

Deterministic modelling of the underlying dynamics of surface wind speed fluctuations with application to grid-connected wind power generation

Report of the Major Research Project submitted to University Grants Commission (UGC) (F. No. 42-30/2013(SR) dated 12 March 2013)

K. Satheesh Kumar

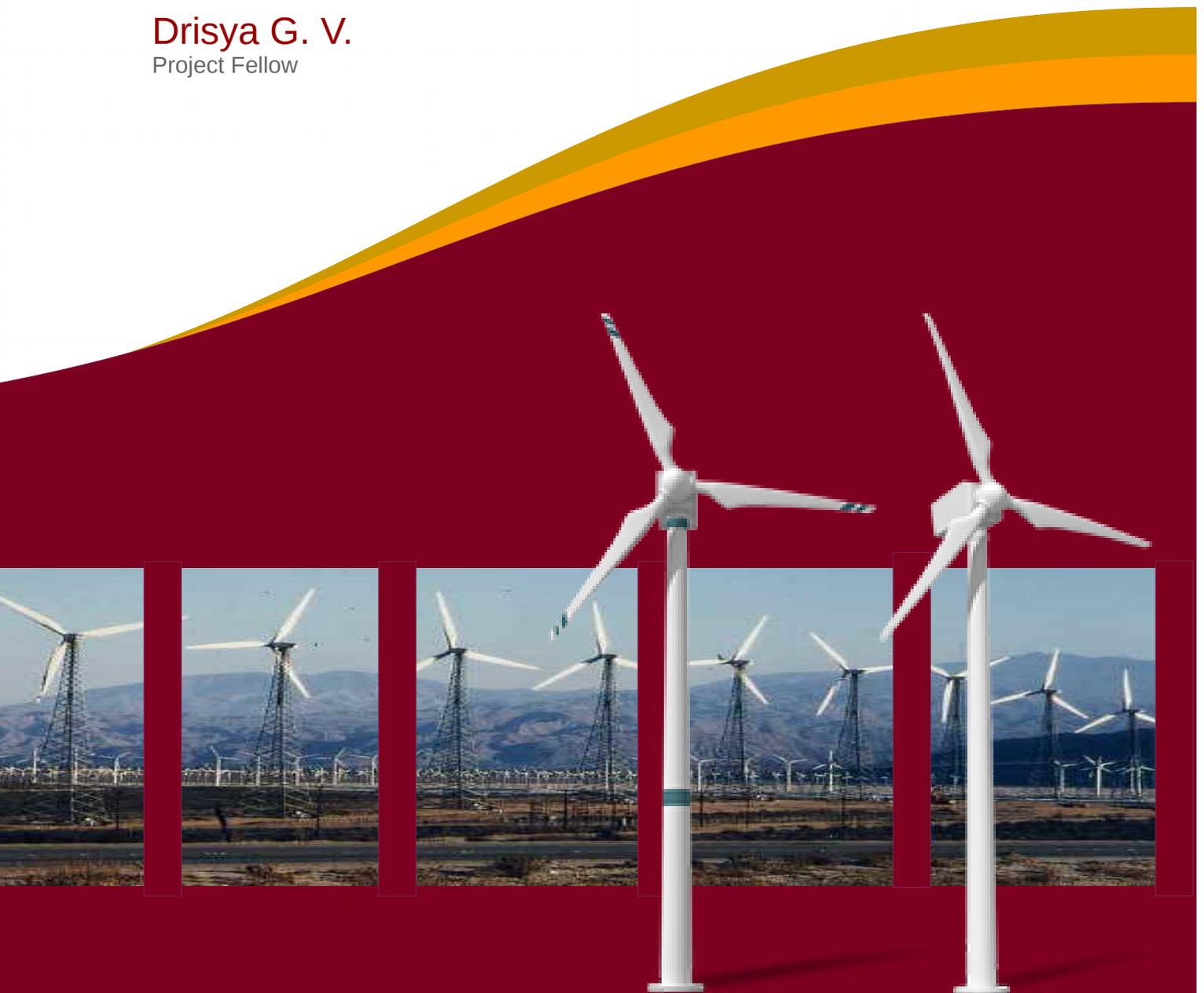
Principal investigator, Department of Futures Studies, University of Kerala

K. Asokan

Co-investigator, Department of Mathematics, College of Engineering, Thiruvananthapuram

Drisya G. V.

Project Fellow



Deterministic modelling
of the underlying dynamics of surface wind speed fluctuations
with application to grid-connected wind power generation

Report
of the Major Research Project submitted to
University Grants Commission (UGC)
(F. No. 42-30/2013(SR) dated 12 March 2013)

K. Satheesh Kumar
Principal Investigator
Department of Futures Studies, University of Kerala
kskumar@keralauniversity.ac.in

K. Asokan
Co-Investigator
College of Engineering, Thiruvananthapuram

Drisya G. V.
Project Fellow

Department of Futures Studies
University of Kerala
Kariavattom PO, Thiruvananthapuram, Kerala - 695 581

Acknowledgements

*W*e are grateful to the University Grants Commission, New Delhi, financial assistance and University of Kerala for the infrastructure facilities and support. We are also grateful to Dr T. R. Ramamohan Retired Chief Scientist, CSIR 4PI, Wind Tunnel Road, Bangalore for stimulating discussions, useful suggestions and encouragement.

K. Satheesh Kumar

K. Asokan

Drisya G. V.

Abstract

Modelling the underlying dynamics of surface wind fluctuation is significant because of its potential impacts on various fields ranging from agriculture to structural engineering. The most important area of application, which recently gained increased attention, is in the energy management sector, as the wind is widely recognised as a clean, economically viable and eco-friendly source of electric power. Many research organisations worldwide are working on different aspects of it with the aim of making wind energy technology more cost-effective and reliable. Wind resource characterization is one of the important research topics reported by *IEA wind* for a long term R&D and it contributes to research needs like site optimisation, design conditions, improvement of wind forecasting techniques *etc.* Wind speed modelling and forecasting is an important aspect of wind power generation and distribution, yet one of the most difficult tasks due to myriads of factors affecting it. The short-term prediction of one to six hours ahead at intervals of 10 minutes are important in power dispatching systems. Although many time series analysis (ANN, ARMA, ARIMA *etc.*) and meteorological modelling (NWP, Prediktor *etc.*) techniques are available to represent the wind speed dynamics, the prediction error could not be reduced significantly compared to the method of persistence. Most of the studies on wind speed prediction reported in the literature are based on statistical methods or the probabilistic distribution of the wind speed data, assuming that the underlying dynamics wind speed fluctuations are stochastic in nature. However, with the advent of chaos theory, it is noted that irregular random-like fluctuations can also arise from a deterministic system with a low dimensional chaotic character. In other words, since the chaotic systems are very much sensitive to the initial condition it may exhibit quite complex behaviour like stochastic systems affected by noise that attributes to the prediction error. Even if the long-term prediction is impossible for a chaotic system, an accurate short-term prediction is always possible using a deterministic model. Therefore it is important to explore the source of these irregular fluctuations. In this work, as a first step, we investigate the deterministic nature of the underlying dynamics of surface wind fluctuations by carrying out a detailed non-linear time series analysis on wind speed data measured at various locations across Indian sub-continent. The results of the analysis strongly suggest that the underlying dynamics is *deterministic, low-dimensional and chaotic* across all locations. These results open up the possibility of developing deterministic models capable of accurate short-term prediction. Interestingly, this is one example of a naturally occurring time series showing chaotic behaviour as most of the chaotic systems are confined to laboratories.

Motivated by the above observations, we investigated further the possibility of developing deterministic prediction models capable of accurate short-term prediction, which is important in various stages of wind energy management such as wind turbine predictive control and wind power scheduling. The statistical analysis of the deterministic model predictions, utilising wind speed measurements at 234 different geographical locations, shows that the predictions are remarkably accurate up to one hour with normalised root mean square error of less than 0.02 and reasonably accurate up to three hours with an error of less than 0.06. Comparison of the results with f-ARIMA model predictions shows that the deterministic models with suitable parameters are capable of returning improved prediction accuracy and capturing the dynamical variations of the actual time series more faithfully. These methods are simple and computationally efficient and require only records of past data for making short-term wind speed forecasts within a practically tolerable margin of errors.

Wind speed oscillations are known to exhibit varying characteristics at different time scales, and a range of models from simple persistence schemes to complex physical models has been used to capture this contrasting behaviour. The recent analysis by our group has shown that a collection of auto-regressive (AR) models fitted separately on frequency components of wind speed time series can significantly increase the prediction accuracy indicating the inability of a single model capturing the entire range of behaviour possibly due to the diverse nature of dynamical characteristics. Therefore, as a further step, we investigated the diverse dynamical characteristics across the wide frequency spectrum of wind speed measurements. The results of the analysis show the variation of stochastic, deterministic and chaotic behaviour apart from the dimensionality of underlying dynamics as well as the degree of fluctuations. Such an analysis would be useful for adopting the most suitable model for fluctuations at a specific range of interest or building hybrid models capturing the entire range of behaviour. It is also demonstrated that a cluster of deterministic models built upon separate frequency components of a wind speed time series can enhance the prediction accuracy as much as 80%, on the average, consistently for predictions up to 12 hours as validated by a statistical analysis of the predictions over a set of locations. The comparison shows a definite advantage of deterministic prediction models over autoregressive models. The f-index introduced in this thesis measure the fluctuations of wind speed over a period and it shows that the observed seasonal variations of prediction errors can be correlated with changes in the f-index of the component series contributed mostly by the lower scales of decomposition.

Contents

List of publications	1
List of Figures	3
List of Tables	9
1 Introduction	1
1.1 Surface wind	1
1.2 Impact of wind dynamics on various fields	3
1.2.1 Agriculture	3
1.2.2 Navigation	3
1.2.3 Structural engineering	3
1.3 Wind energy	4
1.3.1 Energy scenario	4
1.3.2 Issues related to wind energy production and distribution	5
1.3.3 Wind farm - operation and maintenance	7
1.4 Wind speed modelling and forecasting	9
1.4.1 A brief overview of forecasting models	10
1.4.2 Meteorological modelling	11
1.4.3 Time series modelling for wind speed forecasting	12
1.4.4 Machine learning techniques and hybrid models	13
1.5 Nonlinear time series analysis and prediction	16
1.5.1 Phase space methods	17
1.5.2 Quantitative measure of complexity	19
1.5.3 Estimation of topological invariants	20
2 Underlying Dynamics of Wind Speed Fluctuations	23
2.1 Introduction	23
2.2 Chaotic system	26
2.3 Attractor reconstruction	27
2.4 Nonlinear time series analysis	29
2.5 Surrogate data test	40
3 Deterministic Prediction of Surface Wind Speed Variations	45
3.1 Introduction	45
3.2 Analysis of the data	47
3.3 Predictions	51
3.4 Statistical analysis of prediction errors	56
3.5 Comparison with f-ARIMA	60
3.6 Conclusions	63

4	Empirical mode decomposition and chaos based prediction model for wind speed oscillations	65
4.1	Introduction	65
4.2	Methodology	66
4.2.1	Nonlinear time series and prediction tools	66
4.2.2	Empirical mode decomposition	68
4.3	Results and discussions	69
4.4	Conclusion	70
5	Week-ahead predictions of wind speed using simple linear models with wavelet decomposition	71
5.1	Introduction	71
5.2	Maximum Overlap Discrete Wavelet Transform	73
5.3	Wind speed forecasting models	75
5.3.1	New Reference Forecast model	75
5.3.2	Auto-Regressive model	76
5.3.3	Auto-Regressive Fractionally Integrated Moving Average model	81
5.4	Results and discussion	82
5.5	Conclusion	101
6	Diverse dynamical characteristics across the frequency spectrum of wind speed fluctuations	103
6.1	Introduction	103
6.2	Chaotic time series	105
6.3	Wavelet decomposition	106
6.4	Fluctuation Index	109
6.5	Results and discussion	109
6.5.1	Analysis of decomposed data	109
6.5.2	Improved predictions	114
6.5.3	Statistical analysis of predictions	117
6.6	Conclusions	120
7	Summary	123
	Bibliography	127

List of Publications

Journals

1. Drisya, G. V., Kiplangat, D. C., Asokan, K., & Kumar, K. S. (2014), Deterministic prediction of surface wind speed variations. *Ann. Geophys*, 32, 1415-1425. (Chapter 3)
2. Kiplangat, Dennis C, K Asokan, and K Satheesh Kumar (2016). Improved week-ahead predictions of wind speed using simple linear models with wavelet decomposition. *Renewable Energy* 93, pp. 38–44. issn: 18790682. doi: 10.1016/j.renene.2016. 02.054. url: <http://www.sciencedirect.com/science/article/pii/S0960148116301550>. (Chpater 5)
3. Kiplangat, Dennis C, G V Drisya, and K Satheesh Kumar (2016), Day ahead forecast of wind speed intensity using time series partitioning, *International Journal on Computer Science and Engineering* 8.8, pp. 313–319, issn: 09753397. url: <http://www.enggjournals.com/ijcse/doc/IJCSE16-08-08-011.pdf>.

Under communication

4. G. V. Drisya, K. Asokan, K. S. Kumar (2016), Diverse dynamical characteristics across the frequency spectrum of wind speed fluctuations, *Renewable Energy* (Under Revision) (Chapter 6).

Conferences

5. Drisya, G. V., & Kumar, K. S. (2015), Empirical mode decomposition and chaos based prediction model for wind speed oscillations, *Intelligent Computational Systems (RAICS), 2015 IEEE Recent Advances*, pp. 306-311, IEEE. (Chapter 4)
6. G. V. Drisya, D. Kiplangat, K. S. Kumar (2015), Enhancing Accuracy of Deterministic Prediction of Surface Wind Speed Variations using Wavelet Decomposition, *National Conference on Foresight and Technology Futures 2015 - University of Kerala*.
7. Angel, A K, Dennis C Kiplangat, K Satheesh Kumar (2016). Wind Speed Forecasting using Markov Chains. *17th National Conference on Technological Trends - College of Engineering - Trivandrum*.

-
8. Kiplangat, Dennis C, K Satheesh Kumar (2015). Modelling of Wind Speed Fluctuations Across Indian Sub-Continent. *National Conference on Time Series, Analytics and Recent Advances in Statistical - Central University of Rajasthan.*
 9. Kiplangat, Dennis C, G V Drisya, K Satheesh Kumar (2015). Renewable Energy for a Sustainable Future. *National Conference on Foresight and Technology Futures 2015 - University of Kerala.*

List of Figures

1.1	Representation of Global Circulation Model.	2
1.2	Schematic diagram of wind energy production and distribution.	8
2.1	Location at which Daily Mean Wind Speed (DMWS) data is measured.	26
2.2	Time series of the daily mean wind speed (DMWS) in knots, measured across Indian subcontinent at locations (a) Bhuj (b) Ahemadabad (c) Jabalpur (d) Birsa Munda Airport (e) Coimbatore (f) Anantapur (g) Akola (h) Indira Gandhi Airport.	28
2.3	Space-time separation plot for the time series measured at locations (a) Bhuj (b) Ahemadabad (c) Jabalpur (d) Birsa Munda Airport (e) Coimbatore (f) Anantapur (g) Akola (h) Indira Gandhi Airport.	30
2.4	Time series of the daily mean wind speed (DMWS) in knots, measured across Indian subcontinent at locations (a) Bhuj (b) Ahemadabad (c) Jabalpur (d) Birsa Munda Airport. (e) Coimbatore (f) Anantapur (g) Akola (h) Indira Gandhi Airport.	31
2.5	Space-time separation plot for the time series measured at locations (a) Bhuj (b) Ahemadabad (c) Jabalpur (d) Birsa Munda Airport (e) Coimbatore (f) Anantapur (g) Akola (h) Indira Gandhi Airport.	32
2.6	Mutual information of the time series of the DMWS data measured at (a) Bhuj (b) Ahemadabad (c) Jabalpur (d) Birsa Munda Airport (e) Coimbatore (f) Anantapur (g) Akola (h) Indira Gandhi Airport.	34
2.7	The fraction of false nearest neighbours as a function of the embedding dimension m for the DMWS data measured at (a) Bhuj (b) Ahemadabad (c) Jabalpur (d) Birsa Munda Airport (e) Coimbatore (f) Anantapur (g) Akola (h) Indira Gandhi Airport.	35
2.8	The delay representation of the detrended time series (a) Bhuj (b) Ahemadabad (c) Jabalpur (d) Indira Gandhi Airport.	36
2.9	The local slopes $D_2(\epsilon, m)$ for the de-noised and detrended time series of DMWS at (a) Jabalpur (b) Coimbatore (c) Anantapur and (d) Akola. The values of the correlation dimension are given in Table 2.2.	37
2.10	The curve of $S(\delta n)$ for various embedding dimensions m . The maximum Lyapunov exponent of the detrended time series is the slope of the dashed line. (a) Bhuj (b) Ahemadabad (c) Jabalpur (d) Birsa Munda Airport (e) Coimbatore (f) Anantapur (g) Akola (h) Indira Gandhi Airport.	39
2.11	The mean values of the fraction of false nearest neighbours of the surrogates with standard deviation.(a) Bhuj (b) Ahemadabad (c) Jabalpur (d) Birsa Munda Airport (e) Coimbatore (f) Anantapur (g) Akola (h) Indira Gandhi Airport.	41
2.12	Plot of the significance of difference S versus m for different locations	42
2.13	The mean values of $S(\Delta)$ of the surrogates with standard deviation.(a) Bhuj (b) Ahemadabad (c) Jabalpur (d) Birsa Munda Airport (e) Coimbatore (f) Anantapur (g) Akola (h) Indira Gandhi Airport.	43
2.14	Plot of the significance of difference S versus Δn for different locations	44

-
- 3.1 Time series of wind speed at location given by Latitude: 34.98420, Longitude: -104.03971 at 80 metres. 47
- 3.2 The Space time separation plot for the time series for $m = 14$ and $\tau = 85$. Each point in the plot represents a pair of points on the trajectory with their relative separation in time along the horizontal axis and separation in space along the vertical axis. The diurnal variations are evident in this figure. 48
- 3.3 (a) The autocorrelation function of wind speed data. (b) Mutual information of the wind speed data as a function of delay (τ). 49
- 3.4 The fraction of false nearest neighbours as a function of the embedding dimension m for the time series of wind speed with $\tau = 85$, showing that any $m \geq 14$ can be considered optimal. 50
- 3.5 The local slopes $D_2(\varepsilon, m)$ for the wind speed time series with $\tau = 85$ and m ranging from 14 to 16 showing a plateau for small values of ε and giving an estimate of $D_2 = 1.656 \pm 0.008$ 51
- 3.6 The curve of $S(\Delta n)$ for embedding dimensions $m = 14, 15$. The maximum Lyapunov exponent of the time series is the slope of the dashed line 0.15 ± 0.006 52
- 3.7 The power spectrum of wind speed time series as a function of frequency at location Latitude: 34.98420 Longitude: -104.03971. Broadband and exponential decay of the power with frequency are typical characteristics of chaotic signals. 53
- 3.8 Comparison of predicted values with the actual values for LFO and RBF. The symbols are plotted only at every 30 minutes for legibility. (a) Latitude: 42.31925 Longitude: -98.60197 $m = 5, \tau = 8$ (b) Latitude: 43.51076 Longitude: -99.47652, $m = 12, \tau = 2$, 54
- 3.9 (a) The exponential decay of the correlation coefficient between the predicted and actual for every 3 prediction time steps (b) The exponential growth RMS error prediction for every 3 prediction time steps. The arrows show axes for the respective symbols of data points. 55
- 3.10 The geographic locations, denoted by filled circles, where wind speed data was analysed for performance of deterministic model prediction. 56
- 3.11 (a) NRMSE of 1 hour predictions using LFO, for the period from 2004 to 2006 at an interval of 30 days, averaged over 234 locations. (b) NRMSE of 1 hour predictions using LFO for the 234 locations, calculated at an interval 30 days and averaged over the period from 2004 to 2006. (c) & (e) are location averaged NRMSE for 2 and 3 hour predictions respectively. (d) & (f) are time averaged NRMSE for 2 and 3 hour predictions respectively. The error bars in the figures are with respect to the standard error of the mean. The horizontal dotted lines in (b), (d) and (f) represent the mean of the respective time averaged NRMSE values over the entire set of locations. 57
- 3.12 NRMSE, with standard error, of 1 hour prediction for the period of 3 years from 2004 over 234 locations using LFO and RBF. The lines connecting the symbols are to guide the eye. 58
- 3.13 Comparison of predicted values with the actual values for LFO and f -ARIMA. The symbols are plotted only at every 30 minutes for legibility. Latitude: 34.98420, Longitude: -104.03971, (a) $m = 13, \tau = 3$ and (b) $m = 8, \tau = 7$ 59

3.14	Location averaged NRMSE, with standard error, of (a) 1 hour, (b) 3 hour and (c) 6 hour predictions for the period of 3 years from 2004 over 234 locations using LFO and f -ARIMA.	60
4.1	Time series of wind speed at location given by latitude: 47.11332° N longitude: 90.44666° W at 80 m.	67
4.2	Comparison of predicted values with actual values for EMD decomposition and without decomposition for four different locations (a) Latitude: 34.9842° N Longitude: 104.03971° W (b) Latitude: 41.85498° N Longitude: 97.61899° W (c) Latitude: 46.73381° N Longitude: 101.7077° W (d) Latitude: 37.84476° N Longitude: 100.06653° W	68
4.3	Time and location averaged NRMSE for predictions up to 10 hours ahead for Lfo with and without EMD.	70
5.1	Estimates of the prediction variance of the time series that is not explained by the autoregressive model for the three methods.	77
5.2	The geographical locations, denoted by filled circles, for the 234 sites selected for wind speed data analysis and prediction.	83
5.3	The D1–D3 frequency components from the decomposed 10-minute wind speed data for a a time period of one week.	84
5.4	The D4–D6 frequency components from the decomposed 10-minute wind speed data for a a time period of one week.	85
5.5	The D7–D9 frequency components from the decomposed 10-minute wind speed data for a a time period of one week.	86
5.6	The D10–D12 frequency components and the S12 scaling component from the decomposed 10-minute wind speed data for a a time period of one week.	87
5.7	Comparison of predicted values with actual data for AR and ARFIMA models with and without wavelet decomposition. (a) Latitude: 44.34406° N Longitude: 99.61266° W, prediction start time: 2004-12-26 12:10:00 (b) Latitude: 45.39850° N Longitude: 103.51002° W, prediction start time: 2006-03-21 12:10:00	89
5.8	Comparison of predicted values with actual data for AR and ARFIMA models with and without wavelet decomposition. (c) Latitude: 44.95404° N Longitude: 96.60688° W, prediction start time: 2005-09-22 12:10:00 (d) Latitude: 38.67878° N Longitude: 98.59783° W, prediction start time: 2004-01-31 12:10:00	90
5.9	Comparison of location and time averaged root mean squared errors of hourly mean wind speed data between AR and ARFIMA models both combined with wavelet decomposition.	91
5.10	Plots of averaged root mean squared errors for 1 day ahead predictions at (a) different locations and (b) different time intervals.	92
5.11	Plots of averaged root mean squared errors for 2 days ahead predictions at (a) different locations and (b) different time intervals.	93
5.12	Plots of averaged root mean squared errors for 3 days ahead predictions at (a) different locations and (b) different time intervals.	94
5.13	Plots of averaged root mean squared errors for 4 days ahead predictions at (a) different locations and (b) different time intervals.	95

5.14	Plots of averaged root mean squared errors for 5 days ahead predictions at (a) different locations and (b) different time intervals.	96
5.15	Plots of averaged root mean squared errors for 6 days ahead predictions at (a) different locations and (b) different time intervals.	97
5.16	Plots of averaged root mean squared errors for 7 days ahead predictions at (a) different locations and (b) different time intervals.	98
5.17	(a) Location and time averaged RMSE for predictions up to 9 days ahead. (b) Percentage gain of accuracy in terms of root mean squared errors for the <i>AR model</i> and the <i>New Reference Forecast Model (NRFM)</i> method when used in combination with wavelet decomposition.	100
6.1	Diagram showing temporal fluctuations of signal $x(t)$	107
6.2	Delay reconstruction (x_n versus x_{n-d}) for each component of wavelet decomposed time series at location Latitude: 34.9842°N , Longitude: 104.03971°W . The delay d was estimated for each component using mutual information and auto-correlation.	108
6.3	The mean values of the fraction of false nearest neighbours of the surrogates with standard deviation for levels 2-9 at location Latitude: 34.9842°N , Longitude: 104.03971°W	110
6.4	Plot of the significance of difference S versus embedding dimension for levels 4-8.	111
6.5	Variation of Lyapunov exponent for each wavelet component averaged over 212 sites.	111
6.6	Variation of correlation dimension for each wavelet component averaged over three typical sites.	112
6.7	Estimated embedding dimension for each wavelet component.	112
6.8	The average over 40 locations of significance of difference of one step predication error at each level of the original and surrogates.	113
6.9	Comparison of LFO and LFO with wavelet decomposition predicted values with actual values of wind speed. The symbols are plotted only at every one hour for legibility. (a) Latitude: 45.87565 Longitude: -103.30938 (b) Latitude: 34.85728 Longitude: -103.61320	114
6.10	(a) RMSE of predictions for the period from 2004 to 2006 at an interval of 30 days averaged over 212 locations. (b) RMSE at 212 locations averaged over predictions at 30 days apart during the same period.	115
6.11	(a) Location and time averaged RMSE for predictions up to 72 hours ahead. (b) Percentage gain of accuracy in terms of RMSE for LFO method in combination with wavelet decomposition over LFO method. The symbols are plotted only at every two hour for legibility.	116
6.12	(a) Comparison of location and time averaged RMSE of AR model and LFO method both in combination with wavelet decomposition for predictions up to 72 hours ahead. (b) Percentage gain of accuracy in terms of RMSE of predictions of LFO method over AR model, both in combination with wavelet decomposition.	118
6.13	Location averaged RMSE of 12 hour predictions of LFO method with wavelet decomposition at an interval of 30 days and corresponding location averaged f -index. The correlation between RMSE and f -index is 0.697.	119

6.14 (a) Standard Deviation of the wind speed variations for each month at different level and (b) Fluctuation index for each month at different level at location latitude: 34.9842°N, longitude: 104.03971°W. 121

List of Tables

2.1	The geographical locations where wind speed data have been considered for analysis in this thesis.	25
2.2	The estimated values of the correlation dimension at various locations. The values appear to depend more on the local topography rather than geographical location.	37
2.3	The estimated values of the maximum Lyapunov Exponent at various locations. The values do not vary significantly across geographical locations.	38

1

Introduction

Modelling the fluctuations of the Earth's surface wind has a significant role in understanding the dynamics of atmosphere besides its impact on various other fields including agriculture, navigation, structural engineering calculations and reduction of atmospheric pollution. The Recent surge of interest in research and development related to wind power is due to its potential as an alternate source of energy. In this chapter, we give a brief overview of the significance of the wind speed modelling and forecasting, and non-linear time series analysis method.

1.1 Surface wind

Surface wind is one of the major factors that plays a crucial role in climate and weather systems of the earth. It has significant impact on agriculture, navigation, structural engineering calculations and reduction of atmospheric pollution as well as the economy of the region as a alternate energy source (Martin et al., 1999; Elliott, 2004; Banta et al., 2011). Wind is caused by air flowing from high pressure area to low pressure area. Generally the pressure gradient is generated by the unequal heating of the earth's surface. The movement of wind is deflected by the rotation of earth, steering it towards the east or west depending on the hemisphere and the direction of the current. These factors combine to form various global currents of air such as *Trade Winds*, *Westerlies* and *Polar Easterlies*.

The *Trade Winds* blow towards the equator from the north east in the Northern Hemisphere and from the south east in the Southern Hemisphere. They are caused by hot air rising at the Equator—which results in an area of low atmospheric pressure along the Equator known as *doldrums*—and the consequent movement of air from north and south to take its place. The winds are deflected towards the west because of the earth's west-to-east rotation.

The *Westerlies* are the prevailing winds that occur in both hemispheres between 30 and 60 degrees latitude. They blow from high pressure areas in the subtropical area in both hemispheres (known as Horse Latitudes) towards the subpolar low pressure area. Again the rotation of earth steers the *Westerlies* towards the northwest in the northern hemisphere and the southwest in the southern hemisphere.

The *Polar Easterlies* are the dry, cold prevailing winds that blow from the high-pressure areas at the poles towards the low-pressure areas in the subpolar region. They blow in the northeast direction in northern hemisphere and in the southeast direction in southern hemisphere.

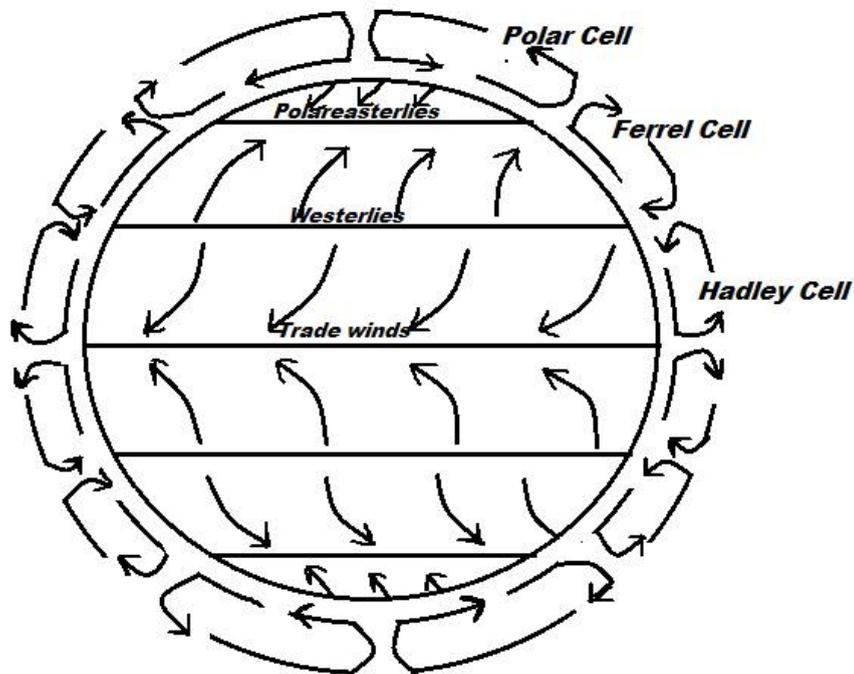


FIGURE 1.1: Representation of Global Circulation Model.

Apart from these major air currents there are numerous local winds generated by local conditions and topography. Common among these are the *land* and *sea breezes* caused by the unequal heating (and cooling in the night) of coastal land and sea water, the *mountain valley breezes* generated by the uneven heating of mountain tops and valley surfaces and the *kebatic winds* originating from radiational cooling of air atop an elevated plateau making it more dense and causing it to flow downwards.

Solar energy falls unequally on the earth surface, the largest near equator and smallest at the poles, is the major driving force of massive movement of air. This atmospheric circulation redistributes the heat energy by transferring warm air from low to high latitude and cold air from high to low latitude. Redistribution of heat is well explained by three main convection cells Known as, Hadley Cell, Ferrel Cell and Polar Cell. Figure. 1.1. shows an idealised view of global atmospheric circulation on earth.

As shown in the Figure. 1.1, convection cell near to equator between 0° (equator) and 30° latitude is the Hadley cell. The air circulation pattern in this cell is, warm air rises above the surface, then near the tropopause, it diverges towards poles and subsequently by the Coriolis effect turned towards the equator. The resulting surface wind in Hadley cell is known as Tradewinds. Ferrel cell found between 30° , and 60° latitude controls the climate types in mid-latitude. About 30° N and S latitude the rising air from equator sinks and some part of the descending flow splits towards poles. Same thing happens about 60° N, and S latitude, the rising air towards poles splits and some flows to opposite direction joining the Ferrel cell. This cell is entirely driven by the other two cells, and the kind of the wind in this cell is known as westerlies. Convection cell located at 60° to 90° of the equator is termed as Polar cell. Since poles receive much less heat from the sun, cold air creates a high pressure causing an air flow towards the equator. At around 60° , due to Coriolis effect air flow is deflected towards poles causing a weak, irregular prevailing wind commonly termed as Easterlies. Three circulation cells combined with Coriolis effect, balance the energy spread between poles and equator and regulate the temperature of the earth.

1.2 Impact of wind dynamics on various fields

The recent surge in research in the area of wind speed modelling and forecasting is due to the extensive penetration of wind power generation as an alternate energy source. However, it is significant in many other fields ranging from agriculture to the structural engineering calculations. In this section, we discuss some of these applications.

1.2.1 Agriculture

Surface wind plays a vital role in our social development, because of its significant impact on almost all the fields of progress and civilisation. In the agricultural sector wind speed and direction has predominant effects, both beneficial and detrimental, on crop production. The wind is favourable, since it helps photosynthesis rates by increasing the supply of carbon dioxide, increases the ethylene production in plants, helps in pollination and so on. The wind can also cause damages to crops for the reasons like, morphological changes in plants, increased crop water requirements due to evapotranspiration, uprooting from soil, etc. Navigation is the another field where wind speed is of high influence.

1.2.2 Navigation

In Aircraft navigation wind speed and direction is used by the pilot for monitoring and controlling the movement of aircraft. According to the direction which wind is blowing navigators fine-tune the speed and angle for takeoffs and landings of the craft. If the direction of wind blow is same as of an airplane, it is known as Tailwind and is regarded as not a favourable condition near runway as its adverse effect in landing and takeoffs. Although during flight tailwind increases the speed of aircraft and reduces the time required to reach the destination, in the course of takeoff airplane may need reduced climb angle and much more runway to get enough lift. In the case of landing, near the ground, there will be a decrease in tailwind causing an inertial effect by which increased lift occurs following a temporary floating of the aircraft and dangerously high landing speeds. For the wind blowing opposite direction of the plane, known as Headwind, is termed as the favourable condition for takeoffs and landing. Headwind helps aircraft to takeoff with sooner lift, lower ground speed and use of the shorter runway. Landing into the headwind also reduces the velocity of aircraft with respect to ground and uses only a short distance to come to a complete stop. Wind speed is also an important factor in aviation as climb and approach performances of aircraft are highly related to it. The speed of headwind increases the angle of climb during takeoff and steepen the angle of approach in the time of landing. The aviation industry uses wind speed forecasting for a safe and economic operation of aircraft.

1.2.3 Structural engineering

Another application of wind speed is in structural engineering, where the analysis is needed to measure the effect of the wind on structures. Understanding of wind behaviour is crucial in all phases of building process. During the design phase itself, inputs from wind engineers are drawn to avoid hurdles that may occur in design and construction stage of the building. Effects like up- lift, tilting and overturning of buildings are three primary effects of wind. Blowing wind causes a load pressure on buildings which can cause damages to the structure with improper

design. As the speed of the wind varies with height, wind load on a structure depends upon the height of the wind above ground level. Dimension and design of the structure are determined by the combined action of external and internal air load acting upon it. A systematic analysis of meteorological wind data history is needed for identifying different types of the wind storm and its impact on buildings. Sometimes computer simulation techniques, which incorporates load points shear force distributions, are also used for preparing a strategic, safe, cost effective and a feasible design of the proposed structure. Structural engineers also think about the direction of the wind for a comfortable environment and the performance of ventilation systems.

1.3 Wind energy

Apart from all these, wind plays a crucial role in the economic development of a country as it is widely recognised as a clean, inexpensive, inextinguishable alternate source of energy. Traditional energy sources like oil, gas, coal and nuclear are known for its production of many greenhouse gases and harmful knock-on effects such as emission of radioactive materials into the environment. According to International Energy Agency, even if 550 million people in the world will be remaining in the dark without any access to energy, by 2040 total demand of electricity may increase by more than 70% (IEA, 2015). Energy sector is considered as the largest and primary source of greenhouse-gas (GHG) emissions which are the heart of global climate change. An obvious consideration in this scenario is, projecting the need of a well defined strategic policies to at least reduce the environmental impact it may cause. By now an energy sector transition is underway in many countries and even though coal remains in leading position, the supportive concern and policies placed renewable resources as the second-largest source of electricity in worldwide. As a fastest growing economy in the world, to meet the growing demand for energy (2.8% per year), India is trying to track the opportunities of safe, sustainable and innovative energy policies. In connection with our pledge in Climate summit, India is looking for a keen cleaner path for energy, aiming to reach 175 GW of installed renewable capacity by 2022 (IEAIndia, 2015).

1.3.1 Energy scenario

As a clean and least cost power source, the wind has already established its own position in the worldwide energy market firmly. Although initial capital cost is very high, in many countries wind energy is being connected to existing electric power grids along with traditional sources, as these initial costs balance out rapidly. All over the world, generation of electricity from the wind has steadily increased over the last few years and at the end of 2013 total installed wind capacity is 318 GW from 90 countries (GWEC, 2014). According to Global Wind Energy Organisation, in a moderate scenario it is estimated that, by 2020, total installed capacity of wind will reach 611 GW and by 2030, 964 GW. Moderate scenario assumes the modest situation, such as emissions reductions agreed by governments will be implemented, planned targets for uptake of wind energy is met in time etc. For the past four years, average global investment only in wind power equipment is nearly €50 billion and the annual investment in 2013 were €44 billion. Not only in the current state of affair but also in future wind energy sector also contribute to the employment rate of a country, as it creates a large number of skilled, semi-skilled and unskilled jobs. Assessment of employment suggests that 14 person/years of employment is created for one megawatt of newly installed capacity. Under a moderate scenario, GWEC assumes by 2015 employment level

created through manufacturing, component supply, wind farm development, construction, transportation, etc. will be 824,000 and will approach 1.1 million by 2020. Wind energy is known for its safe environment beneficial characteristics, and the early deployment of wind power means early reduction of carbon dioxide emission. Considering the regional context of wind energy sector, the “Make in India” campaign projects the replacement of services sector with manufacturing sector as the engine of India’s growth which clearly points out the need of massive amount of Energy to fuel this Indian vision. In India, first wind power development started in the year 1986 and total wind energy installation by 2014 were 21,693 MW providing almost 67% of renewable energy connected to grid. In a moderate scenario which expects an effective implementation of existing wind energy policies by the Ministry of New and Renewable Energy, India is expecting 49 GW of installed capacity by 2020 and 125 GW by 2030. In India by 2020 wind industry, anticipates a yearly investment of €6.6 billion, presumes 123,000 employment positions and reduces 82 million tonnes of CO₂ emissions. Not only in India, but many countries in all over the world could manage to operate with ambitious plans and good market prospects to expand wind energy development (IEA, 2012).

1.3.2 Issues related to wind energy production and distribution

From site optimisation to environmental strategies and planning, deployment of large scale wind power presents a broad range of challenges that must be addressed for a more cost efficient and reliable wind system. The International Energy Agency Wind (IEA Wind), coordinates member countries for a cooperative research on issues affecting wind energy and thereby benefiting the entire wind energy community. IEA Wind frequently lists the issues as the numbered R&D Tasks that are shared among members, usually country governments or international organisations, to participate in collaborative research activities. To direct the efforts on relevant research topics, in 2012 IEA wind reported a long-term R&D needs and categorised these requirements, based on the expected time for research results, as short-term (0–5 years), mid-term (5–10 years), or long-term (10–20 years) tasks (IEA, 2013). Exploiting opportunities in strategic areas like resource, design, operation, integration, and social and environmental impacts can significantly optimise the performance and cost associated with wind power operation. For a long term R&D, IEA Wind identified and documented four general research topic to be pursued by the international wind community, namely, characterising the wind resource, developing next generation wind power technology, wind integration and increasing the social acceptance of wind energy. The document is useful for wind community and research organisations in advanced wind energy technology by preparing their own research agendas.

Wind characterization address the research needs associated with site optimization, operation of wind turbine and power plant, wake loss and performance & output prediction. As wind energy production depends heavily on the site location, good site selection is critical concerning economic and technical feasibility of the project. For a reliable power production system, the decision of where to be located the wind power plant should address the effect of factors like landscape parameters, yearly changing weather patterns and wind characteristics. Geographic information systems (GIS) and Multi-criteria decision making (MCDM) techniques are the commonly used ones in determining the suitability of a particular area for a potential wind power plant. Reducing the performance uncertainties of wind power plants became more crucial as it is important to know rate of variation of produced power on different lead time hours. Since wind power production of a turbine is a direct function of wind speed, improve existing wind speed forecasting model is desirable to bring down energy cost. Technology related research activities

explore design and control component of a wind turbine along with reliability component of wind power plants. In on-shore and off-shore wind turbine design conditions are important for future innovation to reduce cost of energy. The high priority wind turbine characteristics that must be considered includes control system automation, behavior of the generator, hydrodynamic and aerodynamic load distribution on turbine blade, land and marine environment etc. Many software simulation tools, eg. Bladed, QBlade, are also available for assisting design phase of wind turbine in planned environment. Research in rotor architectures emphasis on models with a low cost, lighter, stiffer, and smarter blades aiming to operate through a wind speed of all range. Advanced wind turbine control technologies focus on laser-based radar (LIght Detection And Ranging, LIDAR) and acoustic radar (SOund Detection And Ranging, SODAR) to detect the wind speed and direction across the entire rotor disc of a wind turbine. This improved remote sensing measurements can be applied for evaluating wind flow models, creating a wind atlas, power performance verification and feed forward control of wind turbine. Reliability research concerns developing components with greater lifespan and ability to withstand failure conditions thereby reducing operation and maintenance cost. Because wind is highly variable and uncertain, generation and integration of high wind power has impacts which can be addressed by proper interconnection, planning, and market operations. Penetration of high wind power on grid may cause some operational problems such as insufficient transmission capacity and frequency stability, that reduce the economic value of wind energy system. Wind energy generated in a power plant is first transmitted to a small transmission lines which will transport power to a larger network transmission lines for travelling across a long distance then again transfer to a small distribution line to deliver electricity to the destination. Apart from good electricity storage technologies, proper transmission planning is essential to manage the increased variability and uncertainty associated with wind resource. Technical research related to the internal grid of the power plant and power electronic control are crucial for the reason that before feed into the power grid, electricity must be converted into the correct frequency and voltage. The advances in smart grid architecture include control of integration of heterogeneous systems and devices like distributed energy sources and storage. One of the most appealing benefit of smart grid technology deployment is enabling consumers as active participants in energy use. Consumers will be allowed to monitor the changes in grid conditions and act accordingly making efficient use the transmission system. Impression of non-grid-connected wind power suggest direct use of large-scale wind power output, without transmitting it to grid. Understanding and resolving issues of social acceptance of wind energy can be linked to three key dimensions called, socio-political acceptance, community acceptance and market acceptance. Socio-political acceptance refers to develop methods, tools and policies for acceptance by key stakeholders and policy makers. Generate insight into the acceptance at the local level in such a way as to maximize socio-economic benefit and minimize conflicts are dealt with community acceptance. Market acceptance assess cost driving components and devise the policies with which market parties adopt and support the energy innovations. Though as a renewable energy wind is considered as key solution to reach a sustainable future a moderate level of adverse impact on the environment is also noted. Policies to mitigate environmental impacts must reflect the factors including wind turbine recycling procedures and offshore power plant installation and operation. The overall mission of IEAWind long-term research plan is to encourage and support the policy and technological development to achieve a smart, reliable, cost-effective environment-friendly and market favourable wind power system.

1.3.3 Wind farm - operation and maintenance

In the beginning of the twentieth century itself, the wind had been experimented to utilise as an alternative energy source. After the 1973 oil crisis, fast but step by step technological advancement in wind energy drove it into new dimensions and re-emerged as one of the most important sustainable energy sources. While the first generation wind farms consisted of wind turbine capacity of 50 kW, over the years it has been increased up to an average capacity of 2.5-3 MW. Conversion of the kinetic energy in the wind into electrical energy is achieved by the aerodynamic modelling and the theory of electromagnetic induction. While the mathematical models to describe the aerodynamic forces helps in design processes like deciding the optimal height of the tower, control system design, shape and the number of blades, etc., properties of the electromagnetic induction is used to convert mechanical rotation into electric current. Maximising the aerodynamic efficiency is the primary objective of wind turbine design and according to Betz's law in an ideal wind condition with the help of an infinite number of blades, up to the 59.26% (0.5926 times) of the power available in the wind can be extracted (Singh et al., 2011). Although aerodynamic efficiency increases with the number of blades in a wind turbine, for an operational model, structural and economic consideration limits the blades to only two or three in number and still be able to extract 50% of the available power. Since Wind velocity is high at higher altitudes, desired wind turbine tower height is high as possible, and cost consideration restricts it two to three times the blade length. If the velocity of the wind is V_{wind} and air density is ρ then mathematically, power P_{wind} that can be extracted by a wind turbine with power coefficient C_P and swept area A , can be expressed as

$$P_{wind} = \frac{1}{2} C_P \rho A V_{wind}^3 \quad (1.3.1)$$

C_P represents the efficiency of a wind turbine and usually the values at various wind speeds is provided by the manufacturer. Aerodynamic force exerted on the blade is resolved into drag and lift, where the former component is the force parallel to the direction of flow of the air and the later is the force perpendicular to the direction of relative motion. While horizontal-axis wind turbine (HAWT) uses lift component to rotate the rotor and vertical-axis wind turbine (VAWT) extract power by the drag component. Even though both wind turbines are found in use lift turbine dominates the drag type as it is more efficient in extracting energy per square meter of the swept area. Another important design consideration is the measures to withstand enormous forces applied on the turbine blades by extreme winds or gust. Apart from the cut-out velocity above which turbine will go to a halt state, there are some control strategies for a range of high velocity wind before cut-out. Pitch-controlled wind turbine keeps a constant rotational rotor speed for a very high wind velocity until cut-out speed and hence retain a persistent power output. Another strategy is the stall-controlled turbines, which reduces the energy extraction for high wind speed range before extreme events. This control and safe operation of wind turbine is obtained by an anemometer which will send electronic signal to yaw motor for starting and stopping turbine in accordance with the cut-in and cut-out speed. Prior knowledge of the wind speed is desirable for an optimum performance of the wind turbine and hence to reduce the maintenance cost. It is reported that an efficient yaw control is possible even with a very-short-term wind speed and angle prediction and it is implementable on the real wind turbine hardware (Hure et al., 2015).

A brief explanation of how wind turbine works is shown in Figure. 1.2.

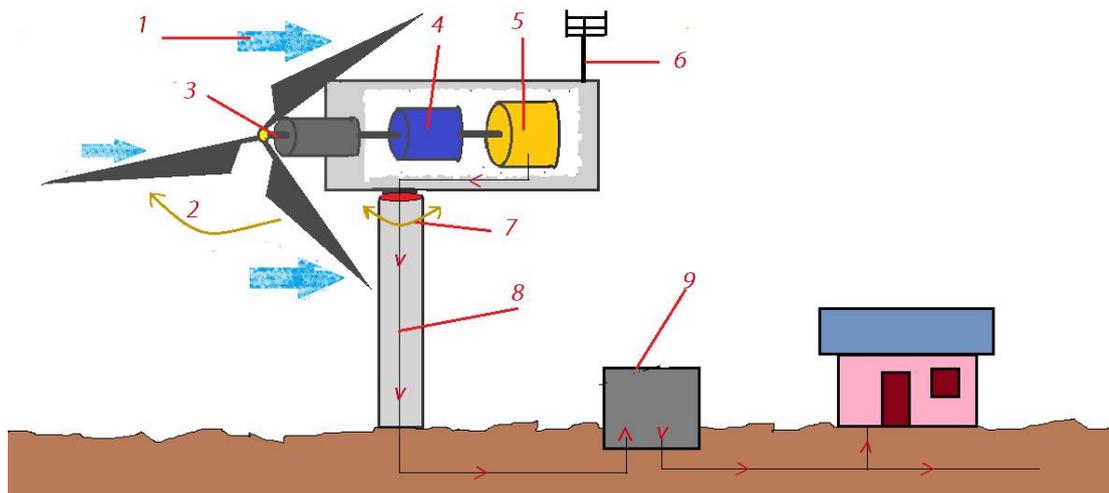


FIGURE 1.2: Schematic diagram of wind energy production and distribution.

(1) The Wind blows towards the turbine blade and (2) aerodynamic force causes the rotor blade to spin (3) and through the hub, energy is transferred to a rotating shaft to convert it into rotational energy. Nacelle is the central part of the turbine which constitutes (4) a gear box: increase the rotational speed of the shaft and (5) a generator: uses the increased rotational speed to produce electrical energy by the theory of electromagnetic induction. For the purpose of controlling the rotors and nacelle (6) anemometer and wind vanes are placed in the back of the nacelle to measure the wind speed and direction. These measurements are used by the (7) yaw motor to keep the rotor facing into the wind despite the changes in direction. Yaw motor also controls cut-in up and cut-out break of the wind turbine. Finally, the generated (8) electricity is transferred through the cables in wind turbine to a (9) substation. In order to minimise the turbulence between wind turbines in a wind farm, wind turbines are installed in a row at a sufficiently large distance, usually beyond seven rotor diameter (Vermeer et al., 2003; Meyers et al., 2012). A power collection system interconnects the energy generated from the single turbine in a wind farm and at a substation, this electric current is transformed into high voltage electric power for connection compatibility with electric power transmission network. Integrating highly variable wind power into the grid is associated with some operational and technical challenges which include power system operating cost, power quality, power imbalances, power system dynamics, and impacts on transmission planning (Georgilakis, 2008). Highly intermittent nature of the wind cause an obvious effect on the wind power production and in turn introduce additional cost overhead. Maintaining the balance between the supply of and demand for electricity in power system grid is one of the main concerns of utility operators. Retaining the grid stability is a highly sophisticated task and traditionally is achieved by utility operators with a great deal of experience accumulated over many years which is now being integrated with smart grid technologies. The services that facilitate and support the continuous flow of electricity so as to balance supply and demand in a stable grid is known as ancillary services and first introduced by the United States Federal Energy Regulatory Commission (FERC). Realising the significance of ancillary services in 2013 Central Electricity Regulatory Commission (CERC) of India has prepared a consultation paper on "Introduction of Ancillary Services in Indian Electricity Market" with the aim of supporting power quality, reliability, grid security and optimum resource utilisation (CERC, 2013). Prior knowledge of wind power can also help utility operators in grid planning at different time scales of interest such as unit-commitment: 1 day to 1 week with 1 h time increments, load-following: 1 h with 5–10 min increments (intra-hour), regulation horizon: 1 min–1 h with 1–5 s increments

(Georgilakis, 2008). Power imbalances on the grid due to variable wind is a major concern for transmission system operator (TSO) and is one of the ancillary services specified in the Open Access Transmission Tariff by FERC Order No. 888 (Wan et al., 2007). Many wind power prediction techniques suitable for different prediction horizon are available in literature and practice that can be utilised to assist ancillary services and thereby reducing the costs associated with maintaining grid stability (Wu et al., 2007; Soman et al., 2010; Kaldellis et al., 2011; Costa et al., 2008; Lei et al., 2009; Wang et al., 2012; Hering et al., 2010). Uneven wind fluctuations affect the electricity market in devising best bidding strategy with minimum possible risk. In electricity spot market, expected power production and market prices for a given time period are proposed by the participants and the electricity spot price settlement is done with the help of an auction system (Gomes et al., 2012). Most of the electricity markets plan midnight to midnight schedules one day ahead according to proposals of market participants and any marginal deviation from schedule will have an adverse impact on wind energy. FERC has devised some energy imbalance pricing rule to keep the system in balance as much as possible. The difference between predicted and actual generated wind energy is inevitable and the application of advanced wind energy forecasting techniques can help the utilities to frame a well-defined market bidding plan. Studies show that with the current forecasting techniques the impact of energy imbalance penalty on wind plant revenue is 2% and any improvement in wind forecasting accuracy can reduce this impact of energy deviation penalty. The intermittent nature and the uneven power production of wind power has a significant impact on grid connected power quality which describe the variability of the voltage level. Low-quality power may cause poor performance of equipment, such as flickering lights, unstable and disrupted power in user end and advanced power electronic systems are used to address these issues. Integration of wind power to traditional transmission lines is the another big challenge faced by many countries. In most areas of the world, large amounts of wind power are generated in remote lowly populated locations where the local consumption is low and have to transport wind power to high load locations where the population and energy consumptions are high. Traditional transmission grid in the remote area is not designed to transmit large amounts of power over long distances and the high-level exploitation of wind power set some problems on inter-area connections (Bindner et al., 2002; Glover et al., 2011). With deregulated power markets, the idea of expanding transmission grid does not seem to work and to avoid possible grid overload and unreliable services, utilities are forced to shut down wind parks leaving a potential amount of wind power untapped and causing huge economic loss (Piwko et al., 2005). Although many economic transmission planning studies such as grid energy storage (Eyer et al., 2010) have been conducted, these time, labour and cost intensive efforts may not result in much action (Georgilakis, 2008).

1.4 Wind speed modelling and forecasting

In the aim of getting a clear idea on chronology and evolution, methodologies underlying state-of-the-art wind speed forecasting models and their application to power systems operations have already reviewed and analysed by many researchers. One of the first largest review series of short-term prediction literature is prepared by ANEMOS (A NExt generation wind resource forecasting system for the large-scale integration of Onshore and off Shore wind farms), a research group supported by the European Commission (Giebel et al., 2003; Giebel et al., 2007; Kariniotakis et al., 2006; Landberg et al., 2003). A comprehensive report on reviews and recommendations of wind power forecasting systems and its integration in operational management tools is prepared by Argonne National Laboratory (Monteiro et al., 2009). Along with international use cases, hybrid

methods and its applicability in reducing average forecasting error are well discussed in Cutler et al., (2008). The investigation done by Costa et al., (2008) pointed out most significant proposals and developments of the wind power short-term prediction in the history of 30 years together with a list of unsolved or even unexploited topics. The Canadian Wind Energy Association (CanWEA) conducted a study on international experiences so as to understand the available techniques and its performance and cost effectiveness (Snodin, 2006). The business case for short-term wind forecasting and significance of setting the wind farm for good predictability is noted by Lei et al., (2009) and Lerner et al., (2009). Apart from all these some interesting books, book chapters and summary of the conference series on this topic are also found useful (Lange et al., 2006; Ernst, 2005; Lange et al., 2012; Fox, 2007; Giebel et al., 2008).

Wind speed forecasting is an interdisciplinary research area requiring skills and knowledge from other branches like meteorology, applied mathematics, artificial intelligence, energy, software engineering, information technology, *etc.*. High variability behaviour of wind happens on all time scales is one of the largest challenges of wind power. Forecasting highly variable wind speed some time ahead, from milliseconds to seconds useful for turbine control and from minutes to weeks is important for the integration of wind power in the electrical grid (Giebel et al., 2011). The following section briefly describes the available methods and techniques together with current research on wind speed forecasting.

1.4.1 A brief overview of forecasting models

In general methods regarding wind speed forecasting can be broken down into two types of models: physical models such as Numerical Weather Prediction model (NWP) and statistical models. The decision of including NWP model is highly depends on prediction horizon since it is reported that after 3-6 hours NWP outperforms all the other models. In most of the commercial use, Model Output Statistics (MOS) Method is used to combine statistical methods with a physical method to post-process the physical method output. Physical considerations of wind speed are used to obtain the best possible estimate, and statistical techniques are applied to reduce the remaining error. Wind speed forecasting models are evaluated based on the error measures that may quantify the difference between estimates of modelled outcome and the actual measurements. Some times statistical association based methods such as Pearson's correlation coefficient (γ) are useful as it gives the statistical co-variation between the actual and predicted. The oldest statistic to measure model accuracy is the mean difference (ME) and Landberg et al., (1994) pointed out the shortcoming of this as a low mean error can also be occurred by averaging negative and positive errors. Other common statistical evaluation criteria are root mean square error (RMSE) and mean absolute error (MAE) of forecasts and is proposed as good estimates by Madsen et al., (2005). Normalised error measures like NRMSE (Normalised RMSE), NMAE (Normalised Mean Absolute Error) is useful for a comparison across different wind farms with different variable time series (Madsen et al., 2005). The error criterion for optimal prediction parameters are characterised by Nielsen et al., (2003) as "The predicted value of the wind power production should be close to the average of the real values. The sum of deviations between the predicted value and real values should be small. The prediction should result in a low cost of the consequences of prediction errors". Analysing the nature of error growth in terms of the spectrum of frequencies in the measurements gives an insight of the suitability methods with reduced error (Vincent et al., 2009). Most frequently when assessing the forecasting performance of a proposed model persistence model, which assumes the future value as the last observed one, is considered as the baseline model. According to many authors the new reference method (Nielsen et al., 1998), a

simple modified version of persistence method with a trend towards the mean of the time series is capable to achieve 10% RMS error improvements over persistence.

1.4.2 Meteorological modelling

Many meteorological organisations have developed mathematically complex and computer intensive numerical weather models in the aim of reproducing the general synoptic characteristics of the weather over large areas and there by predicting its future values. The Weather Research and Forecasting (WRF) model is developed by US meteorological departments National Oceanic and Atmospheric Administration (NOAA) and the National Center for Atmospheric Research (NCAR) and is used to produce forecasts with a 10 km horizontal resolution and a vertical resolution of 80 m. With additional input, cases with specific resolution is also possible in WRF. One such example is the consideration of planetary boundary layer (PBL) which extends from the ground to the bottom of where cumulus clouds form. Its boundary keeps varying according to atmospheric conditions and hence the physical quantities such as flow velocity, temperature, moisture, etc. in this layer is highly variable. To accommodate all the factors interacting with PBL we need a set of physical equations and unfortunately, number equations are fewer than the number of unknowns. Meteorologists addressed these issues with several PBL schemes based on various assumptions about the PBL and surface interaction and ran these different PBL schemes. The WRF model depends on two global meteorological models called, Global Forecast System (GFS) model and North American Mesoscale (NAM) model, for obtaining a set of boundary and initial conditions that must be given to each PBL scheme. The WRF model accepts input from both GFS and NAM model and runs each PBL scheme forming an ensemble of forecasts, which will be then averaged to obtain forecast at a given time. The United Kingdom and Canada also have a well-known atmospheric model designed for their own forecasters and researchers. The model developed by the United Kingdom is known as United Kingdom Meteorological Office (UKMO) model or Met Office Unified model, and Model from Canadian Meteorological Centre (CMC) is Global Environmental Multiscale Model (GEM). GEM is a global model with a resolution of 100km, prediction horizon of 10 days and has a 6 hour cycle. GEM has also got a regional model, which runs 2 times a day with 48hrs ahead prediction and a horizontal resolution of 15km. The unified model of UK runs 2 times in a day with a model duration of 70 mins and prediction horizon of 5 days (Snodin, 2006). NWP ensemble forecasting techniques is an important research area since it tries to confront the chaotic nature of the weather and its effect in physical models. An NWP model is run more than fifty times with several input conditions and the output is compared to check the forecast confidence level. As with all developments, it is noted that repeating NWP many times is a heavy computational process with a high cost. Input conditions for NWP models are usually gathered from a broad range of sources such as meteorological stations, ad-hoc report from aircraft, maritime traffic and satellite-based observations. Although prognostic modelling of the dynamics of atmospheric global circulation and forecasting current state of the atmosphere by system assimilation is efficient in theory, this form of modelling is extremely computationally intensive and case specific which may result in unreliable output in a complex geographic area. To realize the transformation between NWP and site-specific ones, both statistical and physical methods are found in academic research and commercial practice. Statistical methods assumes a systematic difference between a particular site and NWP model and is recognized and eliminated by a multiple regression model with combination of meteorological parameters. Physical model approach aim to increase the resolution of original NWP model with

the aid of a local version incorporating local thermal effects. Another practical difficulty of meteorological model is the exceeding time it takes to integrate and run the system of model equations of atmospheric circulation reproduction. Mostly there is little value for NWP the prediction as by the time forecast is delivered prediction horizon becomes history.

