Trend and Pattern of Wind in India and Australia

Prashanth Gururaja

A Thesis Submitted in Fulfillment of the Requirements for the Degree of

Doctor of Philosophy in Research Methodology

Prince of Songkla University

2020

Trend and Pattern of Wind in India and Australia

Prashanth Gururaja

A Thesis Submitted in Fulfillment of the Requirements for the Degree of

Doctor of Philosophy in Research Methodology

Prince of Songkla University

2020

Thesis Title Trend and Pattern of Wind in India and Australia

Author Mr. Prashanth Gururaja

Major Program Research Methodology

Major Advisor	**Examining Committee:**
…………………………………	…………………………..…..Chairperson
(Emeritus Prof. Dr. Don McNeil)	(Asst. Prof. Dr. Phattrawan Tongkumchum)
	…………………………..…………..
Co-Advisor	(Emeritus Prof. Dr. Don McNeil)
…………………………………	…………………………..…………..
(Asst. Prof. Dr. Rhysa McNeil)	(Asst. Prof. Dr. Rhysa McNeil)
	……………………………….……..
	(Asst. Prof. Dr. Sangdao Wongsai)

The Graduate School, Prince of Songkla University, has approved this thesis as fulfilment of the requirements for the Doctor of Philosophy Degree in Research Methodology.

………………………………………
(Prof. Dr. Damrongsak Faroongsarng)
Dean of Graduate School

This is to certify that the work here submitted is the result of the candidate's investigations. Due acknowledgement has been made of any assistance received.

……………………………..…Signature

(Emeritus Prof. Dr. Don McNeil)

Major Advisor

……………………………..…Signature

(Mr. Prashanth Gururaja)

Candidate

I hereby certify that this work has not been accepted in substance for any other degree, and is not being currently submitted for candidature for any degree.

………………..…………..…Signature

(Mr. Prashanth Gururaja)

Candidate

ชื่อวิทยานิพนธ์ แนวโน้มและรูปแบบของลมในอินเดียและออสเตรเลีย

ผู้เขียน Mr. Prashanth Gururaja

สาขาวิชา วิธีวิทยาการวิจัย

ปีการศึกษา 2562

บทคัดย่อ

การศึกษานี้มีวัตถุประสงค์เพื่อวิเคราะห์แนวโน้มและรูปแบบของลมด้วยวิธีการทางสถิติ ประกอบด้วย 2 การศึกษา ได้แก่ การศึกษาข้อมูลลมทุก 4 ชั่วโมงของประเทศอินเดีย และข้อมูลลมทุก 3 ชั่วโมง ของประเทศออสเตรเลีย การศึกษาแรกใช้ข้อมูลลมสถานี Calcutta ประเทศอินเดีย ปี พ.ศ. 2547- 2551 จาก National Renewable Energy Laboratory การศึกษาที่สองใช้ข้อมูลลม ปี พ.ศ. 2547- 2551 จากสถานีอากาศ 4 แห่ง ได้แก่ Avalon airport, Essendon airport, Point Wilson และ View bank เมลเบิร์นของสำนักอุตุนิยมวิทยาออสเตรเลีย วิเคราะห์ความชุกของลม รูปแบบของลมกระโชก และความเร็วลมที่รุนแรงด้วยตัวแบบการถดถอยโลจิสติกส์ และวิเคราะห์แนวโน้มของความเร็วลม ด้วยตัวแบบการถดถอยพหุคูณ

ผลการศึกษาแรกพบว่ารูปแบบความชุกของลมจะสูงขึ้นในช่วงเวลา 04.00 น. ถึง 16.00 น. ความชุกสูงสุดประมาณ 90% ลมรายเดือนเพิ่มขึ้นอย่างชัดเจน ช่วงเดือนเมษายนถึงเดือนมิถุนายน (90%) และรูปแบบความชุกรายปีสูงกว่าค่าเฉลี่ยโดยรวมระหว่างปี พ.ศ. 2550-2551 (85%)

ความชุกของลมกระโชกมีอัตราการเพิ่มขึ้นเวลา 12.00 น. ถึง 16.00 น. ซึ่งมีความชุกสูงสุดประมาณ 18% สำหรับความชุกลมกระโชกของลมตะวันตกและลมตะวันออกสูงสุดช่วงเวลา 4.00 น. ถึง 16.00 น. ความชุกสูงสุดประมาณ 9% และ 8% ตามลำดับ ในส่วนของลมกระโชครายเดือนมีความชุกสูงขึ้นระหว่างเดือนมกราคมถึงเมษายน (9%) และมีอัตราสูงสุด ปี พ.ศ. 2550 ความเร็วลมเพิ่มขึ้นเวลา 04.00 น. ถึง 12.00 น. และสูงในเดือนมกราคมถึงเมษายน ด้วยความเร็วลมสูงสุด (13 mps) และความเร็วลมประจำปีเพิ่มสูงขึ้นระหว่างปี พ.ศ. 2549-2551 (11 mps)

การศึกษาที่สอง สถานี Avalon airport, Essendon airport และ Viewbank มีรูปแบบความชุกของลมที่คล้ายกันโดยพบความชุกสูงสุดระหว่างเวลา 12.00 น. ถึง 15.00 น. ความชุกของลมแบบรายเดือนสูงกว่าค่าเฉลี่ยในเดือนพฤศจิกายนถึงเดือนมกราคม ส่วนความชุกของลมแบบรายปีไม่แตกต่างจากค่าเฉลี่ยโดยรวมอย่างมีนัยสำคัญ อย่างไรก็ตามสถานี Point Wilson มีความชุกของลมสูงที่สุดเมื่อเทียบdกับสามสถานี (มากกว่า 80%) ความชุกของลมสูงสุดอยู่ช่วงเวลา 15.00 น. ถึง 18.00 น. ส่วนรูปแบบของลมกระโชกไม่แตกต่างจาก ความชุกของลมในทุกสถานี

ความเร็วลมจากทุกสถานีมีความแตกต่างกัน 4 รูปแบบ สถานี Avalon airport ความเร็วลมเพิ่มขึ้นอย่างมากเวลา 6.00 น. ถึง 15.00 น. ด้วยความเร็วลมสูงสุด (มากกว่า 50 Kph) โดยเดือนกันยายนถึงเดือนมกราคมมีความเร็วลม สูงกว่าค่าเฉลี่ยโดยรวม และรูปแบบรายปีสูงกว่าค่าเฉลี่ยโดยรวมระหว่างปี พ.ศ. 2550 ถึง พ.ศ. 2551 (มากกว่า 30%) สถานี Essendon airport ความเร็วลมเพิ่มขึ้น อย่างมากช่วงเวลา 6.00 น. ถึง 12.00 น. ด้วยความเร็วลมสูงสุด (มากกว่า 50 Kph) ความเร็วลมสูงกว่าค่าเฉลี่ยโดยรวมตั้งแต่เดือนกันยายนถึงเดือนมกราคม (มากกว่า 20%) สถานี Point Wilson ความเร็วลมเพิ่มขึ้นอย่างมากเวลา 6.00 น. ถึง 15.00 น. ด้วยความเร็วลมสูงสุด (มากกว่า 40 Kph) ความเร็วลมสูงกว่าค่าเฉลี่ยโดยรวมตั้งแต่เดือนกรกฎาคมถึงเดือนมกราคม (มากกว่า 40%) ปี พ.ศ. 2550 มีความเร็วลมสูงกว่าค่าเฉลี่ยนโดยรวม (มากกว่า 30 Kph) สถานี Viewbank ความเร็วลมเพิ่มอย่างรวดเร็วในเวลา 6.00 น. ถึง 12.00 น. ด้วยความเร็วลมสูงสุด (มากกว่า 40 Kph) มีปริมาณลมมากที่สุดเดือนกันยายน (มากกว่า 30 Kph) และความเร็วสูงสุดปี พ.ศ. 2550 (มากกว่า 30 Kph)

Thesis Title Trend and Pattern of Wind in India and Australia
Author Mr. Prashanth Gururaja
Major Program Research Methodology
Academic Year 2019

ABSTRACT

In this thesis, graphical and statistical methods were used to determine the trends and patterns of wind characteristics. The thesis comprises two studies carried out for different study regions, the 4 hourly wind data from India, and the 3 hourly wind data from Australia.

The first study presents results from wind data for Calcutta station, India. The wind data from 2004-2008 were obtained from the National Renewable Energy Laboratory. In the second study, wind data from 2004-2008 were obtained from 4 weather stations, namely Avalon airport, Essendon airport, Point Wilson, and View bank in Melbourne of the Australia Bureau of Meteorology. For both studies, a logistic regression model was used to investigate the wind prevalence and the patterns of a wind gust, extreme wind velocity. A multiple linear regression model was also fitted to explore the trends in wind speed. A multinomial regression method was then fitted to investigate the wind velocity categories.

For the first study, the wind prevalence patterns revealed higher rates from 4 AM to 4 PM with the maximum prevalence of around 90 %ฯ The monthly wind showed a sharp increase from April to June (90%). The yearly patterns were above the overall mean between 2007 and 2008 (85%).

The wind gust prevalence patterns showed increasing rates during the day between 12 PM and 4 PM with the maximum wind gust prevalence of approximately 18%. The gusty westerly and easterly winds prevalence was mostly higher between 4 AM and 4 PM during the day with the maximum prevalence of approximately 9% and 8%, respectively. Monthly gusty westerly and easterly winds showed a higher prevalence between January and April (9%). Annually, the prevalence showed an increasing rate (8%) in the year 2007. The wind speed patterns revealed a sharp increase from 4 AM to 12 PM with the maximum wind speed (13 mps). The monthly wind speed patterns showed an increased between January and April (13 mps). An increasing rate was observed for the yearly wind speed between 2006 and 2008 (11 mps).

For the second study, the stations of Avalon airport, Essendon airport, and Viewbank revealed quite similar wind prevalence patterns, with the highest prevalence between 12 PM to 3 PM. The monthly patterns were higher than the average from November to January. The yearly patterns were not significantly different from the overall mean. However, the Point Wilson station, the overall percentage of wind prevalence was highest compared with other three stations (over 80%). The prevalence of wind peaked between 3 PM to 6 PM. The wind gust patterns were quite similar to that of wind prevalence for all stations.

There were four distinct wind speed trends. Avalon airport showed a sharp increase from 6 AM to 3 PM with the maximum wind speed (over 50 Kph). The monthly patterns were observed to be higher than the overall mean from September to January. The yearly patterns were above the overall mean between 2007 and 2008 (over 30%). Essendon airport showed a sharp increase from 6 AM to 12 PM with the maximum wind speed (over 50 Kph). The monthly patterns were observed to be higher than the overall mean from September to January (over 20%). The yearly patterns were not different from the overall mean. Point Wilson revealed a sharp increase from 6 AM to 3 PM with the maximum wind speed (over 40 Kph). The monthly patterns were observed to be higher than the overall mean from July to January (over 40%). The yearly patterns were above the overall mean in 2007 (over 30 Kph). Viewbank depicted a sharp increase from 6 AM to 12 PM with the maximum wind speed (over 40 Kph). The highest monthly wind was observed in September (over 30 Kph). The highest yearly speed was observed in 2007(over 30 Kph).

ACKNOWLEDGEMENTS

I would like to thank all of the people who made this dissertation possible. Thank God, wonderful parents, family, and respected friends for their support, patience, and encouragement. I would like to thank my advisor, Emeritus Professor Dr Don McNeil, my co-advisor Asst. Prof. Dr Rhysa McNeil, Asst. Prof. Dr Phattrawan Tongkumchun, and Asst. Prof. Dr Chamnein Choonpradub with humble appreciation for their advice, encouragement, support, and mentoring my skills and support throughout the study. I would like to thank Graduate School, Prince of Songkla University Pattani Campus, for providing the Thailand education hub scholarship (THE-AC) and theses grant. Finally, I would also like to thank the International Relations Department, Faculty of Science and Technology, and International office, Pattani campus providing me essential assistance and guidance for every visa formalities to Thailand.

Prashanth Gururaja

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Background and Rationale

The wind is one of the noteworthy factors in climate science. It plays a pivotal role in climate and weather processes on the Earth's surface and enables climate change, renewable energy, and natural devastation. Wind's profound noteworthy aspects parameters are its magnitude, irregularity diverse patterns, and varying time intervals. The internment movement has enormous damage to ecosystems, fatalities in air transport accidents, and devastation of human-made structures (Jain et al. 2001, Schindler et al., 2012, Jamaludin et al., 2016). Besides, its sporadic behavior can productively harness the energy, improve gain familiarity on global warming, and provide critical figures for weather forecast analysis (Mclnnes et al., 2011). Thus, scientific research on wind characteristics yields beneficial narrative analysis that aid in the improvement of conventional environment research.

The Indian subcontinent comprises of different climatic and geographies. It lies on the Indian Plate, and the northern portion of the Indo-Australian Plate forms the Indian subcontinent. Most of the region is covered around the tropic of cancer; which is the boundary between the tropics and subtropics; passes through the middle of India, the large part of the country is covered by rainforests, monsoon and arid regions regarded as climatically tropical area.

The main aspects of Australia's climate comprise hot sinking air of the subtropical pressure. It is considerably influenced by ocean currents, the ocean Dipole, and El Niño–Southern Oscillation, which results in periodic drought, and also the seasonal tropical low-pressure system that produces cyclones in northern Australia (Kuleshov et al., 2010).

The wind is a natural environment process caused by the large-scale flow of gases. Its intriguing behavior, namely its strength intensity and the direction it is blown, is due to the movement from high pressure to low pressure is mostly due to temperature changes. The divergent patterns of wind play a pivotal role in other climate events such as solar radiation, rainfall, temperature, and sea level anomalies.

Many scientists and researchers have studied wind patterns on the earth's surface using various methodologies. Prediction of the wind speed has been analyzed over a broad geographic region. The diverse distribution of wind speed benefits to effectively harness energy for all the geographic regions (Morgan et al., 2011). Analysis of long-term hourly wind speed trends provides essential spatial and temporal patterns. Also, seasonality, interannual variability, and the linear tendency are few significant aspects of wind speed. It also indicates the El Niño and La Niña events, which change the behavior of global wind circulation and influence wind speeds over the region (Silva and Santos, 2013). Trend analysis showed more significant values for the study groups for all seasons of the year and in the annual average. The additional climate variables such as wind variance and distribution, wind velocity sector zones, the volatility of wind speed and wind direction, renewable energy, wind engineering, marine engineering, aircraft navigation, and engineering which will help in providing wind information to help in making effective policies on renewable energy, enable effective policies to limit hazardous damages should be taken into account for descriptive analysis on the wind.

Investigation of the extreme value of wind speed is pivotal over the various spatial regions. The large wind speed extremes observed on the Earth’s surface are governed by the mesoscale atmospheric phenomena of strong synoptic storms, and gravity waves. However, the concern associated with these studies is the overestimation of extreme wind speed within the inland territory (Kislov and Matveeva, 2016).

Statistical analysis of the temporal and spatial wind patterns and modeling of wind characteristics are pivotal to harness the energy, environmental and ecological modeling, soil erosion and loss of fertile land. The effects of climate change on wind variability are of scientific concern because they agree with contemporary research and forecasting problems in meteorology, climatology, and geography systems.

1.2 Research Objectives

The objectives of this research are to access the spatial and temporal wind patterns and to investigate the trends of wind speed in India and Australia from 2004 to 2008 using an appropriate statistical method.

1.3 Literature Review

Many scholars (Usbeck *et al*., 2010; Friederichs *et al*., 2009; Ali *et al*., 2017) have studied wind characteristics. Most of these studies described the spatial variability based on classification into distinct geographical regions by using the linear regression model, logistic regression model, generalized linear model, extreme value theory, and quantile regression model. These models revealed clear wind patterns in their respective study areas. However, the latter gave a descriptive analysis of the wind pattern of the region.

Wind speed trends have been studied in different geographical regions. He *et al*., (2010) studied the probability distributions of land surface wind speeds over North America using the Weibull probability distribution function method. The Weibull probability distribution function method fitted the data significantly well. Similarly, Letson *et al*. (2019) investigated the characterizing wind gusts in complex terrain Perdigão, Portugal. They used Weibull, log-logistic, log-normal, and Gamma method. The model illustrated good agreement with estimated data. The problem associated with both methods is a large number of parameters that need to be estimated. Cross *et al*., (2015) investigated the impacts of wind speed trends and 30-Year variability in the Pacific Northwest using the ordinary least squares method. The model fitted the data fairly well. The problem associated with this method is the lack of data which would lead to better prediction.

Similarly, Dadaser-Celik and Cengiz (2014) analyzed the wind speed trends over Turkey from 1975 to 2006. They used graphical methods, linear regression, and the Mann – Kendall method. The results revealed that the majority of the stations (73%) in Turkey had average wind speeds between 1.6 and 3.3 m s^{-1} during the 32 years. Lollchund et al. (2014) investigated modeled wind speed for Mauritius by fitting Weibull, Rayleigh, Lognormal, Gamma, Normal and Frechet models. The data fit the models significantly well. Fujibe (2009) investigated the relationship between long-term temperature and wind speed trends in Japan by fitting the Least-squares method. The data fitted the models reasonably well.

Some authors have modeled maximum wind speed using various statistical models. They include a generalized linear model (Yan *et al*., 2002), Weibull parameters method (Hossain, 2019, Ayodele *et al*., 2012), and linear regression model (Abhishek *et al*., 2010; Troccoli *et al*., 2012). Satari *et al*., (2015) investigated the characteristic of wind direction and maximum wind speed over Malaysia for 1999-2008 using the Von mises distribution, multi-sample Watson-William test, and linear – circular correlation coefficient. The results suggested the presence of a prominent direction of the wind that blows in Peninsular Malaysia by the monsoon. This finding may provide useful information on giving a better understanding of the behavior of the wind in Peninsular Malaysia and the potential use of wind as an alternative source of energy.

Most authors have used a logistic regression model to forecast wind speed (Srivastava, 2005; Pang *et al.,* 2019; Sreenivasa *et al.*, 2014). The results from these studies revealed that the logistic regression model was capable of effectively predicting wind speed.

Wind observations are known for their varied patterns. One form of analysis of such data, which gained much attention to the modeling of wind patterns is the linear model. This model is widely used by many scholars (Rozas-Larraondo *et al.,* 2014; Dadaser-Celik and Cengiz, 2014; Wright and Grab, 2017; Hande *et al.*, 2012) to fit spatial wind patterns reasonably well. The linear models are capable of reproducing wind circulation features when varied climate zones are model independently, and for filling an inconsistent data with homogenous data and observations for more extended periods. Some studies by Lydia *et al.,* (2016) and Barhmi *et al.,* (2019) used linear regression for forecasting wind speed in their respective areas and concluded that the models did quite well. Arzu *et al.* (2018) used linear regression to model wind speed Kiribati. The model performed reasonably well. A higher degree of accuracy was noted in forecasting wind speed for the entire study period.

Alternatively, Zhang *et al.* (2014) investigated the annual wind speed probability distribution in the East China Sea using the maximum entropy principle (multinomial model) and the Weibull distribution method. They found that the Maximum entropy distribution performed adequately and accurately in fitting the wind speed frequency distribution with height.

Currently, one of the widely used techniques for modeling wind is Markov chain models, where the wind patterns at a specified site is a function of predictors representing previous periods of wind. Some studies (Petre *et al.*, 2016; Tagliaferri *et al.,* 2016; Elwan and Habibuddin, 2019) used Markov chain models to predict wind speed patterns. Results from these studies revealed that the Markov chain model was capable of estimating wind speed in their respective areas.

Various studies used statistical models to investigate wind patterns over the Indian continent in their respective regions (Jaswal and Koppar, 2013; Kulkarni *et al.*, 2008; Singh *et al.*, 2016; Prema and Rao, 2015) such as logistic regression model. The model was capable of estimating the wind speed patterns reasonably well.

Serval studies in Australia used statistical models to analyze wind patterns. McVicar *et al.*, (2008) described wind speed climatology and trends by fitting the linear regression method. The linear regression model fitted the wind speed data significantly well. Young *et al*., (2012) studied the investigation of trends in extreme value wave height and wind speed in Australia and around the globe. The Gumbel probability distribution function method fitted the data significantly well. Maklad and Glencross-Grant (2014) studied Seasonal Analysis of Potential Wind Power for Armidale NSW, Australia, using the Weibull distribution method. The Weibull method fitted the data reasonably well.

Even though a wide range of studies has focused on wind analysis using yearly data. This study focus on an in-depth analysis of the hourly patterns of wind, wind speed, wind gust, and wind velocity. The results of the analysis of spatial and temporal wind patterns will provide illustrative and enhanced information about wind distribution in the study areas.

1.4 Organization of theses

In this thesis, the five chapters are organized as follows**:** Chapter1 provides an introduction to the background and rationale of wind characteristics of the Earth’s surface, the objectives of the study, and the literature review. Chapter 2 describes all the study areas, data sources and management, variables and path diagram, and methodologies used for this study. Chapter 3 presents the preliminary analysis of wind data, analysis of wind characteristics in the Indian sub-continent. Chapter 4 presents the preliminary analysis of wind data, analysis of wind characteristics in Australia. Chapter 5 concludes and discusses the research findings together with limitations and recommendations for further studies.

CHAPTER 2

METHODOLOGY

This chapter consists of a description of the study area, data sources, data management, variables and path diagram, and various statistical methods for the study. The methods are comprised of the logistic regression, linear regression, and multinomial regression. All data analyses and graphical displays were carried out using the R program (R Development Core Team, 2009).

2.1 Study area

India

India is situated north of the equator between 8°4' and 37°6' north latitude and 68°7' and 97°25' east longitude. The south of India is covered by the Indian Ocean, the Arabian Sea on the southwest and Bay of Bengal on the southeast. India hosts six major climatic subtypes, ranging from arid deserts in the west, alpine tundra and glaciers in the north, and humid tropical regions supporting rain forests in the southwest and the island territories. India forms a well-defined subcontinent, set off from the rest of Asia by the imposing northern mountain rampart of the Himalayas mountain ranges to the west and east. These variant topographic landscapes are prone to volatile winds (Tian *et al.,* 2010; Lakshmanan *et al.,* 2009).

Calcutta was selected from India meteorological stations. The latitude and longitude are 22° 34' N and 88° 24' E (Figure 2.1).

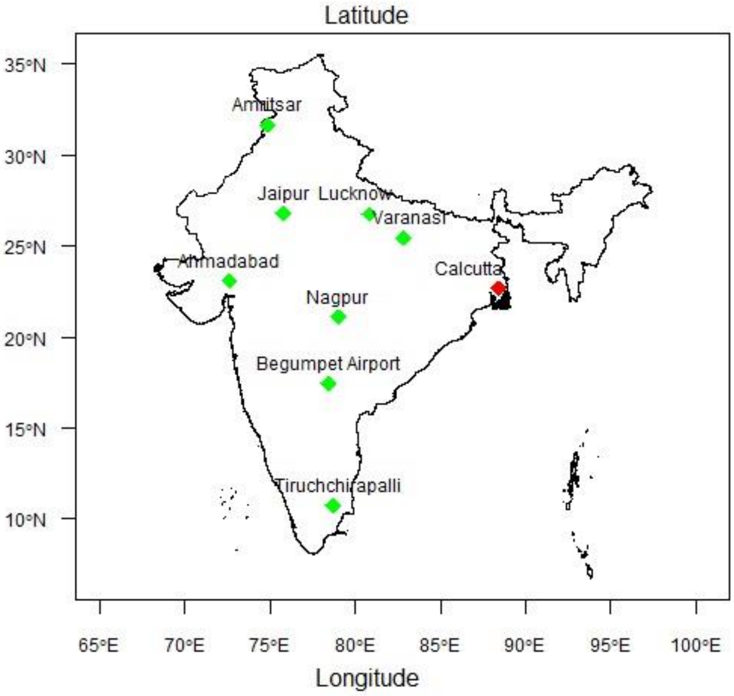


Figure 2.1 Location of 9 Indian meteorological stations

Calcutta was selected due to its consistent hourly wind data. Also, it is in wind and cyclone zone of very ‘high damage risk’ (UNDP, 2006). It has a subtropical climate with moist (rain-bearing) winds and frequent thunderstorms (Mukhopadhyay *et al.* 2009). The winds can vary from frigid winds and also associated with large storm systems such as Nor'westers and Western disturbances (Reddy, 2008). Calcutta, one of the well-known cities, is lying in the north-eastern part of India. The maximum elevation of the city is about 30 feet (9 meters) above sea level, and the area is 40 square miles, i.e. 104 square kilometers. The latitude and longitude are 22° 34' N and 88° 24' E. The Eastward region of the city is formed from the river, the land slopes away to marshes and swamplands. Similar topography on the west bank of the river has confined the metropolitan area largely to strip 3 to 5 miles (5 to 8 kilometer) wide on either bank of the river, the reclamation of the Salt Lake area on the northeastern fringe of the city.

Australia

The island continent of Australia (Figure 2.2) features a wide range of climatic zones, from the tropical regions of the north, through the arid expanses of the interior, to the temperate regions of the south. Australia experiences many of nature's more extreme weather phenomena, including tropical cyclones, severe storms, bushfires, and the occasional tornado. Australia's climate is largely determined by its latitude, with the mainland lying between 10° south to 39°South; extending to at 44° South, and longitudes 112° East and 154° East (Trewin, 2005). Melbourne, one of the popular destinations in Australia, lies on the latitude and longitude 37.8136° S, 144.9631° E.

Further, it is in the vulnerable zone of a significant increase in intense frontal systems which bring extreme winds and dangerous fire conditions (Hasson *et al.*, 2009); increases in the frequency of conditions conducive to thunderstorm development in southern and eastern areas (Allen *et al.*, 2014); and it is prone to hot, cold, gusty, dusty, tornadoes, and thunderstorm (Canning *et al.*, 2010). Melbourne is located at the head of Port Phillip Bay on the southeastern coast, and its area covers 14 square miles (36 square km); Inner Melbourne, 33 square miles (86 square km) which consist of flat terrain, Yarra River and Dandenong ranges to the east. Its location is situated on the boundary of the very hot inland areas and the cool southern ocean. Melbourne is well known for its frequent changes in weather conditions due to its geographical location.

Essendon Airport, Avalon Airport, Point Wilson, and View Bank were selected from Australian meteorological stations (Figure 2.3).

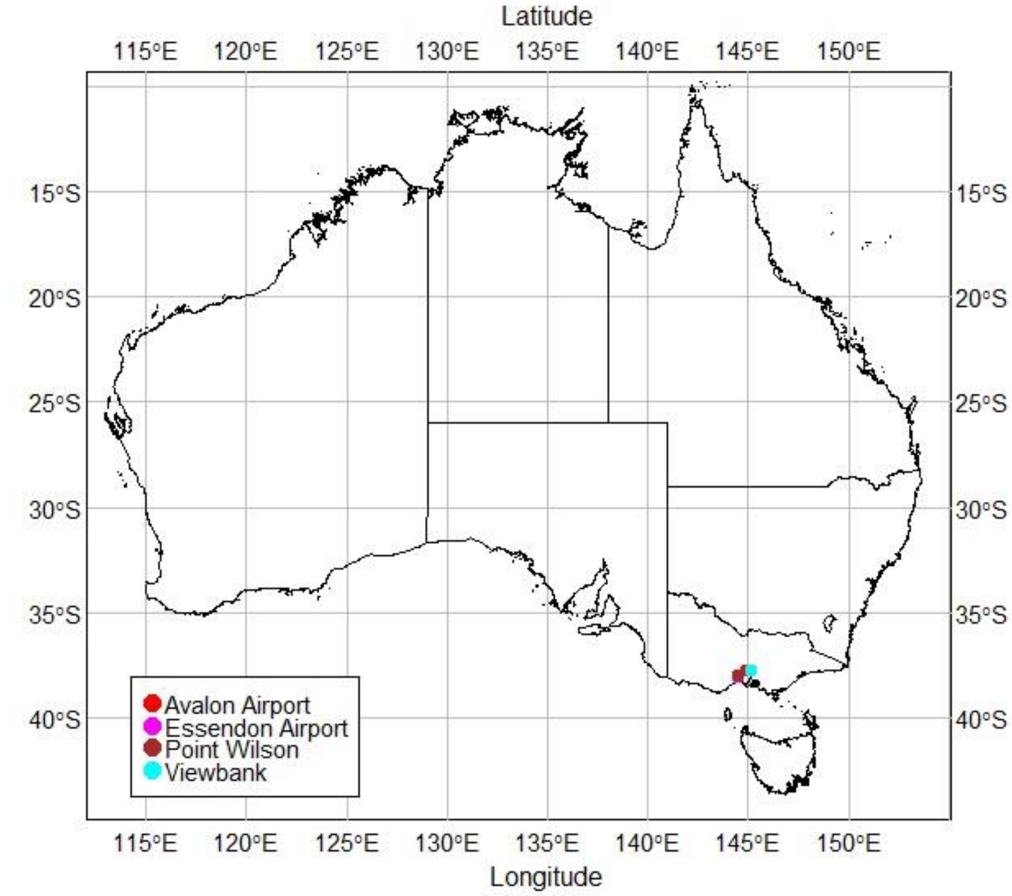


Figure 2.2 Map of Australia geographical landmass

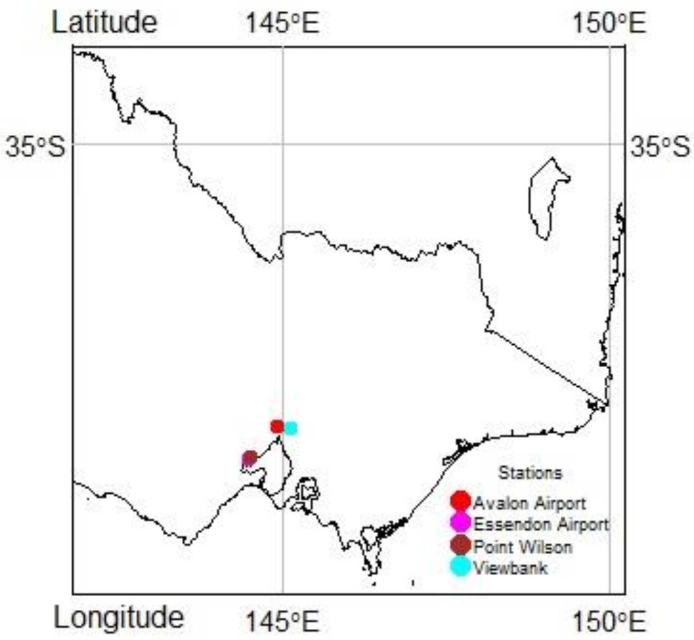


Figure 2.3 Four selected stations in Australia

Table 2.3 shows the cardinal direction, latitude, and longitude for the stations. The stations were selected due to their consistent 3 hourly wind data, and the stations' noteworthy association on the spatial wind pattern zone of westerlies. In this region, prevailing winds in the middle latitudes ranged in between 35 and 65 degrees latitude; blow wind over the poleward areas on the high-pressure area known as the subtropical ridge in the horse latitudes. These prevailing winds mainly blow from the west to the east and bring extra-tropical cyclones with its moving direction. The winds are predominantly from the southwest in the Northern Hemisphere and from the northwest in the Southern Hemisphere. These winds are stronger in the winter due to the lower pressure over the poles.

