Umar Saeed

Resource-Efficient Widely-Linear Collaborative Wind Forecasting for Renewable Energy Generation

School of Electrical Engineering

Thesis submitted for examination for the degree of Master of Science in Technology.

Espoo 23.01.2012

Thesis supervisor:

D.Sc. (Tech.) Stefan Werner

Thesis instructor:

M.Sc. (Tech.) Taneli Riihonen

Aalto University
School of Electrical Engineering

AALTO UNIVERSITY
SCHOOL OF ELECTRICAL ENGINEERING

ABSTRACT OF THE
MASTER'S THESIS

Author: Umar Saeed

Title: Resource-Efficient Widely-Linear Collaborative Wind Forecasting for Renewable Energy Generation

Date: 23.01.2012 Language: English Number of pages:8+52

Department of Signal Processing and Acoustics

Professorship: Signal Processing Code: S-88

Supervisor: D.Sc. (Tech.) Stefan Werner

Instructor: M.Sc. (Tech.) Taneli Riihonen

Wind is an important source of renewable energy with abundant availability at many places of the world. The integration of wind energy into the electricity grid brings new challenges on the system reliability due to its intermittent nature. A robust tool with good accuracy for wind power forecasting is an essential part of wind based power system. Wind power is directly related to the cube of wind speed and hence a small improvement in the forecasting of wind speed provides larger improvement in wind power forecasts. The accuracy of a wind speed forecasting model depends upon the characteristics of wind signal incorporated in the model. Wind direction has significant effect on wind speed forecasts and the complex representation of wind signal provides a convenient tool to include the effect of direction. The complex wind signal is non-circular, and widely-linear modeling of the complex wind signal provides optimal results. Wind signal has strong temporal and spatial correlations as well. Most of the wind forecasting models at present do not incorporate all of these characteristics of the wind signal. In this thesis work, we will develop a widely-linear collaborative wind forecasting model which takes into account the non-circular nature of the complex wind signal and its temporal and spatial correlations as well. The parameters of the model are updated using recursive least squares approach. The forecasting model is further optimized for efficient resource utilization by using a set-membership formulation. A data model for the complex wind signal is developed and the algorithm is tested on the simulated wind data. Simulation results show better forecasting performance with the developed approach. Real world wind data is also used to verify the results.

Keywords: Forecasting, complex wind signal, non-circular, widely-linear, spatial, collaborative, recursive least squares, set-membership.

Preface

First of all, thanks to the Almighty Allah for the countless blessings bestowed upon me throughout my life.

The thanks to my supervisor D.Sc. (Tech.) Stefan Werner are difficult for me to put in words. This thesis would have not been possible without his constant guidance, patience and encouragement.

I acknowledge the constructive comments and reviews from my instructor M.Sc. (Tech.) Taneli Riihonen, which helped me to improve the structure and presentation of the thesis.

Many thanks to my friends Aamir, Pramod, Ali and others for always being there to cheer me up.

Last but not least, my family, whose constant encouragement and emotional support helped me to achieve all what I have in my life.

Otaniemi, 23.01.2012

Umar Saeed

Contents

List of Symbols and Abbreviations vi
List of Figures vii
List of Tables viii

1 Introduction **1**
1.1 Background 1
1.2 Research Problem and Scope 2
1.3 Contributions of the Thesis 3
1.4 Outline of the Thesis 3

2 Spatio-Temporal Analysis of Wind Signal **4**
2.1 Smart Grids and the Need for Wind Power Forecasting 4
2.2 Wind Power versus Wind Speed 5
2.3 Complex Nature of Wind Signal 6
2.4 Non-circularity of Complex Wind Signal 7
2.5 Temporal Analysis of Wind Signal 9
2.6 Spatial Correlation of Wind Signal 9

3 Wind Forecasting and Widely-Linear and Collaborative Processing of Wind Signal **13**
3.1 Overview of Wind Forecasting 13
3.2 Adaptive Algorithms for Short Term Wind Forecasting 15
3.2.1 Widely-Linear Processing of Complex Wind Signal 15
3.3 Performance Gain using Widely-Linear Approach 20
3.4 Collaborative Processing of Wind Signal 21
3.4.1 Multivariate Auto-Regressive Model 21
3.4.2 Decentralized Wind Flow Model and Prediction 22
3.4.3 Other Collaborative Schemes 24

4 Resource-Efficient Widely-Linear Collaborative Processing of Wind Signal **26**
4.1 Collaboration with Widely-Linear Processing 26
4.1.1 Widely-Linear System Model 27
4.1.2 Parameter Estimation 27
4.2 Resource-Efficient Processing 28
4.2.1 Update Equations 31

5 Simulations and Results **34**
5.1 Results with Simulated Wind Data 34
5.1.1 Data Model 35
5.1.2 Experiments and Results 37
5.2 Results with Real World Wind Data 44
5.2.1 Data Description 44
5.2.2 Experiments and Results 47

6 Conclusions **48**

References **49**

List of Symbols and Abbreviations

Symbols

v	wind speed
$\mathbf{W}$	wind power
z	complex valued wind signal
$\mathbf{C}$	auto-covariance matrix
$\mathbf{P}$	pseudo-covariance matrix
$\mathbf{p}$	input cross-correlation vector
p	size of the input data vector
J	objective function
λ	forgetting factor
γ	error bound

Abbreviations

WL	widely linear
pdf	probability density function
RLS	recursive least squares
LMS	least mean squares
AR	auto regressive
OBE	optimal bounding ellipsoids
SMF	set membership filtering

List of Figures

2.1 Power curve of a wind turbine generator. 6
2.2 Wind speed distribution over the direction shown as a polar plot. . . 7
2.3 Real-imaginary plot for real world wind signal (z). 8
2.4 Temporal correlation of real-world wind signal for the selected data (a) and zoomed in for clarity (b). 10
2.5 Spatial correlation of wind signal for selected data (a) and zoomed in for clarity in (b). 12

3.1 RLS and WL-RLS processing of real world wind data. 20
3.2 Effect of direction on collaborative wind forecasting with turbines in a straight line (a) and at an angle θ with each other. 25

5.1 Network of N nodes in straight line with spatial correlation ρ between two adjacent nodes. 36
5.2 The first 300 simulated wind speed data points at first node. 38
5.3 Non-circularity and spatial correlation of the simulated wind data at first node. 39
5.4 MSE behavior at first node with linear and widely-linear RLS. 40
5.5 MSE at first node with single node (no collaboration), two and three nodes collaboration processing. 41
5.6 RMSE plot for linear and WL RLS with no resource saving. 42
5.7 RMSE plot for linear and WL RLS with resource saving. 43
5.8 Effect of γ on output RMSE for different levels of collaboration. . . . 45
5.9 Effect of error threshold γ on system resources. 46

List of Tables

3.1 RMSE for linear and widely-linear modeling. 20

4.1 Linear collaborative RLS processing. 29
4.2 Widely-linear collaborative RLS processing. 29
4.3 Resource efficient linear collaborative RLS processing. 32
4.4 Resource efficient widely-linear collaborative RLS processing. 33

5.1 Resource saving. 44
5.2 RMSE values for linear and widely-linear resource efficient collaboration. 47
5.3 Resource saving. 47

Chapter 1

Introduction

Accurate forecasts of supply and demand in an electrical power system are essential for its smooth operation. Current trends towards renewable energy generation, driven by environmental and economical concerns, are mainly employing the natural phenomenons such as wind and solar energy for electricity generation. Wind is an important source of renewable energy with abundant availability at many places of the world. When wind energy is a part of electrical grid, it brings an uncertainty on the power supply due to its intermittent nature. Thus the need of a robust tool for the wind forecasting is even more evident. In this thesis work we will develop a wind forecasting model which improves the forecasting performance by exploiting the non-circularity and spatial correlation of the wind signal.

In Section 1.1, we will study the background of the wind forecasting problem. Then in Section 1.2, the research problem and scope of the thesis are stated. Section 1.3 consists of the contribution of the thesis work and the outline of the thesis report is given in Section 1.4.

1.1 Background

Wind energy is in fact a transformed form of the solar energy [1, 2]. The sun's radiation heats different parts of the earth surface at different rates. The effect of this uneven heating of earth surface causes hot air to rise up, reducing the pressure at earth surface which draws cooler air to replace it. The moving air, due to its mass, contains kinetic energy, which can be converted to electrical energy using a wind turbine [3].

A wind turbine basically has three key elements called a rotor, a nacelle (containing a gear box, the generator and control and monitoring equipment) and a tower [3]. The tower contains the blades which rotate upon striking of the wind. The rotation of the blades is the mechanism by which the kinetic energy of the wind is transformed into electrical energy [4].

Wind is an intermittent phenomenon and the integration of wind energy into an existing power system brings new challenges. The demand-supply chain in a traditional electrical power system is kept balanced by determining the demand for electricity in advance. When wind energy enters the electricity grid, it brings uncer-

tainty on the supply. In order to balance the supply and demand of the electricity, reliable and high quality forecasts of the wind power must be made in advance. In essence, a robust tool for wind power forecasting is essential for a power system for its reliable and economical operation when wind energy is part of the system [2,4,5].

Wind power could be predicted directly but that is highly dependent upon the types, sizes and the number of turbines in operation. The power output of a wind turbine is proportional to the cube of the wind speed, which means wind speed can be used to drive a prediction of the wind power for a given site. The prediction of wind speed is independent of the physical properties and the configuration of turbine and hence it provides more flexibility for power predictions. A small improvement in the wind speed prediction can provide large improvement in power predictions because of the cubic relationship between them [5].

There are a number wind speed forecasting techniques with varying degrees of accuracy based upon the approach being used. The accuracy of a forecasting technique depends upon the characteristics of the wind signal incorporated in the forecasting model and its complexity.

The power generated by a wind turbine is strongly related to the direction at which the wind strikes the blades of the turbine [6]. The direction of wind is often ignored while making forecasts of wind speed. Neglecting direction reduces the accuracy of the forecasting system and wind signal needs to be treated as a vector wind speed and direction as its components.

Complex-valued processing of wind signal provides an opportunity to exploit the non-circular nature of the complex wind signal [7]. Standard linear processing of the complex signals is based upon the inherent assumption of circularity of the signal and hence it does not provide optimal results for non-circular signals like wind. A widely-linear processing model, which takes into account the non-circularity of the complex wind signal is necessary for optimal results.

Another important characteristic of the wind signal is its spatial correlation at physically close sites. The inertia of the weather system makes wind changes at one place to correlate with nearby places. The spatial correlation of the wind signal at neighboring sites can be incorporated in the forecasting models to improve their performance [8,9].

1.2 Research Problem and Scope

This thesis work is focused on the development of a wind forecasting model which simultaneously takes into account the non-circularity and the spatial correlation of the complex wind signal. Furthermore, a resource-efficient widely-linear collaborative wind forecasting model for N neighboring nodes is developed. The spatial correlation among the neighboring nodes is taken into account by sharing the regressor data from all the nodes with their neighbors. The effect of direction on wind speed forecasts is taken into account using complex wind signal. The widely-linear collaborative wind forecasting model is further optimized for efficient resource utilization by reducing the local processing and communication between the collaborating nodes.

1.3 Contributions of the Thesis

There has been a lot of work in wind forecasting for various purposes including energy generation, weather prediction, air traffic management etc. Forecasting for the purposes of energy generation requires significant accuracy so that the uncertainty in the energy supply can be minimized. In this thesis work, we develop a resource-efficient collaborative model for the forecasting of wind power.

The main contributions of the thesis can be summarized as follows:

- Widely-linear modeling is used for N collaborating nodes to fully exploit the second-order statistics and spatial correlation of the complex wind signal.
- The widely-linear collaborative model is improved for efficient utilization of the system resources by reducing the local processing requirements and the amount of communication between the nodes.

1.4 Outline of the Thesis

The rest of the thesis is organized as follows. In Chapter 2, we study the nature of the complex wind signal and its characteristics using real world wind data. The analysis is used to develop the forecasting model in the following chapters. Chapter 3 provides an overview of existing wind forecasting techniques along with a detailed description of the widely-linear adaptive algorithms for wind forecasting. Towards the end of the chapter, we provide a literature survey of different collaborative techniques for wind forecasting. A review of these techniques will help us to understand the basic considerations and aspects of a collaborative wind forecasting model. In Chapter 4, we derive the widely-linear collaborative wind forecasting model. The widely-linear collaborative wind forecasting model is then optimized for efficient resource utilization using set-membership formulation. Chapter 5 contains the details of simulation setup, the data model for the complex wind signal and real world wind data used for the simulations. The results are discussed highlighting the performance gain of the resource-efficient widely-linear collaborative wind forecasting model.

Chapter 2

Spatio-Temporal Analysis of Wind Signal

This chapter is dedicated to study some characteristics of the wind signal which will be used to develop the wind forecasting model in the rest of the thesis. In this chapter, in Section 2.1, we present the current trends in the power industry, their challenges and the need for wind power forecasting, when wind energy is a part of the power grid. In Section 2.3, we study the relation between wind power with speed. Section 2.4 is devoted to the detailed analysis of the characteristics of wind signal, which are essential for the development of a good forecasting model. First we discuss the complex nature of the wind signal by studying the effect of wind direction on wind speed forecasts. Then the non-circular nature of the complex wind signal is analyzed. Finally the temporal and spatial correlation of the complex wind signal are analyzed with examples from real world data.

2.1 Smart Grids and the Need for Wind Power Forecasting

Conventional power grids have been designed in accordance with the power generation through fossil fuels. With increasing share of the renewable power sources into the conventional power grids, the integration of different kind of power sources and their scheduling has become a problem due to the intermittent nature of the renewable sources [4, 5].

Current trends in the power industry are shifting towards the modernization of power transmission and distribution system by employing the developments in the digital technologies. Thus, smart grids are becoming the solution to the problem of power management issues [10].

Smart grids are defined as an innovative solution to the transmission and distribution system with intelligent monitoring, control, communication and self healing technologies. They upgrade the existing power grid into real-time, two-way communication between power suppliers and their users. With smart grids in action, power suppliers provide customers with the pricing information based on the time

of the day. The customers can use that information to control their power consumption [11].

Since smart grids operate on the forecasted demand and supply of the power from all the sources, which are part of the grid, accurate source forecasts are essential for the smooth operation of the power system and to maintain load and supply in balance. This brings the need for better forecasting models for the renewable energy sources, which are variable in nature due to their dependence upon weather [10, 12].

With increased amount of wind energy into the power generation systems, its impact on the power supply is a critical issue. This impact can bring additional costs on part of the power supplier due to the variable nature of the wind energy. These additional costs can be significantly reduced if wind energy is scheduled using accurate and reliable forecasting methods [13, 14].

2.2 Wind Power versus Wind Speed

As we discussed, wind power forecasting is imperative for its integration into the power generation system. The power output of a wind turbine is directly linked to the speed at which the wind strikes its blades. The power output of the turbine $\mathbf{W}$ is in fact in direct proportion to the cube of the wind speed [1, 4] and the relation given by

$$\mathbf{W}(v) = \frac{1}{2}\rho A v^3 \tag{2.1}$$

can be derived based on the expression for the kinetic energy of the wind, where A is the airflow cross-section in m^2, ρ is the air density in $\mathrm{kg/m}^3$ and v is the air velocity in m/s [1, 6, 15].

The wind turbine cannot convert all of the kinetic energy of the air which strikes its blade, since in that case the air has to become still immediately after the impact which is not the case in the real world. The maximum possible energy extracted from the moving air is 59% of its kinetic energy, which happens when the wind speed after passing from the turbine is 1/3 of its strength as compared to that before the impact [1]. The numerical relation of the power output of a wind turbine and the wind speed at its blades is represented by the power curve [4].

Figure 2.1 shows the power curve for Vestas V82 wind turbine in which the power output from a wind turbine starts at a minimum wind speed 3.5 m/s called the cut-in speed v_{cutin} at which the turbine blades start to rotate. The output power follows a cubic relation until the rated speed (v_{rated}) 13 m/s at which the maximum power output of the turbine, which in this case is 1600 kW, is achieved. The cut-off speed v_{cutoff} is the wind speed at which the turbine blades stop rotating in order to protect it from damage and for V82 wind turbine it is 20 m/s.[1]

[1] www.vestas.com

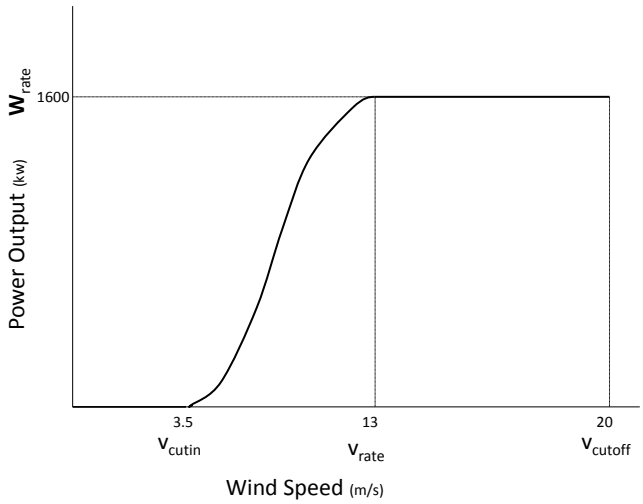


Figure 2.1: Power curve of a wind turbine generator.

2.3 Complex Nature of Wind Signal

Wind turbines are normally built to face the direction of the incoming wind while operating. At stronger winds the direction has less influence, since wind is more concentrated in one direction. For milder winds, the direction of the wind has more prominent impact since it is more spread in area and strikes the face of the turbine blades from broad range of directions [7].

Since wind speed and direction influence the wind turbine power simultaneously, complex-valued representation of the wind signal provides a convenient tool to include the effect of wind direction in wind speed calculations [6, 7]. A complex signal z can be constructed using the speed and direction data in the form $z = ve^{j\phi}$, where v is the wind speed data and ϕ is the wind direction data. Another complex valued representation of the wind signal is achieved using the real and imaginary components of the wind signal in North-East coordinate system. The complex wind signal z is formed by

$$z = z_E + jz_N \tag{2.2}$$

where z_E and z_N are the wind speed measurement data in East and North directions, respectively.