1.4.3 Time series modelling for wind speed forecasting

Observation of state of a system made over time is usually termed as time series and a systematic approach of time series data analysis extracts the characteristics and statistic of the system. Methodologies developed for time series analysis is suitable for historical measurements of any natural system and hence fit to derive a good description of highly intermittent wind speed data. The use of wind speed predictions as an explanatory variable of direct wind power prediction is important and the methods suitable for one can be used for the other also. Corresponding to the prediction horizon wind forecasting can be broadly classified as, very short-term forecasting: From few seconds to 30 minutes ahead, short-term forecasting: From 30 minutes to 6 hours ahead, medium-term forecasting: From 6 hours to 1 day ahead, and long-term forecasting: From 1 day to 1 week ahead. For developing a predictive model of a very short-term to short-term forecast length, time series models are suitable (Box et al., 2013). The capability of Gaussian distribution in explaining random variations in the wind over time is explained by Brown et al., (1984), one of the first papers in this field. Along with transformation to a Gaussian distribution and AR (AutoRegressive) prediction, removal of seasonal and diurnal swings in the AR components and prediction interval is also discussed. Persistence method which assumes $x(t+n) = x(t)$ is widely recognized as the one suitable for short-term prediction method for practical purpose (Nogaret et al., 1994). Use of Kalman Filter for wind speed prediction and its comparison with persistence method shows, for 1-min, averaged data with six previous measurements, prediction accuracy in RMSE can be improved up to 10% for the next time step prediction (Bossanyi, 1985). A systematic analysis of de-trended wind speed data with iterative Box-Jenkins method and central moving average smoothing can achieve 2 hours ahead prediction with reduced RMSE (Fellows et al., 1990). Wavelet decomposition method has also experimented for exploring its capability on modelling and prediction of Wind time series measured at different locations. Wavelet decomposition aid in identifying the wavelet components from the measured time series to form a multivariate time series matrix. Principal Component Analysis (PCA) on this matrix may provide a reliable estimate compared to simple linear regression (Hunt et al., 2001). Usefulness of stochastic models like ARMA and importance of site and month-specific parameters on ARMA models has also been studied and reportedly outperforms persistence method for 1-hour forecast (Torres et al., 2005; Tantareanu, 1992). Assuming wind speed as a stochastic process after removing annual and daily periodicities (Balouktsis et al., 1986) modelled measured data with ARMA and found the Markov Transition Matrices and the coefficients of the ARMA models are identical independent of location. Other reasonably accurate models of stochastic simulation which incorporates autocorrelation, non-Gaussian distribution and diurnal nonstationarity with ARMA is also found in literature (Daniel et al., 1991). The performance of different ARMA models on same set of wind speed data is studied by Schwartz et al., (2002) and concluded that training period highly affects the performance of the model. Investigations on the appropriateness of f-ARIMA (fractional-ARIMA) models for wind speed modelling and prediction manifested a more accurate day ahead prediction compared with persistence method (Kavasseri et al., 2009). A Bayesian framework based modelling of wind speed as an AR process with Markov Chain

Monte Carlo (MCMC) simulated model parameters has also been developed and reported a comparable accuracy with persistence method (Miranda et al., 2006). Markov-switching AR model with time-varying coefficients for modelling and prediction of wind speed is discussed in Pinson et al., (2008). Autoregressive integrated moving average with proper (p,d,q) values can give the best performance in modelling short-term behaviour of wind speed and hence accurate prediction (Palomares-Salas et al., 2009).

1.4.4 Machine learning techniques and hybrid models

Apart from traditional statistical models, Data mining models: extracting knowledge and features from voluminous of data, and Hybrid models: which combines different modeling approaches, are also found in use to solve modeling and prediction of wind speed data. One of the important data mining models is the Artificial Neural Networks (ANN). Dr. Robert Hecht-Nielsen, defines a neural network as "a computing system made up of a number of simple, highly interconnected processing elements, which process information by their dynamic state response to external inputs" (Caudill, 1987). ANN does not depend on any predetermined mathematical model but it is designed to learn from its own learning experience and gets better on each trial. Physical structure of an ANN can be described with two basic components called, processing elements or neurons and connections or link between neurons. These interconnected neurons are organized in layers called input, hidden and output layer. Based on how these neurons are connected a variety of neural network structures, such as Multilayer Perceptrons (MLP), Back Propagation (BP), Radial Basis Function Networks (RBF) etc., developed for numerous potentially important applications can be found. One of the first report on use of ANN with Radial Basis Function (RBF) in predicting wind speed was prepared by Beyer et al., (1994) and they found 10% over persistence on RMSE for one step prediction of either 1-min or 10-min averages. Application of ANN in highly accurate estimation of the wind speed at different heights and terrains is reported (Bechrakis et al., 1998). Possibility of a wind power forecasting model based on recurrent high order neural networks and its implementation into an advanced control system was experimented and presented by Kariniotakis et al., (1996). A widely accepted ANN tool for hourly short-term wind power forecasting upto 3 hours with 5-minute intervals is the Artificial Neural Network Short-Term Load Forecaster(ANNSTLF), developed by the US Electric Power Research Institute . Forecasting of strong winds and gusts with neural network classification was experimented with data measured at Geneva and Sion in Switzerland and concluded that comparing to persistence its performance is improved 1, 6, 12 and 24 hour horizons (Kretzschmar et al., 2004). A comparison of statistical approaches like ARIMA, Moving Averages with Multi-Layered Perceptron neural network approach and reported significantly improved accuracy for ANN over statistical models (Campbell et al., 2005). Neural network model training in an auto-regressive manner using back-propagation and cascade correlation algorithms, on wind speed data measured at two locations give a satisfactory prediction (More et al., 2003). Comparison of the performances of three ANN models, adaptive linear element, back propagation, and radial basis function, and how each model respond to different inputs are investigated by Sfetsos, (2000). Performance evaluation of each model on different metric pointed out a need of post-processing of outputs by combining forecasts from different ANN models. Use of grey predictor GM(1,1), another important data mining model, on wind speed data also gives a reasonable prediction accuracy (El-Fouly et al., 2006). Support Vector Machine (SVM) is another data mining method which extract classification and regression rule from the measured time series data in the aim of modeling and prediction. Support Vector Regression (SVR) construct a model by proper mapping of

data into higher dimensional feature space and utilize this model for predicting the futures values. Ability of SVR in generating an efficient model for wind speed dynamics and its prediction accuracy comparison with ANN- MLP is investigated by Mohandes et al., (2004) and reportedly could produce a favorable output again MLP. Least Square Support Vector Machine (LS-SVM) model, a modified and much simpler version of SVM, is also has been used for improving the prediction accuracy of wind speed data. comparison of back propagation neural network and LSSVM models in wind speed prediction is done by Du et al., (2008) and declared a better performance and accuracy of LSSVM models. Fine tuning LSSVM with the parameters like type of kernel function, the size of training sample, and the settings of parameters aid LS-SVM to outperform the traditional persistence model (Zhou et al., 2011). Fuzzy techniques for time series prediction is also used for modeling and prediction of wind speed variations (Zhu et al., 2012; Damousis et al., 2004). Modelling and prediction with fuzzy rules identifies a fuzzy pattern which specifies a membership degree between 0 and 1 for each data sample. Basic fuzzy models are classified in three: a regression model based analysis, Box-Jenkins model based analysis, a fuzzy reasoning (IF-THEN rule) based analysis. While regression model identifies the fuzzy regression coefficients, Box-Jenkins model uses autocorrelation, autoregression and ARIMA methods, and fuzzy reasoning identifies the relation between time series data using IF-THEN rule. from wind farm history data, powerful Fuzzy models can be derived for prediction purpose and these models are reportedly capable of maintaining a good prediction accuracy. In the aim of increasing wind speed prediction accuracy of time series models researchers experimented some hybrid models, which assumes combinational effect of different techniques. Along with a range of methods Dutton et al., (1999) evaluated a linear autoregressive model combined with adaptive fuzzy logic and could give only a minor improvement over persistence. Sfetsos, in his work (Sfetsos, 2000; Sfetsos, 2002) suggested an adaptive network based Fuzzy inference system (). He also compared linear and non-linear models and concluded that nonlinear models exhibit better performance. Appropriateness of ANFIS model in different geographical location was investigated by Potter et al., (2006) and they measured 30% reduction in mean absolute percentage error (MAPE) over persistence. Model integrating ARIMA and ANN is evaluated on ANN and ARIMA as a separate entity. Forecasting is done in three different regions of Mexico and two error metric, mean square error (MSE) and the mean absolute error (MAE), demonstrated the dominance of hybrid model (Cadenas et al., 2010). Using some statistical measure for seasonal exponential adjustment in wind time series and combining this with back-propagation (BP) neural network model gives a lower mean absolute error than BP alone (Guo et al., 2011). Short-term pattern extraction with Artificial neural network (ANN) and long term pattern extraction with Markov chain (MC) are combined by Kani et al., (2011) to develop a new ANN–MC model, with reduced calculation time and increased prediction accuracy, for a very short-term prediction. Combination Ensemble Empirical Mode Decomposition (EEMD) and the Support Vector Machine (SVM) may decompose the measured time series into several components and applies SVM for each component. These respective estimates are reversed into original form and an observable improvement in prediction accuracy is reported (Hu et al., 2013). Comparison of two hybrid methods Wavelet-LSSVM and Wavelet-NN on different time horizon is performed and concluded as, much easier and simpler Wavelet-LSSVM method performs well than Wavelet-NN (De Giorgi et al., 2014). In order to address nonlinearity and uncertainty in wind speed variations artificial neural network (ANN) and Kalman filter (KF) combination can be used and its forecast comparison with ARIMA model account a better accuracy (Shukur et al., 2015). In Zhang et al., (2012), authors analyzed wind speed with the help of three groups of LS-SVM forecasting models and Fuzzy logic. Univariate, hybrid with ARIMA and multivariate LS-SVM model out is aggregated and defuzzified using an intelligent-agent-based fuzzy group forecasting model.

Apart from these numerous research models, some real operational models are also of great importance. Prediktor (Troen et al., 1990; Landberg et al., 1994) is a short-term wind power prediction tool developed by Risø National Laboratory, Denmark and is known as the father of the modern NWP based forecast. Complete idea of Prediktor is to obtain wind speed and direction, address the local effect and determine the power curve. Simulation of large scale flow is obtained with a numerical weather prediction technique called High Resolution Local Area Modeling (HIRLAM) and the accommodation of local conditions are achieved by an Wind Atlas Analysis and Application Program (WASP) (Troen et al., 1989) tool. An MOS module with at least four month wind history is also present to fine-tune the general errors that may occur in any phase of Prediktor. The Wind Power Prediction Tool (WPPT) is the tool widely used in Denmark since 1994 and has been developed by the Institute for Informatics and Mathematical Modeling (IMM) of the Technical University. WPPT developed in the aim of predicting the wind power production in larger areas by applying statistical methods. Although the tool was designed for multi-step ahead prediction upto 36 hrs its lack of quality limited its accurate prediction horizon only up to 12 hours ahead. Later version integrated HIRLAM forecasts with WPPT and could obtain useful forecasts upto 39 hours ahead (Nielsen et al., 1999). WPPT can incorporate different data sources such as on-line and off-line power production measurements and its aggregations, Multiple NWP forecast providers etc. The model structure will get complex with available data and the area of consideration. Risø National Laboratory and IMM jointly developed a highly flexible software package called Zephyr that combines Prediktor and WPPT to ensure a more accurate forecasts for all prediction horizon (Giebel et al., 2001). Previento is the forecast system developed at Oldenburg University, Germany, in the year 2001 to provide power predictions for a larger area such as covering a whole country (Focken et al., 2001). Its approach is similar to Prediktor and uses data of the German Weather Service instead of HIRLAM. The model is capable of monitoring the weather situation and providing an estimation measure of possible error. eWind is another popular model developed by True Wind Solutions, USA focusing to run in higher resolution. The system has four basic components, 1) a three-dimensional physics-based NWP model: provides a high-resolution simulation and forecast of wind, 2) adaptive statistical model: to adjust the output of physics-based NWP model, 3) plant/wind-turbine output model: defines a fixed or variable (in accordance with the recent atmospheric data) relationship between the wind turbine and the atmospheric variables. 4) forecast delivery system: sending the forecast information to users through any of the option available (Bailey et al., 1999). SIPREOLICO is a prediction tool developed by the University Carlos III, Madrid, Spain, and the transmission system operator Red Electrica de Espana (REE). The tool is based on statistical techniques on online measurements and is capable of providing prediction on every hour upto 36 hrs. The types of inputs SIPREOLICO accepts are HIRLAM output, environmental characteristics of a particular site, both wind speed and power history and online power measurements. Performance of the tool in predicting the wind power for the whole of Spain is reported as satisfactory (Gonzalez et al., 2004; Sánchez et al., 2002). The model developed by University College Cork, Ireland in 2004 is known as HONEYMOON and it uses an ensemble of NWP in order to account the uncertainty associated with wind speed (Möhrlen et al., 2005; Jørgensen et al., 2005; Lang et al., 2002; Lang et al., 2006a; Lang et al., 2006b). A high resolution physical model was developed by TrueWind known as WEFruc – Wind Energy Forecast Rapid Update Cycle and is able to use remote sensing atmospheric data as input (Zack, 2004). LocalPred and RegioPred are the famous tools used in Europe and are capable to fine tune NWP output in accordance with out of other modules like timeseries model, power curve model *etc.* (Martí Perez, 2002). Apart from all these many other operational models emphasizing specific market and the distribution of wind turbines are also found in many countries.

Although many time series analysis and meteorological modelling techniques are available to abstract the dynamics of wind speed, none of these forecasting methods is capable of significantly reducing the prediction error compared to the elementary method of persistence (Sfetsos, 2002). The high fluctuations and variability in wind speed depicts the complex dynamics of the system that attributes the prediction error and therefore it is important to explore the sources of these random like fluctuations. Despite the fact that understanding the nature of dynamics whether it is stochastic or deterministic may help in improving the tools used for prediction, only a few studies have focused on this direction. The centre of interest of most these studies was only to compare stochastic versus deterministic time series models not on exploring the nature of wind speed dynamics. Studies on several time series of wind components and X-band Doppler radar signals measured over area of the ocean surface have found the presence of a low-dimensional dynamical attractor in vertically polarized radar reflectivity and the horizontal surface wind speed (Palmer et al., 1995). They also reported the suitability of deterministic model to achieve a higher short-term prediction correlation coefficient for winds from radar reflectivity. Ragwitz et al., (2007) used a locally low-dimensional prediction scheme to analyse local surface wind velocities for detecting the deterministic structure and reported determinism only in increase of the velocity. As per their result, although prediction error can not be reduced by using a non-linear model instead of a linear stochastic one, a significantly high accurate prediction of intermittent gusts is attainable. There is always a room for new techniques for gaining information and solving this complex system.

Though the idea of the unpredictable behaviour of deterministic systems was introduced in the late 1800s by the French mathematician and theoretical physicist Jules Henri Poincaré, chaos theory became formalised only with the invention of the high-speed computer in the 1950s. Repeated iterations of mathematical formulas with the help of high-precision electronic computers led to the important implications of non-linear systems and the inherently unpredictable nature of dynamical system was first reported by Edward Lorenz as part of his weather simulation study (Lorenz, 1963). Lorenz noted his equation never settled down to equilibrium and moreover very small difference in the current state of the system can cause exponential divergence. Even though the system continued to oscillate in an irregular, aperiodic fashion Lorenz could depict the butterfly-shaped structure of the chaos by means of a three-dimensional view of a set of solutions.

1.5 Nonlinear time series analysis and prediction

The evolution of almost all the natural systems with respect to time are often non-linear and coupled with a large number of variables and parameters. The huge number of parameters make the system very much sensitive which leads to an irregular, unpredictable behaviour. Although formulating a closed solution is difficult, in mathematics many tools and techniques are available for dealing with such a highly uncertain non-linear dynamical systems. Time series analysis is one such technique based on evolution rule, which says the state of a dynamical system follows the past state and in some case randomly occurring events also affects. The state of a system as it evolves over time is recorded as a sequence of data points to form a time series and each data point composed of a set of attributes of the system modelling its behaviour. Time series data can be discrete, in which time is a discrete variable, or continuous where time is a continuous variable. More specifically, discrete time series measures values of variables at distinct points in time and this time series can be represented as $[x(t) : t = 1, 2, 3, \dots, N]$ where t is the time at which state $x(t)$ is measured, whereas in continuous time series data there will be many time points between

any two points in time and it is represented as $[x(t)]$ measured over some time interval $T_0[0, 1]$. For a continuous time system to model the underlying dynamics, it will be more feasible to measure the process an equidistant discrete sequence of observation times. As pointed out by Jones, (1980) even though the state is measured at a discrete interval, this can be handled very conveniently by continuous time series models. Time series analysis techniques are also classified into univariate, multivariate, linear and non-linear. Univariate time series observes the state of only one variable while for multivariate time series simultaneous measurement of more than one variable is involved. The principal objective of time series analysis is to obtain descriptive measures of the system, with these measures analyse and forecast the future values and examine the possibilities of useful control measures. Linear techniques in time series analysis use the concept of linear relation on previously measured data points, and for systems showing dynamic behaviour we assume a non-linear relationship and go for non-linear techniques. Because of its natural temporal ordering and accountability of possible internal structure in the data, non-linear time series based models have many applications in many different areas including numerous dynamical phenomena in nature.

For any dynamical system we usually measure the state of the system $x(t)$ indirectly and for obvious reasons, for complex systems, it is difficult to understand the dynamics from this single time series. In such case, underlying behaviour of the system can be explored by analysing structural characteristics of its time series measurement. Proper reconstruction of the structural changes in dynamics and characterization, from an observed time series $x(t)$ is the very basic step adopted for any non-linear time series analysis (Kaplan et al., 1995). For reconstructing a shadow manifold M' from a single time series $x(t)$ with manifold M , Floris Taken's (Takens, 1981) used time lagged observation and presented the first mathematical proof for this. A non-linear dynamical system with imbalanced but highly complicated steady state behaviour and periodic or almost periodic oscillation is referred as the chaotic system. Since it comes under non-linear systems, its behaviour which is very sensitive to the initial condition can effectively be addressed by the non-linear techniques. For many of the natural phenomena that are chaotic in nature and equation of motion is not known, characteristics of the underlying dynamics, such as dimensions of attractors, entropy, Lyapunov exponents etc. can be effectively obtained once we reconstruct the shadow manifold from the measured time series. Although this reconstructed manifold may not preserve the exact geometric shape, reconstruction retains the characteristics of dynamics hidden in original time series.

1.5.1 Phase space methods

For any dynamical system, states of the system evolve through time are represented by phase space and each point represents the unique state. Phase space is also referred as state space, as the state of the dynamical system usually measured at discrete interval. It is considered as an alternate way to view the system activity where a graphical point in phase space is represented by a number of system variable. Considering a purely deterministic system once its present state is fixed one can determine its future state and establish the state of the system in a phase space or state space. Once the states are specified as points in vector space, the dynamics of the corresponding deterministic system can be studied by studying the dynamics of the corresponding phase space vector points. In theory, the first-order ordinary differential equation can be used to define a deterministic dynamical system and it ensures the existence and uniqueness of the trajectories for certain conditions. A non-deterministic dynamical system is described as the systems with an infinite number of states and the transition of the state can be governed by some rule. Markov

processes are the instance of these kinds in which transition probabilities are used as rules to randomly select the future state. These concepts of infinite states and transition probability limits its memory, in the sense future state may only depend on the current state not on the past. In a Markov chain of order m , the present state is represented by the values of the process during the last m discrete time steps. One can regard a purely deterministic system as a limiting case of Markov Process where the transition is governed by a deterministic rule occurs with probability 1 and all the others with probability 0.

For a deterministic system with finite dimensional vector space \mathbb{R}^m the state of the system at any time can be specified by $x \in \mathbb{R}^m$ and the nature of underlying dynamics by an m dimensional map where time is a discrete variable as

$$x_{n+1} = F(x_n), \quad n \in \mathbb{Z} \quad (1.5.1)$$

or by m first-order ordinary differential equations where time is considered as a continuous one.

$$\frac{d}{dt}x(t) = f(x(t)), \quad t \in \mathbb{R} \quad (1.5.2)$$

Since the system is of manifold M and continuously differentiable with respect to time Eq.1.5.2 is known as flow. Since the right-hand side of Eq.1.5.2 has no dependency on time it is also referred as autonomous. With x_0 or $x(0)$ as initial condition solution of Eq.1.5.2 gives a series of points x_t or $x(t)$ known as the trajectory of dynamical system. This evolution with respect to time approaches either infinity or stay in a bounded area forever similar to a chaotic system. For a bounded dissipative dynamical system, the evolution of a wide variety of initial conditions can be pulled to some subset of phase space which is commonly known as the attractor of the system. Or in other words "An attractor is a set of states (points in the phase space), invariant under the dynamics, towards which neighbouring states in a given basin of attraction asymptotically approach in the course of dynamic evolution" (*Attractor, From MathWorld—A Wolfram Web Resource*). The set of points in phase space that lie on a particular attractor only for a short time is called the basin of attraction of the attractor. Simple examples of attractor are fixed point attractor and limit cycles. In the case of a simple pendulum, the state of the system may evolve finally into the centre bottom position which will be the single stable point for any initial condition and is called a point attractor.

Limit cycle forms an isolated closed trajectory in phase space and all the neighbouring trajectories evolve either towards or away from the closed loop. As time evolve if the state of the system converges towards limit cycle then it is known as stable limit cycle and if the trajectory is moving away from the closed orbits then it is known as unstable limit cycles.

In a non-linear system, the behaviour of the system is described by a non-linear system of equations, in which a non-linear combination of unknown variables or functions such as variables with square or higher power, some sort of threshold function *etc.* are present. In other words, most of the non-linear time series methods are based on the theory of time evolution of the dynamical system which can exhibit deterministic chaos. A non-linear dynamic system may exhibit multiple attractors based on the initial conditions and the state of the system may evolve into basins of attraction of different attractors (Ruelle et al., 1971; Grebogi et al., 1997). The path traced by

an initial condition forms a pattern in phase space known as strange attractors where the nearby states on the attractor at one time diverge from each other exponentially at later times. Apceptly, as the dynamical and geometrical properties of a deterministic system can be well described in the phase space, equations of motions can be approximated by phase space description. Such approximations can be used to understand the geometry of the attractor thereby a deeper understanding of the nature of dynamics. For experimental and naturally occurring chaotic dynamical systems, the equation of motion is usually not known and is only possible to obtain a sequence of scalar measurements or time series. Proper phase space reconstruction can be achieved by embedding the scalar time series measurements into higher dimension after finding the good embedding parameters such as embedding dimension m and time lag τ . The classical embedding theorems promise with appropriate state variables and their respective delay for ideal noise-free data the reconstructed vectors s_n from a measured time series s_n are equivalent to phase space vectors and preserve the characteristics of the attractor. Poincaré surface of sections is another phase space method turns the continuous time flow into a discrete-time map with the help of a section and map in phase space. Poincaré section and Poincaré map are used to visualize the flow in phase space. While Poincaré section identifies a suitable surface in phase space which is crossed by almost every trajectory and the Poincaré map identifies the points where the trajectory intersects the Poincaré surface in a specified direction. Extracting the information about the characteristics of the dynamical system from a measured time series is much simpler if there are some redundancies inside the data. A deterministic system can have a recurrent behaviour in its original state space such as a simple pendulum having a single point attractor which is trivially recurrent for all times. With suitable embedding, a typical non-linear dynamic system exhibits a repetition of trajectory in approximately the same area in phases space known as recurrence and the recurrence plots visualise times at which a state of a dynamical system recurs. The idea of recurrence plot has been introduced by Eckmann et al., (1987) and helps in visualising m dimensional phase space trajectory through a two-dimensional representation of its recurrences. Visual inspection of recurrence plot gives an insight on time series data and embedding space.

1.5.2 Quantitative measure of complexity

Even though the underlying dynamics of the non-linear system is low dimensional and deterministic, as it evolve over time, predictability of the future state becomes limited. However, unpredictability never means the absence of order but occurs on account of the exponential separation of initially nearby trajectories. Chaotic systems are very much sensitive to the initial conditions and even a very small deviation can cause an exponential decay of correlation functions. As the system is predominantly periodic and divergence is exponentially fast, quantification of complexity is possible with the proper averaging of growth exponent. This averaged exponent is known as *Lyapunov exponent* and is denoted by the symbol λ . Consider two points in state space s_{n1} and s_{n2} with distance $\|s_{n1} - s_{n2}\| = \delta_0 \ll 1$, then after Δn time the distance $\delta_{\Delta n}$ between the two trajectories is $\delta_{\Delta n} = \|s_{n1+\Delta n} - s_{n2+\Delta n}\|$. Then the λ can be calculated from the equation

$$\delta_{\Delta n} \simeq \delta_0 e^{\lambda \Delta n}, \quad \delta_{\Delta n} \ll 1, \quad \Delta n \gg 1 \quad (1.5.3)$$

A positive value of λ indicates, as the system evolves over time, nearby trajectories have diverged exponentially - signalling the evidence of chaos. The range is $0 < \lambda < \infty$.

For some dynamical system, the attractor is a stable fixed point and the trajectories approach each other exponentially very fast resulting in a negative Lyapunov exponent. That means $\lambda < 0$ for a stable fixed point. If the attractor is a stable limit cycle, any two trajectories originate from the neighbouring point can approach each other very slower than exponentially and in such case Lyapunov exponent is negative or $\lambda < 0$. If the system is purely random or noise process, then $\lambda = \infty$.

1.5.3 Estimation of topological invariants

The evolution of a dynamical system over time can be well described with the quantification of possibly complicated structures that are hidden in the measured time series. The essential condition to convince the theory of patterns in the measured data is the significant improvement in prediction accuracy, and several non-linear prediction algorithms are available for this purpose. Although predictability is not the sufficient condition to assert the concealed structure in the signal are real not just fluctuations, the statistically significant and better predictability of non-linear methods compared to other techniques can be considered as clear evidence for the non-linear and deterministic structure in it. For a simple deterministic dynamical system represented by Eq.1.5.1 and Eq.1.5.2, with the knowledge of present state(at some time n or t) all the future states can be described and definitely there exist some deterministic forecasting function. But in a real sense, measuring a physical quantity with a hundred percent accuracy is never possible, and the inaccuracy of the present state will be amplified as the system evolve. For a chaotic system even though a very small error is amplified exponentially, the rate of divergence is finite and can have hope for reasonable short-term forecast if we already know the mapping function F . Unfortunately, in real world system, F is unknown and we may go for assumptions about its properties. The very simplest prediction algorithm with the minimal assumption, like F is continuous, searches all past states for the state closest to the current and its future is taken as the forecast value. In other words, for finding x_{N+1} , given the present one x_N , from a list of all past states x_n with $n < N$ find out x_{n0} which is very close to x_N with respect to some norm, then x_{N+1} is also very close to x_{n0+1} and can be taken as the future state. This simplest method is known as ‘‘Lorenz’s method of analogues’’ (Lorenz, 1969). For most of the dynamical systems are measured indirectly we may only have scalar measurements

$$s_n = s(x_n), \quad n = 1, \dots, N \quad (1.5.4)$$

and s is the unknown measurement function. In such situation equivalent state space vectors s_n can be obtained with *phase space reconstruction* method. For predicting the scalar measurement $s_{N+\Delta n}$, use $s_{n0+\Delta n}$ as the forecast where s_{n0} is the closest to s_N in embedding space (Kennel et al., 1992). In this assumption we just ignore the fact that the measurement is valid for a finite resolution and is approximated using a discrete quantity. It implies considering only one closest state in the past is not at all a better solution. For a typical finite resolution size σ consider all the neighbouring points within the radius of σ in phase space and the average of individual predictions may give better predictions (Kantz et al., 2003). For a finite resolution parameter ε , Δn ahead prediction of s_N is done from its neighbourhood $U_\varepsilon(s_N)$ of radius ε . The arithmetic mean of the individual predictions $s_{n+\Delta n}$ of all points $s_n \in U_\varepsilon(s_N)$ is the prediction.

$$\hat{s}_{N+\Delta n} = \frac{1}{|U_\varepsilon(s_N)|} \sum_{s_n \in U_\varepsilon(s_N)} s_{n+\Delta n} \quad (1.5.5)$$

where $|U_\varepsilon(s_N)|$ is the number of elements in the neighbourhood space of radius ε distance. *Zeroth* method is the simplest method of prediction of this kind and is well described in (Kantz et al., 2003).

One of the most striking features of the chaotic attractor is the fractals which refer to the presence of repeating patterns in every scale. This property is usually termed as self-similarity. For non-fractal objects like points, lines and space have integer dimensions, 1, 2 and 3 respectively, and for fractals a change in length is always scaled by a fractional power. Many measures like box counting dimension, entropy dimension, correlation dimension *etc.* are available to quantify the complexity of fractals and correlation dimension is the one used in our study. Calculation of the correlation sum $C(\varepsilon)$ is first step for finding the correlation dimension and it uses counts the number of points in vector space that are separated by a distance less than it measure the mean probability of closeness of two points in some vector space at two different times and ε .

As the number of data points tends to infinity, for small ε we expect zero distance and can be explained by the power law, $C(\varepsilon) \propto \varepsilon^D$ where D is the correlation dimension. Correlation sum from a time series can be calculated with the reconstructed phase space vectors and here the choice of delay embedding, τ , is very much important as the inappropriate choice of τ may result in a poor performance of correlation algorithm. Once the phase space s_n is reconstructed calculate correlation sum $C(m, \varepsilon)$ for a range of dimension m , look for the enough signs of self-similarity and if it is present compute correlation dimension. Results of both correlation sum and correlation dimension can mislead interpret if proper care is not taken. Following the flow of a deterministic system, we can see temporal correlations because consecutive states of deterministic systems stay close as system evolve. In fact, the property of closeness of data in time and space is not only true for the deterministic system but also for the stochastic system and quantitative methods may fail in proper estimation of measurements like correlation dimension. In other words, temporal correlation cause some statistical dependency and when the possible pairs of points in vector space are statistically dependent, the basic form of correlation sum estimator becomes biased in the direction of very small dimension. A stochastic process with infinite dimensions can also give a low dimensional estimate in this way and to avoid this Theiler, (1986) suggested to keep out all the points which are close in time not in geometry. Calculation of correlation sum for close pairs occurring only after a suitable correlation time will reduce the effect of temporal correlation. However, some feature a chaotic system can also be imitated by a linear stochastic process or a series with colour noise. Therefore, one needs to employ techniques such surrogate data test to eliminate such possibility.

2

Underlying Dynamics of Wind Speed Fluctuations

Modelling the intermittency of wind speed has got significant relevance on many fields including the economy of the region as a renewable energy source. Most of the available modelling techniques assume the temporal fluctuations in the wind is due to the stochastic nature of the underlying dynamics and is best described by statistical methods or the probabilistic distribution. The advent of chaos theory have changed the perception about irregular fluctuations of dynamic systems and it has demonstrated that random-like fluctuations can also arise from deterministic chaotic systems. In this chapter, we have analysed the deterministic nature of apparent random fluctuations seen in the daily average wind speed with the help of nonlinear time series analysis tools. Wind speed data measured at nine typical locations over Indian subcontinent from 2005 to 2015 are used. The values of significant chaotic quantifiers obtained from the analysis clearly show the deterministic, low-dimensional and chaotic nature of wind speed dynamics.

2.1 Introduction

Energy and economic growth are considered to be complementary to each other. While energy contribution can stimulate economic growth in both direct and indirect ways by supporting the industrialisation phase of a region, economic development can engender increased demand for energy forcing the authorities to identify and develop new energy technologies. Sustainable energy resources including wind energy technologies are gaining much attention in recent days since the traditional non-renewable resources are depleting at a faster rate and very much hazardous to the environment. As the largest contributor and fastest growing resource among renewables, wind energy is expected to continue its rapid growth for some decades and in worldwide the international wind community is monitoring advancements in any technology with an expectation of increasing the annual energy capture and driving down the cost of wind energy through R&D (IEA, 2013). Characterization of the wind resource has already been identified by IEA as one of the strategic research areas by which remarkable cost reduction and optimal site assessment are possible. Short-term forecasting of wind is an important aspect of wind resource characterization, and an effective representation of the wind resource can improve the forecasting accuracy resulting in a more precise plant performance.

A number of modelling techniques can be found both in literature and practice in the aim of improved wind energy acquisition and utilisation (Kavasseri et al., 2009; Elliott, 2004; Finzi et al., 1984; Celik, 2004). Most of the tools reported for this purpose consider wind speed as a random process and used statistical methods like autoregressive models and probabilistic distribution function to describe the underlying dynamics (Kamal et al., 1997; Kavasseri et al., 2009; Hennessey Jr, 1977; Cadenas et al., 2007; Celik, 2004; Mathew et al., 2011). In 1990 Elman, (1990) introduce the concept of networks with memory which is capable of predicting the future behaviour from its previous responses. This concept is commonly known as artificial neural networks (ANN) and widely used for classification and prediction problems. ANN and its variants are one of the commonly used methods for making short term wind speed and power predictions (Mohandes et al., 1998; Monfared et al., 2009; Mabel et al., 2009; Gomes et al., 2012; Beyer et al., 1994; Bechrakis et al., 1998). Support vector machines are supervised learning techniques for classification and regression and it identifies the best hyperplane with the maximum margin between the two classes (Steinwart et al., 2008). Use of SVM for predicting wind speed and power one step ahead has also been reported (Zhou et al., 2011; Zeng et al., 2011). Various studies on the effect of hybrid models which combine different forecasting models for predicting wind speed can also found in the literature (Soman et al., 2010; Liu et al., 2014; Haque et al., 2013). Despite all the efforts none of the methods performed well enough in comparison with persistence method which assumes the wind speed are persistent (Sfetsos, 2000). Undoubtedly highly intermittent nature of wind speed makes the system much difficult to model and exploring the sources of this random fluctuation can aid in better characterization and modelling of wind resource.

Most of the reported wind speed modelling techniques assume the system as a stochastic process because of the presence of highly irregular fluctuations in the data. Since these fluctuations can also arise from a deterministically chaotic system, it is worth investigating whether the underlying dynamics is stochastic or deterministic. Palmer et al., (1995) analysed several time series of the horizontal wind speed and X-band Doppler radar signals measured concurrently over ocean surface for nonlinearity and turned up with the result of the low-dimensional dynamical behaviour of both the systems. For a limited period, they could also obtain a higher prediction correlation coefficient with a neural network deterministic model. Ragwitz et al., (2007) attempted a comparison between the stochastic and deterministic model for wind speed time series. Even though Ragwitz reported no improvement in prediction accuracy by using nonlinear models, they observed sufficiently great accuracy in predicting wind gusts. Hirata et al., (2008) proposed a prediction framework for wind direction based on a two-dimensional wind vector representation and they observed the time series data used for all these studies are too short. The analysis did by Martin et al., (1999) had used fairly large enough hourly data of wind speed time series measured over 7 years and expressed it as a sum of the deterministic component and a probabilistic or stochastic component. The analysis came up with the evidence of strong 1-year, 24-hour and 12-hour periodicity in deterministic components. These natural diurnal, yearly and semi-diurnal periodicities in wind time series are natural earth cycles and have been already reported by Brett et al., (1991) and Gavalda et al., (1992) and the presence of a periodic component in the data is a clear evidence of deterministic nature of the system. Nevertheless, the authors are vague in stating whether the ostensible random fluctuations are strictly from the stochastic process or arising out of chaotic underlying dynamics.

Apart from the local topography, earth rotation and solar heating are the major causes of surface wind blowing on earth and undoubtedly earth revolution is deterministic in nature. Although

Location	State	Latitude	Longitude
Bhuj	Gujarat	23.287	69.67
Ahemadabad	Gujarat	23.077	72.634
Bhopal	Madhya Pradesh	23.287	77.337
Jabalpur	Madhya Pradesh	23.177	80.052
Birsa Munda Airport	Jharkhand	23.314	85.321
Coimbatore	Tamilnadu	11.031	77.044
Anantapur	Andhra Pradesh	14.583	77.633
Akola	Maharashtra	20.7	77.033
Indira Gandhi Airport	Delhi	28.566	77.103

TABLE 2.1: The geographical locations where wind speed data have been considered for analysis in this thesis.

many authors have argued solar radiation as a stochastic process and hence the wind can be modelled better with both deterministic and stochastic factors, recent studies on several other atmospheric parameters suggest the possibility of other way round. Kumar et al., (2004) shown the strong chaotic nature of underlying dynamics of Total Electron Content (TEC), which is strongly influenced by the solar radiation. Assuming surface wind as a similar parameter Sreelekshmi et al., (2012) have done a preliminary analysis in this direction using 10 year daily mean wind speed (DMWS) measured at Thiruvananthapuram, Kerala, India. Their analysis of wind speed data in Thiruvananthapuram (8.483° N, 76.950° E) using nonlinear time series analysis tools as implemented in the TISEAN package (Hegger et al., 1999) reveals the possibility of the deterministic but chaotic behaviour of the underlying dynamics of apparent random-like oscillations of wind speed measurements. Their assessments are based on a single location. As noted earlier, the wind speed dynamics is highly dependent on local topography. To make the affirmation further stronger, we have done a detailed analysis of DMWS data measured at various locations in India are the results are reported in this chapter. We have also examined the latitudinal and longitudinal variation in chaotic behaviour of wind speed in terms of the measurements of some nonlinear quantifiers like lyapunov exponent, and correlation sum *etc.*. While the investigation of latitudinal variation is done by collecting DMWS data from five locations with a fixed latitude and varying longitude, longitudinal variation is done by collecting DMWS data from five locations with a fixed longitude and varying latitude. Detailed information of the selected locations given in Figure. 2.1 is provided in Table 2.1. The R code used to generate the map Figure. 2.1 is given in Listing 2.1.

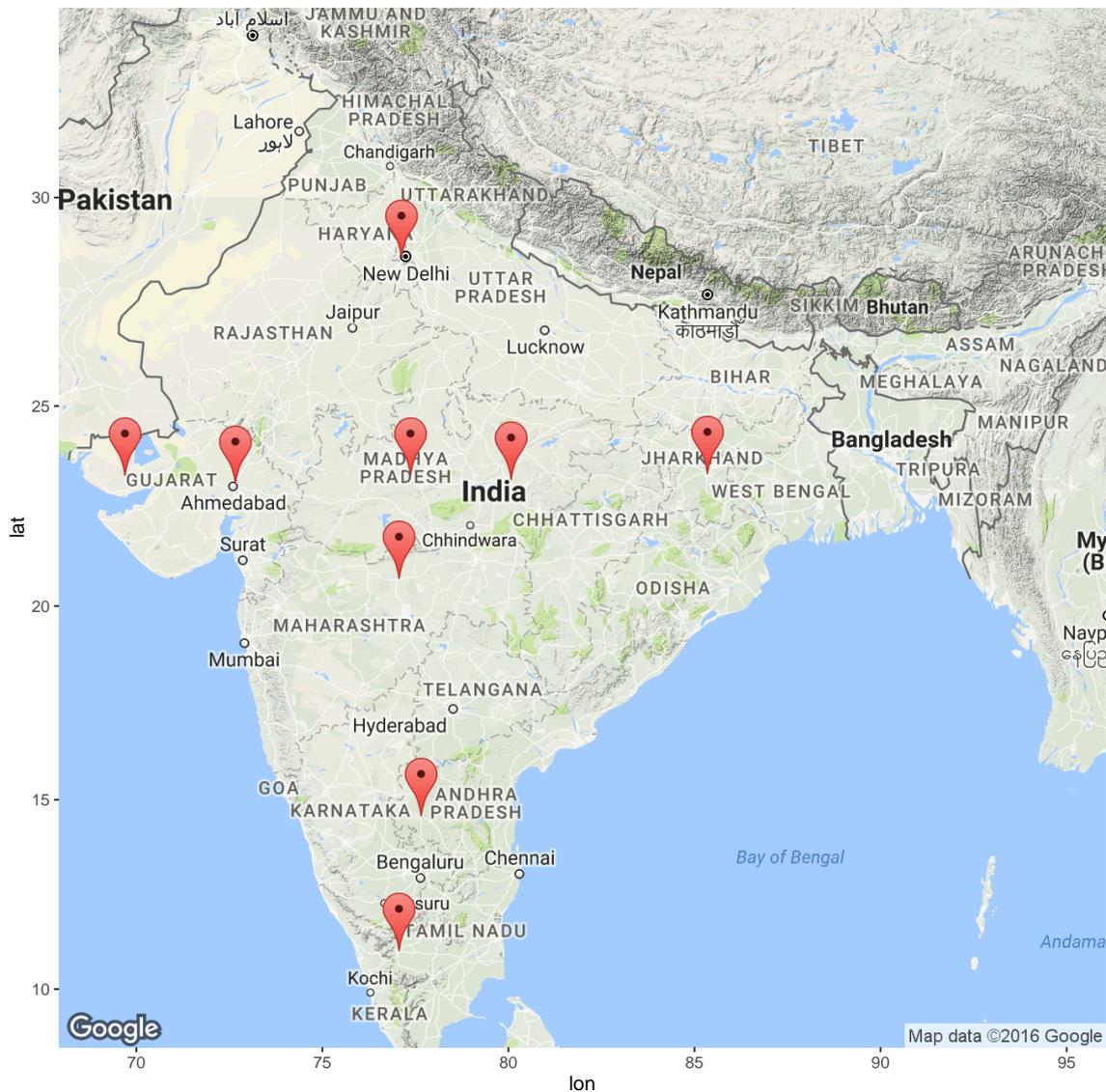


FIGURE 2.1: Location at which Daily Mean Wind Speed (DMWS) data is measured.

```
library(ggmap)
a=c(11.031,20.7,14.583,28.566,23.287,23.287,23.077,23.314,23.177)
b=c(77.044,77.033,77.633,77.103,77.337,69.670,72.634,85.321,80.052)
mark=as.data.frame(cbind(b,a))
ggmap(get_googlemap(center = c(lon = 81.9629,lat = 22), zoom=5,markers = mark))
```

LISTING 2.1: The code in R used to produce the map in Fig 2.1.

2.2 Chaotic system

Daily mean wind speed measurements from different locations in India are plotted in Figure 2.2. Random like temporal fluctuations are clearly evident in all the data and as we discussed above these fluctuations can arise not only from a noisy stochastic system but also from a sensitive low-dimensional deterministic system. Evidence of the chaotic behaviour of the dynamic system are first reported by Lorenz, (1963) and he well explained how a deterministic system is evolving into quite complex unpredictable states in the long term. Hence before concluding fluctuations in

the measured DMWS is purely random a detailed analysis is necessary for exploring the sources of apparently random behaviour. Under some precise condition measures of complexity and predictability for characterising the system evolution are good tool to show whether the system follows a chaotic flow or not.

For any dynamical system time dependence of its states, which are represented by vector points in geometrical space, can be well described by the equation of motion and it is given by

$$\dot{x} \equiv \frac{dx}{dt} = f(x) \quad (2.2.1)$$

where $x(t)$ is the state vector. For a dissipative system as $t \rightarrow \infty$ the trajectories that follow the system evolution will be attracted to a subset of phase space known as *attractor*. Characterization of the attractor is an excellent technique to obtain the detailed information about properties of the system under consideration. Some dynamical systems are highly associated with chaotic behaviour for its hypersensitivity to initial conditions. More precisely, trajectories following the states of the system that are originated from two points which were very close in phase space initially will deviate from each other in an exponential rate. In a longer period, even though the diverging trajectories may evolve separately without depending on each other and move forward in an uncorrelated manner, it will restrict themselves within the limits of a sub set of phase space. *Chaos is the bounded aperiodic behaviour in a deterministic system that shows sensitive dependence on initial conditions* (Alligood et al., 1997). In chaotic system, the adjacent trajectories following the system evolution may spread initially and eventually comes back to remain a bounded region. As $t \rightarrow \infty$ repeated spreading and folding of trajectories happens confining it in a specific region in phase space. This complex region into which the states of the system is attracted as time evolve is known as *attractor*, and it maintenance a definite geometry. Even though underlying dynamics of chaotic system are deterministic in nature due to the property of sensitivity to initial condition predictability is limited to short period. The restriction in long term prediction occurs due to the unavoidable measurement errors which will be amplified as time goes on and can cause exponential divergence of predicted trajectory from original one. Before the advent of chaos theory many chaotic systems producing apparent irregular behaviour were dubbed to be stochastic (Alligood et al., 1997; Ott, 2002).

2.3 Attractor reconstruction

In many real world situations the dynamical system, as given in Eq. 2.2.1, or the state vector $x(t)$ may not be known or available but what is accessible will be the measurements of a variable $y(t)$ equidistant in time *i.e.* a time series. The main objective in analysing such time series is to get insight into the underlying dynamical system. The first and foremost step in time series analysis reconstruct the dynamics of $x(t)$ on the attractor using the methodology known as *attractor reconstruction*, first suggested by Packard et al., (1980). The attractor generated by the m -dimensional delay vector

$$y(t) = (y(t), y(t + \tau), \dots, y(t + (m - 1)\tau)) \quad (2.3.1)$$

constructed from $y(t)$ at time interval (delay) τ is topologically equivalent to attractor of the state vector $x(t)$. The validity of the embedding $x(t) \rightarrow y(t)$ guaranteed by the embedding theorem of Takens, (1981) and its extensions by Sauer et al., (1991) and Sauer et al., (1993) for all values

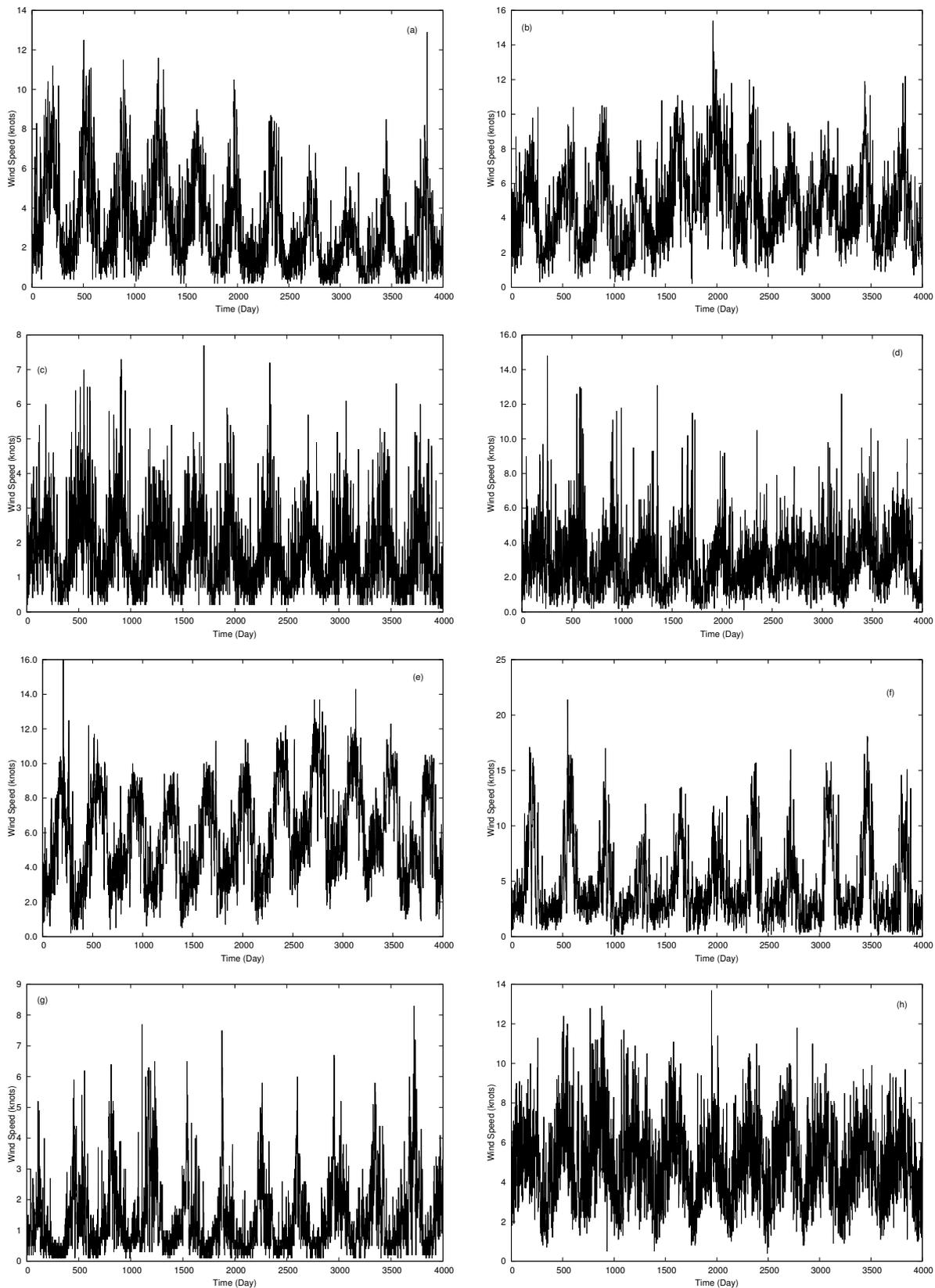


FIGURE 2.2: Time series of the daily mean wind speed (DMWS) in knots, measured across Indian subcontinent at locations (a) Bhuj (b) Ahemadabad (c) Jabalpur (d) Birsa Munda Airport (e) Coimbatore (f) Anantapur (g) Akola (h) Indira Gandhi Airport.

of delay τ and smooth measurement functions $y(t) = h(x(t))$ for $m > 2D$ where D is the box-counting dimension. Hence, the dynamical and geometrical characteristics of the original system $x(t)$ are preserved the reconstructed space and can be computed from the measured time series $y(t)$ (Kantz et al., 2004; Ott et al., 1994). The attractor reconstruction depends on two parameters, the embedding dimension m and delay τ .

2.4 Nonlinear time series analysis

A time series is the measurements of a time dependent variable at equal interval in time. Here we consider daily mean wind speed data, not the wind speeds measured at regular daily intervals. This averaging process contributes an additional additive noise apart from the measurement noise with zero mean and delta correlation. As a first step, we reduced the effect of the noise using the method proposed by Schreiber, (1993). Despite the apparent random-like fluctuations of the DMWS time series the plots in Figure. 2.2 shows strong annual variations. This was further confirmed by the space-time separation plot of each time series. A space-time separation plot depicts the relative separation in time of a pair points on a trajectory along the horizontal axis and their separation in space along the vertical axis. A space-time separation plot is useful in identifying the temporal correlations in a time series (Provenzale et al., 1992). Typical space-time separation plots are given in Figure. 2.3. In an *epoch analysis*, on each time series the modulation effect annual variations was reduced by deducting from each of the data points which are 365 days apart their average value (Kumar et al., 2004). The variation of 28 days arising from lunar influence was evident from the resulting time series and hence the procedure was repeated for each time series to reduce the effect of 28-day variations. Further analysis was carried out on each of the resultant de-noised and detrended time series retaining temporal variations. The plots of the de-noised and detrended time series of eight locations are given in Figure. 2.4 and their space-time separation plots are given in Figure. 2.5.

The first step in the analysis of the de-noised data is the reconstruction of the attractor as per the method discussed in the previous section for which one needs to estimate the optimum embedding parameters - delay τ and embedding dimension m . The choice of τ and m crucial in deriving inferences when the time series is the result of the experimental measurements. Selection of small values of delay shall lead to highly correlated vectors $y(t)$ leading to unduly larger values for the correlation dimension whereas inappropriately large values may lead to fairly uncorrelated components resulting in data randomly distributed in the embedding space (Kantz et al., 2004). Proper choice of the time delay is, therefore, important and a first guess of a suitable delay may be obtained from The autocorrelation function of the sample data y_i given by

$$\rho(\tau) = \frac{\sum_i (y_i - \bar{y})(y_{i+\tau} - \bar{y})}{\sum_i (y_i - \bar{y})^2}, \quad (2.4.1)$$

where \bar{y} is the sample mean, can be utilised to estimate an approximate value of τ (Kantz et al., 2004). A better method to fix the value the delay τ is to calculate the time-delayed mutual information suggested by Fraser et al., (1986). This method also takes into account of the non-linear correlations.

In this method, a quantity called average mutual information is computed for various delays as a measure of the predictability of $y(t + \tau)$ given $y(t)$.

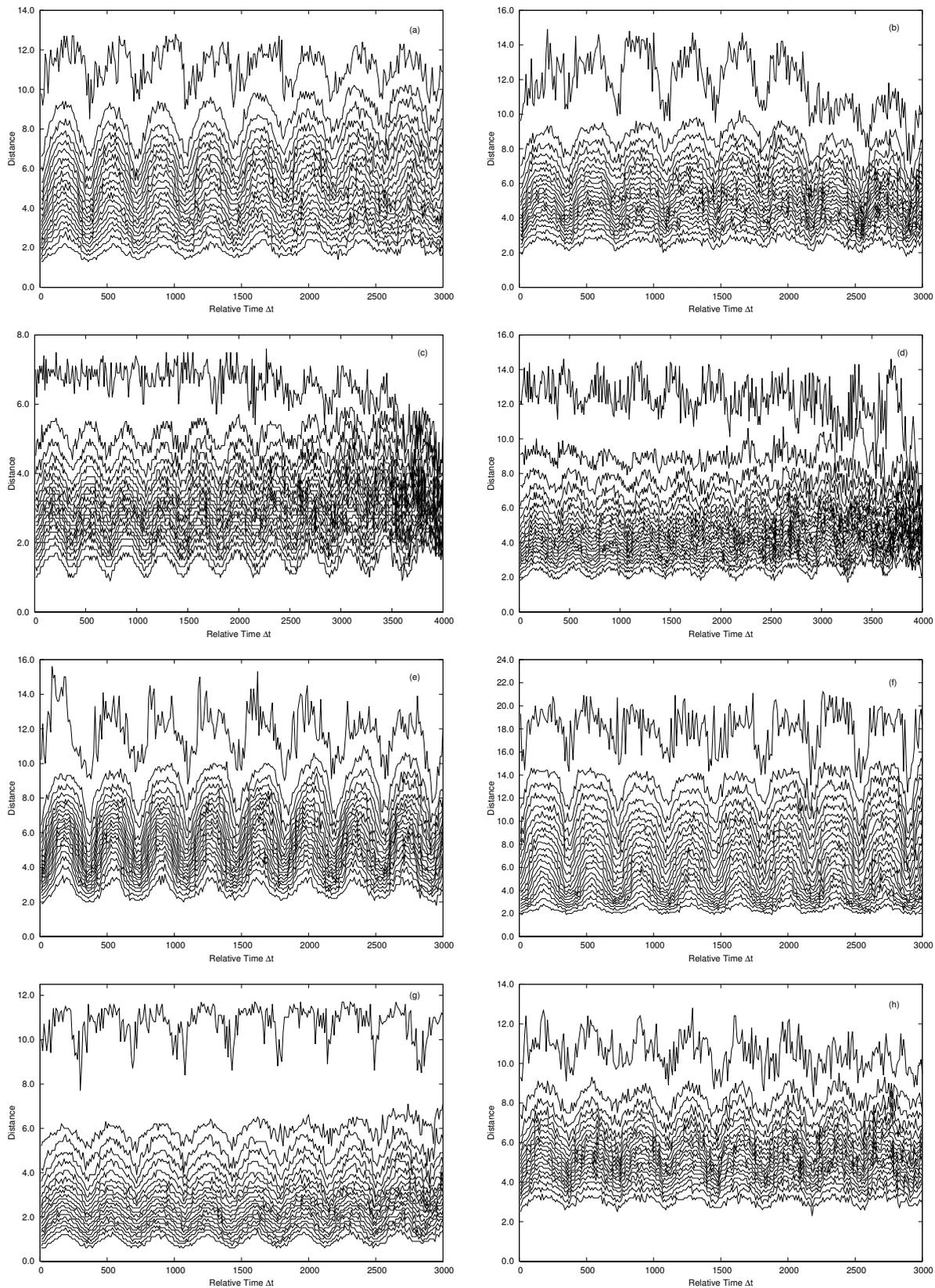


FIGURE 2.3: Space-time separation plot for the time series measured at locations (a) Bhuj (b) Ahemadabad (c) Jabalpur (d) Birsa Munda Airport (e) Coimbatore (f) Anantapur (g) Akola (h) Indira Gandhi Airport.