Table 2.1 Location summary for Australian weather stations.

Station	Geographic direction	Lat. (°S)	Long. (°E)
Essendon Airport	North-West	-37.72	144.90
Avalon Airport	West	-38.03	144.46
Point Wilson	South	-38.01	144.49
View Bank	East	-37.73	145.10

2.2 Data source and management

The data were obtained from the climate bureau data sources of India and Australia. The hourly wind data of the Indian region were obtained from the National Renewable Energy Laboratory (NREL) from 2002 -2008 available at http://www.nrel.gov/international/ra_india.htm. For the Australian region, the 3-hourly wind observations were obtained from the Australian Government Bureau of Meteorology at http://reg.bom.gov.au/reguser/) from the years 2002 -2008.

Data were arranged based on the Earth's surface wind direction patterns for the Indian subcontinent (Wiegand, 2004; Kious and Tilling, 1996; IMD, 2015) which is illustrated in Table 2.2; and wind direction patterns for Melbourne in Australia (BOM, 2011; Breckling, 2012), which is illustrated in Table 2.3.

Table 2.2 Earth’s surface wind direction patterns for the Calcutta in India.

Months	Spatial wind direction patterns in the Indian subcontinent
December to February	North (N), North-East (NE), North-West (NW), South-West (SW) and South (S)
March to May	South (S), North-West (NW), North-East (NE), South-West (SW), and South-East (SE)
June to September	South-West (SW) and North-East (NE)
October to November	North-East (NE), South-East (SE), South (S), South-West (SW) and North-West (NW)

Calcutta in India, the hourly wind observations from weather station resulted in 61,320 observations. The data which did not fall under the wind direction for the particular month (Table 2.2) were discarded from the study resulting in 38,157 observations during the period of study. To smooth out the fluctuations in the hourly wind measurements, 4 hour periods were used for the analysis. The measurements recorded for consecutive 4 hourly periods were considered as per the ancient time metrics of Babylon and Egyptian system. According to this system, a day (24 hours) can be represented in 4 hourly periods which form 1/4th of a day (Gillings, 1972), and is classified into definite time intervals like mid-night, dawn, morning, noon, mid-noon, night (Glickman and Zenk, 2000). There were 9,540 observations for analysis. Afterwards, the wind at least 0.775 mps was considered as the occurrence of wind (Wheeler and Wilkinson, 2004; Water, 2005); wind speed of at least 5 mps (Geer 1996) was classified as gust occurrence.

Table 2.3 Earth's surface wind direction patterns for Melbourne, Australia

Months	Spatial wind direction patterns of Australia
January	South-West (SW), North-West (NW), West (W), South-East (SE)
February	South (S), North-West (NW), West (W), South-East (SE)
March	South (S), South-West(SW), North-East (NE)
April	North (N), East (E), South-East (SE), South-West (SW)
May	West (W), North-West (NW), South-East (SE)
June	West (W), North-West (NW), South-West (SW)
July	North-West (NW), West (W)
August	North-West (NW), West (W)
September	North-West (NW), West (W)
October	West (W), South-West (SW), North-West (NW), South-East (SE)
November	South-East (SE), North-West (NW), West (W), South-East (SE)
December	North-West (NW), West (W), South-West (SW), South-East (SE)

For the Australian study region, the measurements recorded for consecutive 3 hourly periods were associated with a well-known primeval Bronze-Age meteorological system, and ancient time metrics of the Roman system (Graham and Kamm, 2014); and noteworthy heuristic weather measurement instruments of BOM Australia (BOM, 2011). Moreover, eight periods of 3 hours each of the day classified into definite time intervals like mid-night, dawn, morning, forenoon, afternoon, evening, dusk, and night (Glickman and Zenk, 2000).

Figure 2.3 and Figure 2.4 show the data management for hourly wind data for Calcutta from 2002-2008 and 3-hourly wind data for four stations in Melbourne from 2004-2008, respectively. The wind prevalence and gust were analyzed using logistic regression, the wind speed was analyzed linear regression, and wind velocity was analyzed using multinomial logistic regression.

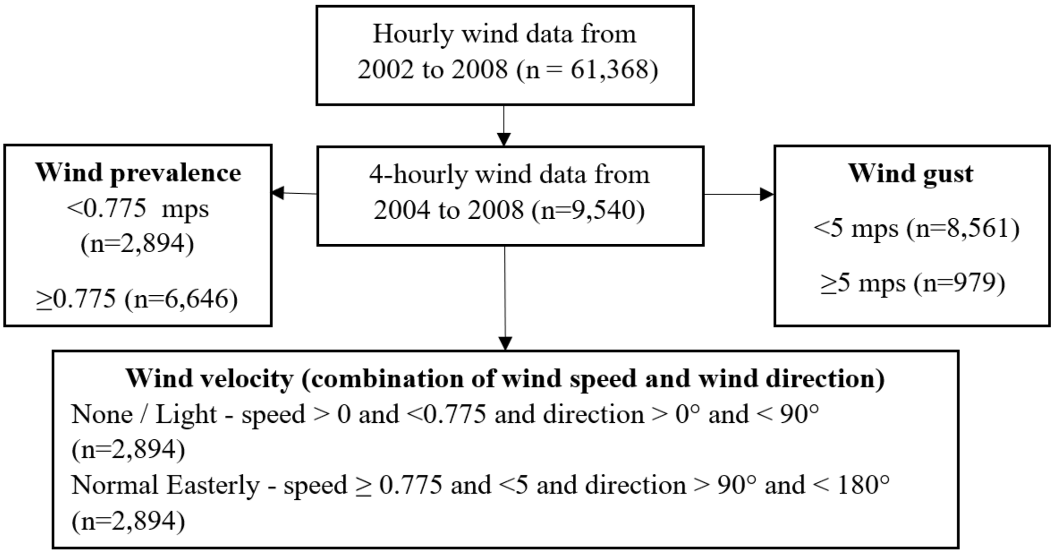


Figure 2.4 The data flow diagram for India.

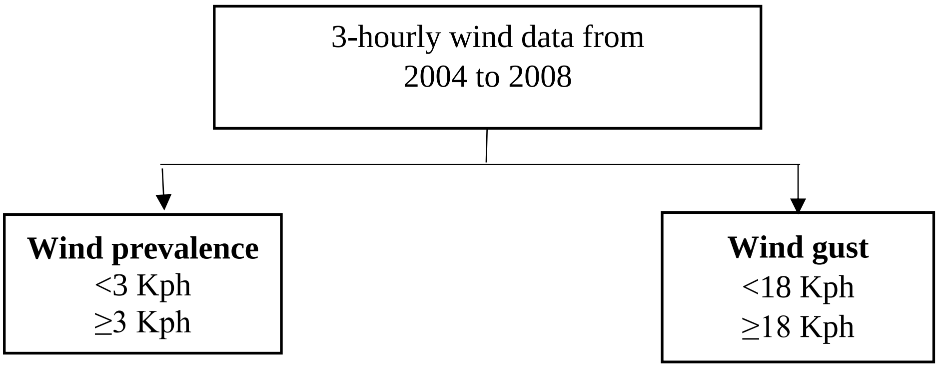


Figure 2.5 The data flow diagram for Australia.

The wind prevalence, wind gust, and wind velocity observations of for four stations were shown in Table 2.4.

Table 2.4 Observations of wind prevalence and wind gust categories for Australia

stations	Sample size	Wind prevalence (Kph)		Wind gust (Kph)	
		<3	≥3	< 18	≥18
Avalon Airport	14,493	4,066	10,427	7,579	6,914
Essendon Airport	14,524	4,493	10,031	7,909	6,615
Point Wilson	14,293	2,495	11,798	5,205	9,088
View Bank	14,586	8,052	6,534	11,318	3,268

2.3 Variables and path diagram

The outcome variables are described in Table 2.5 and Table 2.6 for Indian and Australia study regions, respectively. Table 2.7 describes the predictor variables for the Indian and Australia study regions.

Table 2.5 Dependent variables for the Indian study region.

Outcome Variable	Variable type	Variable Classification
Wind Speed	Continuous	Wind speed ≥ 0.775 mps
Wind Prevalence	Binary	Wind speed < 0.775 mps and ≥ 0.775 mps
Wind Gust	Binary	Wind speed < 5 mps and ≥5 mps
Wind Velocity	Categorical	5 combination of wind speed

Table 2.6 Dependent variables for the Australia study region.

Outcome Variable	Variable type	Variable Classification
Wind speed	Continuous	Wind speed ≥ 3Kph
Wind Prevalence	Binary	Wind speed < 3Kph and ≥ 3Kph
Wind Gust	Binary	Wind speed < 18 Kph and ≥ 18 Kph
Wind velocity	Categorical	5 combination of wind speed

Table 2.7 The predictor variables for the Indian and Australia study region.

Predictor	Variable type	4-hour wind for India	3-hour wind for Australia
Periods	Categorical	1 to 6	1 to 8
Months	Categorical	1 to 12	1 to 12
Years	Categorical	2004 to 2008	2004 to 2008

Path diagram

The path diagram for analyzing data in India and Australia is shown in Figure 2.5. The determinants were period, month and year; the outcome variables were wind prevalence, wind gust prevalence, wind speed, and wind velocity.

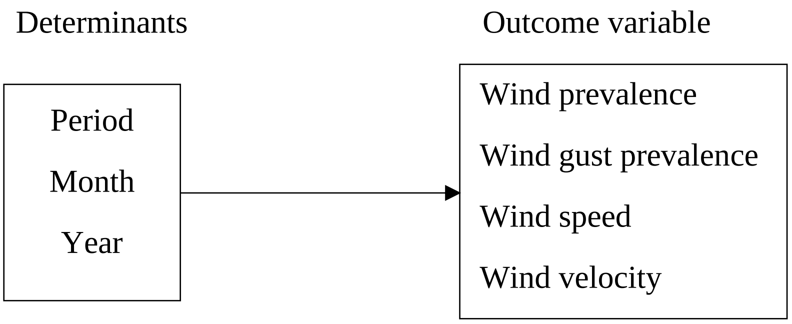


Figure 2.6 Path diagram

2.4 Statistical Methods

Logistic model

Initially, the logistic regression model was fitted to the binary outcome value of 1 for wind or 0 for no wind to determine the percentage of wind prevalence. The general form of the model can be shown by applying the formula,

$$\ln\left(\frac{p_i}{1-p_i}\right) = \alpha + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k \qquad (1)$$

Where p_i denotes the probability of wind prevalence for the i^{th} period in the data set, conditional on the variables x_i where α is the constant, β's are regression coefficients, and x's are the predictors. Model (1) was also used to estimate the prevalence of wind gust using sum contrasts ((Tongkumchum and McNeil, 2009). These models were assessed using the Receiver Operating Characteristic (ROC) curve. The Receiver Operating Characteristic (ROC) curve (Westin 2001) was used as a measure of goodness-of-fit of the model. The ROC curve is popularly known for determining the capability of predicting a binary outcome. The ROC curve plots sensitivity against the false positive rate of the 4 hourly wind prevalence.

Linear Model

Multiple regression was used to describe the trends of wind speed. The model takes the form

$$y_{ijk} = \alpha + \beta_1 x_i + \beta_2 x_j + \beta_3 x_k + \gamma_{t-1}, \qquad (2)$$

where y_{ijk} represents wind speed trends, α is a constant, β*'s* are the regression coefficients x_i, x_j, x_k, are time indicates the predictors 4-hour periods, month and year, γ denotes the regression coefficients of the trimmed lag 1 term and $t - 1$ indicates lag 1 period. The goodness of fit of the model was assessed using the coefficient of determination (R-square) and the Q-Q plots. Further, since the normality assumption was not satisfied as the data were not normally distributed, due to heavy tails; the wind speed values were transformed by square roots, to satisfy the statistical assumption of normality.

Multinomial model

Multinomial logistic regression is used for a categorical dependent variable with two or more levels (van Smeden *et al*., 2017). The multinomial model was fitted to analyze these combined, matched and classified discrete and nominal data of wind speed and direction (wind velocity) categories. The model takes the form

$$p_k = \frac{\exp\left(\alpha_k + \sum_{j=1}^{m} \beta_{jk} x_j\right)}{1 + \sum_{k=1}^{c} \exp\left(\alpha_k + \sum_{j=1}^{m} \beta_{jk} x_j\right)} \tag{3}$$

This equation is extended to situations in which the outcome variable is nominal with more than two categories. If these outcome categories are coded as 0, 1, 2, …, *c*, and p_k is the probability that outcome has the coded value *k*, the model takes the form, for $0 \le k \le c$, where $c = 5$ outcome categories of combined discrete measures of wind speed and direction (wind velocity), i.e. none/light wind, normal easterly, normal westerly, gusty easterly, gusty westerly, x_i is the vector of explanatory variables indicates wind speed and wind direction describing observation j indicates 4-hour periods, β_{jk} is regression coefficients for outcome *k*, *m* is the number of independent variables indicates the predictors 4-hour periods, month and year, α_k is a constant for categories of outcome *k*.

CHAPTER 3

PRELIMINARY ANALYSIS OF WIND SPEED FOR INDIA

This chapter reports on the preliminary analysis of hourly wind data during 2002-2008, and analysis of 4-hourly wind characteristics from 2004 - 2008 in Calcutta, India.

3.1 Hourly wind speed for Calcutta

The hourly average of wind speed in Calcutta is 1.68 mps, with a standard deviation of 1.52 mps, maximum of 15.4 mps. The time series plot (Figure 3.1a and Figure 3.1b) gives a detailed account of annual seasonal patterns of wind during 2002 – 2008 for this station. The horizontal axis scale reveals the wind direction (absolute degree) measure indicating discontinuities removed by extending the compass. The vertical axis scale reveals the hours of the day (marked as black vertical bars), and the three grid blocks illustrate each trimester of the year.

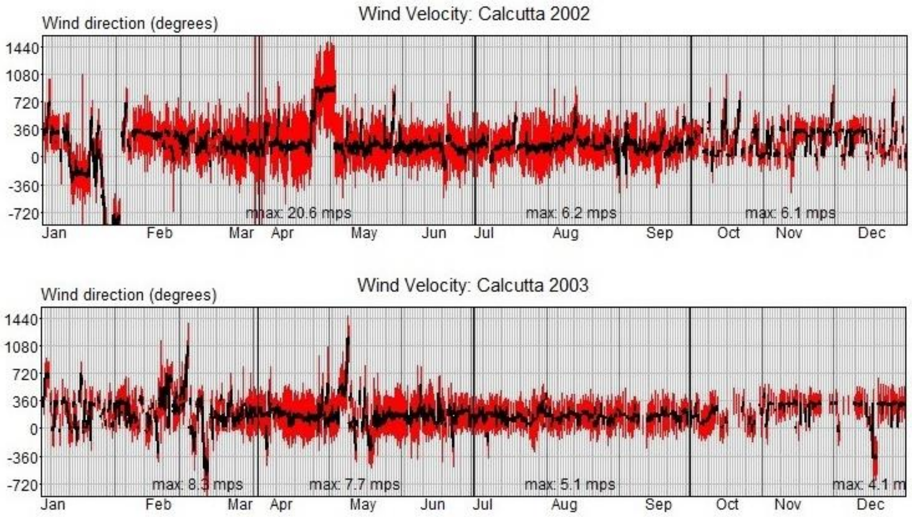


Figure 3.1a Episode-based time-series of wind velocity in absolute degrees for 2002 and 2003

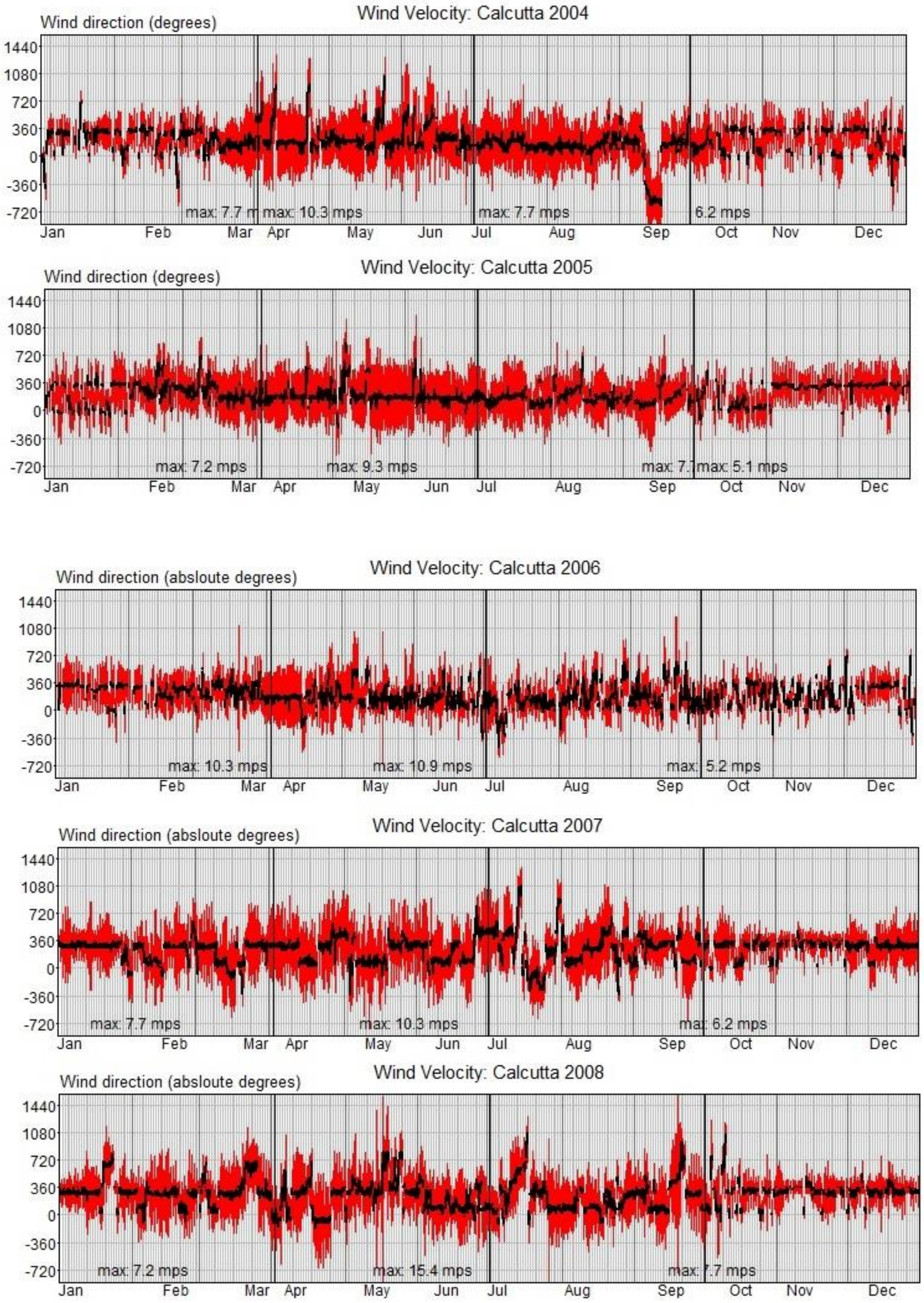


Figure 3.1b Episode-based time-series of wind velocity in absolute degrees for 2004 and 2008

The red line denoted wind speed pattern and black bars indicate wind direction drift. The wind speed patterns revealed apparent seasonal patterns. The years 2002, 2003, 2007 and 2008, the period from January to June illustrated an increasing trend, July to September revealed a constant trend, October to December revealed a low trend.

The year 2005 to 2007, the period from January to March revealed a constant trend, April to September illustrated an increasing trend, October to December revealed a decreasing trend.

3.2 Analysis of the 4-hourly wind speed

The wind observations from the years 2002 and 2003 consisted of larger portions of data with constant wind patterns and were omitted to maintain consistency with data. The section comprises modeling of the wind characteristics of the 4-hourly wind observations during 2004-2008. The results of this study also appear in the manuscripts (Gururaja *et al.*, 2019a; Gururaja *et al.*, 2019b).

Wind prevalence patterns

A logistic regression model was fitted to 4 hourly wind periods for 2004 - 2008.

The model for the 4-hourly mean of wind prevalence was made of 3 predictors. The first predictor was the 4-hourly period of the day and had 6 factors, the second predictor was the months with 12 factors, and the third factor was the year having 5 factors. There were 24 parameters, including the constant term in the model. The parameters were highly significant and influential in the model. ROC curve (Figure 3.2) was used for assessing the goodness of fit of the model; the curve revealed an area under the ROC curve value of 0.75 which shows a sound classification of the prevalence and non-occurrence of wind.

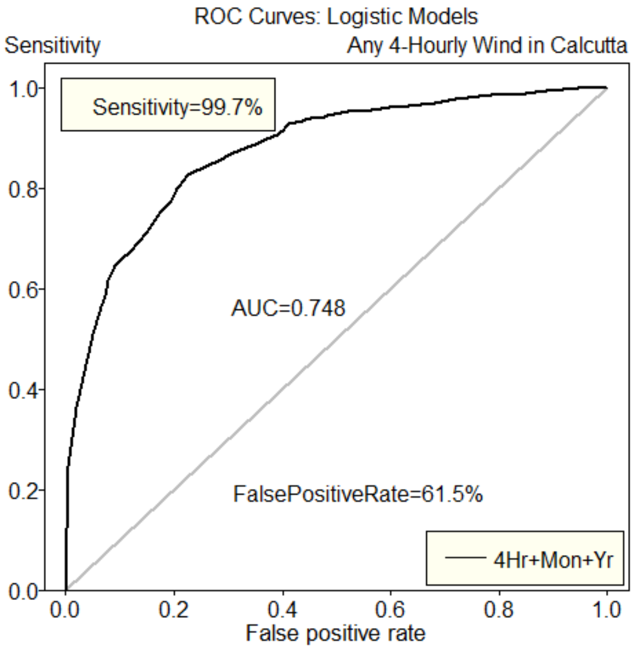
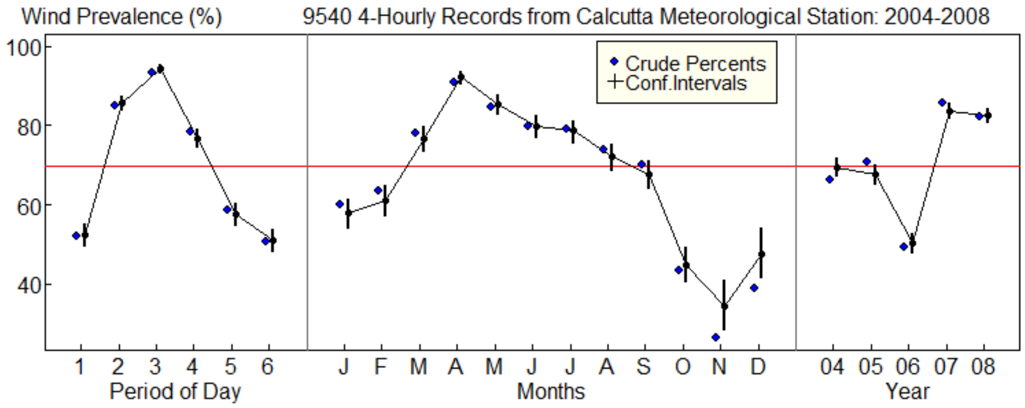


Figure 3.2 ROC curve of the logistic model for a 4 hour period of wind

Besides, the overall accuracy of the classification on the prevalence and non-prevalence of wind was 85%, having a true positive rate of 99.7% with a false positive rate of 61.5%. Further, the plot of confidence intervals (Figure 3.3) illustrates the pattern of wind prevalence percentage for the day, month and year and the crude percentage. The horizontal red line denotes the overall mean of the wind prevalence, which is approximately 70%. Wind speed greater than 0.775 (mps) were classified as wind prevalence. The 95% confidence interval of wind prevalence revealed that the wind prevalence patterns for the periods of the day were significantly different from the overall mean.



Note:- 1 = After 1AM and Before 4AM, 2 = After 5AM and Before 8AM, 3 = After 9AM and Before 12PM
4 = After 1PM and Before 4PM, 5 = After 5PM and Before 8PM, 6 = After 9PM and Before 12PM
The hour denotes the starting hour of the 4-hour time period

Figure 3.3 95% confidence interval plot for percentages of wind prevalence.

The period from dawn (Period 2) to afternoon (Period 4) observed high wind prevalence rates with the maximum prevalence of around 90% observed in the noon (Period 3). The low wind prevalence rates were observed from midnight to dawn (Period 1), evening (Period 5) to night (Period 6) with the minimum prevalence of around 50% in the night (Period 6) respectively. The monthly wind prevalence patterns were observed to have high wind prevalence rates from March to July, with a maximum wind prevalence of 90% observed in April. The moderate wind prevalence rate was observed during August and September with a modest wind prevalence of 70% aligned with the overall mean. Wind prevalence rates were observed to be less than the overall mean during January and February. It increases gradually to attain its maximum (90%) in April, decreases steadily, attains its minimum (30%) in November. The yearly wind prevalence patterns were not significantly different from the overall mean between 2004 and 2005. It reduced considerably to its minimum in 2006 and rose steadily to its maximum in 2007 where it began to reduce slightly. The yearly wind prevalence ranged from 50% to 85% and observed to be higher than the overall mean between 2007 and 2008.

Wind gust prevalence patterns

Model (1) was also used to examine the 4 periods of wind gust prevalence from 2004-2008. The 4-hourly mean wind gust prevalence model was also made of 3 predictors. Assessing the goodness of fit of the model by the ROC curve (Figure 3.4) revealed an area under the ROC curve value of 0.608, which shows the classification of the occurrence and non-occurrence of wind.

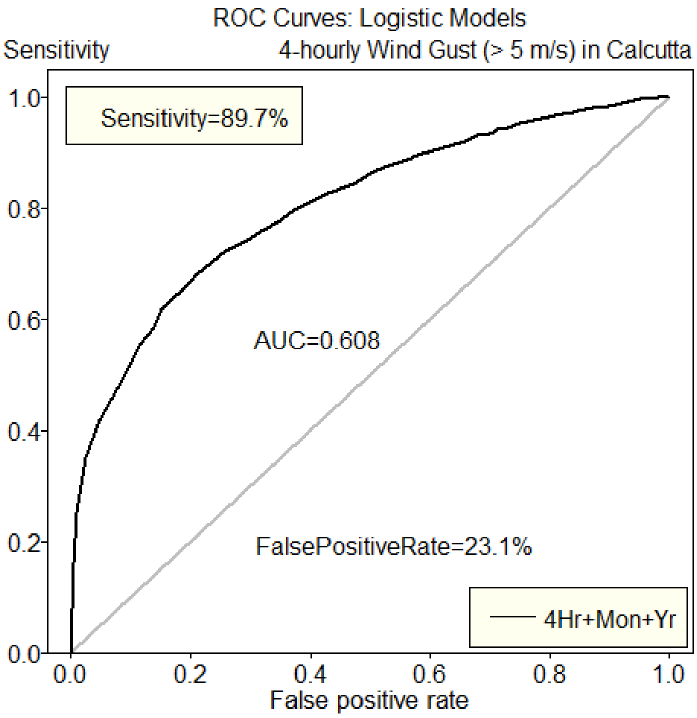


Figure 3.4 ROC curve of the logistic regression model of wind gust prevalence

The overall accuracy of the classification on the prevalence and non-prevalence of wind gust was 81.67% having a true positive rate of 89.7% with a false positive rate of 23.1%. Wind speed greater than 5 mps were classified as wind gust.

In Figure 3.5, the horizontal line denotes the overall mean of the prevalence of wind gusts, which is around 10% during the study period. The wind gust patterns for the periods of day were significantly different from the overall mean.