Figure 2.2 shows the distribution of wind speed over direction using real world wind data from the site Archer City (site#2001) for the month of August, 2007 obtained from the public database of the Alternative Energy Institute (AEI). [1] We

[1] `http://www.windenergy.org`

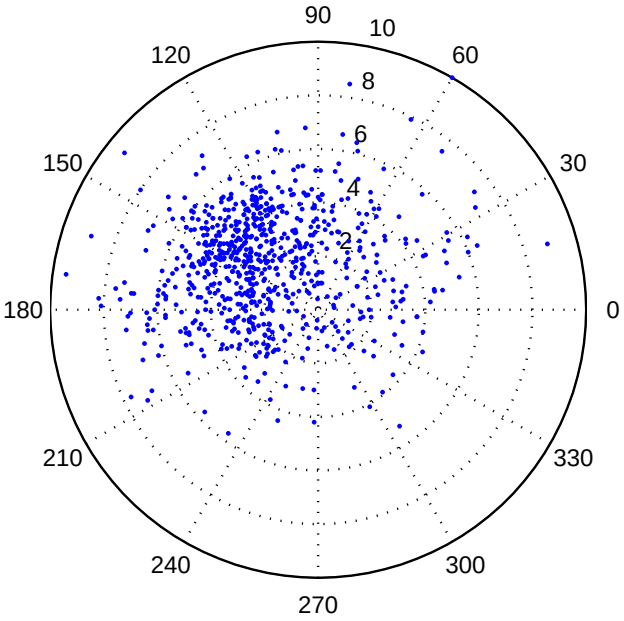


Figure 2.2: Wind speed distribution over the direction shown as a polar plot.

see that wind in certain directions is stronger as compared to others and it is spread in a range of angles. This shows the interdependence between these two variables of the wind signal [7].

2.4 Non-circularity of Complex Wind Signal

The second order statistical properties of complex signals are characterized by their covariance and pseudo-covariance. The covariance of the signal captures the information about the power of the signal whereas its pseudo-covariance characterizes the information about the power difference and the cross-correlation between real and imaginary parts of the complex signal [16]. For a general complex signal $z(k)$, covariance and pseudo-covariance functions are given by

$$c(k_1, k_2) = E[z(k_1)z^*(k_2)] \tag{2.3}$$
$$p(k_1, k_2) = E[z(k_1)z(k_2)] \tag{2.4}$$

where $z(k_1)$ and $z(k_2)$ are the values of signal at time instants k_1 and k_2, respectively.

Standard linear signal processing techniques are developed with an implicit assumption of pseudo-covariance being zero. However, this is not necessarily true and wind signal is a good example when pseudo-covariance is non-zero. Thus, algorithms based on the assumption of zero pseudo-covariance are not optimal for wind signals [17, 18]. Related to the pseudo-covariance, comes the concept of circularity.

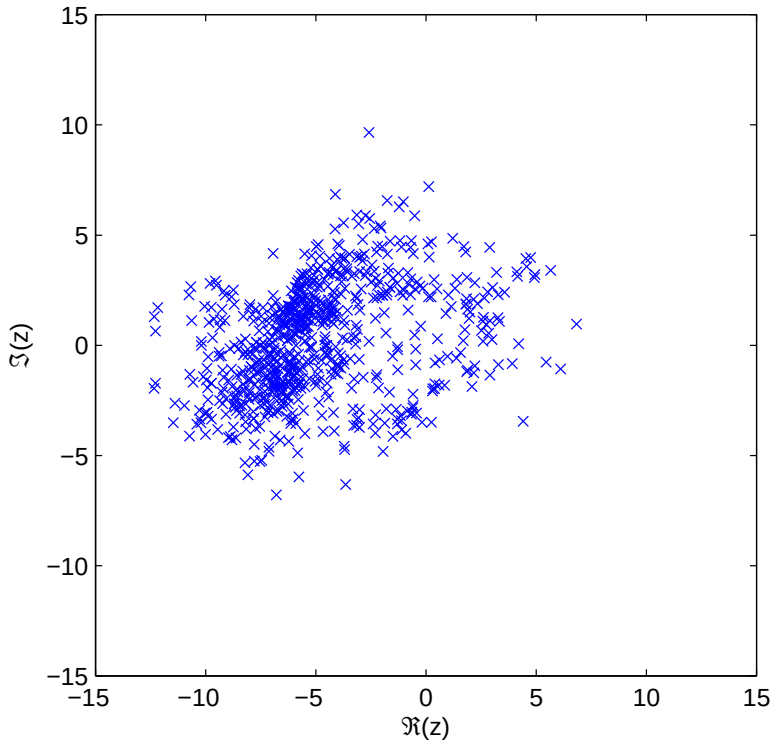


Figure 2.3: Real-imaginary plot for real world wind signal (z).

A complex signal z is circular if its probability density function does not change under any phase transformation. In other words, second order statistics remain the same under rotations of z. In general there is no reason for $p(k1, k2)$ to be zero and complex signals are non-circular which means that $p(k1, k2)$ needs to be taken into account [18, 19].

The non-circular nature of the wind can be verified by a real-imaginary plot [20]. The real-imaginary plot of a complex signal shows the distribution of real and imaginary components of the complex signal relative to each other. For a circular complex signal, the real imaginary plot is symmetric, whereas for non-circular complex signal it is non-symmetric. Non-symmetry of the real-imaginary plot of a complex signal is proportional to the degree of non-circularity and vice-versa.

In Fig. 2.3, real world complex wind data z for the site Archer city is used for the real-imaginary plot. The data set consists of 744 complex wind data points for the month of August 2007. From the plot the non-circular nature of the complex wind signal is evident with high degree of non-symmetry.

2.5 Temporal Analysis of Wind Signal

Weather is a system with memory, making wind a signal with strong temporal correlation. Thus, a measurement of wind signal at a particular time instant k is correlated with p past measurements. The temporal correlation of the complex wind signal can be described in terms of auto-correlation coefficient [21] of its speed and direction which represents the linear correlation among adjacent observations of the speed and direction data in time.

For a given wind data with K measurements $\mathbf{z} = [z(1) \dots z(K)]^T$ with $z(k) = v(k)e^{j\phi(k)}$, the real-valued auto-correlation coefficient for speed and direction represented by, $\hat{A}$, is estimated by the following

$$\hat{A}_x(\tau) = \frac{\frac{1}{K-\tau}\sum_{k=1}^{K-\tau}[x(k)x(k+\tau)] - \hat{m}_x^2}{\hat{\sigma}_x^2}, \tau = 0, \dots, K \tag{2.5}$$

where $x(k) = v(k)$ for wind speed and $x(k) = \phi(k)$ for wind direction, τ is the time lag, $\hat{m}_x$ is the estimated mean of the wind signal and $\hat{\sigma}_x^2$ is its estimated variance.

A plot of the auto-correlation coefficient versus time lag τ shows the decay of the correlation of speed or direction data samples. A slowly decaying auto-correlation coefficient shows long term temporal correlation. Similarly, a fast decaying plot shows short term temporal correlation [6, 21].

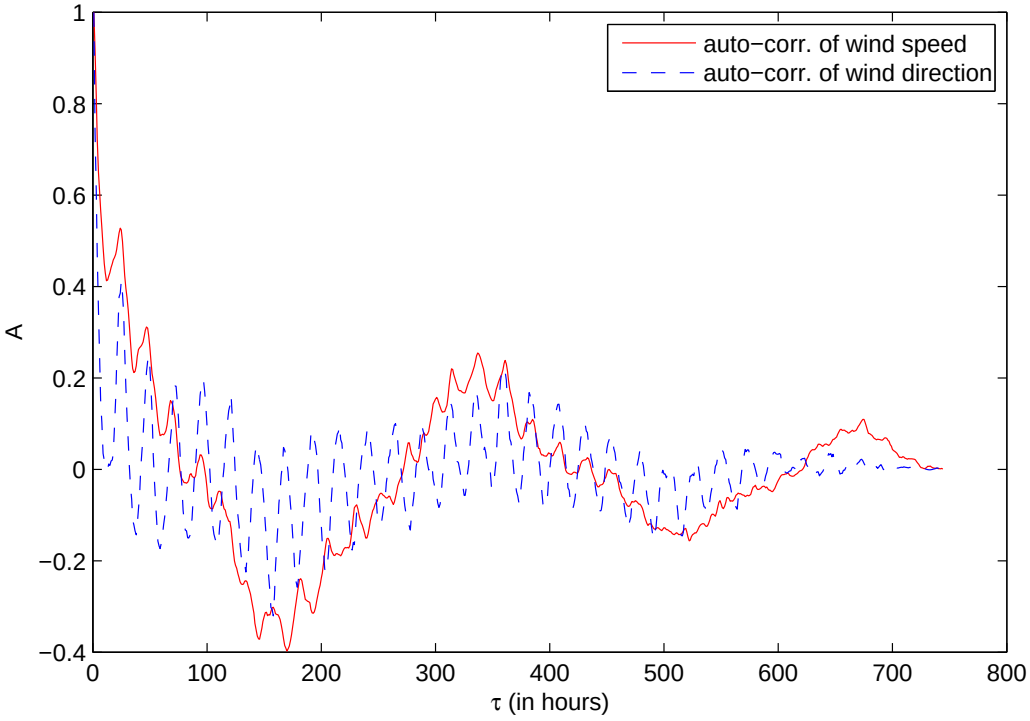
We use real world wind data from the site Archer City mentioned in the previous section to show the temporal correlation of the components of complex wind signal. Figure 2.4(a) shows the plot of temporal correlation of the wind speed and direction versus the time lag. A section of the Figure 2.4(a) is shown in Figure 2.4(b) highlighting the wind speed's temporal correlation for the time lag of up to eight hours which is more than 50% and wind direction has little less than 50% for the time lag of four hours.

2.6 Spatial Correlation of Wind Signal

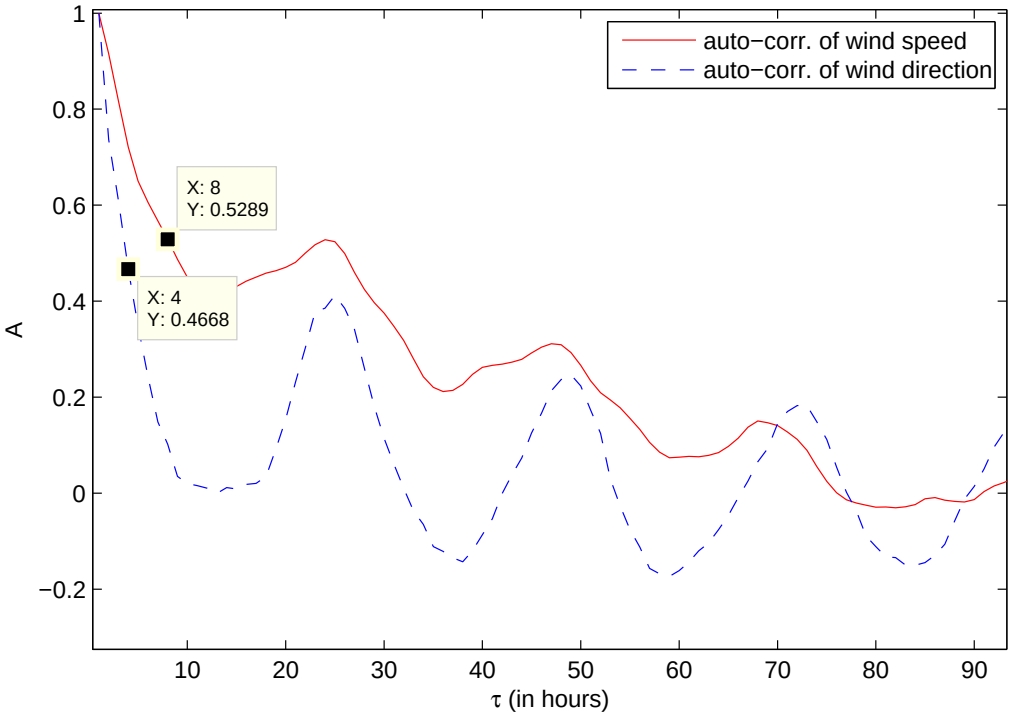
The weather system, due to its inertia, changes slowly over the earth surface. Wind at a particular site is highly correlated with the wind at its neighboring sites. The spatial correlation among neighbors can be described in terms of the cross-correlation coefficient [21]. The spatial correlation of wind signal can be exploited to improve the performance of a wind forecasting system.

For given time series wind data from two different sites with K measurements $\mathbf{z}_1 = [z_1(1) \dots z_1(N)]$, $\mathbf{z}_2 = [z_2(1) \dots z_2(N)]$ where $z_i(k) = v_i(k)e^{j\phi_i(k)}$ for $i = 1, 2$; the cross-correlation coefficient between wind speed and direction represented by $\hat{C}$, is given by

$$\hat{C}_{x_1x_2}(\tau) = \frac{\frac{1}{K-\tau}\sum_{k=1}^{K-\tau}[(x_1(k) - \hat{m}_{x_1})(x_2(k+\tau) - \hat{m}_{x_2})]}{\hat{\sigma}_{x_1}\hat{\sigma}_{x_2}}, \tau = 0, \dots, K \tag{2.6}$$



(a)



(b)

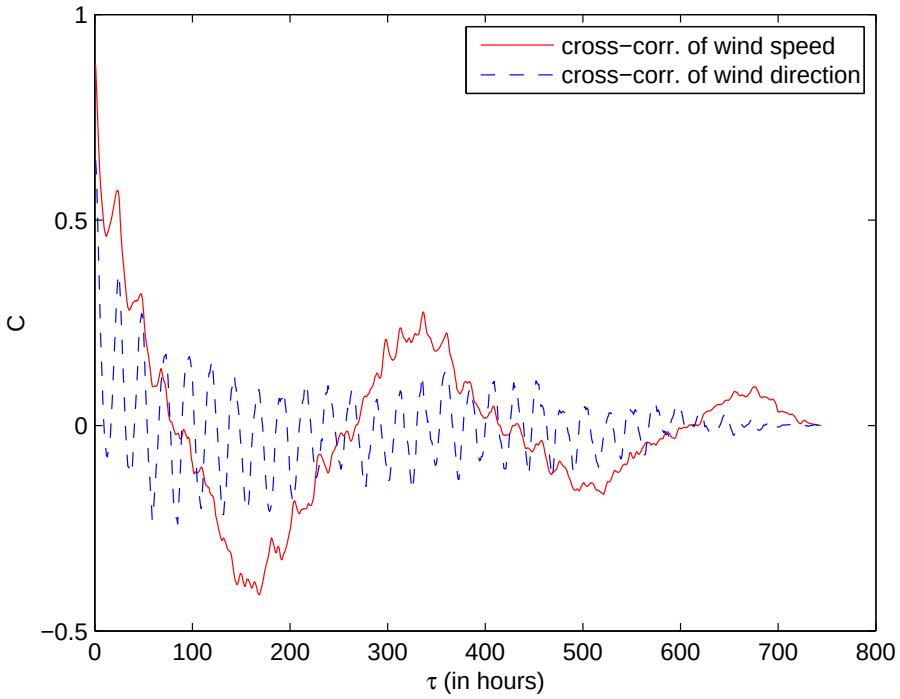
Figure 2.4: Temporal correlation of real-world wind signal for the selected data (a) and zoomed in for clarity (b).

where $x_i(k) = v_i(k)$ for wind speed and $x_i(k) = \phi_i(k)$ for wind direction for $i = 1, 2$ and τ is the time lag, $\hat{m}_{x_1}$ and $\hat{m}_{x_2}$ are the estimated means of the wind data at the two sites and $\hat{\sigma}^2_{x_1}$ and $\hat{\sigma}^2_{x_2}$ are their estimated variances.

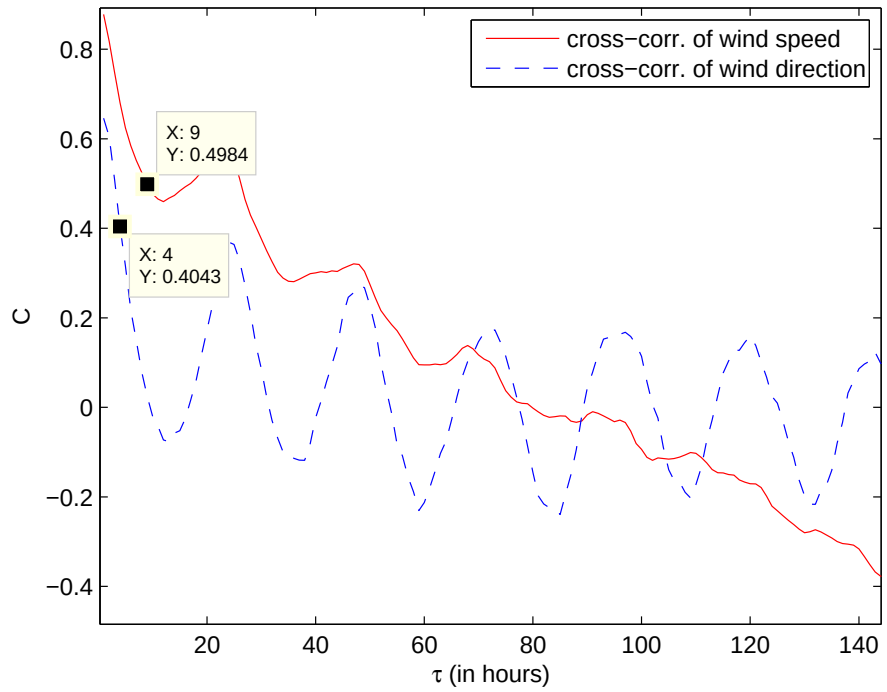
A plot of the cross-correlation coefficient versus the time lag gives a good idea of the spatial correlation of the wind between two sites. A slowly decaying cross-correlation coefficient shows long term spatial correlation. Similarly, a fast decaying cross-correlation coefficient shows short term correlation [5, 22–24].

We use real-world wind data from the two sites with Archer City as mentioned earlier and other site being Young County 1 - Site #1987 with wind speed and direction sensors at the height of 28m to show the spatial correlation of the components of the complex wind signal.

Figure 2.5(a) shows the cross-correlation coefficient between Archer City and Young County-1 versus time lag τ. Figure 2.5(b) shows a section of Figure 2.5(a). From the plots we see that, for the data in question, wind speed and direction from two different sites have high cross-correlation for time lag of up to nine hours in case of speed and four hours in case of direction. The spatial correlation at time lag zero is about 88% for speed and 64% for direction. This spatial correlation can be used in time series modelling of the complex wind signal to improve the forecasting performance.



(a)



(b)

Figure 2.5: Spatial correlation of wind signal for selected data (a) and zoomed in for clarity in (b).

Chapter 3

Wind Forecasting and Widely-Linear and Collaborative Processing of Wind Signal

This chapter starts with a general overview of wind forecasting. Different types of wind forecasting methods and their features are described. Typical considerations in a wind forecasting system such as forecasting horizons and performance measures are also summarized. In Section 3.2, we present the widely-linear filtering model for the complex wind signal. In Section 3.3, we discuss the performance gain achieved with widely-linear approach using real world wind data. Towards the end of the chapter, Section 3.4 is dedicated to the literature review of collaborative wind forecasting. Starting with a general introduction of collaborative wind forecasting, we will describe two collaborative schemes which are relevant to our work.