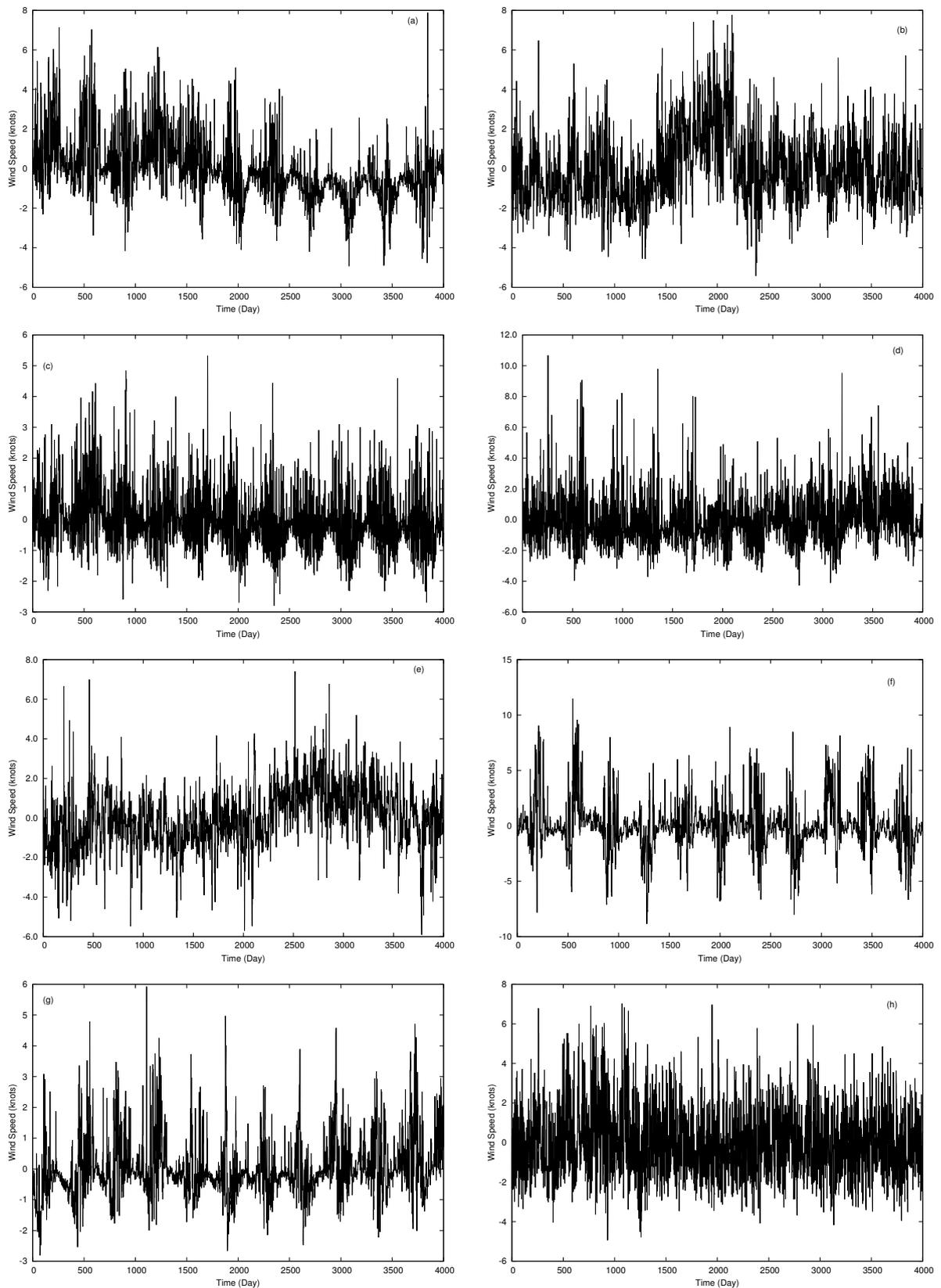


FIGURE 2.4: Time series of the daily mean wind speed (DMWS) in knots, measured across Indian subcontinent at locations (a) Bhuj (b) Ahemadabad (c) Jabalpur (d) Birsa Munda Airport. (e) Coimbatore (f) Anantapur (g) Akola (h) Indira Gandhi Airport.

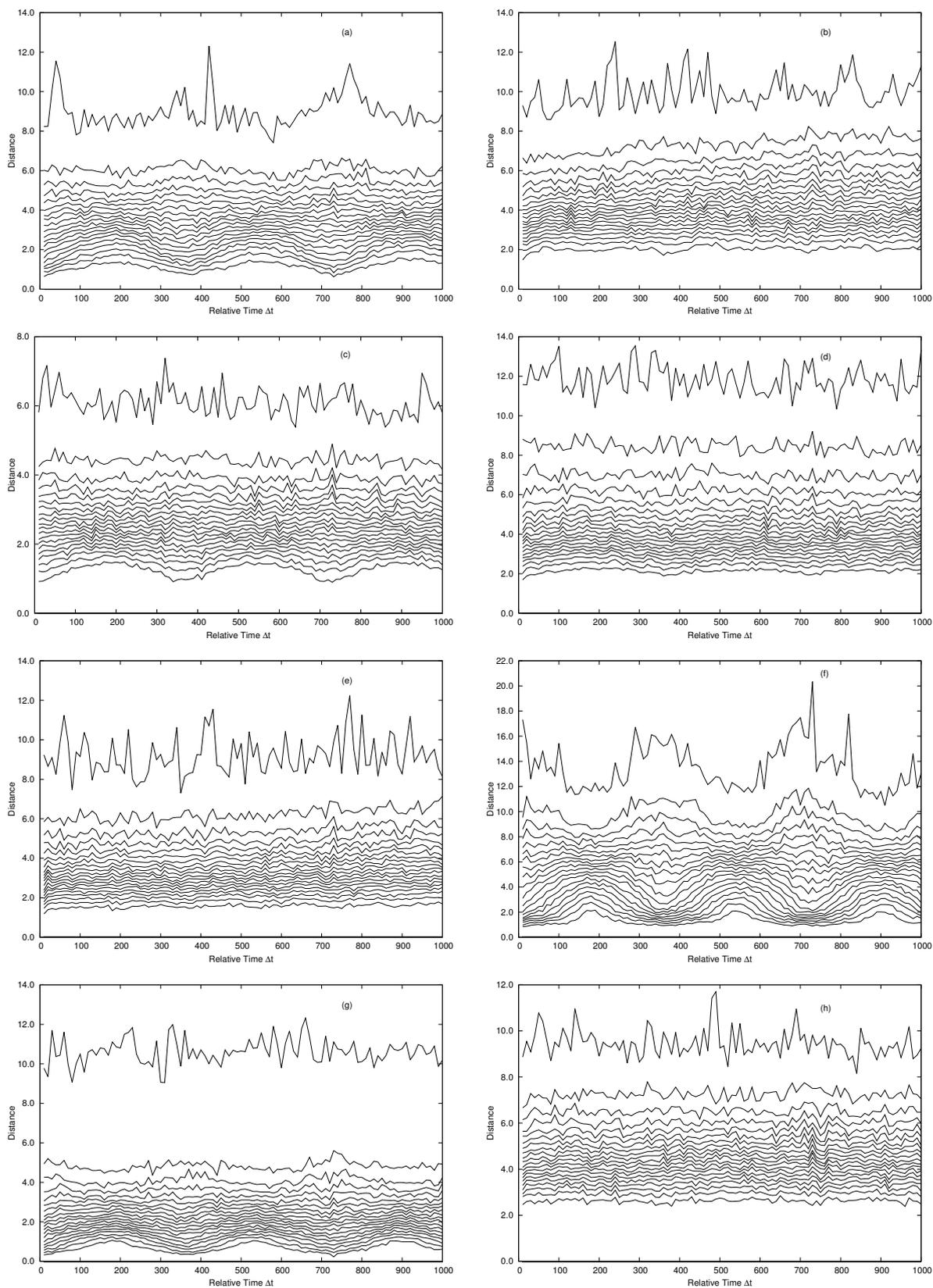


FIGURE 2.5: Space-time separation plot for the time series measured at locations (a) Bhuj (b) Ahemadabad (c) Jabalpur (d) Birsa Munda Airport (e) Coimbatore (f) Anantapur (g) Akola (h) Indira Gandhi Airport.

For a give delay, the mutual information $I(\tau)$ for a given delay τ is calculated by treating the sequences (y_i) and $(y_{i+\tau})$ as values of random variables X and Y according to the formula

$$I(\tau) = \sum_{y \in Y} \sum_{x \in X} p(x,y) \log_2 \left(\frac{p(x,y)}{p(x)p(y)} \right). \quad (2.4.2)$$

Here, $p(x,y)$ is the joint probability mass function of X and Y with marginal $p(x)$ and $p(y)$. The probabilities can be estimated by constructing a histogram of the data points. The average mutual information of a time series $y(t)$ is computed for various values of the delay τ as a measure of the predictability of $y(t + \tau)$ given $y(t)$. The value of the average mutual information where its plot against increasing values of delay τ exhibits a marked minimum can be good estimate of the optimal value of the delay parameter τ . The average mutual information DMWS data indicates $\tau = 1$ can be a good choice for all the locations under study. Typical plots the average mutual information eight different locations are given in Figure. 2.6. The plots of other locations also show similar features. We would like to emphasise that the product $m\tau$ is more important than values of m and τ independently. The precise knowledge of m is only required to estimate the dynamics with minimal computational effort (Kantz et al., 2004).

According to Kantz et al., (2004) the value of embedding dimension m , it should be large enough for the attractor to fully unfold in the embedding space. However, a value that is too large may cause the various algorithms to underperform (Kantz et al., 2004). Kennel et al., (1992) proposed a practical method to choose the right value of the embedding dimension m by calculating the fraction of *false neighbours* as a function of m .

False neighbours arise due to the crossing of trajectories when the attractor can not unfold its true geometry when the value of m is not large enough as a result of projection onto a small dimensional space. To estimate a suitable value for the m fraction of false neighbours are computed in progressively higher dimension until the difference becomes negligible. The first value of m corresponding to the first minimum of the fraction of false neighbours indicates a suitable value for the embedding dimension. We have computed the fraction of false neighbours of DMWS data all locations considered, and plots of eight of them are given in Figure. 2.7. It can be seen that values of $m > 13$ (in same case $m > 12$) the fraction of false neighbours become extremely small. The value $m = 14$ can be selected safely in all these locations. This shows preliminary indication that underlying dynamics is low dimensional in character across the locations.

The delay representation of the time series in four locations with $m = 14$ and $\tau = 1$ are given in Figure. 2.8. The definite structure in these figures is an indication of the deterministic nature of the underlying dynamics. The similar pattern is observed in other locations as well.

A typical feature of a chaotic attractor is its self-similarity. Various dimension estimates, such as the box-counting dimension, the Hausdorff dimension *etc.*, were introduced in the literature to quantify the structure of a chaotic attractor. All these quantities of dimension are the generalisation of the Euclidean definition of dimension to account the self-similar character of chaotic attractors. The dimension estimate of a chaotic attractor need not be an integer. Among all these most popular in literature is the correlation dimension introduced by Grassberger et al., (1983). The correlation integral $C(\epsilon)$, is the probability that a pair of points chosen randomly on the attractor is separated by a distance less than ϵ . It is found that $C(\epsilon) \propto \epsilon^{D_2}$ as $\epsilon \rightarrow 0$. Therefore, the correlation dimension can be computed from the slope of the curve of $\ln C(\epsilon)$ versus $\ln(\epsilon)$ given

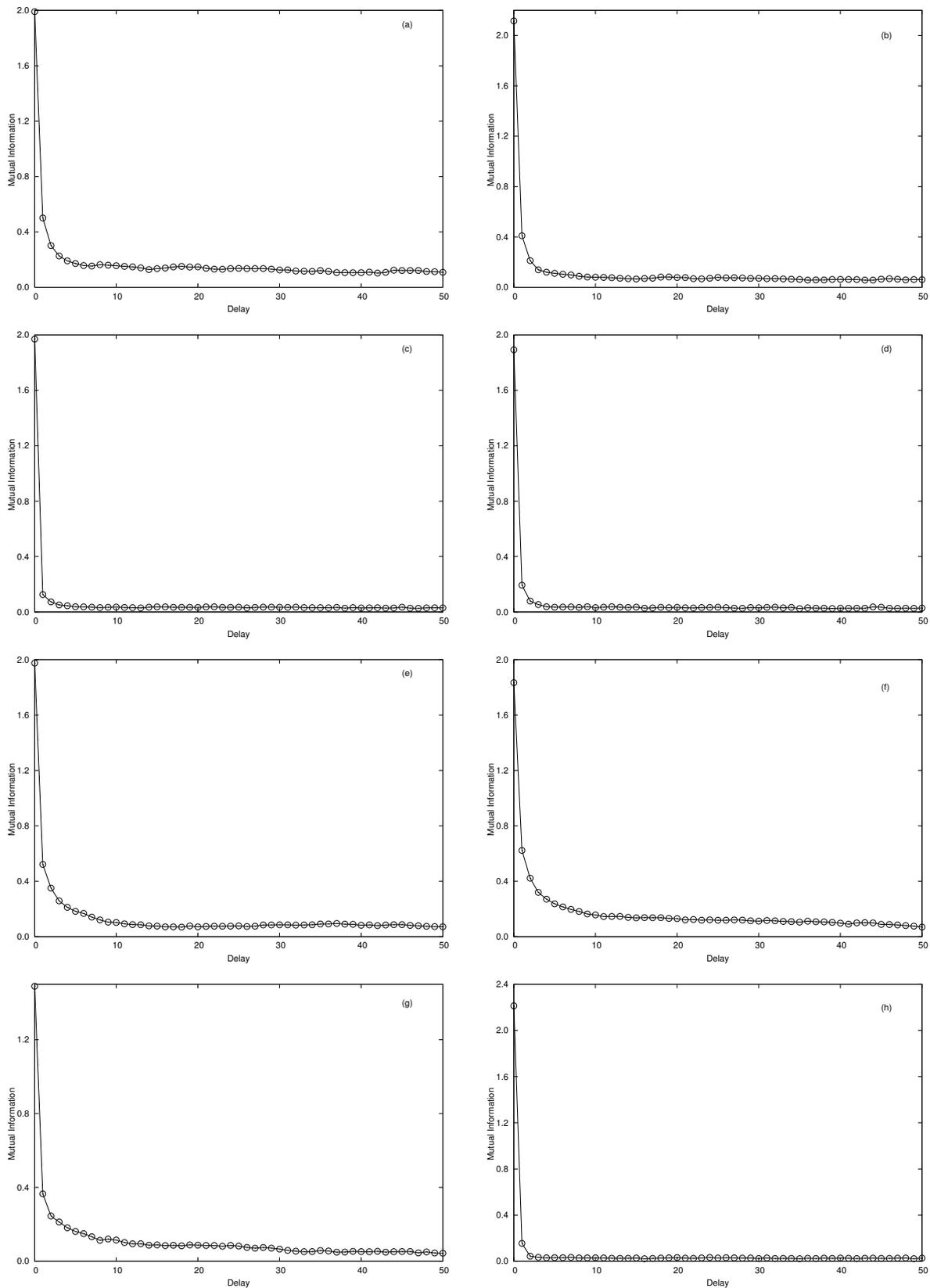


FIGURE 2.6: Mutual information of the time series of the DMWS data measured at (a) Bhuj (b) Ahmedabad (c) Jabalpur (d) Birsa Munda Airport (e) Coimbatore (f) Anantapur (g) Akola (h) Indira Gandhi Airport.

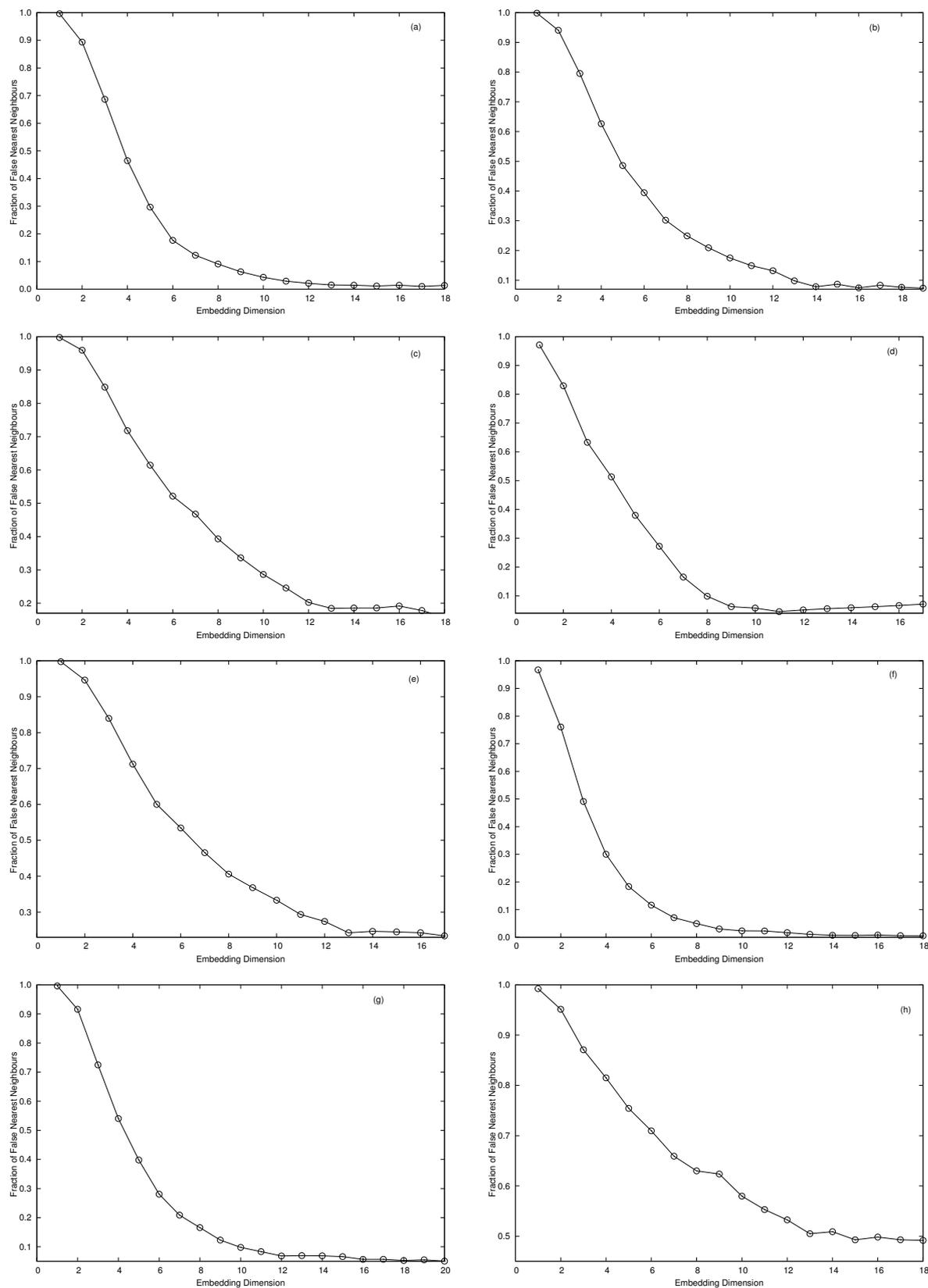


FIGURE 2.7: The fraction of false nearest neighbours as a function of the embedding dimension m for the DMWS data measured at (a) Bhuj (b) Ahemadabad (c) Jabalpur (d) Birsra Munda Airport (e) Coimbatore (f) Anantapur (g) Akola (h) Indira Gandhi Airport.

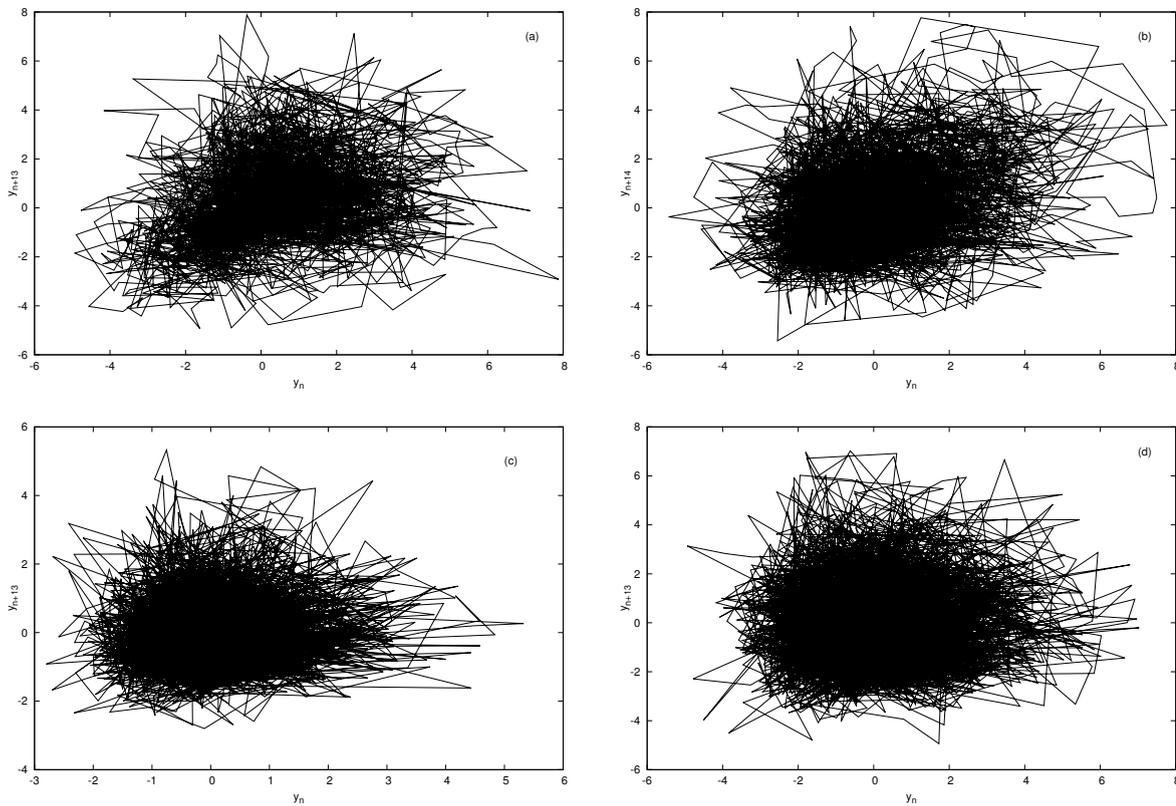


FIGURE 2.8: The delay representation of the detrended time series (a) Bhuj (b) Ahmedabad (c) Jabalpur (d) Indira Gandhi Airport.

by

$$D_2 = \lim_{\varepsilon \rightarrow 0} \frac{d \ln C(\varepsilon)}{d \ln \varepsilon}. \quad (2.4.3)$$

For a single time series and N data points of m -dimensional delay vectors y_i , the correlation integral $C(\varepsilon)$ is approximated by the correlation sum $C(\varepsilon, m)$ given by Kantz et al., (2004)

$$C(\varepsilon, m) = \frac{2}{N(N-1)} \sum_{i=1}^N \sum_{j=i+1}^N \Theta(\varepsilon - \|y_i - y_j\|), \quad (2.4.4)$$

for sufficiently large N , where $\Theta(a) = 1$ if $a > 0$, $\Theta(a) = 0$ if $a \leq 0$. In practice the local slopes

$$D_2(\varepsilon, m) = \frac{d \ln C(\varepsilon, m)}{d \ln \varepsilon} \quad (2.4.5)$$

are computed and plots them as a function of ε for various m ; the value corresponding to a plateau in the curves is estimated as identified as an approximate value of D_2 .

However, only the spatial closeness of points should be accounted for in Eq. 2.4.5 whereas in some cases, it can be affected by the temporal closeness of points as well. To avoid this points which are closer in time by less than a *Theiler window* ω , which is approximately equal to the product of the time delay and the embedding dimension, should be excluded from the computation of the correlation sum (Theiler, 1986). According to Hegger et al., (1999) the value of ω should be chosen generously.

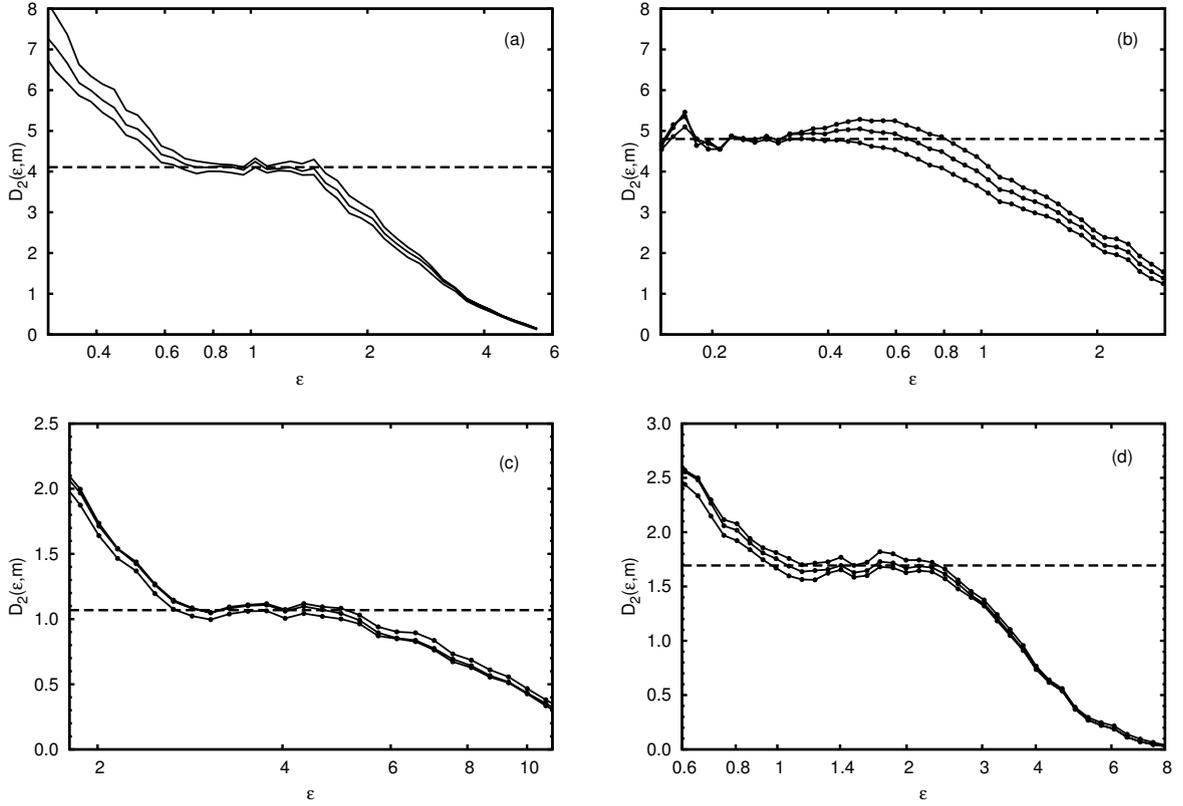


FIGURE 2.9: The local slopes $D_2(\epsilon, m)$ for the de-noised and detrended time series of DMWS at (a) Jabalpur (b) Coimbatore (c) Anantapur and (d) Akola. The values of the correlation dimension are given in Table 2.2.

Location	Latitude (degree)	Longitude (degree)	Correlation dimension
Bhuj	23.287	69.670	2.2488 ± 0.0092
Ahemadabad	23.077	72.634	5.1174 ± 0.0587
Bhopal	23.287	77.337	2.8775 ± 0.0579
Jabalpur	23.177	80.052	4.1087 ± 0.0265
Birsa Munda Airport	23.314	85.321	4.7908 ± 0.0331
Coimbatore	1.031	77.044	4.8001 ± 0.0213
Anantapur	14.583	77.633	1.0685 ± 0.0071
Akola	20.700	77.033	1.6936 ± 0.0116
Indira Gandhi Airport	28.566	77.103	5.9420 ± 0.1863

TABLE 2.2: The estimated values of the correlation dimension at various locations. The values appear to depend more on the local topography rather than geographical location.

The typical plots of the local slopes $D_2(\epsilon, m)$ is given in Figure. 2.9. We have calculated the correlation dimension of the DMWS data at all locations and is given in Table 2.2. The wind dynamics may be affected myriads of factors, but the estimated values of the correlation dimension show that the eventual behaviour characterised by the attractor is low dimensional.

The most striking feature of a chaotic system is its sensitivity to initial conditions. Therefore the trajectories starting from neighbouring initial conditions diverge exponentially as time passes. The average rate of divergence of nearby trajectories is quantified by what known as Lyapunov

Location	Latitude (degree)	Longitude (degree)	Lyapunov exponent
Bhuj	23.287	69.670	0.0434 ± 0.0007
Ahemadabad	23.077	72.634	0.0434 ± 0.0007
Bhopal	23.287	77.337	0.0642 ± 0.0016
Jabalpur	23.177	80.052	0.0675 ± 0.0018
Birsa Munda Airport	23.314	85.321	0.0826 ± 0.0022
Coimbatore	1.031	77.044	0.0984 ± 0.0013
Anantapur	14.583	77.633	0.0778 ± 0.0010
Akola	20.700	77.033	0.0889 ± 0.0011
Indira Gandhi Airport	28.566	77.103	0.0742 ± 0.0015

TABLE 2.3: The estimated values of the maximum Lyapunov Exponent at various locations. The values do not vary significantly across geographical locations.

exponent. Positive Lyapunov exponent is a striking evidence chaotic behaviour of system (Ott, 2002).

The growth of the divergence $\delta(t)$ between two neighbour trajectories is quantified by the maximum Lyapunov exponent λ , so that $\|\delta(t)\| = \|\delta(0)\|e^{\lambda t}$ and hence

$$\lambda = \lim_{t \rightarrow \infty} \frac{1}{t} \ln \frac{\|\delta(t)\|}{\|\delta(0)\|}. \quad (2.4.6)$$

The maximum Lyapunov exponent λ can be estimated by plotting $\ln \delta(t)$ versus t , which should fall nearly in a straight line, the slope of which then gives an estimate of λ . Lyapunov exponents have preserved delay reconstruction as they are invariant under smooth transformations of the attractor, and hence they can be estimated from a time series. There are several algorithms for estimating the maximal Lyapunov exponent from time series. Kantz algorithm (Kantz, 1994; Kantz et al., 2004) is popular in the literature which starts with computing the sum

$$S(\Delta n) = \frac{1}{N} \sum_{n_0=1}^N \ln \left(\frac{1}{\|U(y_{n_0})\|} \sum_{y_n \in U(y_{n_0})} \|y_{n_0+\Delta n} - y_{n+\Delta n}\| \right) \quad (2.4.7)$$

for a point y_{n_0} of the time series in the embedded space and over a neighbourhood $U(y_{n_0})$ of y_{n_0} with diameter ε .

If the plot of $S(\Delta n)$ versus Δn is linear over small Δn , for a reasonable range of ε , and all have identical slope for sufficiently large values of the embedding dimension m , then that slope can be taken as an estimate of the maximum Lyapunov exponent (Kantz, 1994; Kantz et al., 2004).

The curves of $S(\Delta n)$ for $m = 14, 15, 164$ for various locations are given in Figure. 2.10. The estimated values of the maximum Lyapunov exponent are given in Table 2.3. The positive values of the maximum Lyapunov exponents show the underlying dynamics of wind speed variations in all these locations are chaotic in nature.

Many characteristics of a chaotic system are also mimicked by a color noise time series. Pavlos et al., (1992) have noted that phase randomization of a chaotic signal can destroy its profile whereas

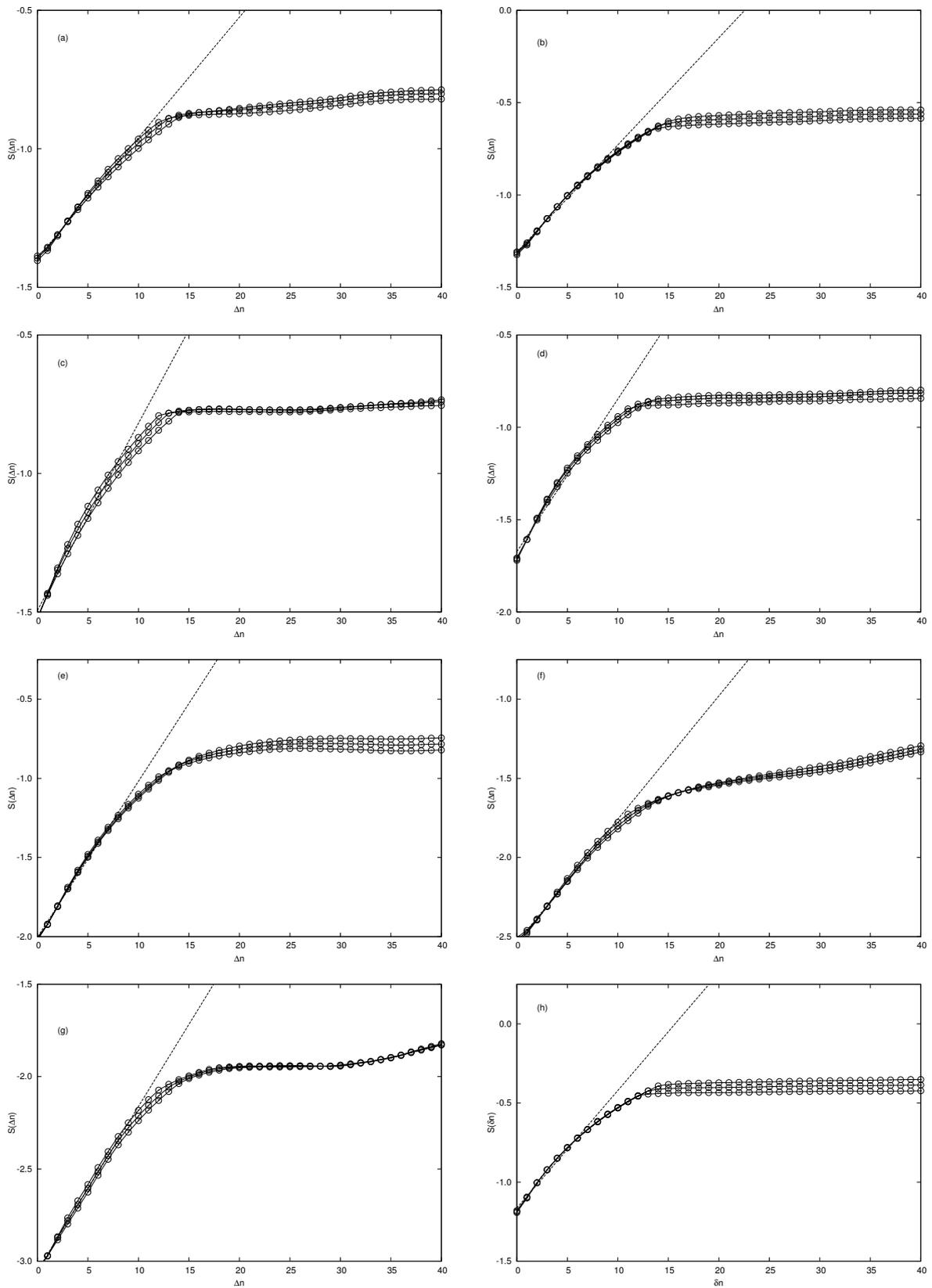


FIGURE 2.10: The curve of $S(\delta n)$ for various embedding dimensions m . The maximum Lyapunov exponent of the detrended time series is the slope of the dashed line. (a) Bhuj (b) Ahmedabad (c) Jabalpur (d) Birsa Munda Airport (e) Coimbatore (f) Anantapur (g) Akola (h) Indira Gandhi Airport.

colour noise time series preserve its profile. We have compared the local slopes of logarithm $D_2(\varepsilon, m)$ of the correlation sum original time series of each location and compared with its phase randomised time series. The phase randomised time series is obtained by obtaining the Fourier series representation original and reconstructing the time series after adding random phase distribution. Essentially we obtained plots like in Fig.11 in Reference (Sreelekshmi et al., 2012). In each location, the chaotic profile of the time series was destroyed by phase randomization indicating the deterministic nature the given time series.

2.5 Surrogate data test

Stochastic systems driven by a linear Gaussian process distorted by some nonlinear process might also exhibit many features of a chaotic system. The main objective of surrogate data test is to ascertain that the complex behaviour exhibit by a time series is not stochastic. We validate further that the results reported in the previous section were not arisen from a linear stochastic process by carrying out surrogate data test on DMWS data at all locations considered.

The method of surrogate data has been widely used for discriminating if the source of apparent random fluctuations in a time series is deterministic or stochastic (Theiler et al., 1992). It is a statistical test to formally reject the hypothesis that the observed time series arose from a linear noise process. The null hypothesis is first formulated that the observed time series is a random process and then an ensemble of random numbers, called surrogate data, consistent with the null hypothesis and otherwise similar to the original data were generated. Then one proceeds to test if a discriminating statistic such as correlation dimension or Lyapunov exponent computed from the surrogate data is significantly different from that of the original data. The null hypothesis is rejected if they significantly different.

For each time series of the DMWS measurements, we generated 40 surrogates by the Amplitude Adjusted Fourier Transform method proposed by Schreiber et al., (1996). The surrogates preserve the amplitude distribution, power spectrum and autocorrelation of the original DMWS-data, so that they can be treated as what the realisations satisfying the null hypothesis. To test null hypothesis we have employed both geometrical and dynamical characteristics such as the fraction of false nearest neighbours, the local slopes of the correlation sums and the curves of $S(\Delta_n)$. The above characteristics are calculated for the data, both original and the surrogate, and the null hypothesis is accepted or rejected depending on the value of the significance of the difference given by Mitschke et al., (1993) and Pavlos et al., (1999)

$$S = \frac{\mu - \mu_{\text{orig}}}{\sigma} \quad (2.5.1)$$

where μ and σ are the mean and standard deviation of the characteristic computed from the surrogates and μ_{orig} is the mean of the characteristic of the original data. It is estimated that we may reject the null hypothesis with 95% confidence level if $S > 2$. In other words if $S > 2$ the observed time series is not a realisation of a linear Gaussian stochastic process with probability 0.95 or more (Pavlos et al., 1999).

In Figure. 2.11 plotted the mean values of the fraction of false nearest neighbours of all surrogates and values one standard deviation away from the mean along that of the original data. The difference between the original data and the surrogates is evident in these figures. The significance of difference S of the data is plotted in Figure. 2.12. It can be noted that the values of S are

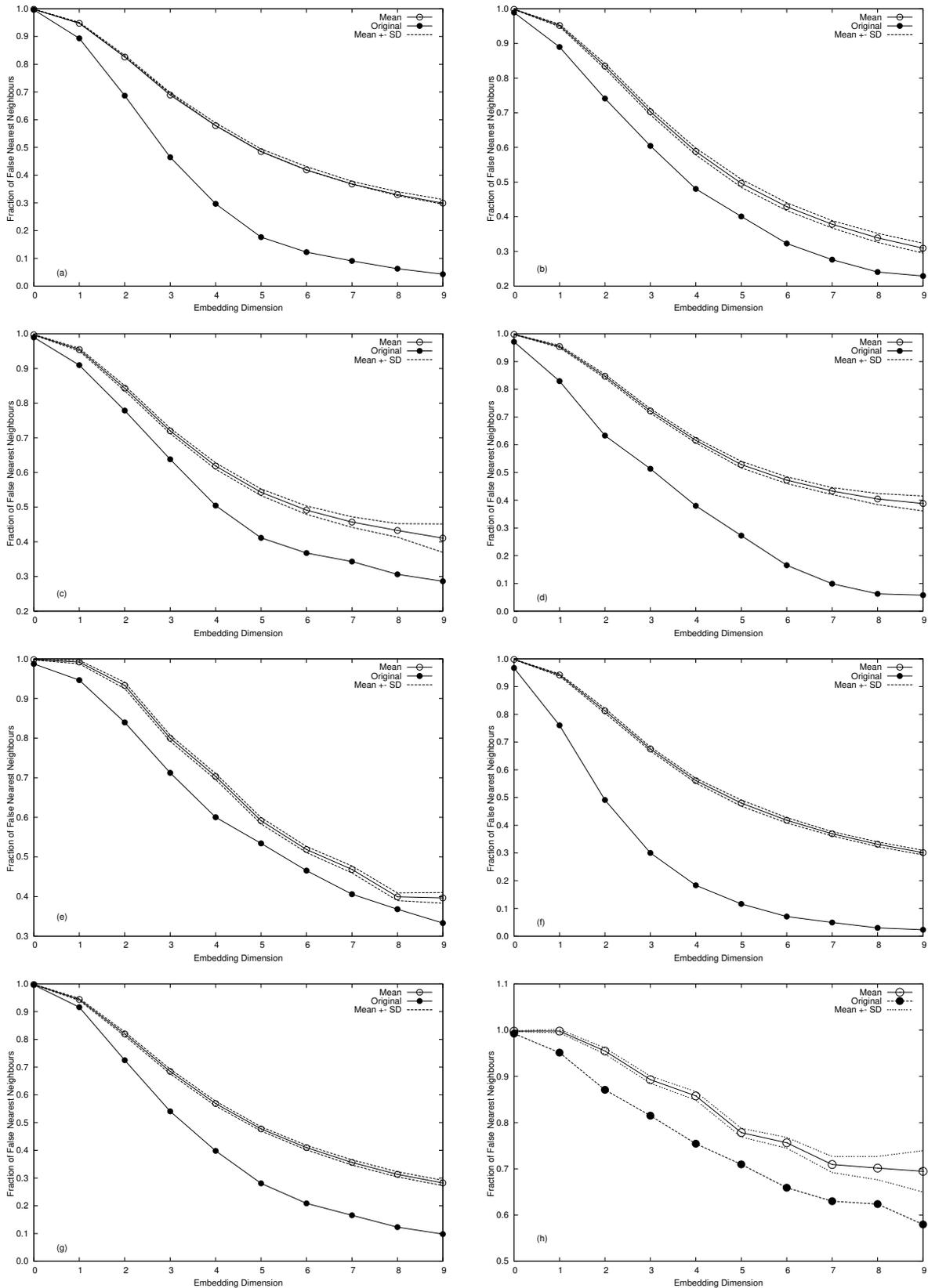


FIGURE 2.11: The mean values of the fraction of false nearest neighbours of the surrogates with standard deviation.(a) Bhuj (b) Ahemadabad (c) Jabalpur (d) Birsa Munda Airport (e) Coimbatore (f) Anantapur (g) Akola (h) Indira Gandhi Airport.

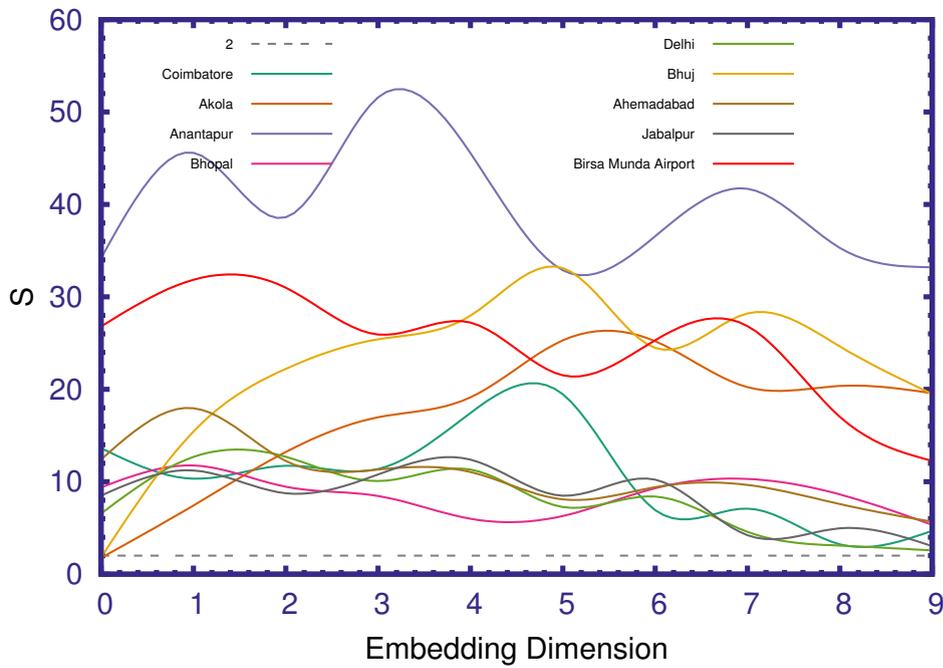


FIGURE 2.12: Plot of the significance of difference S versus m for different locations

remarkably higher than 2 for almost the entire range of values of m considered. Therefore, we can safely reject the null hypothesis.

Next, we compare the original data with its surrogates with $S(\Delta_n)$ of Equation. 2.4.7 as the test statistic. Figure. 2.13 shows the curves of $S(\Delta_n)$, of the surrogates with those of the original data with delay $\tau = 1$ Theiler window $\omega = 25$ and embedding dimension $m = 14$. A strong difference between the values of $S(\Delta_n)$ corresponding to the original data and the surrogates is evident in all the locations. The significance of difference S computed for all locations as shown in Figure. 2.14 is much above 2 for all $\Delta_n \leq 45$ and hence we can reject the null hypothesis with 95% confidence level.

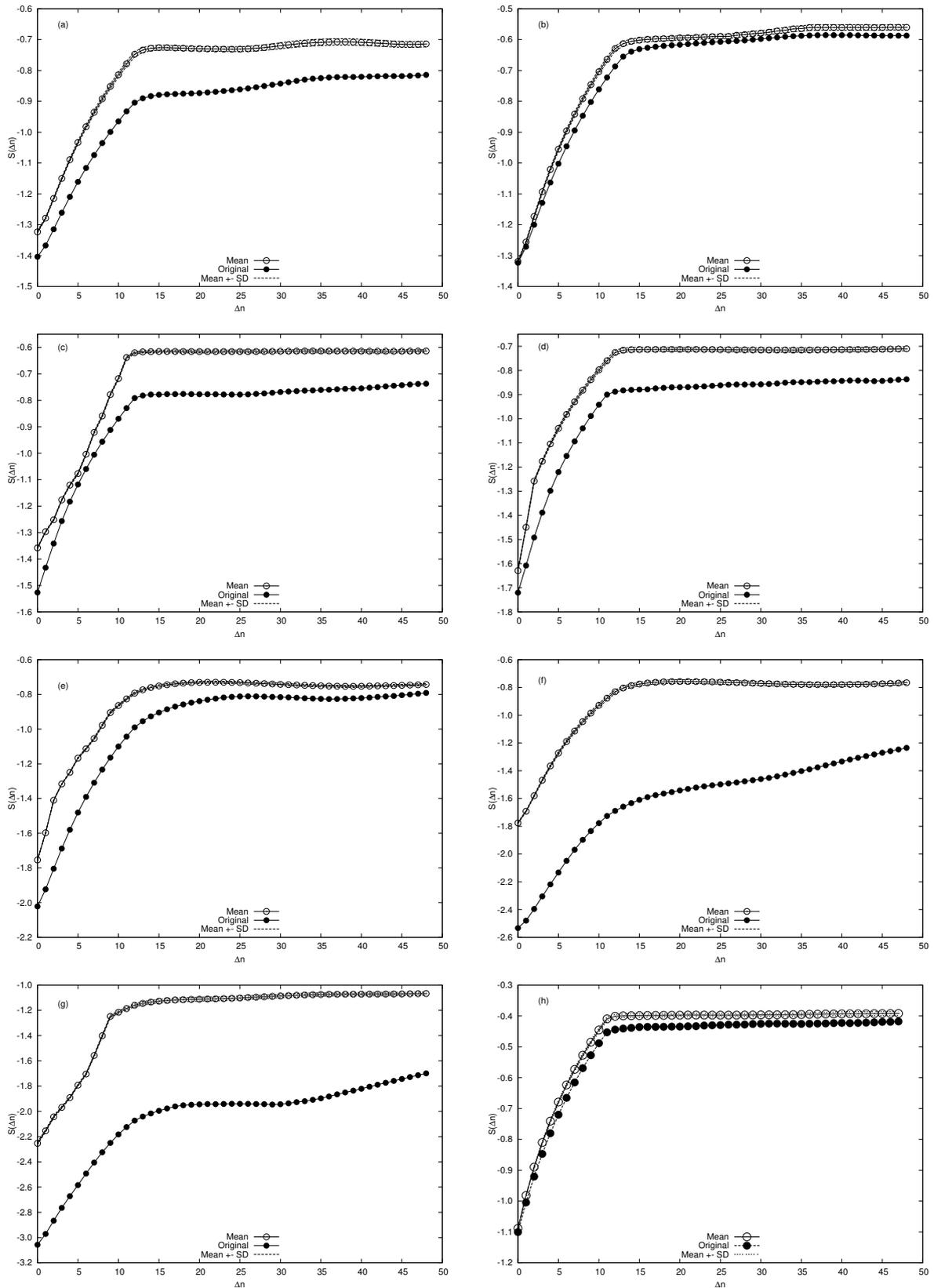


FIGURE 2.13: The mean values of $S(\Delta)$ of the surrogates with standard deviation. (a) Bhuj (b) Ahmedabad (c) Jabalpur (d) Birsa Munda Airport (e) Coimbatore (f) Anantapur (g) Akola (h) Indira Gandhi Airport.

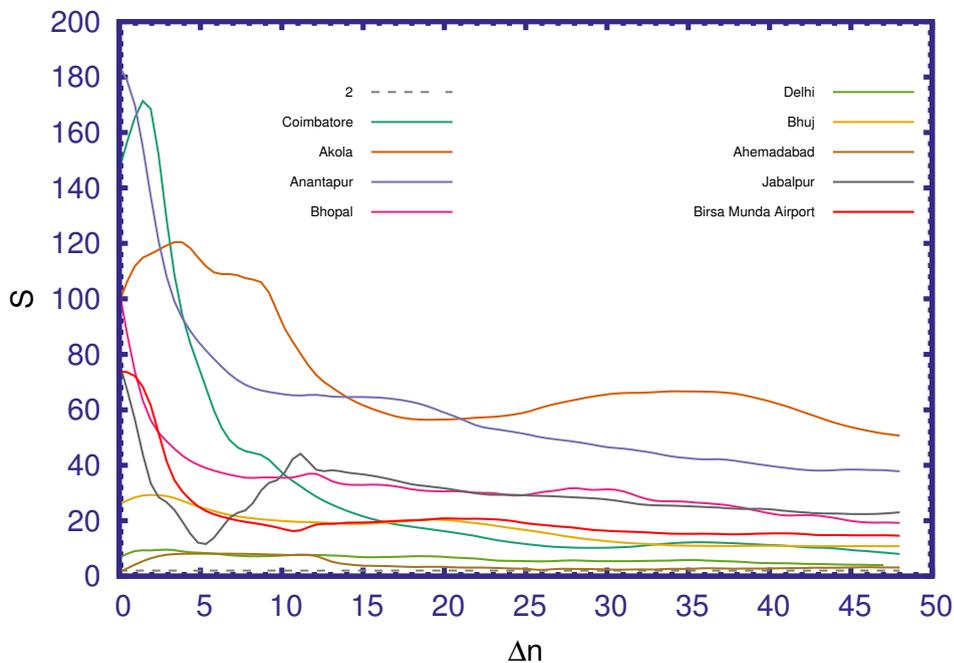


FIGURE 2.14: Plot of the significance of difference S versus Δn for different locations

In conclusion, we can reject the null hypothesis with 95% confidence level and infer that the DMWS-data does not originate from a linear Gaussian process based on the results of these series of statistical tests comparing the DMWS-data with its surrogates using both configurational and dynamical characteristics as the test statistic. These evidences strongly indicate that the results reported in the previous section are not an artefact of a stochastic system but of a system which is indeed deterministic with a low dimensional chaotic attractor. In the next chapter, we demonstrate that deterministic prediction tools can remarkably enhance the accuracy of the prediction which again is an evidence of the deterministic character of the wind speed data.

The long-term predictions of a deterministic but chaotic time series are prone to errors due to the sensitivity of the initial conditions. However, short-term predictions can be made with fairly good accuracy by carefully chosen methods adapted to the data. This will be investigated in the next chapter.

In a preliminary study by Sreelekshmi et al., (2012) using wind speed data from a single location indicated that apparent random-like fluctuations of wind speed data are deterministic. In this chapter we confirm their observations that the underlying dynamics of wind speed fluctuations is *deterministic, low-dimensional and chaotic* by carrying out a detailed analysis of the daily mean wind speed measured at 9 locations across length and breadth of the Indian subcontinent for a period of more than 10 year using tools of nonlinear time series analysis. The results could have remarkable advantage power management in a wind farm.

3

Deterministic Prediction of Surface Wind Speed Variations

Accurate prediction of wind speed is an important aspect of various tasks related to wind energy management such as wind turbine predictive control and wind power scheduling. The most typical characteristic of wind speed data is its persistent temporal variations. Most of the techniques reported in the literature for prediction of wind speed and power are based on statistical methods or probabilistic distribution of wind speed data. In this chapter we demonstrate that deterministic forecasting methods can make accurate short term predictions of wind speed using past data, at locations where the wind dynamics exhibit chaotic behaviour. The predictions are remarkably accurate up to 1 hour with a normalised RMS error of less than 0.02 and reasonably accurate up to 3 hours with error of less than 0.06. Repeated application of these methods at 234 different geographical locations for predicting wind speeds at 30 days intervals for 3 years reveals that the accuracy of prediction is more or less the same across all locations and time periods. Comparison of the results with f-ARIMA model predictions shows that the deterministic models with suitable parameters are capable of returning improved prediction accuracy and capturing the dynamical variations of the actual time series more faithfully. These methods are simple and computationally efficient and require only records of past data for making short term wind speed forecasts within practically tolerable margin of errors.

3.1 Introduction

Wind is widely recognised as a clean, economically viable and eco-friendly source of electric power. Unlike power produced from coal or nuclear energy, wind power production is safe for the environment since it does not produce any greenhouse gases or harmful by-products. Wind is produced by the uneven heating of earth's surface by the sun and is therefore an inextinguishable source of energy. The generation of wind power has increased steadily over the last few years all over the world and as of 2011 the world wide installed capacity of wind power stands at 237 GW. It is estimated that, by 2020, more than 12% of the total demand for electricity could be met from wind energy resources (GWEC, 2012).

Nowadays in many countries, wind energy is being connected to existing electric power grids along with traditional sources. Wind-powered electricity must be used as soon as it enters the

grid, and to determine the additional amount of power to order from traditional sources to meet the demand at the grid, it is important to be able to predict wind power to the order of several minutes to a few hours in advance (Hering et al., 2010). Predictions of wind power, of the order of a couple of hours to a day ahead, are also crucial in liberalised electricity markets where expected power production and market prices are used in devising best bidding strategy with minimum possible risk (Gomes et al., 2012). Short term predictions, ranging from seconds to a few minutes, are useful in operation control of wind turbines and improvement of the power quality of wind farms (Wang et al., 2012).

It is clear from the above discussion that predicting wind power, in the range of at least a few hours ahead, is important both for optimising the performance of wind turbines and for maintaining a cost-effective power distribution system. Wind power is a function of wind speed, so an accurate prediction of wind speed leads to improved predictions of wind power in a given wind farm. For a range of wind speeds, the amount of wind energy produced from a wind turbine is proportional to the cube of wind speed, so any small improvement in short term predictions of wind speed can significantly improve predictions of wind power (Hering et al., 2010). However, with its dependence on topography, climate, seasonal changes, temperature, pressure and a host of other factors and its highly variable and random nature, wind speed is one of the most difficult meteorological parameters to predict. The literature on the various methods of predicting wind speed has grown extensively in recent years, especially in the wake of large scale deployment of wind farms across the globe.

The simplest among the prediction schemes for wind speed is the method of persistence which is based on the assumption that over very small time intervals wind speed does not change appreciably. Its usage is very limited, but is still used in the industry for making very short term predictions (Soman et al., 2010). The physical models, which utilise data of various atmospheric parameters to build up complex mathematical models, furnish another classical way to forecast wind speed. They are useful in identifying recurring patterns and making long term predictions when weather conditions are stable, but the prohibitive computational volume involved in solving such models renders them unreliable for short term predictions (Potter et al., 2006; Candy et al., 2009). Models which use statistical methods for wind speed predictions are also popular in the literature. They include moving average models such as ARMA, ARIMA and its variants fitted to the time series of wind speed (Kamal et al., 1997; Cadenas et al., 2007; Kavasseri et al., 2009) and models based on probability distribution of wind speed (Hennessey Jr, 1977; Celik, 2004; Mathew et al., 2011; Jiang et al., 2013). These models are fairly good in very short term predictions, but do not improve significantly on prediction error compared to the elementary method of persistence. Models based on artificial neural networks, which emulate the parallel distributed processing of human nervous system to adapt by learning from past data, have also been developed by many researchers for making short term predictions of wind speed and power (Mohandes et al., 1998; Cadenas et al., 2007; Bilgili et al., 2007; Monfared et al., 2009). In general, these models outperform the time series models in short term predictions, but their performance edge is not maintained across all locations universally (Soman et al., 2010). Recently researchers have also begun to use hybrid models, which combine different approaches for better forecasting results, such as mixing physical and statistical models or short-term and medium-term models (Soman et al., 2010; Liu et al., 2014; Haque et al., 2013). The central idea of physical approach is to incorporate the physical considerations of local topography into the numerical weather prediction scheme by modelling the local wind profile possibly considering the atmospheric stability. For example, Cassola et al., (2012) use Kalman filtering technique applied to the output of a

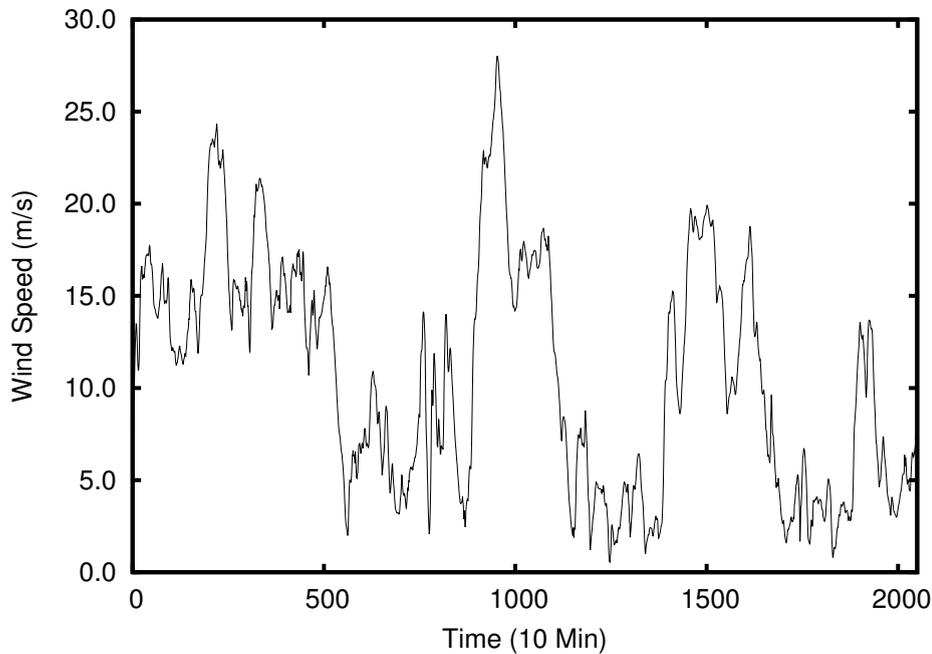


FIGURE 3.1: Time series of wind speed at location given by Latitude: 34.98420, Longitude: -104.03971 at 80 metres.

numerical weather prediction model to improve the accuracy of wind speed forecasts and wind energy output predictions significantly.