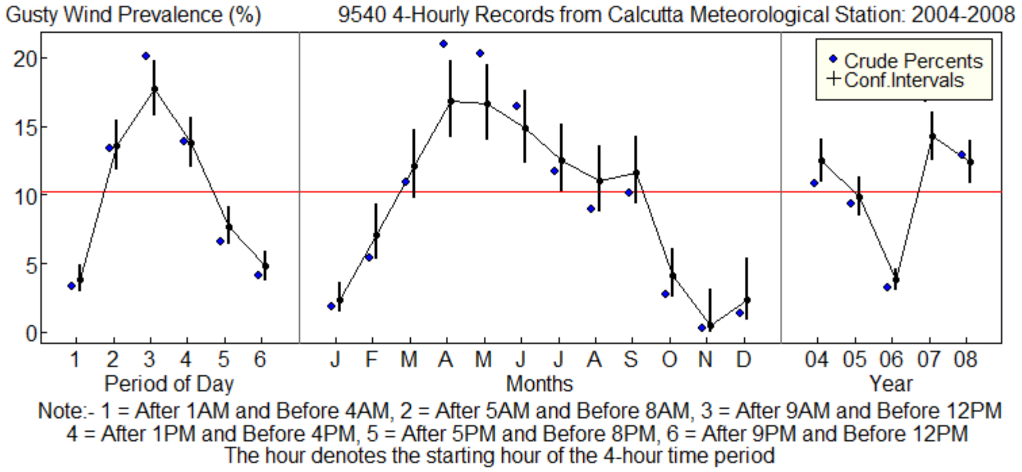


Figure 3.5 The 95% confidence interval plot for the percentage of wind gusts

The period from midnight (Period 1) to noon (Period 3) observed increasing wind gust prevalence rates with the maximum prevalence of around 18% observed in the noon (Period 3). It then decreases steadily from this period to the night. The minimum wind gust prevalence of 4% occurred at midnight (period 1). The monthly wind gust patterns revealed an increasing pattern from January to April, where the maximum was attained. It was constant between April and decreased steadily until August where it increased again until September. It then decreases sharply to its minimum in November. The wind gust prevalence was observed to be mostly above the overall mean between April and July. The yearly wind gust prevalence revealed a decreasing pattern from 2004 to 2006, where it attained a minimum of 4% and increased sharply to its maximum of 14% in 2007. There was a decreasing pattern between 2007 and 2008.

Wind speed trends

The linear regression model (model 2) was fitted to the 4 hourly wind speed to examine the trends. The model for 4 hourly period wind speed from 2004 -2008 was also made of 3 predictors.

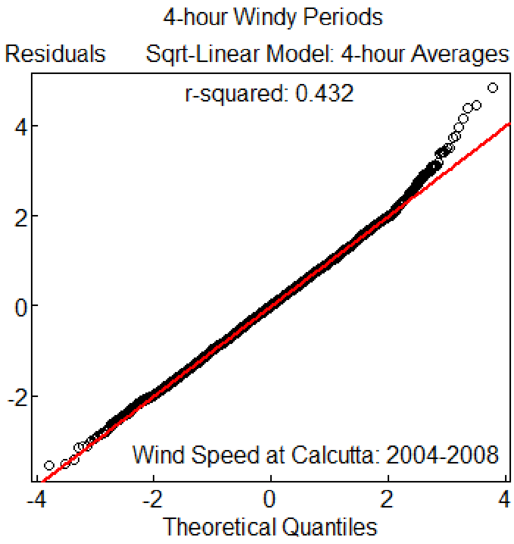


Figure 3.6 Q-Q plots for the linear model of the wind speed

Most of the parameters were highly significant and influential in the model. The coefficient of determination (r-squared) and the quantile-quantile plots was used to evaluate the goodness of fit of the model. The r-squared value of 0.432 shows that 43% of the wind speed patterns are explained by the model (2).

The quantile-quantile (Q-Q) plot illustrates that the residuals from the model are approximately normally distributed (Figure 3.6), which shows that the model fitted the data reasonably well. However, there were some deviations from the upper tail of the model. This deviation may be due to extreme values in the data and some variations that could not be explained by the predictors.

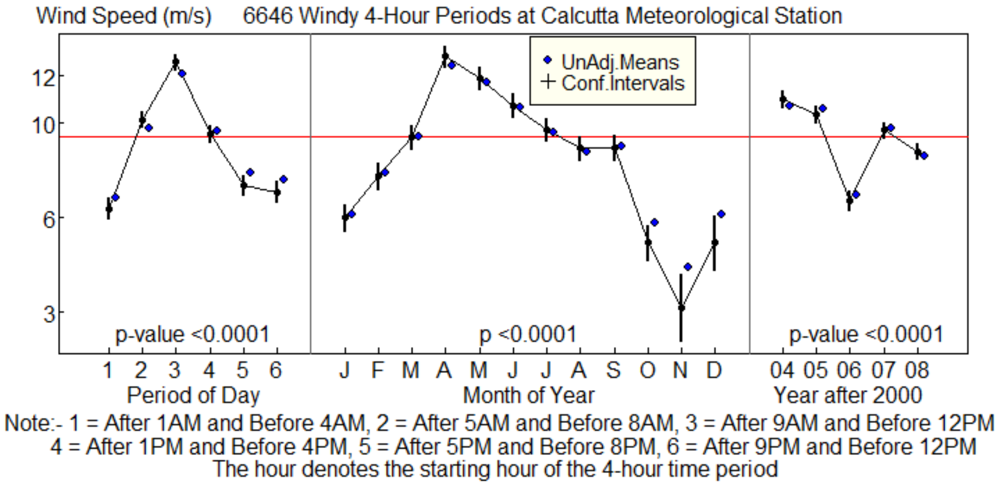


Figure 3.7 95% confidence interval plot of wind speed trends.

Figure 3.7, the horizontal line denotes the overall mean of the wind speed is approximately 9 ms^{-1}. The wind speed trends for the periods of the day were significantly different from the overall mean. The period during dawn (Period 2) to noon (Period 3) had the maximum wind speed of around 13 ms^{-1} observed during noon (Period 3). Moderate wind speed trend was observed during the afternoon (Period 4) with a modest wind speed of approximately 9 ms^{-1} aligning with overall mean. The low wind speed trends were observed during midnight (Period 1) to dawn (Period 2), evening (Period 5) tonight (Period 6) with a minimum wind speed of around 6 ms^{-1} in midnight (Period 1) respectively. The monthly wind speed trends revealed rising trends from April to June with a maximum wind speed of around 13 ms^{-1} observed during April. Moderate wind speed trend was observed during March and July to September with a modest wind speed of approximately 9 ms^{-1} aligning with overall mean. The low wind speed trends were observed from January to February, and October to December with a minimum wind speed of around 3 ms^{-1} in November. The monthly and yearly wind speed trends revealed values either higher, equal to or lower than the overall mean. The yearly wind speed revealed reducing trends during 2004 and 2005 with a maximum wind speed of around 11 ms^{-1} observed during 2004. Moderate wind speed was observed during 2007 with a modest wind speed of approximately 9 ms^{-1} aligning with overall mean. The increasing trends were during 2006 and 2008 with a minimum wind speed of around 7 ms^{-1} in 2006.

Wind velocity categories prevalence

The results of the multinomial regression method applied to model the occurrence probabilities of the 4 hourly wind velocity categories during 2004-2008 have been given in Table 3.1. The model for the wind velocity categories was made of 3 predictors. There were 24 parameters, including the constant term in the model. The model parameters were influential, and most of them were highly significant ($p < 0.01$) in the model.

Table 3.1 gives the model coefficients. The model revealed that a unit increase in wind velocity in period 1 means that, there will be 0.058, 52.446, 49.624 and 55.256 times more likely to stay in the normal easterly, normal westerly, gusty easterly and gusty westerly categories respectively. However, these were not significant. In general, analysis of the various wind categories during periods (1-6) revealed that a unit increase of wind velocity in any period would result in wind velocity that is more likely to stay in the gusty westerly category. These results were highly significant in all periods except in period 1 and 6.

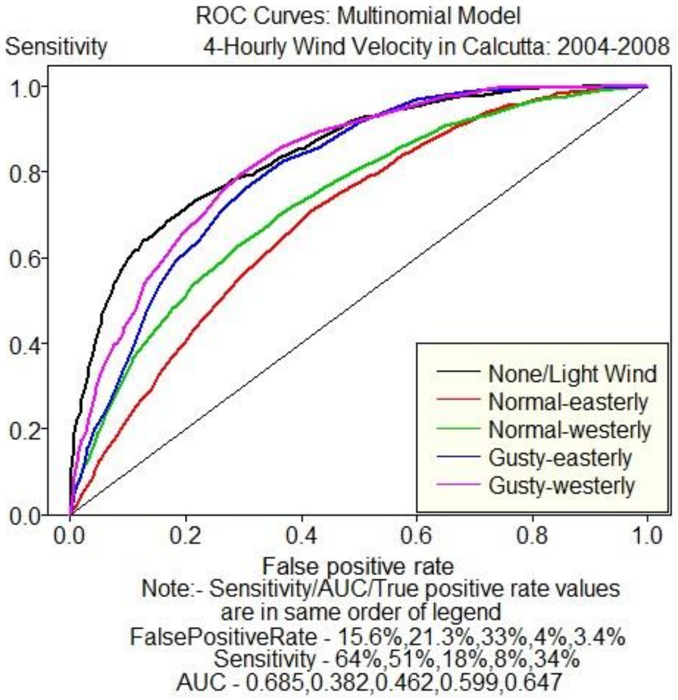
Table 3.1 The results of the multinomial modeling

	Dependent variable (p <0.01)				
	None/light	**Normal easterly**	**Normal westerly**	**Gusty easterly**	**Gusty westerly**
Constant	-3.373	-2.335*	-0.861*	-14.033	-5.633*
Period					
1	-1.188	0.058	52.446	49.624	55.256
2	0.949	1.998*	2.113*	2.936*	3.909*
3	1.990	2.956*	3.158*	4.228*	5.543*
4	0.362	1.548*	1.352*	2.226*	3.577*
5	-0.857	2.600*	0.880*	12.461	3.345*
6	-1.283	-0.044	-0.179*	0.027	0.437
Month Factors					
January	-0.737	0.076	57.780	54.101	61.375
February	-0.542	0.860*	-0.116	8.997	0.991*
March	0.360	1.839*	0.672*	10.642	2.488*
April	1.640	3.243*	1.705*	12.739	4.369*
May	0.932	2.600*	0.880*	12.461	3.345*
June	0.542	2.505*	-0.022	12.236	1.964*
July	0.469	2.317*	0.327*	11.451	2.189*
August	0.113	2.101*	-0.250*	11.226	0.809*
September	-0.130	1.797*	-0.399*	11.125	0.881*
October	-1.475	0.493*	-1.616*	8.839	-1.213*
November	-2.095	-0.486*	-1.671*	-4.300*	-2.665*
December	-1.314	0.120	-0.805*	-4.602*	-0.557
Year Factors					
2004	0.011	0.052	69.497	67.289	71.619
2005	-0.119	0.067	-0.295*	-0.234	-0.728*
2006	-1.094	-0.907*	-1.132*	-2.139*	-2.370*
2007	0.806	0.520*	1.053*	1.126*	0.740*
2008	0.719	0.350*	1.093*	0.604*	0.635*

Note: * means the parameter is significant in the model.

Analysis of the monthly wind pattern from Table 3.1 revealed that a unit increase in wind velocity in January would result in 61.375, 54.101, 57.780 and 0.074 times more likely to stay in gusty westerly, gusty easterly, normal westerly and normal easterly categories respectively, relative to it staying in the none/light wind category. The gusty easterly wind is highly expected from February to December since a unit increase in the wind during these months results in higher ratios relative to other wind categories. Considering the annual factors, a unit increase in wind velocity resulted in the high gusty westerly wind category in 2004, while high normal easterly wind velocity observed during 2005-2006. Also, high gusty easterly wind velocity was observed in 2007, while normal westerly wind velocity was prevalent in 2008.

Further, the goodness of fit of the multinomial model for the occurrence probabilities for the various wind velocity categories was evaluated using the ROC curve (Figure 3.8). Analysis of the ROC curve revealed that unlike the normal easterly and westerly, which show poor classification, the remaining wind categories events did fairly well. The goodness of fit of the models by the ROC curve revealed an area under ROC curve (AUC) values of 0.69, 0.38 and 0.46 for wind velocity categories of none/light, normal easterly and normal westerly respectively. The overall accuracy of the classification on the occurrence and non-occurrence of the various wind categories were 30%, 31%, 28%, 5%, and 4% having a true positive rate of 64%, 51%, 18%, 8%, and 34% with a false positive rate of 15.6%, 21.3%, 33%, 4%, and 3.4% respectively. These indicators depict a reasonable fit of the model to the data.


ROC Curves: Multinomial Model

4-Hourly Wind Velocity in Calcutta: 2004-2008

Note:- Sensitivity/AUC/True positive rate values are in same order of legend

FalsePositiveRate - 15.6%,21.3%,33%,4%,3.4%

Sensitivity - 64%,51%,18%,8%,34%

AUC - 0.685,0.382,0.462,0.599,0.647

Figure 3.8 ROC curves for the multinomial model

The logistic regression models of the 4-hourly mean wind velocity category of gusty westerly and easterly were each made up of 3 predictors. There were 17 parameters, including the constant term in the models. The parameters were highly significant and influential in the models.

Extreme wind velocity categories prevalence

Evaluation of the fitted logistic regression model with the ROC curve (Figure 3.9) shows that the model with 4-hour periods, month and year for investigating gusty easterly and westerly prevalence are reasonable and provide a good description of the main features of the data.

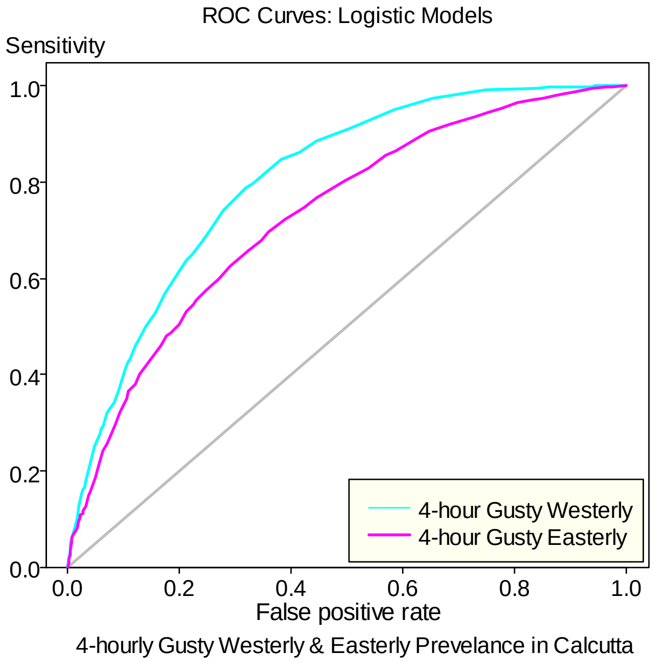


Figure 3.9 ROC curves for the logistic regression model of gusty westerly and easterly

AUC values of 0.60 and 0.65 were observed for the gusty easterly and gusty westerly, respectively. These values show a fair classification of the prevalence and non- prevalence of wind velocity categories.

The plot of confidence intervals illustrates the gusty westerly and easterly prevalence percentage, respectively, for the periods of the day, month and year (Figure 3.10 and Figure 3.11) and the crude percentage. The horizontal line denotes the overall mean of the gusty westerly and easterly prevalence of approximately 5%. The gusty westerly prevalence patterns for the period of the day were significantly different from the overall mean (Figure 3.11). The period from dawn (Period 2) to afternoon (Period 4) observed high westerly gust prevalence rates with the maximum gusty westerly prevalence of around 9% observed in the noon (Period 3). The low gusty easterly prevalence rates were observed from midnight to dawn (Period 1), evening (Period 5) to the night (Period 6) with minimum westerly gust prevalence of around 1% in dawn (Period 1) respectively.

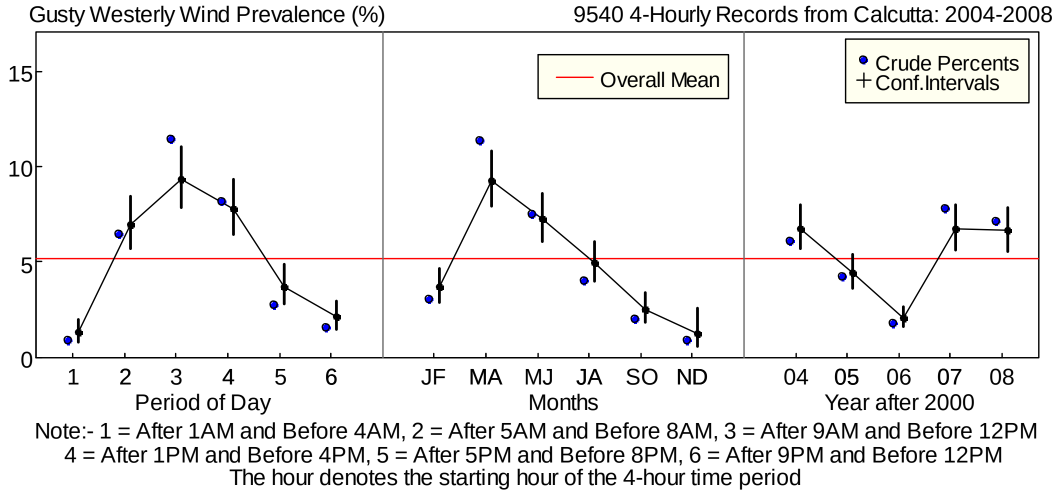


Figure 3.10 95% Confidence interval plot of gusty westerly wind prevalence

The monthly gusty westerly prevalence patterns revealed a sharp increase in prevalence from January to April where it attained its maximum (9%) above the overall mean then it reduced rapidly from April to December to attain its minimum (1%) below the overall mean. The yearly gusty westerly prevalence patterns revealed an instant decrease in gusty westerly prevalence from 2004 to 2006. It then attained a minimum of 2% after 2006. There was a steep increase to attain its maximum of almost 7% in 2007 and a gradual decrease towards 2008. The gusty easterly wind gust patterns for the period of the day were also significantly different from the overall mean (Figure 3.12). During (period 1) at dawn to (period 3) noon illustrated a quick increase of prevalence where it attained its maximum of almost 8% above the overall mean, after that, it decreased rapidly during (period 3) to (period 6) at night below the overall mean. The minimum prevalence was observed in (period 1) at midnight (1%).

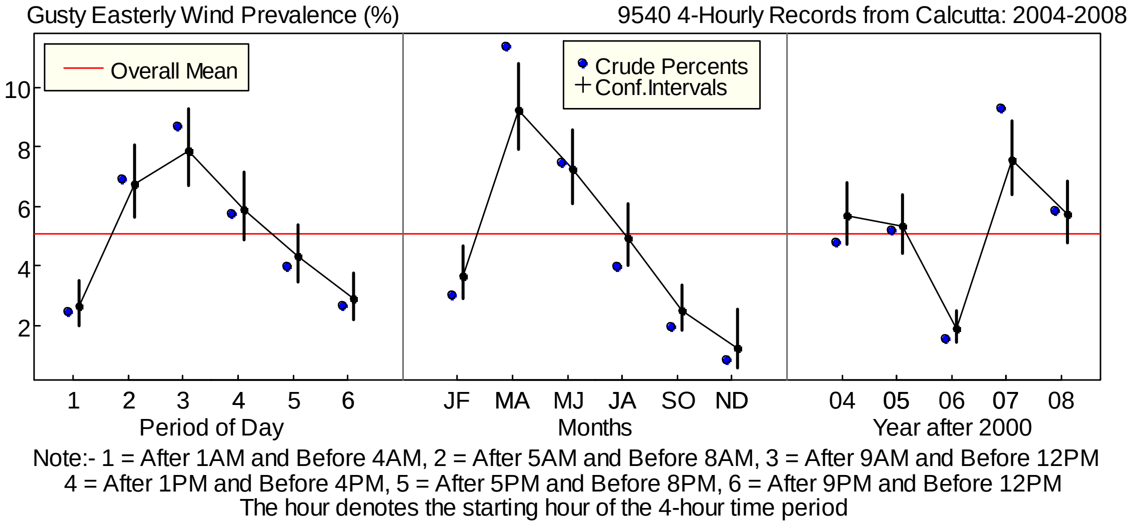


Figure 3.11 95% Confidence interval plot of gusty easterly wind prevalence

The monthly gusty easterly prevalence patterns revealed a sharp increase in prevalence from January to April where it attained its maximum (9%), remained constant from April to October below the overall mean then it reduced briskly from October to December to attain its minimum (2%) below the overall mean. The yearly gusty easterly prevalence patterns revealed decreasing trends during 2004 and 2006, where it attained its minimum of 2% below the overall mean. An immediate increase of up to 7% was observed in 2007, and a graduate decrease towards 2008

CHAPTER 4

PRELIMINARY ANALYSIS OF WIND SPEED FOR AUSTRALIA

This chapter reports on the preliminary analysis of 3-hourly wind data during 2004-2008 in Melbourne, Australia.

4.1 Summary of the 3-hourly wind data

The statistical summaries of the 3-hourly wind for or the 4 selected sample stations for Melbourne, namely Avalon airport, Essendon airport, Point Wilson, and Viewbank during 2004 - 2008 are described in Table 4.1. The 3-hourly mean wind speed ranged from 12.4 Kph to 22.3 Kph. The highest value occurred in Point Wilson while the least occurred in Viewbank. The maximum wind speed was recorded in Point Wilson of 83.2 Kph. The standard deviation of wind speed was ranged from 7 Kph to 10 Kph.

Table 4.1 Numerical summaries of wind speed for selected Meteorological stations.

stations	Wind speed (Kph)			Average Direction (degrees)
	Mean	Standard Deviation	Maximum	
Avalon airport	18.6	10	61.2	201°
Essendon airport	17.7	9	57.2	201°
Point Wilson	22.3	10	83.2	190°
Viewbank	12.4	7	48.2	166°

The time series plot (Figure 4.1a and Figure 4.1b) gives a detailed of wind patterns during 2004-2008 for the station of Avalon airport. In the plot, the horizontal axis scale revealed the month of the year, and the vertical axis scale revealed the wind speed.

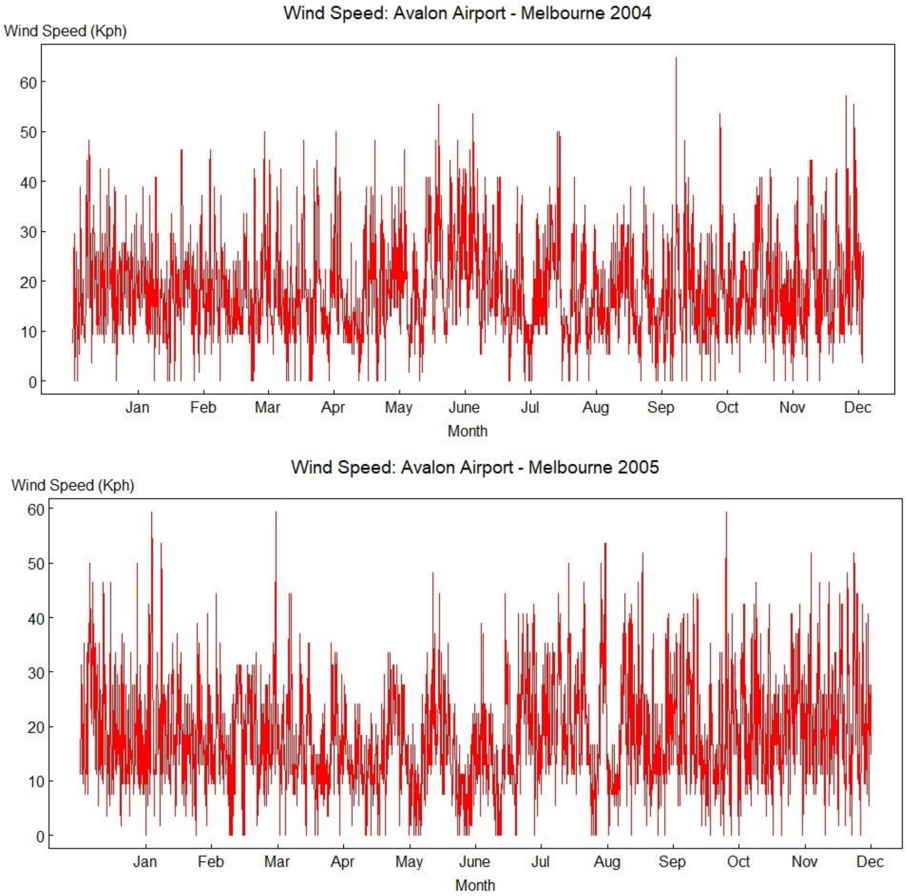


Figure 4.1a The 3-hourly episode-based time-series of wind speed for the station of Avalon airport for the years 2004 to 2005.

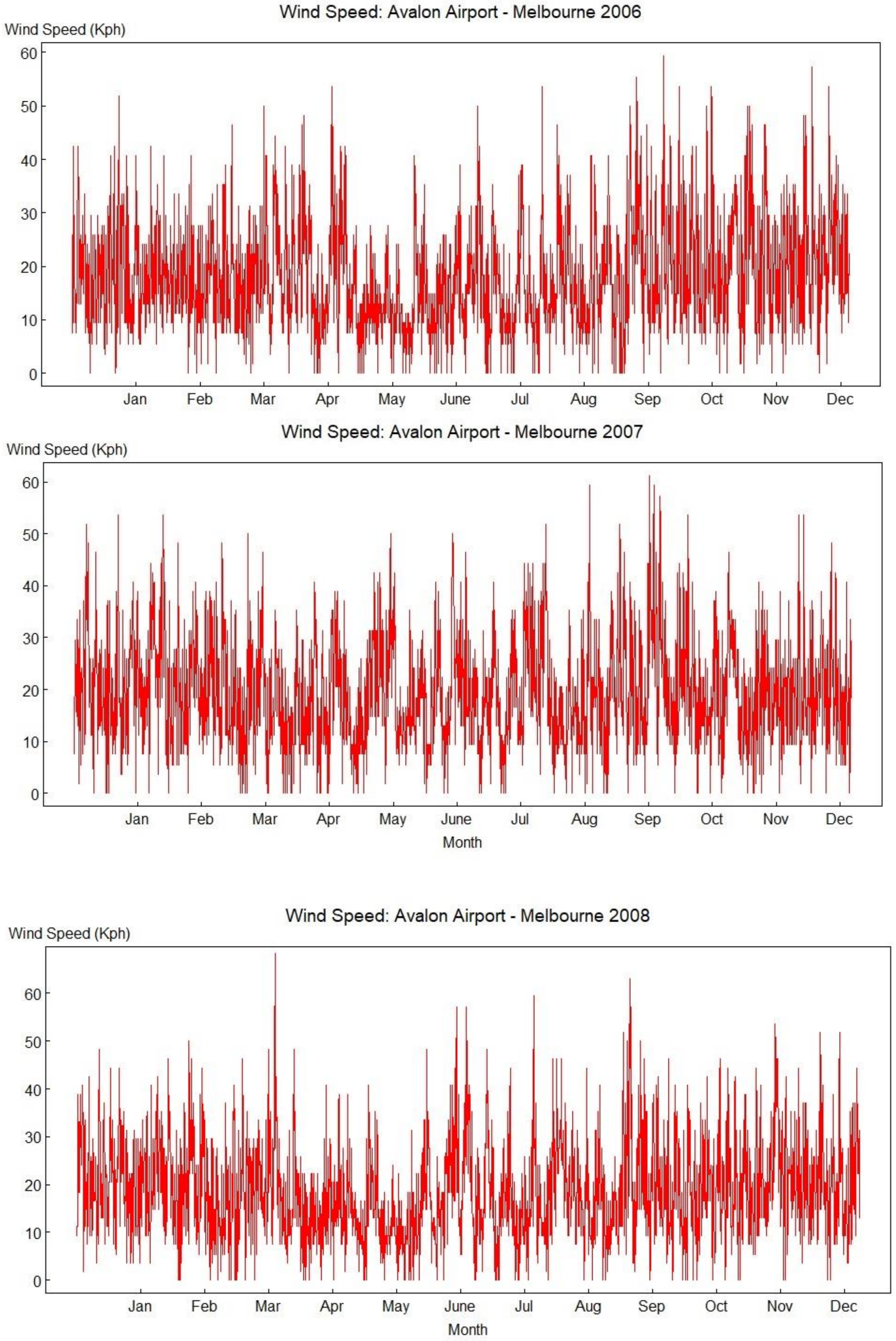


Figure 4.1b. The 3-hourly episode-based time-series of wind speed for the station of Avalon airport for the years 2006 and 2008.

The time-series of wind in the years 2005, 2006, and 2008 had similar patterns. The period from January to May illustrated a constant trend, a decreasing trend was observed in some periods in April, July, and September. The monthly wind in 2004 was high in the middle of the year. The monthly wind patterns of 2007 fluctuated throughout the year.