3.1 Overview of Wind Forecasting

The integration of wind energy into the power grid requires accurate forecasting of the wind power. While the need for wind power forecasting is evident, we focus on wind speed forecasting which can directly be used to estimate power output of the turbines. Complex wind signal is non-circular and also it has strong temporal and spatial correlation. A good forecasting model requires incorporation of all these characteristics of the wind signal to accurately predict the future values of the wind speed.

Different approaches have been used for wind forecasting purposes in the literature. The most basic and simplest technique to forecast the future value of the wind signal is 'persistence' [14, 25, 26] which is based on the simple assumption that wind speed at a future time is same as it is at the time at which forecast is being made. Persistence is used as a reference model to gauge the performance of more sophisticated forecasting models. More advanced models based on numerical weather prediction and statistical methods have been developed to improve the forecasting performance [14, 15, 27].

The suitability of a forecasting model is determined by the *forecasting horizon* which is the time ahead for which the forecast is made. Based on the forecasting horizon, wind forecasting is divided into short term (30min-6hrs), medium term (6hrs-24hrs) and long term (24hrs-7days) forecasts [27, 28].

Numerical weather prediction models rely on physical reasoning and employ topological information to forecast the future values of the weather system and these models were originally designed for weather predictions over large areas and for long time ahead typically ranging from several hours to months ahead. The resolution of numerical weather prediction based models is quite small and for short term forecasting the errors are high. Therefore, these methods are not best suited for short-term wind forecasting [14, 15, 25, 27].

For power scheduling purposes, short-term wind forecasting with a forecasting horizon of one hour are the most commonly used system parameters for forecasting models. Statistical models are best suited for short term wind forecasting and with the incorporation of wind characteristics and terrain information, these models outperform all other kind of methods [5]. Statistical methods are based on the training of model parameters using measurement data and the error signal, which is the difference between the predicted and actual wind signal. These methods are subdivided into two categories: 1) *Time-series based models*, and; 2) *Neural network based models*. The most popular models in time-series based methods are auto-regressive moving average (ARMA) models and their variants, e.g., auto-regressive integrated moving average (ARIMA), *seasonal-* and *fractional-*ARIMA and ARIMA with exogenous input (ARMAX or ARX). Some other time-series models are grey-predictors, linear predictions, exponential smoothing, etc [27, 29]. Neural network based methods are normally used where the available wind data is highly nonlinear and they are computationally intensive. In this thesis work we will focus only on time series based methods.

There is an inherent uncertainty in wind power predictions due to the intermittent nature of wind signal, meaning that there will always be an error in the wind forecasts. Thus, it is important to assess the performance of the forecasting method. The performance of a certain forecasting model is determined by the performance measures such as mean squared error (MSE) and root mean squared error (RMSE) [14, 26, 30, 31].

Performance measures give an account of the accuracy of forecasting model. Evaluation of the forecasting approach is based upon the comparison between the forecasted value of the wind and the actual value of the wind at the time for which prediction is made.

The prediction error defined as the difference [14, 26] between the measured value of the wind and the forecasted value is the most basic measure to look at the forecasting accuracy of the model. It is used in adaptive updating of the model parameters to minimize a function of the error signal like mean of the squared error.

Based on prediction error, there are numerous performance measures which are used in practice. Bias is the average value of the error over whole data set. Mean squared error (MSE) defined as mean of the squared error and root mean square error (RMSE) is square root of the MSE. Mean absolute error is sum of absolute

values of the errors over test data and standard deviation of error (SDE) is the square root of the variance of the error signal [14, 26, 30, 31].

3.2 Adaptive Algorithms for Short Term Wind Forecasting

Statistical models for wind forecasting purposes are based on the exploitation of the statistical properties of the wind signal. In the previous chapter, we studied some properties of the complex wind signal which, in this chapter, will be utilized for the development of statistical models for wind forecasting. The complex representation of wind signal is more natural and convenient for processing. Temporal correlation of the complex wind signal motivates to develop regression based filtering models where the past output of the system is weighted and combined to predict the future value. Widely-linear (WL) filtering models provide a framework to capture full second order statistics of the complex wind signal. Here we present the widely-linear filtering algorithms which fully utilize the second order properties of wind signal. We will start with widely-linear adaptive filtering model and provide the derivation of least squares (LS) and least mean square (LMS) type update algorithms for the filter parameters.

3.2.1 Widely-Linear Processing of Complex Wind Signal

Widely-linear modeling of complex signals is based upon the augmented statistics [18], where the input data $\mathbf{z}(k) = [z(k) \dots z(k-p+1)]^T$ with p samples is appended with its conjugate to form an augmented input vector $\mathbf{z}_a(k)$, which is given by

$$\begin{aligned} \mathbf{z}_a(k) &= [z(k) \dots z(k-p+1), z^*(k) \dots z^*(k-p+1)]^T \\ &= [\mathbf{z}^T(k), \mathbf{z}^H(k)]^T. \end{aligned} \tag{3.1}$$

The output of a widely-linear filter for the desired signal $d(k) = z(k+n)$, in n-step ahead prediction setting, is given by

$$\begin{aligned} y(k) &= \sum_{a=0}^{p-1} h_a z(k-a) + \sum_{a=0}^{p-1} g_a z^*(k-a) \\ &= \mathbf{h}^H(k)\mathbf{z}(k) + \mathbf{g}^H(k)\mathbf{z}^*(k) \\ &= \mathbf{w}^H(k)\mathbf{z}_a(k) \end{aligned} \tag{3.2}$$

where $\mathbf{h} = [h_0 \dots h_{p-1}]^T$, $\mathbf{g} = [g_0 \dots g_{p-1}]^T$ are the widely-linear filter coefficient vectors, and $\mathbf{w}(k) = [\mathbf{h}^T(k), \mathbf{g}^T(k)]^T$.

The objective is to find an estimate for the coefficients of the filter which tries to find those values of $\mathbf{h}(k)$ and $\mathbf{g}(k)$ that make $y(k)$ closest to the desired signal with an instantaneous error signal $e(k) = d(k) - y(k)$, where closeness is defined by an error criterion, also called objective function, which should be realistic enough for the given task and should be analytically tractable.

Two commonly used error criteria are provided below:

- *Deterministic least-squares error criterion*, given by

$$J(k) = \sum_{i=0}^{k} \lambda^{k-i} |e(i)|^2 = \sum_{i=0}^{k} \lambda^{k-i} |d(k) - y(k)|^2 \tag{3.3}$$

  where λ is the exponential weighting factor, or forgetting factor, and should be chosen in the range $0 \ll \lambda \leq 1$.

- *Instantaneous error criterion*, given by

$$J(k) = \frac{1}{2}|e(k)|^2 = \frac{1}{2}|d(k) - y(k)|^2. \tag{3.4}$$

The objective function $J(k)$ is a real-valued scalar function of the complex parameter vectors $\mathbf{h}(k)$ and $\mathbf{g}(k)$. The aim here is to find the parameter vectors $\mathbf{h}(k)$ and $\mathbf{g}(k)$, such that $J(k)$ is minimized. The direction of maximum rate of change of a real-valued function of a complex vector variable is determined by differentiating it with respect to the conjugate parameter vectors, i.e., $\mathbf{h}^*(k)$ and $\mathbf{g}^*(k)$ [18].

In the following derivations we will differentiate and minimize the objective functions with respect to corresponding filter parameters and find the optimum solution for the update of parameter vectors.

The deterministic error criterion for widely-linear least square type problem for complex input data can be given by

$$J(k) = \sum_{i=0}^{k} \lambda^{k-i} |d(i) - \left(\mathbf{h}^H(k)\mathbf{z}(i) + \mathbf{g}^H(k)\mathbf{z}^*(i)\right)|^2 \tag{3.5}$$

$$= \sum_{i=0}^{k} \lambda^{k-i} \left[d(i) - \left(\mathbf{h}^H(k)\mathbf{z}(i) + \mathbf{g}^H(k)\mathbf{z}^*(i)\right)\right] \left[d^*(i) - \left(\mathbf{h}^T(k)\mathbf{z}^*(i) + \mathbf{g}^T(k)\mathbf{z}(i)\right)\right]. \tag{3.6}$$

Now we will differentiate $J(k)$ with respect to $\mathbf{h}^*(k)$ and $\mathbf{g}^*(k)$ and equate the result to zero to find the optimum value of $\mathbf{h}(k)$ and $\mathbf{g}(k)$. Differentiating $J(k)$ with respect to $\mathbf{h}^*(k)$ and $\mathbf{g}^*(k)$ gives

$$\frac{\partial J(k)}{\partial \mathbf{h}^*(k)} = -\sum_{i=0}^{k} \lambda^{k-i} \mathbf{z}(i) \left[d^*(i) - \left(\mathbf{h}^T(k)\mathbf{z}^*(i) + \mathbf{g}^T(k)\mathbf{z}(k)\right)\right] \tag{3.7}$$

$$\frac{\partial J(k)}{\partial \mathbf{g}^*(k)} = -\sum_{i=0}^{k} \lambda^{k-i} \mathbf{z}^*(i) \left[d^*(i) - \left(\mathbf{h}^T(k)\mathbf{z}^*(i) + \mathbf{g}^T(k)\mathbf{z}(k)\right)\right] \tag{3.8}$$

Equating the results to zero

$$-\sum_{i=0}^{k}\lambda^{k-i}\mathbf{z}(i)\mathbf{z}^{H}(i)\mathbf{h}(k)-\sum_{i=0}^{k}\lambda^{k-i}\mathbf{z}(i)\mathbf{z}^{T}(i)\mathbf{g}(k)+\sum_{i=0}^{k}\lambda^{k-i}\mathbf{z}(i)d^{*}(i)]=\begin{bmatrix}0\\ \vdots\\ 0\end{bmatrix} \tag{3.9}$$

$$-\sum_{i=0}^{k}\lambda^{k-i}\mathbf{z}^{*}(i)\mathbf{z}^{H}(i)\mathbf{h}(k)-\sum_{i=0}^{k}\lambda^{k-i}\mathbf{z}^{*}(i)\mathbf{z}^{T}(i)\mathbf{g}(k)+\sum_{i=0}^{k}\lambda^{k-i}\mathbf{z}^{*}(i)d^{*}(i)]=\begin{bmatrix}0\\ \vdots\\ 0\end{bmatrix} \tag{3.10}$$

Equations (3.9) and (3.10) can be written in compact form

$$\mathbf{C}(k)\mathbf{h}(k)+\mathbf{P}(k)\mathbf{g}(k)=\mathbf{p}(k) \tag{3.11}$$

$$\mathbf{P}^{*}(k)\mathbf{h}(k)+\mathbf{C}^{*}(k)\mathbf{g}(k)=\mathbf{p}^{*}(k) \tag{3.12}$$

where

$$\mathbf{C}(k)=\sum_{i=0}^{k}\lambda^{k-i}\mathbf{z}(i)\mathbf{z}^{H}(i) \tag{3.13}$$

$$\mathbf{P}(k)=\sum_{i=0}^{k}\lambda^{k-i}\mathbf{z}(i)\mathbf{z}^{T}(i) \tag{3.14}$$

and

$$\mathbf{p}(k)=\sum_{i=0}^{k}\lambda^{k-i}\mathbf{z}(i)d^{*}(i). \tag{3.15}$$

Now solving (3.11) and (3.12) simultaneously for $\mathbf{h}(k)$ and $\mathbf{g}(k)$ gives

$$\begin{bmatrix}\mathbf{h}(k)\\ \mathbf{g}(k)\end{bmatrix}=\begin{bmatrix}\mathbf{C}(k) & \mathbf{P}(k)\\ \mathbf{P}^{*}(k) & \mathbf{C}^{*}(k)\end{bmatrix}^{-1}\begin{bmatrix}\mathbf{p}(k)\\ \mathbf{p}^{*}(k)\end{bmatrix} \tag{3.16}$$

$$\mathbf{w}(k)=\mathbf{C}_{a}^{-1}(k)\mathbf{p}_{a}(k), \tag{3.17}$$

where $\mathbf{C}_a(k)$ and $\mathbf{p}_a(k)$ are the augmented input covariance matrix and the cross-covariance vector between the augmented input vector and the desired signal.

The direct computation of $\mathbf{C}_a^{-1}(k)$ results in a computational complexity of $\mathcal{O}(p^3)$. The computation of the inverse matrix is avoided by the use of matrix inversion lemma [32]. In the following derivations, we will first express the inverse of $\mathbf{C}_a(k)$ in recursive form and then proceed to derive the parameter update for the filter coefficients based on the *a priori* error $e(k)=d(k)-[\mathbf{h}^{H}(k)\mathbf{z}(k)+\mathbf{g}^{H}(k)\mathbf{z}^{*}(k)]$.

From (3.13) and (3.14) the covariance matrix $\mathbf{C}_a(k)$ can be expressed in the recursive form as

$$\mathbf{C}_{a}(k)=\lambda\mathbf{C}_{a}(k-1)+\mathbf{z}_{a}(i)\mathbf{z}_{a}^{H}(i) \tag{3.18}$$

and by applying the matrix inversion lemma[1], the inverse matrix $\mathbf{S}_a(k)$ is given by

$$\begin{aligned}\mathbf{C}_a^{-1}(k) = \mathbf{S}_a(k) &= \left[\lambda \mathbf{C}_a(k-1) + \mathbf{z}_a(i)\mathbf{z}_a^H(i)\right]^{-1} \\ &= \frac{1}{\lambda}\left[\mathbf{S}_a(k-1) - \frac{\mathbf{S}_a(k-1)\mathbf{z}_a(k)\mathbf{z}_a^H(k)\mathbf{S}_a(k-1)}{\lambda + \mathbf{z}_a^H(k)\mathbf{S}_a(k-1)\mathbf{z}_a(k)}\right] \end{aligned} \tag{3.19}$$

Now (3.9) and (3.10) can be expressed as

$$\begin{aligned}\mathbf{C}(k)\mathbf{h}(k) + \mathbf{P}(k)\mathbf{g}(k) &= \lambda\left[\sum_{i=0}^{k-1}\lambda^{k-i-1}\mathbf{z}(i)d^*(i)\right] + \mathbf{z}(k)d^*(k) \\ &= \lambda\mathbf{p}(k-1) + \mathbf{z}(k)d^*(k) \end{aligned} \tag{3.20}$$

$$\begin{aligned}\mathbf{P}^*(k)\mathbf{h}(k) + \mathbf{C}^*(k)\mathbf{g}(k) &= \lambda\left[\sum_{i=0}^{k-1}\lambda^{k-i-1}\mathbf{z}^*(i)d^*(i)\right] + \mathbf{z}^*(k)d^*(k) \\ &= \lambda\mathbf{p}^*(k-1) + \mathbf{z}^*(k)d^*(k). \end{aligned} \tag{3.21}$$

Since

$$\mathbf{C}(k-1)\mathbf{h}(k-1) + \mathbf{P}(k-1)\mathbf{g}(k-1) = \mathbf{p}(k-1) \tag{3.22}$$

$$\mathbf{P}^*(k-1)\mathbf{h}(k-1) + \mathbf{C}^*(k-1)\mathbf{g}(k-1) = \mathbf{p}^*(k-1) \tag{3.23}$$

Equations (3.20) and (3.21) become

$$\mathbf{C}(k)\mathbf{h}(k) + \mathbf{P}(k)\mathbf{g}(k) = \lambda\left[\mathbf{C}(k-1)\mathbf{h}(k-1) + \mathbf{P}(k-1)\mathbf{g}(k-1)\right] + \mathbf{z}(k)d^*(k) \tag{3.24}$$

$$\mathbf{P}^*(k)\mathbf{h}(k) + \mathbf{C}^*(k)\mathbf{g}(k) = \lambda\left[\mathbf{P}^*(k-1)\mathbf{h}(k-1) + \mathbf{C}^*(k-1)\mathbf{g}(k-1)\right] + \mathbf{z}^*(k)d^*(k). \tag{3.25}$$

Equations (3.24) and (3.25) can be combined in the augmented form as

$$\mathbf{C}_a(k)\mathbf{w}(k) = \lambda\mathbf{C}_a(k-1)\mathbf{w}(k-1) + \mathbf{z}_a(k)d^*(k). \tag{3.26}$$

The *a priori* error is given by

$$e(k) = d(k) - \mathbf{w}^H(k-1)\mathbf{z}_a(k) \tag{3.27}$$

and

$$d(k) = e(k) + \mathbf{w}^H(k-1)\mathbf{z}_a(k). \tag{3.28}$$

We then put the value of $d(k)$ in (3.26) and simplify it for the parameter vector $\mathbf{w}(k)$

$$\begin{aligned}\mathbf{C}_a(k)\mathbf{w}(k) = &\left[\sum_{i=0}^{k}\lambda^{k-i}\mathbf{z}_a(i)\mathbf{z}_a^H(i) - \mathbf{z}_a(k)\mathbf{z}_a^H(k)\right]\mathbf{w}(k-1) \\ &+\mathbf{z}_a(k)\left(e(k) + \mathbf{w}^H(k-1)\mathbf{z}_a(k)\right)^* \end{aligned} \tag{3.29}$$

[1] $(\mathbf{A}+\mathbf{BCD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{C}^{-1} + \mathbf{DA}^{-1}\mathbf{B})^{-1}\mathbf{DA}^{-1}$

$$\mathbf{C}_a(k)\mathbf{w}(k) = \mathbf{C}_a(k)\mathbf{w}(k-1) + e^*(k)\mathbf{z}_a(k) \tag{3.30}$$

Using the inverse relation $\mathbf{S}_a(k) = \mathbf{C}_a^{-1}(k)$, we get

$$\mathbf{w}(k) = \mathbf{w}(k-1) + e^*(k)\mathbf{S}_a(k)\mathbf{z}_a(k). \tag{3.31}$$

Equation (3.31) is the least-square update for widely-linear filter parameters in recursive form.

Now we proceed to derive the filter update equation with stochastic error criterion. The stochastic error function given in (3.4) can be written as

$$J(k) = \left[d(k) - \left(\mathbf{h}^H(k)\mathbf{z}(k) + \mathbf{g}^H(k)\mathbf{z}^*(k)\right)\right]\left[d^*(k) - \left(\mathbf{h}^T(k)\mathbf{z}^*(k) + \mathbf{g}^T(k)\mathbf{z}(k)\right)\right]. \tag{3.32}$$

The generic gradient type update equation for the coefficient vector $\mathbf{h}(k)$ and $\mathbf{g}(k)$ is given by

$$\mathbf{h}(k+1) = \mathbf{h}(k) - \mu\frac{\partial J(k)}{\partial \mathbf{h}^*(k)} \tag{3.33}$$

$$\mathbf{g}(k+1) = \mathbf{g}(k) - \mu\frac{\partial J(k)}{\partial \mathbf{g}^*(k)} \tag{3.34}$$

where μ is a step size controlling the learning and the steady state error.

