/s]	9.82	7.21	6.26	4.89	3.13
	P [W/m ²]	758	305	201	97	28
Height 2 (z = 25 m)	U [m/s]	10.72	8.55	7.67	6.39	4.78
	P [W/m ²]	979	497	361	213	94
Height 3 (z = 36 m)	U [m/s]	11.11	9.16	8.29	7.04	5.45
	P [W/m ²]	1081	602	450	281	138
Height 4 (z = 77 m)	U [m/s]	12.00	10.69	9.79	8.54	6.95
	P [W/m ²]	1362	918	714	483	276
Height 5 (z = 120 m)	U [m/s]	12.61	11.84	10.88	9.57	7.93
	P [W/m ²]	1588	1246	970	669	399

Figure 8.3. Wind Atlas

This atlas is the product of SARIMA filled wind speed and na filled with optimized box-cox transformation seasonality included na.intep wind direction. As you can see from figure 8.3 for our case roughness class 1 and height 3 is the representative of our case.

As a result we can assume 0.738258 GWh is the annual energy production representative with highest recovery rate. So we will model only 2004 year.If we assume linear

Table 8.2. Measured Power Outputs

Years	AEP(kW10min)	kWh	Recovery Rates%
2003-12-17 to end of year	253147.7	42191.29	69
2004	4429552.5	738258.74	62
2005	1273615.8	212269.30	53

relationship is valid between recovery rate and AEP values we can find 1.190739 GWh per year.

Na.interp is the one of function in the forecasts package of R. That function is used often to cleaning and filling procedure. Linear interpolation based process is used to complete some cases.

Table 8.3. Comparison for only 2004 (GWh)

AEP result	Case
1.19	Measured
1.775	Missing Data Estimated Value
1.550	NA interp filled Data Estimated
1.546	NA interp filled VAR model fitted with lambda coefficient
1.602	SARIMA Wind NA direction
1.952	monthly mean filled
1.546	lambda zero SARIMA

- i) VAR, ARIMA and SARIMA models can fill missing values without changing general statistics of data set. (Weibull distributions)
- ii) Filled data is more successful comparatively measured missing data set in AEP estimation.
- iii) VAR, ARIMA and SARIMA can create fitted sets and also forecasted cells.
- iv) Due to lack of recovery date problem our measured value is not certain.
- v) Due to processing limitations, time problem occurred each step model coefficient estimation for all missing values, we could not see directly forecasted wind affect

with SARIMA and VAR models but at least for 2004 we took forecasted values of wind with SARIMA model.

CHAPTER 9

CONCLUSION

This study has investigated the various ways of estimating the wind speed and power generation in Turkey. Our estimates from 4 different stations (ADM; URLA RES, MERRA, CSFR and three different time series models (ARIMA; SARIMA; VAR) indicate several important conclusions.

First, Both SARIMA, ARIMA and VAR models provide an accurate estimate of wind speed as the forecasted values of models fit quite well to the actual data.

Second, multivariate VAR and univariate SARIMA are shown to outperform ARIMA by improving the wind predictions and producing less forecast errors. This result is quite robust across different stations. This has been shown using three performance criteria which indicate the magnitude of forecast errors. In detail, SARIMA and VAR exhibits a better performance in all seasons and for all criteria. One may rank the performance of models as “SARIMA outperforming VAR that performs better than ARIMA”.

Third, among the dynamics of wind speed, air pressure, humidity and wind direction is shown to influence the wind speed significantly. Residual of models seem to provide a satisfactory view for White Noise property that in all stations and models, residuals have a mean close to zero and display a constant variance over time.

Overall, the most important lesson we get from our analysis is that incorporating multivariate models and seasonality component can be useful in improving the forecast accuracy of wind speed that may provide additional benefits to energy harvesting and power production.

Filling procedure code requires a considerable calculation time. However, the code can be written with shortcuts for loops like a fix model parameters. Also Measure-Correlate-Predict module should be run with 10 years ADM case and Urla RES to see affect of MCP improvement in energy estimations via our new suggested methodology after the URLA RES power production records.

For further study the time when the turbines starts power generation, Supervisory Control and Data Acquisition power production data can be investigated for the comparison of energy estimations which data set may create more accurate results comparatively

with measured power production. Another important study area is the multivariate stability analysis, a comparison between stability model and VAR multivariate model could be interesting.

REFERENCES

- Ackerman, S. A., J. A. Knox, et al. (2013). *Meteorology*. Jones & Bartlett Publishers.
- Agency, I. E. (2014). Key world energy statistics,2014. Technical report, IEA.
- Akaike, H. (1974). A new look at the statistical model identification. *Automatic Control, IEEE Transactions on* 19(6), 716–723.
- Arwade, S. R. and M. Gioffre (2014). Validity of stationary probabilistic models for wind speed records of varying duration. *Renewable Energy* 69(0), 74 – 81.
- Baum, C. F. (2013). Quantile regression. URL [http://fmwww. bc. edu/EC-C S 2013](http://fmwww.bc.edu/EC-C S 2013).
- Bivona, S., G. Bonanno, R. Burlon, D. Gurrera, and C. Leone (2011). Stochastic models for wind speed forecasting. *Energy conversion and management* 52(2), 1157–1165.
- Cadenas, E. and W. Rivera (2007). Wind speed forecasting in the south coast of oaxaca, mexico. *Renewable energy* 32(12), 2116–2128.
- Carolin Mabel, M. and E. Fernandez (2008). Analysis of wind power generation and prediction using ann: a case study. *Renewable Energy* 33(5), 986–992.
- Castro, C., M. Mediavilla, L. J. Miguel, and F. Frechoso (2011). Global wind power potential: Physical and technological limits. *Energy Policy* 39(10), 6677–6682.
- Chang, W.-Y. (2014). A literature review of wind forecasting methods. *Journal of Power and Energy Engineering* 2(04), 161.
- Council, G. W. E. (2014). Global wind energy outlook 2014. Technical report, GWEC.
- Crockford, A. and S.-Y. Hui (2007a). *Wind profiles and forests*. Ph. D. thesis, Master Thesis at Risoe DTU.
- Crockford, A. and S.-Y. Hui (2007b). Wind profiles and forests. Master’s thesis, Technical University of Denmark.
- EGENDA (2014). Projects,urla wep. <http://www.egenda.com.tr>,(accessed at December 2014).
- Enercon (2014). Enercon productsheet e-126. url = <http://www.enercon.de/en-en/66.htm>, (accessed January 2014).
- Erdem, E. and J. Shi (2011). Arma based approaches for forecasting the tuple of wind speed and direction. *Applied Energy* 88(4), 1405–1414.

- Foley, A. M., P. G. Leahy, A. Marvuglia, and E. J. McKeogh (2012). Current methods and advances in forecasting of wind power generation. *Renewable Energy* 37(1), 1–8.
- Friendly, M. (2002). Corrgrams: Exploratory displays for correlation matrices. *The American Statistician* 56(4), 316–324.
- Gao, S., Y. He, and H. Chen (2009). Wind speed forecast for wind farms based on armar-ch model. In *Sustainable Power Generation and Supply, 2009. SUPERGEN'09. International Conference on*, pp. 1–4. IEEE.
- Gerdes, G. J. (2005). Energy yield assessment. Technical report, The Global Sustainable Electricity Partnership and Pasific Power Assosiation.
- Gerolimetto, M. (2010). Arima and sarima models.
- Giebel, G., R. Brownsword, G. Kariniotakis, M. Denhard, and C. Draxl (2011). The state-of-the-art in short-term prediction of wind power: A literature overview. Technical report, ANEMOS. plus.
- Gomes, P. and R. Castro (2012). Wind speed and wind power forecasting using statistical models: Autoregressive moving average (arma) and artificial neural networks (ann). *International Journal of Sustainable Energy Development* 1(1/2).
- Gujarati, D. N. (2012). *Basic econometrics*. Tata McGraw-Hill Education.
- Halim, S. and I. N. Bisono (2008). Automatic seasonal auto regressive moving average models and unit root test detection. *International Journal of Management Science and Engineering Management* 3(4), 266–274.
- Hansen, K. S., K. Helgesen Pedersen, and U. Schmidt Paulsen (2006). Online wind turbine measurement laboratory. In *2006 European Wind Energy Conference and Exhibition*.
- Henson, R. (2011). *The rough guide to climate change*. Dorling Kindersley Ltd.
- Hui, Z., L. Bin, and Z. Zhuo-qun (2011). Short-term wind speed forecasting simulation research based on arima-lssvm combination method. In *Materials for Renewable Energy & Environment (ICMREE), 2011 International Conference on*, Volume 1, pp. 583–586. IEEE.
- Hyndman, R. J. and A. B. Koehler (2006). Another look at measures of forecast accuracy. *International journal of forecasting* 22(4), 679–688.
- Jung, J. and R. P. Broadwater (2014). Current status and future advances for wind speed and power forecasting. *Renewable and Sustainable Energy Reviews* 31, 762–777.

- Larsen, G. and K. Hansen (2004). Database on wind characteristics—contents of database bank—revision i. *Risø National Laboratory, Roskilde, DK*.
- Lei, M., L. Shiyan, J. Chuanwen, L. Hongling, and Z. Yan (2009). A review on the forecasting of wind speed and generated power. *Renewable and Sustainable Energy Reviews 13*(4), 915–920.
- Li, G. and J. Shi (2010). On comparing three artificial neural networks for wind speed forecasting. *Applied Energy 87*(7), 2313–2320.
- Ruiz-Aguilar, J., I. Turias, and M. Jiménez-Come (2014). Hybrid approaches based on sarima and artificial neural networks for inspection time series forecasting. *Transportation Research Part E: Logistics and Transportation Review 67*, 1–13.
- Sayın, E. (2003). Physical features of the izmir bay. *Continental Shelf Research 23*(10), 957–970.
- Shoba, G. and G. Shobha. Rainfall prediction using data mining techniques: A survey.
- Sims, C. A. (1980). Macroeconomics and reality. *Econometrica: Journal of the Econometric Society*, 1–48.
- Soman, S., H. Zareipour, O. Malik, and P. Mandal (2010, Sept). A review of wind power and wind speed forecasting methods with different time horizons. In *North American Power Symposium (NAPS), 2010*, pp. 1–8.
- Stull, R. B. (1988). *An introduction to boundary layer meteorology*, Volume 13. Springer.
- Tani, I. (1977). History of boundary layer theory. *Annual review of fluid mechanics 9*(1), 87–111.
- Torres, J., A. Garcia, M. D. Blas, and A. D. Francisco (2005). Forecast of hourly average wind speed with {ARMA} models in navarre (spain). *Solar Energy 79*(1), 65 – 77.
- WAsP (2014). Observational wind atlas. url = <http://www.wasp.dk/Wind-Atlas>, (accessed at December 2014).