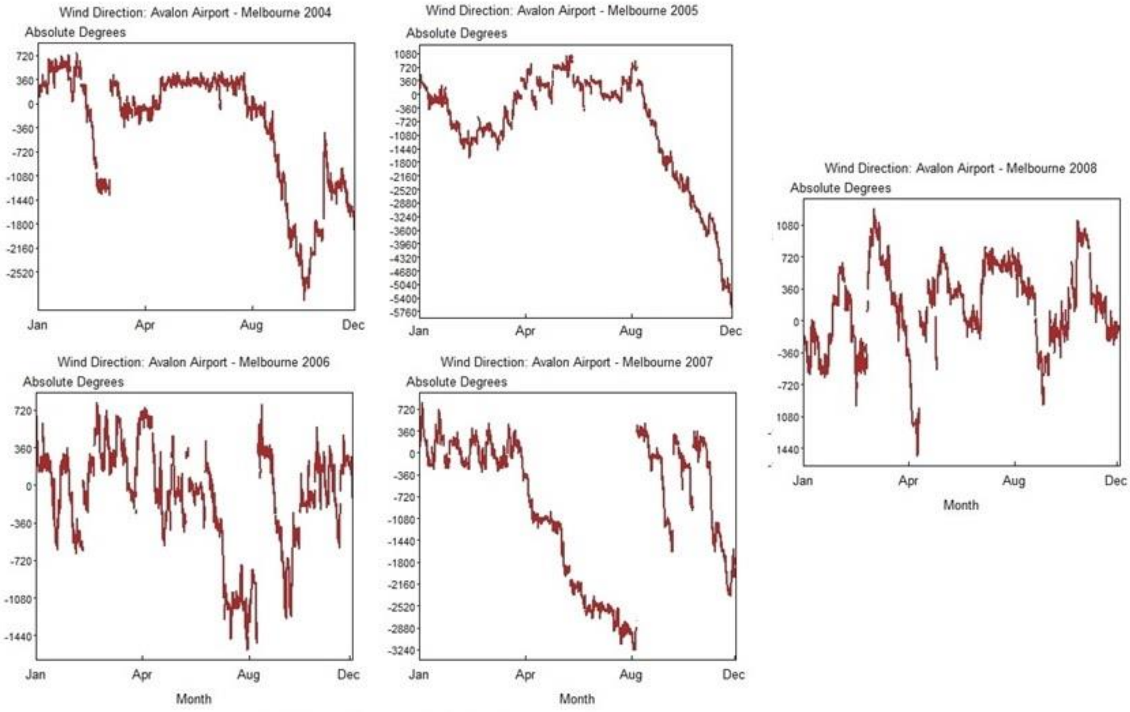


Figure 4.2 The 3-hourly episode-based time-series of wind direction in absolute degrees for the station of Avalon airport for the year 2004 to 2008.

The horizontal axis scale revealed the month of the year, and the vertical axis scale revealed the wind direction (absolute degree) measure indicating discontinuities removed by extending the compass. The wind direction movement illustrated a gradual clockwise drift from January to March, and a fluctuating drift was noticed from April to December. Results for all other stations are shown in Appendix A.

4.2 Modelling of 3- hourly wind characteristics

This sub-section comprises modeling of the 3-hourly wind observations in Avalon airport, Essendon airport, Point Wilson, and Viewbank during 2004-2008.

Wind prevalence patterns

A logistic regression model was fitted to 3 hourly wind prevalence from 2004 - 2008. They are 3 predictors. The first predictor was the 3-hourly period of the day and had 8 factors, the second predictor was the months with 12 factors, and the third factor was the year having 5 factors. There were 26 parameters, including the constant term in the model. The parameters were highly significant and influential in the model. ROC curve (Figure 4.3) was used for assessing the goodness of fit of the model; the plot revealed an area under the ROC curve (AUC) values of 0.450, 0.446, 0.308, and 0.396 respectively.

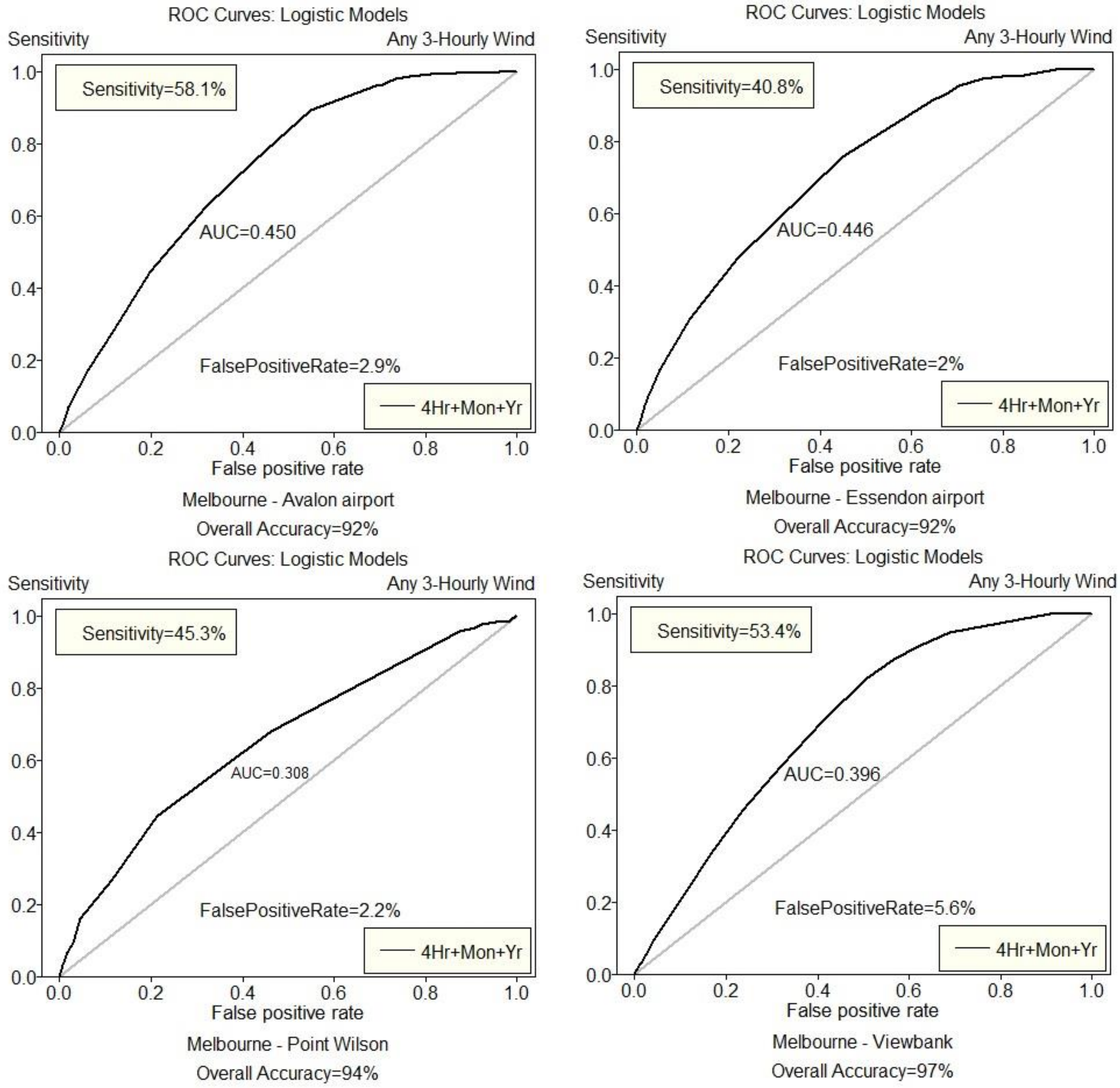


Figure 4.3 ROC curve of the logistic model for 3 hour period of wind for 4 stations

The curve revealed the overall accuracy on the classification on the occurrence and non-occurrence of wind with the rate of 92%, 92%, 94%, and 97%, having a true positive rate of 58.1%, 40.8%, 45.3%, and 53.4% with a false positive rate of 2.9%, 2%, 2.2%, and 5.6% respectively.

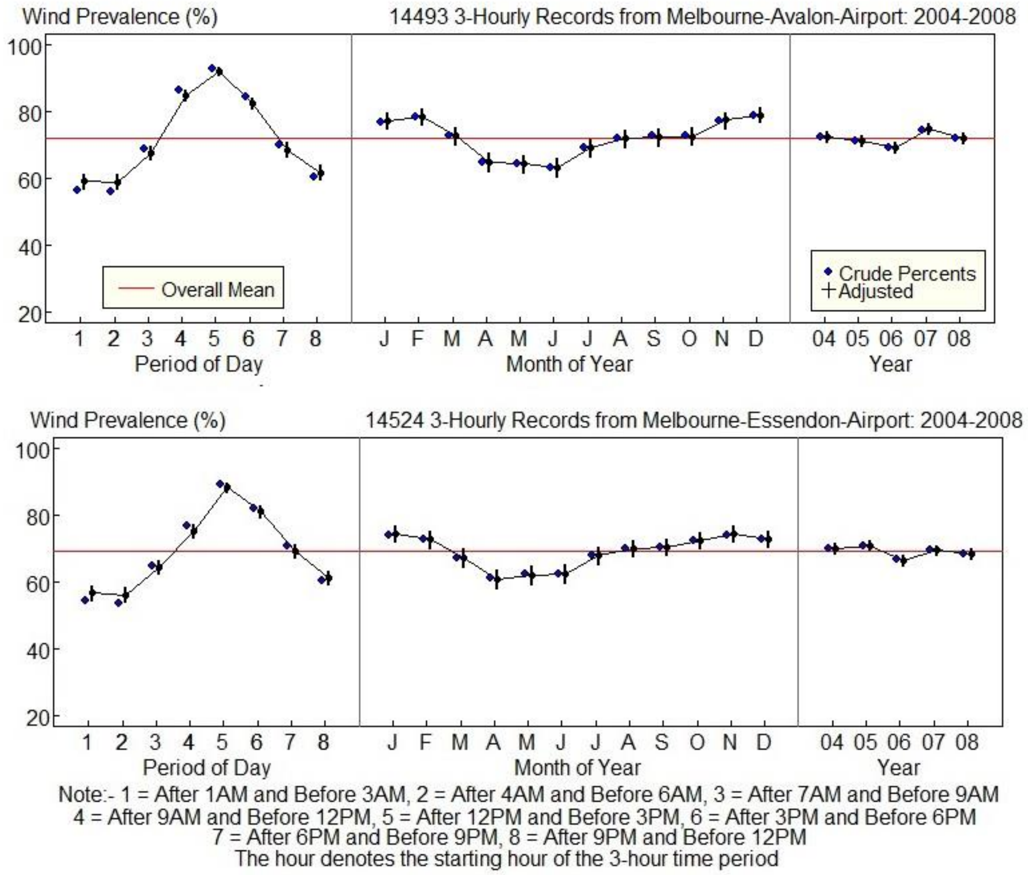
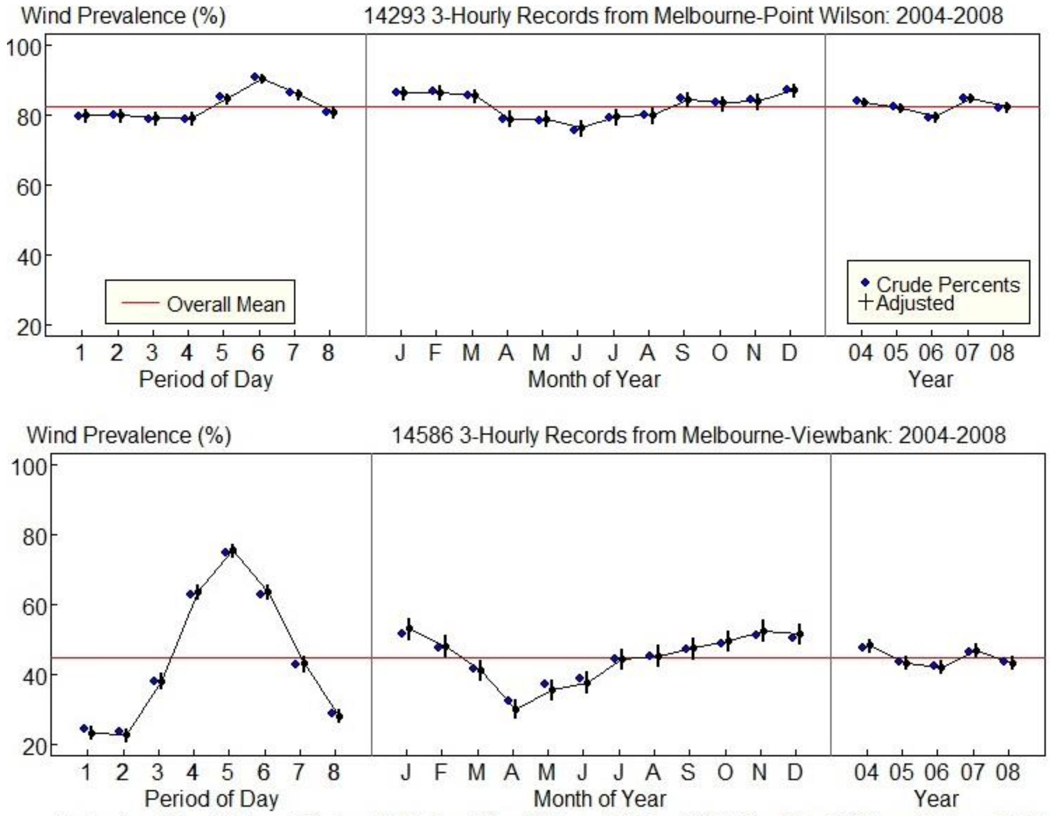


Figure 4.4 The 95% confidence interval for percentages of wind prevalence for Avalon Airport, Essendon Airport.

Figure 4.4 and Figure 4.5 reveals the pattern of wind prevalence percentage for the periods of the day, month and year and the crude percentage for the selected stations. The horizontal red line denotes the overall percentage of wind prevalence of each station. The 95% confidence interval plot of the wind prevalence revealed that the wind prevalence patterns for the period of the day, months and year were significantly different from the overall mean for all stations.

For the station of Avalon airport, Essendon airport and Viewbank, there are quite similar pattern of the wind prevalance. The highest prevalence was in period 5.

The monthty wind prevalance was below the average during March to June. The yearly wind prevalances were not different from the overall mean.



Note:- 1 = After 1AM and Before 3AM, 2 = After 4AM and Before 6AM, 3 = After 7AM and Before 9AM
4 = After 9AM and Before 12PM, 5 = After 12PM and Before 3PM, 6 = After 3PM and Before 6PM
7 = After 6PM and Before 9PM, 8 = After 9PM and Before 12PM
The hour denotes the starting hour of the 3-hour time period

Figure 4.5 The 95% confidence interval for percentages of wind prevalence for Point Wilson, Viewbank.

However, for the Point Wilson station, the overall percentage of wind prevalence was highest compared with other three stations (over 80%). The prevalence of wind was below the average from period 1 to period 4 and peak in period 6. Wind prevalence rates were observed to be less than the overall mean from April to August. It decreases steadily to attain its minimum (75%) in June and increases gradually to attain its maximum (85%) in December.

Wind gust patterns

Model (1) was also used to examine the 3 hourly periods of wind gust prevalence from 2004-2008. The 3-hourly mean wind gust prevalence model was also made of 3 predictors. ROC curve (Figure 4.6) revealed an area under the ROC curve for the selected stations the values of 0.40, 0.399, 0.117, and 0.458 respectively which illustrated classification of the occurrence and non-occurrence of wind. The curve revealed the overall accuracy of the classification on the occurrence and non-occurrence of wind gust was 95.4%, 63.3%, 74.6%, and 73.7%, having a true positive rate of 41%, 43%, 45%, and 49% with a false positive rate of 47.7%, 45.5%, 63.5%, and 22.4% respectively.

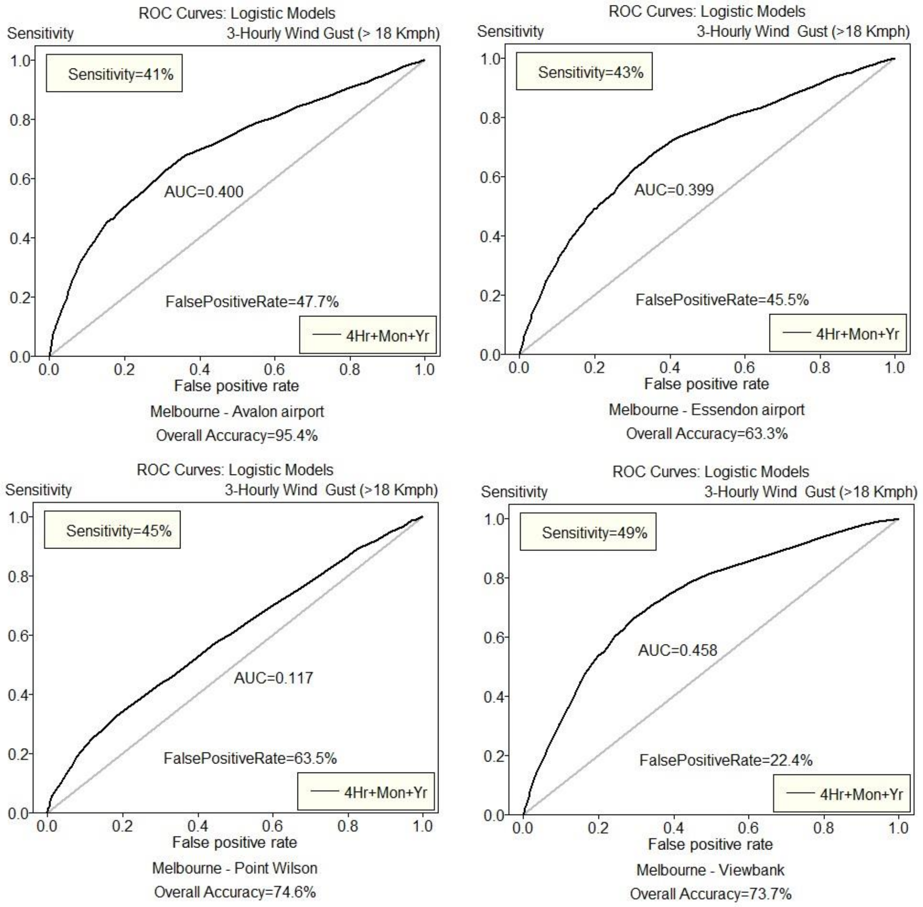
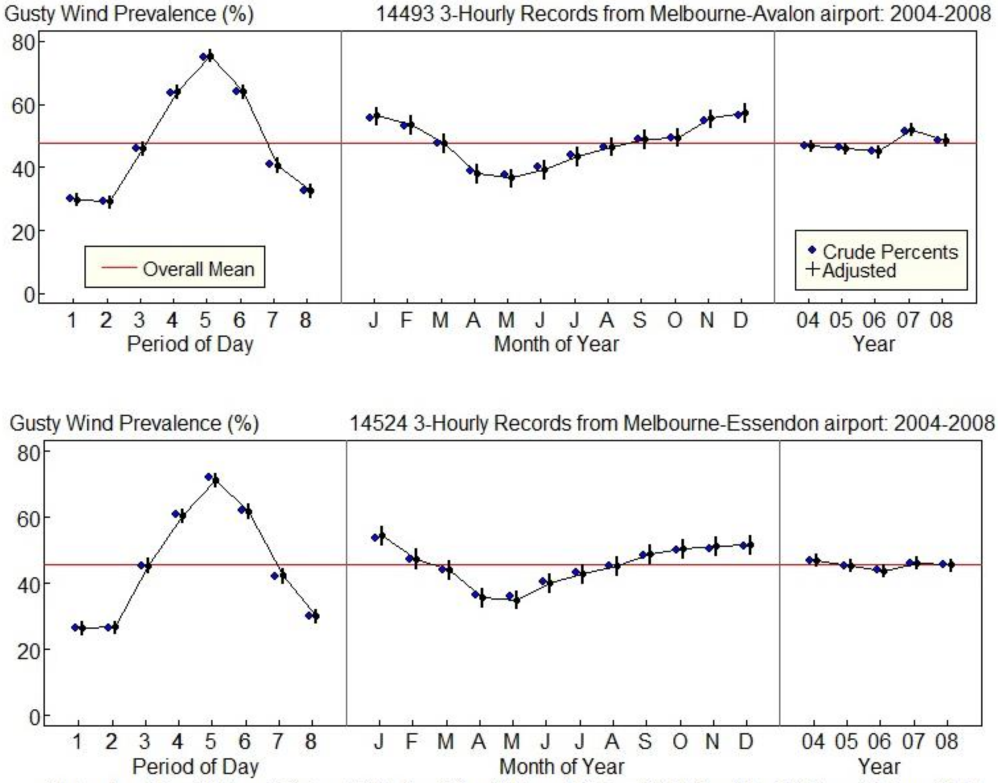


Figure 4.6 ROC curve of the logistic regression model for wind gust prevalence for Essendon Airport.

Figure 4.7 and Figure 4.8, reveals the pattern of gusty wind prevalence percentage for the periods of the day, month and year and the crude percentage for the selected stations. The horizontal line denotes the overall percentage of wind gust prevalence. The wind gust patterns for the period of the day were significantly different from the overall mean. The results were similar to the wind prevalence for all stations.

For the station of Avalon airport, Essendon airport and Viewbank, there were quite similar patterns of the gusty wind prevalance. The highest gusty wind prevalence was in period 5. The monthly gusty wind prevalance was below the average from March to June. The yearly gusty wind prevalences were not different from the overall mean.

For the Point Wilson station, the prevalance of gusty wind was below the average from period 1 to period 4 and peak in period 6. Wind gust prevalence rates were observed to be less than the overall mean from April to June. It decreases steadily to attain its minimum (55%) in May and increases gradually to attain its maximum (70%) in December. The yearly gusty wind prevalence patterns for the selected stations were not significantly different from the overall mean between 2004 and 2005.



Note:- 1 = After 1AM and Before 3AM, 2 = After 4AM and Before 6AM, 3 = After 7AM and Before 9AM
4 = After 9AM and Before 12PM, 5 = After 12PM and Before 3PM, 6 = After 3PM and Before 6PM
7 = After 6PM and Before 9PM, 8 = After 9PM and Before 12PM
The hour denotes the starting hour of the 3-hour time period

Figure 4.7 The 95% confidence interval for the percentage of wind gusts for Avalon Airport, Essendon Airport.

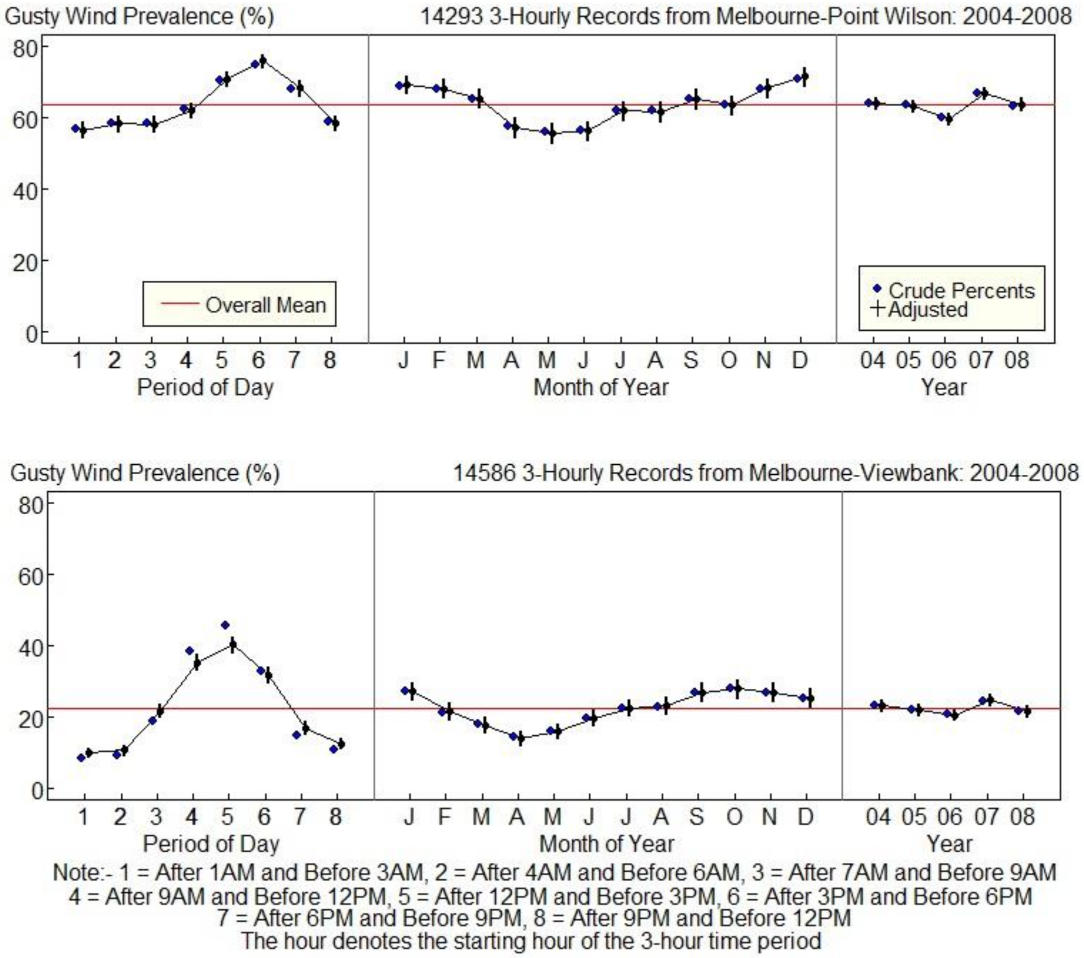


Figure 4.8 The 95% confidence interval for the percentage of wind gusts for Point Wilson, Viewbank.

Wind speed

The linear regression model (model 2) was fitted to the 3 hourly wind speed to examine the trends. The model for 3 hourly period wind speed from 2004 -2008 was also made of 3 predictors. There were 26 parameters, including the constant and AR(1) terms in the model. Most of the parameters were highly significant and influential in the model.

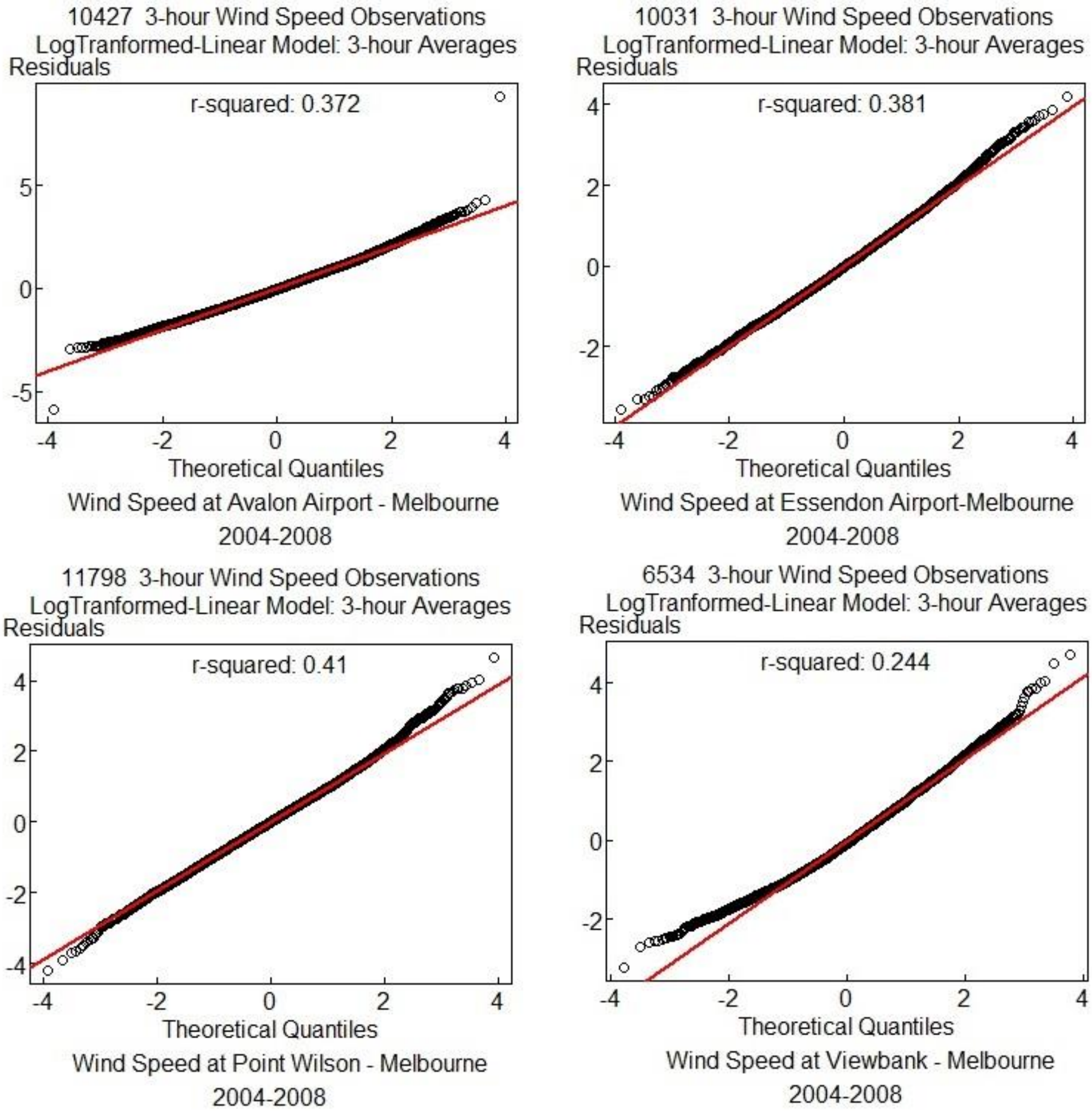


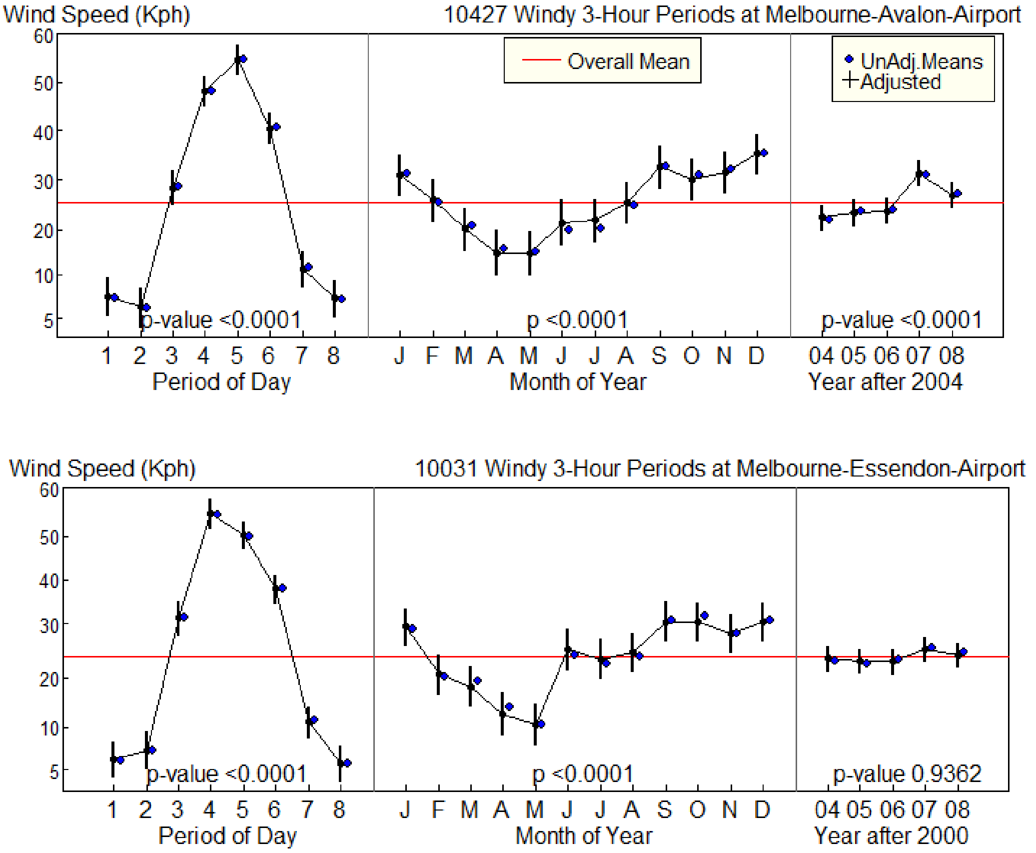
Figure 4.9 Q-Q plot for the linear model of the wind speed for Avalon airport, Essendon airport, Point Wilson, and Viewbank.