Differentiating $J(k)$ with respect to $\mathbf{h}^*(k)$ and $\mathbf{g}^*(k)$gives

$$\begin{aligned}\frac{\partial J(k)}{\partial \mathbf{h}^*(k)} &= -\mathbf{z}(k)[d^*(i) - \mathbf{h}^T(k)\mathbf{z}^*(k) - \mathbf{g}^T(k)\mathbf{z}(k)] \\ &= -e^*(k)\mathbf{z}(k) \qquad (3.35)\\ \frac{\partial J(k)}{\partial \mathbf{g}^*(k)} &= -\mathbf{z}(k)[d^*(i) - \mathbf{h}^T(k)\mathbf{z}^*(k) - \mathbf{g}^T(k)\mathbf{z}(k)] \\ &= -e^*(k)\mathbf{z}^*(k). \qquad (3.36)\end{aligned}$$

From (3.33), (3.34), (3.35) and (3.36) the least mean square type update for widely-linear filter coefficients is given by

$$\mathbf{h}(k+1) = \mathbf{h}(k) + \mu e^*(k)\mathbf{z}(k) \tag{3.37}$$

$$\mathbf{g}(k+1) = \mathbf{g}(k) + \mu e^*(k)\mathbf{z}^*(k). \tag{3.38}$$

The update equations for $\mathbf{h}(k)$ and $\mathbf{g}(k)$ can be combined in augmented form

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu e^*(k)\mathbf{z}_a(k) \tag{3.39}$$

and (3.39) is referred as the widely-linear LMS type update equation.

The widely-linear least-squares and least mean square type filter reduce to their linear counterparts by setting $\mathbf{g}(k) = 0$.

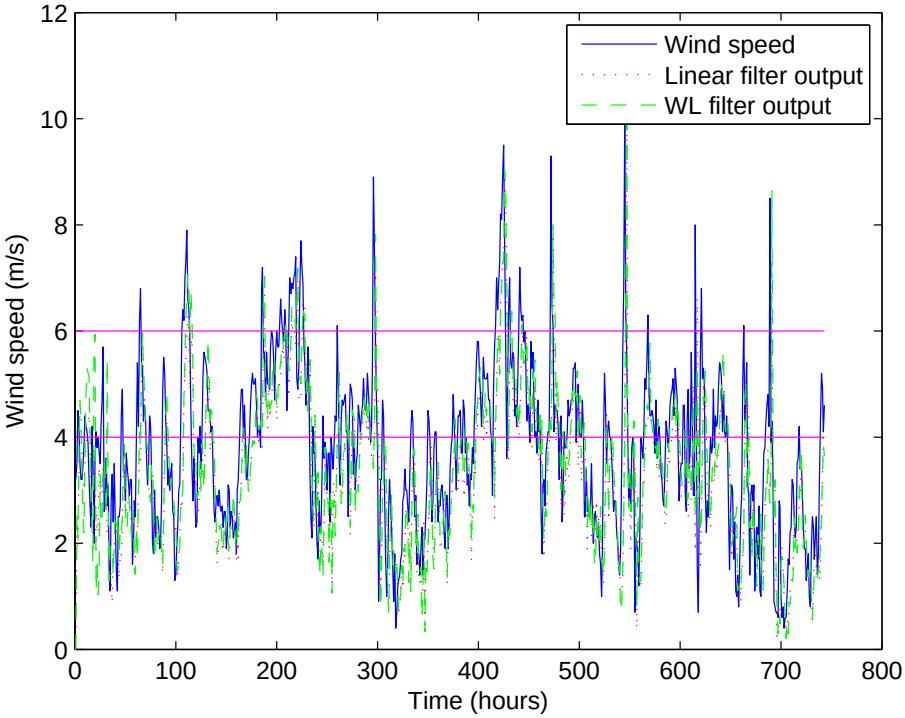


Figure 3.1: RLS and WL-RLS processing of real world wind data.

3.3 Performance Gain using Widely-Linear Approach

Figure 3.1 shows the speed plot of a real word wind data from the site Archer City. The magnitude of wind signal is divided into three regions, i.e., low, medium and high. Wind speeds below 4m/s is called low wind, between 4 m/s and 6 m/s is called medium wind and wind greater than 6 m/s is called high wind. Output plots of one-step ahead prediction with linear and widely linear filters are also shown on the same figure.

Table 3.1: RMSE for linear and widely-linear modeling.

Site/RMSE	linear output	widely-linear output
Archer City	1.50	1.46
Young County1	1.38	1.35
Young County2	1.37	1.35

Table 3.1 shows the RMSE values for the three sites in our study for the overall wind test data set. The RMSE values in case of WL-linear processing are less than those of linear processing, confirming the non-circular nature of the wind signal and the performance gain achieved through WL-linear processing as well.

3.4 Collaborative Processing of Wind Signal

While the need for wind speed forecasting has been emphasized in the previous sections, current forecasting models are mostly developed for the data from a single site without properly taking into account the spatial correlation of the wind signal. As discussed in Chapter 2, wind signals at neighboring sites are correlated. Thus, forecasts for a single site can be improved by incorporating the spatial correlation of neighboring sites into the forecasting model.

Wind data from *upwind* [33] neighboring turbines, where the wind strikes first, in a single wind park can also be utilized to make better forecasts for the turbines at the *downwind* location which receive the same wind that comes from upwind turbine after some time lapse. Since wind does not change much in a relatively small area of a single wind park, most of the work has been done to exploit the wind data from different sites with somewhat larger distances and with good spatial correlation.

This section is dedicated to the literature review of collaborative wind forecasting techniques. In Section 3.4.1, we explain the multi-variable auto-regressive model for collaborative wind forecasting. Section 3.4.2 is dedicated to the decentralized wind flow model for a single wind park. Finally in Section 3.4.3, we summarize some other collaborative wind forecasting techniques.

3.4.1 Multivariate Auto-Regressive Model

Multivariate auto-regressive model has been developed for wind speed forecasting for UK Met-Office weather stations data. The method presented in [34] is summarized here to develop the understanding of a basic spatial correlation based wind forecasting model.

The model uses time-stamped wind speed data from different stations as input and produces a time-series model which utilizes both the temporal and spatial correlations of the wind speed. For the development of this model, data was obtained from the Met-Office stations which offered consistent historical measurements and with the longest records to include all the variations in the wind speeds over the years.

Wind speed at single site n in one-step ahead prediction setting can be represented with a time series, in the form of a *pth* order auto-regressive (AR) model

$$\begin{aligned} y_n(k) &= \sum_{a=1}^{p} h_{n,a} z_n(k-a) + v(k) \\ &= \mathbf{h}_n^T(k)\mathbf{z}_n(k) + \mathbf{v}(k) \end{aligned} \tag{3.40}$$

where $v(k)$ is a random process (zero mean and unit variance), $\mathbf{h}_n(k) = [h_1, \ldots, h_p]^T$ are the auto-regressive coefficients and $\mathbf{z}_n(k) = [z_{k-1}, \ldots, z_{k-p}]^T$ is the regressor data.

For N sites, each with regressor data of length p, the model can be represented

as

$$\mathbf{y}(k) = \begin{bmatrix} y_1(k) \\ \vdots \\ y_N(k) \end{bmatrix} = \sum_{a=1}^{p} \mathbf{H}_a(k)\mathbf{z}(k-a) + \mathbf{v}(k)$$
$$= \mathbf{H}_s(k)\mathbf{z}_s(k) + \mathbf{v}(k) \tag{3.41}$$

where $\mathbf{y}(k) = [y_1(k), \ldots, y_N(k)]^T$ is a $N \times 1$ output vector for N sites at time instant k, $\mathbf{H}_1, \ldots, \mathbf{H}_p$ are the $N \times N$ auto-regressive coefficient matrices which correlate the values of N measurements in $\mathbf{z}(k-a)$ to their own past values and also to the other site's past data values, and $\mathbf{v}(k)$ is an $N \times 1$ vector of white noise and $\mathbf{h}_s(k) = [\mathbf{H}_1(k) \ldots \mathbf{H}_p(k)]$ is a matrix stacked with the matrices $\mathbf{H}_a(k)$ and $\mathbf{z}_s(k) = [\mathbf{z}^T(k-1) \ldots \mathbf{z}^T(k-p)]^T$ is a vector stacked with $\mathbf{z}(k)$ and $\mathbf{v}(k)$ is a vector of white noise.

3.4.2 Decentralized Wind Flow Model and Prediction

The decentralized wind flow model given in [33] is based on the fact that each *downwind* turbine in a wind park is affected by its *upwind* neighboring turbines. Here we summarize the main steps of the decentralized wind flow model.

Prediction Model

The general form of the model is based on discrete time auto-regressive with external input model with upwind turbine's regressor data $\mathbf{z}(k)$ and the output of downwind turbine $y(k)$ in one-step ahead prediction setting is given by

$$y(k) = -\sum_{p=1}^{p_u} h_{up} y(k-p) + \sum_{p=1}^{p_d} h_{dp} z_d(k-p)$$
$$= \mathbf{h}_u^T(k)\mathbf{y}(k) + \mathbf{h}_d^T(k)\mathbf{z}_d(k) \tag{3.42}$$

where p_u is the regressor length of *upwind* turbine with regressor data $\mathbf{z}_u(k) = [z_u(k-1) \ldots z_u(k-p_u)]^T$ and p_d is the regressor length of the *downwind* turbine obtained from previous outputs, $\mathbf{y}(k) = [-y(k-1) \ldots -y(k-p_d)]^T$ and $\mathbf{h}_u(k) = [h_{u1} \ldots h_{up_u}]^T$ and $\mathbf{h}_d(k) = [h_{d1} \ldots h_{dp_d}]^T$ are the parameter vectors for upwind and downwind turbines.

Online Parameter Estimation

In many real-world applications, online estimation of the parameters is required as the system output changes constantly and the measurements are available only when the system is in operation.

Let $\mathbf{h}(k) = [h_{u1} \ldots h_{up_u}, h_{d1} \ldots h_{dp_d}]^T$ be the overall parameter vector of size $p = p_u + p_d$ then the Kalman filtering formulation for the updates of the true parameter vector $\mathbf{h}_\circ(k)$ is given by

$$\mathbf{h}_\circ(k) = \mathbf{h}_\circ(k-1) + \mathbf{v}(k), \tag{3.43}$$

where $\mathbf{h}_\circ(k)$ is the true parameter vector of the same size as $\mathbf{h}(k)$ and $\mathbf{v}(k)$ white noise vector of size p with $p \times p$ covariance matrix $\mathbf{R}_1$.

The general recursive form of the parameter update is given by

$$\mathbf{h}(k) = \mathbf{h}(k-1) + \mathbf{k}(k)(d(k) - y(k)), \tag{3.44}$$

$$y(k) = \mathbf{h}^T(k-1)\mathbf{z}(k), \tag{3.45}$$

where $d(k)$ is the actual output of the system at time k, $y(k)$ is the predicted output, $\mathbf{k}(k)$ is the Kalman gain vector of size $p \times 1$ and $\mathbf{z}(k) = [-y(k-1), \ldots, -y(k-p_u), z_d(k-1), \ldots, z_d(k-p_d)]^T$.

The gain has the following general form

$$\mathbf{k}(k) = \mathbf{Q}(k)\mathbf{z}(k), \tag{3.46}$$

where the $N \times N$ matrix $\mathbf{Q}(k)$ is defined as

$$\mathbf{Q}(t) = \frac{\mathbf{P}(k-1)}{\sigma_e^2 + \mathbf{z}^T(k)\mathbf{P}(k-1)\mathbf{z}(k)}, \tag{3.47}$$

$$\mathbf{P}(k) = \mathbf{P}(k-1) + \mathbf{R}_1 - \frac{\mathbf{P}(k-1)\mathbf{z}(k)\mathbf{z}^T(k)\mathbf{P}(k-1)}{\sigma_e^2 + \mathbf{z}^T(k)\mathbf{P}(k-1)\mathbf{z}(k)}, \tag{3.48}$$

and σ_e^2 is the error covariance and it is a scalar quantity and $\mathbf{P}$ ia a $p \times p$ matrix.

Fusion

When there are several upwind neighboring turbines which are laterally distributed against the wind direction then we have different predictions at downwind turbine resulting from the each upwind turbine. It is then beneficial to fuse all these predictions to generate one global prediction for the downwind turbine. The fusion can be based on the weight given to prediction from each upwind turbine according to the error variance associated with it. The fusion can be made adaptive by using a recursive algorithm for the weight's update.

A simple example to fuse the predictions from N upwind turbines is given by

$$y_g(k) = \sigma_T^2(k) \sum_{i=1}^{N} \sigma_i^{2-1}(k) y_i(k), \tag{3.49}$$

$$\sigma_T^2(k) = \left[\sum_{i=1}^{N} \sigma^{2^{-1}}_i(k) \right]^{-1}, \tag{3.50}$$

where $y_i(k)$ is the ith predicted wind speed and $\sigma_i^2(k)$ is the corresponding error covariance. All quantities in this expression are scalars.

Effect of Direction

The direction of the wind affects performance achieved by collaboration. There are two general cases to consider here. The first case is when the direction of the wind is in line with the two turbines, i.e., wind direction is in parallel to the wind turbine row and the second case is when the wind direction hits the wind turbine row at some constant angle θ.

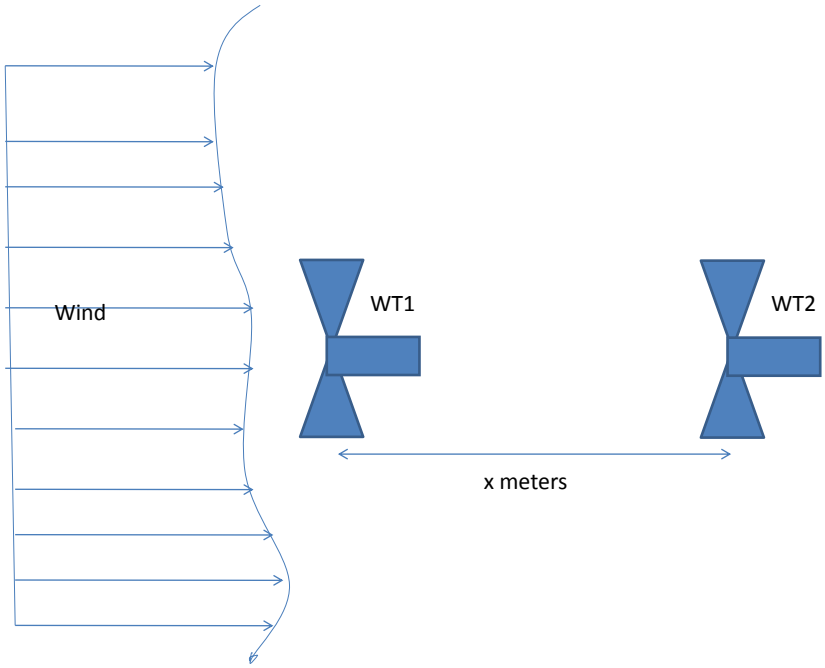
There is very small or negligible benefit achieved by collaboration in the first case while in the second case the forecasting performance is improved considerably. Figure 3.2(a) and Figure 3.2(b) show the two scenarios discussed above. In Figure 3.2(a) the direction of wind is parallel to the line connecting the two wind turbines. In this case same wind signal will be received at the two turbines and hence there is no gain achieved through collaboration. Whereas, in Figure 3.2(b) wind from the upwind turbine hits the downwind turbine at an angle θ. In this case, the measurements at the upwind turbine will be useful for the predictions at downwind turbine. The simulation results for these cases given in [33] confirm the observations made.

3.4.3 Other Collaborative Schemes

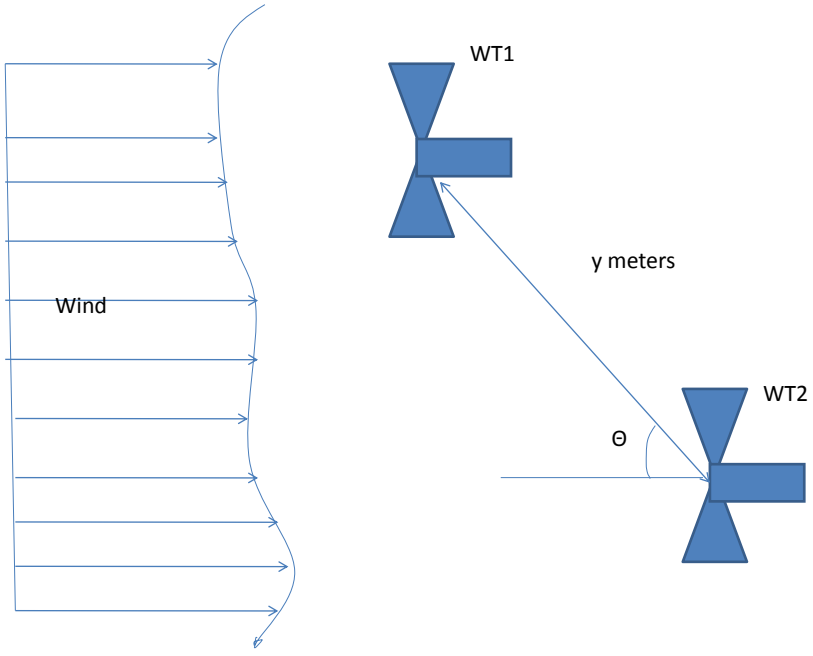
One of the simplest techniques to include the effect of wind direction in spatial correlation studies is to use direction specific forecasting models. Regime-switching space-time diurnal model [35] is an example of such case, where the wind speed forecasting model is switched based on the direction of wind flow. Two different forecasting models are used based on the direction of the wind flow which can be from west to east (westerly regime) or east to west (easterly regime). The main limitation of such kind of model is that only two wind flow directions are assumed, while in practice wind can flow from a wide area with different directions.

The trigonometric direction diurnal model [5] was developed to address this problem by including the wind direction as a covariate in the forecasting model. The inclusion of direction as covariate in wind speed forecasting provides more flexibility and improves the accuracy of the model.

Another model with wind speed and direction data as a bi-variate vector in Cartesian coordinates was developed to improve the spatial correlation-based wind forecasting. The bi-variate Skew-T model [5] uses wind data in Cartesian co-ordinate system from the neighboring sites.



(a)



(b)

Figure 3.2: Effect of direction on collaborative wind forecasting with turbines in a straight line (a) and at an angle θ with each other.

Chapter 4

Resource-Efficient Widely-Linear Collaborative Processing of Wind Signal

The available spatial correlation based wind forecasting techniques, reviewed in Chapter 3 do not take into account the non-circular nature of the complex wind signal. Moreover existing models do not exploit all characteristics of the wind signal such as the effect of direction, non-circularity, spatial correlation and temporal correlation in a single formulation. This chapter presents a widely-linear collaborative wind forecasting model, which takes variable-length regressors from all the sites in the neighborhood of a reference site and appends them with its regressor to make m-step ahead prediction. The coefficients of the model are updated recursively in such a way that the sum of the square of error is minimized. The model is further optimized for efficient resource utilization using a set-membership filtering approach. As a results, we can significantly reduce local processing and communication requirement between neighboring nodes.