As a matter of fact, since the occurrence of wind is highly uncertain in time and space, no single technique can be used universally across all locations and time scales for predicting wind speed, and there is always scope for new methods. In a previous work (Sreelekshmi et al., 2012) we had carried out a detailed analysis of the time series of daily mean wind speed at Thiruvananthapuram, India, which revealed strong evidences for the existence of an underlying system which is deterministic, low-dimensional and chaotic. This means that the apparent random fluctuations found in wind speed data could originate from the chaotic dynamics of the underlying system, and not necessarily due to the system being stochastic as assumed in most of the aforementioned prediction theories, and this could also explain why wind speed predictions becomes erroneous beyond a certain time limit. However, provided wind speed dynamics is chaotic in a given location, we can use existing non-linear prediction schemes developed for chaotic time series to make more accurate short term predictions about wind speed. In this work we apply the methods of non-linear dynamics for forecasting wind speeds at various locations to get fairly accurate predictions up to 3 hours. For the analysis we have used the wind speed data of 10 minutes resolution of the period from January 2004 to January 2007 for 234 locations available from National Renewable Energy Laboratory (<http://www.nrel.gov>), USA.

3.2 Analysis of the data

Figure. 3.1 shows a plot of part of the wind speed data at location given by Latitude: 34.98420, Longitude: -104.03971 at 80 metres. We start with a detailed analysis of the underlying dynamics of wind speed data from this particular site and then move on to other locations. The irregular, random character of the data evident in the figure is typical of systems which are stochastic, but

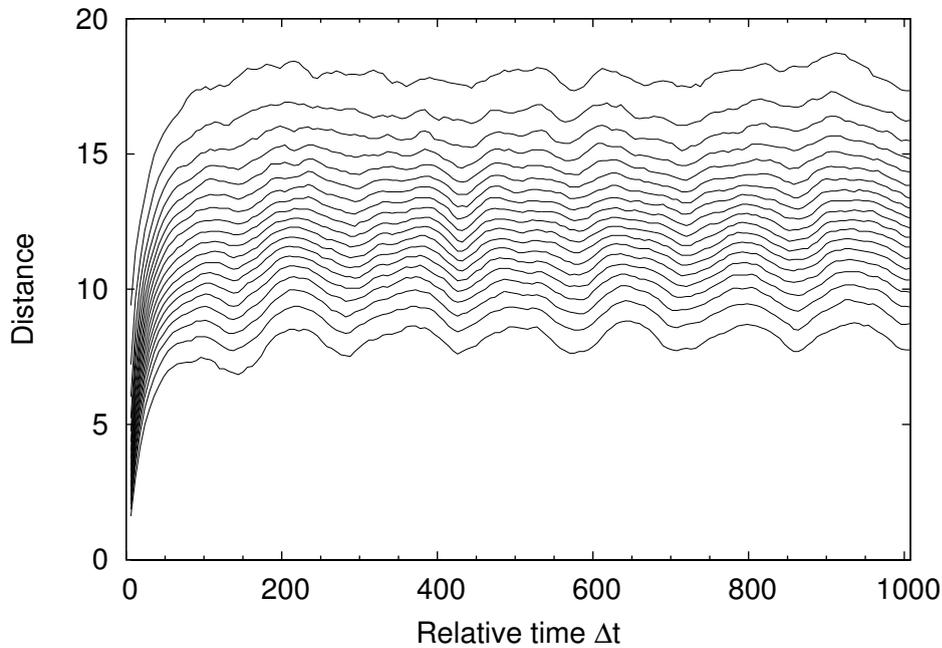


FIGURE 3.2: The Space time separation plot for the time series for $m = 14$ and $\tau = 85$. Each point in the plot represents a pair of points on the trajectory with their relative separation in time along the horizontal axis and separation in space along the vertical axis. The diurnal variations are evident in this figure.

as we have shown previously (Sreelekshmi et al., 2012), these fluctuations could also arise out of an underlying system which is deterministic, low-dimensional and chaotic. Since the stochastic versus deterministic character of wind speed has to be ascertained on a per location basis due to the highly uncertain nature of wind from location to location, we will now carry out a brief analysis of the data, referring the reader to Sreelekshmi et al., (2012) for the finer details of the features of chaotic systems and of the methods applied. The wind speed prediction schemes to be described later in coming sections are based on the results of this analysis.

The first step in the analysis of time series by methods of dynamical system theory is reconstructing the state space dynamics of the original system using the given time series data (Packard et al., 1980). This is done by constructing from the time series $x(t)$ a new vector time series $x(t)$ given by

$$x(t) = (x(t), x(t - \tau), \dots, x(t - (m - 1)\tau)), \quad (3.2.1)$$

where τ is a suitable multiple of the sampling time, called *delay*. Taken's embedding theorem and its extensions (Takens, 1981; Sauer et al., 1991; Sauer et al., 1993) assert that the dynamics of $y(t)$ in the reconstructed phase space will be topologically identical to the dynamics of the original system. In general, if the *cloud* of points generated by $y(t)$, called the *attractor*, fills out the m -dimensional phase space for all values of m , the time series may be considered to be generated by a stationary stochastic process. On the other hand, if the attractor occupies a region of small dimension, for all sufficiently large values of m , it may be an indication that the original system is deterministic and chaotic. On a chaotic attractor, nearby trajectories diverge with time exponentially fast, the rates of which are quantified by the *Lyapunov exponents* in the principal directions. A positive Lyapunov exponent is considered a signature of chaos. Due to the exponential divergence of trajectories, a chaotic attractor usually has a complex structure with a non-integral dimension. A chaotic time series exhibit broad-band spectrum (*c. f.* Figure. 3.7)

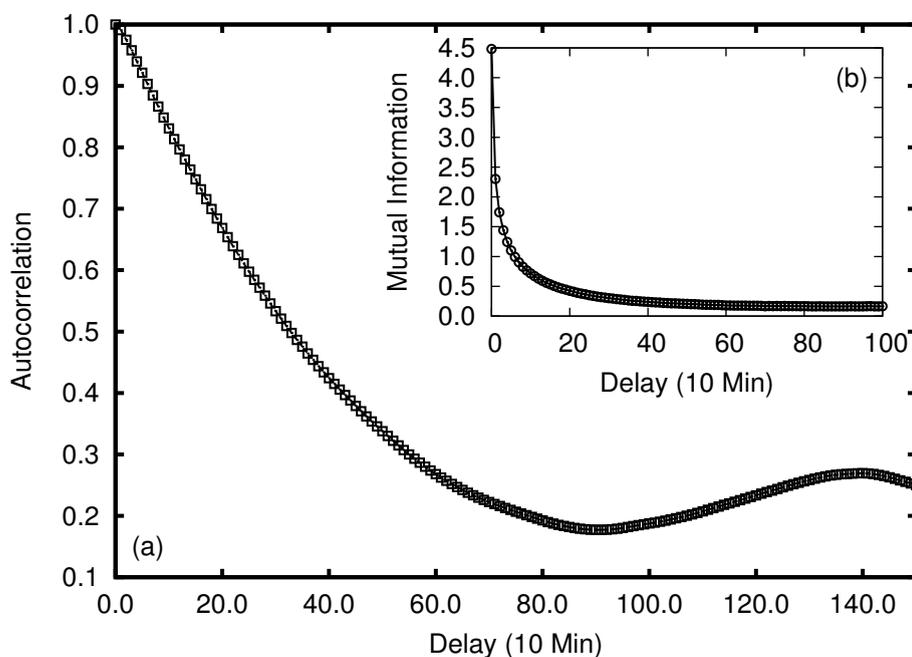


FIGURE 3.3: (a) The autocorrelation function of wind speed data. (b) Mutual information of the wind speed data as a function of delay (τ).

and other characteristics of random time series when analysed using linear stochastic tools, so a very detailed analysis is often necessary to distinguish a chaotic time series from a stochastic data. The space time separation plot given in Figure.3.2 of the time series helps us identify the temporal correlations within the time series and is useful in estimating a reasonable delay. The diurnal variations are evident in this figure. In order to reduce its modulation effects we removed the average diurnal variation by carrying out an *epoch analysis* as discussed by Kumar et al., (2004) and further analysis was carried out on this detrended time series.

In practical applications of embedding theorem, an optimal choice of the delay τ and the embedding dimension m are important. We have used the method of autocorrelation (Kantz et al., 2003) as well as the method of mutual information (Fraser et al., 1986) to arrive at a proper choice of τ . In Figure. 3.3 (a) and (b), which respectively plots the autocorrelation and mutual information of the wind speed data as a function of τ , the first minimum of the autocorrelation curve is observed at around $\tau = 85$ suggesting an optimal value $\tau = 85$ and the mutual information almost level off by $\tau = 75$, suggesting the value of τ can optimally be taken around this value. The value $\tau = 85$ is taken for further analysis. To determine the embedding dimension m , we used the method of false neighbours (Kennel et al., 1992; Kantz et al., 2003) which is based on the idea that in a large enough embedding dimension the fraction of false neighbours, which arise due to crossing of trajectories in a lower than true dimension, would be negligibly small. The fraction of false neighbours for the wind speed data is shown in Figure. 3.4, which suggests that $m = 14$ would be an optimal choice for the embedding dimension as the fraction attains a minimum around $m = 14$. However a precise knowledge of m is desirable only to exploit the determinism in the dynamics with minimal computational effort, and a large value of m will add redundancy and thus degrade the performance of many algorithms such as those for predictions (Kantz et al., 2003). It may also be noted that the parameter values obtained here are consistent with those in our previous analysis of wind speed data from a different geographical location (Sreelekshmi et al., 2012).

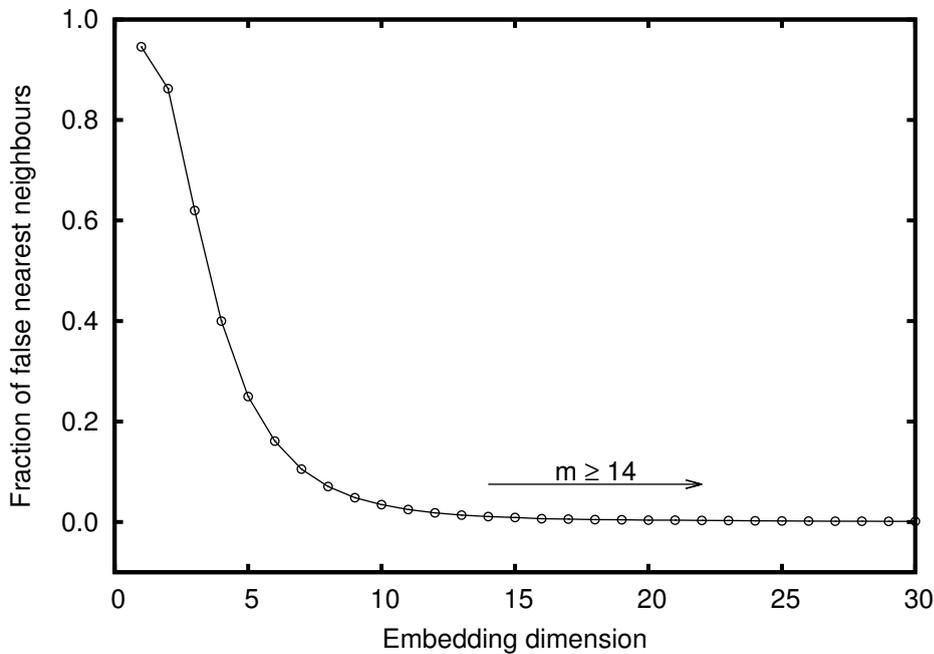


FIGURE 3.4: The fraction of false nearest neighbours as a function of the embedding dimension m for the time series of wind speed with $\tau = 85$, showing that any $m \geq 14$ can be considered optimal.

The dimension of the attractor gives a quantitative measure of the self-similarity of the attractor and also gives an idea of how large or small a region is occupied by the attractor within the embedding space. A standard dimension estimate for time series data is the correlation dimension, introduced by Grassberger et al., (1983), which proceeds by first computing the correlation sum defined by (Hegger et al., 1999)

$$C(\varepsilon, m) = \frac{2}{N(N-1)} \sum_{i=1}^N \sum_{j=i+1}^N \Theta(\varepsilon - \|y_i - y_j\|), \quad (3.2.2)$$

where $\Theta(a) = 1$ if $a > 0$, $\Theta(a) = 0$ if $a \leq 0$, and then the local slopes

$$D_2(\varepsilon, m) = \frac{d \ln C(\varepsilon, m)}{d \ln \varepsilon} \quad (3.2.3)$$

which are estimates of the correlation dimension. Figure. 3.5 plots $D_2(\varepsilon, m)$ versus ε for the wind speed data with the previous choice of delay and for embedding dimensions ranging from 14 to 16. The curves exhibit convergence onto a plateau for a range of values of ε , and the value corresponding to the plateau, $D_2 = 1.656 \pm 0.008$, is an estimate of the correlation dimension for the given data. The low dimensionality of the attractor is an indication of the deterministic character of the underlying dynamics exhibiting chaotic behaviour.

A chaotic system must have at least one positive Lyapunov exponent and to check for this one usually computes the largest Lyapunov exponent in the system which, if found positive, is considered a strong evidence for chaos. We have used the Kantz algorithm (Kantz, 1994) to estimate the largest Lyapunov exponent, which proceeds by computing the stretching factor $S(\Delta n)$, involving

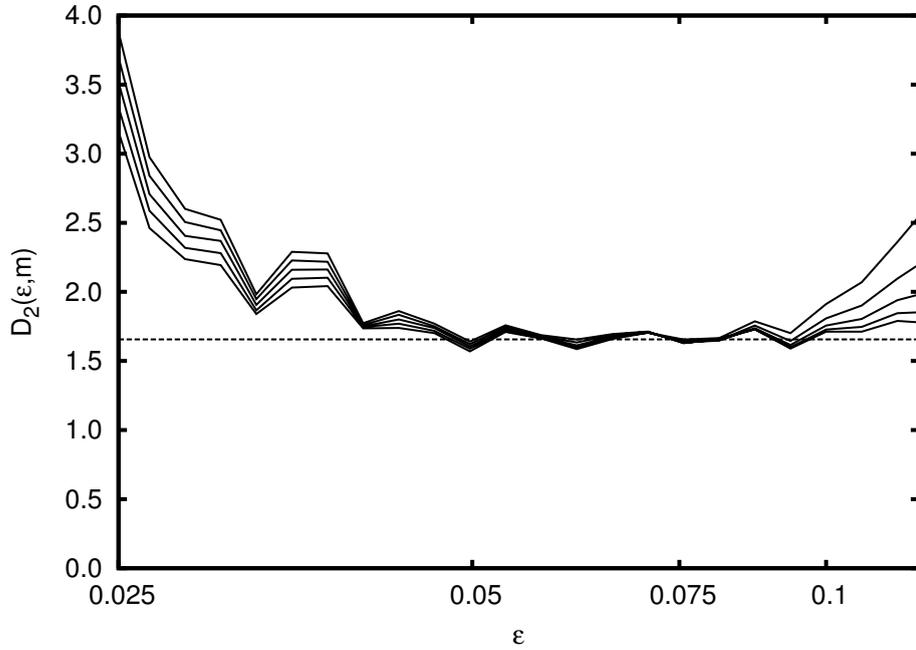


FIGURE 3.5: The local slopes $D_2(\varepsilon, m)$ for the wind speed time series with $\tau = 85$ and m ranging from 14 to 16 showing a plateau for small values of ε and giving an estimate of $D_2 = 1.656 \pm 0.008$.

a reference point y_{n_0} and its neighbours y_n in the embedding space over a neighbourhood $U(y_{n_0})$ of y_{n_0} , defined by

$$S(\Delta n) = \frac{1}{N} \sum_{n_0=1}^N \ln \left(\frac{1}{\|U(y_{n_0})\|} \sum_{y_n \in U(y_{n_0})} \|y_{n_0+\Delta n} - y_{n+\Delta n}\| \right) \quad (3.2.4)$$

For the wind speed data, Figure. 3.6 plots variations of $S(\Delta n)$ with Δn for $m = 14, 15$ and shows a linear growth in the range of $0 < \Delta n < 20$, the slope of which gives an estimate of the maximum Lyapunov exponent, in this case 0.15 ± 0.006 . The positive value of the largest Lyapunov exponent is another evidence for the chaotic dynamics of the underlying system.

The frequency decomposition of the time series, the power spectrum, is broadband and exhibits exponential decay as shown in Figure. 3.7, and this is an indication of the chaotic behaviour of the time series. The first part of the spectrum decays abruptly with an estimated value of -120.29 for the exponent while the second and third parts exhibit slow decay with values -15.80 and -4.07 for the exponent. This could possibly be caused by qualitatively different mechanisms at work in the dynamics of the underlying system.

3.3 Predictions

Due to the random nature of wind speed data, most of the wind speed predictions assume that the data is a realisation of a stochastic process. However, as we have shown in the previous section, the cause of randomness in the wind speed data can also be low-dimensional chaos, in which case

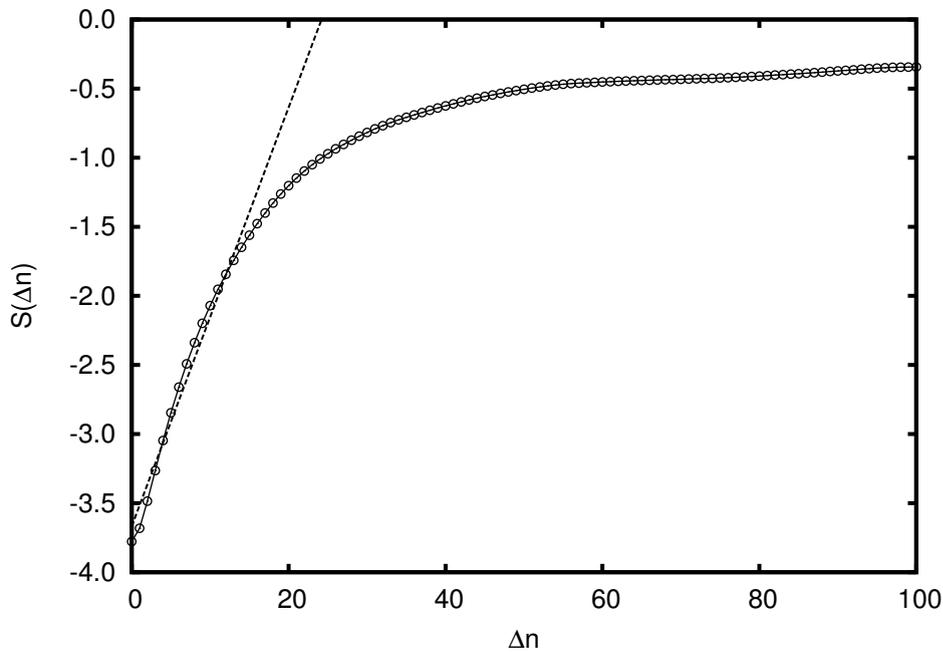


FIGURE 3.6: The curve of $S(\Delta n)$ for embedding dimensions $m = 14, 15$. The maximum Lyapunov exponent of the time series is the slope of the dashed line 0.15 ± 0.006 .

there is a fundamental limit on long term predictions. However, accurate short term predictions can still be made, taking advantage of simple deterministic relationships which may exist within the data, involving a few degrees of freedom. The extent of predictability depends on the local factors affecting wind speed, and in the data we considered accurate predictions up to 3 hours could be made using non-linear prediction tools.

Given a time series x_1, x_2, \dots, x_n , the forecasting methods try to predict values a few time steps ahead, namely x_{n+k} for $k = 1, 2, \dots$. Non-linear forecasting methods are based construction of delay vectors x_n from the time series using eq. (3.2.1);

$$x_n = (x_n, x_{n-\tau}, \dots, x_{n-(m-1)\tau}), \quad (3.3.1)$$

for a suitable integral delay τ and embedding dimension m . If deterministic rules govern the system, we expect a functional relation between x_{n+1} and x_n ;

$$x_{n+1} = F(x_n) \quad (3.3.2)$$

which in delay embedding reduces to

$$x_{n+1} = f(x_n). \quad (3.3.3)$$

If the dynamics is chaotic, F will be non-linear and we try to approximate F or f in various ways to predict an estimate \hat{x}_{n+k} for the actual value x_{n+k} . In chaotic systems these approximations are facilitated by an important property, that a sufficiently long time series produces a sequence of vectors that are dense on the attractor, so that new vectors will be arbitrarily close to some of those already observed. The approximation schemes used for predictions of chaotic time series can be broadly classified into *local* methods and *global* methods.

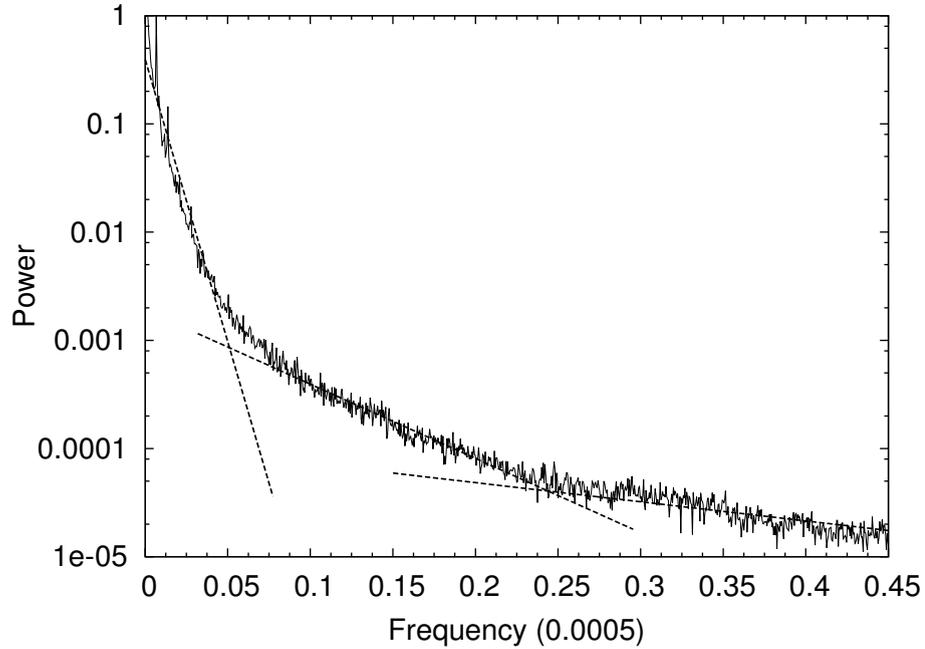


FIGURE 3.7: The power spectrum of wind speed time series as a function of frequency at location Latitude: 34.98420 Longitude: -104.03971. Broadband and exponential decay of the power with frequency are typical characteristics of chaotic signals.

The *local approximation schemes* try to approximate f locally, probably by a different function in each time step, by looking for vectors in the past which are close to x_n in the embedding space and using their future for prediction. The simplest of these is the *zeroth-order approximation*, which uses the average of the futures of the neighbours of x_n in an ε neighbourhood $U_\varepsilon(x_n)$. If there are N neighbours in $U_\varepsilon(x_n)$, the prediction is simply (Kantz et al., 2003)

$$\hat{x}_{n+1} = \frac{1}{N} \sum_{x_j \in U_\varepsilon(x_n)} x_{j+1} \quad (3.3.4)$$

A better method is the *local first order (LFO) approximation*, where instead of taking the average of the neighbours in $U_\varepsilon(x_n)$, a linear model is fitted to these neighbours, so that the prediction takes the form (Kantz et al., 2003)

$$\hat{x}_{n+1} = A_n x_n + b_n. \quad (3.3.5)$$

These local linear models, one for each time step, together generate a non-linear model globally.

The *global models* of prediction try to approximate F by a single function on the whole attractor. One of the popular global models used in predictions of chaotic time series is the *radial basis function (RBF) model* introduced by Lowe et al., (1988). In this, the approximating function F is taken as a linear superposition of a set of radial basis functions $\Phi_i(r)$, with $r > 0$, which are typically bell-shaped with maximum at $r = 0$ and rapidly decaying towards zero with increasing r . For a set of suitably chosen points y_i , called the *centres*, which are reasonably well distributed

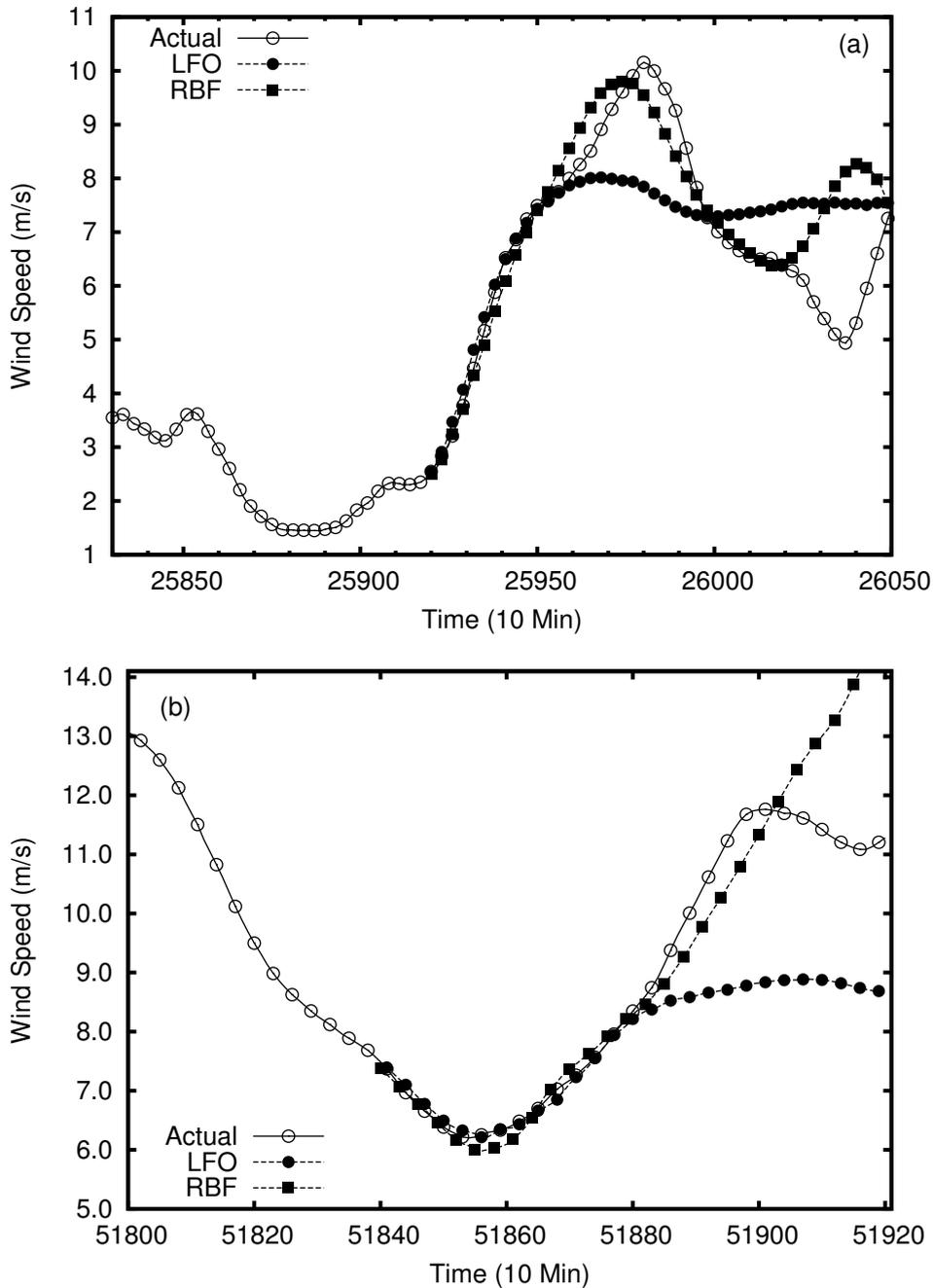


FIGURE 3.8: Comparison of predicted values with the actual values for LFO and RBF. The symbols are plotted only at every 30 minutes for legibility. (a) Latitude: 42.31925 Longitude: -98.60197 $m = 5, \tau = 8$ (b) Latitude: 43.51076 Longitude: -99.47652, $m = 12, \tau = 2$,

on the attractor, the model assumes the form (Kantz et al., 2003)

$$F(x) = \alpha_0 + \sum_{i=1}^p \alpha_i \Phi(\|x - y_i\|). \quad (3.3.6)$$

The basis functions Φ are modelled using Gaussians with their number and width kept fixed throughout the model. This makes the estimation of the constants α_i a linear problem which can be solved using least square method.

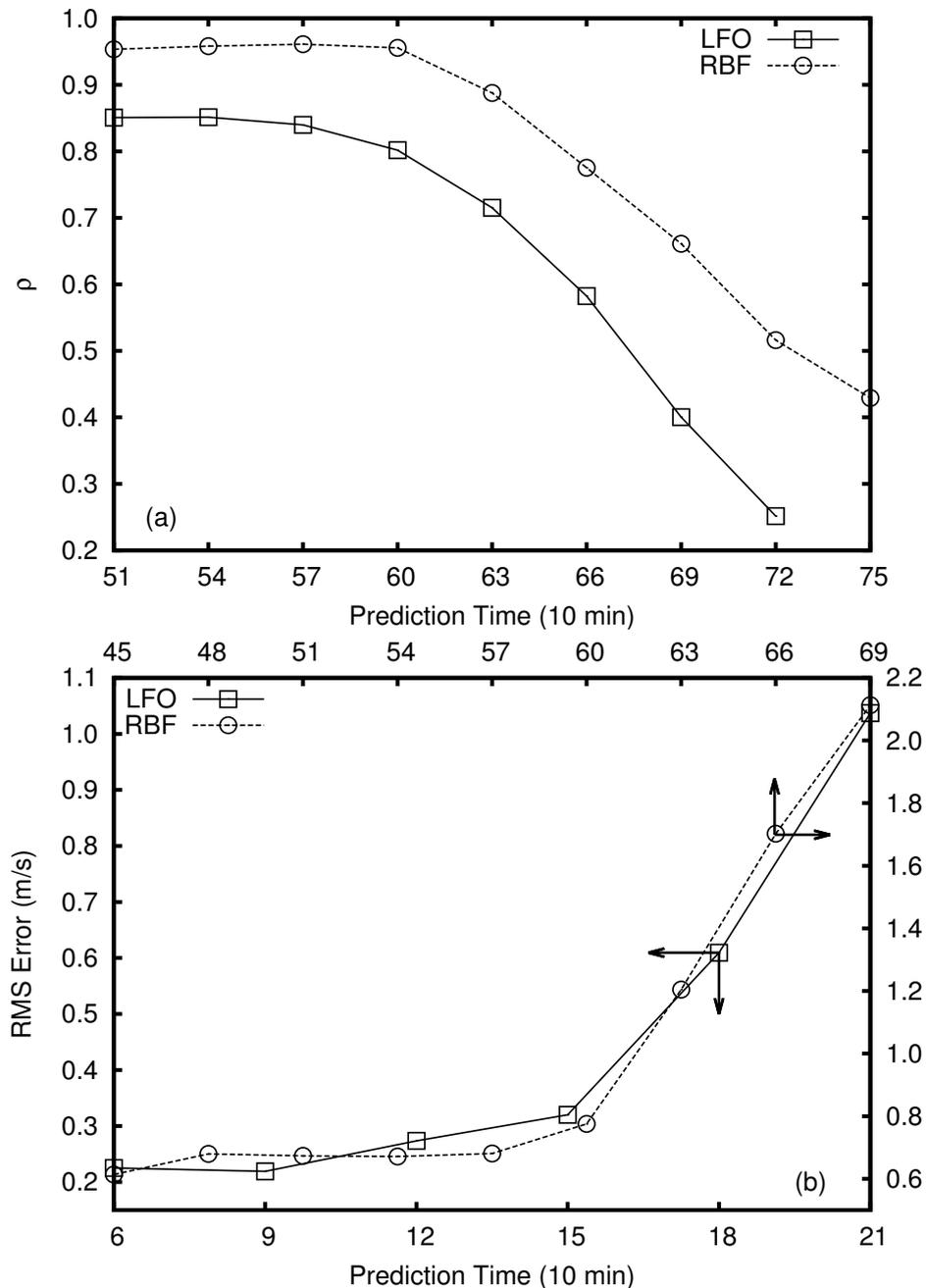


FIGURE 3.9: (a) The exponential decay of the correlation coefficient between the predicted and actual for every 3 prediction time steps (b) The exponential growth RMS error prediction for every 3 prediction time steps. The arrows show axes for the respective symbols of data points.

The advantage of LFO method is its flexibility, but it may not yield desirable performance on parts of the phase space where the points do not span the available space dimensions. On the other hand, global models have the advantage of providing the structure and properties of the underlying system as it can yield closed expressions for the full dynamics. These models can effectively describe the observed process in regions of the space which have been visited by the data, but outside this area, the shape of the model depends heavily on the chosen function (Hegger et al., 1999).

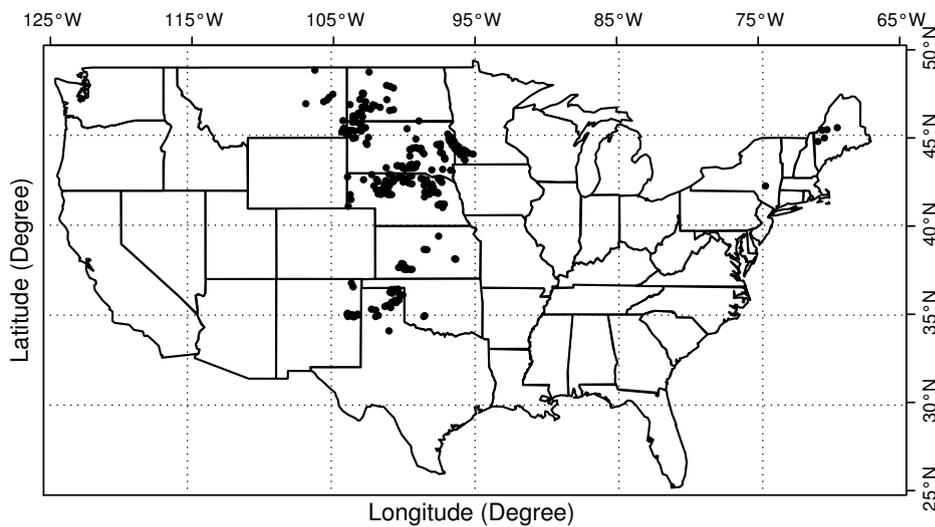


FIGURE 3.10: The geographic locations, denoted by filled circles, where wind speed data was analysed for performance of deterministic model prediction.

Figure. 3.8(a) and (b) shows typical results of wind speed prediction using the above methods, made at a couple of locations for suitable choices of the embedding parameters, and their comparison with actual data. Each prediction uses the available wind speed data for the location up to a specific point of time for modelling, and then employs the model for predicting future values. It is seen from these figures that, provided we use appropriate embedding parameters, the deterministic methods can predict wind speed with remarkable accuracy up to 3 hours, with the RBF method giving fairly accurate predictions for another 15 hours in both the cases. A method for determining the optimal embedding parameters for prediction is discussed in the next section.

Figure. 3.9(a) quantifies the similarity of predicted values with the measured data for a typical case, by plotting the statistical coefficient of correlation between the predicted and actual values as a function of the number of time steps into the future. The correlation coefficients were calculated cumulatively, at the end of every 3 prediction time steps, using all the predicted values available up to that time and the corresponding measured values. The exponential deterioration of the correlation with increasing prediction time is a characteristic feature of deterministic chaos (Sugihara et al., 1990) and provides further evidence of the fact that the erratic fluctuations in wind speed data are caused by the chaotic dynamics of the underlying system and are not an artefact of uncorrelated additive noise. Figure. 3.9(b) shows how the root mean square (RMS) error between the predicted and measured values, again calculated cumulatively every 3 time steps, propagates as we predict further into the future. The exponential growth of the prediction error further substantiates the chaotic nature of the data.

3.4 Statistical analysis of prediction errors

To demonstrate the wider applicability of the deterministic methods for making short term wind speed forecasts we now carry out a statistical analysis of the prediction errors for forecasts made at a total of 234 geographical locations. For the analysis we have considered 10 minutes interval wind speed data for 3 years from 2004, available from National Renewable Energy Laboratory, USA for the 234 locations depicted in Figure. 3.10.

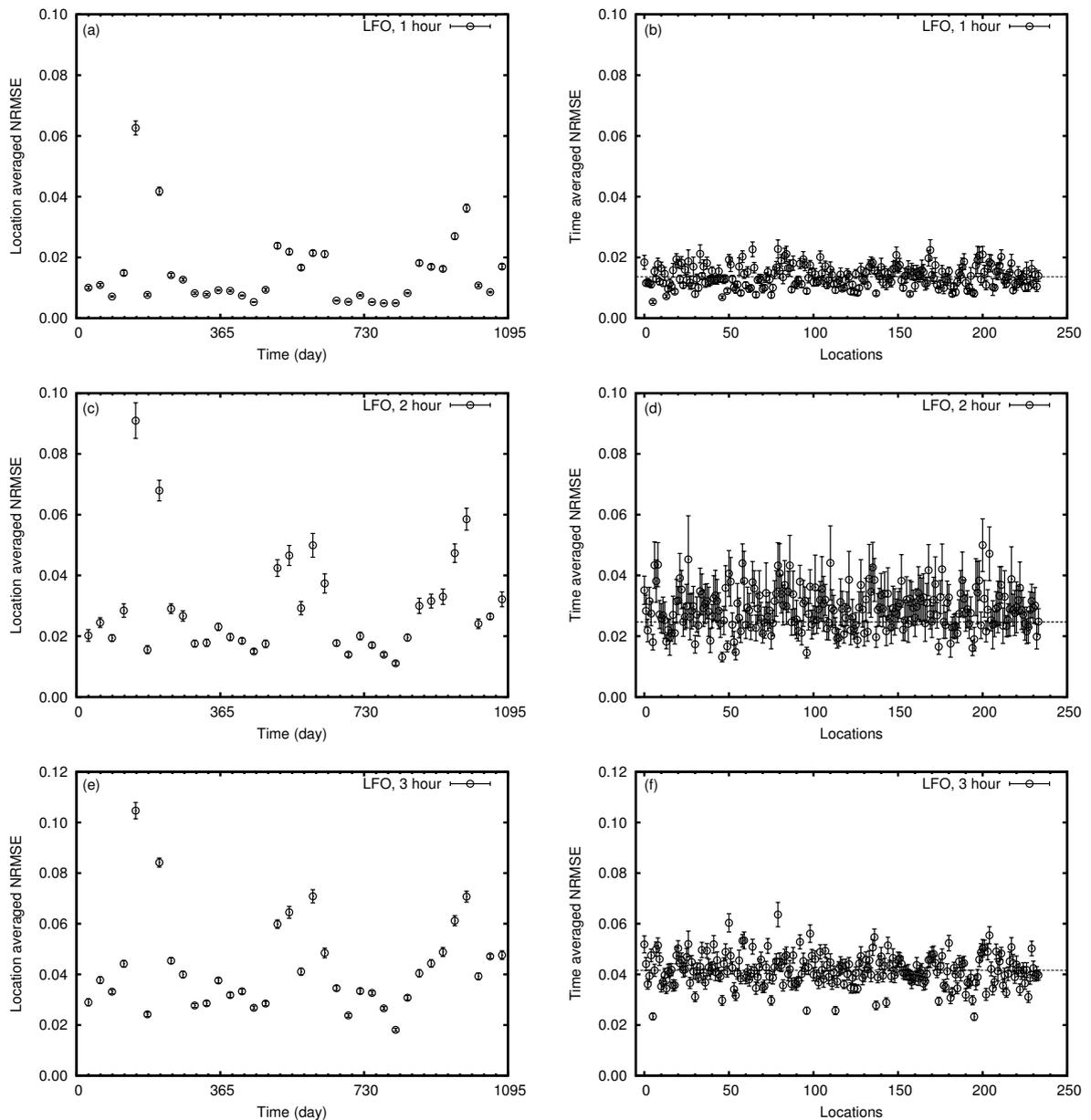


FIGURE 3.11: (a) NRMSE of 1 hour predictions using LFO, for the period from 2004 to 2006 at an interval of 30 days, averaged over 234 locations. (b) NRMSE of 1 hour predictions using LFO for the 234 locations, calculated at an interval 30 days and averaged over the period from 2004 to 2006. (c) & (e) are location averaged NRMSE for 2 and 3 hour predictions respectively. (d) & (f) are time averaged NRMSE for 2 and 3 hour predictions respectively. The error bars in the figures are with respect to the standard error of the mean. The horizontal dotted lines in (b), (d) and (f) represent the mean of the respective time averaged NRMSE values over the entire set of locations.

The optimal choice of the embedding parameters m and d is a major factor affecting the accuracy of prediction. Since the dynamics of wind speed varies over locations, these parameters have to be determined for each location separately. However, the embedding parameters suggested by the autocorrelation function or the fraction of false neighbours need not always give the most accurate predictions (Domenico et al., 2013) and a systematic procedure for determining the most suitable parameters for predictions using a model data is still elusive. For the present analysis, to fix the

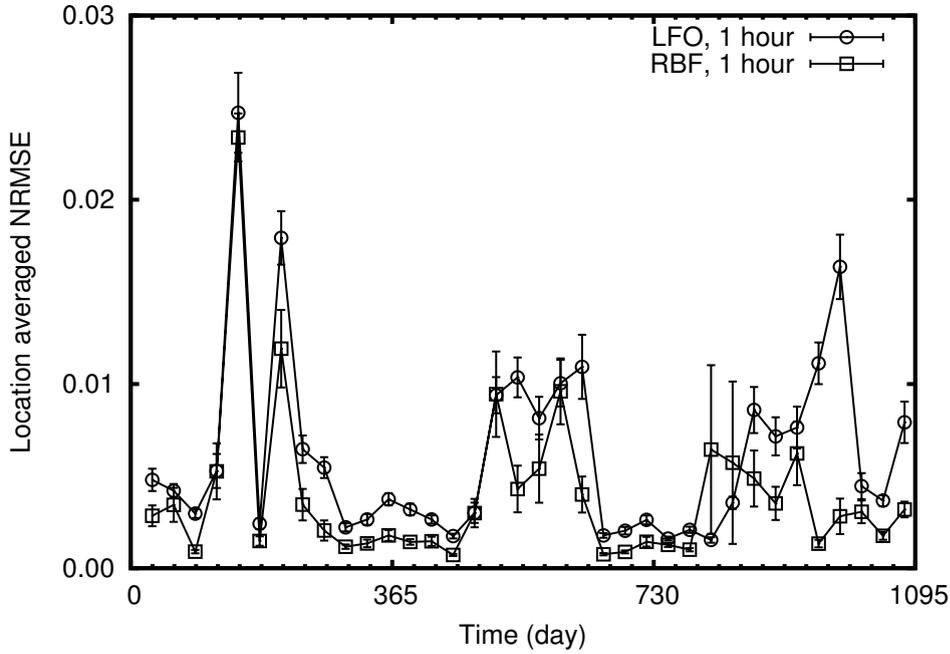


FIGURE 3.12: NRMSE, with standard error, of 1 hour prediction for the period of 3 years from 2004 over 234 locations using LFO and RBF. The lines connecting the symbols are to guide the eye.

optimal parameters at each location, we have used a *test run* procedure as described below. Given a time series of n points x_1, \dots, x_n , we want to predict the next k values x_{n+1}, \dots, x_{n+k} . From the given data set, we take the first $n-3$ data points to form a model for making a test prediction of the next three data points for various values of m and d using one of the deterministic algorithms described earlier. These predicted values $x_{n-2}^p, x_{n-1}^p, x_n^p$ are then compared with the actual data points x_{n-2}, x_{n-1}, x_n to find the RMS error. The values of m and d which yield the minimum RMS error for these three predicted values are selected as the parameter values for the given data set and used for the prediction of values of x_{n+1}, \dots, x_{n+k} .

We use spatial averages of prediction errors over various locations as well as time averages at each location to assess the performance of these methods. Since the range of values of wind speed vary over locations, we have chosen as a measure of the prediction error the root mean squared error normalized over the range of the observed data (NRMSE) given by

$$\text{NRMSE} = \sqrt{\frac{\sum_{n+1}^{n+k} (x_i - x_j^p)^2}{k}} / (x_{\max} - x_{\min}) \quad (3.4.1)$$

where x_j^p are the predicted values. For each location we have calculated NRMSE for 1 hour, 2 hour and 3 hour predictions at intervals of 30 days for a 3 year period from 2004 to 2006. Figure. 3.11(a) depicts NRMSE with error bar for 1 hour predictions averaged over the locations (location averaged NRMSE), computed at 30 days intervals and plotted for 3 year period. Figure. 3.11(b) shows NRMSE for 1 hour prediction for each location averaged over a 3 year time period (time averaged NRMSE) where the horizontal dotted line shows the mean 0.0136 of these values over the locations. Similar estimates of location and time averaged errors for 2 hour predictions are given in Figures.3.11(c) and (d) and for 3 hour predictions in (e) and (f). The mean of

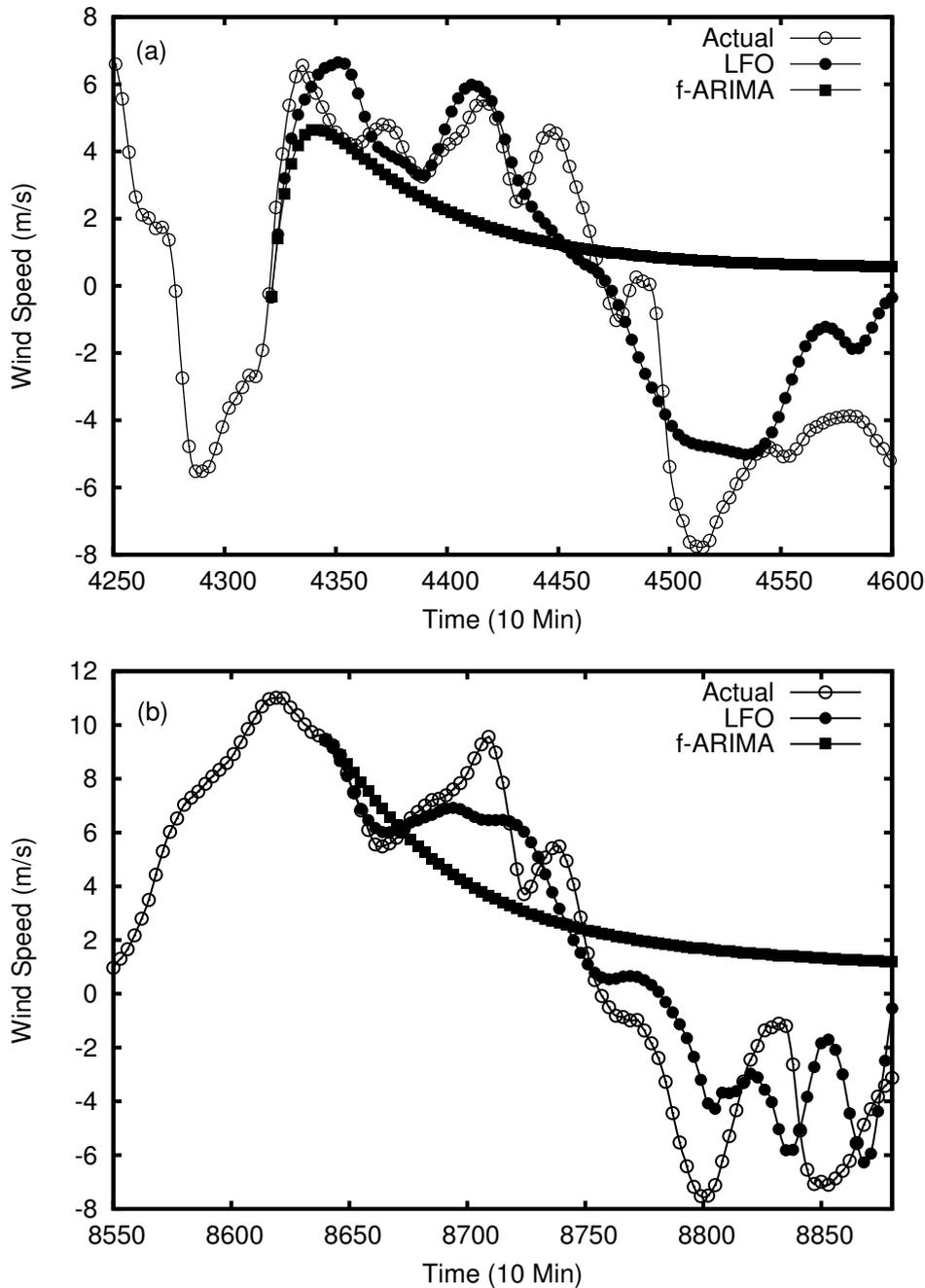


FIGURE 3.13: Comparison of predicted values with the actual values for LFO and *f*-ARIMA. The symbols are plotted only at every 30 minutes for legibility. Latitude: 34.98420, Longitude: -104.03971, (a) $m = 13, \tau = 3$ and (b) $m = 8, \tau = 7$.

the time averaged NRMSE over the locations, indicated in each figure by a dotted line, is 0.0299 for 2 hour predictions and 0.0415 for 3 hour predictions. This shows that the prediction accuracy observed in the typical forecasts shown in Figure. 3.8 are more or less maintained across all locations and various time periods consistently.

Between the deterministic methods, the RBF maintains a consistently slight performance edge over LFO for short term predictions up to one hour, as is clear from Figure. 3.12, which compares the location averaged prediction errors for 1 hour predictions by the two methods.

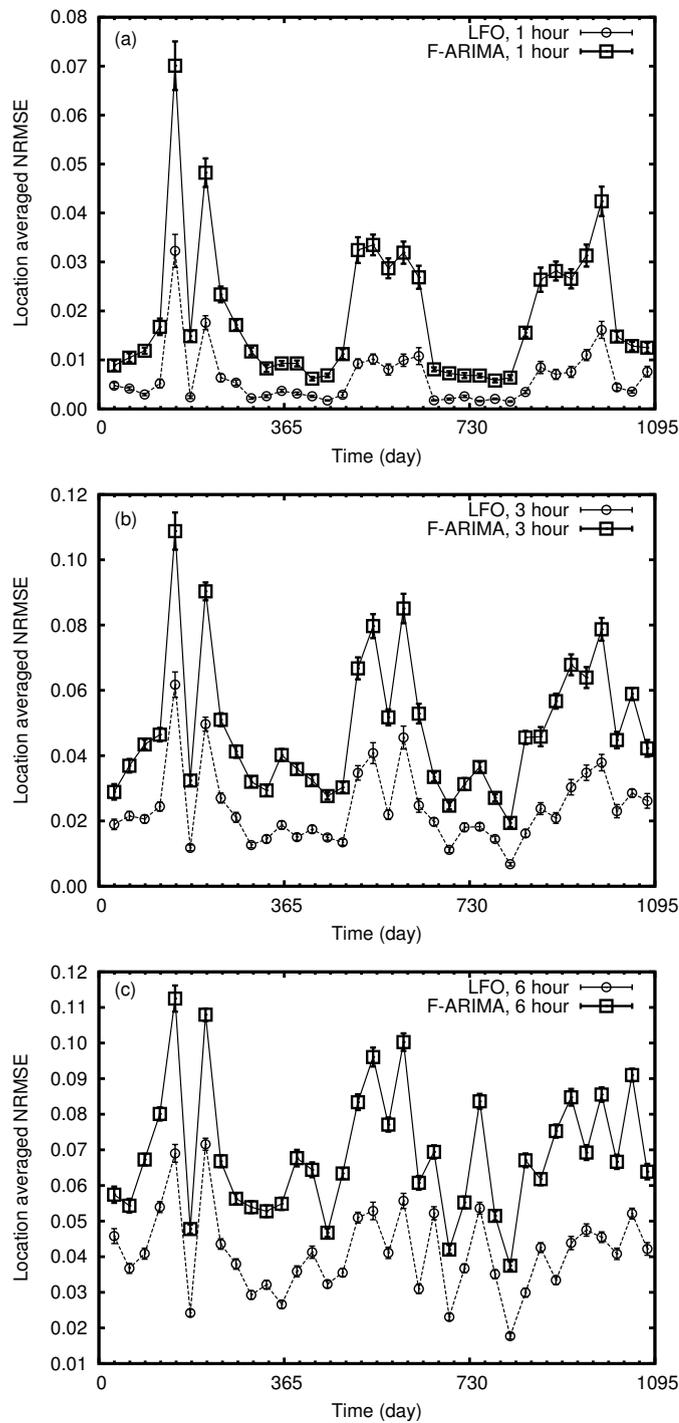


FIGURE 3.14: Location averaged NRMSE, with standard error, of (a) 1 hour, (b) 3 hour and (c) 6 hour predictions for the period of 3 years from 2004 over 234 locations using LFO and *f*-ARIMA.

3.5 Comparison with *f*-ARIMA

Among the various statistical methods used in wind speed prediction ARIMA is a popular model which gives reasonably accurate predictions of wind speed at many locations (Kamal et al., 1997; Cadenas et al., 2007; Kavasseri et al., 2009). An ARIMA(p, d, q) model combines an autoregressive(AR) process of order p , a moving average(MA) process of order q and a differencing

operator of order d into a single model. It has the general form (Box et al., 2013)

$$\Phi(B)\Delta^d x_t = c + \Theta(B)\varepsilon_t \quad (3.5.1)$$

where ε_t is a white noise process and $\Phi(B)$ and $\Theta(B)$ are respectively the autoregressive and moving average operators defined by

$$\begin{aligned} \Phi(B) &= 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p \\ \Theta(B) &= 1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q \end{aligned} \quad (3.5.2)$$

for suitably chosen constants θ_i and ϕ_i and non-negative integers p and q . B is the backward-shift operator so that $Bx_t = x_{t-1}$ and $\Delta = 1 - B$ is the differencing operator and in the general ARIMA model d is an integer.

f -ARIMA is a generalisation of ARIMA where the parameter d is allowed to have a fractional value with the operator $(1 - B)^d$ interpreted to have the binomial expansion (Hosking, 1981)

$$(1 - B)^d = 1 - dB + \frac{d(d-1)}{2!}B^2 + \dots \quad (3.5.3)$$

The possibility of wide range of choices for the parameters p , q , d and the constants ϕ_i and θ_i give the model great flexibility and wider applicability.

One of the features that distinguishes a f -ARIMA process from an ARIMA process is that the former is characterised by a slow decay in its autocorrelation function compared to the latter. This feature makes f -ARIMA model an attractive choice for data sets that exhibit long range correlations such as the wind speed data (Kavasseri et al., 2009).

General characteristics of the predictions by f -ARIMA and how they compare with the predictions by LFO can be seen from fig.3.13(a) and (b). The performance of f -ARIMA is comparable to LFO initially but its predictions deviate from actual values and level off to a steady value after a brief period. In contrast, while the accuracy of prediction of LFO also falls off gradually after 3 hours, it nevertheless captures the essential dynamics of the original time series even further.

For comparing the performance of LFO versus f -ARIMA, we have elected to generate the best possible predictions by both the methods, by experimenting with various values of the parameters which determine the accuracy of prediction. Thus, for a model data set x_1, \dots, x_n , we would generate several trial predictions for the next k data points using various parameter values, compare each of them with the actual observed data x_{n+1}, \dots, x_{n+k} , and choose the one that gives the least prediction error. For the LFO method this might yield better predictions than would be obtained with the embedding parameters selected by the procedure described in the last section. While the latter procedure would be useful in real world applications where there are no future data to compare the predictions with, it need not always give the optimum parameter values giving the most accurate predictions. In fact, we have observed that the LFO predictions obtained here (Figure. 3.14(a), (b)) are marginally better than those in Figure. 3.11(a) and (b), with the location averaged NRMSEs being smaller by 0.8% and 1.8% on the average for 1 and 3 hour predictions respectively.

Figure. 3.14 shows the results of a statistical analysis of the performance of LFO and f -ARIMA, with optimum parameter values, over all the 234 locations described earlier. The figures (a), (b) and (c) compare NRMSE for 1, 3 and 6 hour predictions averaged over all locations computed in intervals of 30 days for a 3 year period. The prediction accuracy of LFO is noticeably better than that of f -ARIMA across all locations and all time periods. For low resolution wind speed data of the kind considered in this work, the accuracy and the longevity of the predictions obtained by the deterministic methods are therefore a significant improvement over existing methods.

3.6 Conclusions

In this work we demonstrate the suitability of deterministic methods in making short term forecasts of wind speed based on past data. These methods are applicable in situations where the underlying dynamics of wind speed is chaotic leading to random like fluctuations in the time series of wind speed. We have applied a couple of chaotic time series prediction tools (one local method and one global method) on records of wind speed data of 10 minute resolution from a total of 234 different geographical locations, at each location making 1 hour, 2 hour and 3 hour predictions at intervals of 30 days for a period of 3 years. The predictions are very accurate for up to 1 hour and fairly accurate for up to 3 hours. A statistical analysis of the prediction errors from these locations reveal that the average prediction error is 1.36% of the range of wind speed for 1 hour predictions, 2.99% for 2 hour predictions and 4.15% for 3 hour predictions.

We have also compared the efficiency of the deterministic methods with predictions by f-ARIMA at each of the above 234 locations on the basis of 6 hour predictions at intervals of 30 days for a period of 3 years. It is seen that, compared to f-ARIMA, the deterministic methods give better prediction accuracy for longer periods of time and capture the dynamics of the fluctuations in the original data more faithfully. These prediction methods are simple and computationally efficient alternatives for short term wind speed forecasts.

4

Empirical mode decomposition and chaos based prediction model for wind speed oscillations

Accurate short-term prediction of wind speed is one of the critical issues faced by wind farm industry so as to plan their trading strategies and energy management. In this chapter we present an empirical mode decomposition (EMD) based chaotic model for short-term predicting wind speed. While EMD technique is used to decompose the measured wind speed time series data into its basic components called intrinsic mode functions (IMF) and residue, chaotic prediction tool is applied on each of them. Prediction result of each component is summed up to reconstruct into original form. Resultant prediction of this hybrid method is compared with that chaos model prediction of given data without decomposition. The comparison results show that prediction accuracy can be remarkably improved by combining EMD and chaos model.