APPENDIX A

MODEL PARAMETERS

Table A.1. Model Parameters for ADM

Months	p	d	q
June00	2	0	0
July00	1	0	0
August00	2	0	0
September00	4	0	1
October00	1	0	2
November00	2	0	0
December00	2	0	0
January01	2	0	0
February01	2	0	0
March01	4	0	1
April01	4	0	0
May01	4	0	1
June01	1	0	0
July01	1	0	0
August01	1	0	0
September01	3	0	1
October01	3	0	1

(a) ARIMA

Months	p	d	q	P	D	Q
June00	2	0	0	3	0	1
July00	1	0	1	2	0	1
August00	2	0	0	1	0	0
September00	2	0	1	3	0	4
October00	2	0	1	3	0	3
November00	2	0	0	0	0	2
December00	2	0	1	1	0	1
January01	2	0	0	3	0	3
February01	1	0	1	4	0	2
March01	1	0	1	2	0	4
April01	3	0	1	3	0	1
May01	3	0	0	3	0	1
June01	4	0	0	2	0	2
July01	2	0	1	4	0	2
August01	3	0	0	2	0	2
September01	2	0	0	1	0	1
October01	2	0	1	3	0	1

(b) SARIMA(s=24)

Months	p
June00	6
July00	4
August00	6
September00	6
October00	6
November00	5
December00	6
January01	4
February01	6
March01	4
April01	4
May01	3
June01	6
July01	5
August01	5
September01	6
October01	6

(c) VAR

Table A.2. Model Parameters for CFSR

Months	p	d	q
June00	2	0	1
July00	2	0	2
August00	2	0	0
September00	2	0	0
October00	2	0	0
November00	4	0	0
December00	2	0	0
January01	2	0	0
February01	3	0	1
March01	2	0	0
April01	4	0	0
May01	2	0	0
June01	3	0	1
July01	2	0	1
August01	2	0	0
September01	2	0	0
October01	2	0	0

(a) ARIMA

Months	p	d	q	P	D	Q
June00	2	0	0	3	0	3
July00	2	0	3	2	0	2
August00	4	0	0	2	0	2
September00	2	0	0	3	0	3
October00	2	0	2	4	0	2
November00	3	0	1	1	0	1
December00	3	0	1	0	0	2
January01	2	0	0	4	0	0
February01	2	0	0	3	0	0
March01	2	0	0	3	0	2
April01	2	0	0	1	0	0
May01	2	0	0	0	0	2
June01	1	0	3	4	0	1
July01	2	0	0	3	0	1
August01	3	0	1	1	0	2
September01	2	0	0	4	0	0
October01	2	0	0	3	0	2

(b) SARIMA (s=24)

Months	p
June00	3
July00	3
August00	6
September00	5
October00	3
November00	6
December00	3
January01	3
February01	6
March01	3
April01	2
May01	2
June01	6
July01	6
August01	6
September01	3
October01	3

(c) VAR

Table A.3. Model Parameters for MERRA

Months	p	d	q
June00	6	0	3
July00	0	2	
August00	5	0	1
September00	5	0	2
October00	2	0	1
November00	3	0	1
December00	4	0	1
January01	3	0	2
February01	3	0	1
March01	2	0	1
April01	2	0	2
May01	2	0	1
June01	3	0	3
July01	5	0	2
August01	4	0	4
September01	2	0	1
October01	3	0	3

(a) ARIMA

Months	p	d	q	P	D	Q
June00	2	0	1	2	0	1
July00	2	0	2	2	0	4
August00	3	0	2	1	0	2
September00	5	0	0	3	0	0
October00	3	0	4	3	0	3
November00	1	0	4	1	0	2
December00	3	0	1	0	1	
January01	4	0	1	1	0	2
February01	3	0	3	4	0	0
March01	3	0	1	2	0	2
April01	2	0	2	1	0	3
May01	2	0	1	2	0	1
June01	2	0	1	4	0	4
July01	3	0	3	3	0	0
August01	5	0	0	1	0	2
September01	3	0	3	2	0	4
October01	5	0	0	3	0	0

(b) SARIMA (s=24)

Months	p
June00	6
July00	6
August00	6
September00	6
October00	6
November00	4
December00	6
January01	6
February01	3
March01	6
April01	6
May01	6
June01	6
July01	6
August01	6
September01	6
October01	6

(c) VAR

Table A.4. Model Parameters for URLA RES

Months	p	d	q
June00	4	0	1
July00	1	0	0
August00	2	0	1
September00	2	0	1
October00	1	0	1
November00	1	0	1
December00	2	0	1
January01	3	0	1
February01	2	0	0
March01	1	0	0
April01	3	0	3
May01	5	0	2
June01	2	0	2
July01	2	0	2
August01	1	0	0
September01	1	0	0
October01	2	0	1

(a) ARIMA

Months	p	d	q	P	D	Q
June00	4	0	3	4	0	0
July00	1	0	0	0	0	2
August00	2	0	1	2	0	1
September00	2	0	1	1	0	0
October00	1	0	1	0	0	0
November00	1	0	2	0	0	1
December00	2	0	1	0	0	0
January01	1	0	1	2	0	0
February01	2	0	0	0	0	1
March01	1	0	0	1	0	2
April01	2	0	0	1	0	0
May01	3	0	2	1	0	0
June01	2	0	0	0	0	2
July01	1	0	0	2	0	0
August01	3	0	1	1	0	0
September01	1	0	1	0	0	1
October01	2	0	1	2	0	0

(b) SARIMA (s=24)

Months	p
June00	3
July00	3
August00	4
September00	3
October00	2
November00	3
December00	3
January01	3
February01	4
March01	3
April01	6
May01	3
June01	6
July01	5
August01	3
September01	3
October01	4

(c) VAR

APPENDIX B

NORDTANK TURBINE POWER CURVE

Table B.1. Power Curve Data of Turbine

Speed [m/s]	Power [kW]	Thrust coefficient
4.00	0	0.8
5.00	27	0.82
6.00	67	0.85
7.00	117	0.82
8.00	179	0.78
9.00	245	0.74
10.00	306	0.68
11.00	362	0.62
12.00	413	0.55
13.00	456	0.49
14.00	490	0.43
15.00	515	0.38
16.00	534	0.32
17.00	546	0.28
18.00	554	0.25
19.00	555	0.21
20.00	550	0.2
21.00	542	0.19
22.00	531	0.17
23.00	520	0.16
24.00	510	0.15
25.00	501	0.14

APPENDIX C

CODES

C.1. URLA RES Accuracy Table and Figures Code

```
require(graphics)
library(astsa)
library(forecast)
library(vars)
library(nlme)
library(stats)
library(xtable)
june0 <- read.table("june00.txt",header=F)
june00 <- na.omit(june0)
P <-june00$V1
T <-june00$V2
H <-june00$V3
W <-june00$V4
D <-june00$V5
nlldata <- cbind(P,W,D,H)
nltest <- VAR(nldata, type = c("const"),season = NULL, exogen = T, lag.max = 6,ic = c("
  AIC"))
nlunivariatetimeseries <- ts(june00$V4, frequency=24, start=c(1), end=c(30))
nlfit1 <- auto.arima(nlunivariatetimeseries, max.q=10,max.d=10,max.p=10,max.order=5,
  start.p=0, start.q=0,seasonal=FALSE,stationary=TRUE,trace=FALSE)
nlfit2 <- auto.arima(nlunivariatetimeseries,max.q=10,max.d=10,max.p=10,max.P=4, max.Q=4,
  max.order=5,start.p=0, start.q=0, start.P=0, start.Q=0,seasonal=TRUE,stationary=
  TRUE,trace=FALSE)
write.table(residuals(nlfit2), file = "fittedS.txt",append = TRUE,col.names = FALSE,row.
  names=FALSE)
write.table(residuals(nlfit1), file = "fittedA.txt",append = TRUE,col.names = FALSE,row.
  names=FALSE)
write.table(residuals(nltest$varresult[[2]]), file = "fittedV.txt",append = TRUE,col.
  names =FALSE,row.names=FALSE)
jpeg('figURLA1.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality =
  100)
plot(ts(W, frequency=24, start=c(1), end=c(31)), col="red",type = "p",pch = 19,lwd=0.5,
  main="_Fitted_Wind_data_&_Raw_Wind_data", sub="JUNE-2000", xlab="Date_(days)", ylab
  ="Wind_Speed_(m/s)")
```

```

points(ts(fitted(n1fit1), frequency=24, start=c(1), end=c(31)),col="black",pch = 4,lwd
=0.5)
points(ts(fitted(n1fit2), frequency=24, start=c(1), end=c(31)),col="blueviolet",pch = 4,
lwd=0.5)
points(ts(n1test$varresult[[2]]$fitted,frequency=24, start=c(1), end=c(31)),col="green",
pch = 4,lwd=0.5)
par(xpd=TRUE)
legend(.5,0,c("RAW","ARIMA","SARIMA","VAR"),lty=c(1,1),col= c("red", "black","blueviolet
","green"), lwd=c(1.5,1.5,1.5,1.5), text.col = "black",horiz=FALSE,merge = TRUE,bt
= "n")
dev.off()
july0 <- read.table("july00.txt",header=F)
july00<- na.omit(july0)
P <-july00$V1
T <-july00$V2
H <-july00$V3
W <-july00$V4
D <-july00$V5
n2data <- cbind(P,W,D, H)
n2ttest <- VAR(n2data, type = c("const"),season = NULL, exogen = T, lag.max = 6,ic = c("
AIC"))
n2univariatetimeseries <- ts(july00$V4, frequency=24, start=c(1), end=c(31))
n2fit1 <- auto.arima(n2univariatetimeseries, max.q=10,max.d=10,max.p=10,max.order=5,
start.p=0, start.q=0,seasonal=FALSE,stationary=TRUE,trace=FALSE)
n2fit2 <- auto.arima(n2univariatetimeseries,max.q=10,max.d=10,max.p=10,max.P=4, max.Q=4,
max.order=5,start.p=0, start.q=0, start.P=0, start.Q=0,seasonal=TRUE,stationary=
TRUE,trace=FALSE)
write.table(residuals(n2fit2), file = "fittedS.txt",append = TRUE,col.names = FALSE,row.
names=FALSE)
write.table(residuals(n2fit1), file = "fittedA.txt",append = TRUE,col.names = FALSE,row.
names=FALSE)
write.table(residuals(n2ttest$varresult[[2]]), file = "fittedV.txt",append = TRUE,col.
names =FALSE,row.names=FALSE)
jpeg('figURLA2.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality =
100)
plot(ts(W, frequency=24, start=c(1), end=c(31)), col="red",type = "p",pch = 19,lwd=0.5,
main="_Fitted_Wind_data_&_Raw_Wind_data", sub="JULY-2000", xlab="Date_(days)", ylab
="Wind_Speed_(m/s)")
points(ts(fitted(n2fit1), frequency=24, start=c(1), end=c(31)),col="black",pch = 4,lwd
=0.5)
points(ts(n2ttest$varresult[[2]]$fitted,frequency=24, start=c(1), end=c(31)),col="green",
pch = 4,lwd=0.5)
points(ts(fitted(n2fit2), frequency=24, start=c(1), end=c(31)),col="blueviolet",pch = 4,
lwd=0.5)
par(xpd=TRUE)
legend(.5,0,c("RAW","ARIMA","SARIMA","VAR"),lty=c(1,1),col= c("red", "black","blueviolet
","green"), lwd=c(1.5,1.5,1.5,1.5), text.col = "black",horiz=FALSE,merge = TRUE,bt

```

```

= "n")
dev.off()
august0 <-read.table("august00.txt",header=F)
august00<- na.omit(august0)
P <-august00$V1
T <-august00$V2
H <-august00$V3
W <-august00$V4
D <-august00$V5
n3data <- cbind(P,W,D, H)
n3test <- VAR(n3data, type = c("const"),season = NULL, exogen = T, lag.max = 6,ic = c("
AIC"))
n3univariatetimeseries <- ts(august00$V4, frequency=24, start=c(1), end=c(31))
n3fit1 <- auto.arima(n3univariatetimeseries, max.q=10,max.d=10,max.p=10,max.order=5,
start.p=0, start.q=0,seasonal=FALSE,stationary=TRUE,trace=FALSE)
n3fit2 <- auto.arima(n3univariatetimeseries,max.q=10,max.d=10,max.p=10,max.P=4, max.Q=4,
max.order=5,start.p=0, start.q=0, start.P=0, start.Q=0,seasonal=TRUE,stationary=
TRUE,trace=FALSE)
write.table(residuals(n3fit2), file = "fittedS.txt",append = TRUE,col.names = FALSE,row.
names=FALSE)
write.table(residuals(n3fit1), file = "fittedA.txt",append = TRUE,col.names = FALSE,row.
names=FALSE)
write.table(residuals(n3test$varresult[[2]]), file = "fittedV.txt",append = TRUE,col.
names =FALSE,row.names=FALSE)
jpeg('figURLA3.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality =
100)
plot(ts(W, frequency=24, start=c(1), end=c(31)), col="red",type ="p",pch = 19,lwd=0.5,
main="_Fitted_Wind_data_&_Raw_Wind_data", sub="AUGUST-2000", xlab="Date_(days)",
ylab="Wind_Speed_(m/s)")
points(ts(fitted(n3fit1), frequency=24, start=c(1), end=c(31)),col="black",pch = 4,lwd
=0.5)
points(ts(n3test$varresult[[2]]$fitted,frequency=24, start=c(1), end=c(31)),col="green",
pch = 4,lwd=0.5)
points(ts(fitted(n3fit2), frequency=24, start=c(1), end=c(31)),col="blueviolet",pch = 4,
lwd=0.5)
par(xpd=TRUE)
legend(.5,0,c("RAW","ARIMA","SARIMA","VAR"),lty=c(1,1),col= c("red", "black","blueviolet
","green"), lwd=c(1.5,1.5,1.5,1.5), text.col = "black",horiz=FALSE,merge = TRUE,bty
= "n")
dev.off()
september0 <- read.table("september00.txt",header=F)
september00<- na.omit(september0)
P <-september00$V1
T <-september00$V2
H <-september00$V3
W <-september00$V4
D <-september00$V5