In Section 4.1, we will derive the widely-linear collaborative wind forecasting model and the update equations for the filter coefficients. In Section 4.2, we will present the resource-efficient formulation of the widely-linear collaborative model and discuss the saving in the communication and processing cycles using this approach.

4.1 Collaboration with Widely-Linear Processing

Current models of collaboration using spatial correlation from different locations are based upon the inherent assumption of the circularity of the wind data. In the following section we derive a widely-linear model for the processing of complex wind signal using augmented statistics [7, 18].

4.1.1 Widely-Linear System Model

We consider a fully connected network of N nodes with measurement data at nth node is given by $\mathbf{z}_n(k) = [z_n(k) \dots z_n(k-p_n+1)]^T, n = 1 \dots N$, with p_n being the regressor length for the node. Each node receives the measurement data from all other nodes in the network and makes its m-steps ahead prediction by exploiting its own auto-correlation and cross-correlation with the other nodes' data using augmented data vectors.

For the desired signal $d(k) = z_n(k+m)$, the output from the node n is given by

$$\begin{aligned} y(k) &= \sum_{n=1}^{N} \sum_{p=0}^{p_n-1} h_{n,p} z_n(k-p) + \sum_{n=1}^{N} \sum_{p=0}^{p_n-1} g_{n,p} z_n^*(k-p) \\ &= \mathbf{h}_1^H(k)\mathbf{z}_1(k) + \cdots + \mathbf{h}_N^H(k)\mathbf{z}_N(k) + \mathbf{g}_1^H(k)\mathbf{z}_1^*(k) + \cdots + \mathbf{g}_N^H(k)\mathbf{z}_N^*(k) \\ &= \mathbf{h}^H(k)\mathbf{z}(k) + \mathbf{g}^H(k)\mathbf{z}^*(k) \\ &= \mathbf{w}^H(k)\mathbf{z}_a(k) \end{aligned} \tag{4.1}$$

where $\mathbf{w}(k) = [\mathbf{h}^T(k), \mathbf{g}^T(k)]^T$ is the combined augmented coefficient vector and $\mathbf{z}_a(k) = [\mathbf{z}^T(k), \mathbf{z}^H(k)]^T$ is the combined augmented data vector.

4.1.2 Parameter Estimation

The parameters of the widely-linear filter are estimated by minimizing an objective function of the instantaneous error signal $e(k) = d(k) - y(k)$. For this purpose we use the least squares objective function, introduced in Section 3.2.1, given by

$$J(k) = \sum_{i=0}^{k} \lambda^{k-i} |d(k) - \mathbf{w}^H(k)\mathbf{z}_a(k)|^2. \tag{4.2}$$

Differentiating $J(k)$ with respect to $\mathbf{w}^*(k)$ and solving for $\mathbf{w}(k)$, we get

$$\begin{aligned} \mathbf{w}(k) &= \mathbf{C}^{-1}(k)\mathbf{p}(k) \\ &= \begin{bmatrix} \mathbf{C}_{11}(k) & \cdots & \mathbf{C}_{1N}(k) & \mathbf{P}_{11}(k) & \cdots & \mathbf{P}_{1N}(k) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{C}_{N1}(k) & \cdots & \mathbf{C}_{NN}(k) & \mathbf{P}_{1N}(k) & \cdots & \mathbf{P}_{1N}(k) \\ \mathbf{P}_{11}^*(k) & \cdots & \mathbf{P}_{1N}^*(k) & \mathbf{C}_{11}^*(k) & \cdots & \mathbf{C}_{1N}^*(k) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{P}_{N1}^*(k) & \cdots & \mathbf{P}_{N1}^*(k) & \mathbf{C}_{1N}^*(k) & \cdots & \mathbf{C}_{NN}^*(k) \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{p}_{11}(k) \\ \vdots \\ \mathbf{p}_{N1}(k) \\ \mathbf{p}_{11}^*(k) \\ \vdots \\ \mathbf{p}_{N1}^*(k) \end{bmatrix} \end{aligned} \tag{4.3}$$

where

$$\mathbf{C}_{mn}(k) = \sum_{i=0}^{k} \lambda^{k-i} \mathbf{z}_m(k) \mathbf{z}_n^H(k)$$

$$\mathbf{P}_{mn}(k) = \sum_{i=0}^{k} \lambda^{k-i} \mathbf{z}_m(k) \mathbf{z}_n^T(k)$$

$$\mathbf{p}_{mn}(k) = \sum_{i=0}^{k} \lambda^{k-i} \mathbf{z}_m(k) d_n^*(k), \quad m = 1 \ldots N, n = 1 \ldots N$$

are the input covariance matrix and pseudo-covariance matrix and cross-covariance vector between input data and desired signal, respectively.

The parameter update equation for the widely-linear collaborative filter is similar in formulation to (3.17) and hence, it can be written in the form

$$\mathbf{w}(k) = \mathbf{w}(k-1) + e^*(k) \mathbf{S}(k) \mathbf{z}_a(k) \tag{4.4}$$

where

$$\mathbf{S}(k) = \frac{1}{\lambda} \left[\mathbf{S}(k-1) - \frac{\mathbf{S}(k-1) \mathbf{z}_a(k) \mathbf{z}_a^H(k) \mathbf{S}(k-1)}{\lambda + \mathbf{z}_a^H(k) \mathbf{S}(k-1) \mathbf{z}_a(k)} \right]$$

. Equation(4.4) is the widely-linear collaborative least-square type update for filter parameter in recursive form. It is evident from the expressions of $\mathbf{C}(k)$, $\mathbf{P}(k)$, $\mathbf{C}^*(k)$, $\mathbf{P}^*(k)$, $\mathbf{p}(k)$ and $\mathbf{p}^*(k)$ that the widely-linear collaborative filtering model takes into account the auto-covariance, pseudo-covariance and cross-covariance of the input data and its conjugate with the desired signal as well as the cross-covariances with the data from all other nodes in the network, thus exploiting the spatial correlation as well as non-circularity of the complex wind signal in a single formulation.

By setting $\mathbf{g}(k) = 0$ and without augmenting the data vector, we can achieve the linear version of the collaborative RLS filtering model. The algorithms for the linear and widely-linear collaborative RLS filtering models are summarized in the Tables 4.1 and 4.2.

4.2 Resource-Efficient Processing

The update for the collaborative RLS filter basically consists of two steps: 1) A communication step where nodes (or sites) share regressor vectors with each other; and 2) An update step where each node computes its parameter update using its own and other nodes regressor vectors. In other words, this requires *constant transmission and processing for each new data measurement* and hence does not render efficient utilization of the system resources, since the new data is not always useful for the parameter update. In particular as the number of nodes increases, local complexity as well as the amount of communication will increase substantially. We will in this section look at a strategy that reduces communication requirements (in terms

Table 4.1: Linear collaborative RLS processing.

Linear collaborative RLS algorithm for node n
Initialization:
$\mathbf{S}(0) = \delta\mathbf{I}$
$\mathbf{h}(0) = 0$
for $k = 1{:}K$
$\{$
$\quad \mathbf{z}(k) = [\mathbf{z}_1^T(k), \ldots, \mathbf{z}_N^T(k)]^T$
$\quad e_n(k) = d_n(k) - \mathbf{h}^H(k)\mathbf{z}(k)$
$\quad G(k) = \mathbf{z}^H(k)\mathbf{S}(k-1)\mathbf{z}(k)$
$\quad \mathbf{S}(k) = \frac{1}{\lambda}\left[\mathbf{S}(k-1) - \frac{\mathbf{S}(k-1)\mathbf{z}(k)\mathbf{z}^H(k)\mathbf{S}(k-1)}{\lambda+G(k)}\right]$
$\quad \mathbf{h}(k) = \mathbf{h}(k-1) + e_n^*(k)\mathbf{S}(k)\mathbf{z}(k)$
$\}$
If necessary, compute:
$y_n(k) = \mathbf{h}^H(k)\mathbf{z}(k)$
$e_n^{'}(k) = d_n(k) - y_n(k)$

Table 4.2: Widely-linear collaborative RLS processing.

Widely-linear collaborative RLS algorithm for node n
Initialization
$\mathbf{S}(0) = \delta\mathbf{I}$
$\mathbf{w}(0) = 0$
for $k = 1{:}K$
$\{$
$\quad \mathbf{z}(k) = [\mathbf{z}_1^T(k), \ldots, \mathbf{z}_N^T(k)]^T$
$\quad \mathbf{z}_a(k) = [\mathbf{z}^T(k)\mathbf{z}^H(k)]^T$
$\quad e_n(k) = d_n(k) - \mathbf{w}^H(k)\mathbf{z}_a(k)$
$\quad G(k) = \mathbf{z}_a^H(k)\mathbf{S}(k-1)\mathbf{z}_a(k)$
$\quad \mathbf{S}(k) = \frac{1}{\lambda}\left[\mathbf{S}(k-1) - \frac{\mathbf{S}(k-1)\mathbf{z}_a(k)\mathbf{z}_a^H(k)\mathbf{S}(k-1)}{\lambda+G(k)}\right]$
$\quad \mathbf{w}(k) = \mathbf{w}(k-1) + e_n^*(k)\mathbf{S}(k)\mathbf{z}_a(k)$
$\}$
If necessary, compute:
$y_n(k) = \mathbf{w}^H(k)\mathbf{z}_a(k)$
$e_n^{'}(k) = d_n(k) - y_n(k)$

of regressor sharing) as well as local processing complexity by selectively updating the parameter estimates. The basic principles underlying the proposed strategy rest on the principles of set-membership filtering (SMF) which feature data-dependent updates of parameter estimates integrated with those of widely-linear adaptive filters.

In SMF, the parameter update is based on a predefined threshold for the magnitude of the error signal for each pair of the input data and desired signal [36–38].

Using this approach, local update and transmission is made only when the error signal is greater than the pre-defined threshold, thus keeping the error below the threshold. There are many set-membership filtering based approaches for the parameters update but optimal bounding ellipsoids (OBE) are one of the most popular. In the following we will present OBE type set-membership formulation of the widely-linear collaborative filtering algorithm.

For a combined augmented input data vector $\mathbf{z}_a(k)$ from a network of N nodes and local desired signal $d(k) = z_n(k+m)$ in m-steps ahead prediction setting at node n and at time instant k with filter parameter vector $\mathbf{w}(k)$, in SMF by setting an upper bound γ on the error, we seek parameter vector $\mathbf{w}$ which keeps the magnitude of the output error below a predefined threshold γ for all input data pairs, i.e.,

$$|e(k)| \leq \gamma, \forall k \tag{4.5}$$

At time instant k, *constraint set* given by

$$H(k) = \{\mathbf{w} : |d(k) - \mathbf{w}^H(k)\mathbf{z}_a(k)|^2 \leq \gamma^2\} \tag{4.6}$$

which defines the set of parameter vectors that for current input data pair $\{d(k), \mathbf{z}_a(k)\}$ obey the bound criterion in (4.5) [36–38].

Let us consider a sequence of data pairs measured at time instants $k = 1$ to $k = K$. The unknown parameter vector remains constant, it must lie in the following set

$$\psi(k) = \cap_{j=1}^{k} H(j) \tag{4.7}$$

referred to as the *exact membership set*. We see that any parameter vector belonging to the exact membership set is a feasible solution for the requirement set in the formulation. This is in contrast with the conventional approach in (4.3) which allows only allows a single solution. The exact membership set defined in (4.7) is a convex polytope in the parameter space and can be tracked more conveniently by using bounding ellipsoids.

Assume that at time instant $k-1$ the exact membership set is outer bounded by an ellipsoid $E(k-1)$ given by

$$E(k-1) = \{\mathbf{w} : [\mathbf{w} - \mathbf{w}(k-1)]^H \mathbf{C}^{-1}(k-1)[\mathbf{w} - \mathbf{w}(k-1)] \leq \sigma^2(k-1)\} \tag{4.8}$$

where $\mathbf{C}^{-1}(k-1)$ is a positive definite matrix which represents the shape, orientation and size of the ellipsoid and $\sigma^2(k-1)$ is a positive number defining the size of the ellipsoid and $\mathbf{w}(k-1)$ is the center of the ellipsoid.

A new ellipsoid $E(k)$ at time instant k that outer bounds the intersection $E(k-1) \cap H(k)$ is obtained by a linear combination of $E(k-1)$ and $H(k)$

$$\begin{aligned} E(k) = \{\mathbf{w} : \alpha(k)[\mathbf{w} - \mathbf{w}(k-1)]^H \mathbf{C}^{-1}(k-1)[\mathbf{w} - \mathbf{w}(k-1)] + \\ \beta(k)|d(k) - \mathbf{w}^H(k)\mathbf{z}_a(k)|^2 \leq \ldots \alpha(k)\sigma^2(k-1) + \beta(k)\gamma^2\} \end{aligned} \tag{4.9}$$

where $\alpha(k)>0$ and $\beta(k) \geq 0$ are time-varying weights will be optimized in the following.

4.2.1 Update Equations

The solution for the OBE type SMF problem formulated above comes out to be in the form of the following recursions

$$\mathbf{S}(k) = \frac{1}{\alpha(k)}\left[\mathbf{S}(k-1) - \frac{\beta(k)\mathbf{S}(k-1)\mathbf{z_a}(k)\mathbf{z_a}^H(k)\mathbf{S}(k-1)}{\alpha(k)+\beta(k)\mathbf{z}_a^H(k)\mathbf{S}(k-1)\mathbf{z}_a(k)}\right] \tag{4.10}$$

$$\mathbf{w}(k) = \mathbf{w}(k-1) + e^*(k)\beta(k)\mathbf{S}(k)\mathbf{z}_a(k) \tag{4.11}$$

$$\sigma^2(k) = \alpha(k)\sigma^2(k-1) + \beta(k)\gamma^2 - \frac{\beta(k)\alpha(k)|e(k)|^2}{\alpha(k)+\beta(k)\mathbf{z}_a^H(k)\mathbf{S}(k-1)\mathbf{z}_a(k)} \tag{4.12}$$

Different choices of $\alpha(k)$ and $\beta(k)$ give different algorithms for the parameter estimates. Setting $\alpha(k) = 1-\lambda(k)$ and $\beta(k) = \lambda(k)$ with $0 \leq \lambda(k) \leq 1$, we get a widely-linear version of the collaborative DH-OBE algorithm [38]. The optimal value for $\lambda(k)$ in DH-OBE is the one that minimizes $\sigma^2(k)$ in (4.12). An update is made only when $\sigma^2(k)+|e(k)|^2 > \gamma^2$ and in this case

$$\lambda(k) = \min(v(k), \lambda_{max}) \tag{4.13}$$

with

$$v(k) = \begin{cases} \lambda_{max} & \text{if } e(k) = 0 \\ \frac{1-\mu(k)}{2} & \text{if } G(n) = 0 \\ \frac{1}{1-G(n)}\left[1-\sqrt{\frac{G(n)}{1+\mu(k)(G(k)-1)}}\right] & \text{if } 1+\mu(k)(G(k)-1) > 0 \\ \lambda_{max} & \text{if } 1+\mu(k)(G(k)-1) \leq 0 \end{cases}$$

$G(k) = \mathbf{z}_a^H(k)\mathbf{S}(k-1)\mathbf{z}_a(k)$, $\mu(k) = \frac{\gamma^2(k)-\sigma^2(k)}{|e(k)|^2}$ and $\lambda_{max} \in (0,1)$, otherwise $\lambda(k) = 0$ and no update is made.

Another algorithm called BEACON is obtained by setting $\alpha(k) = 1$ and $\beta(k) = \lambda(k)$ in (4.10), (4.11) and (4.12). An update is made only when $|e(k)| > \gamma$ in which case the optimal value of $\lambda(k)$ [37] which minimizes (4.12) is given by

$$\lambda(k) = \begin{cases} 0 & \text{if } |e(k)| \leq \gamma \\ \frac{1}{G(k)}\left(\frac{|e(k)|}{\gamma} - 1\right) & \text{if } |e(k)| > \gamma \end{cases}. \tag{4.14}$$

Equation(4.14) shows two cases based on the magnitude of error signal relative to the error bound γ. When error is larger than γ an update for the filter parameters is made with corresponding non-zeros value of $\lambda(k)$ and when the error is less or equal to γ, $\lambda(k)$ is zero, and no parameter update is made. Based on the chosen value of γ, we can save significant amount of processing cycles. In a collaboration scenario, each new data vector at a node is first processed to check if local update is required or not. If an update is required locally the data is broadcasted over the network and similarly any node which receives data from the network utilizes it for its local update (if it is required). The underlying idea is that if new data at any node is useful to cause a local update then it might be useful for other nodes in the network and hence it is transmitted and if the data does not cause a local update, it is not transmitted. Hence, based on the value of γ, significant amount of communications can be saved. The linear and widely-linear versions of BEACON are summarized in Tables 4.3 and 4.4.

Table 4.3: Resource efficient linear collaborative RLS processing.

Resource efficient collaborative RLS algorithm for node n

Initialization:
$\mathbf{S}(0) = \mathbf{I}$
$\mathbf{h}(0) = 0$
for $k = 1{:}K$
{
  $\mathbf{z}_n(k) = [z_1^T(k), \ldots, z_N^T(k)]^T$
  $\mathbf{z}(k) = [\mathbf{z}_1^T(k), \ldots, \mathbf{z}_N^T(k)]^T$
  $e_n(k) = d_n(k) - \mathbf{h}^H(k)\mathbf{z}(k)$
  Check for transmission of new measurement and update
  If $|e_n(k)| > \gamma$
  {
    $G(k) = \mathbf{z}^H(k)\mathbf{S}(k-1)\mathbf{z}(k)$
    $\lambda(k) = \frac{1}{G(k)} \left(\frac{|e_n(k)|}{\gamma} - 1 \right)$
    $\mathbf{S}(k) = \mathbf{S}(k-1) - \frac{\lambda(k)\mathbf{S}(k-1)\mathbf{z}(k)\mathbf{z}^H(k)\mathbf{S}(k-1)}{1+\lambda(k)G(k)}$
    $\mathbf{h}(k) = \mathbf{h}(k-1) + e^*(k)\lambda(k)\mathbf{S}(k)\mathbf{z}(k)$
  }
  Receive data from the network and update the regressor vector or wait for regressors.
  $\mathbf{z}(k) = [\mathbf{z}_1^T(k), \ldots, \mathbf{z}_N^T(k)]^T$
}
If necessary, compute:
$y_n(k) = \mathbf{h}^H(k)\mathbf{z}(k)$
$e_n^{'}(k) = d_n(k) - y_n(k)$

Table 4.4: Resource efficient widely-linear collaborative RLS processing.