4.1 Introduction

As a clean and cheapest power source, wind has already find its own position in worldwide energy market. According to World Wind Energy Association (WWEA) by 2014 the total worldwide wind capacity has reached 326GW meeting 4% of world's electricity demand (WWEA, 2015). Relying on the success stories of the traditional markets, new markets are emerging and Global Wind Energy Council (GWEC) expects by 2050, 20-30% of global electricity supply can be provided by wind power (GWEC, 2014). Although wind energy technologies have advanced significantly, growing importance of wind energy sector alerts researchers and policy makers to consider the essentials that may build and support a well defined, cost efficient wind system. For this purpose International Energy Agency (IEA) identified the four strategic research topic, characterize the wind resource, develop next generation wind power technology and wind integration for wind energy development (IEA, 2013). Accurate short-term wind speed forecasting, which comes under characterization of wind, is an important research area as it can contribute in reducing the wind power plant performance uncertainty. Prior knowledge about wind speed helps utility and wind farm operators to ensure grid stability and achieve favorable trading performances.

Numerous methods, including physical and time series models, are available in literature and practice for predicting wind speed on different time scales (Soman et al., 2010). Based on

methodology used, wind speed modeling techniques are broadly classified as physical models, statistical models, artificial intelligence based models and hybrid models. Prominent physical model, Numerical Weather Prediction (NWP) method, uses mathematical models of the atmospheric parameters like temperature, pressure, surface roughness and obstacles. Statistical such as ARMA, ARIMA models use historical time series data for identifying patterns to make predictions (Kamal et al., 1997; Cadenas et al., 2007; Kavasseri et al., 2009). Other prediction methods like artificial intelligence (AI) models and Hybrid models have also been developed and investigated (Soman et al., 2010; Mohandes et al., 1998; Bilgili et al., 2007; Haque et al., 2013). Although many modeling techniques have been proposed, persistence method which assumes present value as the forecasted is still considered as a bench mark in short-term wind speed prediction (Soman et al., 2010). Apart from these traditional methods it has recently been shown that apparent random oscillation of wind speed data is chaotic suggesting non-linear deterministic prediction schemes can be used for accurate short-term predictions (Sreelekshmi et al., 2012). In our previous work we have analysed the scope of deterministic methods to predict wind speed variations, and demonstrated that these methods are capable of reasonably accurate predictions up to a few hours (Drisya et al., 2014).

In this present work, an attempt is made to combine Empirical Mode Decomposition (EMD) and traditional non-linear time series analysis tools. Two methods are combined on the assumption that EMD shall decompose the data into a finite number of intrinsic mode functions (IMF) and then applying chaotic prediction tools on each IMF may better capture the deterministic character of oscillation giving a better prediction accuracy.

4.2 Methodology

4.2.1 Nonlinear time series and prediction tools

For a chaotic dynamical system, indirect measurement of the system's property is done from a single time series with the help of state space reconstruction (Packard et al., 1980). The reconstruction preserves the properties of the dynamical system. The method of delays reconstructs a new vector $x(t)$ from a measured time series $x(t)$ as

$$x(t) = (x(t), x(t - \tau), \dots, x(t - (m - 1)\tau)) \quad (4.2.1)$$

where τ is called delay or sampling time at which measurement is done and m is the embedding dimension. The embedding theorem of Takens and its extensions asserts, for almost all values of time delay and all smooth measurement functions embedding is valid as long as $m > 2D$ where D is the box-counting dimension of the attractor (Takens, 1981; Sauer et al., 1993). That means with proper reconstruction $x(t)$ preserves characteristics of the system invariants such as Lyapunov exponent, Entropies, Fractal Dimension etc. Non-linear forecasting methods uses reconstructed delay vectors and assumes a functional relation between x_n and x_{n+1} .

$$x_{n+1} = \mathbf{F}x_n \quad (4.2.2)$$

which in delay embedding reduces to

$$x_{n+1} = fx_n \quad (4.2.3)$$

For predicting x_{n+k} both global and local approximation methods are used. Global schemes try to approximate a single function F globally on the whole attractor while local approximation methods approximate F locally, by a different function in each time step. Although global models have the advantage of approximation function which capture the full dynamics, the method is very complex and in outside regions of the space which have been visited by the observed data it depends on the chosen function (Hegger et al., 1999). Because of its flexibility in capturing local behavior of the attractor we chose local approach to study and forecast. Local approximation method suggested by Lorenz was the first one and it approximates future value x_{t+1} by x_T , where x_T is the nearest neighbour in the state space (Lorenz, 1963). A simple modification to this is made by Kantz, known as zeroth order approximation, in which for a collection of N nearest neighbours of x_t within ε distance in phase space, use the average of its futures as the prediction x_{t+1} (Hegger et al., 1999). That means

$$\hat{x}_{n+1} = \frac{1}{N} \sum_{x_j \in U_\varepsilon(x_n)} x_{j+1} \quad (4.2.4)$$

A more accurate method in which a linear model is fitted to the neighbors $U_\varepsilon(x_n)$ of x_n known as local first order method Hegger et al., (1999) represented by

$$\hat{x}_{n+1} = A_n x_n + b_n \quad (4.2.5)$$

Local approximation methods reduces complexity and adapts the local behavior of the attractor precisely.

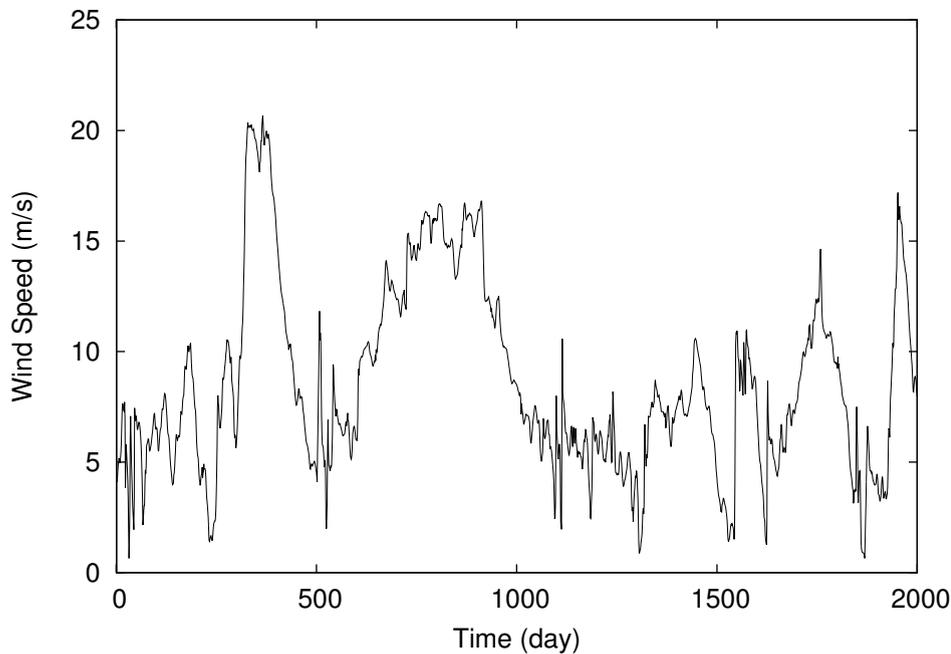


FIGURE 4.1: Time series of wind speed at location given by latitude: 47.11332° N longitude: 90.44666° W at 80 m.

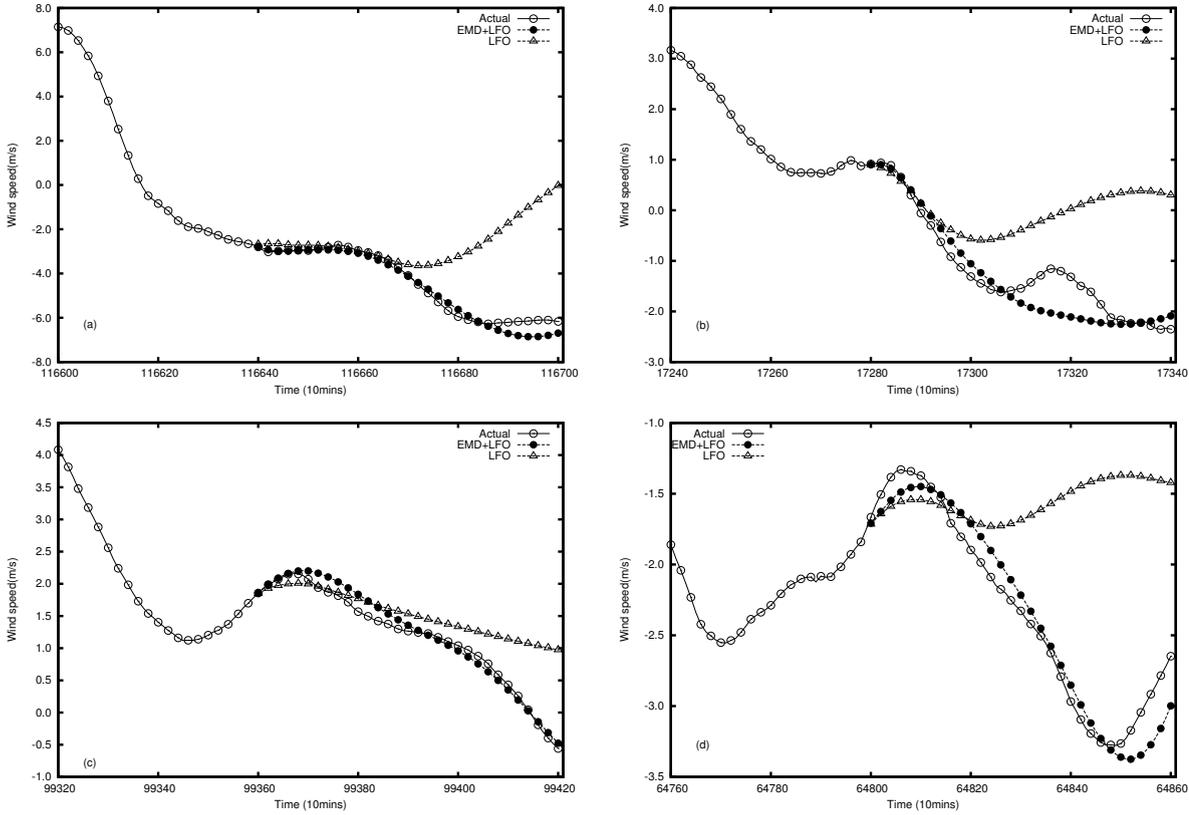


FIGURE 4.2: Comparison of predicted values with actual values for EMD decomposition and without decomposition for four different locations (a) Latitude: 34.9842° N Longitude: 104.03971° W (b) Latitude: 41.85498° N Longitude: 97.61899° W (c) Latitude: 46.73381° N Longitude: 101.7077° W (d) Latitude: 37.84476° N Longitude: 100.06653° W

4.2.2 Empirical mode decomposition

Empirical mode decomposition was first used as a signal decomposition technique for Hilbert-Huang Transform (HHT) (Huang et al., 1998). EMD breaks down the signal $x(t)$ into various components based on its local characteristics and in order to obtain these local details it considers the two consecutive minima (or maxima). It may be noted that a maxima (or minima) exists between these two consecutive minima (or maxima). We can define a local high-frequency part or local detail $d(t)$ corresponding to the oscillation between the two minima (or maxima) passing through the maximum (or minimum) existing between them. One still has to estimate the local low-frequency part or local trend $m(t)$ so that $x(t) = m(t) + d(t)$ between the adjacent minima (or maxima) identified. Suppose that the local consecutive minima are connected by a cubic spline as lower envelope $x_{low}(t)$ and similarly local maxima are connected by a upper envelope $x_{sup}(t)$ and the mean of these two are obtained by

$$m_1(t) = (x_{low}(t) + x_{sup}(t)) / 2 \quad (4.2.6)$$

Then the approximate first Intrinsic Mode Function (IMF) is given by

$$d_1(t) = x(t) - m_1(t) \quad (4.2.7)$$

By treating $d_1 t$ as a new set of data, its upper and lower envelopes are obtained to find the mean $m_{11}(t)$ so that

$$d_{11}(t) = d_1(t) - m_{11}(t). \quad (4.2.8)$$

This process of *shifting* is repeated until the number of zero crossings of $d_{1k}(t)$ is equal to the number of extrema or different at the most by 1 from the number of extrema. The convergent result $c_1(t) = h_{1k}(t)$ with zero local mean is the first IMF.

The first residue $r_1(t) = x(t) - c_1(t)$ is then processed in the same way to obtain $c_2(t)$ so that $r_2(t) = x(t) - c_1(t) - c_2(t)$. The process is repeated till either $c_n(t)$ or $r_n(t)$ becomes smaller than a predetermined value so that

$$x(t) = c_1(t) + c_2(t) + \dots + c_n(t) + r_n(t) \quad (4.2.9)$$

4.3 Results and discussions

For analysis, 10 minutes resolution wind speed data of the period 2004-2007, measured at different geographical locations of USA has been used. The data was obtained from National Renewable Energy Laboratory (<http://www.nrel.gov>), USA. As part of preprocessing of data, a noise reducing technique was applied to reduce the effect of additive noise. This denoised data is used for further analysis. Figure.4.1 shows the apparent random behaviour of wind speed oscillations.

In our previous work we demonstrated the irregular fluctuations exhibited by wind speed data is originated from a low-dimensional, deterministic, chaotic system and hence application of non-linear prediction tools can give accurate short term prediction (Sreelekshmi et al., 2012). This was confirmed by our later analysis (Drisy et al., 2014). Since the system composed of various dynamic factors, it is assumed that better short to mid-term prediction could be possible if we decompose the signal into its basic components and then analysing each component separately. In order to assess this assumption, a simple interpolation smoothing based EMD method is applied to measured wind speed data for getting a family of frequency ordered IMF components and a residue. Decomposition is done with the help of EMD package available for R environment (Kim et al., 2009). A portion of each component is used to build a linear first order model for multi step ahead prediction. Aggregate all the predicted components to reconstruct into its original form. For comparison purpose a prediction model of linear first order model original time series without decomposition is also carried out.

Figure.4.2 depicts the general characteristics of the predictions by linear first order model and EMD based linear first order model model. The results from 4 different locations are plotted along with actual data. The symbols are plotted every 20 minutes for legibility. As we can see both methods are accurate for initial time but as time goes on linear first order model without decomposition deviates from original in a faster rate. From Figure.4.3 it is evident that, the proposed EMD based chaotic prediction method is capable providing better prediction even upto 10 hours, compared to prediction without decomposition. A statistical analysis of prediction errors calculated at 30-day intervals for 3 years is also done over five locations. We have chosen measure of normalised root mean squared error (NRMSE), because of the change in location may affect the range of measured wind speed. For a k step ahead prediction, NRMSE is calculated as

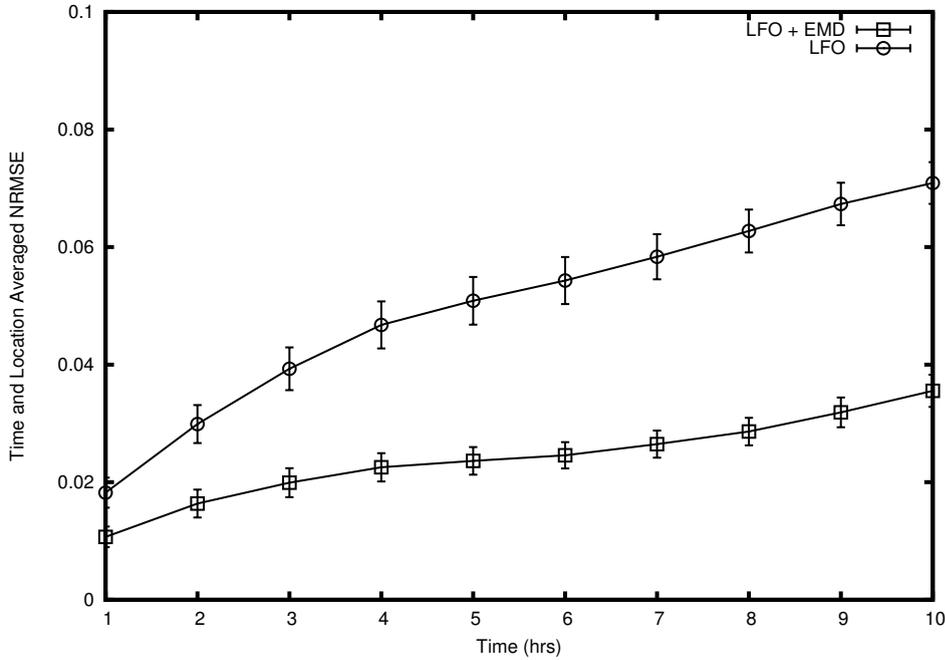


FIGURE 4.3: Time and location averaged NRMSE for predictions up to 10 hours ahead for Lfo with and without EMD.

$$NRMSE = \frac{\sqrt{\frac{\sum_{n+1}^{n+k} (x_i - x_i^p)^2}{k}}}{(x_{max} - x_{min})} \quad (4.3.1)$$

where x_j^p are the predicted values.

Spatial averages of prediction errors over the various locations, as well as time averages over various periods of time at each location, are important measures for getting the clear idea of prediction accuracy since it explain how the model perform along various location as well as at different times. For getting an overall performance of the two models we have calculated both time and location averaged NRMSE upto 10 hours on an hourly basis. From the Figure.4.3 it is clear that EMD based LFO method improves the prediction accuracy significantly and can predict upto 10 hours with a relative error less than 4% of the range of oscillations. For 10 hour prediction the gain of accuracy EMD decomposition is 49.85 % over bare linear first order method.

4.4 Conclusion

Short to medium term prediction of surface wind is critical to the energy management and production strategies of any wind power industry. The apparent unpredictable nature of wind speed oscillations has always been a problem faced by the industry. This has generated lot of enthusiasm research area to develop reliable models to predict wind speed and many techniques have been reported in the literature. Motivated by the capability of deterministic models in providing reliable predications, we investigate strategies to enhance the prediction accuracy. In this chapter we present an empirical mode decomposition (EMD) based chaotic model for short-term prediction of wind speed variations. The comparison of forecast of models with and without EMD decomposition shows that prediction accuracy can be remarkably improved by combining EMD and modeling the wind data based chaos theory.

5

Week-ahead predictions of wind speed using simple linear models with wavelet decomposition

Simple linear methods are widely used for time series modelling and prediction and in particular for the forecast of wind speed variations. Linear prediction models are popular for their simplicity and computational efficiency, but their prediction accuracy deteriorates beyond a few time steps. In this study, we demonstrated that the prediction accuracy of simple auto-regressive (AR) models could be significantly improved, by as much as 60.15% for day-ahead predictions and up to 18.25% for a week ahead forecasts, when combined with suitable time series decomposition. The comparison with new reference forecast model (NRFM) also showed similar accuracy gain of a week ahead predictions. The combined model is capable of forecasting wind speed up to 7 days ahead with an average root mean square error less than 3 m/s. We also compare the performance of AR and ARFIMA models in wind speed prediction and observe that the ARFIMA model is no better than the AR model when used in combination with time series decomposition.

5.1 Introduction

In the recent statistical literature, concern has been the study of long-memory models that capable of going beyond the presence of random walks and unit roots in the univariate time series processes. The autoregressive fractionally integrated moving average (ARFIMA) process is a class of long-memory models, and its primary objective is to explicitly account for persistence to incorporate the long-term correlations in the data (Contreras-Reyes et al., 2013). The recent studies at the Department of Futures Studies - University of Kerala, suggested that the underlying dynamics are deterministic, low-dimensional and chaotic. These findings imply that deterministic models may be more suitable to analyse wind speed fluctuations and may be capable of providing more accurate short-term predictions than the existing stochastic models (Sreelekshmi et al., 2012; Asokan et al., 2012). However, the question of long-term prediction of wind speed intensity remains to be an open problem (Brix et al., 2005).

Wind energy source is considered to be one of the fastest growing alternative sources of energy globally since it is clean, abundant, economically viable and safe for the environment. Each Megawatt hour of wind power generated saves at least 500 kilogrammes of greenhouse effect

gases from being churned into the atmosphere when compared to the electricity produced from gas or coal powered generators thus making the wind power one of the most eco-friendly sources of energy (Nicholson et al., 2011; Clancy et al., 2015). Recently, there has been a steady growth of the generation and use of wind power such that by the end of the year 2013, the total installed capacity of wind energy stood at 318 Gigawatts. According to the European Wind Energy Association, they project that if the growth of wind power generation continues at the current rate, it will account for more than 12% of the total energy demands by the year 2020 (Global Wind Energy Council, 2013). A major factor affecting wind power production is the high variability of the wind speed which is influenced by numerous meteorological factors. These variations do occur at all time scales ranging from seconds to months and even years and being able to predict these fluctuations is a fundamental component in the production and management of wind energy. Based on the forecast period, wind speed predictions are commonly classified as short-term (up to 6 hours ahead), medium term (6 hours up to 1 day ahead) and long-term (1 day up to 1 week ahead) (Soman et al., 2010). Improving the accuracy of predictions at all these time-scales is crucial at various stages of wind energy production and management. For example, short-term forecasts ranging from milliseconds to a few minutes are needed for active turbine control and managing wind energy at electricity grids (Hering et al., 2010; Wang et al., 2012), and forecasts in the range of a few hours up to a few days are useful in energy management and trading, especially in liberalized electricity markets where users device best bidding strategy based on expected power production (Gomes et al., 2012). Long-term forecasts of up to several days ahead are useful in managing the maintenance of wind farms and transmission lines (Aggarwal et al., 2013). The development of methods with improved prediction accuracy of wind speed has been the centre of concern amongst researchers in the recent past.

Various models for wind prediction include physical models, which use complex mathematical equations to describe the physical relationship between different atmospheric parameters and local topography, statistical models which use time series of past data or probability distribution of wind speed for future predictions and also hybrid models which combine physical models with statistical tools (Sfetsos, 2000). An in-depth review of the present status of wind power forecast models, especially of the meteorology based approaches, can be found in (Monteiro et al., 2009) and (Giebel et al., 2003). Provided the meteorological conditions do not change dramatically over a short term, time series models, which use historical wind speed data to predict future values are found to perform reasonably well. They include moving average models such as *ARMA*, *ARIMA* and its variants fitted to the time series of wind speed (Kamal et al., 1997; Cadenas et al., 2007; Kavasseri et al., 2009), models based on artificial neural networks (Mohandes et al., 1998; Cadenas et al., 2007; Bilgili et al., 2007; Monfared et al., 2009) and deterministic prediction models suitable for chaotically varying wind speed dynamics (Sreelekshmi et al., 2012; Drisya et al., 2014). Most of these methods are capable of reasonably accurate predictions up to a few hours, but the range of predictability often varies significantly over topography and other local conditions (Soman et al., 2010).

Wind speed data usually exhibit long-range correlations, and ARFIMA models are especially suited for making short to medium term forecasts of such data. Kavasseri and Sreetharaman Kavasseri et al., 2009 have applied ARFIMA models to forecast hourly average wind speeds up to a period of two days ahead, with an improvement of prediction accuracy up to 42% compared to the simple method of persistence. In this study, we demonstrated that decomposition of the wind speed data into selected frequency components before applying the forecasting technique can dramatically improve the accuracy and longevity of prediction. The decomposition of wind

speed data is achieved by the use of wavelet transform method, while the actual forecasts on the component series are made by simple prediction tools such as auto-regressive (AR) model and ARFIMA model.

With the inherent variability of the wind resource, it is often valuable to be able to forecast wind speed for some time ahead. For example, it may be useful, from a controls standpoint, to be able to predict the very short-term turbulent variations from a few seconds to a few minutes. Or, in the case of a wind turbine or wind farm operator, the capacity to effectively integrate the wind energy into a grid may be affected by the predictability of the output from the wind turbine(s). In this case, wind speed, or power production forecasts, might be needed for the next few hours or even for one to two days ahead.

Wind speed forecasting is very fundamental in determining the resultant wind power generated from a wind farm, that is, the power output from a wind turbine is directly proportional to the cube of wind speed. Thus the accurate prediction of wind speed can result in better wind power forecast. Most of the current research studies are focusing on the development of prediction models that can guarantee improved long-term wind speed prediction accuracy.

5.2 Maximum Overlap Discrete Wavelet Transform

Wavelet transform enables us to decompose time series data into different frequency components and then study each corresponding component with a resolution matched to its scale. Whereas traditional Fourier transform methods use superposition of sines and cosines of different amplitudes and frequency to represent functions, the wavelet transform does this using a collection of wavelet functions, all of which can be generated by scaling and translating a single base wavelet called *mother wavelet*. The mother wavelet and all wavelets generated from it are, unlike sines and cosines, localised in space and the given function or data approximation by a series of scaled and translated versions of these localised functions. This transformation allows processing of data at different scales or resolutions, with lower levels giving finer details of the high-frequency components and higher levels yielding grosser features of the low-frequency components of the data. Measured data such as wind speed data are inherently multi-scale due to contributions from events occurring with different localisations in time and frequency, and wavelets are more suited for analysis of this kind of data.

Mathematically, a mother wavelet is a square integrable function $\psi(t)$, which satisfies the *admissibility condition* (Daubechies et al., 1992),

$$0 < c_\psi = \int_{-\infty}^{\infty} \frac{|\hat{\psi}(\omega)|}{|\omega|} < \infty \quad (5.2.1)$$

where $\hat{\psi}(\omega)$ is the Fourier transform of $\psi(t)$ and preferably a regularity condition which requires that $\psi(t)$ be fast decaying or be non-zero only on a finite interval (Daubechies et al., 1992). To decompose a given function, wavelet transforms use a family of wavelet functions $\psi_{s,\tau}(t)$ obtained from the mother wavelet $\psi(t)$ by dilations and translations;

$$\psi_{s,\tau}(t) = \frac{1}{\sqrt{s}} \psi\left(\frac{t-\tau}{s}\right), \quad s, \tau \in \mathbb{R}, s \neq 0 \quad (5.2.2)$$

where s is the scale parameter and τ is the location parameter. The continuous wavelet transform of the function $x(t)$ is then defined by,

$$W_x(s, \tau) = \int_{-\infty}^{\infty} x(t) \psi_{s, \tau}^*(t) dt \quad (5.2.3)$$

where $*$ denotes the complex conjugate. The admissibility condition of $\psi(t)$ ensures that $x(t)$ can be recovered from $\psi_{s, \tau}(t)$ by the inverse transform (Daubechies et al., 1992)

$$x(t) = \frac{1}{c_\psi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_x(s, \tau) \psi_{s, \tau}(t) \frac{d\tau ds}{s^2}, \quad (5.2.4)$$

and this is called the wavelet decomposition of $x(t)$.

The variation of the scale and location parameter of the wavelet over a continuum of values in the continuous wavelet decomposition leads to undesirable redundancy in the calculation of wavelet coefficients. In practical applications it is more convenient to sample the parameters s and τ on a discrete set of values in the scale-time plane. This leads to *discrete wavelets* defined for suitably chosen grid points on the $s - \tau$ plane,

$$\psi_{j, k}(t) = s_0^{-j/2} \psi(s_0^{-j} t - k\tau_0), \quad j, k \in \mathbb{Z} \quad (5.2.5)$$

where $s_0 > 1$ and τ_0 are fixed dilation and translation factors (Daubechies et al., 1992). The so called *dyadic sampling* corresponds to the choice $s_0 = 2$ and $\tau_0 = 1$. The discrete wavelet transform(DWT) is then defined by,

$$W_x(i, j) = \int_{-\infty}^{\infty} x(t) \psi_{j, k}^*(t) dt. \quad (5.2.6)$$

If the set of wavelets $\psi_{j, k}(t)$ forms an orthogonal basis, the above transform can be inverted leading to the *discrete wavelet decomposition* of $x(t)$ given by,

$$x(t) = \frac{1}{c_\psi} \sum_{j, k \in \mathbb{Z}} W_x(j, k) \psi_{j, k}(t). \quad (5.2.7)$$

The DWT is especially suited for time series data sampled at equal intervals of time. We use a specific version of DWT, called Maximal Overlap DWT (MODWT) which has some advantages over traditional DWT. First, it is well defined for all sample sizes N , unlike DWT which requires N to be multiple of J for a complete decomposition of J scales. MODWT is highly redundant over DWT and also non-orthogonal, but the redundancy allows better comparison of the series with its decomposition (Percival et al., 2006).

At each scale J , the MODWT transforms an N dimensional vector X , which represents the given data, into $J + 1$ new vectors each of dimension N . These vectors consist of J vectors W_1, W_2, \dots, W_J of MODWT wavelet coefficients corresponding to the scales τ_j , $j = 1, 2, \dots, J$ and a vector V_J containing the so-called MODWT scaling coefficients. We can invert this procedure and recover the original vector X from these wavelet and scaling coefficients. This leads to a decomposition known as multi-resolution analysis(MRA), expressed as (Percival et al., 2006; Percival et al.,

2004),

$$X = \sum_{j=1}^J D_j + S_J \quad (5.2.8)$$

The vector D_j contains the details of the data associated with the average variations on a scale of τ_j and is computed exclusively from the wavelet coefficients in W_j . On the other hand, S_J is calculated from the scaling coefficients in V_J and is associated with the averages at scales $2\tau_J$ and higher, which separates the smoother part of the data (Percival et al., 2006; Percival et al., 2004). Thus the MRA expresses the given data as a sum of a smoother part and a set of component parts giving details of the variations at various scales. In this study, the original wind speed data were decomposed into twelve levels after which the decomposed frequencies were predicted independently using the *ARFIMA* model. From the findings, it was realised that the higher the level, the better the prediction. The independent predictions are then summed up together to form the final prediction (R Core Team, 2014).

The study focused on two sets of data for modelling. In the first dataset is the original 10-minute wind speed data which was modelled and predicted without any transformation. This was achieved through the "*forecast*" and "*stats*" packages found in "*R-Statistical software*" that were used in fitting the *ARFIMA* and *AR* models to the 10-minute wind speed data. The optimal parameters p , d and q for *ARFIMA* and p for *AR* and the respective predictions were accomplished using the inbuilt functions in these packages (R Core Team, 2014; George Athanasopoulos et al., 2014). In the second dataset, the original 10-minute wind speed data were subjected to wavelet transformation. Wavelet transformation was achieved using the *Maximal Overlap Discrete Wavelet Transform (MODWT)* built in the "*waveslim*" package using the *modwt()* function which performs level J decomposition of the input vector using the non-decimated discrete wavelet transform. The *imodwt()* function performs the reconstruction of the time series from its maximal overlap discrete wavelet transform using inverse transformation. Level J specifies the depth of the decomposition and it must be a number less than or equal to $\log_2(\text{length}(x))$ (Whitcher et al., 2007).

5.3 Wind speed forecasting models

The two statistical methods that were considered in the prediction of wind speed, namely; *the Auto-Regressive (AR) model* and *the Auto-Regressive Fractionally Integrated Moving Average (ARFIMA) model*. Conventionally, persistence model has been the sole benchmark method used in the assessment of how good the forecast performance of a model is ranging from short-term to long-term forecast, but, Nielsen et al., (1998) suggested that it is unreasonable to apply persistence model as a benchmark method when the forecast time horizon is beyond short-term. He proposed *a new reference forecast model* that should be considered when evaluating the forecast performance of a model with forecast time exceeding the short-term limit.

5.3.1 New Reference Forecast model

The new reference forecast model was proposed by Nielsen et al., (1998) as a preferable benchmark option to persistence model when gauging the forecast accuracy of the model exceeding the short-term duration. The persistence model is a commonly used reference forecasting method

that assumes that the current wind speed observation will remain unchanged for some time ahead and is defined by,

$$P_{t+k} = P_t + \varepsilon_{t+k} \quad (5.3.1)$$

where P denotes the wind speed, t denotes the time index, k denotes the look-ahead time, and ε denotes the error term. The persistence model forecast is then given by, $\hat{P}_{t+k} = P_t$, which implies that after k time steps ahead, the wind speed will still be equal to the current wind speed (Foley et al., 2012).

The k -step ahead forecast obtained from a new reference forecast model is given by,

$$\hat{P}_{t+k} = a_k P_t + (1 - a_k) \bar{P} \quad (5.3.2)$$

where P_t denotes the current wind speed observation and \bar{P} represents the wind speed average defined by,

$$\bar{P} = \frac{1}{N} \sum_{t=1}^N P_t \quad (5.3.3)$$

which means that with small k , then a_k will be approaching one and hence the reference will almost equivalent to persistence method. But, with a large k and zero correlation, then a_k will be approaching zero with the forecast reversing towards the mean thus making it logical to describe a_k as the correlation coefficient between P_{t+k} and P_t as follows,

$$a_k = \frac{\frac{1}{N} \sum_{t=1}^{N-k} (P_t - \bar{P})(P_{t+k} - \bar{P})}{\frac{1}{N} \sum_{t=1}^{N-k} (P_t - \bar{P})^2} \quad (5.3.4)$$

which implies that a_k will lower the mean squared error for the new reference model (Nielsen et al., 1998).

5.3.2 Auto-Regressive model

An Auto-Regressive model is just like a multiple regression model whereby the dependent variable depends on a set of predictor variables. In an Auto-Regressive model, the predictor variables comprises of lagged time series. It is a popular model for time series prediction which uses a linear combination of p past observations and a random error. For instance, an AR(p) model is defined by,

$$X_t = c + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \varepsilon_t \quad (5.3.5)$$

which in generalised form will be given by,

$$X_t = c + \sum_{i=1}^p \phi_i X_{t-i} + \varepsilon_t \quad (5.3.6)$$

where c is a constant, ϕ_i^s are the suitably estimated model coefficients, $t \in \{p, \dots, n\}$, and $\{\varepsilon_t\}$ is independent and identically distributed random variable at time t (Box et al., 2008).

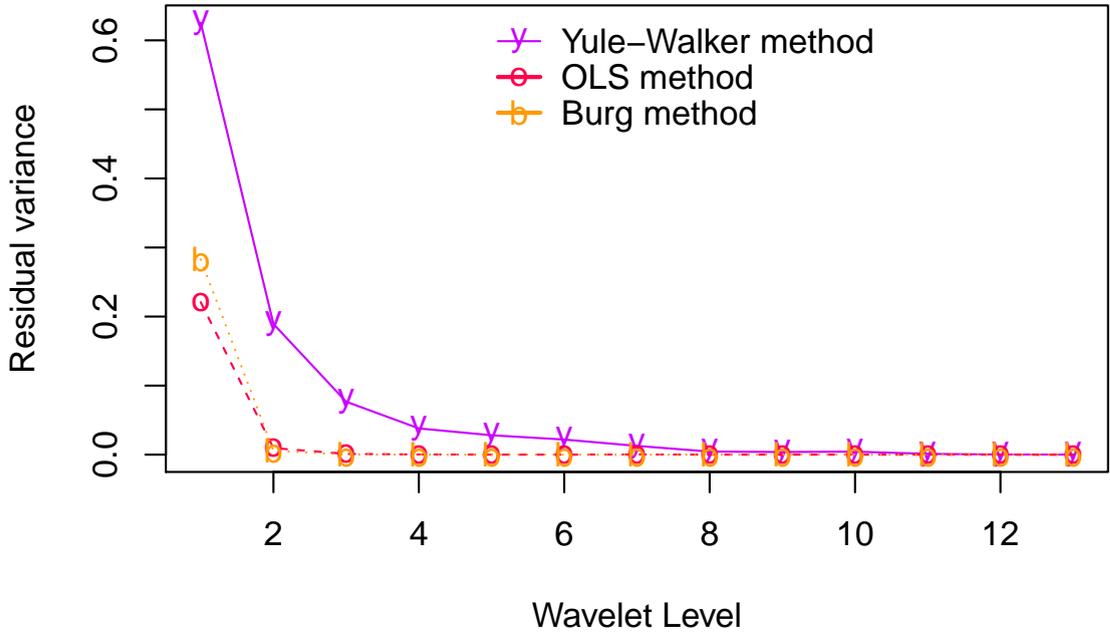


FIGURE 5.1: Estimates of the prediction variance of the time series that is not explained by the autoregressive model for the three methods.

Levinson-Durbin algorithm in AR parameter estimation

Several methods have been put forward for estimating parameters of an Auto-Regressive model, namely; *Yule-Walker method*, *Ordinary least squares method*, *Maximum likelihood estimation method* and the *Burg's method*. On comparing the location and time averaged error variance for four methods listed above, the results showed a decreasing residual variance with decreasing frequency such that *Yule-Walker method* exhibited a higher error variance while the *Maximum likelihood estimation method* failed to converge for some of the wavelet decomposed frequency components. On the other hand, the *ordinary least squares method* and the *Burg's method* showed reasonable level of robustness in estimating the parameters with *Burg's method* performing better than the *ordinary least squares method* as shown in Fig. 5.1.

Due to the robustness exhibited by the Burg's method, it was adopted for the estimation of the parameters as implemented in the *stats* package of *R* (De Hoan et al., 1996; R Core Team, 2014). Burg method fits an Auto-Regressive model to the input data through minimization of the forward and the backward prediction errors while constraining the AR parameters to satisfy the Levinson-Durbin recursion thus providing more stable and robust model parameter estimates (Tamazin et al., 2013). The ultimate goal in forecasting is to forecast the value of X_{t+1} given the past observations $X_t, X_{t-1}, X_{t-2}, \dots, X_2, X_1$, and the best linear predictor is defined by,

$$X_{t+1|t} = P_{X_1, \dots, X_t}(X_{t+1}) = X_{t+1|t, \dots, 1} = \sum_{i=1}^t \phi_{t,i} X_{t+1-i} \quad (5.3.7)$$

where $\phi_{t,i}^s$ are chosen such the the mean squared error is minimized and are solutions to the equation,

$$\begin{pmatrix} \phi_{t,i} \\ \vdots \\ \phi_{t,t} \end{pmatrix} = \mathbf{E} [X_i X_j]^{-1} \mathbf{E} [X_{t-j} X_{t+1}] \quad (5.3.8)$$

Using the standard methods, like *Gauss-Jordan elimination*, to solve this system of equations requires $O(t^3)$ operations. However, since X_t is a stationary time series, thus $\mathbf{E}[X_i X_j]$ is a Toeplitz matrix, by using this information in the 1940's Norman Levinson proposed an algorithm which reduced the number of operations to $O(t^2)$. In 1960's, Jim Durbin adapted the algorithm to time series and improved it.

To outline the algorithm, the best linear predictor of X_{t+1} given the past observations $X_t, X_{t-1}, X_{t-2}, \dots, X_2, X_1$ as defined in equation 5.3.7 is given by,

$$X_{t+1|t} = \sum_{i=1}^t \phi_{t,i} X_{t+1-i} \quad (5.3.9)$$

The mean squared error is $r(t+1) = \mathbf{E}[X_{t+1} - X_{t+1|t}]^2$ and given that the second order stationary covariance structure, the idea of the *Levinson-Durbin algorithm* is to recursively estimate $\phi_{t,i}; i = 1, \dots, t$ given that $\phi_{t-1,i}; i = 1, \dots, t-1$ which are the coefficients of the best linear predictor of X_t given $X_{t-1}, X_{t-2}, \dots, X_2, X_1$. Suppose that the auto-covariance function $c(k) = \text{cov}[X_0, X_k]$ is known, then the *Levinson-Durbin algorithm* is calculated using the following recursion,

- *Step 1:*

$$\phi_{1,1} = \frac{c(1)}{c(0)}$$

and,

$$\begin{aligned} r(2) &= \mathbf{E}[X_2 - X_{2|1}]^2 \\ &= \mathbf{E}[X_2 - \phi_{1,1} X_1]^2 \\ &= 2c(0) - 2\phi_{1,1}c(1) \end{aligned} \quad (5.3.10)$$

- *Step 2:* For $i = t$,

$$\begin{aligned} \phi_{t,t} &= \frac{c(t) - \sum_{i=1}^{t-1} \phi_{t-1,i} c(t-i)}{r(t)} \\ \phi_{t,i} &= \phi_{t-1,i} - \phi_{t,t} \phi_{t-1,t-i} \text{ for } i \in \{1, 2, \dots, t-2, t-1\}, \text{ and} \\ r(t+1) &= 2c(0) - 2\phi_{1,1}c(1) \end{aligned} \quad (5.3.11)$$

The recursion in equation 5.3.11 above can be verified using a proof based on projections as described below; Suppose that $\{X_t\}$ is a stationary time series with mean zero and auto-covariance, $c(k) = \mathbf{E}[X_k X_0]$. Let $P_{X_t, X_{t-1}, \dots, X_3, X_2}(X_1)$ denote the best optimum linear predictor of X_1 given $X_t, X_{t-1}, \dots, X_3, X_2$ and $P_{X_t, X_{t-1}, \dots, X_3, X_2}(X_{t+1})$ denote the best optimum linear predictor of X_{t+1}

given $X_t, X_{t-1}, \dots, X_3, X_2$. According to stationarity, the following predictors have the same coefficients,

$$\begin{aligned} X_{t|t-1} &= \sum_{i=1}^{t-1} \phi_{t-1,i} X_{t-i} \\ P_{X_t, X_{t-1}, \dots, X_3, X_2}(X_{t+1}) &= \sum_{i=1}^{t-1} \phi_{t-1,i} X_{t+1-i} \\ P_{X_t, X_{t-1}, \dots, X_3, X_2}(X_1) &= \sum_{i=1}^{t-1} \phi_{t-1,i} X_{i+1} \end{aligned} \quad (5.3.12)$$

The three relations defined in equation 5.3.12 above are important components of the proof. By recalling the objective to derive the coefficients of the best linear predictor of

$$P_{X_t, X_{t-1}, \dots, X_2, X_1}(X_{t+1})$$

based on the coefficients of the best linear predictor $P_{X_{t-1}, X_{t-2}, \dots, X_2, X_1}(X_t)$. To implement this, $s\bar{p}(X_t, X_{t-1}, \dots, X_2, X_1)$ is partitioned into two orthogonal spaces as follows,

$$s\bar{p}(X_t, X_{t-1}, \dots, X_2, X_1) = s\bar{p}(X_t, X_{t-1}, \dots, X_2, X_1) \oplus s\bar{p}(X_1 - P_{X_t, X_{t-1}, \dots, X_3, X_2}(X_1)) \quad (5.3.13)$$

Therefore due to no correlation, we get the partition,

$$\begin{aligned} X_{t+1|t} &= P_{X_t, X_{t-2}, \dots, X_3, X_2}(X_{t+1}) + P_{X_1 - P_{X_t, X_{t-1}, \dots, X_3, X_2}(X_1)}(X_{t+1}) \\ &= \sum_{i=1}^{t-1} \phi_{t-1,i} X_{t+1-i} + \phi_{t,t}(X_1 - P_{X_t, X_{t-1}, \dots, X_3, X_2}(X_1)) \\ &= \sum_{i=1}^{t-1} \phi_{t-1,i} X_{t+1-i} + \phi_{t,t}(X_1 - \sum_{i=1}^{t-1} \phi_{t-1,i} X_{i+1}) \end{aligned} \quad (5.3.14)$$

By evaluation the expression $\phi_{t,t}$, we get,

$$\begin{aligned} \phi_{t,t} &= \frac{\mathbf{E}[X_{t+1}(X_1 - P_{X_t, X_{t-1}, \dots, X_3, X_2}(X_1))]}{\mathbf{E}[X_1 - P_{X_t, X_{t-1}, \dots, X_3, X_2}(X_1)]^2} \\ &= \frac{\mathbf{E}[X_{t+1} - P_{X_t, X_{t-1}, \dots, X_3, X_2}(X_{t+1})(X_1 - P_{X_t, X_{t-1}, \dots, X_3, X_2}(X_1))]}{\mathbf{E}[X_1 - P_{X_t, X_{t-1}, \dots, X_3, X_2}(X_1)]^2} \end{aligned} \quad (5.3.15)$$

The denominator is the mean squared error $r(t)$, that is,

$$r(t) = \mathbf{E}[X_1 - P_{X_t, X_{t-1}, \dots, X_3, X_2}(X_1)]^2 = \mathbf{E}[X_t - P_{X_{t-1}, X_{t-2}, \dots, X_2, X_1}(X_t)]^2 \quad (5.3.16)$$

hence by simplification we get,

$$\begin{aligned}
 \phi_{t,t} &= \frac{\mathbf{E} [X_{t+1}(X_1 - P_{X_t, X_{t-1}, \dots, X_3, X_2}(X_1))]}{r(t)} \\
 &= \frac{c(0) - \mathbf{E} [X_{t+1}P_{X_t, X_{t-1}, \dots, X_3, X_2}(X_1)]}{r(t)} \\
 &= \frac{c(0) - \sum_{i=1}^{t-1} \phi_{t-1,i}c(t-i)}{r(t)}
 \end{aligned} \tag{5.3.17}$$

which is the same as the first expression in equation 5.3.11 of the *Levinson-Durbin algorithm*.

To obtain the recursion for $\phi_{t,i}$, equation 5.3.14 is used for simplification,

$$\begin{aligned}
 X_{t+1|t} &= \sum_{i=1}^t \phi_{t,i}X_{t+1-i} \\
 &= \sum_{i=1}^{t-1} \phi_{t-1,i}X_{t+1-i} + \phi_{t,t}(X_1 - \sum_{i=1}^{t-1} \phi_{t-1,i}X_{i+1})
 \end{aligned}$$

and by comparing the coefficients we get,

$$\phi_{t,i} = \phi_{t-1,i} - \phi_{t,t}\phi_{t-1,t-i} \text{ for } i \in \{1, 2, \dots, t-2, t-1\}$$

which is the same as the second expression in equation 5.3.11. To get the recursion for the mean squared prediction error, we note that by orthogonality of $\{X_t, X_{t-1}, \dots, X_3, X_2$ and $X_1 - P_{X_t, X_{t-1}, \dots, X_3, X_2}(X_1)$, equation 5.3.14 is used for simplification to get,

$$\begin{aligned}
 r(t+1) &= \mathbf{E} [X_{t+1} - X_{t+1|t}]^2 \\
 &= \mathbf{E} [X_{t+1} - P_{X_t, X_{t-1}, \dots, X_3, X_2}(X_{t+1}) - \phi_{t,t}X_1 - P_{X_t, X_{t-1}, \dots, X_3, X_2}(X_1)]^2 \\
 &= r(t) + \phi_{t,t}^2 r(t) - 2\phi_{t,t}r(t)\phi_{t,t} \\
 &= r(t) [1 - \phi_{t,t}^2]
 \end{aligned}$$

which gives the final expression of equation 5.3.11 of the *Levinson-Durbin algorithm*.

For time series having predominantly deterministic character, Autoregressive model is expected to perform better in prediction compared to other linear methods. On the other hand, a moving average (MA) model uses a linear combination of past errors. A generalised moving average model $MA(q)$ of order q has the form,

$$x_t = \mu + \sum_{j=1}^q \theta_j \varepsilon_{t-j} + \varepsilon_t \tag{5.3.18}$$

where θ_j 's are suitably chosen model coefficients, μ is the mean of the series and the random errors ε_{t-j} are assumed to be independent and identically distributed with zero mean and constant variance (Box et al., 2008).

An autocorrelation function of the residuals is plotted to check how well the model fits the data, and the results showed a Gaussian white noise in all the selected locations. The autocorrelation functions of the residuals for all the sites cut off at lag 1, which implies that they are independent and identically distributed (Granger et al., 2014). A proper model fitting and selection determines how robust the model would be in producing reliable forecasts and several methods have been put forward to determine the model's accuracy and in this case, the root mean squared error was used. Suppose that X_t denote the observed wind speed and \hat{X}_t the forecast wind speed, then the root mean squared error is given by,

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (\hat{X}_t - X_t)^2}{n}} \quad (5.3.19)$$

where $t = 1, 2, 3, \dots, n$.

5.3.3 Auto-Regressive Fractionally Integrated Moving Average model

In time series analysis, historical data are very crucial in the development of forecasting models used in the prediction of future values. An *Auto-Regressive Fractionally Integrated Moving Average* model is a special case of an *ARIMA* process. *ARIMA* model is a crucial process used in time series analysis, and it incorporates three terms, namely;

1. *Auto-Regressive (AR) term* : – the *Auto Regressive* term is a linear regression of the current observed time series value with lagged values of the time series. It captures the dependency of present value and its nearest prior values.
2. *Integrated (I) term* : – refers to the reverse process of differencing to produce the forecast and,
3. *Moving Average (MA) term* : – the *Moving Average* term captures the influence of random shocks. In the real world, any set of time series data is affected by several random factors resulting in random shocks and may memorise the previously received random shocks for a while.

An auto-regressive model of order p and a moving average model of order q can be effectively combined to form the more useful $ARMA(p, q)$ model, which has the general form,

$$x_t = c + \varepsilon_t + \sum_{i=1}^p \phi_i x_{t-i} + \sum_{j=1}^q \theta_j \varepsilon_{t-j} \quad (5.3.20)$$

Using the operators,

$$\begin{aligned} \Phi(B) &= 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p \\ \Theta(B) &= 1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q \end{aligned} \quad (5.3.21)$$

the definition of $ARMA(p, q)$ model can be written as,

$$\Phi(B)x_t = c + \Theta(B)\varepsilon_t \quad (5.3.22)$$

where B is the backward-shift operator so that $Bx_t = x_{t-1}$.

ARMA models are used in forecasting time series of stationary processes. Time series of non-stationary processes are best modelled using integrated ARMA models or ARIMA models, which additionally uses differencing operation to remove stationarity. The $ARIMA(p, d, q)$ model has the general form,

$$\Phi(B)\Delta^d x_t = c + \Theta(B)\varepsilon_t \quad (5.3.23)$$

where $\Delta = 1 - B$ is the differencing operator and d is an integer (Box et al., 2008).

The combination of the *AR* and the *MA* terms results in an *ARMA* model, this model assumes that the data are stationary, *i.e.* the statistical properties of data do not change over time. Although in real life, this assumption does not hold for most of the real time series data. Thus, by introducing the Integrated term, as in *ARIMA* removes the impact of non-stationary data by differencing. Usually, the first-order differencing is usually sufficient, although higher-order differencing can still be implemented if need be.

Consider an $ARIMA(p, d, q)$ model, *i.e.* p – is the order of the autoregressive components, d – is the number of differencing operators, and q – the highest order of the moving average term. The only difference between the *ARFIMA* model and the *ARIMA* model is that d is allowed to take fractionally continuous value instead of an integer value within the range of $(-0.5, 0.5)$. *ARFIMA* is a generalization of *ARIMA* where the parameter d is allowed to have a fractional value with the operator $(1 - B)^d$ interpreted to have the binomial expansion (Brockwell et al., 2009),

$$(1 - B)^d = 1 - dB + \frac{d(d-1)}{2!}B^2 + \dots \quad (5.3.24)$$

The possibility of a wide range of choices for the parameters p , d , q and the constants ϕ_i and θ_j give the model great flexibility and wider applicability.

One of the features that distinguish a *ARFIMA* process from an *ARIMA* process is that the former is characterised by a slow decay in its auto-correlation function compared to the latter. This feature makes *ARFIMA* model an attractive choice for data sets that exhibit long-range correlations such as the wind speed data (Kavasseri et al., 2009).

Suppose that $\{x_t\}$ and $\{x_{t-i}\}$ denote the current and the previous 10-minute wind speeds (m/s) respectively whereas $\{\varepsilon_t\}$ and $\{\varepsilon_{t-j}\}$ denote the current and the previous error terms respectively at time t with $i, j \in [1, 2, 3, \dots]$ being their respective lags. Let B denote a *backward shift operator* or *time lag operator* such that $B^i x_t = x_{t-i}$, then, *fractional-ARIMA* model can be expressed as follows;

$$\left(1 - \sum_{i=1}^p \phi_i B^i\right)(1 - B)^d x_t = \left(1 + \sum_{j=1}^q \theta_j B^j\right)\varepsilon_t \quad (5.3.25)$$

where ϕ_i and θ_j are constants, $\varepsilon_t \sim iid(0, \sigma^2)$ and $d \in (-0.5, 0.5)$.

5.4 Results and discussion

In earlier works, we have demonstrated that random like fluctuations found in time series of wind speed could arise from an underlying chaotic dynamics (Sreelekshmi et al., 2012), and

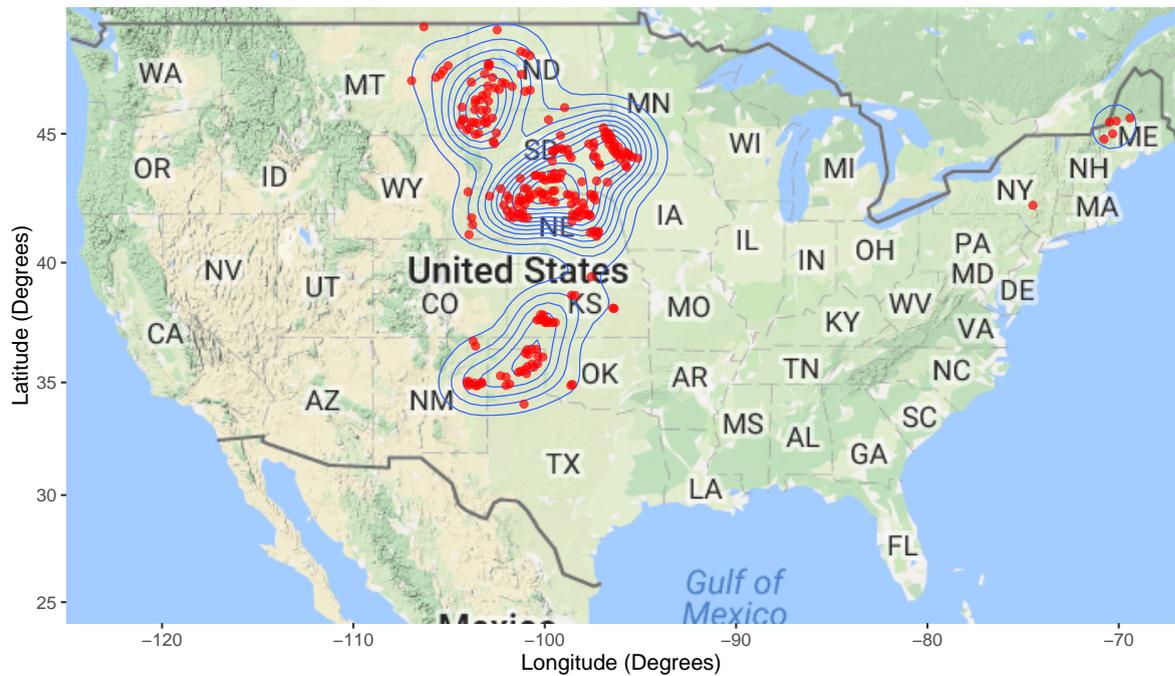


FIGURE 5.2: The geographical locations, denoted by filled circles, for the 234 sites selected for wind speed data analysis and prediction.

that in such situations, deterministic forecasting methods can make significantly accurate short-term predictions of wind speed (Drisya et al., 2014). However, the chaotic behaviour inherently limits the possibility of accurate long-term predictions using chaotic prediction methods due to the exponential divergence of nearby trajectories. Wind speed time series data is the result of an interplay between numerous dynamical factors of various scales and frequencies and with a systematic procedure to keep track of the various frequencies it could be possible to bypass these limitations to some extent and make fairly accurate long-term predictions taking advantage of the underlying determinism. We demonstrated that wavelet decomposition of wind speed data combined with simple auto-regressive prediction models could make long-term predictions as far as a week ahead possible with root mean square error below 3 m/s.

For developing wind speed prediction models, higher resolution wind speed data are desirable. In this analysis, we used wind speed data with 10-minute resolution for the period from January 2004 to January 2007 recorded at 234 different locations in the USA ranging from latitude 34.05911°N , longitude $106.95718^{\circ}\text{W}$ to latitude 48.84354°N , longitude 69.40916°W as available from National Renewable Energy Laboratory (<http://www.nrel.gov>), USA as shown in Fig. 5.2.

Our prediction method starts with an MRA of the given time series, thus decomposing it into time series of various scales by applying MODWT with Daubechies wavelet of order 8. We have set $J = 12$ giving rise to 12 detail series D_j , $j = 1, 2, \dots, 12$ (at scales 2^j , $j = 1, 2, \dots, 12$) and the smooth series S_{12} of variations at scales greater than 2^{12} . A part of this component series is then selected as the model data for prediction which is then used to forecast several time steps into the future using AR or ARFIMA method. The resulting series are then combined, again using MRA, to reconstruct the original series along with the predicted values as shown in figures 5.6 - 5.3.

From figures 5.3 - 5.6, it can be noted that as the level increases, the determinism in the data also increases. This relationship explains the reason behind improved accuracy on long-term

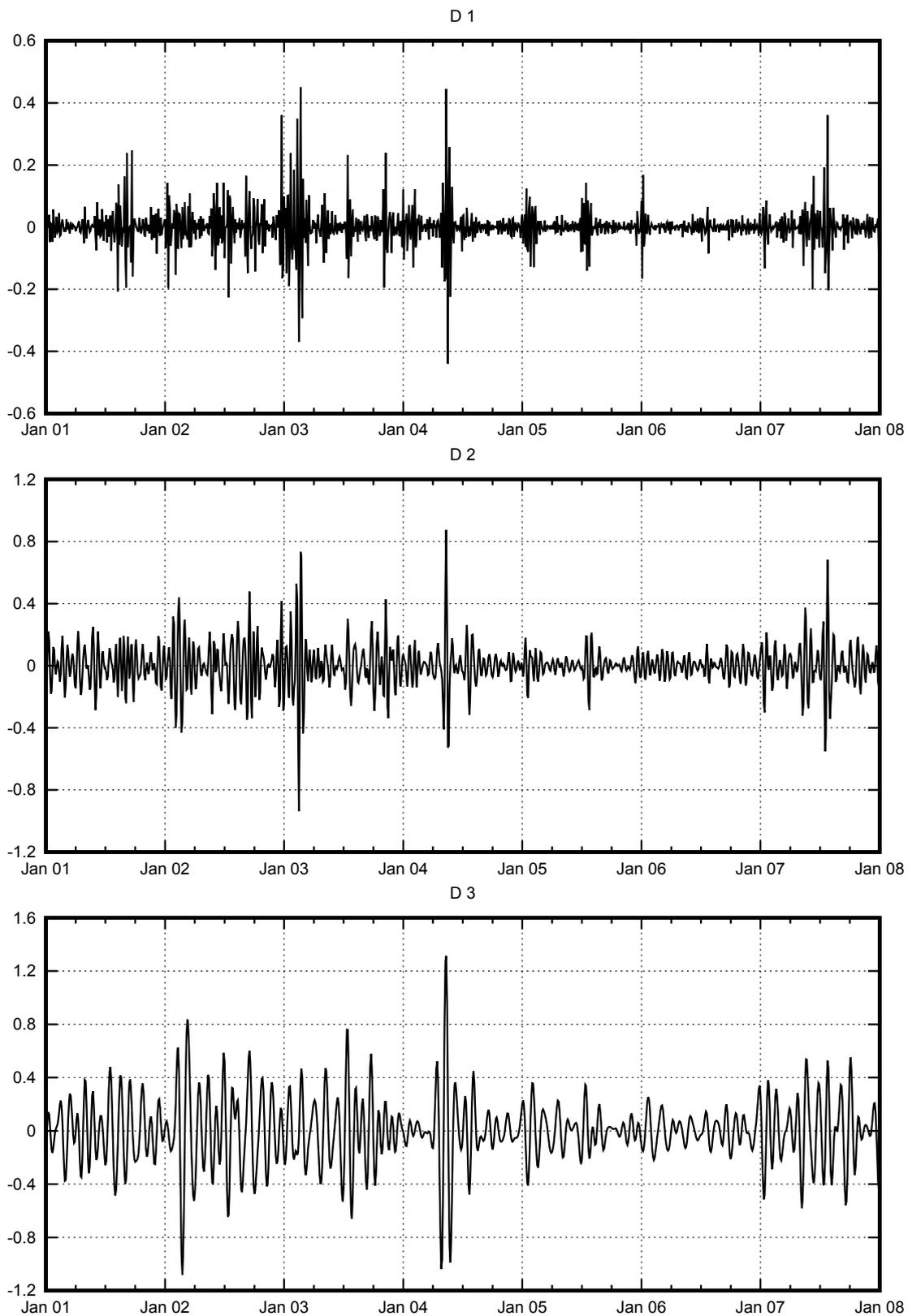


FIGURE 5.3: The D1–D3 frequency components from the decomposed 10-minute wind speed data for a time period of one week.