```

```

n4data <- cbind(P,W,D, H)
n4test <- VAR(n4data, type = c("const"),season = NULL, exogen = T, lag.max = 6,ic = c("
  AIC"))
n4univariatetimeseries <- ts(september00$V4, frequency=24, start=c(1), end=c(30))
n4fit1 <- auto.arima(n4univariatetimeseries, max.q=10,max.d=10,max.p=10,max.order=5,
  start.p=0, start.q=0,seasonal=FALSE,stationary=TRUE,trace=FALSE)
n4fit2 <- auto.arima(n4univariatetimeseries,max.q=10,max.d=10,max.p=10,max.P=4, max.Q=4,
  max.order=5,start.p=0, start.q=0, start.P=0, start.Q=0,seasonal=TRUE,stationary=
  TRUE,trace=FALSE)
write.table(residuals(n4fit2), file = "fittedS.txt",append = TRUE,col.names = FALSE,row.
  names=FALSE)
write.table(residuals(n4fit1), file = "fittedA.txt",append = TRUE,col.names = FALSE,row.
  names=FALSE)
write.table(residuals(n4test$varresult[[2]]), file = "fittedV.txt",append = TRUE,col.
  names =FALSE,row.names=FALSE)
jpeg('figURLA4.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality =
  100)
plot(ts(W, frequency=24, start=c(1), end=c(31)), col="red",type = "p",pch = 19,lwd=0.5,
  main="_Fitted_Wind_data_&_Raw_Wind_data", sub="September-2000", xlab="Date_(days)",
  ylab="Wind_Speed_(m/s) ")
points(ts(fitted(n4fit1), frequency=24, start=c(1), end=c(31)),col="black",pch = 4,lwd
  =0.5)
points(ts(n4test$varresult[[2]]$fitted,frequency=24, start=c(1), end=c(31)),col="green",
  pch = 4,lwd=0.5)
points(ts(fitted(n4fit2), frequency=24, start=c(1), end=c(31)),col="blueviolet",pch = 4,
  lwd=0.5)
par(xpd=TRUE)
legend(.5,0,c("RAW","ARIMA","SARIMA","VAR"),lty=c(1,1),col= c("red", "black","blueviolet
  ", "green"), lwd=c(1.5,1.5,1.5,1.5), text.col = "black",horiz=FALSE,merge = TRUE,bty
  = "n")
dev.off()
october0 <- read.table("october00.txt",header=F)
october00<- na.omit(october0)
P <-october00$V1
T <-october00$V2
H <-october00$V3
W <-october00$V4
D <-october00$V5
n5data <- cbind(P,W,D, H)
n5test <- VAR(n5data, type = c("const"),season = NULL, exogen = T, lag.max = 6,ic = c("
  AIC"))
n5univariatetimeseries <- ts(october00$V4, frequency=24, start=c(1), end=c(31))
n5fit1 <- auto.arima(n5univariatetimeseries, max.q=10,max.d=10,max.p=10,max.order=5,
  start.p=0, start.q=0,seasonal=FALSE,stationary=TRUE,trace=FALSE)
n5fit2 <- auto.arima(n5univariatetimeseries,max.q=10,max.d=10,max.p=10,max.P=4, max.Q=4,
  max.order=5,start.p=0, start.q=0, start.P=0, start.Q=0,seasonal=TRUE,stationary=
  TRUE,trace=FALSE)

```

```

write.table(residuals(n5fit2), file = "fittedS.txt", append = TRUE, col.names = FALSE, row.
names=FALSE)
write.table(residuals(n5fit1), file = "fittedA.txt", append = TRUE, col.names = FALSE, row.
names=FALSE)
write.table(residuals(n5test$varresult[[2]]), file = "fittedV.txt", append = TRUE, col.
names =FALSE, row.names=FALSE)
jpeg('figURLA5.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality =
100)
plot(ts(W, frequency=24, start=c(1), end=c(31)), col="red",type ="p",pch = 19,lwd=0.5,
main="_Fitted_Wind_data_&_Raw_Wind_data", sub="October-2000", xlab="Date_(days)",
ylab="Wind_Speed_(m/s) ")
points(ts(fitted(n5fit1), frequency=24, start=c(1), end=c(31)),col="black",pch = 4,lwd
=0.5)
points(ts(n5test$varresult[[2]]$fitted,frequency=24, start=c(1), end=c(31)),col="green",
pch = 4,lwd=0.5)
points(ts(fitted(n5fit2), frequency=24, start=c(1), end=c(31)),col="blueviolet",pch = 4,
lwd=0.5)
par(xpd=TRUE)
legend(.5,0,c("RAW","ARIMA","SARIMA","VAR"),lty=c(1,1),col= c("red", "black","blueviolet
","green"), lwd=c(1.5,1.5,1.5,1.5), text.col = "black",horiz=FALSE,merge = TRUE,bty
= "n")
dev.off()
november0 <- read.table("november00.txt",header=F)
november00<- na.omit(november0)
P <-november00$V1
T <-november00$V2
H <-november00$V3
W <-november00$V4
D <-november00$V5
n6data <- cbind(P,W,D, H)
n6test <- VAR(n6data, type = c("const"),season = NULL, exogen = T, lag.max = 6,ic = c("
AIC"))
n6univariatetimeseries <- ts(november00$V4, frequency=24, start=c(1), end=c(30))
n6fit1 <- auto.arima(n6univariatetimeseries, max.q=10,max.d=10,max.p=10,max.order=5,
start.p=0, start.q=0,seasonal=FALSE,stationary=TRUE,trace=FALSE)
n6fit2 <- auto.arima(n6univariatetimeseries,max.q=10,max.d=10,max.p=10,max.P=4, max.Q=4,
max.order=5,start.p=0, start.q=0, start.P=0, start.Q=0,seasonal=TRUE,stationary=
TRUE,trace=FALSE)
write.table(residuals(n6fit2), file = "fittedS.txt",append = TRUE,col.names = FALSE,row.
names=FALSE)
write.table(residuals(n6fit1), file = "fittedA.txt",append = TRUE,col.names = FALSE,row.
names=FALSE)
write.table(residuals(n6test$varresult[[2]]), file = "fittedV.txt",append = TRUE,col.
names =FALSE, row.names=FALSE)
jpeg('figURLA6.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality =
100)
plot(ts(W, frequency=24, start=c(1), end=c(31)), col="red",type ="p",pch = 19,lwd=0.5,

```

```

    main="_Fitted_Wind_data_&_Raw_Wind_data", sub="November-2000", xlab="Date_(days)",
    ylab="Wind_Speed_(m/s) ")
points(ts(fitted(n6fit1), frequency=24, start=c(1), end=c(31)),col="black",pch = 4,lwd
=0.5)
points(ts(n6test$varresult[[2]]$fitted,frequency=24, start=c(1), end=c(31)),col="green",
pch = 4,lwd=0.5)
points(ts(fitted(n6fit2), frequency=24, start=c(1), end=c(31)),col="blueviolet",pch = 4,
lwd=0.5)
par(xpd=TRUE)
legend(.5,0,c("RAW","ARIMA","SARIMA","VAR"),lty=c(1,1),col= c("red", "black","blueviolet
","green"), lwd=c(1.5,1.5,1.5,1.5), text.col = "black",horiz=FALSE,merge = TRUE,bty
= "n")
dev.off()
december0 <- read.table("december00.txt",header=F)
december00<- na.omit(december0)
P <-december00$V1
T <-december00$V2
H <-december00$V3
W <-december00$V4
D <-december00$V5
n7data <- cbind(P,W,D, H)
n7test <- VAR(n7data, type = c("const"),season = NULL, exogen = T, lag.max = 6,ic = c("
AIC"))
n7univariatetimeseries <- ts(december00$V4, frequency=24, start=c(1), end=c(31))
n7fit1 <- auto.arima(n7univariatetimeseries, max.q=10,max.d=10,max.p=10,max.order=5,
start.p=0, start.q=0,seasonal=FALSE,stationary=TRUE,trace=FALSE)
n7fit2 <- auto.arima(n7univariatetimeseries,max.q=10,max.d=10,max.p=10,max.P=4, max.Q=4,
max.order=5,start.p=0, start.q=0, start.P=0, start.Q=0,seasonal=TRUE,stationary=
TRUE,trace=FALSE)
write.table(residuals(n7fit2), file = "fittedS.txt",append = TRUE,col.names = FALSE,row.
names=FALSE)
write.table(residuals(n7fit1), file = "fittedA.txt",append = TRUE,col.names = FALSE,row.
names=FALSE)
write.table(residuals(n7test$varresult[[2]]), file = "fittedV.txt",append = TRUE,col.
names =FALSE,row.names=FALSE)
jpeg('figURLA7.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality =
100)
plot(ts(W, frequency=24, start=c(1), end=c(31)), col="red",type="p",pch = 19,lwd=0.5,
main="_Fitted_Wind_data_&_Raw_Wind_data", sub="December-2000", xlab="Date_(days)",
ylab="Wind_Speed_(m/s) ")
points(ts(fitted(n7fit1), frequency=24, start=c(1), end=c(31)),col="black",pch = 4,lwd
=0.5)
points(ts(n7test$varresult[[2]]$fitted,frequency=24, start=c(1), end=c(31)),col="green",
pch = 4,lwd=0.5)
points(ts(fitted(n7fit2), frequency=24, start=c(1), end=c(31)),col="blueviolet",pch = 4,
lwd=0.5)
par(xpd=TRUE)

```

```

legend(.5,0,c("RAW","ARIMA","SARIMA","VAR"),lty=c(1,1),col= c("red", "black","blueviolet",
  "green"), lwd=c(1.5,1.5,1.5,1.5), text.col = "black",horiz=FALSE,merge = TRUE,bty
  = "n")

dev.off()

january1 <- read.table("january01.txt",header=F)
january01<- na.omit(january1)

P <-january01$V1
T <-january01$V2
H <-january01$V3
W <-january01$V4
D <-january01$V5

n8data <- cbind(P,W,D, H)
n8test <- VAR(n8data, type = c("const"),season = NULL, exogen = T, lag.max = 6,ic = c("
  AIC"))

n8univariatetimeseries <- ts(january01$V4, frequency=24, start=c(1), end=c(31))
n8fit1 <- auto.arima(n8univariatetimeseries, max.q=10,max.d=10,max.p=10,max.order=5,
  start.p=0, start.q=0,seasonal=FALSE,stationary=TRUE,trace=FALSE)
n8fit2 <- auto.arima(n8univariatetimeseries,max.q=10,max.d=10,max.p=10,max.P=4, max.Q=4,
  max.order=5,start.p=0, start.q=0, start.P=0, start.Q=0,seasonal=TRUE,stationary=
  TRUE,trace=FALSE)

write.table(residuals(n8fit2), file = "fittedS.txt",append = TRUE,col.names = FALSE,row.
names=FALSE)

write.table(residuals(n8fit1), file = "fittedA.txt",append = TRUE,col.names = FALSE,row.
names=FALSE)

write.table(residuals(n8test$varresult[[2]]), file = "fittedV.txt",append = TRUE,col.
names =FALSE,row.names=FALSE)

jpeg('figURLA8.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality =
  100)

plot(ts(W, frequency=24, start=c(1), end=c(31)), col="red",type = "p",pch = 19,lwd=0.5,
  main="_Fitted_Wind_data_&_Raw_Wind_data", sub="January-2001", xlab="Date_(days)",
  ylab="Wind_Speed_(m/s) ")

points(ts(fitted(n8fit1), frequency=24, start=c(1), end=c(31)),col="black",pch = 4,lwd
  =0.5)

points(ts(n8test$varresult[[2]]$fitted,frequency=24, start=c(1), end=c(31)),col="green",
  pch = 4,lwd=0.5)

points(ts(fitted(n8fit2), frequency=24, start=c(1), end=c(31)),col="blueviolet",pch = 4,
  lwd=0.5)

par(xpd=TRUE)

legend(.5,0,c("RAW","ARIMA","SARIMA","VAR"),lty=c(1,1),col= c("red", "black","blueviolet",
  "green"), lwd=c(1.5,1.5,1.5,1.5), text.col = "black",horiz=FALSE,merge = TRUE,bty
  = "n")

dev.off()

february1 <- read.table("february01.txt",header=F)
february01<- na.omit(february1)

P <-february01$V1
T <-february01$V2
H <-february01$V3