The coefficient of determination(r-squared) and the quantile-quantile plots was used to evaluate the goodness of fit of the model. The plot (Figure 4.9) revealed an r-squared value of 0.372, 0.381, 0.41, and 0.244 respectively shows that 37%, 38%, 41%, and 24% of the wind speed patterns are explained by the model (2). The quantile-quantile (Q-Q) plot illustrates that the residuals from the model are approximately normally distributed, which shows that the model fitted the data reasonably well. However, there were some deviations from the upper tail of the model. This deviation may be due to extreme values in the data and some variations that could not be explained by the predictors.

Figure 4.10 and Figure 4.11 reveals the wind speed trend for the periods of the day, month and year and the crude mean for the selected stations. The horizontal red line denotes the overall mean of wind speed. The 95% confidence interval plot of the wind speed revealed that the wind speed trends for the period of the day were significantly different from the overall mean.

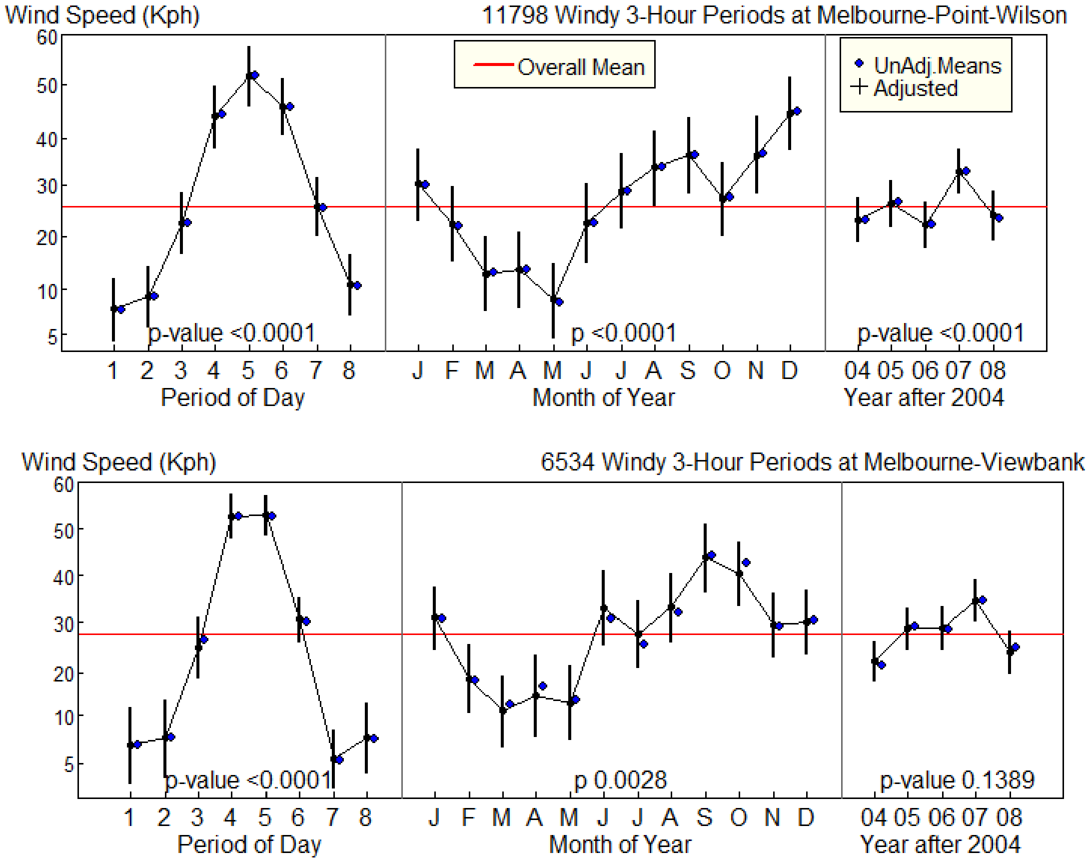
Four distinct trends were observed of the wind speed for selected stations, Avalon airport revealed significant increasing trends from period 2 to period 5. The declining trends were apparent from period 6 to period 8. The wind speed was higher than the overall mean from period 3 to 6 (7 am to 6 pm). The wind speed trends were observed to be higher than the overall mean from September to January. The yearly wind speed trend was mostly below the overall mean from 2004 to 2006.

Essendon airport revealed a highest wind speed in period 4. The declining trends were apparent from period 5 to period 8. The wind speed trends were observed to be higher than the overall mean from September to January. It decreases gradually to attain its minimum in May and increases steadily to attain its maximum in September. The yearly wind speed trend was not difference from the overall mean.



Note:- 1 = After 1AM and Before 3AM, 2 = After 4AM and Before 6AM, 3 = After 7AM and Before 9AM
4 = After 9AM and Before 12PM, 5 = After 12PM and Before 3PM, 6 = After 3PM and Before 6PM
7 = After 6PM and Before 9PM, 8 = After 9PM and Before 12PM
The hour denotes the starting hour of the 3-hour time period

Figure 4.10 The 95% confidence interval of wind speed trends for Avalon Airport, Essendon Airport.



Note:- 1 = After 1AM and Before 3AM, 2 = After 4AM and Before 6AM, 3 = After 7AM and Before 9AM
4 = After 9AM and Before 12PM, 5 = After 12PM and Before 3PM, 6 = After 3PM and Before 6PM
7 = After 6PM and Before 9PM, 8 = After 9PM and Before 12PM
The hour denotes the starting hour of the 3-hour time period

Figure 4.11 The 95% confidence interval of wind speed trends for Point Wilson, Viewbank.

Point Wilson revealed the wind speed was higher than the overall mean from period 4 to 6. The declining trends were apparent from period 6 to period 8. The wind speed trends were observed to be higher than the overall mean from July to January. It decreases gradually to attain its minimum in May and increases steadily to attain its maximum in December. The yearly wind speed trend was mostly below the overall mean but above the overall mean in 2007, where the maximum occurred. Viewbank revealed significant increasing trends with the highest speed during period 4 and 5. The highest wind speed in September. The yearly wind speed was high in 2007.

CHAPTER 5

CONCLUSIONS AND DISCUSSIONS

This chapter summarized the findings of the wind patterns and trends in India and Australia. Discussions of the findings, limitations, and recommendations are also presented.

5.1 Conclusions and discussions

Region of India

The statistical models were used to investigate the pattern of the 4-hourly wind observations during 2004-2008 in Calcutta. The logistic regression was used to model the wind prevalence and wind gust prevalence. The overall wind prevalence was observed to be 70% throughout the year. The period of the day factors revealed the wind prevalence above the overall mean between 5 am to 4 pm (period 2 and period 4). The observed monthly maximum wind prevalence of 90% in March and June and the minimum wind prevalence occurred between October and December. The yearly wind prevalence was mostly below the overall mean from 2004 to 2006 but above the overall mean between 2007 and 2008, where the maximum occurred (85%). The results of wind gust prevalence were similar to that of wind prevalence. This observed pattern due to stronger breeze during earlier periods of the morning to afternoon (Zhong 1992). The observed wind prevalence below the overall mean during the other periods of the day due to the lighter land breeze during the night (Mandal *et al*. 2013). The observed monthly maximum wind prevalence of 90% during March and June may be due to the high rates of wind such as breeze, storm during the first and second quarter of the year (Gadgil 2003, Rajeevan *et al*., 2012). However, the minimum wind prevalence occurred between October and December which may be as a result of the lighter rates of windblown such as squall on the surface of the Earth (Gadgil 2003, Rajeevan *et al*., 2012). The yearly wind prevalence was mostly below the overall mean from 2004 to 2006 but above the overall mean between 2007 and 2008, where the maximum occurred (85%). This pattern was partly because of stronger breeze and monsoon winds during the initial period of the decadal period from the year 2000, and lighter and weak monsoon winds from the final period of decade 2000 (Kumar *et al*., 2006).

Linear regression model for the wind speed revealed a maximum wind speed during 9 am to 12 pm (period 3), and the minimum wind speed during 5 pm to 4 am (period 5, 6 and 1). The monthly wind speed trend showed maximum wind speed in April, then decreasing trend from April to November (the minimum wind speed). The observed trend due to the more vigorous intensity of sea breeze and prevailing wind during early periods of the morning to afternoon and low intensity of land breeze and prevailing wind during the night (Chaudhuri *et al*., 2013). The monthly wind speed trend showed an increasing trend between January and April and reducing trend from April to November; these observed trends partly due to vigorous variation of wind speed near to the surface to the land (Torralba *et al*., 2017), global warming (Mclnnes *et al*., 2011) strong monsoon (Loo *et al*., 2015), and weaker monsoon (Vishnu *et al*., 2016).

The categories of wind velocity comprised of none/light, normal easterly, normal westerly, gusty westerly, and gusty easterly. The multinomial regression method used to describe the wind velocity categories pattern fits the data reasonably well. These patterns of wind velocity categories were due to discrete wind variability rates surrounding over Indian subcontinent landmass.

The prevalence of gusty westerly and gusty easterly patterns was investigated. The gusty westerly and easterly prevalence patterns for the periods of the day, month and year illustrated distinct, varied patterns. The overall gusty westerly and easterly prevalence were observed to be almost 5% all through the year. The periods of the day factors revealed that the gusty westerly prevalence was above the overall mean between 5 am to 4 pm and below the overall mean during the other periods of the day. These patterns of gusty westerly may be due to wind-induced ocean swell. The monthly maximum gusty westerly prevalence of 9% was observed during February and July. Also, the minimum gusty westerly prevalence occurred between August and December (1%) and may be due to global warming and ocean warming (Mojgan *et al*., 2017).

The yearly gusty westerly wind prevalence was mostly below the overall mean during most periods of 2005 to 2006 but was above the overall mean during 2004 while the maximum of 7% occurred between 2007 and 2008. This pattern was partly because of the monsoon and coastal currents (Rao *et al*., 2012).

Region of Australia

The 3-hourly wind patterns of four stations, Essendon airport, Avalon airport, Point Wilson, and Viewbank, in Australia during 2004-2008 were investigated. The logistic regression model fitted wind prevalence and wind gust pattern reasonably well. The period of the day factors revealed the wind prevalence above the overall mean between periods 4 and 6 for the observatories of Essendon airport, Avalon airport, and Viewbank, and for period 5 to period 8 in the observatory of Viewbank. For the observatories of Avalon airport, Essendon airport, and Point Wilson the observed monthly maximum wind prevalence of 95% during January and February, and the minimum wind prevalence occurred in July, April, and July. For these three stations, the yearly wind prevalence was not significantly different from the overall mean between 2004 and 2005 but above the overall mean between 2007 where the maximum occurred (40% to 85%). The results of the 95% confidence interval plot for the wind gust was similar to that of wind prevalence. These observed patterns were due to variability of the surface winds produced by the Asian Monsoon, seasonal variations in the westerly winds in both hemispheres (Hemer, 2010), and temporal atmospheric circulation variations (Troccoli *et al*., 2012) on the north and eastern Australian region.

Linear regression model of wind speed revealed that wind speed of Avalon airport and Essendon airport stations were above the average from 7 am to 6 pm. The wind speed trends were observed to be higher than the overall mean from September to January. The yearly wind speed trend was mostly below the overall mean from 2004 to 2006, but above the overall mean between 2007 and 2008, where the maximum occurred.

Point Wilson and Viewbank stations revealed the wind speed was higher than the overall mean from 9 am to 6 pm. The wind speed trends were observed to be higher than the overall mean from July to January. The yearly wind speed trend was mostly below the overall mean but above the overall mean in 2007where the maximum occurred.

These observed patterns were mainly due to significant variation of wind speed rates. Research by (Kossin *et al*., 2014) narrates the cyclonic intensity and their frequency changes. Maximum wind speeds are now occurring at higher latitudes than in the past; and strong westerly winds and storms encircle Antarctica (Kemp and Dear, 1976; Sturman and Tapper, 2006), and easterly equatorial trade winds (Hendon *et al*., 2014) over central Australia.

5.2 Limitations and recommendations

1. There are nine meteorological stations in India, but we could obtain consistent hourly wind speed data from only one station. Further studies should include all nine stations so that a more comprehensive estimate of the overall wind patterns in India can be derived.

2. This study modeled 4- and 3-hourly wind data using common climatic predictors but did not consider other variables that could affect wind patterns and trends such as air pressure, solar radiation, and humidity.

REFERENCES

Abhishek, A., Lee, J.Y., Keener, T.C. and Yang, Y.J. 2010. Long-term wind speed variations for three Midwestern US cities. Journal of the Air and Waste Management Association. 60(9), 1057-1064.

Ali, M. E.K., Hassan, M.Z., Ali, A.S. and Kumar, J. 2017. Prediction of Wind Speed Using Real Data: An Analysis of Statistical Machine Learning Techniques. In 2017 4th Asia-Pacific World Congress on Computer Science and Engineering (APWC on CSE) (pp. 259-264). IEEE.

Allen, J.T., Karoly, D.J. and Walsh, K.J. 2014. Future Australian severe thunderstorm environments. Part II: The influence of a strongly warming climate on convective environments. Journal of Climate, 27(10), pp.3848-3868.

Arzu, A., Kutty, S.S., Ahmed, M.R., and Khan, M.G. 2018. Wind speed forecasting using regression, time series and neural network models: a case study of Kiribati. In Australasian Fluid Mechanics Conference (pp. 1-4). Australasian Fluid Mechanics Society.

Ayodele, T.R., Adisa, A., Jimoh, J.L.M, and John T.A. 2012. Statistical analysis of wind speed and wind power potential of Port Elizabeth using Weibull parameters. Journal of Energy in Southern Africa 23(2), 30-38.

Barhmi, S., Elfatni, O. and Belhaj, I. 2019. Forecasting of wind speed using multiple linear regression and artificial neural networks. Energy Systems, 1-12.

BOM. 2011. Climate, Available at: http://www.bom.gov.au/jsp/ncc/climate_average wind-velocity/index.jsp [January 25, 2017].

Breckling, J. 2012. The analysis of directional time series: applications to wind speed and direction (Vol. 61), Springer Science & Business Media, Berlin, Germany.

Canning, S., Thiele, F. and Mitchell, M. 2010. Indigenous cultural heritage and history within the Metropolitan Melbourne Investigation Area. A report to the Victorian Environmental Assessment Council. Available at: http://www.veac.vic.gov.au/reports/Indigenous%20Cultural%20Heritage%20and% 20History% 20within% 20the% 20VE AC% 20Melbourne% 20Metropolitan% 20Investigation% 20Area. pdf

Chaudhuri, S. and Middey, A. 2013. Nowcasting lightning flash rate and peak wind gusts associated with severe thunderstorms using remotely sensed TRMM-LIS data. International Journal of Remote Sensing, 34(5): 1576-1590.

Cross, B.D., Kohfeld, K.E., Bailey, J. and Cooper, A.B. 2015. The Impacts of Wind Speed Trends and 30-Year Variability in Relation to Hydroelectric Reservoir Inflows on Wind Power in the Pacific Northwest. PloS one, 10(8), e0135730.

Dadaser-Celik., Filiz,. and Cengiz, E. 2014. Wind speed trends over Turkey from 1975 to 2006. International journal of climatology 34(6) 2: 1913-1927.

Elwan, A.A. and Habibuddin, H.M. 2019. Wind speed forecasting using First Order Markov Transition with Regime Switch and Time Duration. In IOP Conference Series: Earth and Environmental Science (Vol. 322, No. 1, p. 012016). IOP Publishing.

Friederichs, P., Göber, M., Bentzien, S., Lenz, A. and Krampitz, R. 2009. A probabilistic analysis of wind gusts using extreme value statistics. Meteorologische Zeitschrift. 18(6), 615-629.

Fujibe, F. 2009. Relation between long-term temperature and wind speed trends at surface observation stations in Japan. SOLA, 5, 81-84.

Gadgil, S. 2003. The Indian monsoon and its variability. Annual Review of Earth and Planetary Sciences, 31: 429-467.

Geer, I.W. 1996. Glossary of Weather and Climate : With Related Oceanic and Hydrologic Terms.

Gillings, R.J. 1972. Mathematics in the Time of the Pharaohs, Dover Publications Inc, New York, USA.

Glickman, T.S. and Zenk, W. 2000. Glossary of Meteorology, American Meteorological Society, Boston, USA.

Graham, A. and Kamm, A. 2014. The Romans: an introduction. Routledge.

Gururaja, P., McNeil, N. and Owusu, B.E. 2019a. Statistical Modelling of 4-hourly Wind Patterns in Calcutta, India. Nature Environment and Pollution Technology, 18(1), pp.73-80.

Gururaja, P, McNeil, N. and Owusu, B.E. 2019b. Statistical modelling of wind velocity in Calcutta, India. Ecology, Environment and Conservation, 25 (3): 1088-1095.

Hande, L.B., Siems, S.T. and Manton, M.J. 2012. Observed trends in wind speed over the Southern Ocean. Geophysical Research Letters, 39(11).

Hasson, A.E.A., Mills, G.A., Timbal, B. and Walsh, K. 2009. Assessing the impact of climate change on extreme fire weather events over southeastern Australia. Climate Research, 39(2), pp.159-172.

He, Y., Monahan, A.H., Jones, C.G., Dai, A., Biner, S., Caya, D. and Winger, K. 2010. Probability distributions of land surface wind speed over North America. Journal of Geophysical Research: Atmospheres, 115(D4).

Hemer, M.A. 2010. Historical trends in Southern Ocean storminess: Long-term variability of extreme wave heights at Cape Sorell, Tasmania. Geophysical Research Letters 37, no. 18.

Hendon, H.H., Lim, E.P. and Nguyen, H. 2014. Seasonal variations of subtropical precipitation associated with the southern annular mode. Journal of climate 27(9): 3446-3460.

Hossain, M.M. 2019. Probability Modeling of Monthly Maximum Sustained Wind Speed in Bangladesh. Statistics, Optimization & Information Computing, 7(1), 75-84.

IMD. 2015. Indian Meterogical department, Monsoon, Available at: http://www.imd.gov.in/pages/monsoon_main.php [February 5, 2017].

Jain, A., Srinivasan, M. and Hart, G.C. 2001. Performance-based design extreme wind loads on a tall building. Structural Design of Tall Buildings, 10(1): 9-26.

Jamaludin, A.R., Yusof, F., Kane, I.L. and Norrulasikin, S.M. 2016. A comparative study between conventional ARMA and Fourier ARMA in modeling and forecasting

wind speed data. In Advances in Industrial and Applied Mathematics: Proceedings of 23rd Malaysian National Symposium of

Mathematical Sciences (SKSM23) (Vol. 1750, p. 60022).

Jaswal, A.K. and Koppar, A.L. 2013. Climatology and trends in near-surface wind speed over India during 1961–2008. Mausam, 64(3), 417-436.

Kemp, P. and Dear I. 1976. The Oxford companion to ships and the sea. Vol. 627. Oxford University Press, London.

Kious, W.J. and Tilling, R.I. 1996. This Dynamic Earth: The Story of Plate Tectonics. DIANE Publishing, Darby, USA.

Kislov, A. and Matveeva, T. 2016. An extreme value analysis of wind speed over the European and Siberian Parts of Arctic Region. Atmospheric and Climate Sciences, 6(2), 205-223.

Kossin, J.P., Emanuel, K.A. and Vecchi, G.A. 2014. The poleward migration of the location of tropical cyclone maximum intensity. Nature 509, no. 7500: 349-352.

Kuleshov, Y., Fawcett, R., Qi, L., Trewin, B., Jones, D., McBride, J. and Ramsay, H. 2010. Trends in tropical cyclones in the South Indian Ocean and the South Pacific Ocean. Journal of Geophysical Research: Atmospheres, 115(D1).

Kulkarni, M.A., Patil, S., Rama, G.V. and Sen, P.N. 2008. Wind speed prediction using statistical regression and neural network. Journal of Earth System Science, 117(4), 457-463.

Kumar, K.K., Rajagopalan, B., Hoerling, M., Bates, G. and Cane, M. 2006. Unraveling the mystery of Indian monsoon failure during El Niño. Science, 314(5796): 115-119.

Lakshmanan, N., Gomathinayagam, S., Harikrishna, P., Abraham, A. and Ganapathi, S. C. 2009. Basic wind speed map of India with long-term hourly wind data. Current science. 96(7). 911-922.

Letson, F., Barthelmie, R.J., Hu, W. and Pryor, S.C. 2019. Characterizing wind gusts in complex terrain. Atmospheric Chemistry and Physics, 19(6), 3797-3819.

Lollchund, R.M., Boojhawon, R. and Rughooputh, S.D. 2014. Statistical modelling of wind speed data for Mauritius. International Journal of Renewable Energy Research. 4(4), 1056-1064.

Loo, Y.Y., Billa, L. and Singh, A. 2015. Effect of climate change on seasonal monsoon in Asia and its impact on the variability of monsoon rainfall in Southeast Asia.
Geoscience Frontiers, 6(6): 817-823.

Lydia, M., Kumar, S.S., Selvakumar, A.I. and Kumar, G. E. P. 2016. Linear and non-linear autoregressive models for short-term wind speed forecasting. Energy Conversion and Management, 112, 115-124.

Maklad, Y. and Glencross-Grant, R. 2014. A seasonal analysis of potential wind power for Armidale NSW, Australia. International Journal of Energy Economics and Policy 4.1, 92-107.

Mandal, S., Choudhury, B.U., Mondal, M. and Bej, S. 2013. Trend analysis of weather
variables in Sagar Island, West Bengal, India: a long-term perspective (1982 2010).

Mclnnes, K.L., Erwin, T.A. and Bathols. J.M. 2011. Global Climate Model projected changes in 10 m wind speed and direction due to anthropogenic climate change. Atmospheric Science Letters, 12(4): 325-333.

McVicar, T.R., Thomas, G.V.N., Ling, T.L., Michael, L.R., David, P.R., Lucrezia, R. and Randall, J.D. 2008. Wind speed climatology and trends for Australia, 1975–2006: Capturing the stilling phenomenon and comparison with near-surface reanalysis output. Geophysical Research Letters 35 (20).

Mojgan, G.M., Mehdi, M.M. and Reza, B.M. 2017. The Trend of Changes in Surface Wind in the Indian Ocean, in the Period from 1981 to 2015, Using Reanalysis Data, NCEP/NCAR. Open Journal of Marine Science, 7(4): 445-457.

Morgan, E.C., Lackner, M., Vogel, R.M. and Baise, L.G. 2011. Probability distributions for offshore wind speeds. Energy Conversion and Management, 52(1), 15-26.

Mukhopadhyay, P., Mahakur, M. and Singh, H.A. 2009. The interaction of large scale and mesoscale environment leading to formation of intense thunderstorms over Kolkata Part I: Doppler radar and satellite observations. Journal of Earth System Science, 118(5), 441-466.

Pang, G., He, J., Huang, Y. and Zhang, L. 2019. A Binary Logistic Regression Model for Severe Convective Weather with Numerical Model Data. Advances in Meteorology, 2019.

Petre, I.A., Rebenciuc, M. and Ciucu, S.C. 2016. The use of Markov chains in forecasting wind speed: Matlab source code and applied case study. Computational Methods in Social Sciences, 4(2), 44.

Prema, V. and Rao, K.U. 2015. Time series decomposition model for accurate wind speed forecast. Renewables: Wind, Water, and Solar 2.1,18.

Quantick, H.R. 2001. Climatology for Airline, Blackwell Science Alden press, Oxford, UK.

R Development Core Team. 2008. A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria.

Rajeevan, M., Unnikrishnan, C.K., Bhate, J., Kumar, K.N. and Sreekala, P.P. 2012. Northeast monsoon over India: variability and prediction. Meteorological Applications, 19(2): 226-236.

Rao, A.D., Mahapatra, D.K., Babu, S.V. and Pandey, S. 2012. Spatial-temporal variation and mechanism associated with mini-cold pool off the southern tip of India during summer and winter monsoon season. Water Quality Research Journal. 47 (3-4) : 333-352.

Reddy, S.J. 2008. Climate change: myths and realities, JCT, Hyderabad, India.

Rozas-Larraondo, P., Inza, I., and Lozano, J.A. 2014. A method for wind speed forecasting in airports based on nonparametric regression. Weather and Forecasting, 29(6), 1332-1342.

Satari, S.Z., Zubairi, Y.Z., Hussin, A.G. and Hassan, S.F. 2015. Some statistical characteristic of Malaysian wind direction recorded at maximum wind speed: 1999–2008. Sains Malaysiana, 44(10), 1521-1530.

Schindler, D., Bauhus, J. and Mayer, H. 2012. Wind effects on trees. European Journal of Forest Research, 131(1): 159-163.

Singh, A., Gurtej, K., Jain, G., Nayyar, F. and Tripathi, M.M. 2016. Short term wind speed and power forecasting in Indian and UK wind power farms. In 2016 IEEE 7th Power India International Conference (PIICON) (pp. 1-5). IEEE.

Sreenivasa, S.C., Agarwal, S.K. and Kumar, R. 2014. Short term wind forecasting using logistic regression driven hypothesis in artificial neural network. In 2014 6th IEEE Power India International Conference (PIICON) (pp. 1-6). IEEE.

Srivastava, N. 2005. A logistic regression model for predicting the occurrence of intense geomagnetic storms.

Sturman, A.P. and Tapper, N.J. 2006. The weather and climate of New Zealand and Australia. Oxford University Press, Melbourne.

Tagliaferri, F., Hayes, B.P., Viola, I.M. and Djokić, S.Z. 2016. Wind modelling with nested Markov chains. Journal of Wind Engineering and Industrial Aerodynamics, 157, 118-124.

Tian, H., Banger, K., Bo, T. and Dadhwal, V.K. 2014. History of land use in India during 1880–2010: Large-scale land transformations reconstructed from satellite data and historical archives. Global and Planetary change, 121, 78-88.

Tongkumchum, P. and McNeil, D. 2009. Confidence intervals using contrasts for regression model. Songklanakarin Journal of Science and Technology. 31(2), 151-156.

Torralba, V., Doblas-Reyes, F.J. and Gonzalez-Reviriego, N. 2017. Uncertainty in recent

near-surface wind speed trends: a global reanalysis intercomparison. Environmental Research Letters, 12(11): 114019.

Torres Silva, D.S.A., Silva, S. and Moisés, C. 2013. Seasonality, interannual variability, and linear tendency of wind speeds in the Northeast Brazil from 1986 to 2011. The Scientific World Journal, 2013.

Troccoli, A., Muller, K., Coppin, P., Davy, R., Russell, C. and Hirsch, A.L. 2012. Long term wind speed trends over Australia. Journal of Climate. 25(1), 170-183.

UNDP. 2006. Hazard profile of Indian districts, Available at: https://web.archive.org/web/20060519100611/http://www.undp.org.in/dmweb/hazardprofile.pdf [August 23, 2006].

Usbeck, T., Wohlgemuth, T., Pfister, C., Volz, R., Beniston, M. and Dobbertin, M. 2010. Wind speed measurements and forest damage in Canton Zurich (Central Europe) from 1891 to winter 2007. International Journal of Climatology. 30(3), 347-358.

van Smeden, M., de Groot, J.A., Nikolakopoulos, S., Bertens, L.C.M., Moons, K.G.M. and Reitsma, J.B. 2017. A generic nomogram for multinomial prediction models: theory and guidance for construction. Diagnostic and prognostic research. 1: 8

Venables, W.N. and Ripley, B.D. 2002. Modern Applied Statistics with S Fourth edition.

Vishnu, S., Francis, P.A., Shenoi, S.S.C. and Ramakrishna, S.S.V.S. 2016. On the decreasing trend of the number of monsoon depressions in the Bay of Bengal. Environmental Research Letters, 11(1): 14011.

Water, F.L. 2005. The Beaufort Wind Scale.