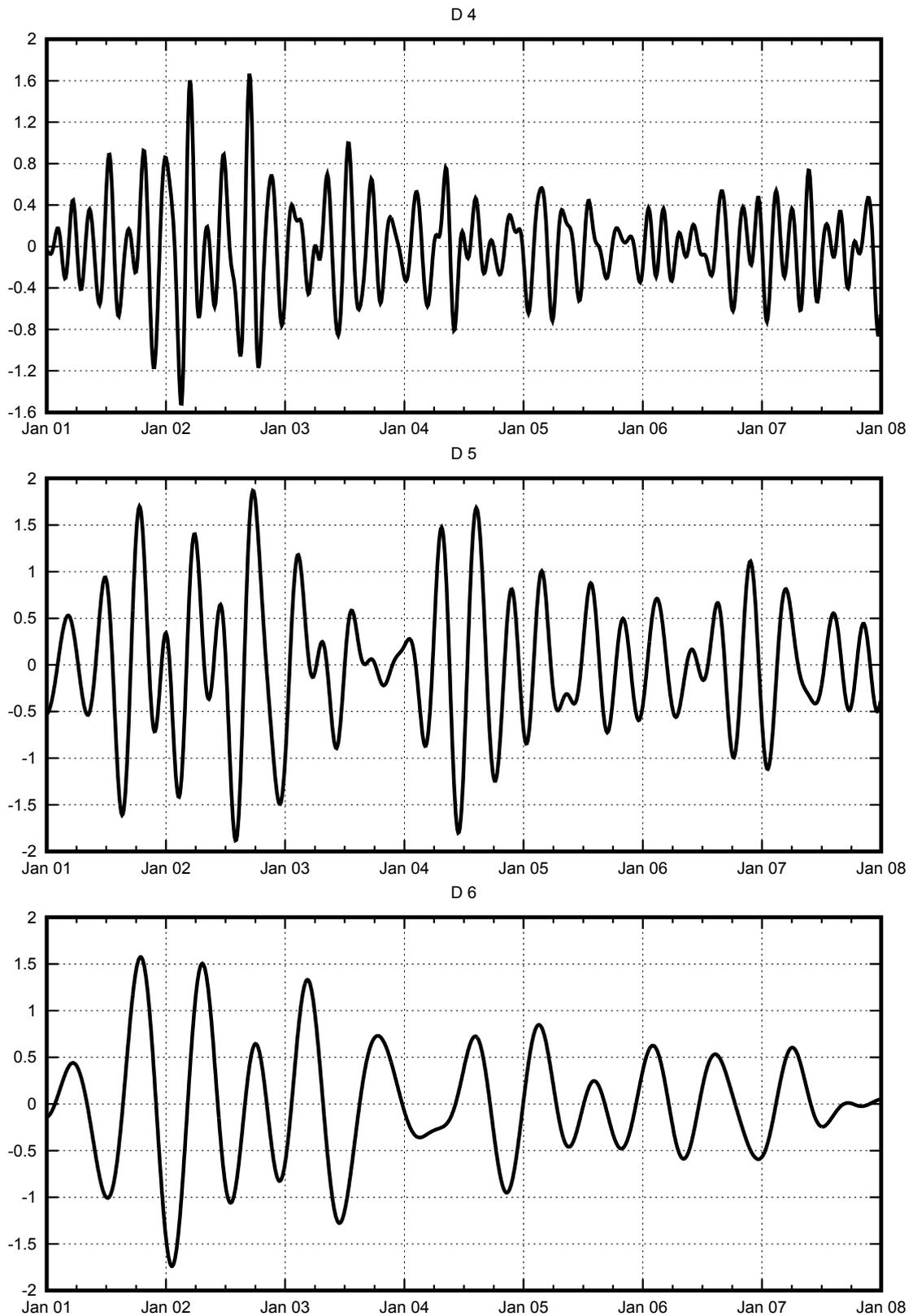


FIGURE 5.4: The D4–D6 frequency components from the decomposed 10-minute wind speed data for a time period of one week.

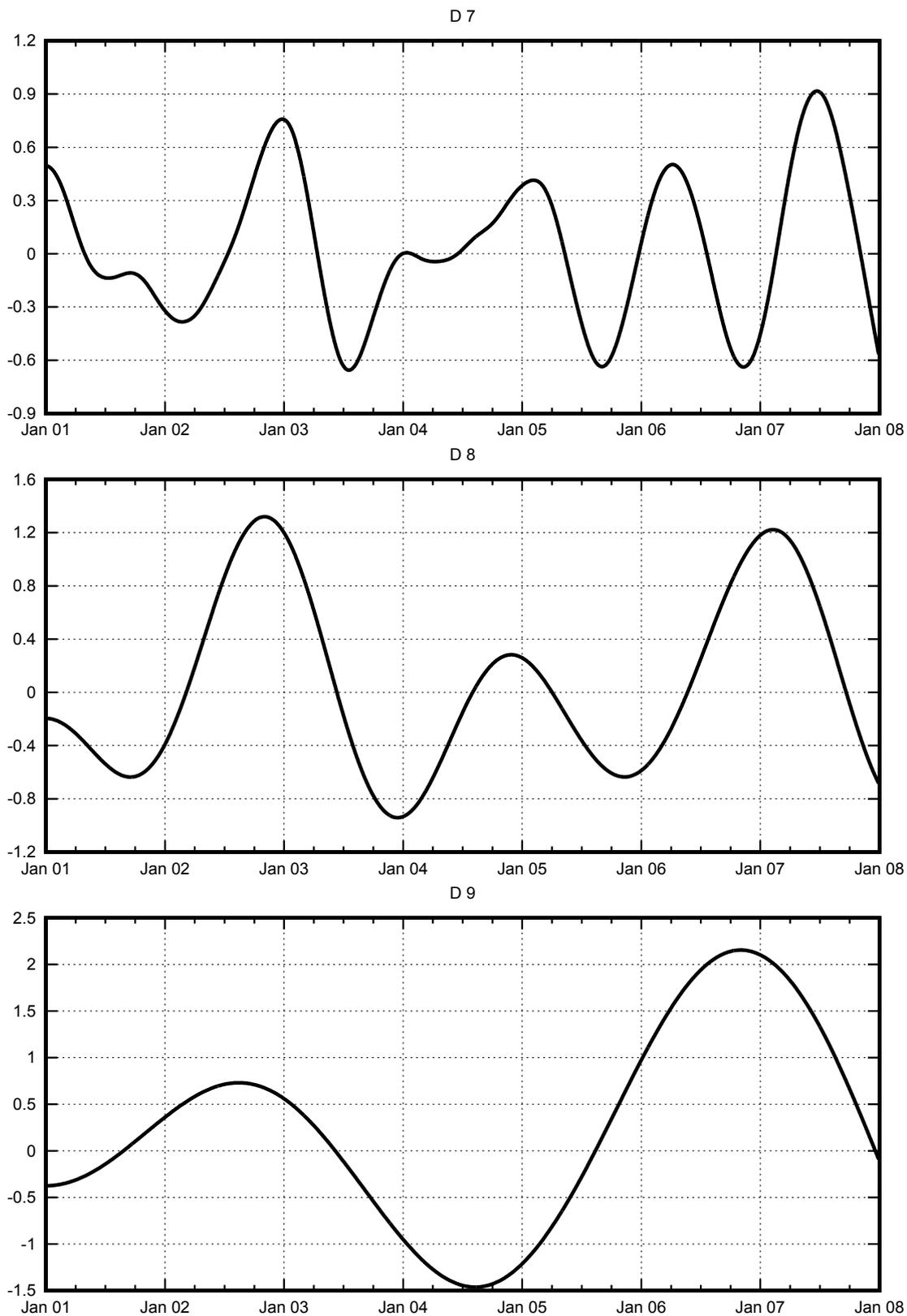


FIGURE 5.5: The D7–D9 frequency components from the decomposed 10-minute wind speed data for a time period of one week.

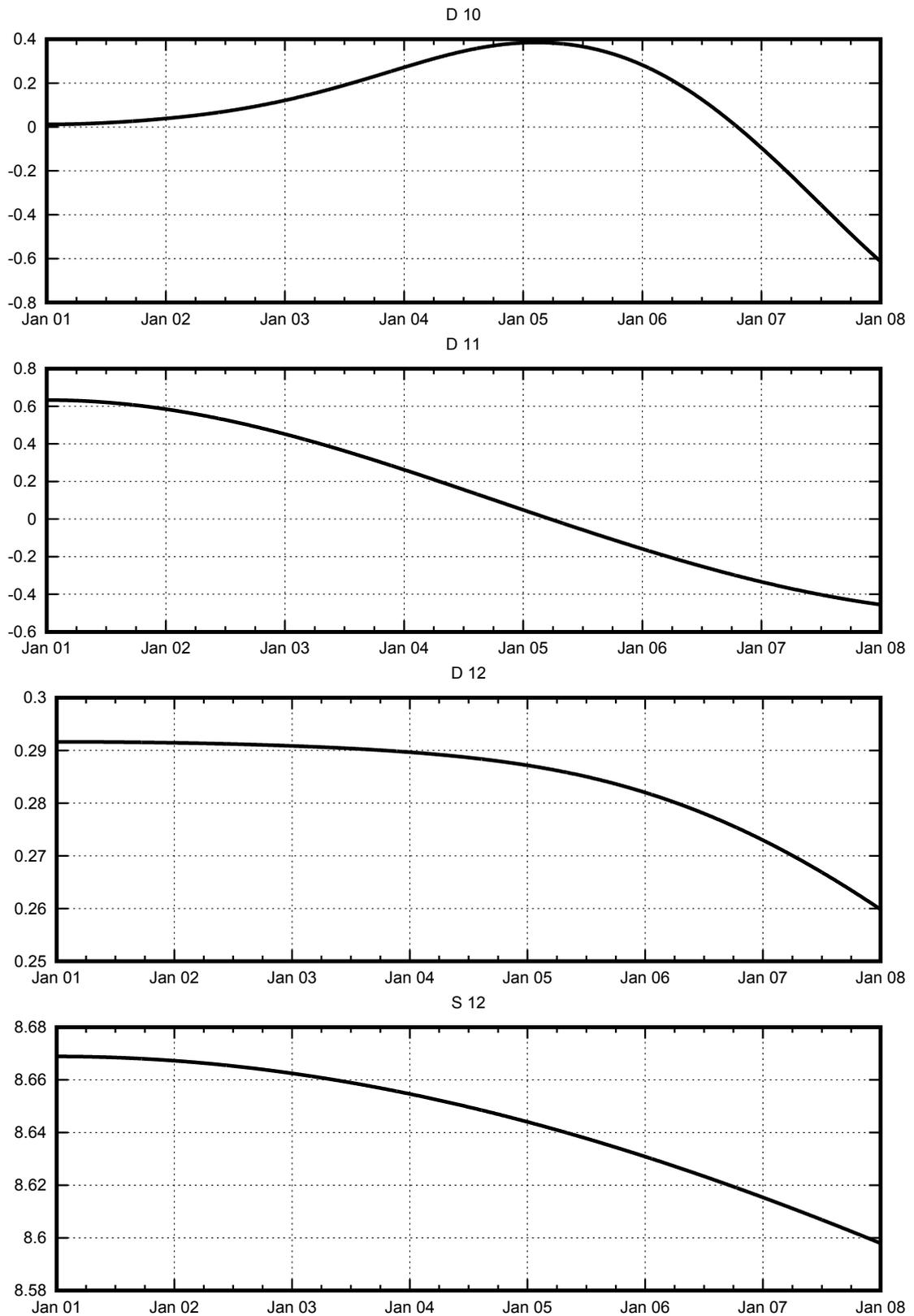


FIGURE 5.6: The D10–D12 frequency components and the S12 scaling component from the decomposed 10-minute wind speed data for a time period of one week.

prediction achieved at lower frequencies as compared to the higher frequencies having a faster rate of mean reversion as described in the following analogy.

Figures 5.7 and 5.8 show the results of predictions made by AR and ARFIMA models, for some hundred hours into the future, on the wind speed data from four different locations, plotted along with the actual data. Also plotted in the figures are the predictions by each of these models in combination with wavelet decomposition technique as described earlier. These plots clearly show that both the AR and ARFIMA predictions are remarkably improved when combined with wavelet decomposition. Whereas the forecasts by the plain models diverge from the actual data after a few time steps, when combined with wavelet decomposition, they yield reasonably accurate predictions for a longer period and pick up the dynamics of the original time series more or less reasonably.

To further investigate the reliability of this technique combining prediction models with wavelet decomposition, and also to demonstrate their wider applicability in wind speed prediction, we have carried out a statistical analysis of the wind speed forecasts made at a range of different geographical locations. The analysis consists of computing and comparing wind speed prediction errors across the 234 different locations mentioned above. We use spatial averages of prediction errors over the various sites, as well as time averages over different periods of time at each location, to compare the prediction accuracies of the forecast models when used directly and in combination with wavelet decomposition. The prediction errors are measured using the root mean squared error (RMSE) defined as follows. Suppose that from a given time series of $n+k$ observed values x_1, x_2, \dots, x_{n+k} , we choose the first n values as constituting the model data for prediction and obtain the forecasts $x_{n+1}^p, x_{n+2}^p, \dots, x_{n+k}^p$ for the next k values. The root mean squared error is then given by,

$$RMSE = \sqrt{\frac{\sum_{i=n+1}^{n+k} (x_i - x_i^p)^2}{k}} \quad (5.4.1)$$

For each location we obtained predictions for 1 to 9 days ahead at intervals of 30 days for the 3-year period from 2004 to 2006, using the AR or ARFIMA model alone at first and then in combination with wavelet decomposition. Wind speed data of previous 30 days were used to build suitable models for predictions in all cases. The time-averaged RMSE for each location was computed by averaging over the RMSEs of the predictions at intervals of 30 days at the location. The location averaged RMSE at a particular time was calculated by averaging over the root mean squared errors of predictions at various locations at the time.

To compare the performance of AR and ARFIMA models in the presence of wavelet decomposition, we computed both location and time averaged root mean squared errors for predictions up to 6 hours ahead, which evaluates to 0.01627287 ± 0.00017 for AR model and to 0.05017721 ± 0.00075 for ARFIMA model respectively. We have repeated this comparison for hourly mean wind speed data, and the results are shown in Fig. 5.9. These results indicate that when used in combination with wavelet decomposition, the AR model fares much better, with considerably less computational cost, than the ARFIMA model. The ARFIMA model also shows the tendency to diverge for certain scales of the data. Hence the rest of the analysis was carried out using exclusively the AR model as the base prediction tool in combination with wavelet decomposition.

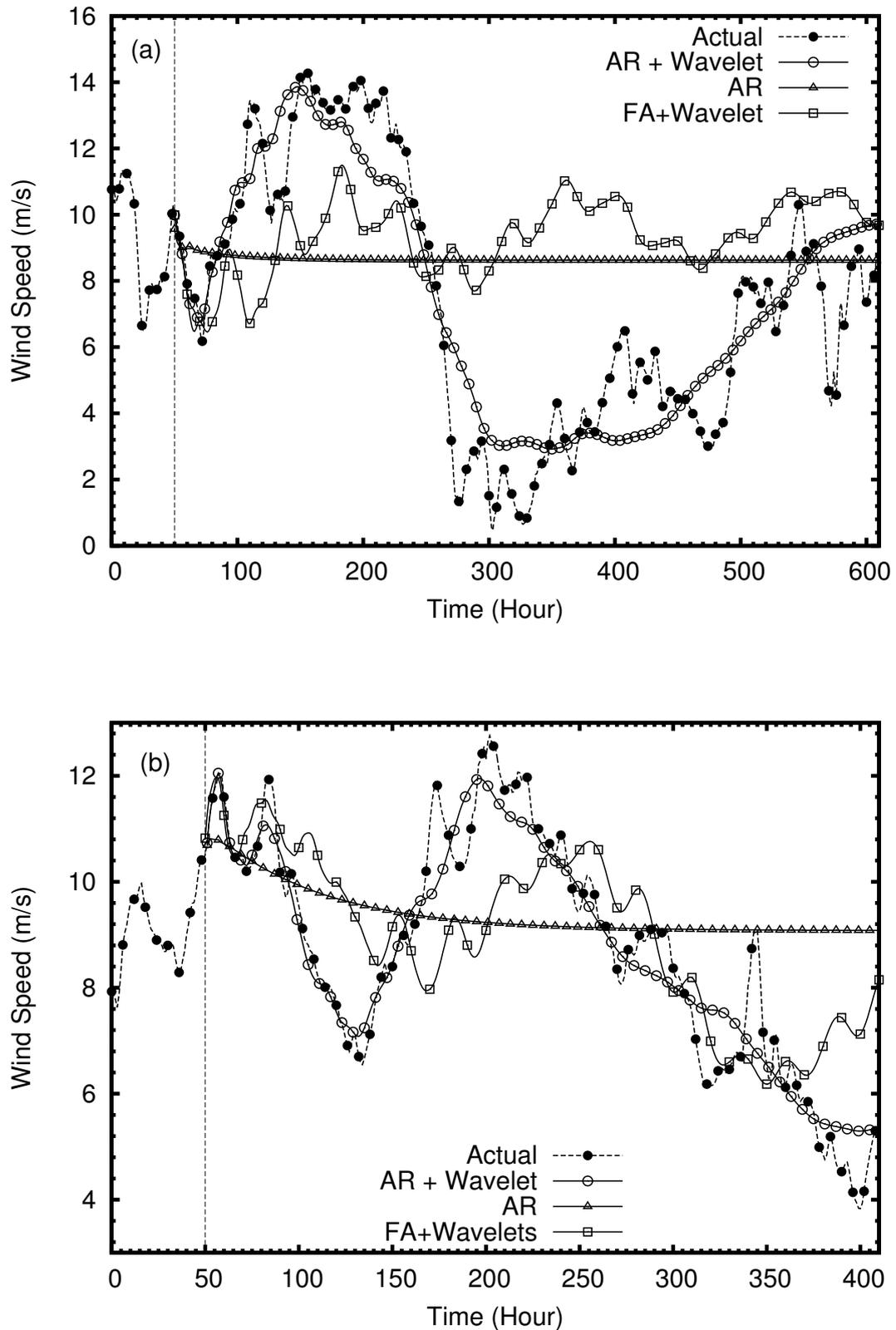


FIGURE 5.7: Comparison of predicted values with actual data for AR and ARFIMA models with and without wavelet decomposition. (a) Latitude: 44.34406°N Longitude: 99.61266°W , prediction start time: 2004-12-26 12:10:00 (b) Latitude: 45.39850°N Longitude: $103.51002^{\circ}\text{W}$, prediction start time: 2006-03-21 12:10:00

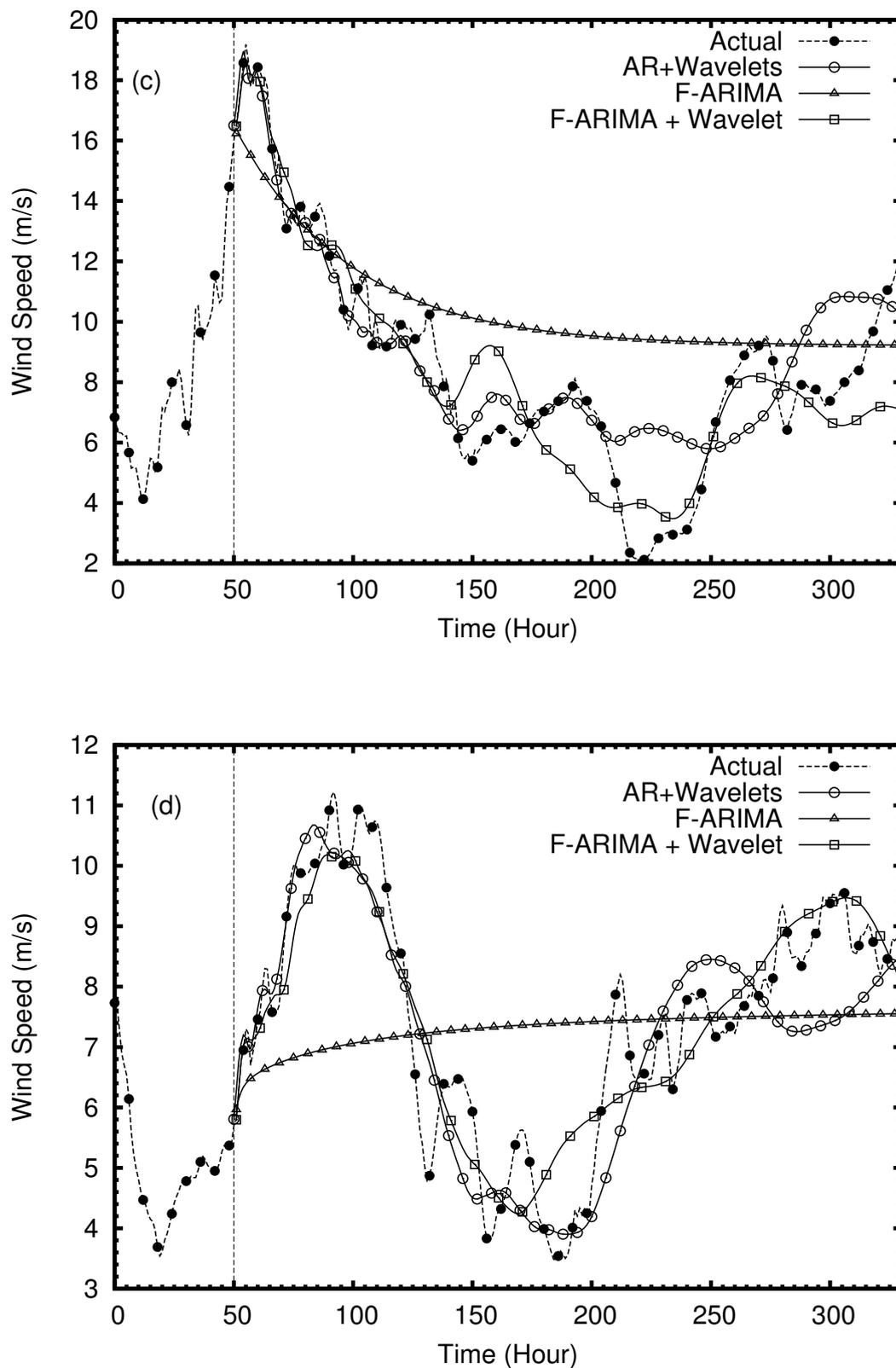


FIGURE 5.8: Comparison of predicted values with actual data for AR and ARFIMA models with and without wavelet decomposition. (c) Latitude: 44.95404°N Longitude: 96.60688°W, prediction start time: 2005-09-22 12:10:00 (d) Latitude: 38.67878°N Longitude: 98.59783°W, prediction start time: 2004-01-31 12:10:00

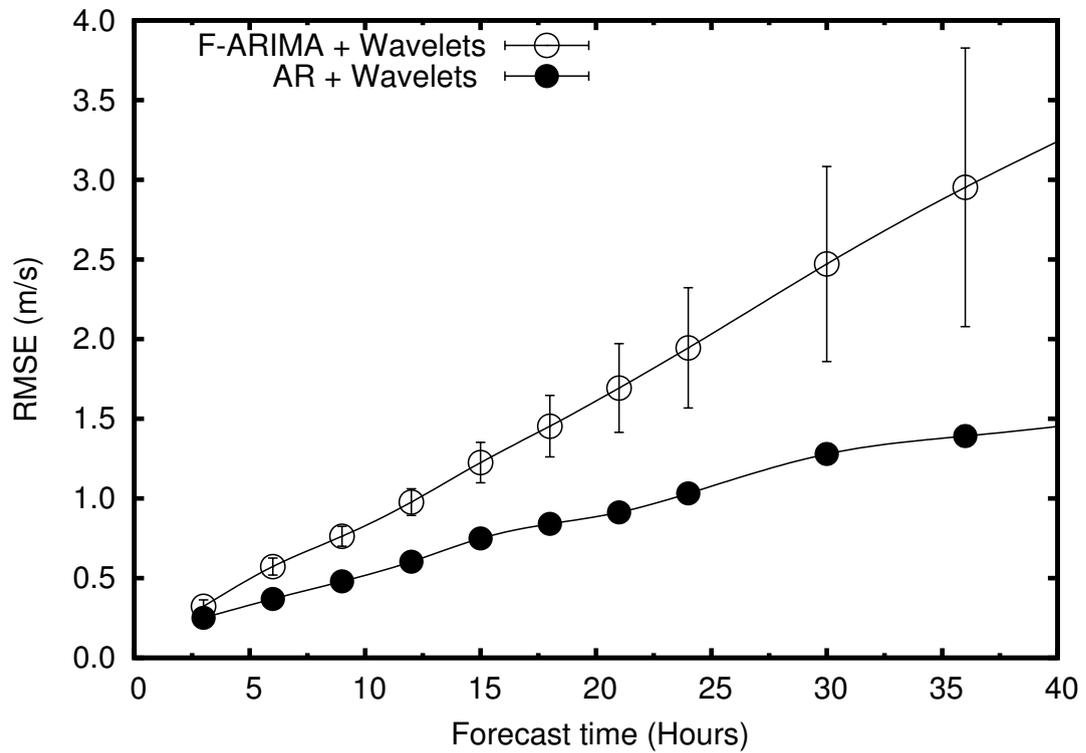


FIGURE 5.9: Comparison of location and time averaged root mean squared errors of hourly mean wind speed data between AR and ARFIMA models both combined with wavelet decomposition.

Figures 5.10–5.16 below show the average root mean squared errors for 1 day to 7 days ahead predictions for both the *location-averaged root mean squared errors* and the *time-averaged root mean squared errors* respectively.

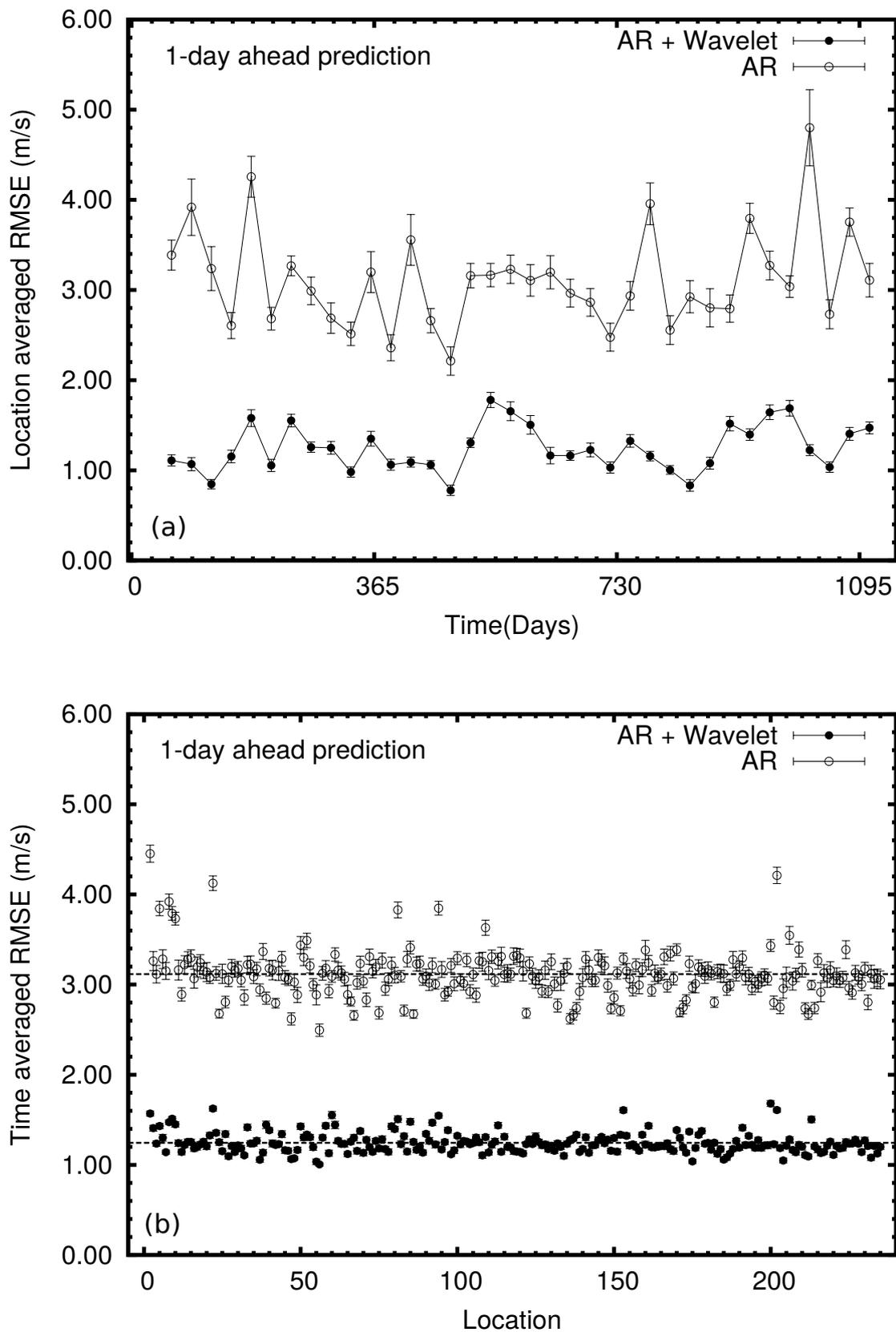


FIGURE 5.10: Plots of averaged root mean squared errors for 1 day ahead predictions at (a) different locations and (b) different time intervals.

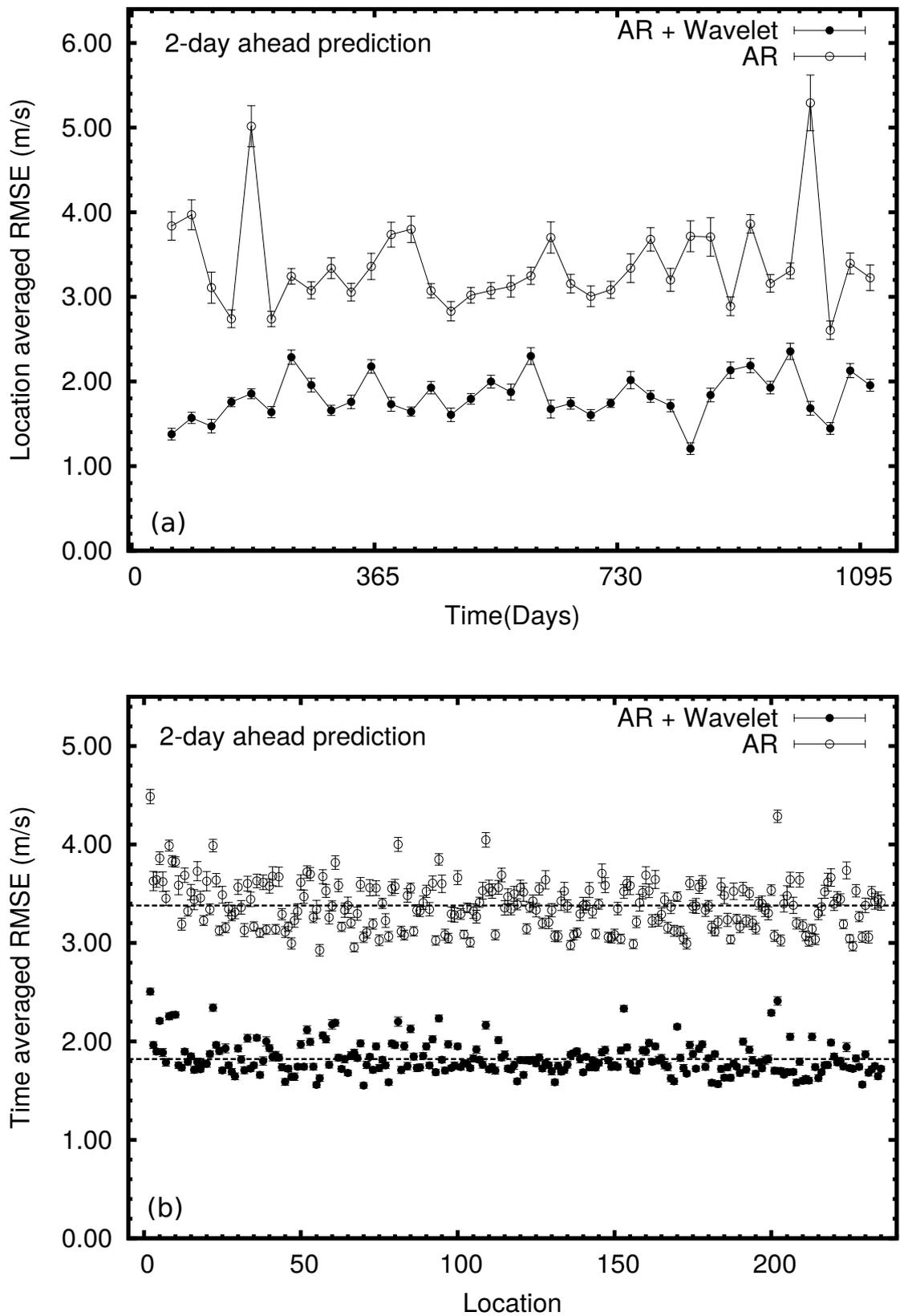


FIGURE 5.11: Plots of averaged root mean squared errors for 2 days ahead predictions at (a) different locations and (b) different time intervals.

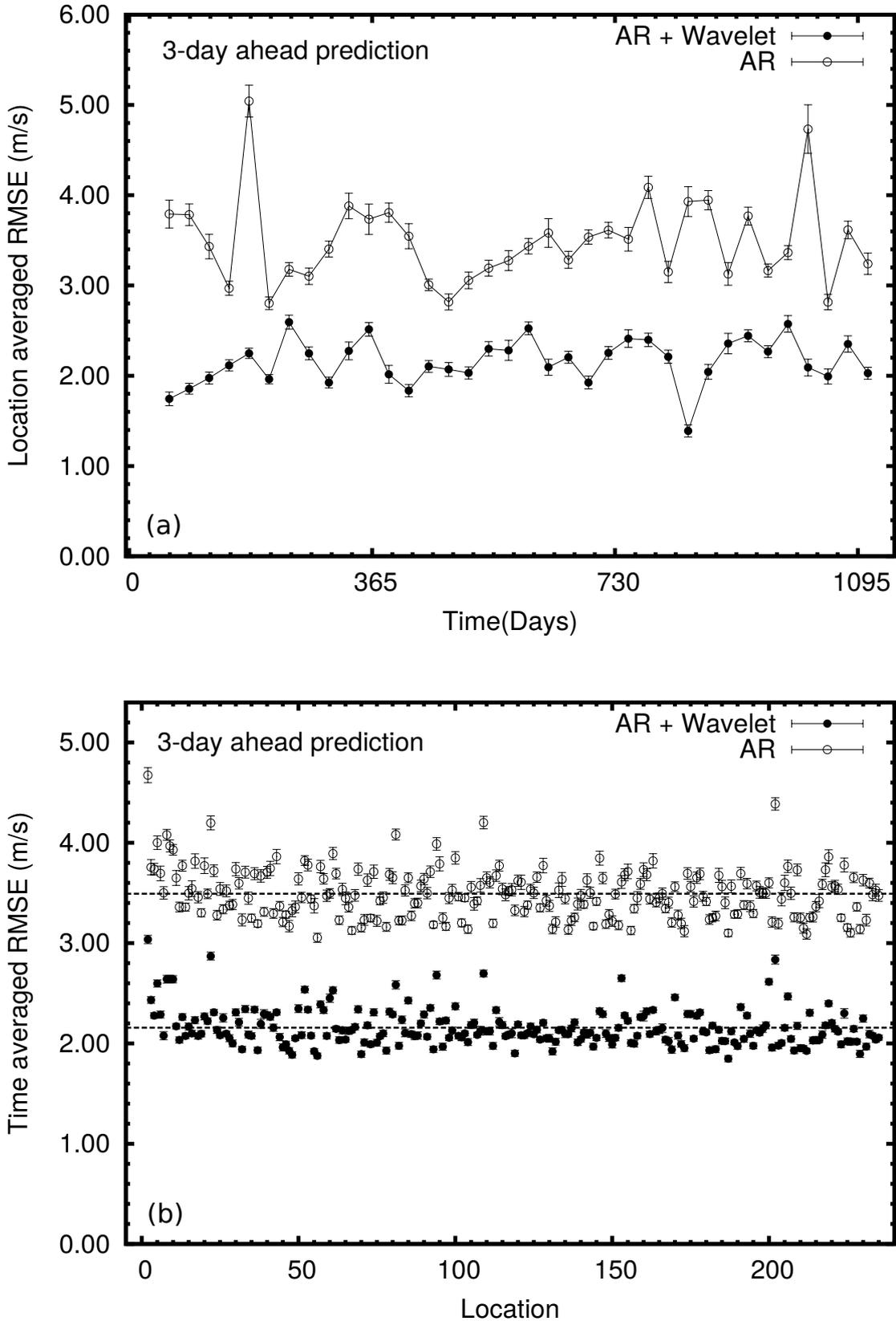


FIGURE 5.12: Plots of averaged root mean squared errors for 3 days ahead predictions at (a) different locations and (b) different time intervals.

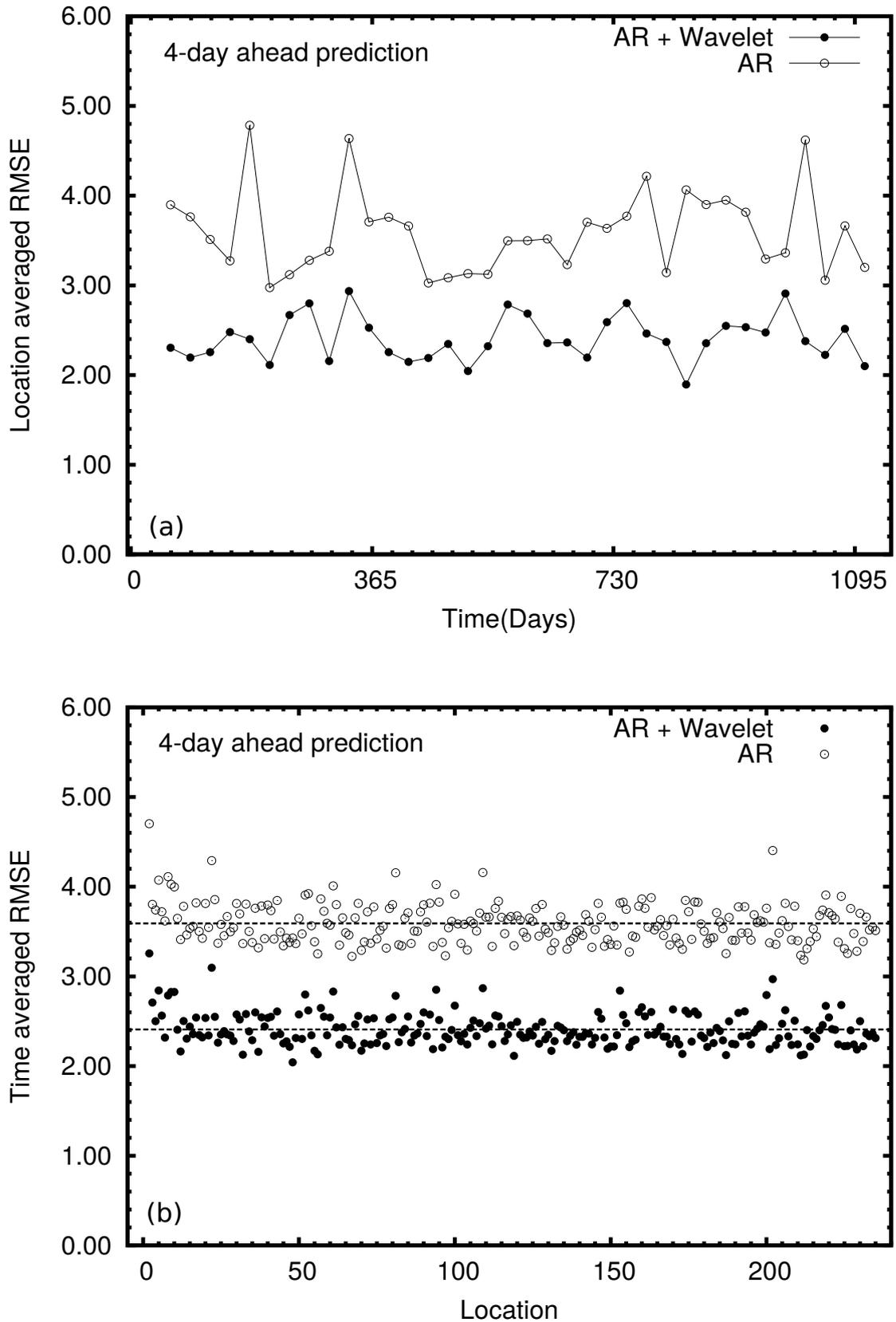


FIGURE 5.13: Plots of averaged root mean squared errors for 4 days ahead predictions at (a) different locations and (b) different time intervals.

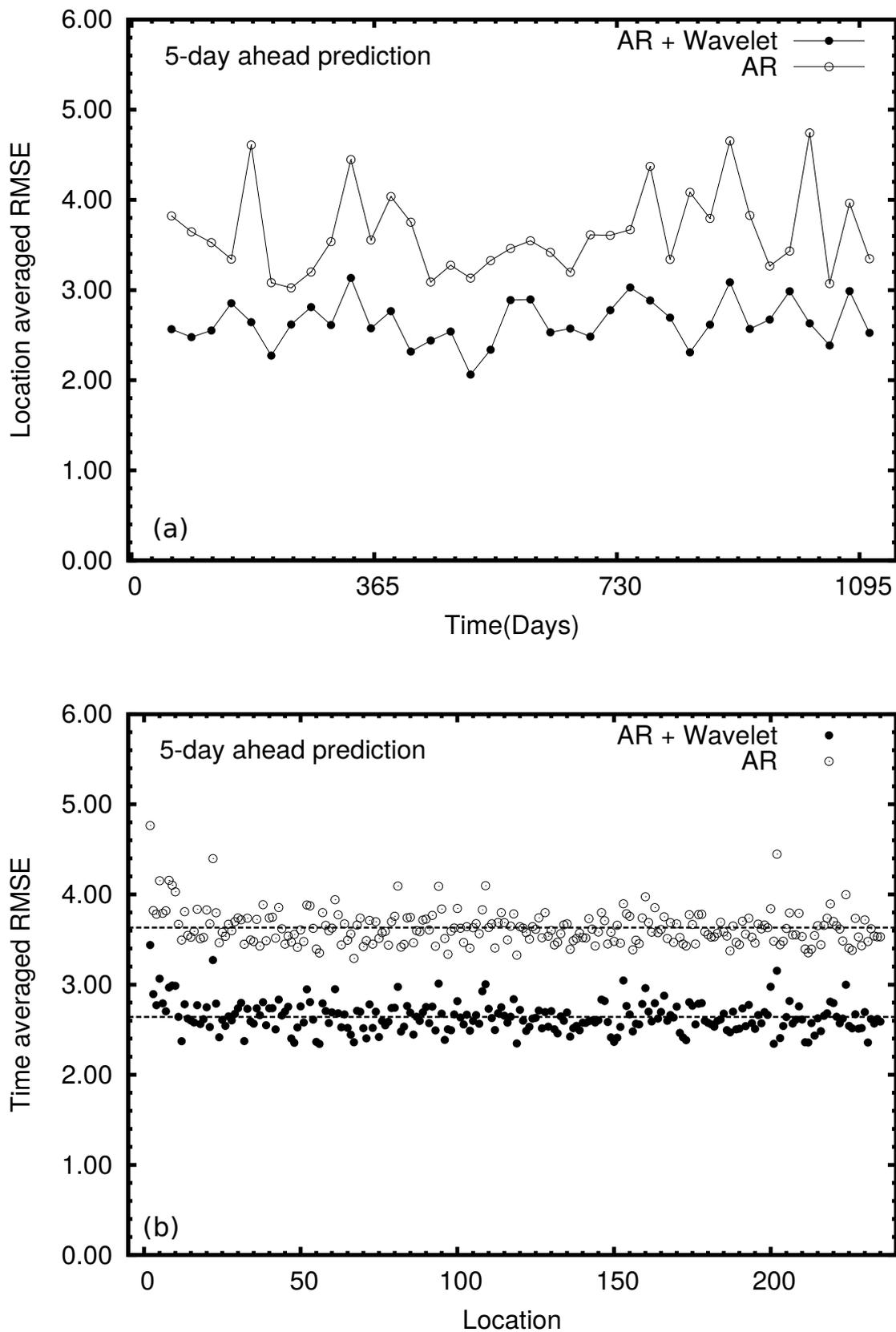


FIGURE 5.14: Plots of averaged root mean squared errors for 5 days ahead predictions at (a) different locations and (b) different time intervals.

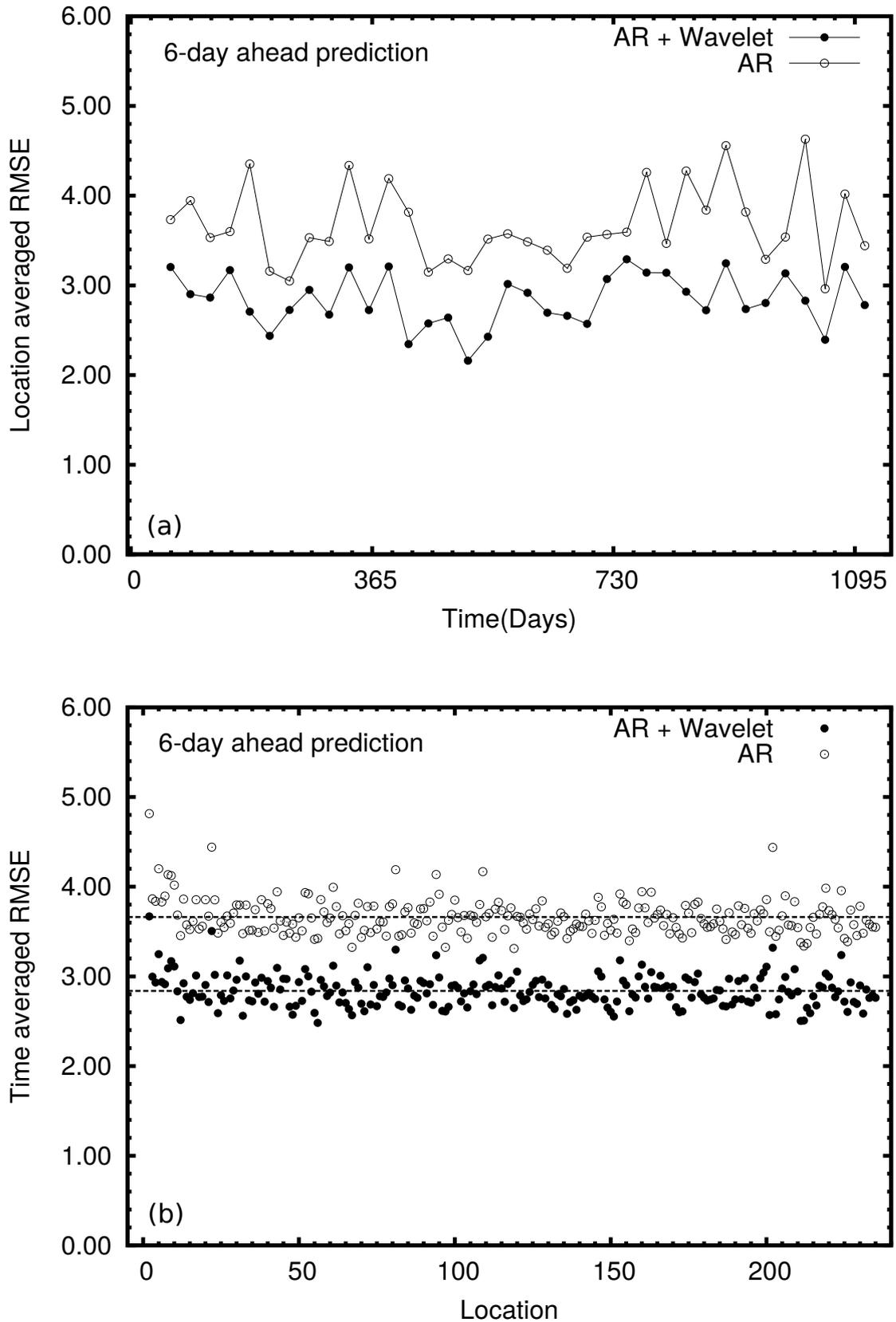


FIGURE 5.15: Plots of averaged root mean squared errors for 6 days ahead predictions at (a) different locations and (b) different time intervals.

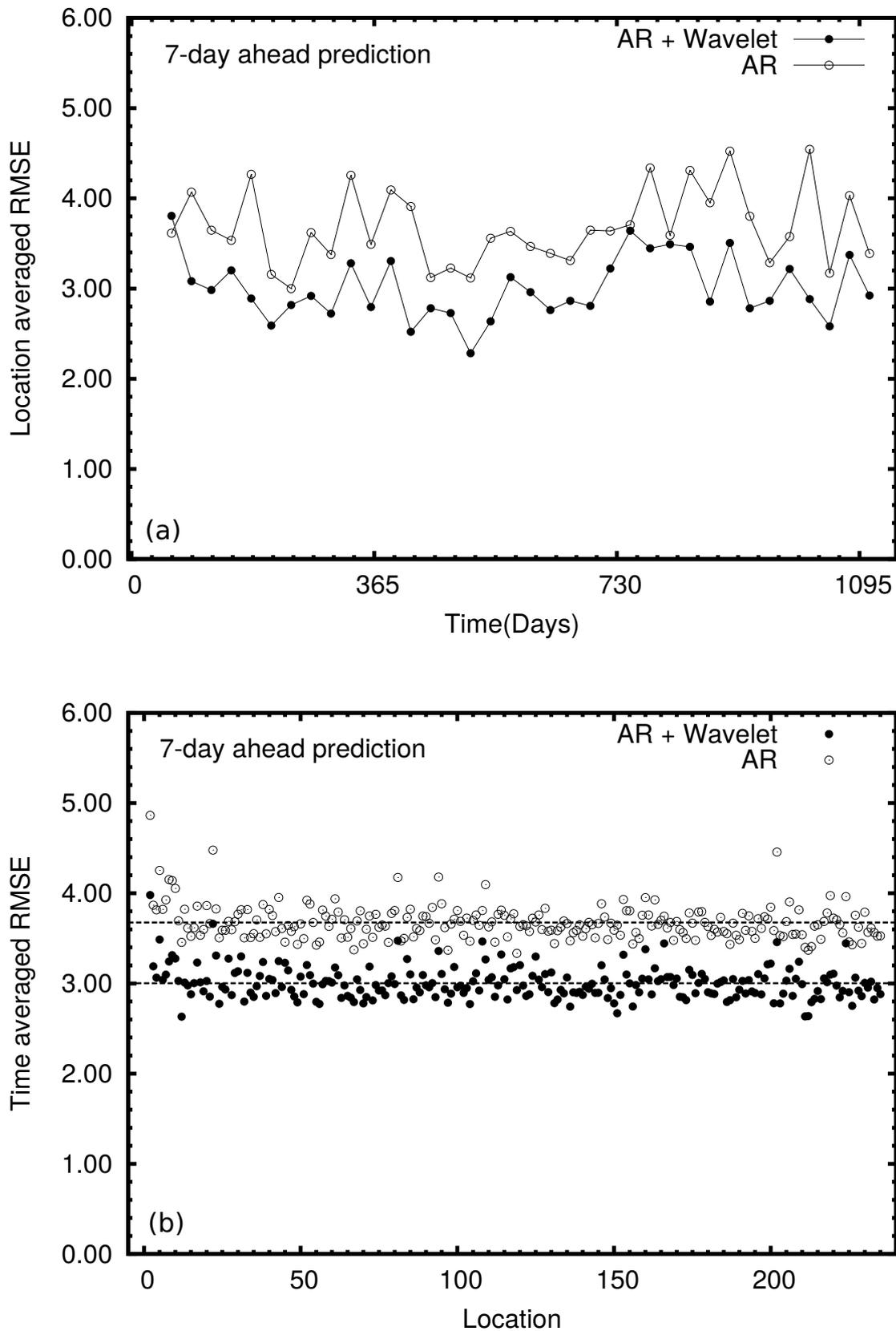


FIGURE 5.16: Plots of averaged root mean squared errors for 7 days ahead predictions at (a) different locations and (b) different time intervals.

The results clearly demonstrated that wavelet decomposition can significantly improve the prediction accuracy of the AR model and that the performance edge of the combined model is more or less maintained at the same level across all locations and all time periods. The order of AR model was between 34 to 36 in almost all cases. In predictions up to 3 days ahead, the combined model is accurate to within an average relative error of about 7–8%, which is roughly 5–6% less than what AR model produces when used directly. For 4 to 9 days ahead predictions the accuracy is certainly lower than the earlier set of forecasts, but the combined method of AR with wavelet decomposition continues to deliver much better accuracy than the individual AR method.

Fig. 5.17(a) shows the overall performance of forecasting methods for predictions up to 9 days ahead by plotting the time, and location averaged root mean squared errors against the number of days of predictions. For comparison, we have also computed the RMSE of predictions using the *New Reference Forecast Model (NRFM)*, which is an effective modification of the persistence forecast model for forecast lengths of more than a few hours (Nielsen et al., 1998). The AR prediction based on wavelet decomposition can make relatively accurate predictions for up to a week ahead, consistently across all locations and seasons, with an average relative error of about 11% or less. As shown in the Fig. 5.17(b) the major benefit of the method proposed in this paper is the significant gain in accuracy of the AR model obtained by the use of wavelet decomposition. Regarding location and time averaged RMSE, this gain is 60.15% for 1-day prediction and 46.24% for 2-days prediction. Comparison of the combined model with the *new reference forecast model* also shows similar performance advantage. For the week ahead forecasts, the use of wavelet decomposition returns an average 18.25% of accuracy gain for the AR model. The exponential decay of the accuracy gain, as seen in the figure, may be due to the chaotic behaviour of underlying dynamics of wind speed variations (Sreelekshmi et al., 2012; Drisya et al., 2014). Prediction accuracy of a chaotic time series is bound to decay exponentially as the time length of the predicted value increases. However, decomposition of time series before the prediction is seen to slow down this decay process appreciably.

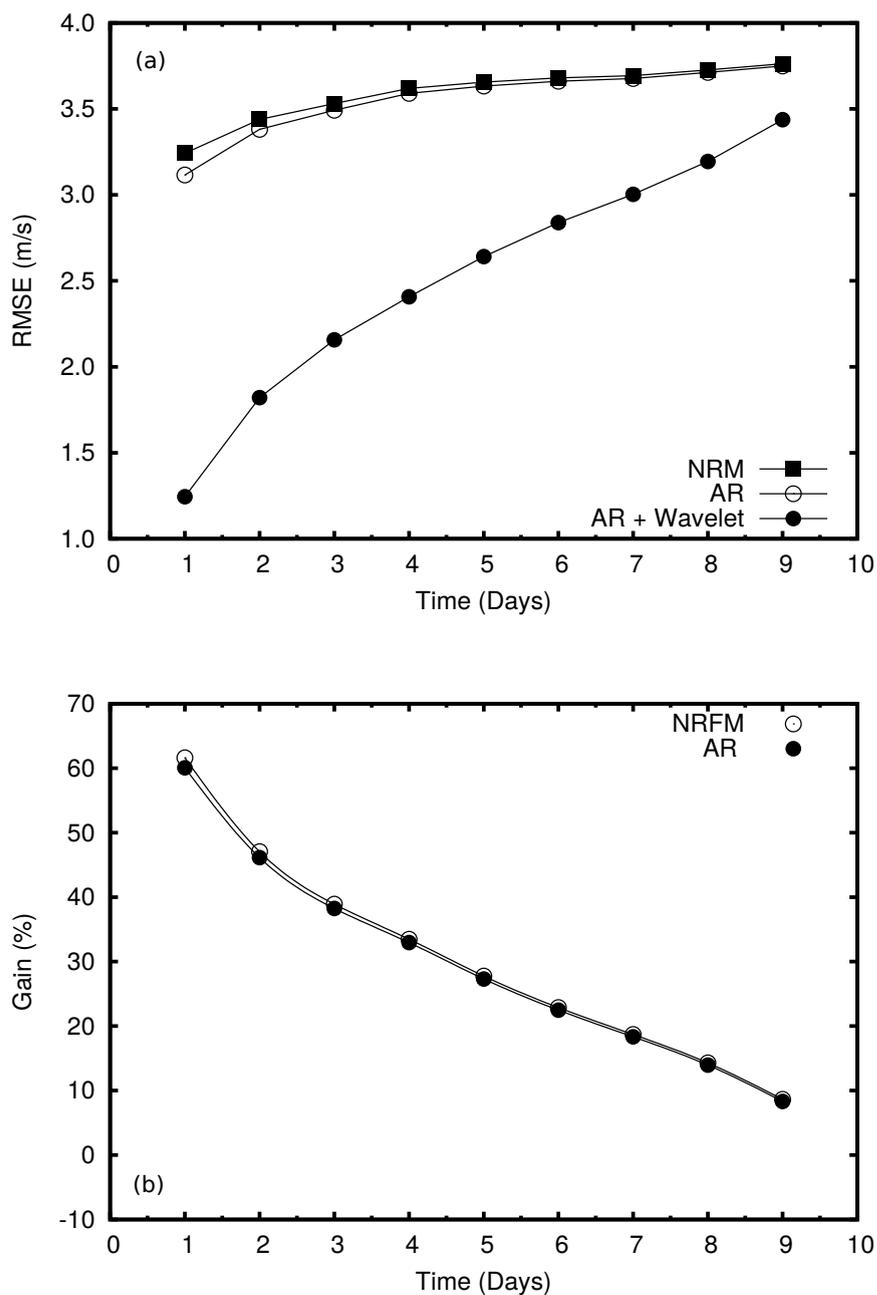


FIGURE 5.17: (a) Location and time averaged RMSE for predictions up to 9 days ahead. (b) Percentage gain of accuracy in terms of root mean squared errors for the *AR model* and the *New Reference Forecast Model (NRFM)* method when used in combination with wavelet decomposition.