Resource Efficient WL collaborative RLS algorithm for node n

Initialization:
$\mathbf{S}(0) = \mathbf{I}$
$\mathbf{w}(0) = 0$
for $k = 1{:}K$
{
$\mathbf{z}_n(k) = [z_1^T(k), \ldots, z_N^T(k)]^T$
$\mathbf{z}(k) = [\mathbf{z}_1^T(k), \ldots, \mathbf{z}_N^T(k)]^T$
$\mathbf{z}_a(k) = [\mathbf{z}^T(k)\mathbf{z}^H(k)]^T$
$e_n(k) = d_n(k) - \mathbf{w}^H(k)\mathbf{z}_a(k)$
Check for transmission of new measurement and update
If $|e_n(k)| > \gamma$
{
$G(k) = \mathbf{z}_a^H(k)\mathbf{S}(k-1)\mathbf{z}_a(k)$
$\lambda(k) = \frac{1}{G(k)}\left(\frac{|e_n(k)|}{\gamma} - 1\right)$
$\mathbf{S}(k) = \mathbf{S}(k-1) - \frac{\lambda(k)\mathbf{S}(k-1)\mathbf{z}_a(k)\mathbf{z}_a^H(k)\mathbf{S}(k-1)}{1+\lambda(k)G(k)}$
$\mathbf{w}(k) = \mathbf{w}(k-1) + e^*(k)\lambda(k)\mathbf{S}(k)\mathbf{z}_a(k)$
}
Receive data from the network and update the regressor vector
or wait for regressors.
$\mathbf{z}(k) = [\mathbf{z}_1^T(k), \ldots, \mathbf{z}_N^T(k)]^T$
$\mathbf{z}_a(k) = [\mathbf{z}^T(k)\mathbf{z}^H(k)]^T$
}
If necessary, compute:
$y_n(k) = \mathbf{w}^H(k)\mathbf{z}_a(k)$
$e_n^{'}(k) = d_n(k) - y_n(k)$

Chapter 5

Simulations and Results

This chapter is dedicated to the implementation of the resource-efficient widely-linear collaborative wind forecasting algorithm, which was developed in the previous chapter, and discussion of results for simulated and real-world wind data. A data model for the highly-changing complex wind signal, which incorporates its noncircularity, spatial correlation and the temporal correlation, is presented. The non-circularity and the spatial correlation of the simulated complex wind signal are verified using real-imaginary plot and the spatial correlation profile of first wind data site, hereby called node, with all other nodes in the network is also presented. The MSE plots are used to show the convergence and the steady state behavior of the algorithm in single node and collaboration cases with linear and widely-linear processing. Simulations are performed for the algorithm with single node, collaboration in linear and widely-linear settings, with no resource saving and resource saving, and the results are discussed to highlight the performance gain achieved with the resource-efficient widely-linear processing of the complex wind signal. Real world wind data is also used to verify these results.

The chapter is organized into two major sections. Section 5.1 is related to the simulated wind data where, first we will first present the data model for noncircular and spatially correlated complex wind signal. Then the simulated wind data is analyzed for its non-circularity and spatial correlation. Next, the convergence and steady state MSE behavior of the algorithm are analyzed and finally the algorithm is applied on simulated wind data in single node and collaboration settings and the results are discussed. Section 5.2 contains the site information about the real world wind data and the simulation results.

5.1 Results with Simulated Wind Data

In this section we will present the simulated wind data model, analyze its properties and perform different experiments with the resource-efficient widely-linear collaborative wind forecasting algorithm.

5.1.1 Data Model

Here we present a model for complex wind data at N nodes, with spatial correlation coefficient of ρ between two neighbors. The complex wind signal at each node is obtained by combining the north-south (z_N) and east-west (z_E) components [7, 39] of the wind speed signal in North-East coordinate system

$$z = z_E + iz_N \tag{5.1}$$

where the wind speed v is given by $v = \sqrt{z_E^2 + z_N^2}$ and wind direction by $\theta = \tan^{-1}\left(\frac{z_N}{z_E}\right)$. The complex wind signal becomes noncircular when there is difference between the variances of the north and east components, where the individual components are normally distributed [40].

In order to generate spatially correlated complex wind signal for a network of N nodes, we adopt the following procedure. Let $\mathbf{x}(k) = [x_1(k), \ldots, x_N(k)]$ and $\mathbf{y}(k) = [y_1(k), \ldots, y_N(k)]$, where $x_n(k)$ and $y_n(k)$ are Normal random variables,

$$\begin{aligned} &x_n(k) \sim \mathcal{N}(0,1), \\ &y_n(k) \sim \mathcal{N}(0,1), \qquad n = 1 \ldots N, k = 1 \ldots K. \end{aligned} \tag{5.2}$$

The spatial correlation between the nodes is created using a spatial correlation matrix $\mathbf{R}$ of size $N \times N$

$$\tilde{\mathbf{x}}(k) = \mathbf{x}(k)\mathbf{R}^{\frac{1}{2}}, \tag{5.3}$$

$$\tilde{\mathbf{y}}(k) = \mathbf{y}(k)\mathbf{R}^{\frac{1}{2}}, \quad k = 1 \ldots K, \tag{5.4}$$

where $\mathbf{R}$ in case of N nodes assembled in straight line such as shown in Figure 5.1, is given by

$$\mathbf{R} = \begin{bmatrix} 1 & \rho & \rho^2 & \ldots & \rho^{N-1} \\ \rho & 1 & \rho & \ldots & \rho^{N-2} \\ \rho^2 & \rho & 1 & \ldots & \rho^{N-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{N-1} & \rho^{N-2} & \ldots & \rho & 1 \end{bmatrix}. \tag{5.5}$$

The next step is to model the temporal correlation of the wind data using a temporal correlation coefficient α

$$\tilde{\tilde{\mathbf{x}}}(k) = \alpha\tilde{\tilde{\mathbf{x}}}(k) + (1-\alpha)\tilde{\mathbf{x}}(k-1), \tag{5.6}$$

$$\tilde{\tilde{\mathbf{y}}}(k) = \alpha\tilde{\tilde{\mathbf{y}}}(k) + (1-\alpha)\tilde{\mathbf{y}}(k-1), \quad k = 2 \ldots K. \tag{5.7}$$

The final step to create the simulated wind data, with the desired mean vectors $\mu_{\mathcal{R}(\mathbf{z})}$ and $\mu_{\mathcal{I}(\mathbf{z})}$ and variance vectors $\sigma^2_{\mathcal{R}(\mathbf{z})}$ and $\sigma^2_{\mathcal{I}(\mathbf{z})}$ for real and imaginary components of the wind signal at N nodes, consists of first normalization of the spatially and temporally correlated data at each node to zero mean and unit variance and then assigning the desired mean and variance values.

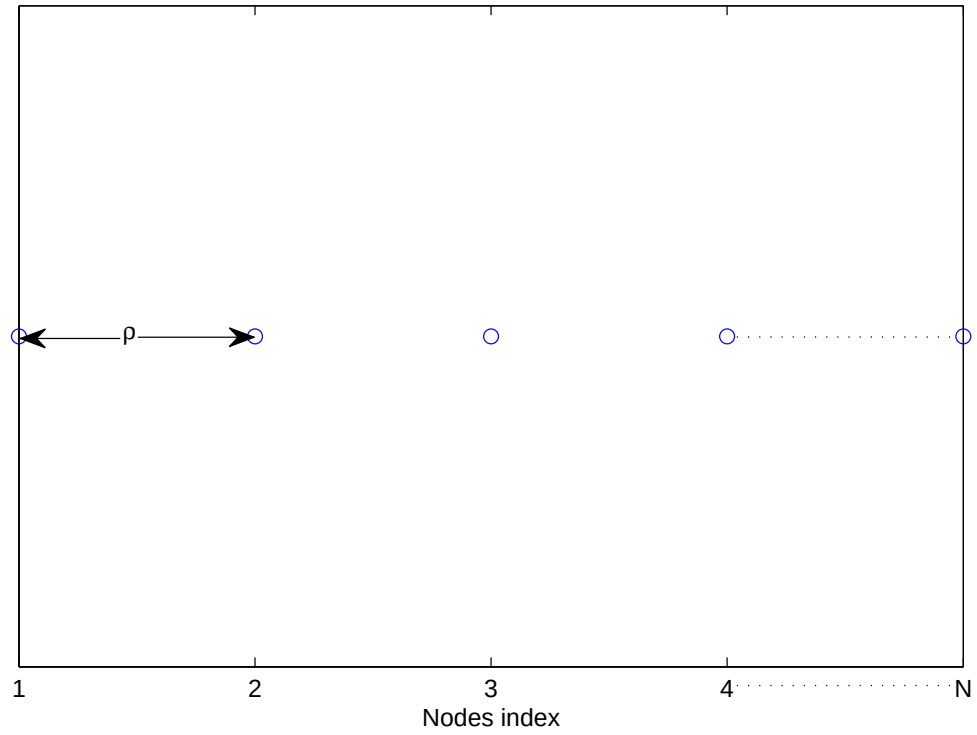


Figure 5.1: Network of N nodes in straight line with spatial correlation ρ between two adjacent nodes.

Let

$$\mu_{\tilde{\tilde{\mathbf{x}}}_n} = \mathbb{E}[\tilde{\tilde{\mathbf{x}}}_n], \tag{5.8}$$

$$\mu_{\tilde{\tilde{\mathbf{y}}}_n} = \mathbb{E}[\tilde{\tilde{\mathbf{y}}}_n], \tag{5.9}$$

$$\sigma^2_{\tilde{\tilde{\mathbf{x}}}_n} = \mathbb{E}[(\tilde{\tilde{\mathbf{x}}}_n - \mu_{\tilde{\tilde{\mathbf{x}}}_n})^2], \tag{5.10}$$

$$\sigma^2_{\tilde{\tilde{\mathbf{y}}}_n} = \mathbb{E}[(\tilde{\tilde{\mathbf{y}}}_n - \mu_{\tilde{\tilde{\mathbf{y}}}_n})^2], \quad n = 1 \ldots N. \tag{5.11}$$

be the mean and variance vectors of the spatially and temporally correlated data at N nodes, then it can be converted to zero mean and unit variance by

$$\mathbf{x}^{\circ}_n = \frac{\tilde{\tilde{\mathbf{x}}}_n - \mu_{\tilde{\tilde{\mathbf{x}}}_n}}{\sigma_{\tilde{\tilde{\mathbf{x}}}_n}}, \tag{5.12}$$

$$\mathbf{y}^{\circ}_n = \frac{\tilde{\tilde{\mathbf{y}}}_n - \mu_{\tilde{\tilde{\mathbf{y}}}_n}}{\sigma_{\tilde{\tilde{\mathbf{y}}}_n}}, \quad n = 1 \ldots N. \tag{5.13}$$

The real and imaginary components of the complex wind data with desired means and variances at N nodes are achieved by

$$\mathcal{R}(\mathbf{z}_n) = \sigma_{\mathcal{R}(\mathbf{z}_n)}\mathbf{x}^{\circ}_n + \mu_{\mathcal{R}(\mathbf{z}_n)}, \tag{5.14}$$

$$\mathcal{I}(\mathbf{z}_n) = \sigma_{\mathcal{R}(\mathbf{z}_n)}\mathbf{y}^{\circ}_n + \mu_{\mathcal{R}(\mathbf{z}_n)}, \quad n = 1 \ldots N. \tag{5.15}$$

where $\mathbf{Z} = \mathcal{R}(\mathbf{Z}) + i\mathcal{I}(\mathbf{Z})$ is the matrix containing spatially and temporally correlated data at N nodes with the desired means and variances.

5.1.2 Experiments and Results

Simulation Setup

We used $K = 5000$ number of samples and $N = 5$ number of nodes in our simulations. Considering a highly changing wind signal, the spatial correlation coefficient was set to $\rho = 0.85$ and the temporal correlation coefficient α of 0.05 was used. The mean value $\mu = 3.29$ and variance $\sigma^2 = 2.34$ for wind speed [7] are distributed between the real and imaginary parts with the variance of real part being $\frac{3}{4}\sigma^2$ and that of imaginary part $\frac{1}{4}\sigma^2$. All simulations are run for 1-step ahead prediction and the steady state RMSE is calculated for the iterations $1000 - 5000$. The MSE plots to study the transient and study state behavior are generated with $10,000$ independent number of trials. The regressor size of three is used for single node. The performance measure used in our simulations is root mean square error (RMSE). For M number of independent trials and K data samples, the RMSE is calculated by

$$RMSE = \sqrt{\frac{1}{MK}\sum_{k=0}^{K}\sum_{m=1}^{M}|e_m(k)|^2} \tag{5.16}$$

where $e_m(k)$ is the instantaneous error at iteration k of the realization m.

Data Analysis

In this section, we will study the non-circularity and spatial correlation of the simulated wind data.

Figure 5.2 shows the magnitude of simulated wind signal representing wind speed at the first node. The wind speed signal is highly changing representing a wind signal with fast variations in the high wind speed region. Wind signal with fast variations has high degree of non-circularity and spatial correlation. Next we will verify these properties of the simulated wind data.

Real-imaginary plot of the wind signal at the first node is shown in Figure 5.3(a). The plot is highly non-symmetric along x and y-axis, because of the difference in the variances of real and imaginary parts, thus representing a high degree of non-circularity of the wind signal.

Now, let us study the spatial correlation of the first node with all other nodes in the network. There are five nodes in the network and in Figure 5.3(b), we show the spatial correlation coefficient of the first node with node two, three, four and five at 0 time lag. From the plot we notice, that the spatial correlation of first node decreases exponentially in relation to the distance between first node and other nodes which is in agreement to the spatial correlation matrix of the network shown in the previous section.

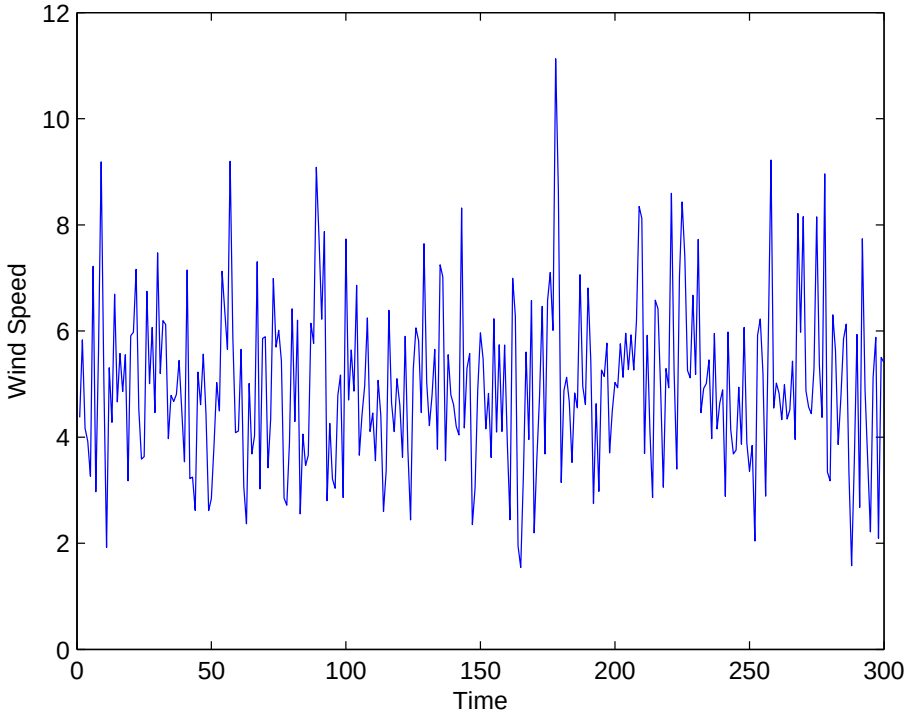


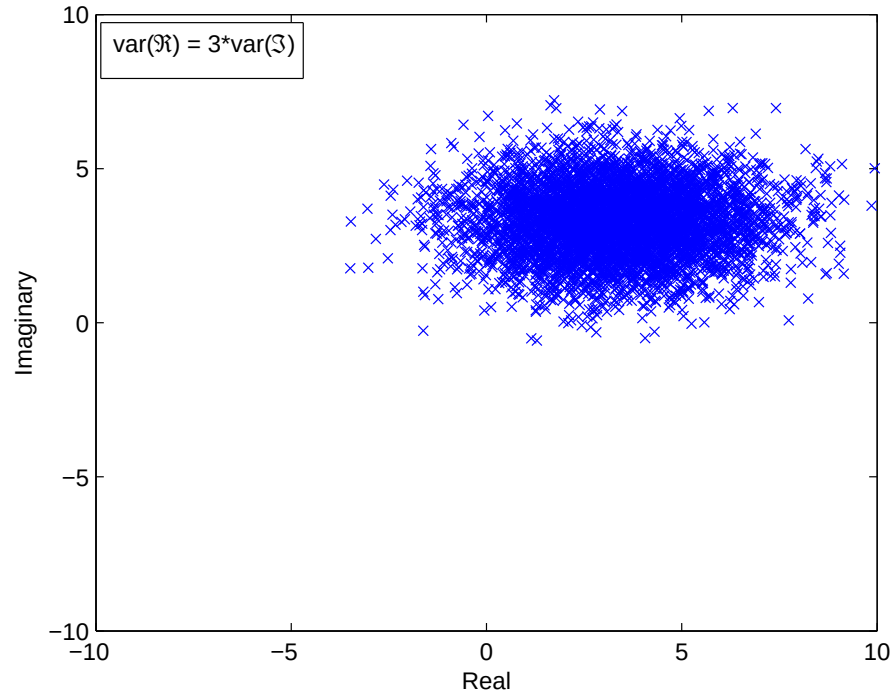
Figure 5.2: The first 300 simulated wind speed data points at first node.

In this section, we studied the non-circular and spatial correlation properties of the complex wind signal. These properties will be exploited in the following sections to improve the wind forecasting.