```



```

W <-february01$V4
D <-february01$V5
n9data <- cbind(P,W,D, H)
n9test <- VAR(n9data, type = c("const"),season = NULL, exogen = T, lag.max = 6,ic = c("
AIC"))
n9univariatetimeseries <- ts(february01$V4, frequency=24, start=c(1), end=c(28))
n9fit1 <- auto.arima(n9univariatetimeseries, max.q=10,max.d=10,max.p=10,max.order=5,
start.p=0, start.q=0,seasonal=FALSE,stationary=TRUE,trace=FALSE)
n9fit2 <- auto.arima(n9univariatetimeseries,max.q=10,max.d=10,max.p=10,max.P=4, max.Q=4,
max.order=5,start.p=0, start.q=0, start.P=0, start.Q=0,seasonal=TRUE,stationary=
TRUE,trace=FALSE)
write.table(residuals(n9fit2), file = "fittedS.txt",append = TRUE,col.names = FALSE,row.
names=FALSE)
write.table(residuals(n9fit1), file = "fittedA.txt",append = TRUE,col.names = FALSE,row.
names=FALSE)
write.table(residuals(n9test$varresult[[2]]), file = "fittedV.txt",append = TRUE,col.
names =FALSE,row.names=FALSE)
jpeg('figURLA9.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality =
100)
plot(ts(W, frequency=24, start=c(1), end=c(31)), col="red",type="p",pch = 19,lwd=0.5,
main="_Fitted_Wind_data_&_Raw_Wind_data", sub="February-2001", xlab="Date_(days)",
ylab="Wind_Speed_(m/s) ")
points(ts(fitted(n9fit1), frequency=24, start=c(1), end=c(31)),col="black",pch = 4,lwd
=0.5)
points(ts(n9test$varresult[[2]]$fitted,frequency=24, start=c(1), end=c(31)),col="green",
pch = 4,lwd=0.5)
points(ts(fitted(n9fit2), frequency=24, start=c(1), end=c(31)),col="blueviolet",pch = 4,
lwd=0.5)
par(xpd=TRUE)
legend(.5,0,c("RAW","ARIMA","SARIMA","VAR"),lty=c(1,1),col= c("red", "black","blueviolet
","green"), lwd=c(1.5,1.5,1.5,1.5), text.col = "black",horiz=FALSE,merge = TRUE,bty
= "n")
dev.off()
march1 <- read.table("march01.txt",header=F)
march01<- na.omit(march1)
P <-march01$V1
T <-march01$V2
H <-march01$V3
W <-march01$V4
D <-march01$V5
n10data <- cbind(P,W,D, H)
n10test <- VAR(n10data, type = c("const"),season = NULL, exogen = T, lag.max = 6,ic = c
("AIC"))
n10univariatetimeseries <- ts(march01$V4, frequency=24, start=c(1), end=c(31))
n10fit1 <- auto.arima(n10univariatetimeseries, max.q=10,max.d=10,max.p=10,max.order=5,
start.p=0, start.q=0,seasonal=FALSE,stationary=TRUE,trace=FALSE)
n10fit2 <- auto.arima(n10univariatetimeseries,max.q=10,max.d=10,max.p=10,max.P=4, max.Q

```

```

=4, max.order=5, start.p=0, start.q=0, start.P=0, start.Q=0, seasonal=TRUE, stationary=
TRUE, trace=FALSE)
write.table(residuals(n10fit2), file = "fittedS.txt", append = TRUE, col.names = FALSE, row
.names=FALSE)
write.table(residuals(n10fit1), file = "fittedA.txt", append = TRUE, col.names = FALSE, row
.names=FALSE)
write.table(residuals(n10test$varresult[[2]]), file = "fittedV.txt", append = TRUE, col.
names =FALSE, row.names=FALSE)
jpeg('figURLA10.jpg', width = 800, height = 400, units = "px", pointsize = 12, quality =
100)
plot(ts(W, frequency=24, start=c(1), end=c(31)), col="red", type="p", pch = 19, lwd=0.5,
main="_Fitted_Wind_data_&_Raw_Wind_data", sub="March-2001", xlab="Date_(days)",
ylab="Wind_Speed_(m/s)")
points(ts(fitted(n10fit1), frequency=24, start=c(1), end=c(31)), col="black", pch = 4, lwd
=0.5)
points(ts(n10test$varresult[[2]]$fitted, frequency=24, start=c(1), end=c(31)), col="green"
, pch = 4, lwd=0.5)
points(ts(fitted(n10fit2), frequency=24, start=c(1), end=c(31)), col="blueviolet", pch =
4, lwd=0.5)
par(xpd=TRUE)
legend(.5, 0, c("RAW", "ARIMA", "SARIMA", "VAR"), lty=c(1,1), col= c("red", "black", "blueviolet
", "green"), lwd=c(1.5, 1.5, 1.5, 1.5), text.col = "black", horiz=FALSE, merge = TRUE, bty
= "n")
dev.off()
april1 <- read.table("april01.txt", header=F)
april01 <- na.omit(april1)
P <- april01$V1
T <- april01$V2
H <- april01$V3
W <- april01$V4
D <- april01$V5
n11data <- cbind(P, W, D, H)
n11test <- VAR(n11data, type = c("const"), season = NULL, exogen = T, lag.max = 6, ic = c
("AIC"))
n11univariatetimeseries <- ts(april01$V4, frequency=24, start=c(1), end=c(30))
n11fit1 <- auto.arima(n11univariatetimeseries, max.q=10, max.d=10, max.p=10, max.order=5,
start.p=0, start.q=0, seasonal=FALSE, stationary=TRUE, trace=FALSE)
n11fit2 <- auto.arima(n11univariatetimeseries, max.q=10, max.d=10, max.p=10, max.P=4, max.Q
=4, max.order=5, start.p=0, start.q=0, start.P=0, start.Q=0, seasonal=TRUE, stationary=
TRUE, trace=FALSE)
write.table(residuals(n11fit2), file = "fittedS.txt", append = TRUE, col.names = FALSE, row
.names=FALSE)
write.table(residuals(n11fit1), file = "fittedA.txt", append = TRUE, col.names = FALSE, row
.names=FALSE)
write.table(residuals(n11test$varresult[[2]]), file = "fittedV.txt", append = TRUE, col.
names =FALSE, row.names=FALSE)
jpeg('figURLA11.jpg', width = 800, height = 400, units = "px", pointsize = 12, quality =

```

```

100)
plot(ts(W, frequency=24, start=c(1), end=c(31)), col="red",type ="p",pch = 19,lwd=0.5,
     main="_Fitted_Wind_data_&_Raw_Wind_data", sub="April-2001", xlab="Date_(days)",
     ylab="Wind_Speed_(m/s)")
points(ts(fitted(n11fit1), frequency=24, start=c(1), end=c(31)),col="black",pch = 4,lwd
       =0.5)
points(ts(n11test$varresult[[2]]$fitted,frequency=24, start=c(1), end=c(31)),col="green"
       ,pch = 4,lwd=0.5)
points(ts(fitted(n11fit2), frequency=24, start=c(1), end=c(31)),col="blueviolet",pch =
       4,lwd=0.5)
par(xpd=TRUE)
legend(.5,0,c("RAW", "ARIMA", "SARIMA", "VAR"),lty=c(1,1),col= c("red", "black", "blueviolet"
    ,"green"), lwd=c(1.5,1.5,1.5,1.5), text.col = "black",horiz=FALSE,merge = TRUE,bty
    = "n")
dev.off()
may1 <- read.table("may01.txt",header=F)
may01<- na.omit(may1)
P <-may01$V1
T <-may01$V2
H <-may01$V3
W <-may01$V4
D <-may01$V5
n12data <- cbind(P,W,D, H)
n12test <- VAR(n12data, type = c("const"),season = NULL, exogen = T, lag.max = 6,ic = c
    ("AIC"))
n12univariatetimeseries <- ts(may01$V4, frequency=24, start=c(1), end=c(31))
n12fit1 <- auto.arima(n12univariatetimeseries, max.q=10,max.d=10,max.p=10,max.order=5,
    start.p=0, start.q=0,seasonal=FALSE,stationary=TRUE,trace=FALSE)
n12fit2 <- auto.arima(n12univariatetimeseries,max.q=10,max.d=10,max.p=10,max.P=4, max.Q
    =4, max.order=5,start.p=0, start.q=0, start.P=0, start.Q=0,seasonal=TRUE,stationary=
    TRUE,trace=FALSE)
write.table(residuals(n12fit2), file = "fittedS.txt",append = TRUE,col.names = FALSE,row
    .names=FALSE)
write.table(residuals(n12fit1), file = "fittedA.txt",append = TRUE,col.names = FALSE,row
    .names=FALSE)
write.table(residuals(n12test$varresult[[2]]), file = "fittedV.txt",append = TRUE,col.
    names =FALSE,row.names=FALSE)
jpeg('figURLA12.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality =
    100)
plot(ts(W, frequency=24, start=c(1), end=c(31)), col="red",type ="p",pch = 19,lwd=0.5,
     main="_Fitted_Wind_data_&_Raw_Wind_data", sub="May-2001", xlab="Date_(days)", ylab=
     "Wind_Speed_(m/s)")
points(ts(fitted(n12fit1), frequency=24, start=c(1), end=c(31)),col="black",pch = 4,lwd
       =0.5)
points(ts(n12test$varresult[[2]]$fitted,frequency=24, start=c(1), end=c(31)),col="green"
       ,pch = 4,lwd=0.5)
points(ts(fitted(n12fit2), frequency=24, start=c(1), end=c(31)),col="blueviolet",pch =

```

```

      4, lwd=0.5)
par(xpd=TRUE)
legend(.5, 0, c("RAW", "ARIMA", "SARIMA", "VAR"), lty=c(1,1), col= c("red", "black", "blueviolet",
  "green"), lwd=c(1.5,1.5,1.5,1.5), text.col = "black", horiz=FALSE, merge = TRUE, btj
  = "n")
dev.off()
june1 <- read.table("june01.txt", header=F)
june01<- na.omit(june1)
P <-june01$V1
T <-june01$V2
H <-june01$V3
W <-june01$V4
D <-june01$V5
n13data <- cbind(P,W,D, H)
n13test <- VAR(n13data, type = c("const"), season = NULL, exogen = T, lag.max = 6, ic = c
  ("AIC"))
n13univariatetimeseries <- ts(june01$V4, frequency=24, start=c(1), end=c(30))
n13fit1 <- auto.arima(n13univariatetimeseries, max.q=10,max.d=10,max.p=10,max.order=5,
  start.p=0, start.q=0, seasonal=FALSE, stationary=TRUE, trace=FALSE)
n13fit2 <- auto.arima(n13univariatetimeseries, max.q=10,max.d=10,max.p=10,max.P=4, max.Q
  =4, max.order=5, start.p=0, start.q=0, start.P=0, start.Q=0, seasonal=TRUE, stationary=
  TRUE, trace=FALSE)
write.table(residuals(n13fit2), file = "fittedS.txt", append = TRUE, col.names = FALSE, row
  .names=FALSE)
write.table(residuals(n13fit1), file = "fittedA.txt", append = TRUE, col.names = FALSE, row
  .names=FALSE)
write.table(residuals(n13test$varresult[[2]]), file = "fittedV.txt", append = TRUE, col.
  names =FALSE, row.names=FALSE)
jpeg('figURLA13.jpg', width = 800, height = 400, units = "px", pointsize = 12, quality =
  100)
plot(ts(W, frequency=24, start=c(1), end=c(31)), col="red", type = "p", pch = 19, lwd=0.5,
  main="_Fitted_Wind_data_&_Raw_Wind_data", sub="June-2001", xlab="Date_(days)", ylab
  ="Wind_Speed_(m/s) ")
points(ts(fitted(n13fit1), frequency=24, start=c(1), end=c(31)), col="black", pch = 4, lwd
  =0.5)
points(ts(n13test$varresult[[2]]$fitted, frequency=24, start=c(1), end=c(31)), col="green"
  , pch = 4, lwd=0.5)
points(ts(fitted(n13fit2), frequency=24, start=c(1), end=c(31)), col="blueviolet", pch =
  4, lwd=0.5)
par(xpd=TRUE)
legend(.5, 0, c("RAW", "ARIMA", "SARIMA", "VAR"), lty=c(1,1), col= c("red", "black", "blueviolet",
  "green"), lwd=c(1.5,1.5,1.5,1.5), text.col = "black", horiz=FALSE, merge = TRUE, btj
  = "n")
dev.off()
july1 <- read.table("july01.txt", header=F)
july01<- na.omit(july1)
P <-july01$V1