Westin, L.K. 2001. Receiver operating characteristic (ROC) analysis: evaluating discriminance effects among decision support systems. UMINF report, Department of Computer Science, Umea University, Sweden: Available from: http://www.cs.umu.se/research/ reports/2001/018/part1.pdf [December, 08 2017]

Wiegand, P. 2004. Oxford International Student Atlas, Oxford University Press, Chennai, India.

Wheeler, D. and Wilkinson C. 2004. From calm to storm: the origins of the Beaufort wind scale. The Mariner's Mirror 90, 2. 187-201.

Wright, M.A. and Grab, S.W. 2017. Wind speed characteristics and implications for wind power generation: Cape regions, South Africa. South African Journal of Science, 113(7-8), 1-8.

Yan, Z., Steven, B., Richard, E.C., Valerie I. and Howard, W. 2002. An analysis of daily maximum wind speed in northwestern Europe using generalized linear models. Journal of Climate 15, 15, 2073-2088.

Young, I.R., Vinoth, J., Zieger, S. and Babanin, A.V. 2012. Investigation of trends in extreme value wave height and wind speed. Journal of Geophysical Research: Oceans, 117(C11).

Zhang, H., Yu, Y.J. and Liu, Z.Y. 2014. Study on the Maximum Entropy Principle applied to the annual wind speed probability distribution: A case study for observations of intertidal zone anemometer towers of Rudong in East China Sea. Applied Energy. 114, 931-938.

Zhong, S. and Takle, E.S. 1992. An observational study of sea- and land-breeze circulation in an area of complex coastal heating. Journal of Applied Meteorology, 31(12): 1426-1438.

APPENDIX A

Summary of the 3-hourly wind data for all stations.

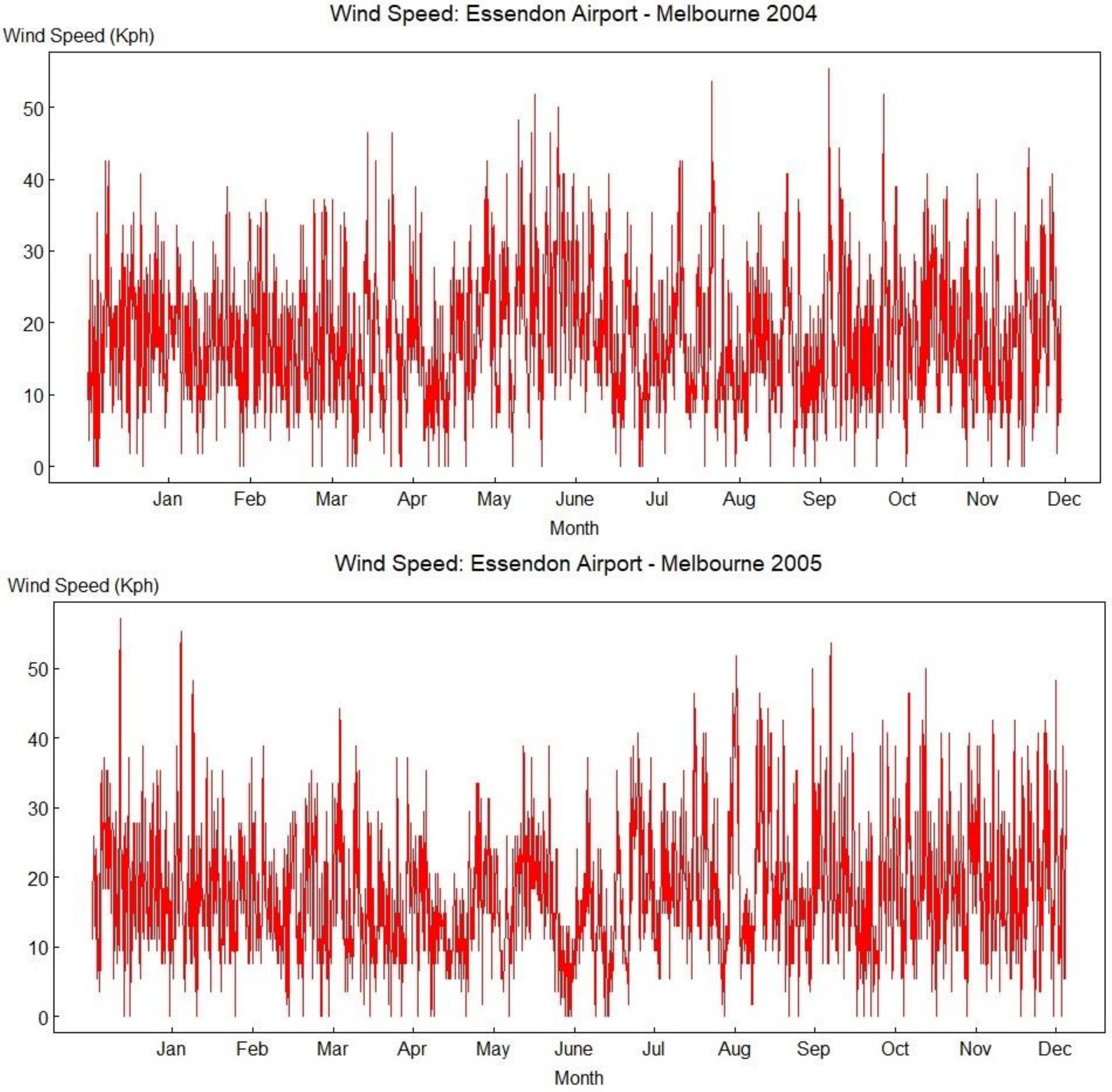


Figure 1 The 3-hourly episode-based time-series graphs of wind speed for the station of Essendon airport for the years 2004 and 2005.

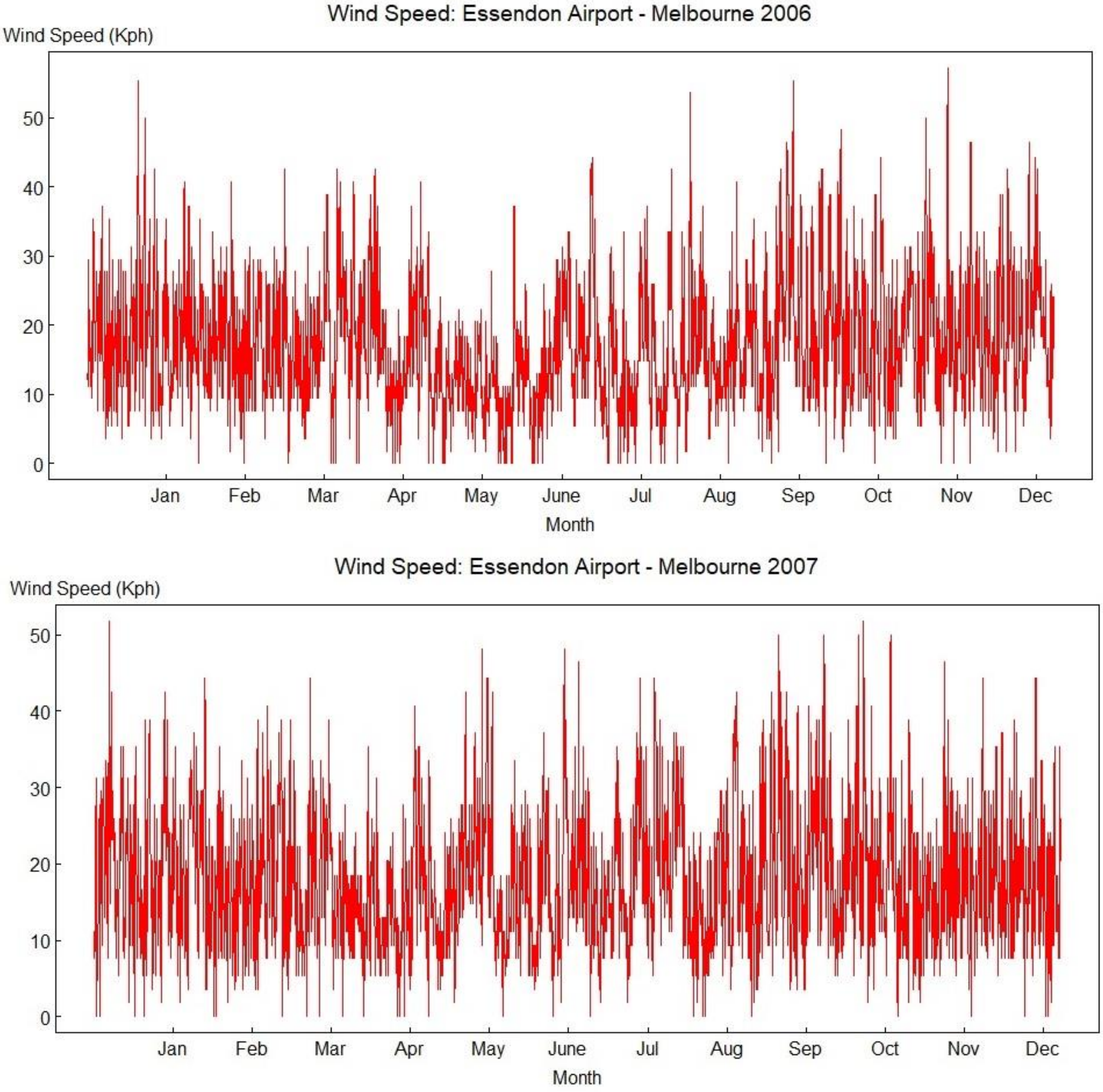


Figure 2 The 3-hourly episode-based time-series graphs of wind speed for the station of Essendon airport for the years 2006 and 2007.

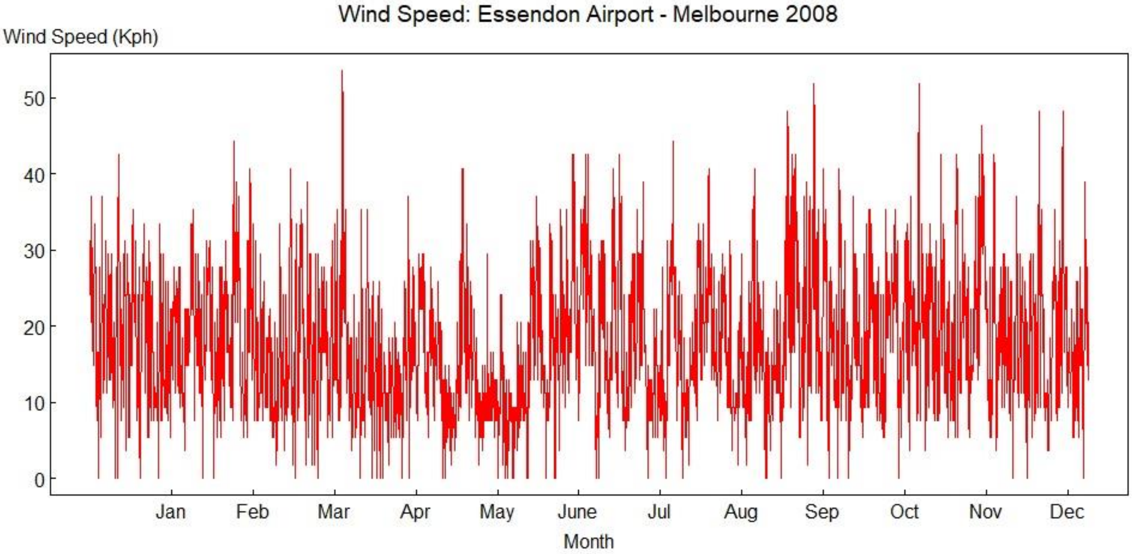


Figure 3 The 3-hourly episode-based time-series graphs of wind speed for the station of Essendon airport for the year 2008.

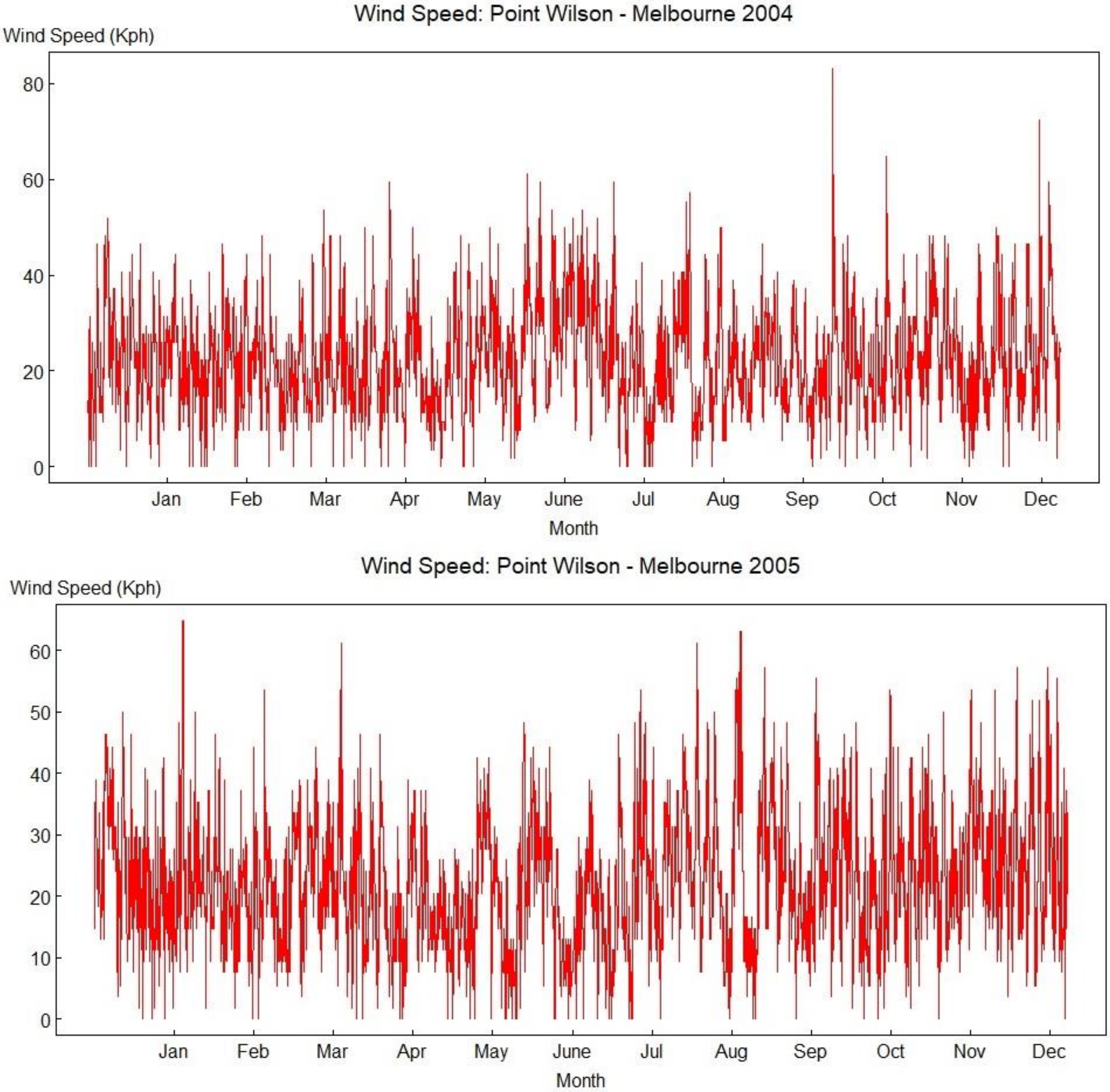
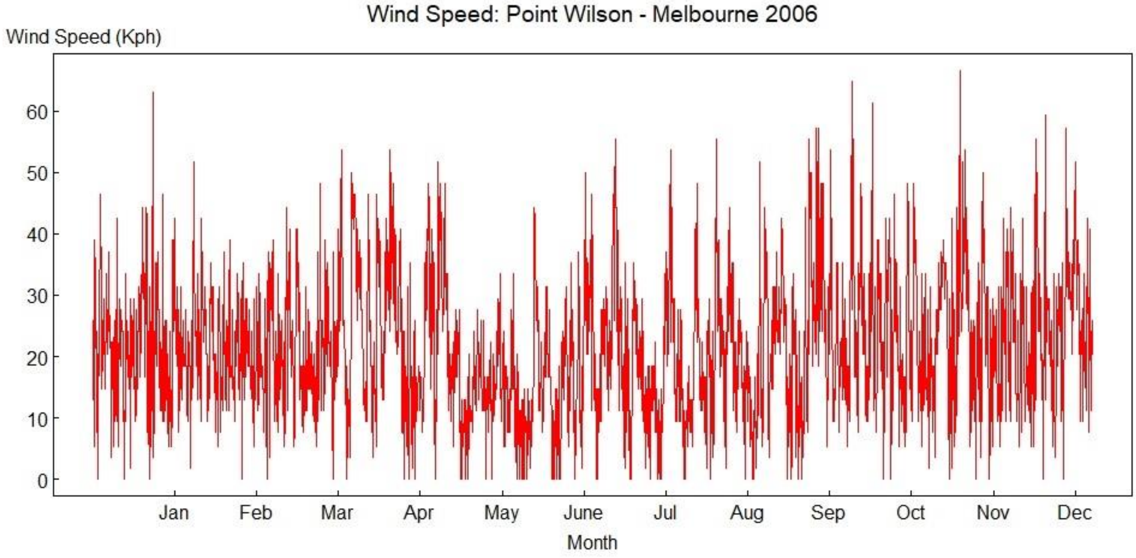


Figure 4 The 3-hourly episode-based time-series graphs of wind speed for the station of Point Wilson for the years 2004 and 2005.



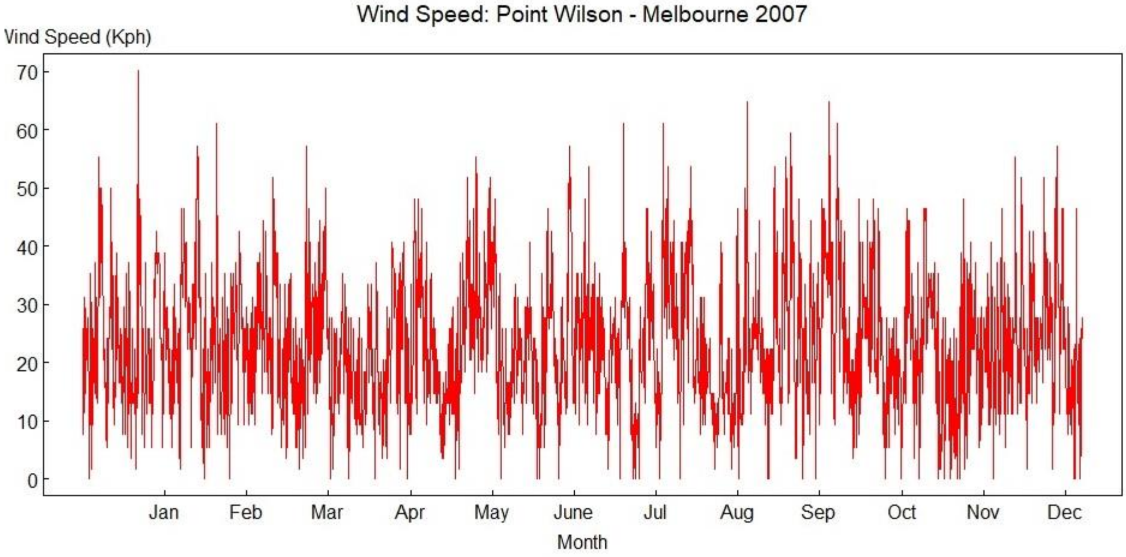


Figure 5 The 3-hourly episode-based time-series graphs of wind speed for the station of Point Wilson for the year 2006 and 2007.

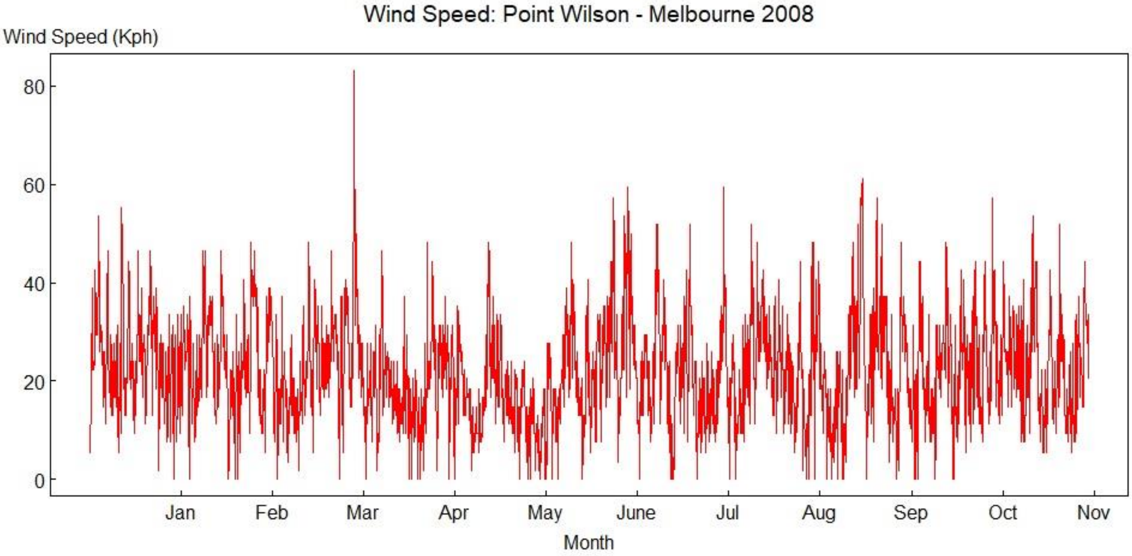


Figure 6 The 3-hourly episode-based time-series graphs of wind speed for the station of Point Wilson for the year 2008.

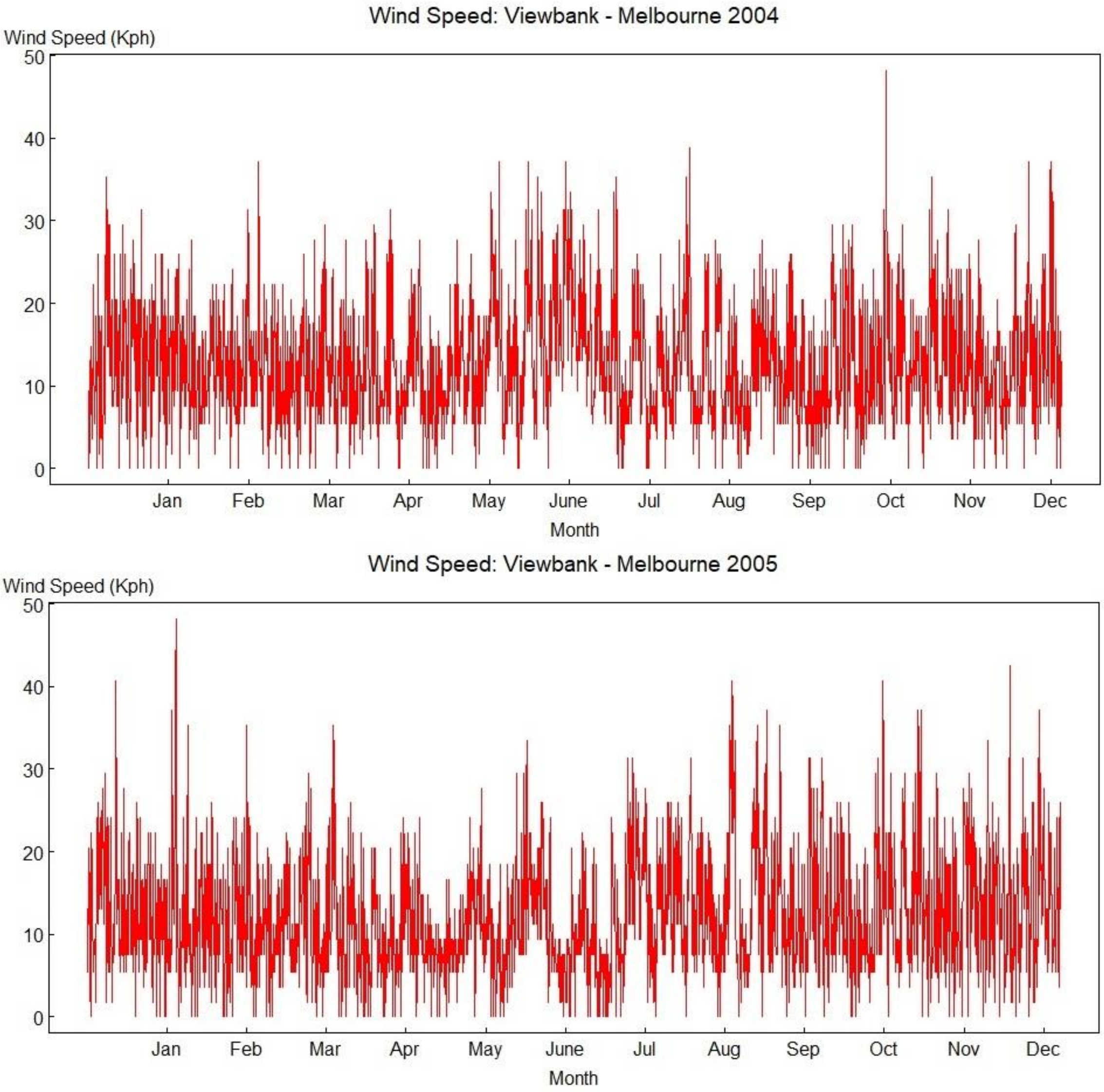


Figure 7 The 3-hourly episode-based time-series graphs of wind speed for the station of Viewbank for the years 2004 and 2005.

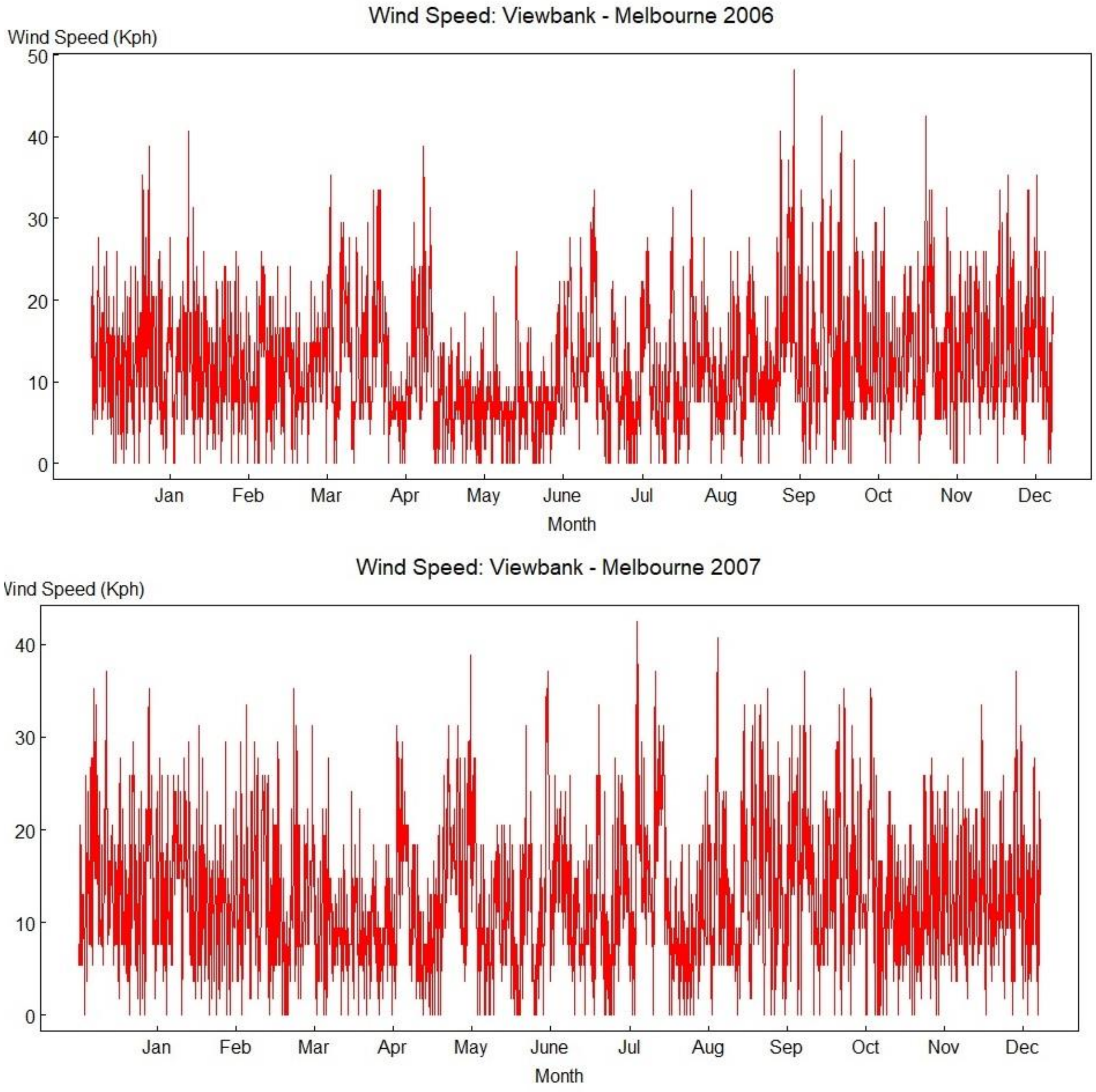


Figure 8 The 3-hourly episode-based time-series graphs of wind speed for the station of Viewbank for the year 2006 and 2007.