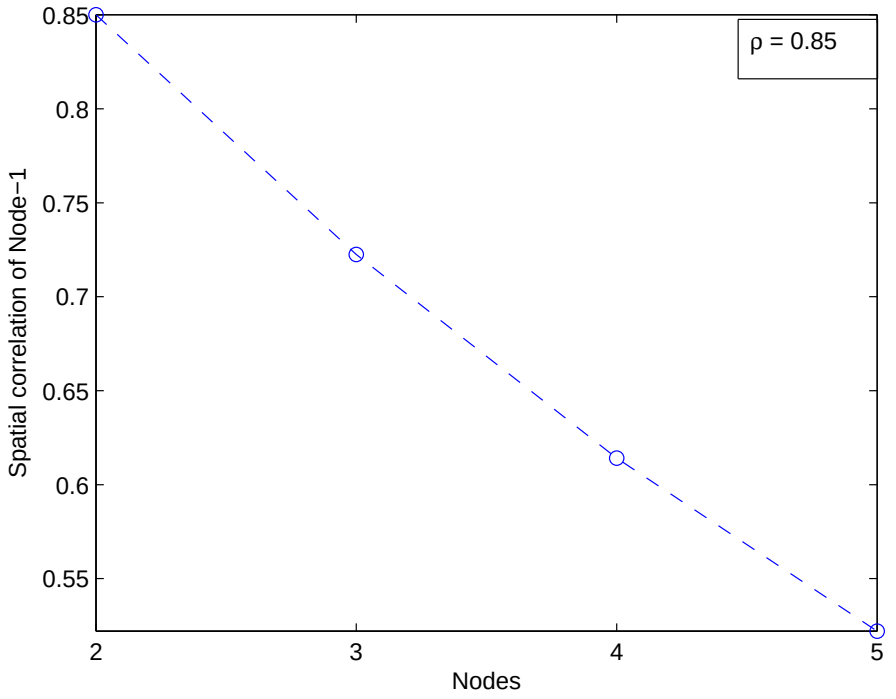
MSE Analysis of the Algorithms

In this section, we will look into the MSE behavior of the widely-linear collaborative forecasting algorithm in single node and collaboration cases. First, we will see the convergence and the steady state MSE for a single node with linear and widely-linear processing. Figure 5.4 shows the MSE behavior of single node for linear and widely linear cases. The convergence behavior shown in Figure 5.4(a) depicts the convergence in the case of linear processing is faster than widely linear but the steady state MSE in case of widely-linear processing is lower than the linear case which highlighted in Figure 5.4(b) more clearly. That means better forecasting performance is achieved with the widely-linear processing of the non-circular complex wind signal.

Figure 5.5 shows the MSE behavior in the case of three nodes collaboration with linear RLS processing. The convergence of the MSE curves in Figure 5.5(a) shows that the convergence speed in case of single node is faster than two nodes and three collaborating nodes and similarly convergence in case of two collaborating nodes is faster than three collaborating nodes meaning that as the level of collaboration increases, the convergence speeds slower. The steady state MSE in Figure 5.5(b) shows that the MSE in the case when three nodes collaborate is lower than when

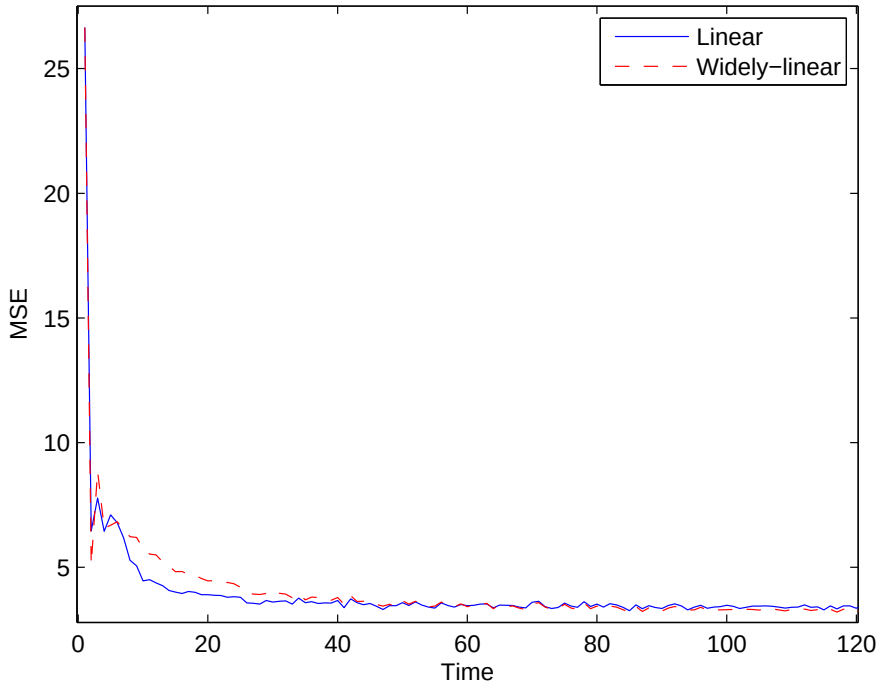


(a)

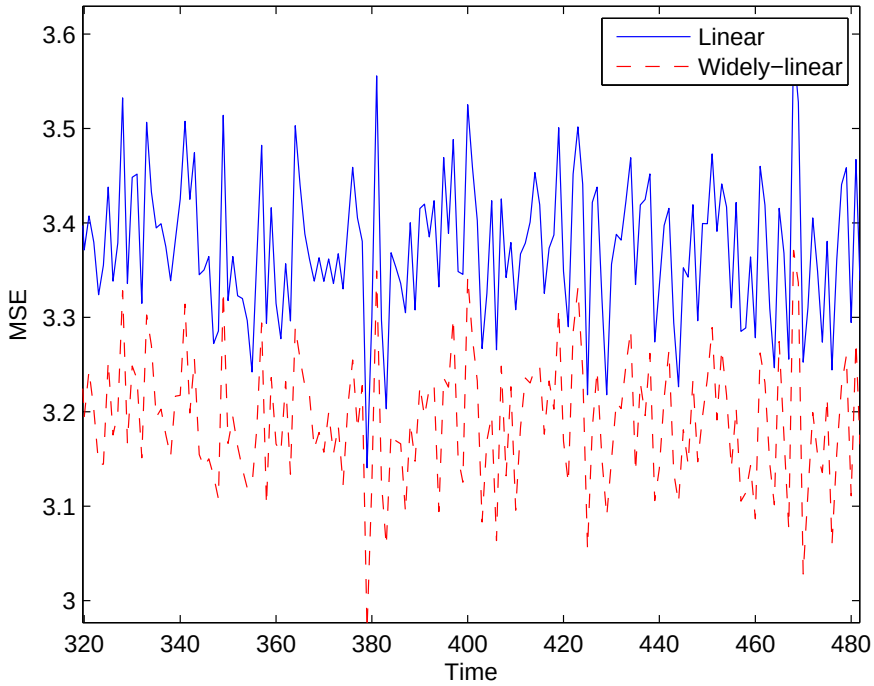


(b)

Figure 5.3: Non-circularity and spatial correlation of the simulated wind data at first node.

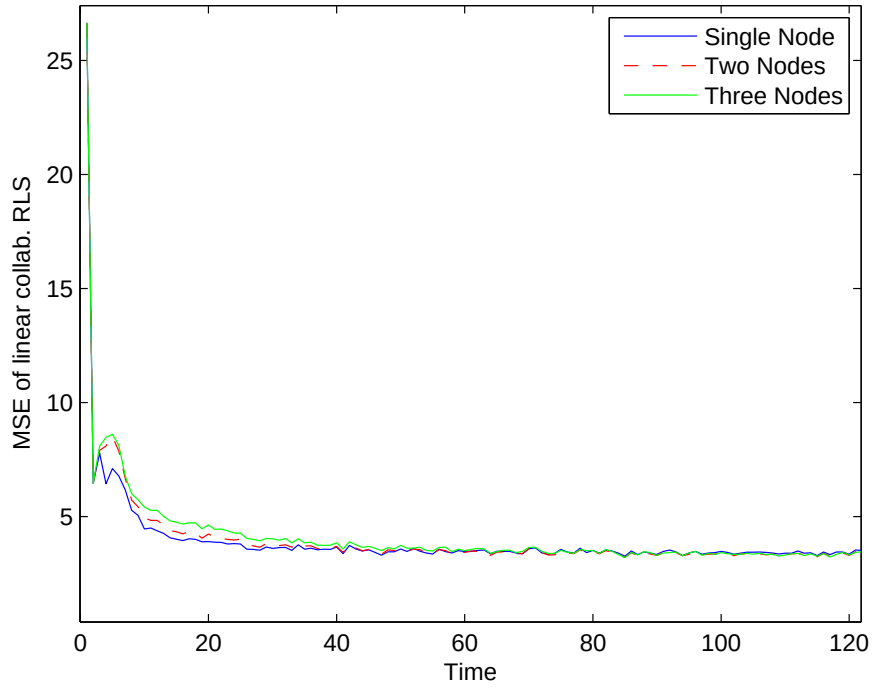


(a)

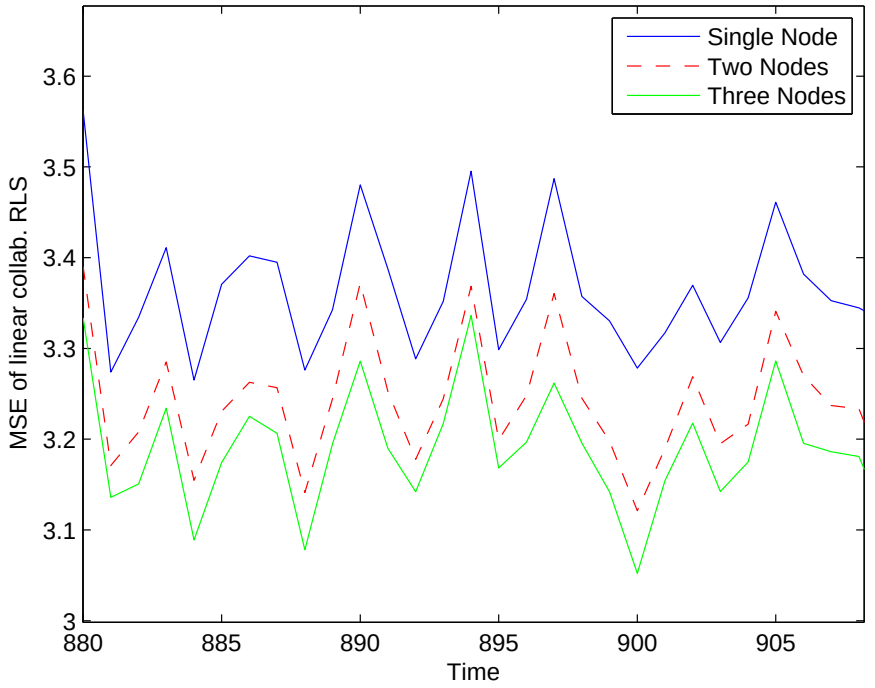


(b)

Figure 5.4: MSE behavior at first node with linear and widely-linear RLS.



(a)



(b)

Figure 5.5: MSE at first node with single node (no collaboration), two and three nodes collaboration processing.

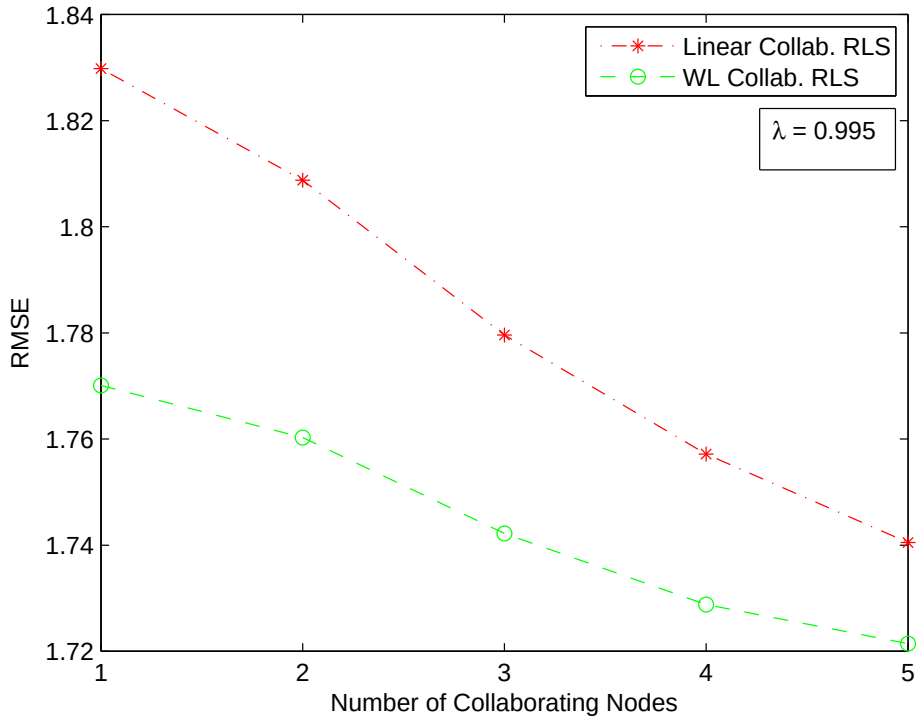


Figure 5.6: RMSE plot for linear and WL RLS with no resource saving.

two nodes collaborate, and single node processing without collaboration.

From this section, we conclude that for non-circular and spatially correlated complex wind data, we can achieve better forecasting results by using widely-linear processing and collaboration. In the following sections, we will use the widely-linear and collaboration techniques jointly, to further improve the wind forecasting performance.

Collaborative RLS Processing

In this section we discuss the results of linear and widely-linear collaborative RLS algorithm on the simulated wind data.

Figure 5.6 shows the output curves of RMSE versus the number nodes for linear and widely-linear collaborative processing. From the figure we observe that the RMSE with widely-linear processing is lower than that of linear processing for all levels of collaboration. The RMSE for both linear and widely-linear cases decrease as the level of collaboration increases. Thus, we can infer that better performance is achieved with widely-linear and collaborative processing than a single node linear processing and no collaboration.

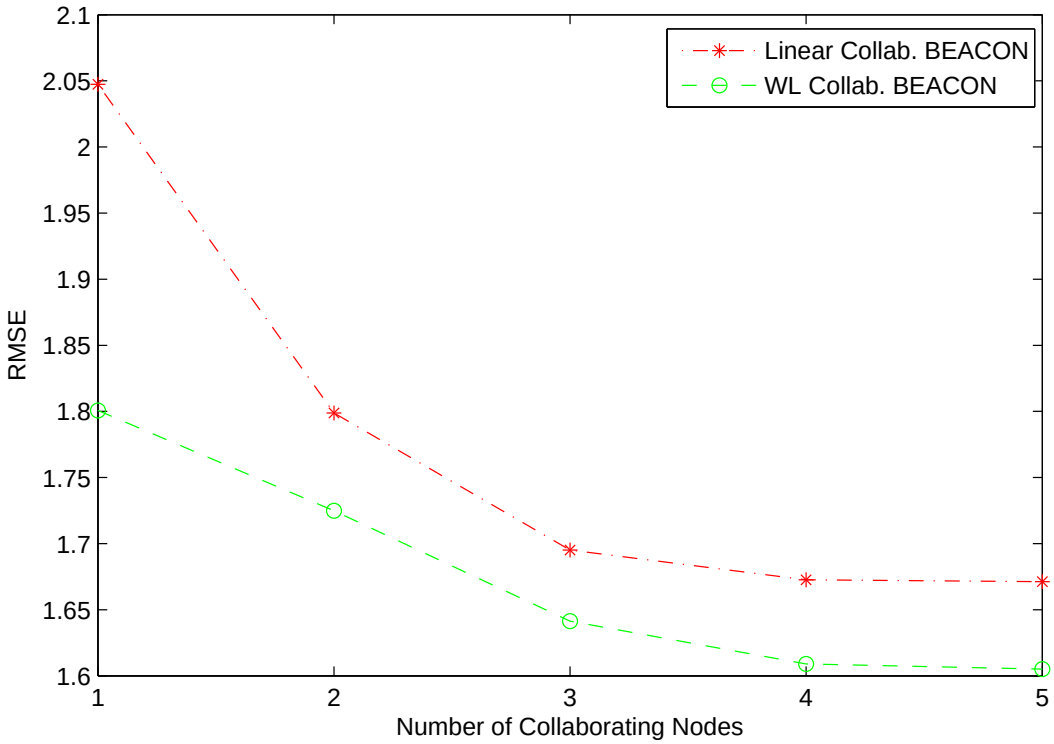


Figure 5.7: RMSE plot for linear and WL RLS with resource saving.

Resource-Efficient Processing

Figure 5.7 shows RMSE versus the number of collaborating nodes for linear and widely-linear collaborative BEACON algorithms using simulated wind data. Similar to collaborative RLS processing, the RMSE decreases with increasing number of collaborating nodes and widely-linear processing gives lower RMSE values than linear processing. Comparing these results with the results in the previous figure, we notice that as the level of collaboration increases the RMSE values in case of resource-efficient processing become even lower than those of collaborative RLS processing and the reason for this improvement is that in resource efficient processing only the useful data is utilized for an update, thus making the estimates even better. Along with better results significant amount reduction in communication and processing is also achieved, which is tabulated below.

Table 5.1 shows the number of communication and processing cycles for single node processing (no collaboration) and two or more nodes collaboration with linear and widely-linear processing. From the table we observe that the number of communication and processing cycles are just around $1-2\%$. Thus a significant saving in communication and local processing which is a critical consideration in the systems with limited resources.

The gain in resource savings increases monotonically with the increasing value of error threshold γ. Next, we will study the effect of γ on RMSE and the resource

usage for a range of the values of γ.

Table 5.1: Resource saving.

Setting	Linear				Widely-Linear			
Resource	Proc.		Comm.		Proc.		Comm.	
Usage	cycles	%	cycles	%	cycles	%	cycles	%
Single Node	14	1.4	-	-	12	1.2	-	-
Two Nodes	24	2.4	24	2.4	23	2.3	23	2.3
Three Nodes	9	0.9	9	0.9	9	0.9	9	0.9
Four Nodes	12	1.2	12	1.2	13	1.3	13	1.3
Five Nodes	9	0.9	9	0.9	9	0.9	9	0.9

Figure 5.8 shows the effect of γ values on RMSE for single node processing (no collaboration) and two or more nodes collaboration with linear (Figure 5.8(a)) and widely-linear (Figure 5.8(b)) processing. From the figures we observe that the RMSE values, for both linear and widely-linear cases, are minimum for a particular value of γ and increases exponentially with the increasing value ofγ. Thus the selection of error threshold directly effects the performance of the algorithm.

In Figure 5.9, we show the relation between the value of γ and the resource usage for the linear and widely-linear collaboration. From the Figure 5.9(a) and Figure 5.9(b), we observe that the resource usage decreases monotonically with increasing value of γ for any level of collaboration in the cases of linear and widely-linear processing both.