```

```

T    <-july01$V2
H    <-july01$V3
W    <-july01$V4
D    <-july01$V5
n14data    <-    cbind(P,W,D, H)
n14test <- VAR(n14data, type = c("const"),season = NULL, exogen = T, lag.max = 6,ic = c
("AIC"))
n14univariatetimeseries <- ts(july01$V4, frequency=24, start=c(1), end=c(31))
n14fit1 <- auto.arima(n14univariatetimeseries, max.q=10,max.d=10,max.p=10,max.order=5,
start.p=0, start.q=0,seasonal=FALSE,stationary=TRUE,trace=FALSE)
n14fit2 <- auto.arima(n14univariatetimeseries,max.q=10,max.d=10,max.p=10,max.P=4, max.Q
=4, max.order=5,start.p=0, start.q=0, start.P=0, start.Q=0,seasonal=TRUE,stationary=
TRUE,trace=FALSE)
write.table(residuals(n14fit2), file = "fittedS.txt",append = TRUE,col.names = FALSE,row
.names=FALSE)
write.table(residuals(n14fit1), file = "fittedA.txt",append = TRUE,col.names = FALSE,row
.names=FALSE)
write.table(residuals(n14test$varresult[[2]]), file = "fittedV.txt",append = TRUE,col.
names =FALSE,row.names=FALSE)
jpeg('figURLA14.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality =
100)
plot(ts(W, frequency=24, start=c(1), end=c(31)), col="red",type = "p",pch = 19,lwd=0.5,
main="_Fitted_Wind_data_&_Raw_Wind_data", sub="July-2001", xlab="Date_(days)", ylab
="Wind_Speed_(m/s)")
points(ts(fitted(n14fit1), frequency=24, start=c(1), end=c(31)),col="black",pch = 4,lwd
=0.5)
points(ts(n14test$varresult[[2]]$fitted,frequency=24, start=c(1), end=c(31)),col="green"
,pch = 4,lwd=0.5)
points(ts(fitted(n14fit2), frequency=24, start=c(1), end=c(31)),col="blueviolet",pch =
4,lwd=0.5)
par(xpd=TRUE)
legend(.5,0,c("RAW","ARIMA","SARIMA","VAR"),lty=c(1,1),col= c("red", "black","blueviolet
","green"), lwd=c(1.5,1.5,1.5,1.5), text.col = "black",horiz=FALSE,merge = TRUE,bty
= "n")
dev.off()
august1 <- read.table("august01.txt",header=F)
august01<- na.omit(august1)
P    <-august01$V1
T    <-august01$V2
H    <-august01$V3
W    <-august01$V4
D    <-august01$V5
n15data    <-    cbind(P,W,D, H)
n15test <- VAR(n15data, type = c("const"),season = NULL, exogen = T, lag.max = 6,ic = c
("AIC"))
n15univariatetimeseries <- ts(august01$V4, frequency=24, start=c(1), end=c(31))
n15fit1 <- auto.arima(n15univariatetimeseries, max.q=10,max.d=10,max.p=10,max.order=5,

```

```

      start.p=0, start.q=0, seasonal=FALSE, stationary=TRUE, trace=FALSE)
n15fit2 <- auto.arima(n15univariatetimeseries, max.q=10, max.d=10, max.p=10, max.P=4, max.Q
  =4, max.order=5, start.p=0, start.q=0, start.P=0, start.Q=0, seasonal=TRUE, stationary=
  TRUE, trace=FALSE)
write.table(residuals(n15fit2), file = "fittedS.txt", append = TRUE, col.names = FALSE, row
  .names=FALSE)
write.table(residuals(n15fit1), file = "fittedA.txt", append = TRUE, col.names = FALSE, row
  .names=FALSE)
write.table(residuals(n15test$varresult[[2]]), file = "fittedV.txt", append = TRUE, col.
  names =FALSE, row.names=FALSE)
jpeg('figURLA15.jpg', width = 800, height = 400, units = "px", pointsize = 12, quality =
  100)
plot(ts(W, frequency=24, start=c(1), end=c(31)), col="red", type = "p", pch = 19, lwd=0.5,
  main="Fitted_Wind_data_&Raw_Wind_data", sub="August-2001", xlab="Date_(days)",
  ylab="Wind_Speed_(m/s)")
points(ts(fitted(n15fit1), frequency=24, start=c(1), end=c(31)), col="black", pch = 4, lwd
  =0.5)
points(ts(n15test$varresult[[2]]$fitted, frequency=24, start=c(1), end=c(31)), col="green"
  , pch = 4, lwd=0.5)
points(ts(fitted(n15fit2), frequency=24, start=c(1), end=c(31)), col="blueviolet", pch =
  4, lwd=0.5)
par(xpd=TRUE)
legend(.5, 0, c("RAW", "ARIMA", "SARIMA", "VAR"), lty=c(1,1), col= c("red", "black", "blueviolet
  ", "green"), lwd=c(1.5, 1.5, 1.5, 1.5), text.col = "black", horiz=FALSE, merge = TRUE, bty
  = "n")
dev.off()
september01 <- read.table("september01.txt", header=F)
september1 <- na.omit(september01)
P <- september01$V1
T <- september01$V2
H <- september01$V3
W <- september01$V4
D <- september01$V5
n16data <- cbind(P, W, D, H)
n16test <- VAR(n16data, type = c("const"), season = NULL, exogen = T, lag.max = 6, ic = c
  ("AIC"))
n16univariatetimeseries <- ts(september01$V4, frequency=24, start=c(1), end=c(30))
n16fit1 <- auto.arima(n16univariatetimeseries, max.q=10, max.d=10, max.p=10, max.order=5,
  start.p=0, start.q=0, seasonal=FALSE, stationary=TRUE, trace=FALSE)
n16fit2 <- auto.arima(n16univariatetimeseries, max.q=10, max.d=10, max.p=10, max.P=4, max.Q
  =4, max.order=5, start.p=0, start.q=0, start.P=0, start.Q=0, seasonal=TRUE, stationary=
  TRUE, trace=FALSE)
write.table(residuals(n16fit2), file = "fittedS.txt", append = TRUE, col.names = FALSE, row
  .names=FALSE)
write.table(residuals(n16fit1), file = "fittedA.txt", append = TRUE, col.names = FALSE, row
  .names=FALSE)
write.table(residuals(n16test$varresult[[2]]), file = "fittedV.txt", append = TRUE, col.

```

```

names =FALSE, row.names=FALSE)
jpeg('figURLA16.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality =
100)
plot(ts(W, frequency=24, start=c(1), end=c(31)), col="red",type="p",pch = 19,lwd=0.5,
main="_Fitted_Wind_data_&_Raw_Wind_data", sub="September-2001", xlab="Date_(days)",
ylab="Wind_Speed_(m/s)")
points(ts(fitted(n16fit1), frequency=24, start=c(1), end=c(31)),col="black",pch = 4,lwd
=0.5)
points(ts(n16test$varresult[[2]]$fitted,frequency=24, start=c(1), end=c(31)),col="green"
,pch = 4,lwd=0.5)
points(ts(fitted(n16fit2), frequency=24, start=c(1), end=c(31)),col="blueviolet",pch =
4,lwd=0.5)
par(xpd=TRUE)
legend(.5,0,c("RAW","ARIMA","SARIMA","VAR"),lty=c(1,1),col= c("red", "black","blueviolet
","green"), lwd=c(1.5,1.5,1.5,1.5), text.col = "black",horiz=FALSE,merge = TRUE,bty
= "n")
dev.off()
october1 <-read.table("october01.txt",header=F)
october01<- na.omit(october1)
P <-october01$V1
T <-october01$V2
H <-october01$V3
W <-october01$V4
D <-october01$V5
n17data <- cbind(P,W,D, H)
n17test <- VAR(n17data, type = c("const"),season = NULL, exogen = T, lag.max = 6,ic = c
("AIC"))
n17univariatetimeseries <- ts(october01$V4, frequency=24, start=c(1), end=c(31))
n17fit1 <- auto.arima(n17univariatetimeseries, max.q=10,max.d=10,max.p=10,max.order=5,
start.p=0, start.q=0,seasonal=FALSE,stationary=TRUE,trace=FALSE)
n17fit2 <- auto.arima(n17univariatetimeseries,max.q=10,max.d=10,max.p=10,max.P=4, max.Q
=4, max.order=5,start.p=0, start.q=0, start.P=0, start.Q=0,seasonal=TRUE,stationary=
TRUE,trace=FALSE)
write.table(residuals(n17fit2), file = "fittedS.txt",append = TRUE,col.names = FALSE,row
.names=FALSE)
write.table(residuals(n17fit1), file = "fittedA.txt",append = TRUE,col.names = FALSE,row
.names=FALSE)
write.table(residuals(n17test$varresult[[2]]), file = "fittedV.txt",append = TRUE,col.
names =FALSE, row.names=FALSE)
jpeg('figURLA17.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality =
100)
plot(ts(W, frequency=24, start=c(1), end=c(31)), col="red",type="p",pch = 19,lwd=0.5,
main="_Fitted_Wind_data_&_Raw_Wind_data", sub="October-2001", xlab="Date_(days)",
ylab="Wind_Speed_(m/s)")
points(ts(fitted(n17fit1), frequency=24, start=c(1), end=c(31)),col="black",pch = 4,lwd
=0.5)
points(ts(n17test$varresult[[2]]$fitted,frequency=24, start=c(1), end=c(31)),col="green"

```

```

, pch = 4, lwd=0.5)
points(ts(fitted(n17fit2), frequency=24, start=c(1), end=c(31)), col="blueviolet", pch =
4, lwd=0.5)
par(xpd=TRUE)
legend(.5, 0, c("RAW", "ARIMA", "SARIMA", "VAR"), lty=c(1,1), col= c("red", "black", "blueviolet",
"green"), lwd=c(1.5, 1.5, 1.5, 1.5), text.col = "black", horiz=FALSE, merge = TRUE, btj
= "n")
dev.off()
dogruluk1<- accuracy(n1test$varresult[[2]])
dogruluk2<- accuracy(n1fit1)
dogruluk4<- accuracy(n2test$varresult[[2]])
dogruluk5<- accuracy(n2fit1)
dogruluk7<- accuracy(n3test$varresult[[2]])
dogruluk8<- accuracy(n3fit1)
dogruluk10<- accuracy(n4test$varresult[[2]])
dogruluk11<- accuracy(n4fit1)
dogruluk13<- accuracy(n5test$varresult[[2]])
dogruluk14<- accuracy(n5fit1)
dogruluk16<- accuracy(n6test$varresult[[2]])
dogruluk17<- accuracy(n6fit1)
dogruluk19<- accuracy(n7test$varresult[[2]])
dogruluk20<- accuracy(n7fit1)
dogruluk22<- accuracy(n8test$varresult[[2]])
dogruluk23<- accuracy(n8fit1)
dogruluk25<- accuracy(n9test$varresult[[2]])
dogruluk26<- accuracy(n9fit1)
dogruluk28<- accuracy(n10test$varresult[[2]])
dogruluk29<- accuracy(n10fit1)
dogruluk31<- accuracy(n11test$varresult[[2]])
dogruluk32<- accuracy(n11fit1)
dogruluk34<- accuracy(n12test$varresult[[2]])
dogruluk35<- accuracy(n12fit1)
dogruluk37<- accuracy(n13test$varresult[[2]])
dogruluk38<- accuracy(n13fit1)
dogruluk40<- accuracy(n14test$varresult[[2]])
dogruluk41<- accuracy(n14fit1)
dogruluk43<- accuracy(n15test$varresult[[2]])
dogruluk44<- accuracy(n15fit1)
dogruluk46<- accuracy(n16test$varresult[[2]])
dogruluk47<- accuracy(n16fit1)
dogruluk49<- accuracy(n17test$varresult[[2]])
dogruluk50<- accuracy(n17fit1)
d1<-accuracy(n1fit2)
d2<-accuracy(n2fit2)
d3<-accuracy(n3fit2)
d4<-accuracy(n4fit2)
d5<-accuracy(n5fit2)