5.5 Conclusion

Accurate prediction of wind speed is an important aspect of the control and management of electricity produced at the wind farms, and consequently, wind speed forecasting has emerged as a major research area in recent years. The majority of the techniques reported in the literature for wind speed forecasting use linear time series models. While attractive for their computational efficiency and simplicity, their prediction accuracy beyond a few time steps is feeble. In this study, we have demonstrated through numerical computations that the prediction accuracy of simple linear models can be remarkably improved, even for forecasts beyond a week, by properly decomposing the time series before prediction. Wind speed forecasts by an auto-regressive model combined with the wavelet-based decomposition of time series are found to be accurate within an average error of 7% to 8% for predictions up to 3 days ahead. A statistical analysis of the predictions made at 234 different locations reveals that time series decomposition improves the accuracy of AR models by an average 60.15% for day-ahead forecasts and up to 18.23% for week-ahead predictions. Since the entire analysis has been carried out on high-resolution data of 10-minute intervals, the results reported here are of greater practical relevance.

6

Diverse dynamical characteristics across the frequency spectrum of wind speed fluctuations

Wind speed oscillations are known to exhibit varying characteristics at different time scales and a range of models from simple persistence scheme to complex physical models has been used to capture this contrasting behaviour. Our recent analysis has shown that a collection of autoregressive (AR) models fitted separately on frequency component of wind speed time series can significantly increase the prediction accuracy indicating the inability of a single model capturing the entire range of behaviour possibly due to the diverse nature of dynamical characteristics. In this chapter we report the results of the investigation of diverse dynamical characteristics across the wide frequency spectrum wind speed measurements. The results show the variation of stochastic, deterministic and chaotic behaviour apart from the dimensionality of underlying dynamics as well as the degree of fluctuations. Such an analysis would be useful for adopting most suitable model for fluctuations at a specific range of interest or building hybrid models capturing the entire range of behaviour. It is also demonstrated that a cluster of deterministic models built upon separate frequency components of a wind speed time series can enhance the prediction accuracy as much as 80%, on the average, consistently for predictions up to 12 hours as validated by a statistical analysis of the predictions over a set of locations. The comparison shows definite advantage of deterministic prediction models over autoregressive models. The f -index introduced in this chapter to measure the fluctuations of wind speed over a period of time shows that the observed seasonal variations of prediction errors can be correlated with changes in the f -index of the component series contributed mostly by the lower scales of decomposition.

6.1 Introduction

The use of wind as a feasible, eco-friendly source of alternate energy has increased steadily over the past several years. There are many factors that make wind energy attractive over thermal or nuclear energy; it is available in abundance, pollution-free, sustainable in the long term and comparably cheaper to produce with minimum recurring costs. It is estimated that if the growth rate of wind power production continues at the present pace, it would account for about 12% of the total energy demands in the next five years (GWEC, 2012). A major impedance to harnessing wind energy to its full potential or the large scale deployment of windmills is the indeterminate nature of wind. Wind power is a function of wind speed which exhibits fluctuations at all time-scales due to numerous meteorological factors, and being able to predict wind speed and power accurately is a key factor affecting production and transmission of wind energy at various stages.

Short term forecasts ranging from milliseconds to a few hours ahead are useful in the operation and control of wind turbines and optimal utilization of wind power at electric power grids (Wang et al., 2012). Power trading based on bidding is common in liberalized electricity markets where in many countries wind power is also being connected to existing electric power grids, and forecasts of the expected availability of wind power for the next few hours or days ahead are needed to devise the best bidding strategy (Wang et al., 2012; Hering et al., 2010). Proper maintenance of wind farms and transmission lines require fairly accurate forecasts of wind power up to several days ahead (Aggarwal et al., 2013; Lei et al., 2009). Short and medium term predictions of wind speed and power has therefore emerged as an important research area in recent times.

The most important factor requiring accurate prediction for a reliable forecast of wind power is the wind speed, for which numerous models and methods have been proposed in the literature. The classical methods based on physical models built using various atmospheric parameters, though useful for long term predictions under stable atmospheric conditions, are unsuitable for the needs of the wind power industry which require predictions at much smaller time scales than what these models are meant for (Potter et al., 2006). Most of the methods reported in the literature use statistical models which use moving averages of past data or their probability distribution for making future predictions (Kavasseri et al., 2009; Hennessey Jr, 1977; Celik, 2004; Jiang et al., 2013; Shu et al., 2015). These methods are attractive for their universal applicability across most topographical conditions but do not significantly improve on the prediction accuracy compared to the elementary method of persistence (Sfetsos, 2002; De Giorgi et al., 2011). A simple modified version of persistence method known as reference forecast model is available for forecast length more than a few hours (Nielsen et al., 1998). Models based on artificial neural networks are reported to offer better prediction accuracy (Liu et al., 2013; Mohandes et al., 1998; Bilgili et al., 2007; Monfared et al., 2009), but this improved performance is not maintained across all locations (Soman et al., 2010). Attempts to improve the prediction accuracy of existing methods have also led to the development of hybrid models which combine different approaches to get over the weakness of each, such as mixing short-term and medium-term models or physical and statistical models. This often leads to better results than using each method separately (Soman et al., 2010; Liu et al., 2014; Haque et al., 2015; Liu et al., 2012; Celik et al., 2013; Liu et al., 2015).

In earlier works we had shown that the random like fluctuations in wind speed could also be due to an underlying dynamics which is deterministic and chaotic (Sreelekshmi et al., 2012). Wind speed dynamics is one of a few systems exhibiting chaotic behaviour outside laboratories. Applied on a time series of past data sampled at intervals of 10 minutes, the deterministic non-linear prediction methods could predict wind speed with remarkable accuracy up to one hour and reasonably good accuracy up to three hours. The performance of these models were found to be consistent across different topographical locations and over various seasons (Drisya et al., 2014). Since a chaotic attractor is an unlimited reservoir of unstable periodic orbits, a chaotic wind speed time series may exhibit a wide frequency spectrum and the numerous factors that affect the wind speed dynamics could manifest at various levels of this frequency spectrum. Some of these factors such as those arising from the rotation of earth and its revolution around the sun would be mostly deterministic, while those due to temperature or pressure variations could be possibly stochastic and the contributions by measurement errors might be purely noise. The interactions among coupled deterministic subsystems could also lead to chaotic dynamics resulting in behaviour which might easily be mistaken as stochastic. A frequency-level analysis of wind speed time series is therefore necessary to reveal the nature of the underlying dynamics that is predominant at various

frequency ranges. As of now, such an analysis at the frequency level, for understanding the key nature of the dynamics at various frequency ranges is not available in the literature. Apart from this most of these studies have considered wind speed fluctuations as a single time series, except for a couple of recent attempts to improve prediction accuracy using wavelet decomposition (Kiplangat et al., 2016; Tascikaraoglu et al., 2016). Our recent analysis shows that a cluster of AR models representing wind speed fluctuations at separate frequency range can better predict oscillations up to a week ahead in comparison to a single AR model for the entire frequency range indicating the possible diversity of dynamical characteristics at different frequency components (Kiplangat et al., 2016). The objective of this work is two-fold: to investigate the variation of dynamical characteristics and to demonstrate that a cluster of deterministic models, one for each frequency component, can better predict future oscillations compared to a single deterministic model fitted to wind speed time series for the entire range of frequencies. We also investigate if a proper deterministic model such as linear first order (LFO) method has an edge over AR model for short term prediction when fitted to frequency components independently. More specifically, we use wavelet decomposition to split the time series into component series of various frequencies (or inversely as scales/levels) and analyse each component for its deterministic characteristics, chaotic behaviour, dimensionality of underlying dynamics, variation of predictability with seasons, degree of fluctuations and the effect of fluctuations on the prediction accuracy *etc.* We report the results of the detailed analysis of the nature fluctuations of wind speed across its wide frequency spectrum and attempt to identify the regimes where the dynamical characters are predominantly stochastic, deterministic and chaotic. We also demonstrate that the performance of the deterministic prediction tools can be significantly improved by suitably combining them with wavelet decomposition. The additional gain of accuracy obtained by the use of wavelet filter on the deterministic prediction scheme is above 75% on the average throughout a 12-hours ahead prediction period. The comparison of prediction of linear first order (LFO) method and auto regressive (AR) model, both in combination with wavelet decomposition, shows about 60% gain of accuracy of LFO method over AR model for the interim range. We also introduce an index to quantify the fluctuation of wind speed over a period and investigate how it is correlated with prediction accuracy across seasons and frequencies. The analysis has been carried out on the wind speed data of 10 minute resolution available from National Renewable Energy Laboratory web site (<http://www.nrel.gov>).

6.2 Chaotic time series

The bounded aperiodic behaviour of a system that is sensitive to initial conditions is called chaotic (Ott, 2002). Sensitive to initial conditions means that trajectories diverge from one another exponentially fast, so that initially nearby trajectories may end up at far away places on the attractor within a short time. The time series of a dynamic quantity generated by the system may resemble the output of a stochastic system that exhibits random fluctuations. It is possible to detect and analyse chaos in a system by systematically investigating a single time series generated from the system, using the technique of attractor reconstruction with delay co-ordinates (Packard et al., 1980). There are several characteristics of the reconstructed system which can be used to analyse the dynamics of the underlying system in great detail.

There may be many variables contributing to the dynamics of a chaotic system but the trajectories eventually converge to an attractor, which though complex in its structure, may be low-dimensional with a non-integral dimension. Using delay co-ordinates, it is possible to calculate

an estimate for the dimension of the attractor of the underlying (Grassberger et al., 2004). A preferred choice for such an estimate is the correlation dimension (Hegger et al., 1999) which is computed from the correlation sum defined by

$$C(\varepsilon, m) = \frac{2}{N(N-1)} \sum_{i=1}^N \sum_{j=i+1}^N \Theta(\varepsilon - \|y_i - y_j\|),$$

where $\Theta(a) = 1$ if $a > 0$, $\Theta(a) = 0$ if $a \leq 0$. y_i is the delay vector for the time series $x(t)$, given by

$$\mathbf{y}(t) = (x(t), x(t - \tau), \dots, x(t - (m-1)\tau)),$$

for appropriate choices of the *delay* τ and the *embedding dimension* m . The correlation dimension is then calculated from the local slopes of $C(\varepsilon, m)$ versus ε for various values of m . For a time series coming from a deterministic chaotic system the correlation dimension may be much smaller compared to the embedding dimension, but if it originates from a stochastic system the two values may be of comparable magnitude.

The exponential rate of divergence of trajectories can be numerically measured by the so called Lyapunov exponents in the principal directions, and the existence of a positive Lyapunov exponent is considered a strong evidence for chaos in the system. Kantz algorithm can be used to compute an estimate for the largest Lyapunov exponent of a system from a time series originating from it (Kantz, 1994). This proceeds by first computing a stretching factor $S(\Delta n)$ defined by

$$S(\Delta n) = \frac{1}{N} \sum_{n_0=1}^N \ln \left(\frac{1}{\|U(y_{n_0})\|} \times \sum_{y_n \in U(y_{n_0})} \|y_{n_0+\Delta n} - y_{n+\Delta n}\| \right)$$

where y_{n_0} is a fixed point in the embedding space and y_n are the points in a neighbourhood $U(y_{n_0})$ of y_{n_0} . If the plot of $S(\Delta n)$ versus Δn shows a linear growth for various values of m then its slope gives an estimate for the maximum Lyapunov exponent.

6.3 Wavelet decomposition

Wavelet transform allows us to convert a data or function in the time domain into a representation involving different layers of frequency levels. The frequency components can then be observed and analysed at a resolution matched to its frequency. This is made possible by using, for the decomposition of the data, a basis consisting of simple functions called wavelets, which are small waves localized both in time and scale, all of which can be generated by scaling and translating a single base wavelet called ‘mother wavelet’. While the lower scale components give the finer microscopic details of the data at higher resolutions, the higher scales yield the grosser features at lower resolutions. Wavelets are most suited for the study of data such as that of wind speed which are inherently multi-scale due to contributions from numerous atmospheric and topographic factors.

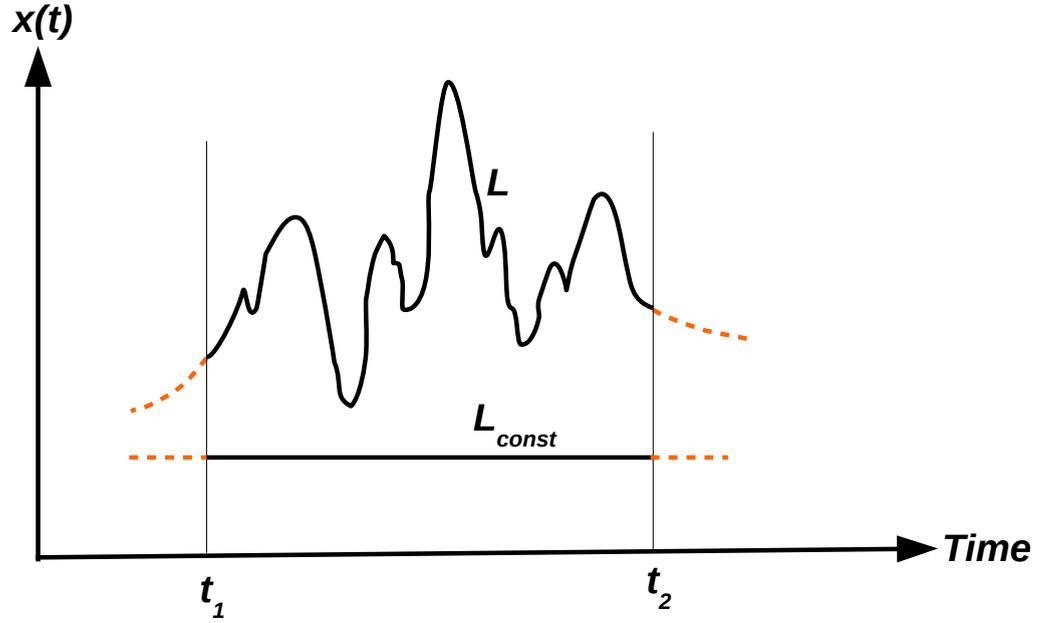


FIGURE 6.1: Diagram showing temporal fluctuations of signal $x(t)$.

Mathematically, a wavelet is a square integrable function $\psi(t)$, which is non-zero only on a finite interval and satisfies the *admissibility condition* (Daubechies et al., 1992),

$$0 < c_\psi = \int_{-\infty}^{\infty} \frac{|\hat{\psi}(\omega)|}{|\omega|} < \infty$$

where $\hat{\psi}(\omega)$ is the Fourier transform of $\psi(t)$. Wavelet transform uses a family of orthonormal functions $\psi_{j,k}(t)$ which are all generated by suitable dilations and translations of an appropriate *mother wavelet* $\psi(t)$;

$$\psi_{j,k}(t) = s^{-j/2} \psi(s^{-j}t - k\tau), \quad j, k \in \mathbb{Z}$$

where $s > 1$ and τ are fixed dilation and translation factors. The so called *dyadic sampling* corresponds to the choice $s = 2$ and $\tau = 1$.

The discrete wavelet transform (DWT) of a function $x(t)$ in the time domain is a discrete system of numbers defined by

$$W_x(i, j) = \int_{-\infty}^{\infty} x(t) \psi_{j,k}^*(t) dt, \quad (i, j \in \mathbb{Z}).$$

Conversely, if the set of wavelets $\psi_{j,k}(t)$ forms an orthogonal basis, the function $x(t)$ can be expressed as a linear combination of these transforms, called the *discrete wavelet decomposition* of $x(t)$, given by

$$x(t) = \frac{1}{c_\psi} \sum_{j,k \in \mathbb{Z}} W_x(j, k) \psi_{j,k}(t).$$

The discrete wavelet transform is well suited for analysis of data given in the form of a time series. If the data is to be decomposed into J levels of different scales, DWT requires the sample size N of the data to be a multiple of J . This restriction is removed in a special form of DWT called

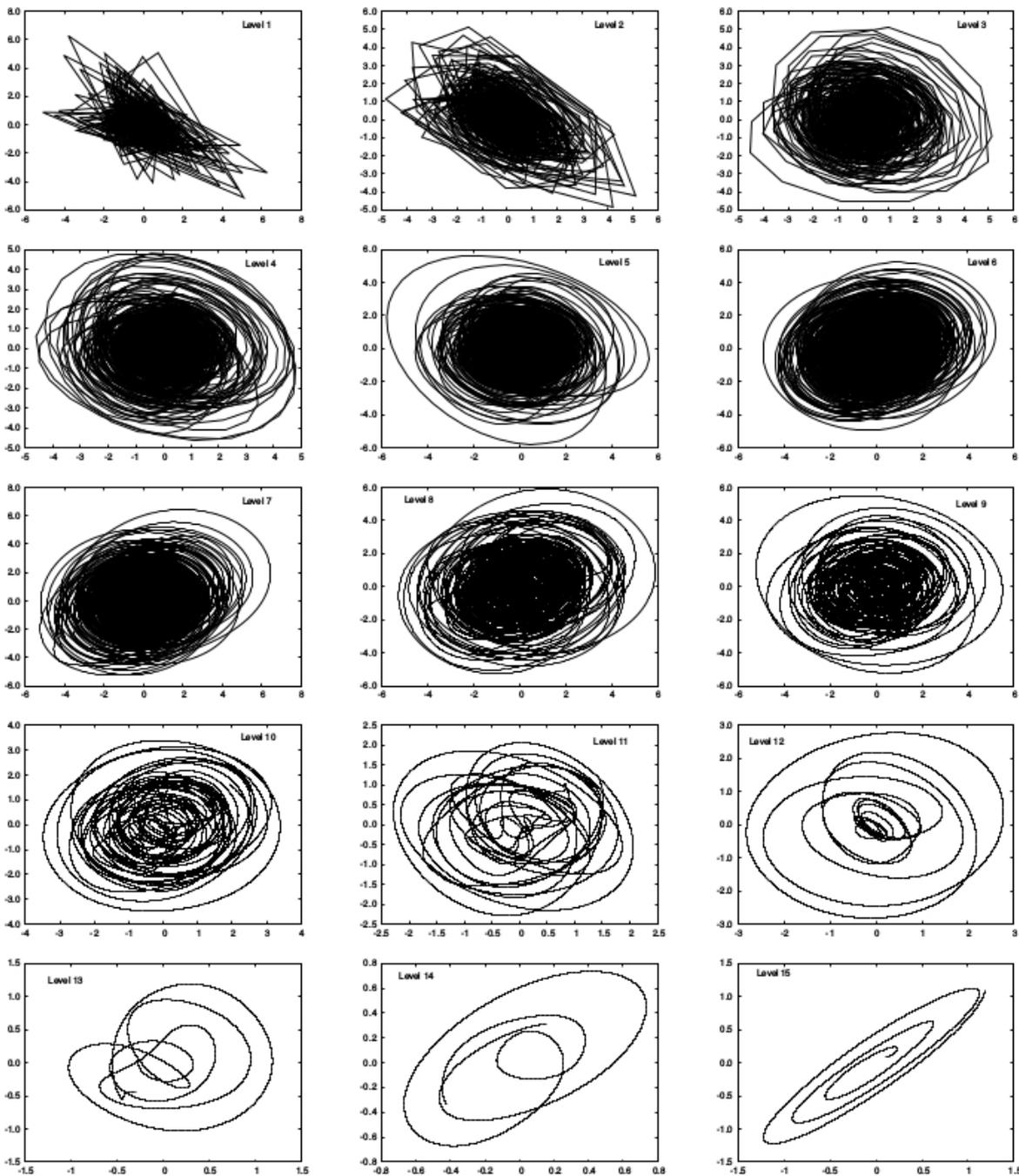


FIGURE 6.2: Delay reconstruction (x_n versus x_{n-d}) for each component of wavelet decomposed time series at location Latitude: 34.9842°N , Longitude: 104.03971°W . The delay d was estimated for each component using mutual information and auto-correlation.

Maximal Overlap DWT (MODWT), but this makes MODDWT highly redundant over traditional DWT. This redundancy incurred by the use of MODWT is, however, amply compensated by the possibility of better comparison of the series with its decomposition and MODWT has more or less become a standard in analysis of data using wavelet decomposition (Percival et al., 2006).

The decomposition using MODWT can be formulated as a multi-resolution analysis (MRA) in which the data is expressed as a sum of a smoother component and a set of components that give

details of the data at progressively higher resolutions. For an N -dimensional vector X representing the given data such a decomposition corresponding to J scales, $\tau_j, j = 1, 2, \dots, J$ reads as (Percival et al., 2006)

$$X = \sum_{j=1}^J D_j + S_J$$

The vector D_j captures the average variations on a scale of τ_j , whereas S_J contains the smoother component of the data involving the averages at scales $2\tau_j$ and higher (Percival et al., 2006).

6.4 Fluctuation Index

In general signals having lesser temporal variations are amenable to better modelling and prediction. The degree of fluctuations may vary with different periods or seasons depending on the local dynamics. To quantify the temporal fluctuations, we introduce an index named *fluctuation index*, shortly *f-index*, defined for a signal $x(t)$ over a period (t_1, t_2) as the fraction of the excess curve length of the signal over the length of a signal of constant magnitude. Thus,

$$\text{f-index} = \frac{L - L_{const}}{L_{const}} \quad (6.4.1)$$

where

$$L = \int_{t_1}^{t_2} \sqrt{1 + \left(\frac{dx}{dt}\right)^2} dt \quad (6.4.2)$$

is the arc-length of the signal over the period considered and $L_{const} = t_2 - t_1$ is the corresponding length of a signal of constant magnitude. (c. f. Figure. 6.1). In the next section we analyze how the *f-index* of the wind speed series affects the prediction accuracy and its seasonal characteristics.

6.5 Results and discussion

6.5.1 Analysis of decomposed data

The fluctuations in wind speed dynamics is the result of the interplay between numerous factors at various frequencies, partly random and partly deterministic. As a first step towards better understanding the nature of dynamics at various frequency ranges, we decompose the time series of wind speed into component series of suitably chosen frequency levels using wavelet transform. This is achieved by an MRA of the given time series using MODWT with Daubechies wavelet of order 8 and number of scales $J = 15$, which decomposes it into 16 time series, comprising of 15 detail series D_j at scales $2^j, j = 1, 2, \dots, 15$ and a smooth series S_{15} . Each of these time series then embodies the dynamical properties of the wind speed series at a certain band of frequencies active in the system. More insight into the nature of the dynamics at each frequency subsystems may be obtained by analysing each component series separately. Accordingly, for each component series we reconstruct the attractor of the underlying system in a suitable embedding space using the standard technique of attractor reconstruction with delay co-ordinates. More details about these methods and the subtleties involved are discussed in (Sreelekshmi et al., 2012; Drisya et al., 2014).

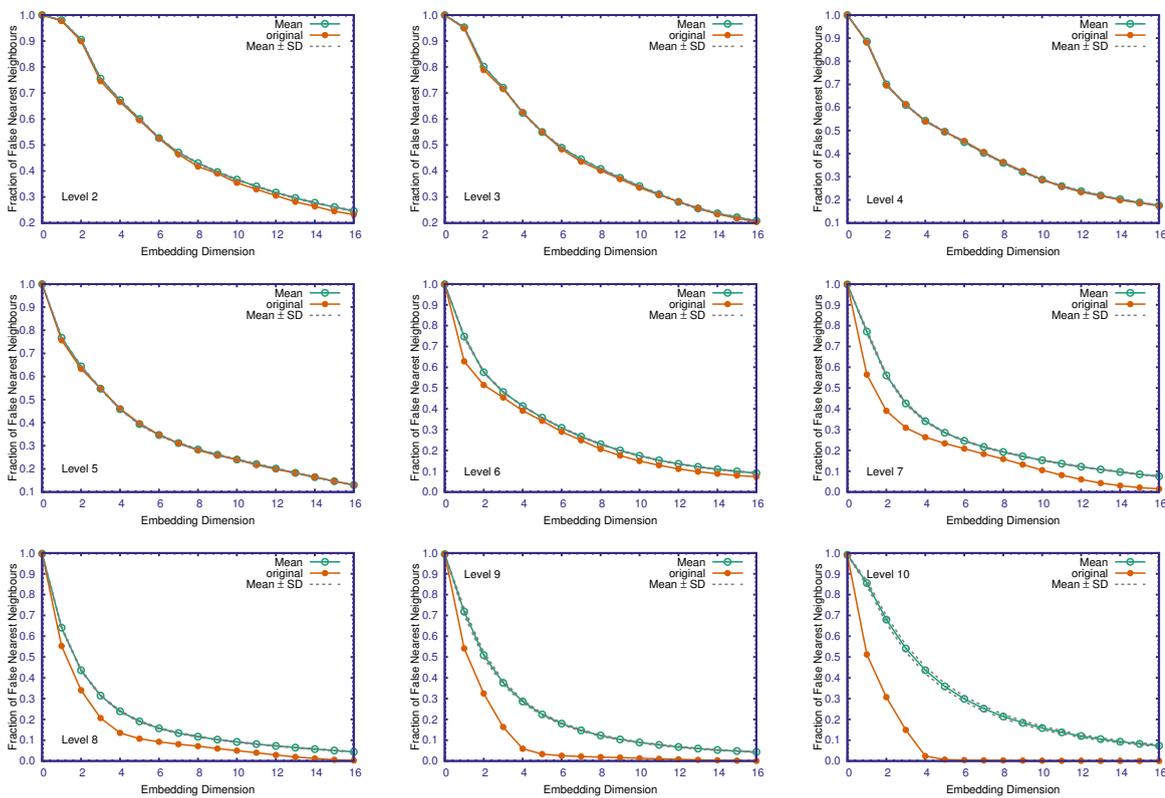


FIGURE 6.3: The mean values of the fraction of false nearest neighbours of the surrogates with standard deviation for levels 2-9 at location Latitude: 34.9842°N , Longitude: $104.03971^{\circ}\text{W}$.

Figure. 6.2 shows the delay representations of the attractor of the underlying system corresponding to each of the component time series constructed with suitable choices of the embedding parameters. A progressive variation in the structure of the attractor from more complex at higher frequency ranges to nearly periodic towards the lower end of the frequency spectrum can be clearly seen in the figure. The dynamics at higher frequency bands (lower levels) is evidently more complex than at lower frequencies. The complex structure of the attractors, at lower levels, could safely be assumed to be due to noise processes. However, the presence of complex dynamics at the intermediate ranges of frequencies could be more due to an underlying system which is deterministic and chaotic. The possibility of chaotic dynamics of wind speed time series has been demonstrated elsewhere (Sreelekshmi et al., 2012; Drisya et al., 2014), and here we can demarcate the frequency ranges that corresponds to stochastic, deterministic and chaotic behaviour of the underlying system by analysing the individual time series obtained by wavelet decomposition. To discriminate chaotic dynamics from stochastic behaviour we carried out a surrogate data test on each of the component series obtained by decomposition of a typical wind speed time series. The surrogate data test is a formal statistical method which involves testing the null hypothesis that the given time series is a linear Gaussian noise process. As first step in this direction we compared the fraction of false neighbours of the original time series of each frequency component was compared with 40 surrogates obtained using the algorithm of Schreiber and Schmitz (Schreiber et al., 1996) for the location given by Latitude: 34.9842°N , Longitude: $104.03971^{\circ}\text{W}$. Figure. 6.3 shows the fraction of false nearest neighbours for each of the original component time series along with the mean of the corresponding values with standard deviation for the 40 surrogates. The validity of the null hypothesis is tested based on the value of the significance of

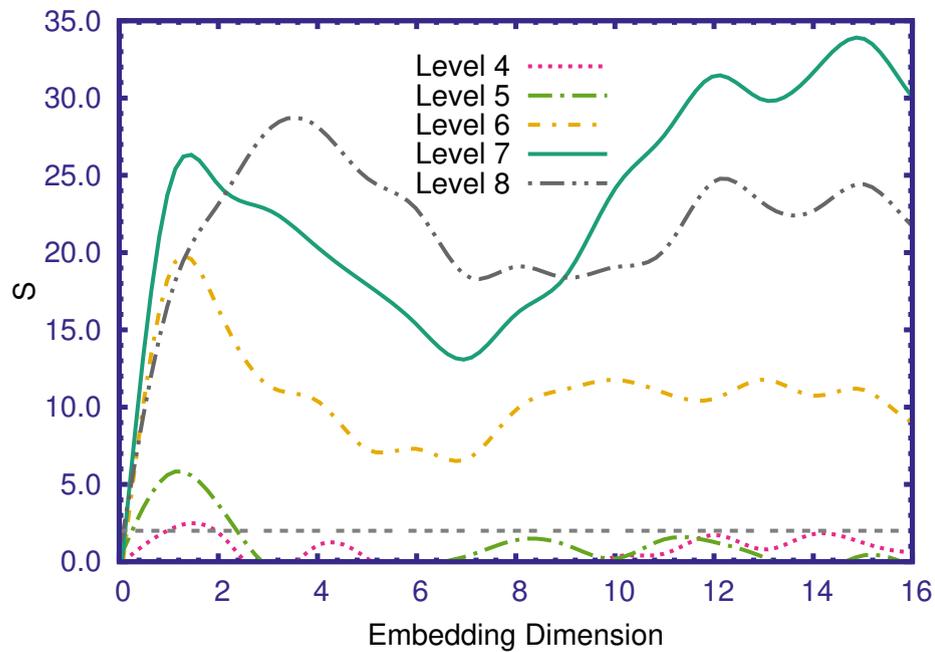


FIGURE 6.4: Plot of the significance of difference S versus embedding dimension for levels 4-8.

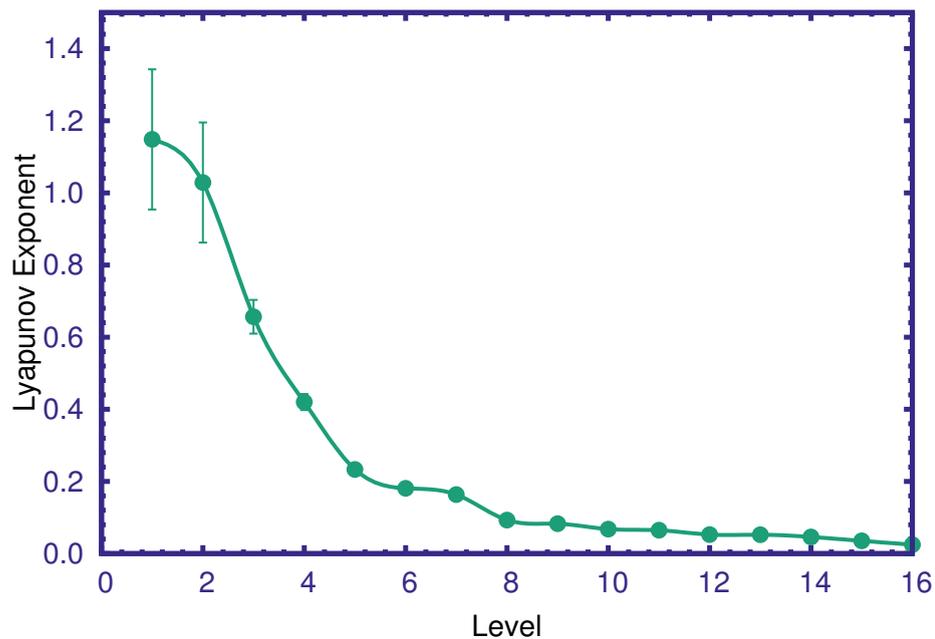


FIGURE 6.5: Variation of Lyapunov exponent for each wavelet component averaged over 212 sites.

difference computed according to (Mitschke et al., 1993)

$$S = \frac{\mu - \mu_{\text{orig}}}{\sigma} \quad (6.5.1)$$

where μ and σ mean and standard deviation of fraction of false neighbours calculated for surrogates and μ_{orig} mean of the original. The null hypothesis can be rejected if $S > 2$ with 95% confidence level. The plot of S for various levels is given in Figure 6.4.

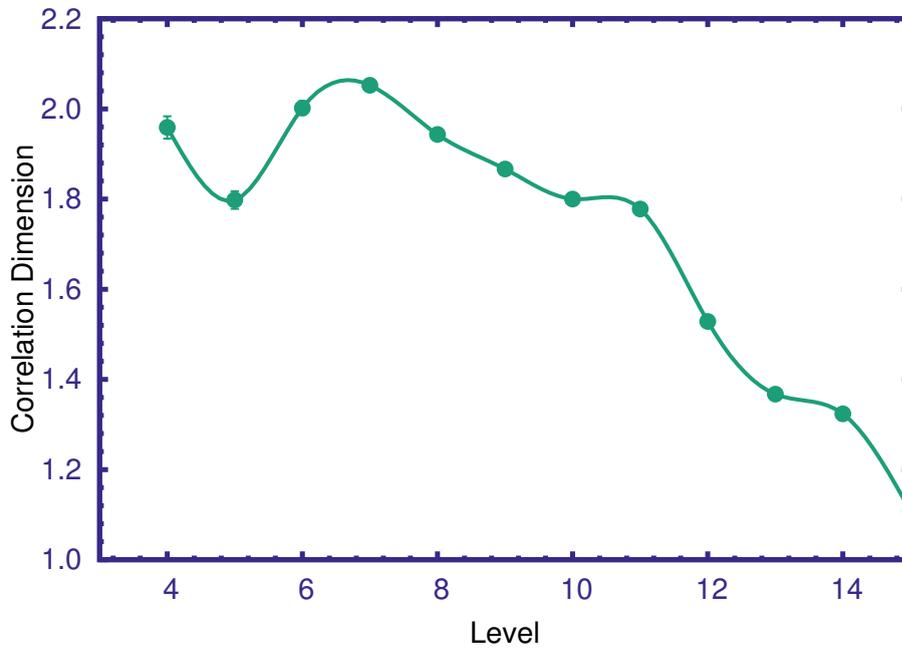


FIGURE 6.6: Variation of correlation dimension for each wavelet component averaged over three typical sites.

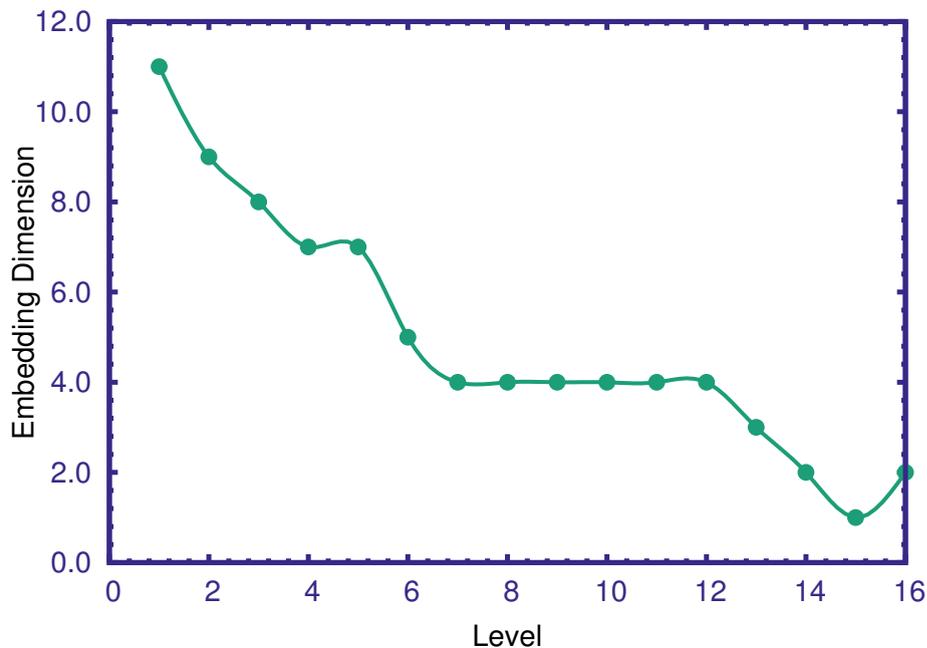


FIGURE 6.7: Estimated embedding dimension for each wavelet component.

It is seen that the values for the original data and the surrogate means start to differ from level five onwards. This indicates that while the high frequency variations up to level five are generally stochastic and may be considered as contributions by noise processes, the lower frequency variations embodied in the component series beyond level five are essentially deterministic in nature. This analysis, therefore, helps us to delineate frequency ranges of variations which arise due to random sources from those with underlying dynamics deterministic and chaotic.

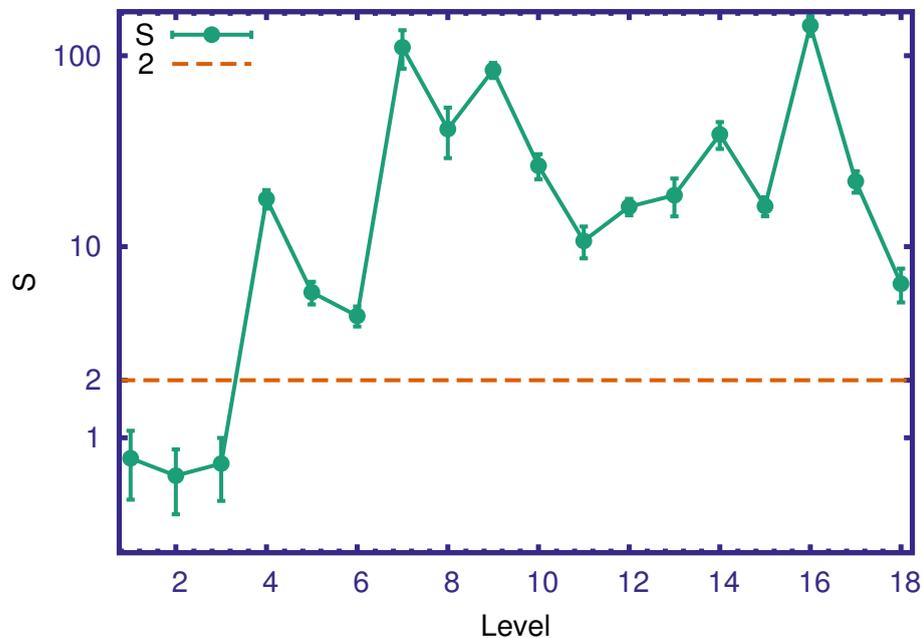


FIGURE 6.8: The average over 40 locations of significance of difference of one step prediction error at each level of the original and surrogates.

In order to validate this statistically we have computed the significance difference one step prediction error of the original and surrogates for each level at 40 locations. The average over 40 locations of S with error bar for each level is plotted in Figure. 6.8. It can be seen from the figure that fluctuations up to level 3 may have arisen from noise process.

Further confirmation as to the chaotic nature of the intermediate frequency components as well as a quantitative description of degree of chaoticity may be obtained by computing the Lyapunov exponents of the component subsystems. For sample computations we decomposed the wind speed time series from three typical sites using wavelet transform and computed the maximum Lyapunov exponents at each level and averaged over the three locations. Figure. 6.5 plots these averages for the various levels of decomposition. It can be seen from Figure. 6.5 that Lyapunov exponents decrease with increasing level, from a slightly high positive value at lower levels to nearly zero around higher levels. At lower levels these results corroborates the results of surrogate data test that the higher frequency variations are due to noise. At the intermediate levels this provides further evidence of a deterministic system exhibiting chaotic dynamics and at the lowest frequencies the dynamics is deterministic but not chaotic. Moreover, the variation of Lyapunov exponents with decomposition levels also suggests that more accurate predictions into longer periods of time would be possible at longer frequency ranges. This means that for long term wind speed predictions, which the wind power industry needs for long term energy management, using low frequency components as model data would yield better results. These results are also supported by the estimates of correlation dimension at various levels shown in Figure. 6.6. The diminishing dimensionality with increasing level is evident in the figure. The estimated embedding dimensions for a typical site also shows similar behaviour with a plateau region after level 7 as seen in Figure. 6.7. The plateau region indicates that the dynamics at these levels could have more or less identical characteristics.

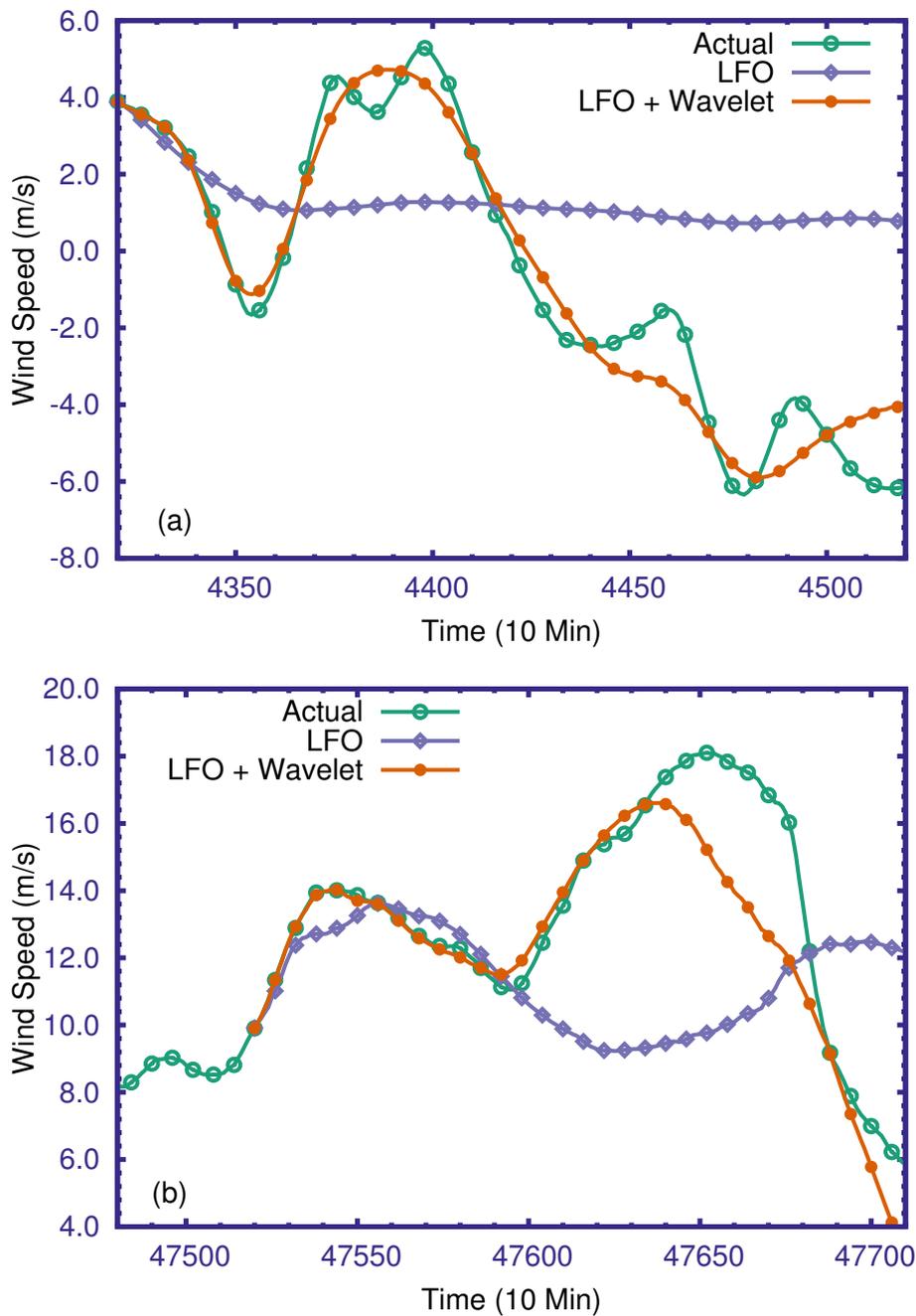


FIGURE 6.9: Comparison of LFO and LFO with wavelet decomposition predicted values with actual values of wind speed. The symbols are plotted only at every one hour for legibility. (a) Latitude: 45.87565 Longitude: -103.30938 (b) Latitude: 34.85728 Longitude: -103.61320.

6.5.2 Improved predictions

Given a series of n observations x_1, x_2, \dots, x_n , the time series methods utilize these observed values for predicting the likely values of the series a few time steps into the future, viz., x_{n+k} , $k = 1, 2, \dots$. Most of these methods use suitable combinations of the past values or past errors or their probability distributions to predict future values. Non-linear methods, on the other hand, try to approximate the evolution of the delay vector

$$x_n = (x_n, x_{n-\tau}, \dots, x_{n-(m-1)\tau}),$$

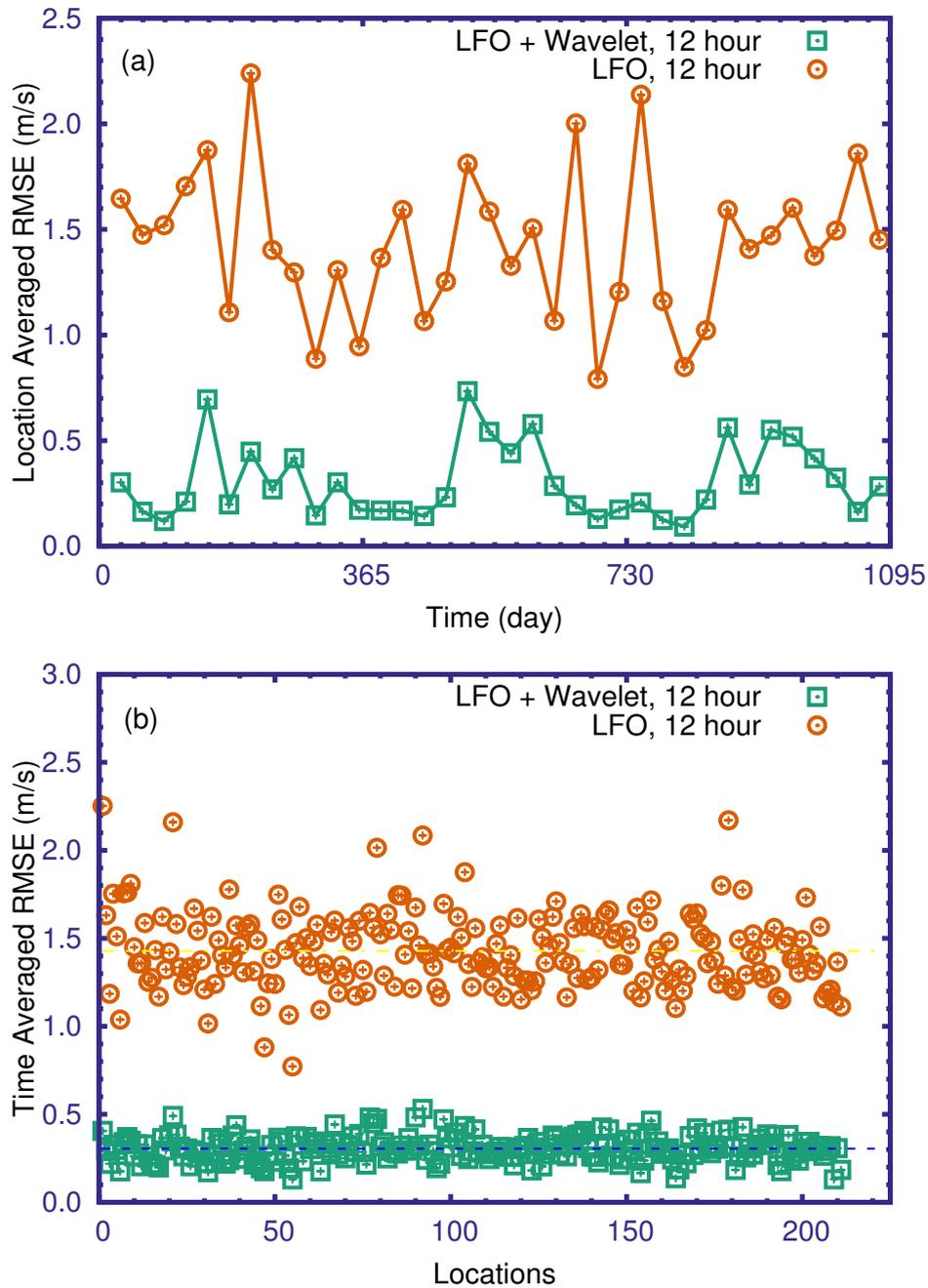


FIGURE 6.10: (a) RMSE of predictions for the period from 2004 to 2006 at an interval of 30 days averaged over 212 locations. (b) RMSE at 212 locations averaged over predictions at 30 days apart during the same period.

in the phase space. Thus for a deterministic system we can expect a functional relation of the form

$$x_{n+1} = f(x_n).$$

where f would be non-linear for a chaotic system. The local first order (LFO) method approximates f in each time step by a linear model fitted to the neighbors of x_n in an ε neighbourhood $U_\varepsilon(x_n)$, and has the form

$$x_{n+1} \approx A_n \cdot x_n + b_n.$$

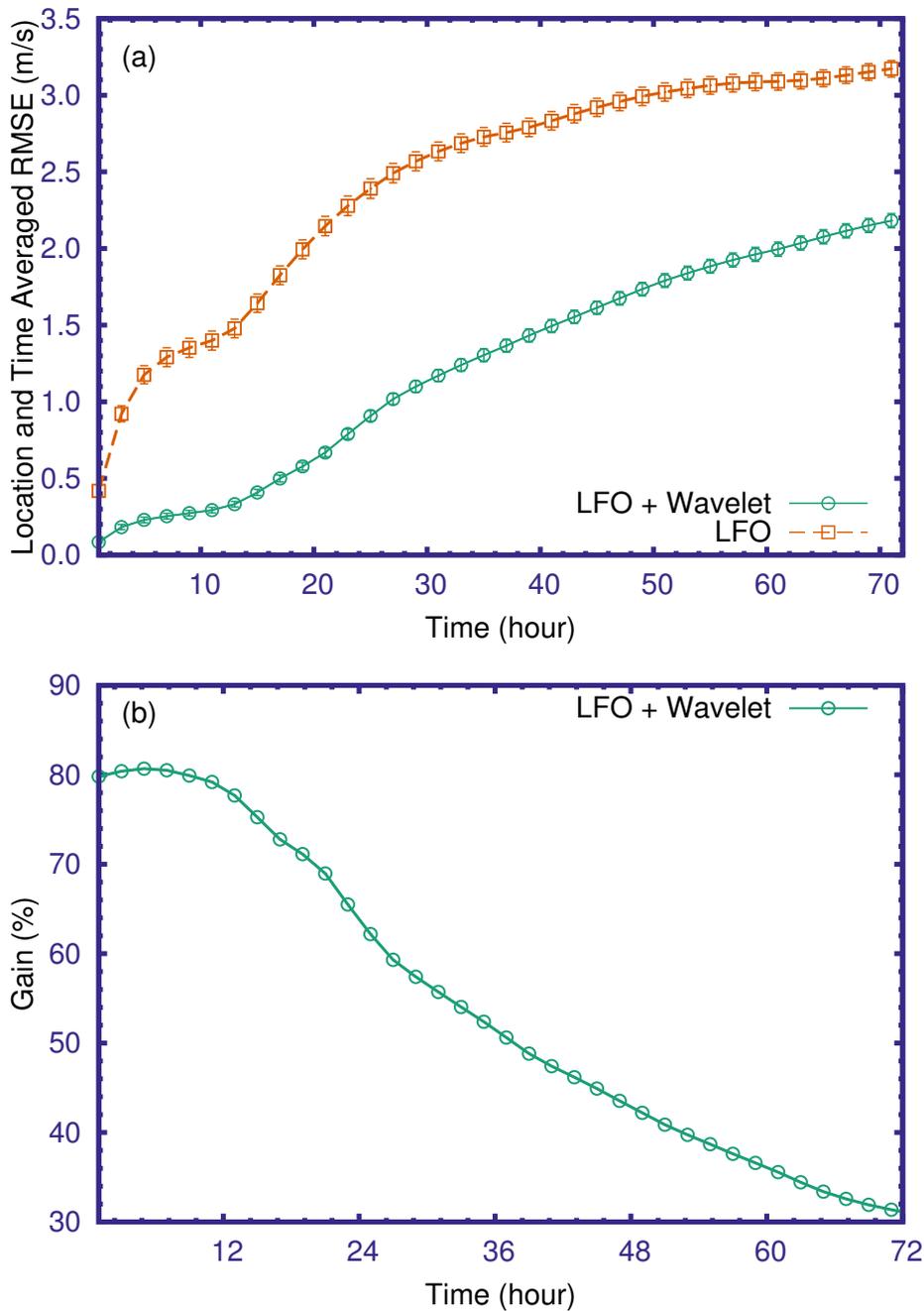


FIGURE 6.11: (a) Location and time averaged RMSE for predictions up to 72 hours ahead. (b) Percentage gain of accuracy in terms of RMSE for LFO method in combination with wavelet decomposition over LFO method. The symbols are plotted only at every two hour for legibility.

The method proceeds in steps, using a possibly different local linear model for each time step, thus giving a non-linear model globally.

In previous works (Drisy et al., 2014) we have shown that where the wind speed dynamics is chaotic, the non-linear deterministic methods can yield accurate short term predictions of wind speed up to 3 hours ahead with better prediction accuracy and longer duration of prediction than the traditional f-ARIMA methods. Here we demonstrate that the accuracy of the deterministic prediction schemes can be further improved by using wavelet decomposition of the model data.

The possibility of using wavelet decomposition for improving prediction accuracy of existing models had been demonstrated earlier for the case of simple linear forecasting models of wind speed (Kiplangat et al., 2016). To make comparisons easier, we use the same data set used in our previous analysis, namely the wind speed data of 10 minute resolution, for the period from January 2004 to January 2007, for 212 locations available from National Renewable Energy Laboratory (<http://www.nrel.gov>), USA.

For each set of predictions, we use a set of 4320 past data points to build the model and predict several time steps into the future, first using the LFO method alone and then using LFO combined with wavelet decomposition. In the combined method, the given time series is first decomposed into several component series at various scales (levels) using wavelet transform MRA as described earlier. Each component series thus obtained is then predicted several time steps into the future using LFO method, and the resulting series are then combined using MRA again to reconstruct the original series along with the predictions.

In applying the LFO method, the choice of the delay τ and embedding dimension m are important, which will also vary depending on the location of the data. The traditional methods based on autocorrelation function and fraction of false nearest neighbours need not always give the best of predictions (Domenico et al., 2013), so we used optimum values of τ and m for comparison of the prediction accuracy. Figure. 6.9 shows a comparison of typical predictions, for about 30 hours ahead at a couple of locations, first using LFO method alone and then by using wavelet decomposition before applying the prediction algorithm as described before. The results clearly show remarkable improvement in the performance of LFO when combined with wavelet decomposition.