5.2 Results with Real World Wind Data

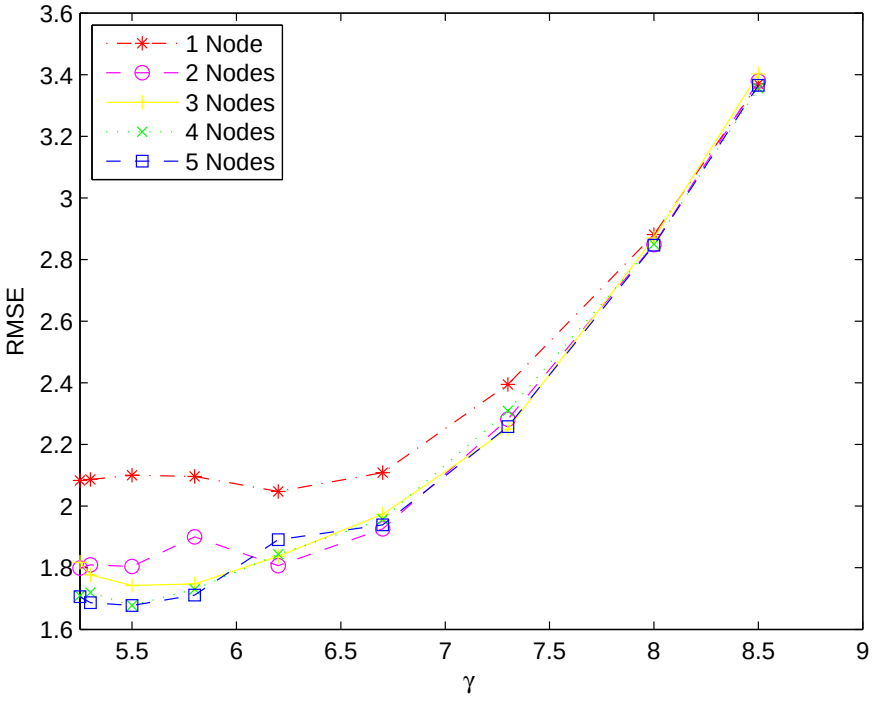
5.2.1 Data Description

The wind data used in our work is freely available at the website of Alternative Energy Institute (AEI) which was formed in 1977 at West Texas University (now West Texas A&M University).[1]

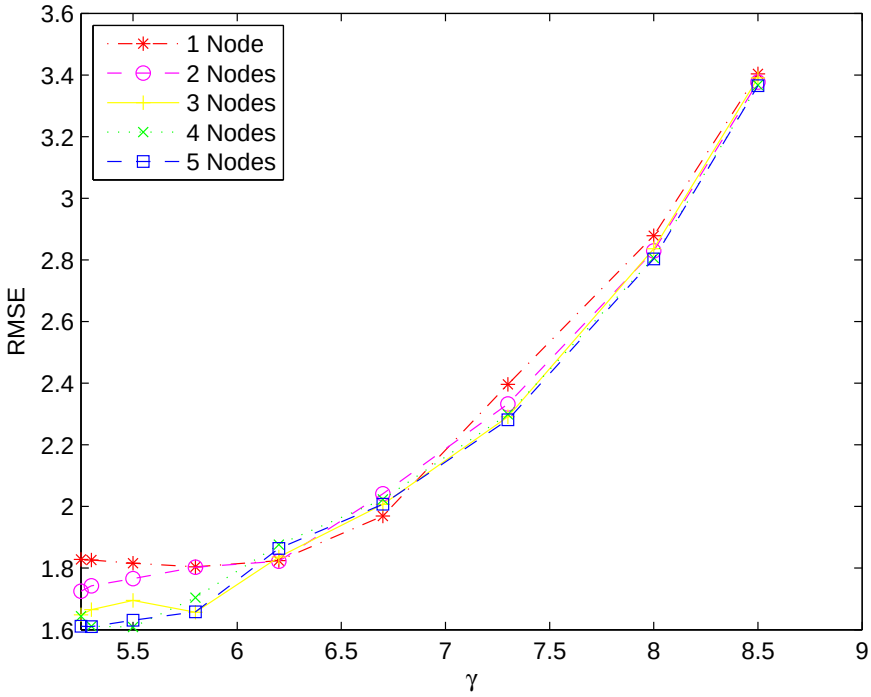
Two sites, selected for data collection, are based on their relative location which is small enough to have good spatial correlation, availability of data for the same time scale and the relative height of the wind measurement sensors.

- Clay Archer County - Site #1358 with wind speed sensors at the height of 20 m, 40 m, 45 m and 60 m. Wind direction sensors are placed at a height of 40 m and 50 m.

- Young County 2 - Site #1988 with wind speed sensors at the height of 18m and 28m and wind direction sensors at the height of 28m.

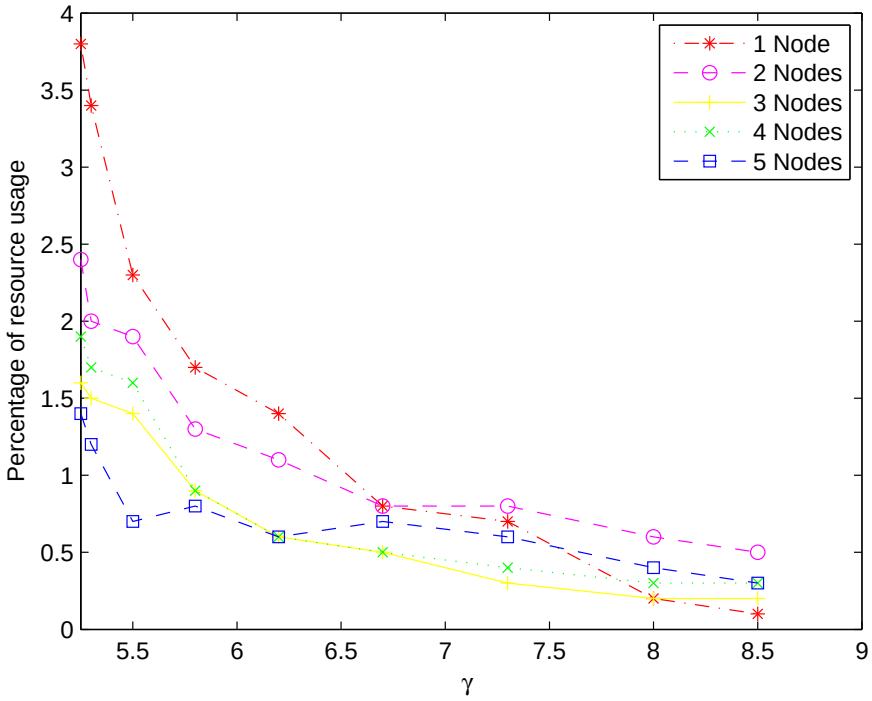


(a) Linear processing

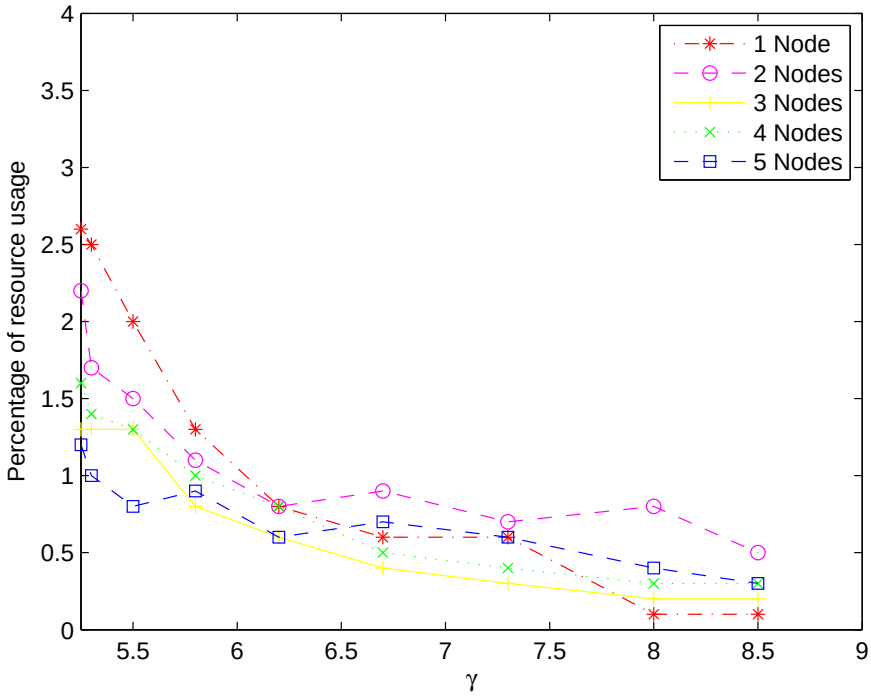


(b) Widely-linear processing

Figure 5.8: Effect of γ on output RMSE for different levels of collaboration.



(a) Linear processing



(b) Widely-linear processing

Figure 5.9: Effect of error threshold γ on system resources.

Table 5.2: RMSE values for linear and widely-linear resource efficient collaboration.

Setup	No Resource Saving		With Resource Saving	
	Linear	Widely-Linear	Linear	Widely-Linear
Site1	1.43	1.40	1.77	1.54
Site1-Site2	1.37	1.36	1.38	1.35

Table 5.3: Resource saving.

Setting	Linear				Widely-Linear			
Resource	Proc.		Comm.		Proc.		Comm.	
Usage	cycles	%	cycles	%	cycles	%	cycles	%
Site1	62	11.27	-	-	73	13.27	-	-
Site1-Site2	99	18	99	18	121	22	121	22

5.2.2 Experiments and Results

The linear and the widely-linear collaborative RLS algorithms are applied on the real world wind data mentioned in Section 5.2.1 for the month of July, 2007 with 550 samples for the filter training the remaining samples as the test data. The results are shown in Table 5.2 and we observe, that the RMSE value for a single node is improved in case of widely-linear processing and further improved when two nodes are collaborating. The combined results of widely-linear with two nodes collaboration gives much improved result for RMSE as compared to the single node with linear processing.

The resource efficient linear and the widely-linear collaborative BEACON algorithms with γ value of (4.4) have similar performance as their RLS counterparts but with significant reduction in communication and processing cycles.

Table 5.3 shows the number of communication and processing cycles in the cases of single processing (no collaboration) and two nodes collaboration. We see that the number of communication and processing cycles are less than 25% as compared to the case with no resource saving, thus a significant gain in resource saving is achieved.

[1] `http://www.windenergy.org`

Chapter 6

Conclusions

This thesis works presented a novel method of wind speed forecasting which exploits the non-circularity as well the spatial correlation of the complex wind signal.

Widely-linear modeling was used to include the non-circularity, and regressor data from all nodes was gathered at each node and utilized to incorporate the spatial correlation of the wind signal. The widely-linear collaborative model was then further optimized for efficient resource utilization using set-membership filtering.

Simulated wind data was used to study the performance gains achieved with this approach. The generated data was made non-circular and spatially and temporally correlated. The resource efficient widely-linear collaborative algorithm was applied on the non-circular and spatially correlated simulated wind data in different settings. Results with no collaboration (single node), collaboration in linear setting and widely-linear with resource-saving and without resource-efficient processing were discussed and compared and it was shown that the widely-linear collaborative model is superior in performance in comparison to a single with linear processing. The resource-efficient processing provided significant saving in communications and processing cycles. Real world wind data was also used to show the results in linear and widely-linear collaboration cases.

Bibliography

[1] W. A. Edelstein, C. J. Walcek, D. L. Cox, and L. C. Davis, "Wind energy," Panel on Public Affairs (POPA), American Physical Society, Report, 2003.

[2] S. Pal, "Wind energy - an innovative solution to global warming?" in *1st International Conference on the Developments in Renewable Energy Technology*, 2010.

[3] A. Reeves, "Wind energy for electric power," Renewable Energy Policy Project, Report, 2003.

[4] Y. Zhang and K. W. Chan;, "The impact of wind forecasting in power system reliability," in *Electric Utility Deregulation and Restructuring and Power Technologies*, 2008.

[5] A. S. Hering and M. G. Genton, "Powering up with space-time wind forecasting," *Journal of the American Statistical Association*, vol. 105, pp. 92–104, Mar 2010.

[6] S. L. Goh, D. Popovic, and D. Mandic, "Complex-valued estimation of wind profile and wind power," in *Proceedings of the 12th IEEE Mediterranean Electrotechnical Conference*, May 2004.

[7] D. Mandic, S. Javidi, S. Goh, A. Kuh, and K. Aihara, "Complex valued prediction of wind profile using augmented complex statistics," *Renewable Energy*, vol. 34, pp. 196–201, 2009.

[8] J. Tastu, P. Pinson, E. Kotwa, H. Madsen, and H. A. Nielsen, "Spatio-temporal analysis and modeling of short-term wind power forecast errors," *Wind Energy*, vol. 14, no. 1, pp. 43–60, 2011.

[9] M. Bhaskar, A. Jain, and N. Venkata Srinath, "Wind speed forecasting: Present status," in *International Conference on Power System Technology*, 2010.

[10] C. Potter, A. Archambault, and K. Westrick, "Building a smarter smart grid through better renewable energy information," in *IEEE/PES Power Systems Conference and Exposition*, 2009.

[11] "The smart grid," Argonne National Laboratory, Tech. Rep., 2010.

[12] S. V. and M. D. Vadari, "Investigating smart grid solutions to integrate renewable," Battelle Energy Technology, Tech. Rep., 2009.

[13] M. Ahlstrom and R. Zavadil, "The role of wind forecasting in grid operations & reliability," in *IEEE/PES Transmission and Distribution Conference and Exhibition: Asia and Pacific*, 2005.

[14] B. A. and W. J., "Wind power forecasting and electricity market operations," Decision and Information Sciences Division, Argonne National Laboratory, Tech. Rep.

[15] F. Castellanos and N. James, "Average hourly wind speed forecasting with ANFIS," in 11^{th} *Americas Conference on Wind Engineering*, San Juan, Puerto Rico, June 2009.

[16] D. Dini and D. P. Mandic, "A class of widely linear complex Kalman filters," University Defense Research Centre (UDRC), Tech. Rep. TR-IC-UDRC-WL-Kalman/0257-09-10-2, 2011.

[17] E. Ollila, V. Koivunen, and H. Poor, "A robust estimator and detector of circularity of complex signals," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, may 2011, pp. 3620–3623.

[18] D. P. Mandic and S. L. Goh, *Complex Valued Nonlinear Adaptive Filters: Noncircularity, Widely Linear and Neural Models.* Wiley Online Library, 2009.

[19] B. Picinbono and P. Bondon, "Second-order statistics of complex signals," *IEEE Transactions on Signal Processing*, vol. 45, no. 2, pp. 411–420, Feb 1997.

[20] S. Javidi, B. Jelfs, and D. Mandic, "Blind extraction of noncircular complex signals using a widely linear predictor," in *IEEE/SP 15th Workshop on Statistical Signal Processing*, Sep 2009, pp. 501–504.

[21] P. Chen, T. Pedersen, B. Bak-Jensen, and Z. Chen, "ARIMA-based time series model of stochastic wind power generation," *IEEE Transactions on Power Systems*, vol. 25, no. 2, pp. 667–676, May 2010.

[22] R. Toriumi, H. Katsuchi, and N. Furuya, "A study on spatial correlation of natural wind," *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 87, no. 2-3, pp. 203–216, 2000.

[23] M. Fujimura and J. Maeda, "Cross-correlation of fluctuating components of wind speed based on strong wind measurement," in *The Seventh Asia-Pacific Conference on Wind Engineering*, Nov 2009.

[24] T. Imai, K. Kusunoki, Y. Hono, T. Takemi, K. Araki, T. Fukuhara, K. Bessho, S. Hoshino, and T. Shibata, "Spatial correlation and temporal fluctuation of wind," in *The Seventh Asia-Pacific Conference on Wind Engineering*, 2009.

[25] S. Watson, L. Landberg, and J. Halliday, "Application of wind speed forecasting to the integration of wind energy into a large scale power system," *IEE Proceedings on Generation, Transmission and Distribution*, vol. 141, no. 4, pp. 357–362, Jul 1994.

[26] R. H. Madsen, "A protocol for standardizing the performance evaluation of short-term wind power prediction models," Technical University of Denmark, IMM, Ricard Petesens Plads, Bld. 321, 2800 Lyngby, Denmark, Tech. Rep., 2004.

[27] S. Soman, H. Zareipour, O. Malik, and P. Mandal, "A review of wind power and wind speed forecasting methods with different time horizons," in *North American Power Symposium*, Sep 2010, pp. 1–8.

[28] G. Boyle, "Renewable electricity and the grid: The challenge of variability," Earthscan, Tech. Rep., 2007.

[29] G. Sideratos and N. Hatziargyriou, "An advanced statistical method for wind power forecasting," *IEEE Transactions on Power Systems*, vol. 22, no. 1, pp. 258–265, Feb. 2007.

[30] J. J. Traiteur, "A short term ensemble wind-speed forecasting system for wind power applications," Ph.D. dissertation, University of Illinois at Urbana-Champaign, 2011.

[31] A. Boone, "Simulation of short-term wind speed forecast errors using a multivariate ARMA(1,1) time-series model," Ph.D. dissertation, Royal Institute of Technology, Stockholm, Sweden, 2005.

[32] P. S. R. Diniz, *Adaptive Filtering: Algorithms and Practical Implementation.* Amazon, 2008.

[33] M. Soltani, S. M. Shakeri, J. Grunnet, T. Knudsen, and T. Bak, "Wind farm decentralized dynamic modeling with parameters," Aalborg University, Tech. Rep., 2010.

[34] M. Miranda and R. Dunn, "Spatially correlated wind speed modelling for generation adequacy studies in the uk," in *Power Engineering Society General Meeting, 2007. IEEE*, June 2007, pp. 1–6.

[35] T. Gneiting, K. Larson, K. Westrick, M. G. Genton, and E. Aldrich, "Calibrated probabilistic forecasting at the stateline wind energy center," *Journal of the American Statistical Association*, vol. 101, pp. 968–979, Sep 2006.

[36] S. Gollamudi, S. Nagaraj, S. Kapoor, and Y.-F. Huang, "Set-membership filtering and a set-membership normalized LMS algorithm with an adaptive step size," *IEEE Signal Processing Letters*, vol. 5, no. 5, pp. 111–114, May 1998.

[37] S. Nagaraj, S. Gollamudi, S. Kapoor, and Y.-F. Huang, "BEACON: an adaptive set-membership filtering technique with sparse updates," *IEEE Transactions on Signal Processing*, vol. 47, no. 11, pp. 2928–2941, Nov 1999.

[38] S. Gollamudi, S. Kapoor, S. Nagaraj, and Y.-F. Huang, "Set-membership adaptive equalization and an updator-shared implementation for multiple channel communications systems," *IEEE Transactions on Signal Processing*, vol. 46, no. 9, pp. 2372 –2385, Sep 1998.

[39] X. Yili, B. Jelfs, M. Van Hulle, J. Principe, and D. Mandic, "An augmented echo state network for nonlinear adaptive filtering of complex noncircular signals," *IEEE Transactions On Neural Networks*, vol. 22, no. 1, pp. 74–83, Jan 2011.

[40] H. Aksoy, Z. F. Toprak, A. Aytek, and N. E. Ünal, "Stochastic generation of hourly mean wind speed data," *Renewable Energy*, vol. 29, no. 14, pp. 2111–2131, Mar 2004.