```



```

d6<-accuracy (n6fit2)
d7<-accuracy (n7fit2)
d8<-accuracy (n8fit2)
d9<-accuracy (n9fit2)
d10<-accuracy (n10fit2)
d11<-accuracy (n11fit2)
d12<-accuracy (n12fit2)
d13<-accuracy (n13fit2)
d14<-accuracy (n14fit2)
d15<-accuracy (n15fit2)
d16<-accuracy (n16fit2)
d17<-accuracy (n17fit2)
totalA<- rbind(dogruluk2, dogruluk5, dogruluk8, dogruluk11, dogruluk14, dogruluk17, dogruluk20
, dogruluk23, dogruluk26, dogruluk29, dogruluk32, dogruluk35, dogruluk38, dogruluk41,
, dogruluk44, dogruluk47, dogruluk50)
totalV<- rbind(dogruluk1, dogruluk4, dogruluk7, dogruluk10, dogruluk13, dogruluk16, dogruluk19
, dogruluk22, dogruluk25, dogruluk28, dogruluk31, dogruluk34, dogruluk37, dogruluk40,
, dogruluk43, dogruluk46, dogruluk49)
totalS<- rbind(d1, d2, d3, d4, d5, d6, d7, d8, d9, d10, d11, d12, d13, d14, d15, d16, d17)
model1<-rbind(n1fit1, n2fit1, n3fit1, n4fit1, n5fit1, n6fit1, n7fit1, n8fit1, n9fit1, n10fit1,
, n11fit1, n12fit1, n13fit1, n14fit1, n15fit1, n16fit1, n17fit1)
model2<-rbind(n1fit2, n2fit2, n3fit2, n4fit2, n5fit2, n6fit2, n7fit2, n8fit2, n9fit2, n10fit2,
, n11fit2, n12fit2, n13fit2, n14fit2, n15fit2, n16fit2, n17fit2)
model3<-rbind(n1test$varresult[[2]], n2test$varresult[[2]], n3test$varresult[[2]], n4test$
varresult[[2]], n5test$varresult[[2]], n6test$varresult[[2]], n7test$varresult[[2]],
, n8test$varresult[[2]], n9test$varresult[[2]], n10test$varresult[[2]], n11test$varresult
[[2]], n12test$varresult[[2]], n13test$varresult[[2]], n14test$varresult[[2]], n15test$
varresult[[2]], n16test$varresult[[2]], n17test$varresult[[2]])
write.table(model1, file = "Asummary.txt", col.names = TRUE, row.names=FALSE)
write.table(model2, file = "Ssummary.txt", col.names = TRUE, row.names=FALSE)
write.table(model3, file = "Vsummary.txt", col.names = TRUE, row.names=FALSE)
exo1<-june00$V2[1:24]
exo1=matrix(exo1, nrow=24, ncol=1)
exo2<-july00$V2[1:24]
exo2=matrix(exo2, nrow=24, ncol=1)
exo3<-august00$V2[1:24]
exo3=matrix(exo3, nrow=24, ncol=1)
exo4<-september00$V2[1:24]
exo4=matrix(exo4, nrow=24, ncol=1)
exo5<-october00$V2[1:24]
exo5=matrix(exo5, nrow=24, ncol=1)
exo6<-november00$V2[1:24]
exo6=matrix(exo6, nrow=24, ncol=1)
exo7<-december00$V2[1:24]
exo7=matrix(exo7, nrow=24, ncol=1)
exo8<-january01$V2[1:24]
exo8=matrix(exo8, nrow=24, ncol=1)

```

```

exo9<-february01$V2[1:24]
exo9=matrix(exo9,nrow=24,ncol=1)
exo10<-march01$V2[1:24]
exo10=matrix(exo10,nrow=24,ncol=1)
exo11<-april01$V2[1:24]
exo11=matrix(exo11,nrow=24,ncol=1)
exo12<-may01$V2[1:24]
exo12=matrix(exo12,nrow=24,ncol=1)
exo13<-june01$V2[1:24]
exo13=matrix(exo13,nrow=24,ncol=1)
exo14<-july01$V2[1:24]
exo14=matrix(exo14,nrow=24,ncol=1)
exo15<-august01$V2[1:24]
exo15=matrix(exo15,nrow=24,ncol=1)
exo16<-september01$V2[1:24]
exo16=matrix(exo16,nrow=24,ncol=1)
exo17<-october01$V2[1:24]
exo17=matrix(exo17,nrow=24,ncol=1)
va1<-predict(n1test, n.ahead = 24, ci = 0.95,dumvar=exo1)
va2<-predict(n2test, n.ahead = 24, ci = 0.95,dumvar=exo2)
va3<-predict(n3test, n.ahead = 24, ci = 0.95,dumvar=exo3)
va4<-predict(n4test, n.ahead = 24, ci = 0.95,dumvar=exo4)
va5<-predict(n5test, n.ahead = 24, ci = 0.95,dumvar=exo5)
va6<-predict(n6test, n.ahead = 24, ci = 0.95,dumvar=exo6)
va7<-predict(n7test, n.ahead = 24, ci = 0.95,dumvar=exo7)
va8<-predict(n8test, n.ahead = 24, ci = 0.95,dumvar=exo8)
va9<-predict(n9test, n.ahead = 24, ci = 0.95,dumvar=exo9)
va10<-predict(n10test, n.ahead = 24, ci = 0.95,dumvar=exo10)
va11<-predict(n11test, n.ahead = 24, ci = 0.95,dumvar=exo11)
va12<-predict(n12test, n.ahead = 24, ci = 0.95,dumvar=exo12)
va13<-predict(n13test, n.ahead = 24, ci = 0.95,dumvar=exo13)
va14<-predict(n14test, n.ahead = 24, ci = 0.95,dumvar=exo14)
va15<-predict(n15test, n.ahead = 24, ci = 0.95,dumvar=exo15)
va16<-predict(n16test, n.ahead = 24, ci = 0.95,dumvar=exo16)
va17<-predict(n17test, n.ahead = 24, ci = 0.95,dumvar=exo17)
a1<-forecast(n1fit1,h=24)
a2<-forecast(n2fit1,h=24)
a3<-forecast(n3fit1,h=24)
a4<-forecast(n4fit1,h=24)
a5<-forecast(n5fit1,h=24)
a6<-forecast(n6fit1,h=24)
a7<-forecast(n7fit1,h=24)
a8<-forecast(n8fit1,h=24)
a9<-forecast(n9fit1,h=24)
a10<-forecast(n10fit1,h=24)
a11<-forecast(n11fit1,h=24)
a12<-forecast(n12fit1,h=24)

```

```

a13<-forecast (nl3fit1,h=24)
a14<-forecast (nl4fit1,h=24)
a15<-forecast (nl5fit1,h=24)
a16<-forecast (nl6fit1,h=24)
a17<-forecast (nl7fit1,h=24)
sa1<-forecast (nlfit2,h=24)
sa2<-forecast (n2fit2,h=24)
sa3<-forecast (n3fit2,h=24)
sa4<-forecast (n4fit2,h=24)
sa5<-forecast (n5fit2,h=24)
sa6<-forecast (n6fit2,h=24)
sa7<-forecast (n7fit2,h=24)
sa8<-forecast (n8fit2,h=24)
sa9<-forecast (n9fit2,h=24)
sa10<-forecast (n10fit2,h=24)
sa11<-forecast (n11fit2,h=24)
sa12<-forecast (n12fit2,h=24)
sa13<-forecast (n13fit2,h=24)
sa14<-forecast (n14fit2,h=24)
sa15<-forecast (n15fit2,h=24)
sa16<-forecast (n16fit2,h=24)
sa17<-forecast (n17fit2,h=24)

jpeg('figforecast1.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality
= 100)
plot(ts(july00$V4[1:24], frequency=1, start=c(1), end=c(24)), col="red",type = "p",pch =
19,lwd=0.5,main="_Forecasted_Wind_data_&_Next_Month_Recorded_Wind_data", sub="July
-01-2000", xlab="Date", ylab="Wind_Speed_(m/s)")

points(ts(a1$mean[1:24], frequency=1, start=c(1), end=c(24)),col="black",pch = 4,lwd
=0.5)

points(ts(val$fcst$W[1:24],frequency=1, start=c(1), end=c(24)),col="green",pch = 4,lwd
=0.5)

points(ts(sa1$mean[1:24], frequency=1, start=c(1), end=c(24)),col="blueviolet",pch = 4,
lwd=0.5)

dev.off()

jpeg('figforecast2.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality
= 100)
plot(ts(august00$V4[1:24], frequency=1, start=c(1), end=c(24)), col="red",type = "p",pch
= 19,lwd=0.5,main="_Forecasted_Wind_data_&_Next_Month_Recorded_Wind_data", sub="
August-01-2000", xlab="Date", ylab="Wind_Speed_(m/s)")

points(ts(a2$mean[1:24], frequency=1, start=c(1), end=c(24)),col="black",pch = 4,lwd

```

```

=0.5)

points(ts(va2$fcst$W[1:24], frequency=1, start=c(1), end=c(24)), col="green", pch = 4, lwd
=0.5)

points(ts(sa2$mean[1:24], frequency=1, start=c(1), end=c(24)), col="blueviolet", pch = 4,
lwd=0.5)

dev.off()

jpeg('figforecast3.jpg', width = 800, height = 400, units = "px", pointsize = 12, quality
= 100)
plot(ts(september00$V4[1:24], frequency=1, start=c(1), end=c(24)), col="red", type="p",
pch = 19, lwd=0.5, main="_Forecasted_Wind_data_&_Next_Month_Recorded_Wind_data", sub=
"September-01-2000", xlab="Date", ylab="Wind_Speed_(m/s)")

points(ts(a3$mean[1:24], frequency=1, start=c(1), end=c(24)), col="black", pch = 4, lwd
=0.5)

points(ts(va3$fcst$W[1:24], frequency=1, start=c(1), end=c(24)), col="green", pch = 4, lwd
=0.5)

points(ts(sa3$mean[1:24], frequency=1, start=c(1), end=c(24)), col="blueviolet", pch = 4,
lwd=0.5)

dev.off()

jpeg('figforecast4.jpg', width = 800, height = 400, units = "px", pointsize = 12, quality
= 100)
plot(ts(october00$V4[1:24], frequency=1, start=c(1), end=c(24)), col="red", type="p", pch
= 19, lwd=0.5, main="_Forecasted_Wind_data_&_Next_Month_Recorded_Wind_data", sub="
October-01-2000", xlab="Date", ylab="Wind_Speed_(m/s)")

points(ts(a4$mean[1:24], frequency=1, start=c(1), end=c(24)), col="black", pch = 4, lwd
=0.5)

points(ts(va4$fcst$W[1:24], frequency=1, start=c(1), end=c(24)), col="green", pch = 4, lwd
=0.5)

points(ts(sa4$mean[1:24], frequency=1, start=c(1), end=c(24)), col="blueviolet", pch = 4,
lwd=0.5)

dev.off()

```

```

jpeg('figforecast5.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality
= 100)
plot(ts(november00$V4[1:24], frequency=1, start=c(1), end=c(24)), col="red",type ="p",
  pch = 19,lwd=0.5,main="_Forecasted_Wind_data_&_Next_Month_Recorded_Wind_data", sub=
  "November-01-2000", xlab="Date", ylab="Wind_Speed_(m/s)")

points(ts(a5$mean[1:24], frequency=1, start=c(1), end=c(24)),col="black",pch = 4,lwd
=0.5)

points(ts(va5$fcst$W[1:24],frequency=1, start=c(1), end=c(24)),col="green",pch = 4,lwd
=0.5)

points(ts(sa5$mean[1:24], frequency=1, start=c(1), end=c(24)),col="blueviolet",pch = 4,
  lwd=0.5)

dev.off()

jpeg('figforecast6.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality
= 100)
plot(ts(december00$V4[1:24], frequency=1, start=c(1), end=c(24)), col="red",type ="p",
  pch = 19,lwd=0.5,main="_Forecasted_Wind_data_&_Next_Month_Recorded_Wind_data", sub=
  "December-01-2000", xlab="Date", ylab="Wind_Speed_(m/s)")

points(ts(a6$mean[1:24], frequency=1, start=c(1), end=c(24)),col="black",pch = 4,lwd
=0.5)

points(ts(va6$fcst$W[1:24],frequency=1, start=c(1), end=c(24)),col="green",pch = 4,lwd
=0.5)

points(ts(sa6$mean[1:24], frequency=1, start=c(1), end=c(24)),col="blueviolet",pch = 4,
  lwd=0.5)

dev.off()

jpeg('figforecast7.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality
= 100)
plot(ts(january01$V4[1:24], frequency=1, start=c(1), end=c(24)), col="red",type ="p",pch
= 19,lwd=0.5,main="_Forecasted_Wind_data_&_Next_Month_Recorded_Wind_data", sub="
  January-01-2001", xlab="Date", ylab="Wind_Speed_(m/s)")

points(ts(a7$mean[1:24], frequency=1, start=c(1), end=c(24)),col="black",pch = 4,lwd
=0.5)

points(ts(va7$fcst$W[1:24],frequency=1, start=c(1), end=c(24)),col="green",pch = 4,lwd