6.5.3 Statistical analysis of predictions

We now extend the preceding analysis of the predictions using LFO with and without wavelet decomposition to a total of 212 different locations depicted in fig, and carry out a statistical analysis of the prediction errors to determine the consistency and practical applicability of the technique. Three types of averages of forecast errors are used in these calculations, namely the spatial averages of errors over the different locations, the time averages at each location over different periods of time and both time and location averaged forecast errors. Since the range of values of wind speed exhibit considerable variations over locations and time we use as a measure of prediction error the root mean squared error (RMSE) defined as follows. Suppose from a given time series of $n+k$ observed values x_1, x_2, \dots, x_{n+k} , we use the first n values for building the model to forecast the the next k values $x_{n+1}^p, x_{n+2}^p, \dots, x_{n+k}^p$, then the RMSE is then given by

$$\text{RMSE} = \sqrt{\frac{\sum_{i=n+1}^{n+k} (x_i - x_i^p)^2}{k}}$$

At each location we obtained forecasts for up to 72 hours, at several points of time 30 days apart, for a period of 3 years form 2004 to 2006, first using LFO method alone and then using LFO combined with wavelet decomposition, and computed the corresponding RMSEs. Each set of these errors are averaged over the various locations and Figure. 6.10(a) shows errors for 12 hour prediction for various points of time, 30 days apart, for the 3 year period. Figure. 6.10(b) shows the RMSEs averaged over time at each location and plotted against the locations. As is

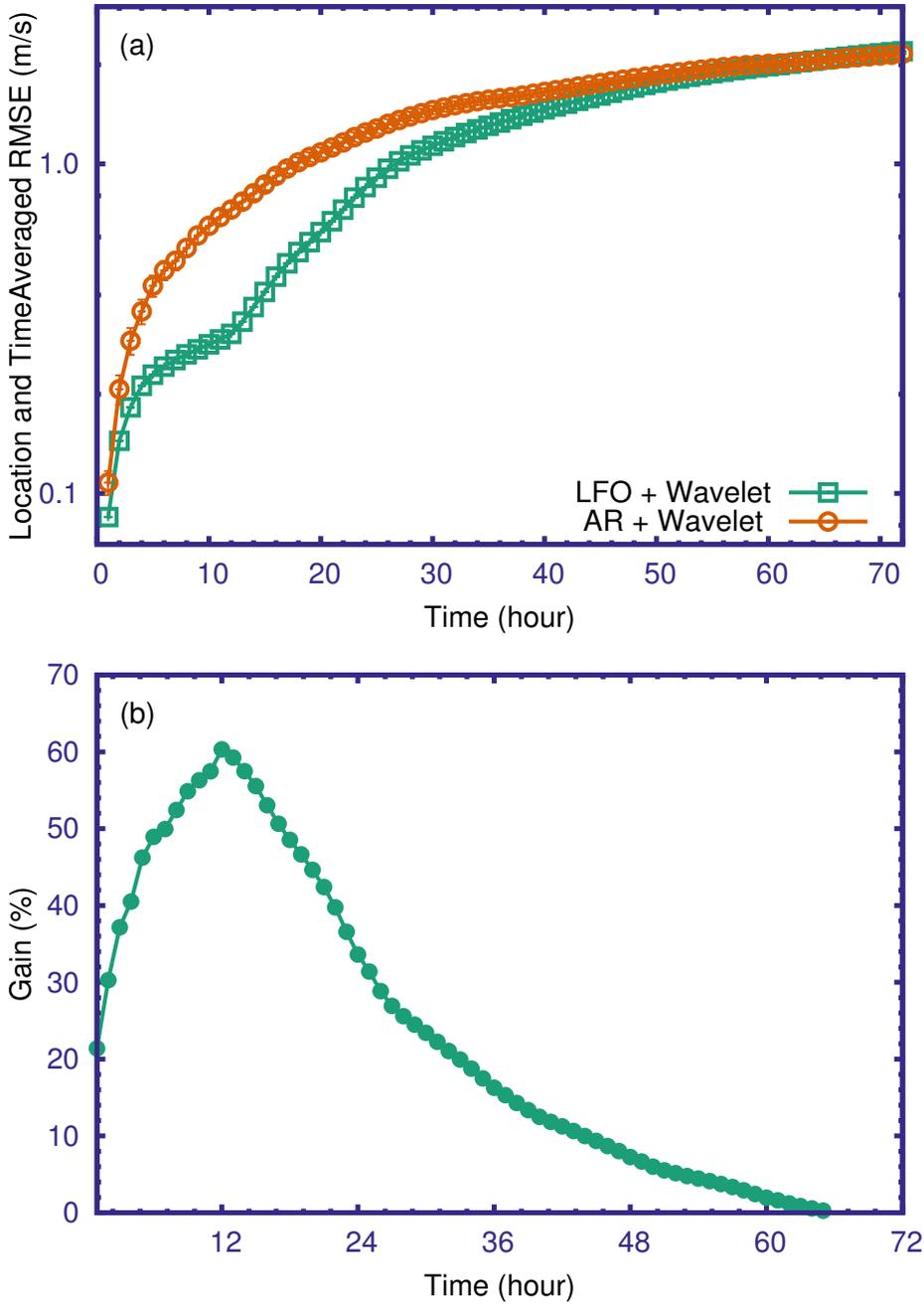


FIGURE 6.12: (a) Comparison of location and time averaged RMSE of AR model and LFO method both in combination with wavelet decomposition for predictions up to 72 hours ahead. (b) Percentage gain of accuracy in terms of RMSE of predictions of LFO method over AR model, both in combination with wavelet decomposition.

clear from the figures the combined prediction method consistently maintains better accuracy across all locations and over all periods of time comprising various seasons. Figure. 6.11(a) shows the RMSEs averaged over both time and location for predictions up to 72 hours ahead, and Figure. 6.11(b) the gain in accuracy, as measured by the averaged RMSE, earned by the use of wavelet decomposition before applying prediction algorithm. The combined method returns an average 80% gain in accuracy for 12 hours ahead predictions but beyond 12 hours, while the performance edge is maintained, the gain in accuracy gradually falls off. The improvement in the

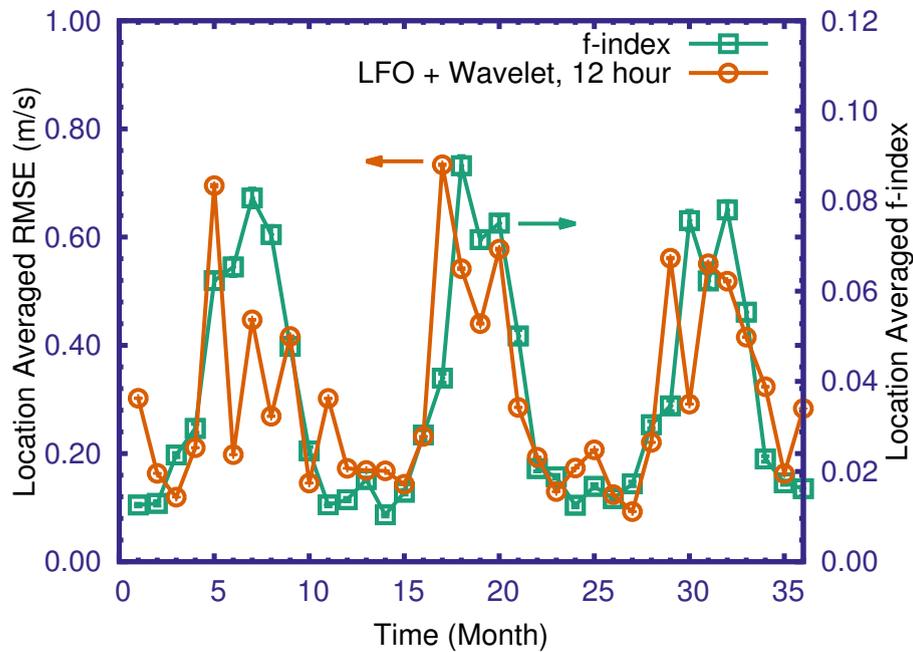


FIGURE 6.13: Location averaged RMSE of 12 hour predictions of LFO method with wavelet decomposition at an interval of 30 days and corresponding location averaged f -index. The correlation between RMSE and f -index is 0.697.

performance of LFO when paired with wavelet decomposition can be explained on the basis of Figure. 6.5, which shows that the maximum Lyapunov exponent decreases as the decomposition scale increases (frequency decreases). Smaller values of Lyapunov exponents indicate better predictability, hence component series at higher scales can be predicted more accurately for longer periods. So when forecasts on component series are made separately and recombined, the overall prediction accuracy benefits from the better predictions on the higher level components.

We have also compared the performance of the LFO model with that of the Auto Regressive (AR) model for prediction Kiplangat et al., (2016), both in combination with wavelet decomposition. We might expect the LFO method to perform better as it captures the dynamics in the embedded space whereas the AR model is based on a linear combination of previous data values which does not take into account the dynamics or dimension of the phase space of the underlying system. Figure. 6.12 plots RMSE of predictions of both methods and gain of prediction accuracy of LFO method over AR model. As can be noted from these figures, for short term predictions LFO method is comparably better than the AR model, but with increasing prediction time the predictions by the LFO falls off to the level of AR model.

A closer look at Figure. 6.10(a) and Figure. 14 in reference Drisya et al., (2014) would reveal smaller variations in the prediction errors with change in seasons, the prediction error being the lowest during winter. This may be explained in terms of the seasonal variations in the extent of fluctuations exhibited by the component series, especially at lower scales dominated by noise. The degree of fluctuations of the series at each scale of decomposition may be quantified by the monthly standard deviation of variations, or more faithfully by using the fluctuation index (f -index) introduced earlier. For comparison, the location averaged RMSE of predictions of LFO method in combination with wavelet decomposition along with f -index of model data (without decomposition) is plotted in Figure. 6.13. The estimated value of correlation between these two

is 0.697. As is evident from the figure, the prediction accuracy is comparatively higher in winter where f -index relatively low. To further examine the impact of seasonal variations on different frequency components, we have computed the standard deviation and f -index of corresponding model data for each components and plotted in Figure. 6.14. It is seen that at lower scales there is considerable variations in the range of fluctuations, highest during summer and lowest in winter. This could be due to the impact of solar radiations, the effect of which is the highest during summer, and its diurnal nature which contributes to the high frequency variations in the lower scale components. This means that the seasonal variations in the prediction accuracy of wind speed could be more due to the frequent fluctuations at the lower scale components resulting from solar activity, and could be related to the seasonal variations of the fluctuation index meaningfully. This can in turn be utilized for developing better season dependent models for wind speed forecast.

6.6 Conclusions

The dynamical system underlying the apparent random oscillations of wind speed could be the resultant of many coupled sub-systems. These subsystems may vary from very complex to simple periodic phenomena. We have presented an analysis of wind speed fluctuations based on the dynamics embodied by various ranges of constituent frequencies. The study bases on the wind speed data from a set of locations and uses wavelet transform technique to decompose the original time series into a set of constituent series at various scales associated with different frequency ranges. Each of these component series were then analysed using tools of non-linear analysis such as attractor reconstruction, Lyapunov exponents and correlation dimension along with surrogate data analysis to differentiate stochastic behaviour from deterministic dynamics. The results provide a broad classification of the dynamics at various scales, with the dynamics up to level five showing purely random character while beyond five the dynamics is essentially deterministic. Even in the deterministic realm, the dynamics is complex with strange attractors and positive Lyapunov exponents in the intermediate range of frequencies showing that the underlying dynamics is chaotic at these levels. The lowest range of frequencies at the highest scales of decomposition, however, are almost non-chaotic and possibly results from periodic phenomena.

Another advantage of using wavelet decomposition of wind speed series is that when used in conjunction with deterministic time series prediction methods, it can significantly improve the prediction accuracy of the latter. Where wind speed dynamics is predominantly deterministic, forecast tools such as LFO have been shown to be very effective in short to medium term wind speed predictions. However, as we have shown in this work, using wavelet decomposition prior to prediction can significantly improve the prediction accuracy as well as duration of prediction. The gain in accuracy obtained by the use of wavelet decomposition goes as much as 80% on the average on predictions up to 12 hours ahead. A statistical analysis of the predictions made at a total of 212 different locations asserts that this improved performance of the new model is consistent across different locations and periods of time. The LFO method in comparison with AR model, both in combination with decomposition, is better upto 20 - 60% over the AR model for up to a day ahead prediction. However, with increasing prediction time, the predictions by the LFO method falls off to the level of AR model. The analysis also reveals variations of the accuracy of predictions on a finer scale with change in seasons, presumably due to the diurnal nature of solar activity, which can be related to the degree of fluctuations of the component series at lower levels of decomposition.

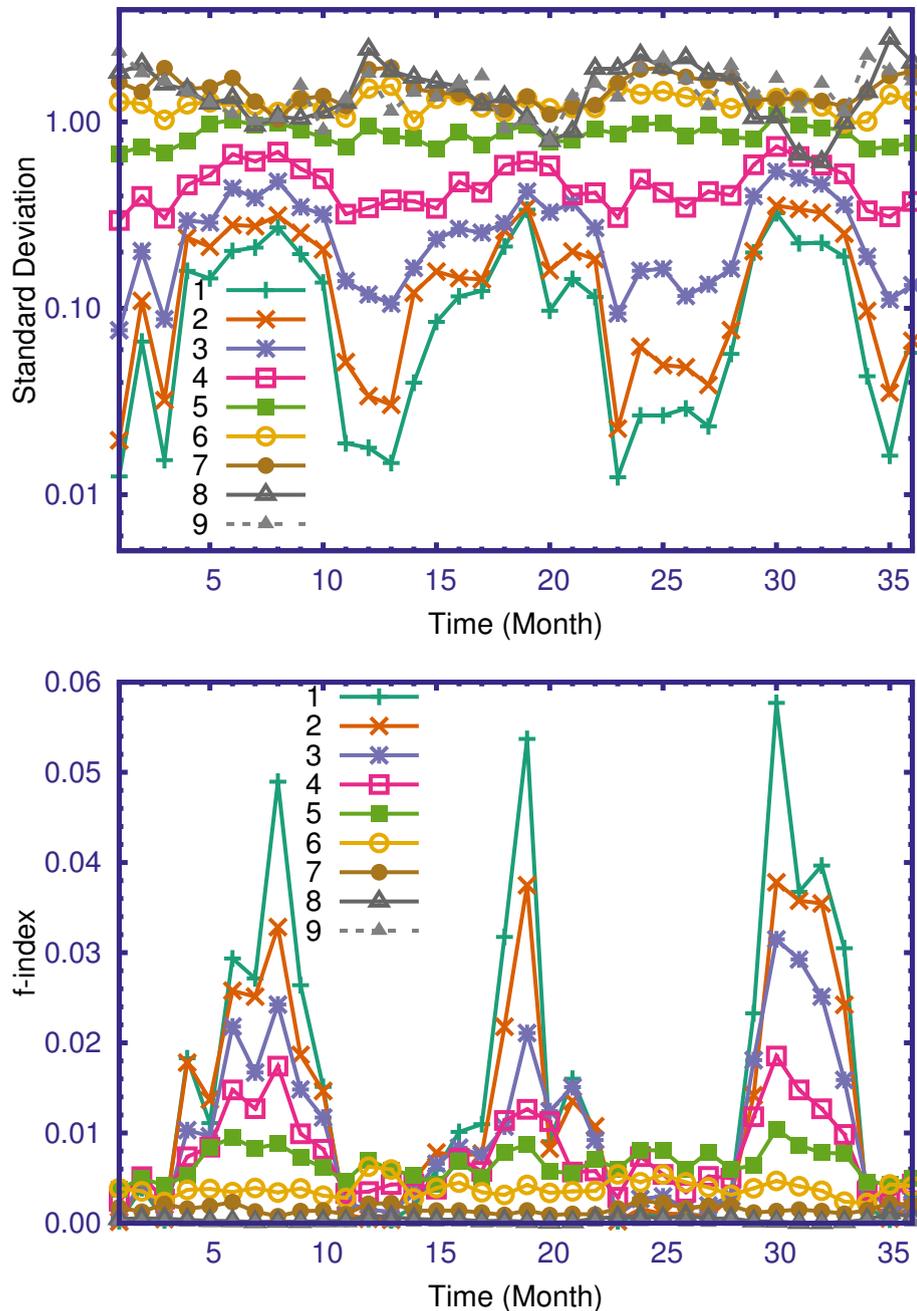


FIGURE 6.14: (a) Standard Deviation of the wind speed variations for each month at different level and (b) Fluctuation index for each month at different level at location latitude: 34.9842°N , longitude: $104.03971^{\circ}\text{W}$.

The change in the fundamental character of the dynamics of wind speed data across the decomposition levels also opens up the possibility of using hybrid models for prediction, where the component series at each level of frequency range can be predicted using tools most appropriate for the dynamical characters exhibited by that particular series, and then reconstructed to yield better forecast results. This will be explored in future works.

7

Summary

In this work, (a) we have carried out a detailed non-linear time series analysis of the daily mean wind speed time series for eleven years of nine locations across Indian subcontinent with the objective to investigate the dynamical characteristics of the underlying dynamical system, (b) we demonstrated that deterministic forecasting methods could make accurate short-term predictions of wind speed using past data, at locations where the wind dynamics exhibit chaotic behaviour and (c) we showed the existence of diverse dynamical characteristics across the frequency spectrum of wind speed fluctuations and demonstrated that a cluster of deterministic models built upon separate frequency components of a wind speed time series can enhance the prediction accuracy as much as 80%, on the average, consistently for predictions up to 12 hours as validated by a statistical analysis of the predictions over a set of locations. We know that the wind speed variations are affected by myriads of factors. A good number of tools have been developed for the prediction of wind speed variations. A significant share of these models assumes that wind speed oscillations are stochastic in nature. However, results of the analysis of wind speed data from several locations in India over a period of one decade strongly suggest that the fluctuations are indeed deterministic, low-dimensional and chaotic opening the possibility of developing accurate short-term prediction tools based on deterministic models. Irrespective of the location, the estimated values of the correlation dimension and the fraction of false neighbours clearly shows the low-dimensionality of the system. The wind speed time series under consideration yield a positive maximum Lyapunov exponent revealing the chaotic behaviour of the system. The detailed surrogate data test has also been carried out on the wind speed data from several locations, and the results rule out the possibility of the underlying dynamics being stochastic, suggesting the deterministic chaotic nature as the reason for the apparent temporal fluctuations. The colour noise test conducted also corroborates the deterministic character of the system. The most interesting point is that the wind is one of the natural system showing chaotic behaviour. The analysis fails to show any significant dependence of the degree of chaos on the variation of latitude and longitude.

After establishing the deterministic character of the system, we next attempted to develop better deterministic short-term prediction models. In the third chapter, we demonstrate the suitability of deterministic methods in making short-term forecasts of wind speed based on past data. For the analysis, we utilised finer time resolution wind speed data than the daily mean. These methods are applicable in situations where the underlying dynamics of wind speed is chaotic leading to random like fluctuations in the time series of wind speed. We have applied a couple of chaotic

time series prediction tools (one local method and one global method) on the records of the wind speed data of 10-minute resolution from a total of 234 different geographical locations. At each location, we obtained 1-hour, 2-hour and 3-hour predictions at intervals of 30 days for a period of 3 years. The predictions are very accurate for up to 1 hour and fairly accurate for up to 3 hours. A statistical analysis of the prediction errors from these locations reveals that the average prediction error is 1.36% of the range of wind speed for 1-hour predictions, 2.99% for 2-hour predictions and 4.15% for 3-hour predictions.

We have also compared the efficiency of the deterministic methods with predictions from the f-ARIMA model. For each of the 234 locations 6 hour ahead prediction is obtained with both methods at intervals of 30 days for a period of 3 years. It is observed that, compared to f-ARIMA, the deterministic methods give better prediction accuracy for longer periods of time and capture the dynamics of the fluctuations in the original data more faithfully. These prediction methods are simple and computationally efficient alternatives for short-term wind speed forecasts.

In chapter 6 we investigate the dynamics of the wind speed oscillation at different frequency levels. The dynamical system underlying the apparent random oscillations of wind speed could be the resultant of many coupled subsystems. These subsystems may vary from very complex to simple periodic phenomena. We have presented a detailed examination of wind speed fluctuations based on the dynamics embodied by various ranges of constituent frequencies. The analysis made use of the wind speed data from a set of locations and applied a wavelet transform technique to decompose the original time series into a set of constituent series at various scales associated with different frequency ranges. Each of these component series was then analysed using tools of non-linear time series analysis such as attractor reconstruction, Lyapunov exponents and correlation dimension along with surrogate data test to differentiate stochastic behaviour from deterministic dynamics. The results provide a broad classification of the dynamics at various scales, with the dynamics up to level five showing purely random character while beyond five the dynamics are essentially deterministic. Even in the deterministic realm, the dynamics are complex with strange attractors and positive Lyapunov exponents in the intermediate range of frequencies showing that the underlying dynamics is chaotic at these levels. The lowest range of frequencies at the highest scales of decomposition, however, are almost non-chaotic and possibly results from periodic phenomena.

Another advantage of using wavelet decomposition of wind speed series is that when used in conjunction with deterministic time series prediction methods, it can significantly improve the prediction accuracy of the latter. Where wind speed dynamics is predominantly deterministic, forecast tools such as LFO have been shown to be very effective in short to medium term wind speed predictions. However, as we have shown in this work, using wavelet decomposition prior to prediction can significantly improve the prediction accuracy as well as prediction horizon. On an average for predictions up to 12 hours, the gain in accuracy obtained by combining LFO forecast on wavelet-transformed time series goes as much as 80%. A statistical analysis of the predictions made at a total of 212 different locations asserts that this improved performance of the new model is consistent across different locations and periods of time. The LFO method in comparison with AR model, both models combined with decomposed series, is better up to 20 - 60% over the AR model for up to a day ahead prediction. However, with increasing prediction time, the predictions by the LFO method falls off to the level of AR model. The analysis also reveals variations of the accuracy of predictions on a finer scale with a change in seasons, presumably due to the diurnal nature of solar activity, which can be related to the degree of fluctuations of the component series at lower levels of decomposition.

The change in the fundamental character of the dynamics of wind speed data across the decomposition levels also opens up the possibility of using hybrid models for prediction. The component series at each level of the frequency range can be predicted using the most appropriate tools for the specific dynamical characters exhibited by that particular series and then reconstructed to yield better forecasting results. This can be explored in future works.

Bibliography

- Aggarwal, SK and Meenu Gupta (2013). "Wind power forecasting: A review of statistical models-wind power forecasting". In: *International Journal of Energy Science* 3.1.
- Alligood, Kathleen T, Tim D Sauer, and James A Yorke (1997). "Chaos: An Introduction to Dynamical Systems Springer Verlag". In: *New York*.
- Asokan, K and K Satheesh Kumar (2012). "Modelling the Wind Speed Oscillation Dynamics". In: *Eco-friendly Computing and Communication Systems*. Springer, pp. 311–318.
- Bailey, B, MC Brower, and J Zack (1999). "Short-term wind power forecasting". In: *Proceedings of the European Wind Energy Conference*, pp. 1–5.
- Balouktsis, A, D Tsanakas, and G Vachtsevanos (1986). "Stochastic simulation of hourly and daily average wind speed sequences". In: *Wind Engineering* 10.1, pp. 1–11.
- Banta, Robert M, Christoph J Senff, Raul J Alvarez, Andrew O Langford, David D Parrish, Michael K Trainer, Lisa S Darby, R Michael Hardesty, Bryan Lambeth, J Andrew Neuman, et al. (2011). "Dependence of daily peak O₃ concentrations near Houston, Texas on environmental factors: Wind speed, temperature, and boundary-layer depth". In: *Atmospheric Environment* 45.1, pp. 162–173.
- Bechrakis, DA and PD Sparis (1998). "Wind speed prediction using artificial neural networks". In: *Wind Engineering* 22.6, pp. 287–296.
- Beyer, HG, T Degner, J Hausmann, M Hoffmann, and P Rujan (1994). "Short term prediction of wind speed and power output of a wind turbine with neural net networks". In: *Proc. of EWEC*. Vol. 94.
- Bilgili, Mehmet, Besir Sahin, and Abdulkadir Yasar (2007). "Application of artificial neural networks for the wind speed prediction of target station using reference stations data". In: *Renewable Energy* 32.14, pp. 2350–2360.
- Bindner, Henrik and Per Lundsager (2002). "Integration of wind power in the power system". In: *IECON 02 [Industrial Electronics Society, IEEE 2002 28th Annual Conference of the]*. Vol. 4. IEEE, pp. 3309–3316.
- Bossanyi, EA (1985). "Short-term wind prediction using Kalman filters". In: *Wind Engineering* 9.1, pp. 1–8.
- Box, G E P, G M Jenkins, and G C Reinsel (2008). *Time Series Analysis: Forecasting and Control*.
- Box, George EP, Gwilym M Jenkins, and Gregory C Reinsel (2013). *Time series analysis: forecasting and control*. John Wiley & Sons.
- Brett, Arthur C and Stanton E Tuller (1991). "The autocorrelation of hourly wind speed observations." In: *Journal of Applied Meteorology* 30, pp. 823–833.
- Brix, A, S Jewson, and C Ziehmann (2005). "Weather Derivative Valuation". In: *Cambridge University Press*.
- Brockwell, Peter J and Richard A Davis (2009). *Time series: theory and methods*. Springer Science & Business Media.
- Brown, Barbara G, Richard W Katz, and Allan H Murphy (1984). "Time series models to simulate and forecast wind speed and wind power". In: *Journal of climate and applied meteorology* 23.8, pp. 1184–1195.
- Cadenas, Erasmo and Wilfrido Rivera (2007). "Wind speed forecasting in the South Coast of Oaxaca, Mexico". In: *Renewable energy* 32.12, pp. 2116–2128.

- Cadenas, Erasmo and Wilfrido Rivera (2010). "Wind speed forecasting in three different regions of Mexico, using a hybrid ARIMA-ANN model". In: *Renewable Energy* 35.12, pp. 2732–2738.
- Campbell, PRJ and K Adamson (2005). "A novel approach to wind forecasting in the United Kingdom and Ireland". In: *International Journal Of Simulation Systems* 6.12-13, pp. 1–10.
- Candy, Brett, Stephen J English, and Simon J Keogh (2009). "A comparison of the impact of quikscat and windsat wind vector products on met office analyses and forecasts". In: *Geoscience and Remote Sensing, IEEE Transactions on* 47.6, pp. 1632–1640.
- Cassola, Federico and Massimiliano Burlando (2012). "Wind speed and wind energy forecast through Kalman filtering of Numerical Weather Prediction model output". In: *Applied Energy* 99, pp. 154–166.
- Caudill, Maureen (1987). "Neural networks primer, part I". In: *AI expert* 2.12, pp. 46–52.
- Celik, Ali N and Mohan Kolhe (2013). "Generalized feed-forward based method for wind energy prediction". In: *Applied Energy* 101, pp. 582–588.
- Celik, Ali Naci (2004). "A statistical analysis of wind power density based on the Weibull and Rayleigh models at the southern region of Turkey". In: *Renewable energy* 29.4, pp. 593–604.
- CERC (2013). *Introduction of Ancillary Services in Indian Electricity Market, April 2013*. <http://cercind.gov.in/2013/whatsnew/SP13.pdf>.
- Clancy, J M, F Gaffney, J P Deane, J Curtis, and B P O Gallachoir (2015). "Fossil fuel and CO₂ emissions savings on a high renewable electricity system - A single year case study for Ireland". In: *Energy Policy* 83, pp. 151–164. ISSN: 03014215. DOI: 10.1016/j.enpol.2015.04.011.
- Contreras-Reyes, Javier E and Wilfredo Palma (2013). "Statistical analysis of autoregressive fractionally integrated moving average models in R". In: *Springer-Verlag Berlin Heidelberg*.
- Costa, Alexandre, Antonio Crespo, Jorge Navarro, Gil Lizcano, Henrik Madsen, and Everaldo Feitosa (2008). "A review on the young history of the wind power short-term prediction". In: *Renewable and Sustainable Energy Reviews* 12.6, pp. 1725–1744.
- Cutler, Nicholas J, Jeffrey D Kepert, Hugh R Outhred, and Iain F MacGill (2008). "Characterizing wind power forecast uncertainty with numerical weather prediction spatial fields". In: *Wind Engineering* 32.6, pp. 509–524.
- Damousis, Ioannis G, Minas C Alexiadis, John B Theocharis, and Petros S Dokopoulos (2004). "A fuzzy model for wind speed prediction and power generation in wind parks using spatial correlation". In: *Energy Conversion, IEEE Transactions on* 19.2, pp. 352–361.
- Daniel, AR and AA Chen (1991). "Stochastic simulation and forecasting of hourly average wind speed sequences in Jamaica". In: *Solar energy* 46.1, pp. 1–11.
- Daubechies, Ingrid et al. (1992). *Ten lectures on wavelets*. Vol. 61. SIAM.
- De Giorgi, Maria Grazia, Antonio Ficarella, and Marco Tarantino (2011). "Error analysis of short term wind power prediction models". In: *Applied Energy* 88.4, pp. 1298–1311.
- De Giorgi, Maria Grazia, Stefano Campilongo, Antonio Ficarella, and Paolo Maria Congedo (2014). "Comparison between wind power prediction models based on wavelet decomposition with least-squares support vector machine (LS-SVM) and artificial neural network (ANN)". In: *Energies* 7.8, pp. 5251–5272.
- De Hoon, M J L, THJJ der Hagen, H Schoonewelle, and H Van Dam (1996). "Why Yule-Walker should not be used for autoregressive modelling". In: *Annals of nuclear energy* 23.15, pp. 1219–1228.
- Domenico, Manlio De, Mohammad Ali Ghorbani, Oleg Makarynsky, Dina Makarynska, and Hakimeh Asadi (2013). "Chaos and reproduction in sea level". In: *Applied Mathematical Modelling* 37.6, pp. 3687–3697.

- Drisya, GV, DC Kiplangat, K Asokan, and K Satheesh Kumar (2014). “Deterministic prediction of surface wind speed variations”. In: *Ann. Geophys* 32, pp. 1415–1425.
- Du, Ying, Ji-ping Lu, Qing Li, and YL Deng (2008). “Short-term wind speed forecasting of wind farm based on least square-support vector machine”. In: *Power System Technology* 32.15, pp. 62–66.
- Dutton, Geoff A, Georges Kariniotakis, Jackie A Halliday, and Eric Nogaret (1999). “Load and wind power forecasting methods for the optimal management of isolated power systems with high wind penetration”. In: *Wind Engineering* 23.2, p–69.
- Eckmann, J-P, S Oliffson Kamphorst, and David Ruelle (1987). “Recurrence plots of dynamical systems”. In: *EPL (Europhysics Letters)* 4.9, p. 973.
- El-Fouly, THM, EF El-Saadany, and MMA Salama (2006). “Grey predictor for wind energy conversion systems output power prediction”. In: *IEEE Transactions on Power Systems* 3.21, pp. 1450–1452.
- Elliott, AJ (2004). “A probabilistic description of the wind over Liverpool Bay with application to oil spill simulations”. In: *Estuarine, Coastal and Shelf Science* 61.4, pp. 569–581.
- Elman, Jeffrey L (1990). “Finding structure in time”. In: *Cognitive science* 14.2, pp. 179–211.
- Ernst, Bernhard (2005). “Wind power forecast for the German and Danish networks”. In: *Wind Power in Power Systems*, pp. 365–381.
- Eyer, Jim and Garth Corey (2010). “Energy storage for the electricity grid: Benefits and market potential assessment guide”. In: *Sandia National Laboratories* 20.10, p. 5.
- Fellows, A and D Hill (1990). “Wind and load forecasting for integration of wind power into a meso-scale electrical grid”. In: *Proceedings of the European community Wind Energy conference*, pp. 10–14.
- Finzi, G, P Bonelli, and G Bacci (1984). “A Stochastic Model of Surface Wind Speed for Air Quality Control Purposes.” In: *Journal of Applied Meteorology* 23, pp. 1354–1361.
- Focken, Ulrich, Matthias Lange, and Hans-Peter Waldl (2001). “Previento-a wind power prediction system with an innovative upscaling algorithm”. In: *Proceedings of the European Wind Energy Conference, Copenhagen, Denmark*. Vol. 276. Citeseer.
- Foley, Aoife M, Paul G Leahy, Antonino Marvuglia, and Eamon J McKeogh (2012). “Current methods and advances in forecasting of wind power generation”. In: *Renewable Energy* 37.1, pp. 1–8.
- Fox, Brendan (2007). *Wind power integration: connection and system operational aspects*. Vol. 50. Iet.
- Fraser, AM and HL Swinney (1986). “Using mutual information to find independent coordinates for strange attractors”. In: *Phys. Rev. A* 33, pp. 1134–1140.
- Gavaldà, Jna, J Massons, J Camps, and F Diaz (1992). “Statistical and spectral analysis of the wind regime in the area of Catalonia”. In: *Theoretical and applied climatology* 46.2, pp. 143–152.
- George Athanasopoulos, Rob J Hyndman with contributions from, Slava Razbash, Drew Schmidt, Zhenyu Zhou, Yousaf Khan, Christoph Bergmeir, and Earo Wang (2014). *forecast: Forecasting functions for time series and linear models*. R package version 5.3. URL: <http://CRAN.R-project.org/package=forecast>.
- Georgilakis, Pavlos S (2008). “Technical challenges associated with the integration of wind power into power systems”. In: *Renewable and Sustainable Energy Reviews* 12.3, pp. 852–863.
- Giebel, Gregor, Lars Landberg, Torben Skov Nielsen, and Henrik Madsen (2001). “The ZEPHYR-project: The next generation prediction system”. In: *Proc. of the 2001 European Wind Energy Conference, EWEC’01, Copenhagen, Denmark*, pp. 777–780.

- Giebel, Gregor, Lars Landberg, Georges Kariniotakis, and Richard Brownsword (2003). “State-of-the-art Methods and software tools for short-term prediction of wind energy production”. In: *EWEC 2003 (European Wind Energy Conference and exhibition)*, CD-ROM.
- Giebel, Gregor and Georges Kariniotakis (2007). “Best practice in short-term forecasting. A Users Guide”. In: *European Wind Energy Conference and Exhibition 2007*, 5–pages.
- Giebel, Gregor, Pierre Pinson, Ulrich Focher, Matthias Lange, Ronni Meyer, and George Kariniotakis (2008). “Best practice in the use of short-term forecasting—Results from 2 workshops organized by the POW’WOW Coordination Action”. In: *Proceedings of the European Wind Energy Conference 2008*.
- Giebel, Gregor, Richard Brownsword, George Kariniotakis, Michael Denhard, and Caroline Draxl (2011). *The state-of-the-art in short-term prediction of wind power: A literature overview*. Tech. rep. ANEMOS. plus.
- Global Wind Energy Council (2013). *Global wind statistics*. URL: <http://www.gwec.net/wp-content/uploads/2014/02/GWEC-PRstats-2013{\%}5C{\%}EN.pdf>.
- Glover, J Duncan, Mulukutla Sarma, and Thomas Overbye (2011). *Power System Analysis & Design, SI Version*. Cengage Learning.
- Gomes, Pedro and Rui Castro (2012). “Wind Speed and Wind Power Forecasting using Statistical Models: AutoRegressive Moving Average (ARMA) and Artificial Neural Networks (ANN)”. In: *International Journal of Sustainable Energy* 1.1/2.
- Gonzalez, Gerardo, Belen Diaz-Guerra, Fernando Soto, Sara Lopez, Ismael Sanchez, Julio Usaola, Monica Alonso, and Miguel G Lobo (2004). “SIPREÓLICO-Wind power prediction tool for the Spanish peninsular power system”. In: *Proceedings of the CIGRÉ 40th general session and exhibition, París, France*.
- Granger, Clive William John and Paul Newbold (2014). *Forecasting economic time series*. Academic Press.
- Grassberger, Peter and Itamar Procaccia (1983). “Measuring the strangeness of strange attractors”. In: *Physica D: Nonlinear Phenomena* 9.1, pp. 189–208.
- (2004). “Measuring the strangeness of strange attractors”. In: *The Theory of Chaotic Attractors*. Springer, pp. 170–189.
- Grebogi, Celso and Ying-Cheng Lai (1997). “Controlling chaotic dynamical systems”. In: *Systems & control letters* 31.5, pp. 307–312.
- Guo, Zhen-hai, Jie Wu, Hai-yan Lu, and Jian-zhou Wang (2011). “A case study on a hybrid wind speed forecasting method using BP neural network”. In: *Knowledge-based systems* 24.7, pp. 1048–1056.
- GWEC (2012). *Global Wind Energy Outlook*. http://www.gwec.net/wp-content/uploads/2012/11/GWEO_2012_lowRes.pdf.
- (2014). *Global Wind Energy Outlook, 2014*. http://www.gwec.net/wp-content/uploads/2014/10/GWEO2014_WEB.pdf.
- Haque, Ashraf Ul, Paras Mandal, Julian Meng, and Michael Negnevitsky (2013). “Wind speed forecast model for wind farm based on a hybrid machine learning algorithm”. In: *International Journal of Sustainable Energy* ahead-of-print, pp. 1–14.
- (2015). “Wind speed forecast model for wind farm based on a hybrid machine learning algorithm”. In: *International Journal of Sustainable Energy* 34.1, pp. 38–51.
- Hegger, Rainer, Holger Kantz, and Thomas Schreiber (1999). “Practical implementation of nonlinear time series methods: The TISEAN package”. In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 9.2, pp. 413–435.
- Hennessey Jr, Joseph P (1977). “Some aspects of wind power statistics”. In: *Journal of Applied Meteorology* 16.2, pp. 119–128.

- Hering, Amanda S and Marc G Genton (2010). “Powering up with space-time wind forecasting”. In: *Journal of the American Statistical Association* 105.489, pp. 92–104.
- Hirata, Yoshito, Danilo P Mandic, Hideyuki Suzuki, and Kazuyuki Aihara (2008). “Wind direction modelling using multiple observation points”. In: *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* 366.1865, pp. 591–607.
- Hosking, Jonathan RM (1981). “Fractional differencing”. In: *Biometrika* 68.1, pp. 165–176.
- Hu, Jianming, Jianzhou Wang, and Guowei Zeng (2013). “A hybrid forecasting approach applied to wind speed time series”. In: *Renewable Energy* 60, pp. 185–194.
- Huang, Norden E, Zheng Shen, Steven R Long, Manli C Wu, Hsing H Shih, Quanan Zheng, Nai-Chyuan Yen, Chi Chao Tung, and Henry H Liu (1998). “The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis”. In: *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*. Vol. 454. 1971. The Royal Society, pp. 903–995.
- Hunt, Katherine and Guy P Nason (2001). “Wind speed modelling and short-term prediction using wavelets”. In: *Wind Engineering* 25.1, pp. 55–61.
- Hure, Nikola, Robi Turnar, Mario Vasak, and Goran Bencic (2015). “Optimal wind turbine yaw control supported with very short-term wind predictions”. In: *Industrial Technology (ICIT), 2015 IEEE International Conference on*. IEEE, pp. 385–391.
- IEA (2012). *IEA WIND 2011 ANNUAL REPORT*. http://ieawind.org/annual_reports_PDF/2011/201120IEA20Wind20AR_1_small.pdf.
- (2013). *LONG-TERM RESEARCH AND DEVELOPMENT NEEDS FOR WIND ENERGY FOR THE TIME FRAME 2012 to 2030*. https://www.ieawind.org/long-term20reports/IEA20Long20Term20R_D_Approved20July2023202013.pdf.
- (2015). *WORLD ENERGY OUTLOOK 2015 FACTSHEET Global energy trends to 2040*. http://www.worldenergyoutlook.org/media/weowebiste/2015/WEO2015_Factsheets.pdf.
- IEAIndia (2015). *India Energy Outlook, World Energy Outlook Special Report*. http://www.worldenergyoutlook.org/media/weowebiste/2015/IndiaEnergyOutlook_WEO2015.pdf.
- Jiang, Yu, Zhe Song, and Andrew Kusiak (2013). “Very short-term wind speed forecasting with Bayesian structural break model”. In: *Renewable Energy* 50, pp. 637–647.
- Jones, Richard H (1980). “Maximum likelihood fitting of ARMA models to time series with missing observations”. In: *Technometrics* 22.3, pp. 389–395.
- Jørgensen, JU and Corinna Möhrle (2005). *HONEYMOON-A high resolution numerical wind energy model for on-and offshore forecasting using ensemble predictions*.
- Kaldellis, John K and D Zafirakis (2011). “The wind energy (r) evolution: A short review of a long history”. In: *Renewable Energy* 36.7, pp. 1887–1901.
- Kamal, Lalarukh and Yasmin Zahra Jafri (1997). “Time series models to simulate and forecast hourly averaged wind speed in Quetta, Pakistan”. In: *Solar Energy* 61.1, pp. 23–32.
- Kani, SA Pourmousavi and MM Ardehali (2011). “Very short-term wind speed prediction: a new artificial neural network–Markov chain model”. In: *Energy Conversion and Management* 52.1, pp. 738–745.
- Kantz, Holger (1994). “A robust method to estimate the maximal Lyapunov exponent of a time series”. In: *Physics Letters A* 185.1, pp. 77–87.
- Kantz, Holger and Thomas Schreiber (2003). *Nonlinear time series analysis*. Vol. 2000. Cambridge university press Cambridge.

- Kantz, Holger and Thomas Schreiber (2004). *Nonlinear time series analysis*. Vol. 7. Cambridge university press.
- Kaplan, Daniel and Leon Glass (1995). “Time-series analysis”. In: *Understanding nonlinear dynamics*. Springer, pp. 278–358.
- Kariniotakis, Georges, J Halliday, R Brownsword, Ignacio Marti, Ana Maria Palomares, I Cruz, H Madsen, TS Nielsen, Henrik Aa Nielsen, Ulrich Focken, et al. (2006). “Next Generation Short-Term Forecasting of Wind Power—Overview of the ANEMOS Project.” In: *European Wind Energy Conference, EWEC 2006*, 10–pages.
- Kariniotakis, GN, GS Stavrakakis, and EF Nogaret (1996). “Wind power forecasting using advanced neural networks models”. In: *Energy conversion, ieee transactions on* 11.4, pp. 762–767.
- Kavasseri, Rajesh G and Krithika Seetharaman (2009). “Day-ahead wind speed forecasting using f-ARIMA models”. In: *Renewable Energy* 34.5, pp. 1388–1393.
- Kennel, Matthew B, Reggie Brown, and Henry DI Abarbanel (1992). “Determining embedding dimension for phase-space reconstruction using a geometrical construction”. In: *Physical review A* 45.6, p. 3403.
- Kim, Donghoh and Hee-Seok Oh (2009). “EMD: a package for empirical mode decomposition and Hilbert spectrum”. In: *The R Journal* 1.1, pp. 40–46.
- Kiplangat, Dennis C, K Asokan, and K Satheesh Kumar (2016). “Improved week-ahead predictions of wind speed using simple linear models with wavelet decomposition”. In: *Renewable Energy* 93, pp. 38–44.
- Kretschmar, Ralf, Pierre Eckert, Daniel Cattani, and Fritz Eggimann (2004). “Neural network classifiers for local wind prediction”. In: *Journal of Applied Meteorology* 43.5, pp. 727–738.
- Kumar, K Satheesh, CVA Kumar, Benny George, G Renuka, and C Venugopal (2004). “Analysis of the fluctuations of the total electron content (TEC) measured at Goose Bay using tools of nonlinear methods”. In: *Journal of Geophysical Research: Space Physics* 109.A2.
- Landberg, Lars and Simon J Watson (1994). “Short-term prediction of local wind conditions”. In: *Boundary-Layer Meteorology* 70.1-2, pp. 171–195.
- Landberg, Lars, Gregor Giebel, Henrik Aalborg Nielsen, Torben Nielsen, and Henrik Madsen (2003). “Short-term Prediction—An Overview”. In: *Wind Energy* 6.3, pp. 273–280.
- Lang, S, C Möhrlen, J Jørgensen, BO Gallachóir, and E McKeogh (2002). “Application of a Multi-Scheme Ensemble Prediction System for wind power forecasting in Ireland and comparison with validation results from Denmark and Germany”. In: *system* 2001, p. 5.
- Lang, S, C Mohrlen, J Jorgensen, BO Gallachóir, and E McKeogh (2006a). “Aggregate forecasting of wind generation on the irish grid using a multi-scheme ensemble prediction system”. In: *international solar energy society uk section-conference-c*. Vol. 85, p. 89.
- Lang, S, C Möhrlen, J Jørgensen, BO Gallachóir, and E McKeogh (2006b). “Forecasting total wind power generation on the Republic of Ireland Grid with a Multi-Scheme Ensemble Prediction System”. In: *vectors* 10, p. 11.
- Lange, Bernhard, Kurt Rohrig, Florian Schlögl, Ümit Cali, and Rene Jursa (2012). “Wind power forecasting”. In: *Renewable electricity and the grid*, pp. 95–120.
- Lange, Matthias and Ulrich Focken (2006). *Physical approach to short-term wind power prediction*. Springer.
- Lei, Ma, Luan Shiyan, Jiang Chuanwen, Liu Hongling, and Zhang Yan (2009). “A review on the forecasting of wind speed and generated power”. In: *Renewable and Sustainable Energy Reviews* 13.4, pp. 915–920.

- Lerner, Jeff, M Grundmeyer, and M Garvert (2009). “The role of wind forecasting in the successful integration and management of an intermittent energy source”. In: *Energy Central Topics Newsletter* 3.8.
- Liu, Da, Dongxiao Niu, Hui Wang, and Leilei Fan (2014). “Short-term wind speed forecasting using wavelet transform and support vector machines optimized by genetic algorithm”. In: *Renewable Energy* 62.0, pp. 592–597. ISSN: 0960-1481.
- Liu, Hui, Hong-qi Tian, and Yan-fei Li (2012). “Comparison of two new ARIMA-ANN and ARIMA-Kalman hybrid methods for wind speed prediction”. In: *Applied Energy* 98, pp. 415–424.
- Liu, Hui, Hong-qi Tian, Di-fu Pan, and Yan-fei Li (2013). “Forecasting models for wind speed using wavelet, wavelet packet, time series and Artificial Neural Networks”. In: *Applied Energy* 107, pp. 191–208.
- Liu, Hui, Hong-qi Tian, Xi-feng Liang, and Yan-fei Li (2015). “Wind speed forecasting approach using secondary decomposition algorithm and Elman neural networks”. In: *Applied Energy* 157, pp. 183–194.
- Lorenz, Edward N (1963). “Deterministic nonperiodic flow”. In: *Journal of the atmospheric sciences* 20.2, pp. 130–141.
- (1969). “Atmospheric predictability as revealed by naturally occurring analogues”. In: *Journal of the Atmospheric sciences* 26.4, pp. 636–646.
- Lowe, David and D Broomhead (1988). “Multivariable functional interpolation and adaptive networks”. In: *Complex systems* 2, pp. 321–355.
- Mabel, M Carolin and E Fernandez (2009). “Estimation of energy yield from wind farms using artificial neural networks”. In: *Energy Conversion, IEEE Transactions on* 24.2, pp. 459–464.
- Madsen, Henrik, Pierre Pinson, George Kariniotakis, Henrik Aa Nielsen, and Torben Nielsen (2005). “Standardizing the performance evaluation of shortterm wind power prediction models”. In: *Wind Engineering* 29.6, pp. 475–489.
- Martí Perez, I (2002). “Wind forecasting activities”. In: *Proceedings of the First IEA Joint Action Symposium on Wind Forecasting Techniques, Norrköping, Sweden*, pp. 11–20.
- Martin, M, LV Cremades, and JM Santabarbara (1999). “Analysis and modelling of time series of surface wind speed and direction”. In: *International journal of climatology* 19.2, pp. 197–209.
- Mathew, Thomas, Smitha George, Abhijit Sarkar, and Sujit Basu (2011). “Prediction of mean winds at an Indian coastal station using a data-driven technique applied on Weibull Parameters”. In: *The International Journal of Ocean and Climate Systems* 2.1, pp. 45–54.
- Meyers, Johan and Charles Meneveau (2012). “Optimal turbine spacing in fully developed wind farm boundary layers”. In: *Wind Energy* 15.2, pp. 305–317.
- Miranda, Marcos S and Rod W Dunn (2006). “One-hour-ahead wind speed prediction using a Bayesian methodology”. In: *Power Engineering Society General Meeting, 2006. IEEE. IEEE*, 6–pp.
- Mitschke, F and M Dämmig (1993). “Chaos versus noise in experimental data”. In: *International Journal of Bifurcation and Chaos* 3.03, pp. 693–702.
- Mohandes, MA, TO Halawani, S Rehman, and Ahmed A Hussain (2004). “Support vector machines for wind speed prediction”. In: *Renewable Energy* 29.6, pp. 939–947.
- Mohandes, Mohamed A, Shafiqur Rehman, and Talal O Halawani (1998). “A neural networks approach for wind speed prediction”. In: *Renewable Energy* 13.3, pp. 345–354.
- Möhrlen, Corinna and Jess U Jørgensen (2005). “Verification of ensemble prediction systems for a new market: wind energy”. In: *ECMWF Special Project Interim Report* 3.1.

- Monfared, Mohammad, Hasan Rastegar, and Hossein Madadi Kojabadi (2009). "A new strategy for wind speed forecasting using artificial intelligent methods". In: *Renewable energy* 34.3, pp. 845–848.
- Monteiro, C, R Bessa, V Miranda, A Botterud, J Wang, G Conzelmann, et al. (2009). *Wind power forecasting: state-of-the-art 2009*. Tech. rep. Argonne National Laboratory (ANL).
- More, Anurag and MC Deo (2003). "Forecasting wind with neural networks". In: *Marine structures* 16.1, pp. 35–49.
- Nicholson, Martin, Tom Biegler, and Barry W Brook (2011). "How carbon pricing changes the relative competitiveness of low-carbon baseload generating technologies". In: *Energy* 36.1, pp. 305–313. ISSN: 03605442. DOI: 10.1016/j.energy.2010.10.039.
- Nielsen, Claus S, Hans F Ravn, and Camilla Schaumburg-Müller (2003). "Two wind power prognosis criteria and regulating power costs". In: *4th Int'l Workshop on Large-Scale Integration of Wind Power and Transmission Networks for Offshore Wind Farms*.
- Nielsen, Torben Skov, Alfred Joensen, Henrik Madsen, Lars Landberg, and Gregor Giebel (1998). "A new reference for wind power forecasting". In: *Wind energy* 1.1, pp. 29–34.
- Nielsen, Torben Skov, Henrik Madsen, and John Tofting (1999). "Experiences with statistical methods for wind power prediction". In: *Proceedings From Ewec'99*.
- Nogaret, E, G Stavrakakis, JC Bonin, G Kariniotakis, B Papadidas, G Contaxis, M Papadopoulos, N Hatziargyriou, S Papatthanassiou, J Garopoulos, et al. (1994). "Development and implementation of an advanced control system for medium size wind-diesel systems". In: *Proceedings of the EWEC*. Vol. 94, pp. 599–604.
- Ott, Edward (2002). *Chaos in dynamical systems*. Cambridge university press.
- Ott, Edward, Tim Sauer, and James A Yorke (1994). *Coping with Chaos: Analysis of Chaotic Data and the Exploitation of Chaotic Systems*, 418 pp.
- Packard, Norman H, James P Crutchfield, J Doyne Farmer, and Robert S Shaw (1980). "Geometry from a time series". In: *Physical Review Letters* 45.9, pp. 712–716.
- Palmer, AJ, RA Kropfli, and CW Fairall (1995). "Signatures of deterministic chaos in radar sea clutter and ocean surface winds." In: *Chaos (Woodbury, NY)* 5.3, p. 613.
- Palomares-Salas, JC, JJG De la Rosa, JG Ramiro, J Melgar, A Agüera, and A Moreno (2009). "Comparison of models for wind speed forecasting". In: *Proc. International Conference on Computational Science*.
- Pavlos, GP, GA Kyriakou, AG Rigas, PI Liatsis, PC Trochoutsos, and AA Tsonis (1992). "Evidence for strange attractor structures in space plasmas". In: *Annales geophysicae*. Vol. 10. 5. Copernicus, pp. 309–322.
- Pavlos, GP, MA Athanasiu, D Kugiumtzis, N Hatzigeorgiu, AG Rigas, ET Sarris, et al. (1999). "Nonlinear analysis of magnetospheric data Part I. Geometric characteristics of the AE index time series and comparison with nonlinear surrogate data". In: *Nonlinear Processes in Geophysics* 6.1, pp. 51–65.
- Percival, Donald B, Muyin Wang, and James E Overland (2004). "An introduction to wavelet analysis with applications to vegetation time series". In: *Community Ecology* 5.1, pp. 19–30.
- Percival, Donald B and Andrew T Walden (2006). *Wavelet methods for time series analysis*. Vol. 4. Cambridge university press.
- Pinson, Pierre and Henrik Madsen (2008). "Probabilistic forecasting of wind power at the minute time-scale with markov-switching autoregressive models". In: *Probabilistic Methods Applied to Power Systems, 2008. PMAPS'08. Proceedings of the 10th International Conference on*. IEEE, pp. 1–8.

- Piwko, Richard, Dale Osborn, Robert Gramlich, Gary Jordan, David Hawkins, and Kevin Porter (2005). "Wind energy delivery issues [transmission planning and competitive electricity market operation]". In: *Power and Energy Magazine, IEEE* 3.6, pp. 47–56.
- Potter, Cameron W and Michael Negnevitsky (2006). "Very short-term wind forecasting for Tasmanian power generation". In: *Power Systems, IEEE Transactions on* 21.2, pp. 965–972.
- Provenzale, A, Leonard A Smith, R Vio, and G Murante (1992). "Distinguishing between low-dimensional dynamics and randomness in measured time series". In: *Physica D: nonlinear phenomena* 58.1, pp. 31–49.
- R Core Team (2014). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. URL: <http://www.R-project.org/>.
- Ragwitz, M and H Kantz (2007). "Detecting non-linear structure and predicting turbulent gusts in surface wind velocities". In: *EPL (Europhysics Letters)* 51.6, p. 595.
- Ruelle, David and Floris Takens (1971). "On the nature of turbulence". In: *Communications in mathematical physics* 20.3, pp. 167–192.
- Sánchez, I, J Usaola, O Ravelo, C Velasco, J Domínguez, MG Lobo, G González, F Soto, B Diaz-Guerra, and M Alonso (2002). "Sipreólico—a wind power prediction system based on flexible combination of dynamic models. application to the spanish power system". In: *Poster on the World Wind Energy Conference in Berlin, Germany*.
- Sauer, Tim, James A Yorke, and Martin Casdagli (1991). "Embedology". In: *Journal of Statistical Physics* 65.3, pp. 579–616.
- Sauer, Tim and A Yorke James (1993). "How many delay coordinates do you need?" In: *International Journal of Bifurcation and Chaos* 3.03, pp. 737–744.
- Schreiber, Thomas (1993). "Extremely simple nonlinear noise-reduction method". In: *Physical Review E* 47.4, p. 2401.
- Schreiber, Thomas and Andreas Schmitz (1996). "Improved surrogate data for nonlinearity tests". In: *Physical Review Letters* 77.4, pp. 635–638.
- Schwartz, M and M Milligan (2002). "Statistical Wind Forecasting at the US National Renewable Energy Laboratory". In: *Proceedings of the First IEA Joint Action Symposium on Wind Forecasting Techniques, Norrköping, Sweden*, 115–124B.
- Sfetsos, A (2000). "A comparison of various forecasting techniques applied to mean hourly wind speed time series". In: *Renewable energy* 21.1, pp. 23–35.
- (2002). "A novel approach for the forecasting of mean hourly wind speed time series". In: *Renewable Energy* 27.2, pp. 163–174.
- Shu, ZR, QS Li, and PW Chan (2015). "Investigation of offshore wind energy potential in Hong Kong based on Weibull distribution function". In: *Applied Energy* 156, pp. 362–373.
- Shukur, Osamah Basheer and Muhammad Hisyam Lee (2015). "Daily wind speed forecasting through hybrid KF-ANN model based on ARIMA". In: *Renewable Energy* 76, pp. 637–647.
- Singh, Mohit and Surya Santoso (2011). *Dynamic models for wind turbines and wind power plants*. National Renewable Energy Laboratory.
- Snodin, Helen (2006). *Short-term wind energy forecasting: Technology and Policy, 2006*. http://s3.amazonaws.com/zanran_storage/www.canwea.ca/ContentPages/43316176.pdf.
- Soman, Saurabh S, Hamidreza Zareipour, Om Malik, and Paras Mandal (2010). "A review of wind power and wind speed forecasting methods with different time horizons". In: *North American Power Symposium (NAPS), 2010*. IEEE, pp. 1–8.
- Sreelekshmi, RC, K Asokan, and K Satheesh Kumar (2012). "Deterministic nature of the underlying dynamics of surface wind fluctuations". In: *Annales Geophysicae-Atmospheres Hydro-spheresand Space Sciences* 30.10, p. 1503.

- Steinwart, Ingo and Andreas Christmann (2008). *Support vector machines*. Springer Science & Business Media.
- Sugihara, George and Robert M May (1990). *Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series*.
- Takens, Floris (1981). “Detecting strange attractors in turbulence”. In: *Dynamical systems and turbulence, Warwick 1980*. Springer, pp. 366–381.
- Tamazin, M, A Noureldin, and M J Korenberg (2013). “Robust Modeling of Low-Cost MEMS Sensor Errors in Mobile Devices Using Fast Orthogonal Search”. In: *Journal of Sensors 2013*.
- Tantareanu, Cristian (1992). *Wind Prediction in Short Term: A first step for a better wind turbine control*. Nordvestjysk Folkecenter for Vedvarende Energi.
- Tascikaraoglu, Akin, Borhan M Sanandaji, Kameshwar Poolla, and Pravin Varaiya (2016). “Exploiting sparsity of interconnections in spatio-temporal wind speed forecasting using Wavelet Transform”. In: *Applied Energy* 165, pp. 735–747.
- Theiler, James (1986). “Spurious dimension from correlation algorithms applied to limited time-series data”. In: *Physical review A* 34.3, p. 2427.
- Theiler, James, Stephen Eubank, André Longtin, Bryan Galdrikian, and J Doynne Farmer (1992). “Testing for nonlinearity in time series: the method of surrogate data”. In: *Physica D: Nonlinear Phenomena* 58.1, pp. 77–94.
- Torres, Jose Luis, Almudena Garcia, Marian De Blas, and Adolfo De Francisco (2005). “Forecast of hourly average wind speed with ARMA models in Navarre (Spain)”. In: *Solar Energy* 79.1, pp. 65–77.
- Troen, I and L Landberg (1990). “Short term prediction of local wind conditions”. In: *European Community Wind Energy Conference. Proceedings*. HS Stephens and Associates.
- Troen, IB and Erik L Petersen (1989). “European wind atlas. Published for the Commission of the European Communities, Directorate General for Science”. In: *Research and Development (Brussels, Belgium) by Risø National Laboratory, Roskilde, Denmark*.
- Vermeer, LJ, Jens Nørkær Sørensen, and A Crespo (2003). “Wind turbine wake aerodynamics”. In: *Progress in aerospace sciences* 39.6, pp. 467–510.
- Vincent, Claire Louise, Caroline Draxl, Gregor Giebel, Pierre Pinson, Jess Jørgensen, and Corinna Möhrlen (2009). “Spectral verification of a mesoscale ensemble”. In: *2009 European Wind Energy Conference and Exhibition*.
- Wan, Yih-huei, Michael R Milligan, and Brendan Kirby (2007). *Impact of energy imbalance tariff on wind energy*. National Renewable Energy Laboratory.
- Wang, Xiaolan and LI Hui (2012). “Multiscale prediction of wind speed and output power for the wind farm”. In: *Journal of Control Theory and Applications* 10.2, pp. 251–258.
- Weisstein, Eric W. *Attractor*, From MathWorld—A Wolfram Web Resource. URL: <http://mathworld.wolfram.com/Attractor.html> (visited on 12/01/2016).
- Whitcher, Brandon and Maintainer Brandon Whitcher (2007). “The waveslim Package”. In: *waveslim.pdf*, p. 1.
- Wu, Yuan-Kang and Jing-Shan Hong (2007). “A literature review of wind forecasting technology in the world”. In: *Power Tech, 2007 IEEE Lausanne*. IEEE, pp. 504–509.
- WWEA (2015). *The World Wind Energy Association, 2014 Half-/Year Report*. http://www.wwindea.org/webimages/WWEA_half_year_report_2014.pdf.
- Zack, JC (2004). “A new technique for short-term wind energy forecasting: a rapid update cycle with a physics-based atmospheric model”. In: *CD-Proc. of the 2004 Global Windpower Conference, Chicago, Illinois (USA)*.

- Zeng, Jianwu and Wei Qiao (2011). "Support vector machine-based short-term wind power forecasting". In: *Power Systems Conference and Exposition (PSCE), 2011 IEEE/PES*. IEEE, pp. 1–8.
- Zhang, Qian, Kin Keung Lai, Dongxiao Niu, Qiang Wang, and Xuebin Zhang (2012). "A fuzzy group forecasting model based on least squares support vector machine (LS-SVM) for short-term wind power". In: *Energies* 5.9, pp. 3329–3346.
- Zhou, Junyi, Jing Shi, and Gong Li (2011). "Fine tuning support vector machines for short-term wind speed forecasting". In: *Energy Conversion and Management* 52.4, pp. 1990–1998.
- Zhu, Bo, Min-you Chen, Neal Wade, and Li Ran (2012). "A prediction model for wind farm power generation based on fuzzy modeling". In: *Procedia Environmental Sciences* 12, pp. 122–129.

UJGC - WERP



The underlying dynamics of the apparent random like fluctuations of wind speed measurements are deterministic, low dimensional and chaotic. Deterministic models show significant improvement in short-term prediction accuracy of wind speed oscillations. Week-ahead predictions with reliable accuracy. Wind speed fluctuations exhibit diverse dynamical characteristics across its frequency spectrum.