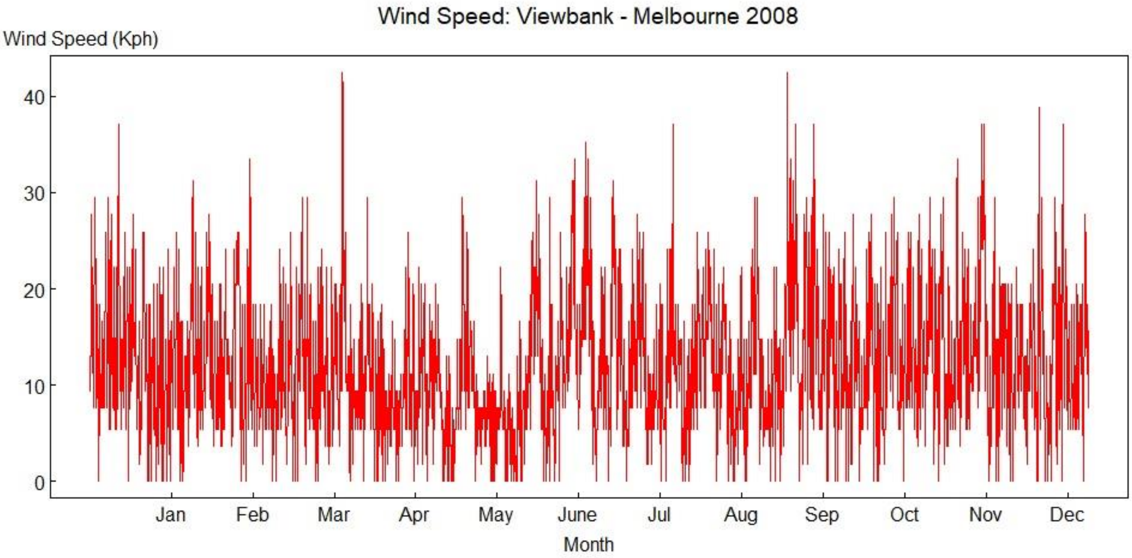


Figure 9: The 3-hourly episode-based time-series graphs of wind speed for the station of Viewbank for the year 2008.

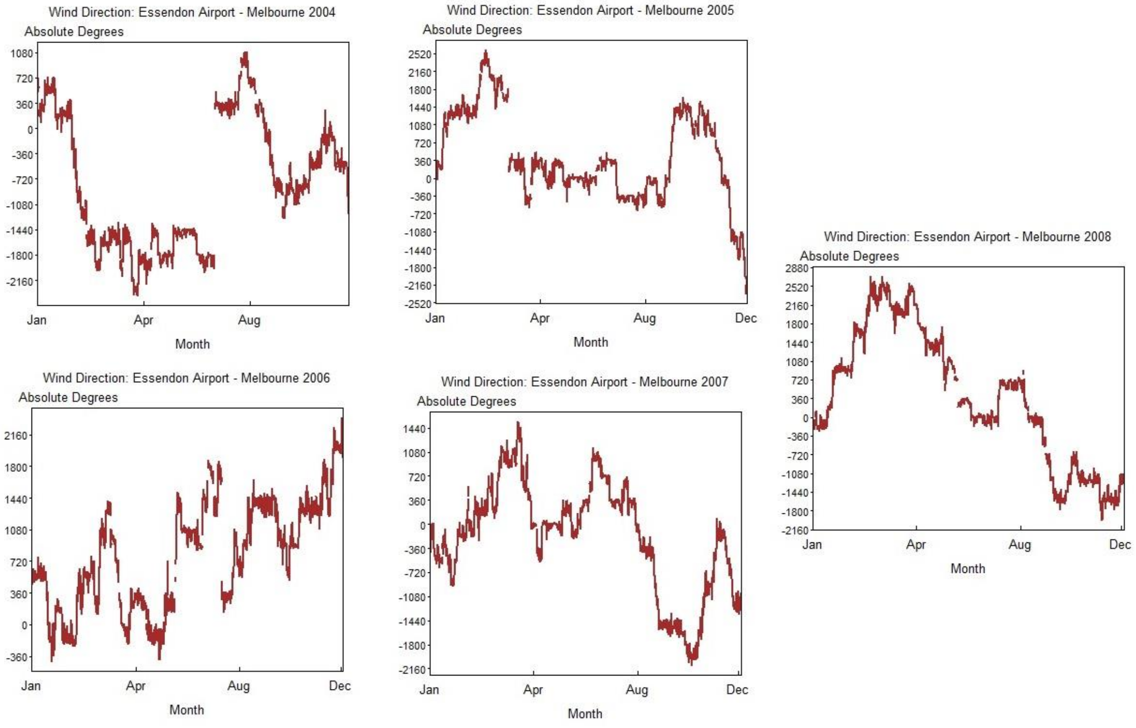


Figure 10 The 3-hourly episode-based time-series graphs of wind direction in absolute degrees for the station of Essendon airport for the year 2004 to 2008.

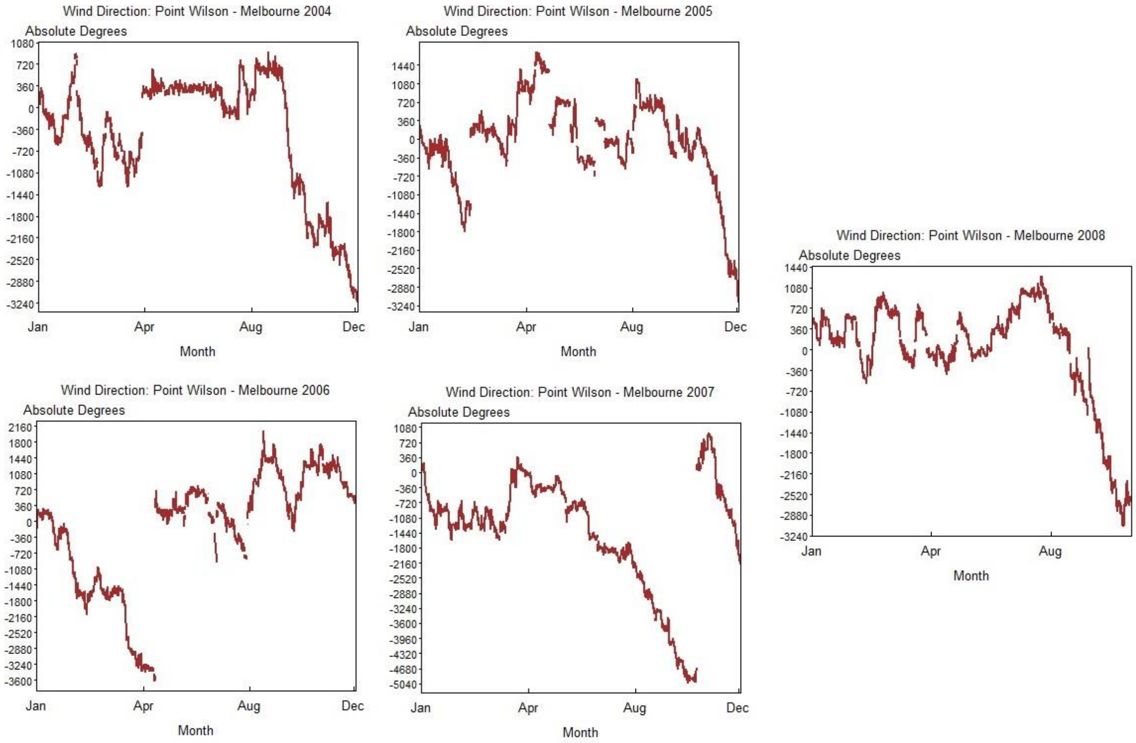


Figure 11 The 3-hourly episode-based time-series graphs of wind direction in absolute degrees for the station of Point Wilson for the year 2004 to 2008.

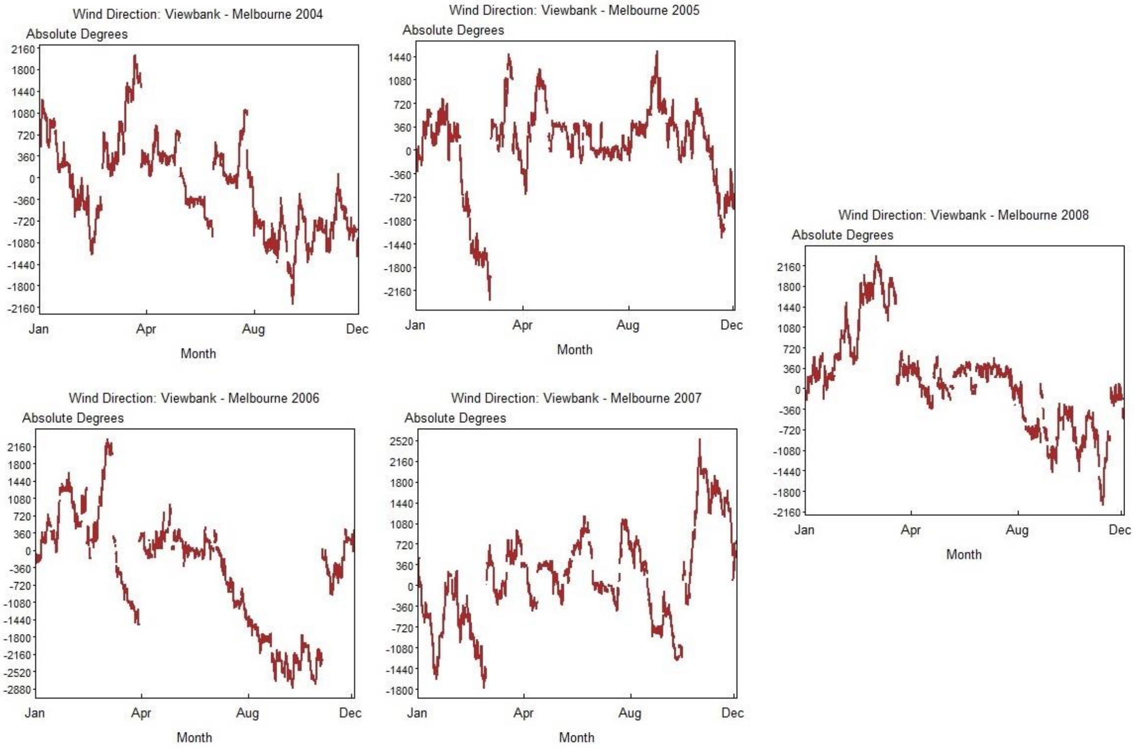


Figure 12 The 3-hourly episode-based time-series graphs of wind direction in absolute degrees for the station of Viewbank for the year 2004 to 2008.

APPENDIX B

Article I: Statistical Modelling of 4-hourly Wind Patterns in Calcutta, India.

Nature Environment and Pollution Technology *An International Quarterly Scientific Journal*	p-ISSN: 0972-6268 e-ISSN: 2395-3454	Vol. 18	No. 1	pp.73-80	2019

Original Research Paper **Open Access**

Statistical Modelling of 4-hourly Wind Patterns in Calcutta, India

Prashanth Gururaja*, Nittaya McNeil*(**)† and Bright Emmanuel Owusu*(***)
*Department of Mathematics and Computer Science, Faculty of Science and Technology, Prince of Songkla University, Mueang Pattani, 94000, Thailand
**Centre of Excellence in Mathematics, Commission on Higher Education (CHE), Ministry of Education, Ratchathewi, Bangkok, 10400, Thailand
***Department of Information and Communication Technology/Mathematics, Faculty of Science and Technology, Presbyterian University College, Ghana
†Corresponding author: Nittaya McNeil

Nat. Env. & Poll. Tech.
Website: www.neptjournal.com

Received: 05-05-2018
Accepted: 02-08-2018

Key Words:
Wind prevalence
Wind patterns
Logistic and linear regression
Speed and gust

ABSTRACT

The rising energy demands of the world and minimal availability of conservative energy sources have considerably increased the role of non-conventional sources of energy like solar and wind. The modelling of the prevalence of wind speed and trends helps in estimating the energy produced from wind farms. This study uses statistical models to analyse the wind patterns in India. Hourly wind data during 2004-2008 were obtained from the National Renewable Energy Laboratory for the study. A logistic regression model was initially used to investigate 4-hourly wind prevalence and the pattern of wind gust. A linear regression model was then applied to investigate wind speed trends. The 4-hour periods of the day, months, and year factors of wind were used as the independent variables in the statistical analysis. The results showed that wind prevalence was mostly higher between 4 AM to 4 PM within the day (90%). Analysis of the monthly wind prevalence revealed higher percentage between April to June (90%), while higher annual prevalence was revealed in 2007 and 2008 (85%). Wind speed trends (4-hourly periods) was observed to increasing from 4 AM to 4 PM within which the maximum occurred. Monthly wind speed was seen to be increasing from January to April where it attains the maximum (13 ms^{-1}) and reduces it to minimum in November (3 ms^{-1}). Annual mean wind speed has reduced by 4 ms^{-1} between 2004 and 2008.

INTRODUCTION

Wind plays a pivotal role in climate and weather on Earth's surface. The movement of wind on Earth's surface has diverse ranges of magnitudes, different directions and various time intervals from minute to few hours. Wind circulation on the Earth's surface commonly known as wind prevalence describes the motion of wind with internment momentarily rates (Greene et al. 2010). The wind speed intriguing behaviour has a vast impact on global climate and weather conditions (Csavina et al. 2014) nurturing of ecosystems, and human life imbalance (Mitchell 2012). Wind gusts rapid and internment speeds have an enormous influence on climate and weather events (Cheng et al. 2014). In addition, its extreme can cause enormous damages or destruction to human-made structures, devastation to humans and ecosystems, and dangerous situations in aircrafts (Jain et al. 2001, Schindler et al. 2012, Jamaludin et al. 2016). It plays a significant factor in wind characteristics often combined and denoted as wind velocity. Further, six major wind belts, three each in the northern and southern hemisphere, encircle the surface of the Earth. These belts lie from pole to equator namely polar easterlies, westerlies, and trade winds describe the wind pattern on the Earth's surface (Myrl 2012).

The variant behaviours of wind cause a difference in absorption of solar energy among the climate zones of earth and atmospheric circulation by differential heating between equator and poles (Chapin et al. 2002). Therefore, wind with its volatile behaviour can be used productively to harness energy, operate transport and warfare equipment. Thus, statistical analysis of wind characteristics will be beneficial.

In recent decades, many studies have investigated wind patterns over the Earth's surface. A study by Klink (2002) investigated trends and interannual variability of wind speed distributions in Minnesota. The research results showed reducing wind speed trend below the overall mean for most of the stations. However, only few stations showed increasing trends above the overall mean. Nchaba et al. (2016) studied long-term austral summer wind speed trends over southern Africa. The results revealed evidence of a decline in wind speeds, results in deceleration of mid latitude westerly, and Atlantic southeasterly winds with a poleward shift in the subtropical anticyclone. Increasing trends in the annual frequency of summer circulation weather types, and the weakening of the subtropical continental heat. Another study by Klink (1999) studied on trends in mean monthly maximum and minimum surface wind speeds in the coterminous United States, 1961 to 1990. The results showed reducing trends in the maximum and minimum winds observed in spring and summer reduced. The decreasing wind speed occurred in western and southeastern United States may be due to variable topography and high pressure mostly all through the year. However, central and the northeastern United States having gentle topography and are near common storm tracks had the highest maximum and minimum winds. Also, a study by Bett et al. (2013) investigated European wind variability over 140 years. The results revealed strong highly-variable wind in the north-east Atlantic. Also, findings do not show any clear long-term trends in the wind speeds across Europe and the variability between decades is large. However, the year 2010 had the lowest mean wind speed on record for this region. Mohanan et al. (2011) investigated the probability distribution of land surface wind speeds. The results illustrated daytime surface wind speeds to be broadly consistent, whereas night time surface wind speeds are more positively skewed. However, in the mid latitudes, these strongly positive skewnesses are shown to be associated with conditions of strong surface stability and weak lower-tropospheric wind shear. Moreover, research by Troccoli et al. (2012) studied long-term wind speed trends over Australia. The results showed that light winds tend to increase more rapidly than the mean winds, whereas strong winds increase less rapidly than the mean winds. The trends in both light and strong winds vary in line with the mean winds. Chen et al. (2013) investigated wind speed trends over China, quantifying the magnitude and assessing causality. The results exhibited pronounced downward trends especially in the upper percentiles and during spring. The warm and cold Arctic Oscillation and El Niño-Southern Oscillation phases have significant influence on the probability distribution of wind speeds. Thus, internal climate variability is a major source of both interannual and long-term variability.

Reddy et al. (2015) used statistical analysis to estimate wind energy over Gadanki, India from 2007 to 2012 using Weibull, Rayleigh and Gamma parametric methods. They showed that Weibull distribution fitted well. A study by Gupta & Biswas (2010) analysed wind velocity of Silchar (Assam, India) from 2003-2007 by applying Rayleigh's and Weibull methods. The results revealed that average wind velocity in Silchar is about 3.11 kmh^{-1}, which is considerably low. Lakshmanan et al. (2009) studied on the basic wind speed map of India with long-term hourly wind data from 1983-2005 using Gumbel method. The study results showed that certain regions require improvement to higher wind zones from the current benchmark wind zones. They also suggested the need to revised basic wind speed map with upgraded wind speed zones for India. Even though many aspects of wind in India have been analysed, not much has been on the description of wind prevalence rates, wind speed trends and wind gust patterns. The objectives of this study are to describe the wind and wind gust prevalences and analyse the wind speed patterns using statistical models.

MATERIALS AND METHODS

India has nine meteorological stations specified by National Renewable Energy Laboratory (NREL). Hourly wind observations from Calcutta during 2004-2008 were obtained from NREL (https://www.nrel.gov/international/ra_india.html) for the study. This station was selected due to its subtropical climate with moist (rain-bearing) winds and frequent thunderstorms (Mukhopadhyay et al. 2009). The winds can vary from frigid winds, may be related to large storm systems such as Nor'westers and Western disturbances (Reddy 2008). It is also in wind and cyclone zone of very 'high damage risk' (UNDP 2006).

Data were arranged based on the Earth's surface wind direction patterns for Indian subcontinent (Wiegand 2004, Kious & Tilling 1996, IMD 2015). The data, which did not fall under the wind direction for the particular month, were discarded from the study resulting in 38,157 observations during the period of study. The hourly wind occurrence had fluctuations that were smoothened by accumulating data to 4 hour periods, and these were used for the analysis. The measurements recorded for consecutive 4-hourly periods were associated as per the well-known ancient time metrics of Babylon and Egyptian system which can be represented in 4-hourly periods of 6 intervals of a day (Gillings 1972). Moreover, 6 periods of 4 hours each of the day classified into clear time intervals like mid-night, dawn, morning, noon, mid-noon, night (Glickman & Zenk 2000) which is indicated as a period of the day, and other predictors are months and year. Observations on February 29 were discarded to have uniform number of observations for each year and also associated with heuristic 365 day climate data observations, therefore totalling to 9,540 observations for the study.

STATISTICAL METHODS

Logistic regression was initially fitted to model the prevalence of 4-hourly wind. The 4-hourly wind speed less than 0.775 ms^{-1} was kept as the non-occurrence of wind, and at least 0.775 ms^{-1} was considered as the occurrence of wind.

In modelling the prevalence of 4-hourly wind, if p_i is the probability of wind for the ith period in the data set, restricted on the variables x_i, then the model takes the form:

$$\ln\left(\frac{p_i}{1-p_i}\right) = \alpha + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_m x_m, \qquad \ldots(1)$$

Where, α is a constant, $\beta_i\,(i = 1, 2, 3, \ldots, m)$ are the slope coefficients of the model and $x_i\,(i = 1, 2, 3, \ldots, m)$ are the predictors and m is the number of predictors. The predictors are the 4-hourly periods, month and year factors.

Model (1) was also used to estimate the prevalence of wind gust. Wind speed of at least 5 ms^{-1} (Geer 1996) was classified as gust occurrence and classified as the non-occurrence of a gust if otherwise. The modelling process of estimating the wind gust occurrence is similar to that of the wind prevalence. The predictors are also the 4-hourly periods, month and year factors. These models were assessed using the receiver operating characteristic (ROC) curve.

The ROC curve is popularly known for determining the capability of prediction a binary outcome (Westin 2001). The ROC curve plots sensitivity against the false positive rate of the 4-hourly wind prevalence.

Multiple regression was then used to describe the trends of wind speed of at least 0.775 ms^{-1}. The model takes the form

$$y_{ijk} = \alpha + \beta_1 x_i + \beta_2 x_j + \beta_3 x_k + \gamma_{t-1}, \qquad \ldots(2)$$

Where, y_{ijk} represents wind speed trends, α is a constant, β's are the regression coefficients, x_i, x_j, x_k is time indicating the predictors 4-hour periods, month and year, γ denotes the regression coefficients of the trimmed lag 1 term and t-1 indicates lag 1 period. The goodness of fit of the model was assessed using the coefficient of determination (R-square) and the Q-Q plots.

Further, since the normality assumption was not satisfied as the data were not normally distributed, due to heavy tails; the wind speed values were transformed by square roots, to satisfy the statistical assumption of normality. Subsequently, the wind speed observation anomalies were filtered by removing the auto correlations at lag 1 term (Chatfield 1996). The filtered autocorrelations at lag 1 term were then added to fit linear regression model (Venables & Ripley 2002) that investigated the wind speed trends.

The sum contrasts (Tongkumchum & McNeil 2009) were applied to obtain 95% confidence intervals (CI) to compare the fitted model means with the overall wind means. This contrast gives criteria to classify levels of the factor into three groups, according to whether each relating CI is greater than or equal to, or is less than the overall mean. All analysis was done in R (R Development Core Team 2008)

RESULTS

A logistic regression model was fitted to 4-hourly wind periods for 2004-2008.

The model for the 4-hourly mean of wind prevalence was made of 3 predictors. The first predictor was the 4-hourly period of the day and had 6 factors, the second predictor was the months with 12 factors, and the third factor was the year having 5 factors. There were 24 parameters including the constant term in the model. The parameters were highly significant and influential in the model.

Assessing the goodness of fit of the model by the ROC curve (Fig. 1) revealed an area under ROC curve (AUC) value of 0.75 which show a sound classification of the prevalence and non-occurrence of wind.

The overall accuracy of the classification on the occurrence and non-occurrence of wind was 85% having a true positive rate of 99.7% with a false positive rate of 61.5%.

The plot of confidence intervals illustrates the pattern of wind prevalence percentage for period of the day, month and year and the crude percentage (Fig. 2). The horizontal red line denotes the overall mean of the wind prevalence, which is approximately 70%.

Analysis of the 95% confidence interval plots revealed that the wind prevalence patterns for period of the day were significantly different from the overall mean.

The period from dawn (period 2) to afternoon (period 4) observed high wind prevalence rates with the maximum

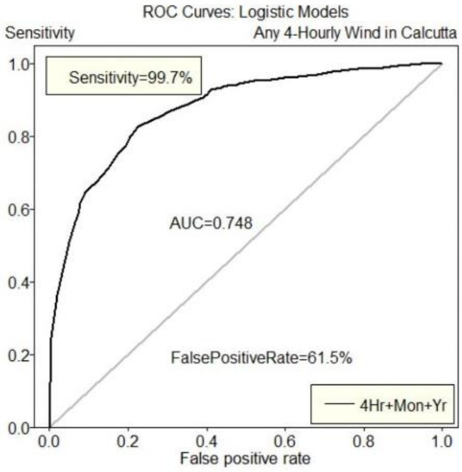


Fig. 1: ROC curve of the logistic model for 4-hour period of wind to investigate the prevalence of wind.

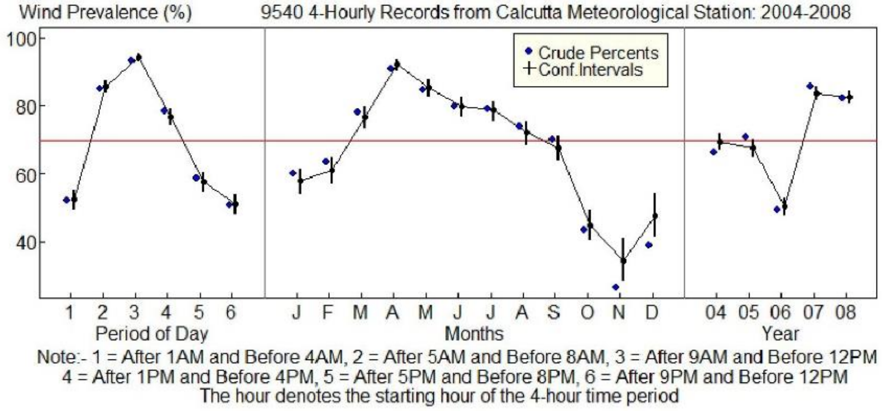


Fig. 2: Logistic model confidence interval plot for percentages of wind prevalence.

prevalence of around 90% observed in noon (period 3). The low wind prevalence rates were observed from midnight to dawn (period 1), evening (period 5) tonight (period 6) with the minimum prevalence of around 50 % in the night (period 6) respectively.

The monthly wind prevalence patterns were observed to have high wind prevalence rates from March to July with maximum wind prevalence of 90% observed in April. Moderate wind prevalence rate was observed during August and September with modest wind prevalence of 70% aligned with the overall mean. Wind prevalence rates were observed to be less than the overall mean during January and February. It increases gradually to attain its maximum (90%) in April, decreases steadily, attain its minimum (30%) in November.

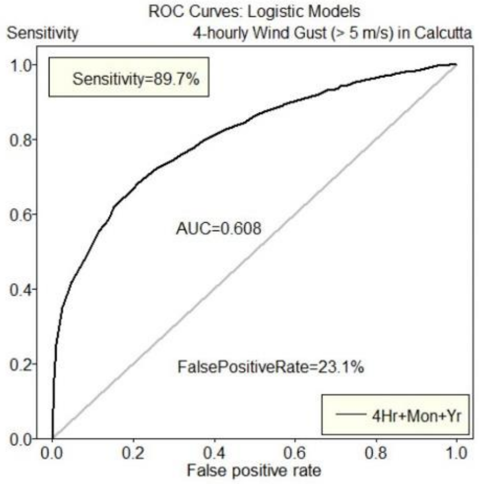


Fig. 3: ROC curve of the logistic regression model for the 4-hourly wind gust prevalence.

The yearly wind prevalence patterns were not significantly different from the overall mean between 2004 and 2005. It reduced considerably to its minimum in 2006 and rose steadily to its maximum in 2007, where it began to reduce slightly. The yearly wind prevalence ranged from 50 to 85% and observed to be higher than the overall mean between 2007 and 2008.

Model (1) was also used to examine the 4-hourly period wind gust prevalence from 2004-2008. The 4-hourly mean wind gust prevalence model was also made of 3 predictors.

Assessing the goodness of fit of the model by the ROC curve (Fig. 3) revealed an area under ROC curve (AUC) value of 0.61 which show classification of the occurrence and non-occurrence of wind. The overall accuracy of the classification on the occurrence and non-occurrence of wind gust was 81.67% having a true positive rate of 89.7% with a false positive rate of 23.7%.

In Fig. 4, the horizontal line denotes the overall mean of the prevalence of wind gusts, which is around 10 % during the study period. The wind gust patterns for period of the day were significantly different from the overall mean. The period from midnight (period 1) to noon (period 3) observed increasing wind gust prevalence rates with the maximum prevalence of around 18% observed in noon (period 3). It then decreases steadily from this period to the night. The minimum wind gust prevalence of 4% occurred in midnight (period 1).

The monthly wind gust patterns revealed increasing pattern from January to April, where the maximum was attained. It was constant between April and decreases steadily until August, where it increased again until September. It then decreases sharply to minimum in November. The wind

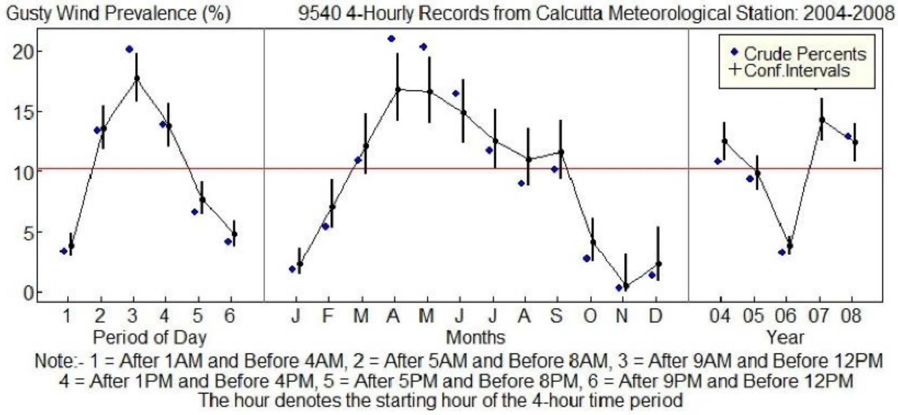


Fig. 4: The 95% confidence interval plot for the percentage of wind gusts using logistic regression.

gust prevalence was observed to be mostly above the overall mean between March and September.

The yearly wind gust prevalence revealed decreasing pattern from 2004 to 2006, where it attained the minimum of 4% and increased sharply to its maximum of 14% in 2007. There was a decreasing pattern between 2007 and 2008.

The linear regression model (model 2) was fitted to the 4-hourly wind speed to examine the trends. The model for 4-hourly period wind speed from 2004 -2008 was also made of 3 predictors. There were 25 parameters including the constant and AR(1) terms in the model. Most of the parameters were highly significant and influential in the model. The coefficient of determination (r-squared) and the quantile-quantile plots was used to evaluate the goodness of fit of the model. The r-squared value of 0.432 shows that 43% of the wind speed patterns are explained by the model (2).