```

```

=0.5)

points(ts(sa7$mean[1:24], frequency=1, start=c(1), end=c(24)),col="blueviolet",pch = 4,
      lwd=0.5)

dev.off()

jpeg('figforecast8.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality
     = 100)
plot(ts(february01$V4[1:24], frequency=1, start=c(1), end=c(24)), col="red",type ="p",
     pch = 19,lwd=0.5,main="_Forecasted_Wind_data_&_Next_Month_Recorded_Wind_data", sub=
     "February-01-2001", xlab="Date", ylab="Wind_Speed_(m/s)")

points(ts(a8$mean[1:24], frequency=1, start=c(1), end=c(24)),col="black",pch = 4,lwd
      =0.5)

points(ts(va8$fcst$W[1:24],frequency=1, start=c(1), end=c(24)),col="green",pch = 4,lwd
      =0.5)

points(ts(sa8$mean[1:24], frequency=1, start=c(1), end=c(24)),col="blueviolet",pch = 4,
      lwd=0.5)

dev.off()

jpeg('figforecast9.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality
     = 100)
plot(ts(march01$V4[1:24], frequency=1, start=c(1), end=c(24)), col="red",type ="p",pch =
     19,lwd=0.5,main="_Forecasted_Wind_data_&_Next_Month_Recorded_Wind_data", sub="
     March-01-2001", xlab="Date", ylab="Wind_Speed_(m/s)")

points(ts(a9$mean[1:24], frequency=1, start=c(1), end=c(24)),col="black",pch = 4,lwd
      =0.5)

points(ts(va9$fcst$W[1:24],frequency=1, start=c(1), end=c(24)),col="green",pch = 4,lwd
      =0.5)

points(ts(sa9$mean[1:24], frequency=1, start=c(1), end=c(24)),col="blueviolet",pch = 4,
      lwd=0.5)

dev.off()

jpeg('figforecast10.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality
     = 100)

```

```

plot(ts(april01$V4[1:24], frequency=1, start=c(1), end=c(24)), col="red",type ="p",pch =
  19,lwd=0.5,main="_Forecasted_Wind_data_&_Next_Month_Recorded_Wind_data", sub="
  April-01-2001", xlab="Date", ylab="Wind_Speed_(m/s)")

points(ts(a10$mean[1:24], frequency=1, start=c(1), end=c(24)),col="black",pch = 4,lwd
  =0.5)

points(ts(va10$fcst$W[1:24],frequency=1, start=c(1), end=c(24)),col="green",pch = 4,lwd
  =0.5)

points(ts(sa10$mean[1:24], frequency=1, start=c(1), end=c(24)),col="blueviolet",pch = 4,
  lwd=0.5)

dev.off()

jpeg('figforecast11.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality
  = 100)
plot(ts(may01$V4[1:24], frequency=1, start=c(1), end=c(24)), col="red",type ="p",pch =
  19,lwd=0.5,main="_Forecasted_Wind_data_&_Next_Month_Recorded_Wind_data", sub="May
  -01-2001", xlab="Date", ylab="Wind_Speed_(m/s)")

points(ts(a11$mean[1:24], frequency=1, start=c(1), end=c(24)),col="black",pch = 4,lwd
  =0.5)

points(ts(va11$fcst$W[1:24],frequency=1, start=c(1), end=c(24)),col="green",pch = 4,lwd
  =0.5)

points(ts(sa11$mean[1:24], frequency=1, start=c(1), end=c(24)),col="blueviolet",pch = 4,
  lwd=0.5)

par(xpd=TRUE)
legend(.5,3,c("RAW","ARIMA","SARIMA","VAR"),lty=c(1,1),col= c("red","black","blueviolet"
  ,"green"), lwd=c(1.5,1.5,1.5,1.5), text.col = "black",horiz=FALSE,merge = TRUE,bty =
  "n")
dev.off()

jpeg('figforecast12.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality
  = 100)
plot(ts(june01$V4[1:24], frequency=1, start=c(1), end=c(24)), col="red",type ="p",pch =
  19,lwd=0.5,main="_Forecasted_Wind_data_&_Next_Month_Recorded_Wind_data", sub="June
  -01-2001", xlab="Date", ylab="Wind_Speed_(m/s)")

points(ts(a12$mean[1:24], frequency=1, start=c(1), end=c(24)),col="black",pch = 4,lwd
  =0.5)

points(ts(va12$fcst$W[1:24],frequency=1, start=c(1), end=c(24)),col="green",pch = 4,lwd

```

```

=0.5)

points(ts(sa12$mean[1:24], frequency=1, start=c(1), end=c(24)),col="blueviolet",pch = 4,
      lwd=0.5)

dev.off()

jpeg('figforecast13.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality
     = 100)
plot(ts(july01$V4[1:24], frequency=1, start=c(1), end=c(24)), col="red",type ="p",pch =
     19,lwd=0.5,main="_Forecasted_Wind_data_&_Next_Month_Recorded_Wind_data", sub="July
     -01-2001", xlab="Date", ylab="Wind_Speed_(m/s)")

points(ts(a13$mean[1:24], frequency=1, start=c(1), end=c(24)),col="black",pch = 4,lwd
      =0.5)

points(ts(val13$fcst$W[1:24],frequency=1, start=c(1), end=c(24)),col="green",pch = 4,lwd
      =0.5)

points(ts(sa13$mean[1:24], frequency=1, start=c(1), end=c(24)),col="blueviolet",pch = 4,
      lwd=0.5)

dev.off()

jpeg('figforecast14.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality
     = 100)
plot(ts(august01$V4[1:24], frequency=1, start=c(1), end=c(24)), col="red",type ="p",pch
     = 19,lwd=0.5,main="_Forecasted_Wind_data_&_Next_Month_Recorded_Wind_data", sub="
     August-01-2001", xlab="Date", ylab="Wind_Speed_(m/s)")

points(ts(a14$mean[1:24], frequency=1, start=c(1), end=c(24)),col="black",pch = 4,lwd
      =0.5)

points(ts(val14$fcst$W[1:24],frequency=1, start=c(1), end=c(24)),col="green",pch = 4,lwd
      =0.5)

points(ts(sa14$mean[1:24], frequency=1, start=c(1), end=c(24)),col="blueviolet",pch = 4,
      lwd=0.5)

dev.off()

jpeg('figforecast15.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality
     = 100)
plot(ts(september01$V4[1:24], frequency=1, start=c(1), end=c(24)), col="red",type ="p",
     pch = 19,lwd=0.5,main="_Forecasted_Wind_data_&_Next_Month_Recorded_Wind_data", sub=
     "September-01-2001", xlab="Date", ylab="Wind_Speed_(m/s)")

```



```

points(ts(a15$mean[1:24], frequency=1, start=c(1), end=c(24)),col="black",pch = 4,lwd
=0.5)

points(ts(va15$fcst$W[1:24],frequency=1, start=c(1), end=c(24)),col="green",pch = 4,lwd
=0.5)

points(ts(sa15$mean[1:24], frequency=1, start=c(1), end=c(24)),col="blueviolet",pch = 4,
lwd=0.5)

dev.off()

jpeg('figforecast16.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality
= 100)
plot(ts(october01$V4[1:24], frequency=1, start=c(1), end=c(24)), col="red",type ="p",pch
= 19,lwd=0.5,main="Forecasted_Wind_data_&_Next_Month_Recorded_Wind_data", sub="
October-01-2001", xlab="Date", ylab="Wind_Speed_(m/s)")

points(ts(a16$mean[1:24], frequency=1, start=c(1), end=c(24)),col="black",pch = 4,lwd
=0.5)

points(ts(va16$fcst$W[1:24],frequency=1, start=c(1), end=c(24)),col="green",pch = 4,lwd
=0.5)

points(ts(sa16$mean[1:24], frequency=1, start=c(1), end=c(24)),col="blueviolet",pch = 4,
lwd=0.5)

dev.off()

jpeg('figforURLAA1.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality
= 100)
plot(forecast(n1fit1,h=24))
dev.off()
jpeg('figforURLAS1.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality
= 100)
plot(forecast(n1fit2,h=24))
dev.off()
jpeg('figforURLAA2.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality
= 100)
plot(forecast(n2fit1,h=24))
dev.off()
jpeg('figforURLAS2.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality
= 100)
plot(forecast(n2fit2,h=24))
dev.off()

```

```

jpeg('figforURLAA3.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality
    = 100)
plot(forecast(n3fit1,h=24))
dev.off()
jpeg('figforURLAS3.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality
    = 100)
plot(forecast(n3fit2,h=24))
dev.off()
jpeg('figforURLAA4.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality
    = 100)
plot(forecast(n4fit1,h=24))
dev.off()
jpeg('figforURLAS4.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality
    = 100)
plot(forecast(n4fit2,h=24))
dev.off()
jpeg('figforURLAA5.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality
    = 100)
plot(forecast(n5fit1,h=24))
dev.off()
jpeg('figforURLAS5.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality
    = 100)
plot(forecast(n5fit2,h=24))
dev.off()
jpeg('figforURLAA6.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality
    = 100)
plot(forecast(n6fit1,h=24))
dev.off()
jpeg('figforURLAS6.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality
    = 100)
plot(forecast(n6fit2,h=24))
dev.off()
jpeg('figforURLAA7.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality
    = 100)
plot(forecast(n7fit1,h=24))
dev.off()
jpeg('figforURLAS7.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality
    = 100)
plot(forecast(n7fit2,h=24))
dev.off()
jpeg('figforURLAA8.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality
    = 100)
plot(forecast(n8fit1,h=24))
dev.off()
jpeg('figforURLAS8.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality
    = 100)
plot(forecast(n8fit2,h=24))

```

```

dev.off()
jpeg('figforURLAA9.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality
     = 100)
plot(forecast(n9fit1,h=24))
dev.off()
jpeg('figforURLAS9.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality
     = 100)
plot(forecast(n9fit2,h=24))
dev.off()
jpeg('figforURLAA10.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality
     = 100)
plot(forecast(n10fit1,h=24))
dev.off()
jpeg('figforURLAS10.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality
     = 100)
plot(forecast(n10fit2,h=24))
dev.off()
jpeg('figforURLAA11.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality
     = 100)
plot(forecast(n11fit1,h=24))
dev.off()
jpeg('figforURLAS11.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality
     = 100)
plot(forecast(n11fit2,h=24))
dev.off()
jpeg('figforURLAA12.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality
     = 100)
plot(forecast(n12fit1,h=24))
dev.off()
jpeg('figforURLAS1.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality
     = 100)
plot(forecast(n12fit2,h=24))
dev.off()
jpeg('figforURLAA13.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality
     = 100)
plot(forecast(n13fit1,h=24))
dev.off()
jpeg('figforURLAS13.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality
     = 100)
plot(forecast(n13fit2,h=24))
dev.off()
jpeg('figforURLAA14.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality
     = 100)
plot(forecast(n14fit1,h=24))
dev.off()
jpeg('figforURLAS14.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality
     = 100)

```

```

plot(forecast(nl4fit2,h=24))
dev.off()
jpeg('figforURLAA15.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality
    = 100)
plot(forecast(nl5fit1,h=24))
dev.off()
jpeg('figforURLAS15.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality
    = 100)
plot(forecast(nl5fit2,h=24))
dev.off()
jpeg('figforURLAA16.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality
    = 100)
plot(forecast(nl6fit1,h=24))
dev.off()
jpeg('figforURLAS16.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality
    = 100)
plot(forecast(nl6fit2,h=24))
dev.off()
jpeg('figforURLAA17.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality
    = 100)
plot(forecast(nl7fit1,h=24))
dev.off()
jpeg('figforURLAS17.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality
    = 100)
plot(forecast(nl7fit2,h=24))
dev.off()
totalA<-round(totalA,digits=3)
totalV<-round(totalV,digits=3)
totalS<-round(totalS,digits=3)
print(xtable(totalA), type="latex", file="outputA.tex")
print(xtable(totalV), type="latex", file="outputV.tex")
print(xtable(totalS), type="latex", file="outputS.tex")
write.table(totalA, file = "accuracyonefileA.txt", col.names = TRUE,row.names=FALSE)
write.table(totalV, file = "accuracyonefileV.txt", col.names = TRUE,row.names=FALSE)
write.table(totalS, file = "accuracyonefileS.txt", col.names = TRUE,row.names=FALSE)
resV <-read.table("fittedV.txt",header=F)
resA <-read.table("fittedA.txt",header=F)
resS <-read.table("fittedS.txt",header=F)
jpeg('resfark1VURLA.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality
    = 100)
plot(ts(resV),type ="l",main='URLA-VAR', xlab='Observations', ylab='Residuals')
dev.off()
jpeg('resfark2URLA.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality
    = 100)
plot(ts(resA),type ="l",main='URLA-ARIMA', xlab='Observations', ylab='Residuals')
dev.off()
jpeg('resfarkSURLA.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality

```

```

= 100)
plot(ts(resS),type ="l",main='URLA-SARIMA', xlab='Observations', ylab='Residuals')
dev.off()

```

C.2. URLA RES IRF Code

```

require(graphics)
library(astsa)
library(forecast)
library(vars)
library(nlme)
library(stats)
library(circular)
june0 <- read.table("june00.txt", header=F)
july0 <- read.table("july00.txt", header=F)
august0 <- read.table("august00.txt", header=F)
september0 <- read.table("september00.txt", header=F)
october0 <- read.table("october00.txt", header=F)
november0 <- read.table("november00.txt",header=F)
december0 <- read.table("december00.txt",header=F)
january1 <- read.table("january01.txt",header=F)
february1 <- read.table("february01.txt",header=F)
march1 <- read.table("march01.txt",header=F)
april1 <- read.table("april01.txt",header=F)
may1 <- read.table("may01.txt",header=F)
june1 <- read.table("june01.txt",header=F)
july1 <- read.table("july01.txt",header=F)
august1 <- read.table("august01.txt",header=F)
september01 <- read.table("september01.txt",header=F)
october1 <- read.table("october01.txt",header=F)
datam <- rbind(june0, july0, august0, september0, october0, november0, december0, january1,
               february1, march1, april1, may1, june1, july1, august1, september01, october1)
newdata <- na.omit(datam)
P <-ts(newdata$V1)
T <-ts(newdata$V2)
H <-ts(newdata$V3)
W <-ts(newdata$V4)
D <-ts(newdata$V5)
sum <- cbind(P,W,D, H)
total <- VAR(sum, type = c("const"),season = NULL, exogen = cbind(T),lag.max = 6,ic = c
("AIC"))
irf.P<- irf(total, impulse = c ("P"), response = c("W"), n.ahead=24,otho =TRUE,
cumulative =TRUE, boot=FALSE, ci=0.95)

```

```

irf.H<- irf(total, impulse = c("H"), response = c("W"), n.ahead=24,otho =TRUE,
  cumulative =TRUE, boot=FALSE, ci=0.95)
irf.D<- irf(total, impulse = c("D"), response = c("W"), n.ahead=24,otho =TRUE,
  cumulative =TRUE, boot=FALSE, ci=0.95)
irf.W<- irf(total, impulse = c("W"), response = c("W"), n.ahead=24,otho =TRUE,
  cumulative =TRUE, boot=FALSE, ci=0.95)
jpeg('irfP.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality = 100)
plot(irf.P)
dev.off()
jpeg('irfH.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality = 100)
plot(irf.H)
dev.off()
jpeg('irfD.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality = 100)
plot(irf.D)
dev.off()
jpeg('irfW.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality = 100)
plot(irf.W)
dev.off()
jpeg('fig1.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality = 100)
plot( W, col='black', type='l', main='Wind_Speed_Data', xlab='Time', ylab='Wind_speed_(m
/s)',xaxt='n')
text(seq(1, 12281, by = 740), par("usr")[3] - 0.2, labels =c("06-2000", "07-2000", "
08-2000","09-2000", "10-2000", "11-2000","12-2000", "01-2001", "02-2001","03-2001",
"04-2001", "05-2001","06-2001", "07-2001", "08-2001","09-2001", "10-2001"), srt =
90, pos = 2, xpd = TRUE)
dev.off()
jpeg('fig2.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality = 100)
plot( T, col='black', type='l', main='Temperature_Data', xlab='Time', ylab='Temperature_
(C)',xaxt='n')
text(seq(1, 12281, by = 740), par("usr")[3] - 0.2, labels =c("06-2000", "07-2000", "
08-2000","09-2000", "10-2000", "11-2000","12-2000", "01-2001", "02-2001","03-2001",
"04-2001", "05-2001","06-2001", "07-2001", "08-2001","09-2001", "10-2001"), srt =
90, pos = 2, xpd = TRUE)
dev.off()
jpeg('fig3.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality = 100)
plot( P, col='black', type='l', main='Pressure_Data', xlab='Time', ylab='Pressure_(hPa)'
,xaxt='n')
text(seq(1, 12281, by = 740), par("usr")[3] - 0.2, labels =c("06-2000", "07-2000", "
08-2000","09-2000", "10-2000", "11-2000","12-2000", "01-2001", "02-2001","03-2001",
"04-2001", "05-2001","06-2001", "07-2001", "08-2001","09-2001", "10-2001"), srt =
90, pos = 2, xpd = TRUE)
dev.off()
jpeg('fig4.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality = 100)
plot( H, col='black', type='l', main='Humidity_Data', xlab='Time', ylab='_%_to_%_RH
',xaxt='n')
text(seq(1, 12281, by = 740), par("usr")[3] - 0.2, labels =c("06-2000", "07-2000", "
08-2000","09-2000", "10-2000", "11-2000","12-2000", "01-2001", "02-2001","03-2001",

```

```

      "04-2001", "05-2001", "06-2001", "07-2001", "08-2001", "09-2001", "10-2001"), srt =
      90, pos = 2, xpd = TRUE)
dev.off()
jpeg('fig5.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality = 100)
univariatetimeseries <- ts(W, frequency=24, start=c(1), end=c(12432))
plot( acf(univariatetimeseries, 2400,na.action = na.contiguous),main='',xlab='Lag_(in_
      days)')
dev.off()
jpeg('fig6.jpg',width = 800, height = 400, units = "px", pointsize = 12,quality = 100)
univariatetimeseries <- ts(W, frequency=24, start=c(1), end=c(12432))
plot( pacf(univariatetimeseries,2400, xlim=c(1,20),na.action = na.contiguous),main='',
      xlab='Lag_(in_days)')
dev.off()

```

C.3. NORDTANK Filler

```

library(astsa)
library(forecast)
library(vars)
library(lubridate)
library(zoo)
nordtank.2 <- read.csv("~/Google_Drive/nursedata/nordtank-2.csv")
#nordtank.2<-nordtank.2[8837:116063,1:7]
date<-strptime(nordtank.2$run_name,"%Y%m%d%H%M")
WS35_5<-nordtank.2$WS35_5
WD33<-nordtank.2$WD33
T03<-nordtank.2$T03
P03<-nordtank.2$P03
POWER<-nordtank.2$POWER
STAT<-nordtank.2$STAT
nd<-data.frame(date,WS35_5,WD33,T03,P03,POWER,STAT)
#grid <- data.frame(date=seq.POSIXt(as.POSIXct("2004-04-08 13:40:00"),to=as.POSIXct
      ("2008-04-08 13:30:00"), by="10 mins"))
grid <- data.frame(date=seq.POSIXt(as.POSIXct("2003-12-17_16:10:00"),to=as.POSIXct("
      2008-04-08_13:30:00"), by="10_mins"))
namevector<-c("WS35_5","WD33")
datand<- merge(grid, nd, by="date", all.x=TRUE)
datef<-datand$date
WS35_5<-datand$WS35_5
WS35_5<-na.interp(WS35_5)
#W<-datand$WS35_5
checkp <- aggregate(nordtank.2, list(Date=format(date, "%Y-%m")),mean)
month<-month(as.POSIXlt(datand$date))
year<-year(as.POSIXlt(datand$date))

```

```

time<-cbind(month,year)
d<-na.aggregate(datand,as.yearmon)
check<-mean(datand[1:2064,2],na.rm=TRUE)
m<-cbind(datef[1:2064],datand$WS35_5[1:2064])
case<-na.aggregate(m,as.yearmon)
WD33<-datand$WD33
WD33<-na.interp(WD33)
D<-aggregate(datand$WD33, list(Date=format(datand$date, "%Y-%m")),mean)
T03<-datand$T03
T03<-na.interp(T03)
T<-aggregate(datand$T03, list(Date=format(datand$date, "%Y-%m")),mean)
P03<-datand$P03
P03<-na.interp(P03)
P<-aggregate(datand$P03, list(Date=format(datand$date, "%Y-%m")),mean)
POWER<-datand$POWER
POWER<-na.interp(POWER)
STAT<-datand$STAT
STAT<-na.interp(STAT)
naraw<-cbind(datef,WS35_5,WD33)
nldata <- cbind(WS35_5,WD33,P03)
nvunivariatetimeseries <- ts(WS35_5)
ndunivariatetimeseries <- ts(WD33)
fitSARIMAv <- auto.arima(nvunivariatetimeseries,max.q=10,max.d=10,max.p=10,max.P=4, max.Q=4, max.order=5,start.p=0, start.q=0, start.P=0, start.Q=0,seasonal=TRUE,stationary=TRUE,trace=FALSE)
fitSARIMAd <- auto.arima(ndunivariatetimeseries,max.q=10,max.d=10,max.p=10,max.P=4, max.Q=4, max.order=5,start.p=0, start.q=0, start.P=0, start.Q=0,seasonal=TRUE,stationary=TRUE,trace=FALSE)
fitVAR <- VAR(nldata, type = c("const"),season = NULL, exogen = cbind(T03), lag.max = 10,ic = c("AIC"))
Wind<-fitVAR$varresult[[1]]$fitted
Dir<-fitVAR$varresult[[2]]$fitted
WindS<-fitted(fitSARIMAv)
DirS<-fitted(fitSARIMAd)
fitARIMAv <- auto.arima(nvunivariatetimeseries,max.q=10,max.d=10,max.p=10,max.P=4, max.Q=4, max.order=5,start.p=0, start.q=0, start.P=0, start.Q=0,seasonal=FALSE,stationary=TRUE,trace=FALSE)
fitARIMAd <- auto.arima(ndunivariatetimeseries,max.q=10,max.d=10,max.p=10,max.P=4, max.Q=4, max.order=5,start.p=0, start.q=0, start.P=0, start.Q=0,seasonal=FALSE,stationary=TRUE,trace=FALSE)
WindA<-fitted(fitARIMAv)
WindA<-round(WindA,digits=2)
DirA<-fitted(fitARIMAd)
DirA<-round(DirA,digits=2)
setA<-data.frame(datef,WindA,DirA)
write.table(setA, file = "DATASETARIMA112.txt")

```



```

setS<-data.frame(datef,WindS,DirS)
write.table(setS, file = "DATASETSARIMA112.txt")

setV<-data.frame(datef[11:226637],Wind,Dir)
write.table(setV, file = "DATASETVAR112.txt")

naraw<-data.frame(datef,WS35_5,WD33)
write.table(naraw, file = "DATASETNA112.txt")

```

C.4. NORDTANK Cleaner Filler

```

library(astsa)
library(forecast)
library(vars)

nordtank.2 <- read.csv("~/Google_Drive/nursedata/nordtank-2.csv")
date<-strptime(nordtank.2$run_name,"%Y%m%d%H%M")
WS35_5<-nordtank.2$WS35_5
WD33<-nordtank.2$WD33
T03<-nordtank.2$T03
P03<-nordtank.2$P03
POWER<-nordtank.2$POWER
STAT<-nordtank.2$STAT
nd<-data.frame(date,WS35_5,WD33,T03,P03,POWER,STAT)
grid <- data.frame(date=seq.POSIXt(as.POSIXct("2003-12-17_16:10:00"),to=as.POSIXct("
  2008-12-17_16:10:00"), by="10_mins"))
datand<- merge(grid, nd, by="date", all.x=TRUE)
input<-datand[1:263091,1:7]

cleaner<-function(x) {
nlunivariatetimeseries2 <- ts(x$WS35_5)
#n1<-match(NA,(nlunivariatetimeseries2))
b<-which(is.na(x$WS35_5), arr.ind=TRUE)
b<-na.pass(b)
a<-as.numeric(b[1]-1)
#b<-as.numeric(b)
#if (b[1]<0) break
#n1<-which(is.na(x$WS35_5[1]))
fit <- auto.arima(nlunivariatetimeseries2[1:a],max.q=10,max.d=10,max.p=10,max.P=4, max.Q
  =4, max.order=5,start.p=0, start.q=0, start.P=0, start.Q=0,seasonal=TRUE,stationary=
  TRUE,trace=FALSE)
pred<- forecast(fit,h=1)
#nlunivariatetimeseries2[1:(b[1]-1)] <- fitted(fit)
nlunivariatetimeseries2[b[1]] <- pred$mean[1]
nlunivariatetimeseries2<-round(nlunivariatetimeseries2,digits=4)

```

```
return(nlunivariatetimeseries2)

}

for(i in 1:nrow(input)) {
  repeat { input[i + 1] <- cleaner(input); write.table(input$WS35_5, file = "
    windarimaforecastedSARIMA.txt"); if (input[i] <= nrow(input)) break; }
}
}
```