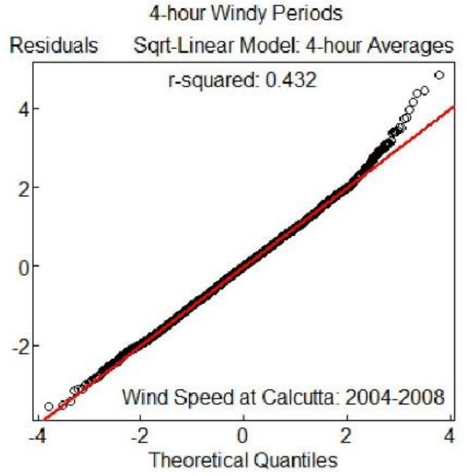


Fig. 5: Q-Q plots for the linear model of the wind speed.

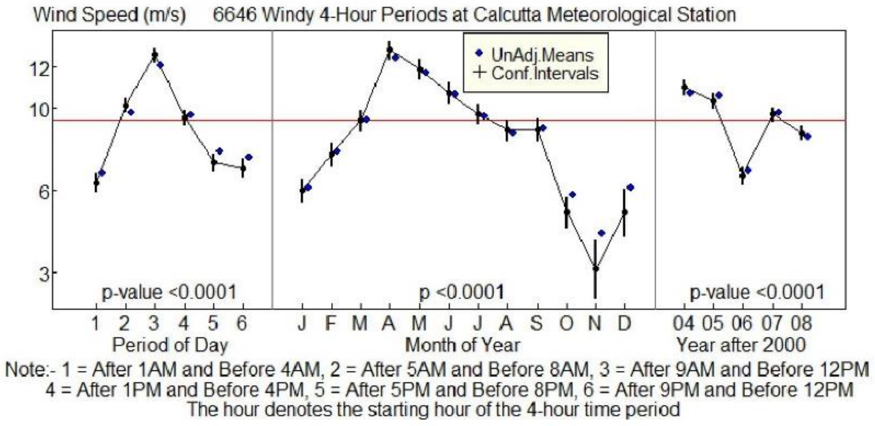


Fig. 6: Linear model confidence interval plot of wind speed trends.

The quantile-quantile (Q-Q) plot illustrates that the residuals from the model are approximately normally distributed (Fig. 5) which shows that the model fitted the data reasonably well. However, there were some deviations from the upper tail of the model. This deviation may be due to extreme values in the data and some variations that could not be explained by the predictors.

Fig. 6, the horizontal line denotes the overall mean of the wind speed is approximately 9 ms^{-1}. The wind speed trends for the period of the day were significantly different from the overall mean. The period during dawn (period 2) to noon (period 3) had the maximum wind speed of around 13 ms^{-1} observed during noon (period 3). Moderate wind speed trend was observed during the afternoon (period 4) with a modest wind speed of approximately 9 ms^{-1} aligning with overall mean. The low wind speed trends were observed during midnight (period 1) to dawn (period 2), evening (period 5) to night (period 6) with a minimum wind speed of around 6 ms^{-1} in midnight (period 1) respectively.

The monthly wind speed trends revealed rising trends during April to June with a maximum wind speed of around 13 ms^{-1} observed during April. Moderate wind speed trend was observed during March and July to September with a modest wind speed of approximately 9 ms^{-1} aligning with overall mean. The low wind speed trends were observed during January to February, and October to December with a minimum wind speed of around 3 ms^{-1} in November. The monthly and yearly wind speed trends revealed values either greater, equal to or lower than the overall mean. The yearly wind speed revealed increasing trends during 2004 and 2005 with a maximum wind speed of around 11 ms^{-1} observed during 2004. Moderate wind speed was observed during 2007 with a modest wind speed of approximately 9 ms^{-1} aligning with overall mean. The reducing trends were observed during 2006 and 2008 with a minimum wind speed of around 7 ms^{-1} in 2006.

CONCLUSION AND DISCUSSION

Statistical models were used to investigate the 4-hourly wind observations of Calcutta, India during 2004-2008. The wind and wind gust prevalence patterns together with the wind speed trends were investigated. The wind prevalence pattern, wind speed trend and wind gust prevalence for the period of the day, month and year illustrated distinct divergent patterns. Logistic regression model fitted wind prevalence and wind gust pattern reasonably well. The overall wind prevalence was observed to be 70% all through the year. The period of the day factors revealed the wind prevalence above the overall mean between period 2 and period 4. This pattern may be due to stronger breeze during early morning to afternoon (Zhong 1992). The observed wind prevalence below the overall mean during the other periods of the day may be due to lighter land breeze during the night (Mandal et al. 2013). The observed monthly maximum wind prevalence of 90% during March and June may be due to the high rates of wind such as breeze, storm during the first and second quarter of the year (Gadgil 2003, Rajeevan et al. 2012). However, the minimum wind prevalence occurred between October and December which may be as a result of the lighter rates of wind blown such as squall on the surface of the Earth (Gadgil 2003, Rajeevan et al. 2012). The yearly wind prevalence was mostly below the overall mean during 2004 to 2006, but above the overall mean between 2007 and 2008 where the maximum occurred (85%). This pattern was partly because of stronger breeze and monsoon winds during the initial period of the decadal period from the year 2000, and lighter and weak monsoon winds from the final period of decade 2000 (Kumar et al. 2006). The results of the 95% confidence interval plot for the wind gust was similar to that of wind prevalence.

Linear regression model revealed an increasing trend for the period of the day factors period 1 to period 3 and decreasing trend between period 3 and period 6. The observed trend may be due to the stronger intensity of sea breeze and prevailing wind during early morning to afternoon and low intensity of land breeze and prevailing wind during the night (Chaudhuri et al. 2013). The monthly wind speed trend showed increasing trend between January and April and reducing trend from April to November. These trends observed partly due to vigorous variation of wind speed near to the surface of the land (Torralba et al. 2017), global warming (McInnes et al. 2011) strong monsoon (Loo et al. 2015), and weaker monsoon (Vishnu et al. 2016). The yearly wind speed trend displayed reducing trend between 2004 and 2006 and increasing trend between 2007 and 2008. This observed trends may be due to the effect of global warming and ocean warming on monsoon patterns (Mojgan et al. 2017).

This study has provided a description of wind trends and patterns, which helps in obtaining descriptive and knowledgeable information of wind circulation over Calcutta, India. These trends are vital in the development of wind resources distribution in several areas of India. Future studies on wind could include data on dewpoint, wind chill and heat index along with different statistical methods.

ACKNOWLEDGMENT

The work has been supported by the Higher Education Research Promotion and Thailand's Education Hub for Southern Region of ASEAN Countries Project Office of the Higher

Education Commission under grant no. THE/033. We thank Emeritus Prof. Dr. Don McNeil for supervising this work. Thanks to Centre of Excellence in Mathematics, Commission on Higher Education, Thailand.

REFERENCES

Bett, P.E., Thornton, H.E. and Clark, R.T. 2013. European wind variability over 140 yr. Advances in Science and Research, 10(1): 51-58.

Chapin, F.S., Matson III, P.A. and Mooney, H. 2002. Principles of Terrestrial Ecosystem Ecology. Springer-Verlag, New York, New York, USA.

Chatfield, C. 1996. The Analysis of Time Series. Chapman and Hall: Melbourne.

Chaudhuri, S. and Middey, A. 2013. Nowcasting lightning flash rate and peak wind gusts associated with severe thunderstorms using remotely sensed TRMM-LIS data. International Journal of Remote Sensing, 34(5): 1576-1590.

Chen, L., Li, D. and Pryor, S.C. 2013. Wind speed trends over China: quantifying the magnitude and assessing causality. International Journal of Climatology, 33(11): 2579-2590.

Cheng, C.S., Lopes, E., Fu, C. and Huang, Z. 2014. Possible impacts of climate change on wind gusts under downscaled future climate conditions: Updated for Canada. Journal of Climate, 27(3): 1255-1270.

Csavina, J., Field, J., Félix, O., Corral-Avitia, A.Y., Sáez, A.E. and Betterton, E.A. 2014. Effect of wind speed and relative humidity on atmospheric dust concentrations in semi-arid climates. Science of the Total Environment, 487(1): 82-90.

Gadgil, S. 2003. The Indian monsoon and its variability. Annual Review of Earth and Planetary Sciences, 31: 429-467.

Geer, I.W. 1996. Glossary of Weather and Climate: With Related Oceanic and Hydrologic Terms. Boston, Mass: American Meteorological Society, pp. 272.

Greene, S., Morrissey, M. and Johnson, S.E. 2010. Wind climatology, climate change, and wind energy. Geography Compass, 4(11): 1592-1605.

Gillings, R.J. 1972. Mathematics in the Time of the Pharaohs. Dover Publications Inc, New York, USA.

Glickman, T.S. and Zenk, W. 2000. Glossary of Meteorology. American Meteorological Society, Boston, USA.

Gupta, R. and Biswas, A. 2010. Wind data analysis of Silchar (Assam, India) by Rayleigh's and Weibull methods. Journal of Mechanical Engineering Research, 2(1): 010-024.

IMD 2015. Indian Meteorological Department, Monsoon. Available at http://www.imd.gov.in/pages/monsoon main.php [February 5, 2017].

Jain, A., Srinivasan, M. and Hart, G.C. 2001. Performance based design extreme wind loads on a tall building. Structural Design of Tall Buildings, 10(1): 9-26.

Jamaludin, A.R., Yusof, F., Kane, I.L. and Norrulasikin, S.M. 2016. A comparative study between conventional ARMA and Fourier ARMA in modeling and forecasting wind speed data. In: Advances in Industrial and Applied Mathematics, Proceedings of 23rd Malaysian National Symposium of Mathematical Sciences (SKSM23), 1750: 60022.

Kious, W.J. and Tilling, R.I. 1996. This Dynamic Earth: The Story of Plate Tectonics. DIANE Publishing, Darby, USA.

Klink, K. 2002. Trends and interannual variability of wind speed distributions in Minnesota. Journal of Climate, 15(22): 3311-3317.

Klink, K. 1999. Trends in mean monthly maximum and minimum surface wind speeds in the coterminous United States, 1961 to 1990. Climate Research, 13(3): 193-205.

Kumar, K.K., Rajagopalan, B., Hoerling, M., Bates, G. and Cane, M. 2006. Unraveling the mystery of Indian monsoon failure during El Niño. Science, 314(5796): 115-119.

Lakshmanan, N., Gomathinayagam, S., Harikrishna, P., Abraham, A. and Ganapathi, S.C. 2009. Basic wind speed map of India with long-term hourly wind data. Current Science, 96(7): 911-922.

Loo, Y.Y., Billa, L. and Singh, A. 2015. Effect of climate change on seasonal monsoon in Asia and its impact on the variability of monsoon rainfall in Southeast Asia. Geoscience Frontiers, 6(6): 817-823.

Mandal, S., Choudhury, B.U., Mondal, M. and Bej, S. 2013. Trend analysis of weather variables in Sagar Island, West Bengal, India: a long-term perspective (1982-2010). Current Science, pp. 947-953.

McInnes, K.L., Erwin, T.A. and Bathols. J.M. 2011. Global climate model projected changes in 10 m wind speed and direction due to anthropogenic climate change. Atmospheric Science Letters, 12(4): 325-333.

Mitchell, S.J. 2012. Wind as a natural disturbance agent in forests: a synthesis. Forestry, 86(2): 147-157.

Mojgan, G.M., Mehdi, M.M. and Reza, B.M. 2017. The trend of changes in surface wind in the Indian Ocean, in the period from 1981 to 2015, using reanalysis data, NCEP/NCAR. Open Journal of Marine Science, 7(4): 445-457.

Mohanan, A.H., He, Y., Mcfarlane, N. and Dai, A. 2011. The probability distribution of land surface wind speeds. Journal of Climate, 24(15): 3892-3909.

Mukhopadhyay, P., Mahakur, M. and Singh, H.A. 2009. The interaction of large scale and mesoscale environment leading to formation of intense thunderstorms over Kolkata Part I: Doppler radar and satellite observations. Journal of Earth System Science, 118(5): 441-466.

Myrl, S. 2012. The Earth's Prevailing Wind Belts. Greensboro: Mark Twain Media.

Nchaba, T., Mpholo, M. and Lennard, C. 2016. Long-term austral summer wind speed trends over southern Africa: Austral summer wind speed trends over Southern Africa. International Journal of Climatology, 37(6): 2850-2862.

R Development Core Team 2008. A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria.

Rajeevan, M., Unnikrishnan, C.K., Bhate, J., Kumar, K.N. and Sreekala, P.P. 2012. Northeast monsoon over India: variability and prediction. Meteorological Applications, 19(2): 226-236.

Reddy, G.K.K., Reddy, S.V., Sarojamma, B. and Ramkumar, T.K. 2015. Statistical analysis for wind energy estimation over Gadanki, India. Research Journal of Engineering and Technology, 1(2): 30-40.

Reddy, S.J. 2008. Climate Change: Myths and Realities. JCT, Hyderabad, India.

Venables, W.N. and Ripley, B.D. 2002. Modern Applied Statistics with S Fourth edition. Springer.

Tongkumchum, P. and McNeil, D. 2009. Confidence intervals using contrasts for regression model. Songklanakarin Journal of Science and Technology, 31(2): 151-156.

Torralba, V., Doblas-Reyes, F.J. and Gonzalez-Reviriego, N. 2017. Uncertainty in recent near-surface wind speed trends: a global reanalysis intercomparison. Environmental Research Letters, 12(11): 114019.

Troccoli, A., Muller, K., Coppin, P., Davy, R., Russell, C. and Hirsch, A.L. 2012. Long term wind speed trends over Australia. Journal of Climate, 25(1): 170-183.

UNDP 2006. Hazard profile of Indian districts, Available at https://web.archive.org/web/20060519100611/http://www.undp.org.in/dmweb/hazardprofile.pdf [August 23, 2006].

Schindler, D., Bauhus, J. and Mayer, H. 2012. Wind effects on trees. European Journal of Forest Research, 131(1): 159-163.

Vishnu, S., Francis, P.A., Shenoi, S.S.C. and Ramakrishna, S.S.V.S. 2016. On the decreasing trend of the number of monsoon depressions in the Bay of Bengal. Environmental Research Letters, 11(1): 14011.

Westin, L.K. 2001. Receiver operating characteristic (ROC) analysis: evaluating discriminanceeffects among decision support systems. UMINF report, Department of Computer Science, Umea University, Sweden: Available from: https://www8.cs.umu.se/research/reports/2001/018/part1.pdf [December, 08 2017]

Wiegand, P. 2004. Oxford International Student Atlas. Oxford University Press, Chennai, India.

Zhong, S. and Takle, E.S. 1992. An observational study of sea- and land-breeze circulation in an area of complex coastal heating. Journal of Applied Meteorology, 31(12): 1426-1438.

APPENDIX C

Article II: Statistical Modelling of Wind Velocity in Calcutta, India.

Eco. Env. & Cons. 25 (3) : 2019; pp. (1088-1095)
Copyright@ EM International
ISSN 0971–765X

Statistical modelling of wind velocity in Calcutta, India

Prashanth Gururaja[1], Nittaya McNeil [1,2*] and Bright Emmanuel Owusu[3]

[1]Faculty of Science and Technology, Prince of Songkla University, Pattani Campus, Muang Pattani, 94000, Thailand
[2]Centre of Excellence in Mathematics, Commission on Higher Education, Ministry of Education, Ratchathewi, Bangkok, 10400 Thailand
[3]Department of Mathematics, Faculty of Science and Technology, Presbyterian University College, Ghana

(Received 27 January, 2019; accepted 3 April, 2019)

ABSTRACT

Knowledge of weather and climate events such as thunderstorms, cyclones and hurricane play a vital role in the management of natural hazards. This study uses statistical models to describe patterns of wind velocity in India. Hourly wind observations during 2004-2008 were obtained from the National Renewable Energy Laboratory for the study. The multinomial model was used to describe the patterns of wind velocity categories and logistic regression to investigate the prevalence of extreme wind velocity categories. The multinomial model fitted the data quite well in describing the prevalence patterns of wind categories. Analysis of the logistic regression model revealed that gusty westerly and easterly winds prevalence were mostly higher between 4 AM to 4 PM and 4 AM to 12 PM within the day (9% and 8%) respectively. Monthly gusty westerly and easterly winds have higher percentage between January to April (9%) and January to March (9%) while higher prevalence was revealed in 2007 (7%) and 2004 (8%) respectively.

Key words : *Statistical modelling, Gusty prevalence, Wind velocity categories.*

Introduction

The dynamic and combined form of wind speed and wind direction often popularly known as wind velocity form a quint essential aspect of wind characteristics. Wind velocity prominent nature is its intensity, impactand vigorous speed. It plays a vital role in the environment, climate, and weather on Earth's surface. The wind velocity movement pattern on Earth's surface scales in diverse ranges of magnitudes, various directions and rapid intensity with time intervals from minute to few hours. The more interesting nature of wind velocity on the surface of the Earth is often referred to as wind velocity gradient or wind shear. The divergent movement patterns of wind velocity aids ingaining knowledgeable information on climate and weather, illustrate valued figures in weather analysis and forecast, provide critical data to aircraft navigation systems, improve familiarity on global warming, climate change and renewable energy. Wind velocity act as chief driving forces for other climatological factors such as rain, solar radiation, and sand (Udo *et al.*, 2008; Daut *et al.*, 2011). Besides, its sporadic behaviour can cause damages to human-built structures, human life, ecosystems andpossibly affect climate change and fatal situations in aeroplanes (Wang *et al.*, 2013; Schindler *et al.*, 2012; Suomi, 2017). However, the importance of wind velocity and its volatile behaviour can be used productively.

Numerous studies over the globe have investi-

***Corresponding author's email :** nittaya.ch@psu.ac.th

gated wind velocity patterns. A study by Pérez Foguet (2014) investigated the characterisation of local wind patterns in complex mountain valleys in Peru. The study showed that wind speed and direction were found to be relevant in wind pattern characterisation. In particular, both parameters have proven helpful in identifying and quantifying the prevailing winds during cold dawns and thermal inversion periods. Mistral and tramontane wind speed and wind direction patterns in regional climate simulations in southern France and by Obermann *et al*. (2018) in the northwestern Mediterranean Sea displayed mistral and (or) tramontane events. Most of the simulations underestimated (by 13 % on average) the wind speed over the Mediterranean Sea. The underestimation effect is most substantial at the lateral borders of the main flow-the flow width. All the simulations of this study show a clockwise wind direction bias over the sea during the mistral and tramontane events. It is essential to know the other wind factors such as wind velocity analysis and trends associated with the global wind velocity patterns (Galdies, 2011). Wind velocity influence the ecosystems (Dinulica *et al*., 2016), measurement of wind events (DePaul and Sheih, 1986). It also essential in the production of renewable energy (Simley *et al*., 2016). Further, various studies have described the significance of wind velocity patterns on the surface of the Earth. These studies include the analysis of long-term trends of wind velocity (Keevallik *et al*., 2008), seasonal trends (Ahmadi and Udo, 2016) and interannual and multidecadal intervals (Ganske *et al*., 2016).

Hewston and Dorling (2011) examined the observed daily maximum wind gusts in the UK. They revealed an increase in maximum daily gust across the UK between 1959 and the early 1990s and a decrease in gust speeds since 1993. There was a very slight increase in wind gust over the longer periods during 1959-2001. The severity of individual windstorm events was observed in recent years. However, the daily cycle is identified from the station observations in the timing of the daily maximum gust speeds, with an afternoon peak occurring between 12:00-15:00, exhibiting spatial and intra-annual variations. In addition, Bardal and Saetran (2016) investigated wind gust factors in a coastal wind climate at Froya Island, Norway. They revealed that the peak factor of gustiness decreases with increasing turbulence strength and average time. However, gustiness is less dependent on atmospheric stability. Das and Bhattacharjee (2016) examined the inter-relationship between wind direction and wind speed for five meteorological stations across Assam, India. The results showed that wind from northeast direction during both morning and evening and the association between wind speed and direction for the winter season is non-significant for all the five meteorological stations.

Statistical methods are widely applied for the analysis of wind velocity. Non-homogeneous Markov-switching autoregressive models used to describe wind time series observations for the Island of Ouessant (France) by Ailliot and Monbet (2012) described the properties of the data quite well. Moreover, investigation on the tree damage risk factors associated with large, infrequent wind disturbances of Carolina forests using logistic regression revealed that the model well predicted the tree damage increases with spatial scale. They explained the importance of examining broad geographical patterns when assessing risk factors for broad-scale disturbances such as windstorms Xi *et al*. (2008). Furthermore, the extreme value statistical analysis of meteorological parameters observed at Kudankulam site during 2004-2014 (Prabhu *et al*., 2016) estimated wind gust at 10 meters height. They revealed variations in the estimated meteorological parameter.

The present study assesses and describe the occurrence of wind velocity patterns and investigate extreme wind velocity categories of gusty westerly and easterly using statistical models.

Materials and Methods

Hourly wind observations from 2004-2008 were obtained for Calcutta from National Renewable Energy Limited (NREL) USA (https://www.nrel.gov/international/ra_india.html) for the study. The hourly wind observations were arranged on the basis of the wind direction patterns for Indian subcontinent (Wiegand, 2004; Kious and Tilling 1996; IMD, 2015). The annual wind direction patterns for the Indian subcontinent are described as north, northeast, north-west, south-west and south from December to February. It also includes south, north-west, north-east, south-west and south-east from March to May and south-west and north-east from June to September, and north-east, south-east, south, southwest and north-west from October to November respectively.

The data which did not fall under the wind direction for the particular month were discarded from the study resulting in 38,157 observations during the period of the study. The wind speed values and wind direction values were discrete, combined, and nominally ordered on the heuristic basis and classified as wind velocity. The data were classified into the following forms of wind velocity, namely None/Light, Normal-easterly, Normal-westerly, Gusty-easterlyand Gusty-westerlyas described in the left panel of Table 1.

Gusty-easterly and Gusty-westerly categories were classed as extreme wind velocity categories due to its extreme wind speed heuristically (> 5 mps) combined with direction. The naming convention of wind velocity categories is given on the left panel of Table 1. The category was labelled as none/light due tothe low magnitude of wind speed. It is very highly improbable to measure the wind direction for very low magnitude wind heuristically. Therefore the magnitude of wind speed and wind direction is omitted from naming convention of this wind velocity category. However, other wind velocity categories mentioned in Table 1 is termed as heuristic wind speed and wind direction values.

The hourly wind velocity categories had fluctuations that were smoothened by computing the 4 hour periods, and these were used for the analysis. The consecutive 4 hourly periods measurement for wind velocity recorded were notably associated with the primitive time metrics of Babylon and Egyptian system which can be represented in 4 hourly periods of 6 intervals of a day (Gillings, 1972). The station of Calcutta was selected due to its consistent 4-hourly wind data and it is also in the wind and cyclone zone with a very high damage risk (UNDP, 2006). Observations on February 29 were omitted to have an equal number of observations for each year. Therefore there were 9,540 observationsfor the study.

The four hourly wind average observations were modelled using the six periods of 4 hours each of the day, classified into clear time intervals like midnight, dawn, morning, noon, mid-noon, night (Glickman and Zenk, 2000) and indicated as a period of the day, months and year factors as the predictors.

Statistical methods

The multinomial model was applied to investigate the probabilities of occurrence of the various wind velocity categories with the predictors. The model takes the form

$$p_k = \frac{\exp\left(\alpha_k + \sum_{j=1}^{m} \beta_{jk} x_j\right)}{1 + \sum_{k=1}^{c} \exp\left(\alpha_k + \sum_{j=1}^{m} \beta_{jk} x_j\right)} \quad .. (1)$$

where p_k is the probability of the wind velocity categories, k indicates a number of wind velocity categories explained as when $c = 5$ outcome categories of wind velocity. The x_j is the matrix indicating 4-hour periods, month and year, β_{jk} is regression coefficients for outcome variable k, and the vector j, m indicates the number of predictors while α_k is a constant for the categories of outcome k.

In the investigation of the prevalence of extreme wind, a separate logistic regression model was fitted to 4-hourly extreme wind velocity categories of gusty easterly and gusty westerly observations. The monthly data had fluctuations due to seasonal effects. To smoothen these fluctuations out, bimestrial readings of the month were used. The logistic regression model takes the form:

$$\ln\left(\frac{p}{1-p}\right) = \alpha + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 \quad .. (2)$$

In formula (2), p is the probability of wind velocity prevalence for categories of gusty westerly and gusty easterly, α is constant, β's are regression co-

Table 1. Wind velocity classification

wind speed and direction classification	wind speed (mps) and direction values
None /Light	wind speed ≥ 0 and < 0.775, wind direction > 0° and < 90°
Normal-easterly	wind speed ≥ 0.775 and < 5, wind direction ≥ 90° and < 180°
Normal-westerly	wind speed ≥ 0.775 and < 5, wind direction ≥ 180° and < 360°
Gusty-easterly	wind speed ≥ 5 ,wind direction ≥ 90° and < 180°
Gusty-westerly	wind speed ≥ 5,wind direction ≥ 180° and < 360°

efficients, and x_1, x_2,x_3 are the predictors for the factors 4- hour periods, month (bi-monthly) and year respectively.

The sum contrasts were applied to attain 95% confidence intervals (CI) to compare the fitted model means with the overall wind means. This contrast provides the scale to classify levels of the factor into three groups, according to whether each relating CI is greater than or equal to, or is less than the overall mean (Tongkumchum and McNeil, 2009). All analysis was done in R(R Development Core Team, 2008).

The Receiver Operating Characteristic (ROC) curve provided a measure of goodness-of-fit for the model. The ROC curve is widely known for determining the capability of prediction of a binary outcome. It plots the sensitivity (probability of positive outcomes correctly predicted by the model) against the false positive rate (probability of all outcomes incorrectly predicted by the model) of the 4 hourly wind velocity categories.

Results

The results of the multinomial regression method applied to model the occurrence probabilities of the 4 hourly wind velocity categories during 2004-2008 have been given in Table 2. The model for the wind velocity categories was made of 3 predictors, the 4-hourly period of the day, the months, and the years. There were 24 parameters including the constant term in the model. The parameters were highly significant (p<0.01) in the model. The overall fitted multinomial model was significant ($p < 0.01$).

Table 2 gives the relative risk ratios of the model coefficients. These ratios which are the exponentiated values of the logit coefficients provides a description of the model coefficients. Table

Table 2. The results of the multinomial modelling

	none/light	Normal-easterly	Normal-westerly	Gusty-easterly	Gusty-westerly
Constant	-3.373	-2.335^{*}	-0.861^{*}	-14.033	-5.633^{*}
Period Factors (4hrs)					
1:1 AM-4 AM	-1.188	0.058	52.446	49.624	55.256
2: 5 AM-8 AM	0.949	1.998^{*}	2.113^{*}	2.936^{*}	3.909^{*}
3: 9 AM-12 AM	1.990	2.956^{*}	3.158^{*}	4.228^{*}	5.543^{*}
4: 1 PM-4 PM	0.362	1.548^{*}	1.352^{*}	2.226^{*}	3.577^{*}
5: 5 PM-8 PM	-0.857	2.600^{*}	0.880^{*}	12.461	3.345^{*}
6:9 PM-12 PM	-1.283	-0.044	-0.179^{*}	0.027	0.437
Month Factors					
January	-0.737	0.076	57.780	54.101	61.375
February	-0.542	0.860^{*}	-0.116	8.997	0.991^{*}
March	0.360	1.839^{*}	0.672^{*}	10.642	2.488^{*}
April	1.640	3.243^{*}	1.705^{*}	12.739	4.369^{*}
May	0.932	2.600^{*}	0.880^{*}	12.461	3.345^{*}
June	0.542	2.505^{*}	-0.022	12.236	1.964^{*}
July	0.469	2.317^{*}	0.327^{*}	11.451	2.189^{*}
August	0.113	2.101^{*}	-0.250^{*}	11.226	0.809^{*}
September	-0.130	1.797^{*}	-0.399^{*}	11.125	0.881^{*}
October	-1.475	0.493^{*}	-1.616^{*}	8.839	-1.213^{*}
November	-2.095	-0.486^{*}	-1.671^{*}	-4.300^{*}	-2.665^{*}
December	-1.314	0.120	-0.805^{*}	-4.602^{*}	-0.557
Year Factors					
2004	0.011	0.052	69.497	67.289	71.619
2005	-0.119	0.067	-0.295^{*}	-0.234	-0.728^{*}
2006	-1.094	-0.907^{*}	-1.132^{*}	-2.139^{*}	-2.370^{*}
2007	0.806	0.520^{*}	1.053^{*}	1.126^{*}	0.740^{*}
2008	0.719	0.350^{*}	1.093^{*}	0.604^{*}	0.635^{*}

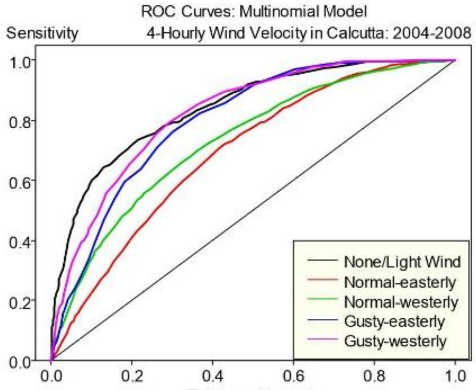
Note: * the parameter is significant

2, revealed that, the none/light wind as the reference category, if all variables are kept constant, a unit increase in wind velocity in period 1 means that, there will be 0.058, 52.446, 49.624 and 55.256 times more likely to stay in the normal-easterly, normal-westerly, gusty-easterly and gusty-westerly categories respectively relative to it staying in the none/light wind category. However, these relative ratios were not significant. In general, analysis of the various wind categories during periods (1-6) revealed that a unit increase of wind velocity in any period will result in wind velocity that is more likely to stay in the gusty-westerly category. These results were highly significant in all periods except in period 1 and 6.

For the monthly wind pattern, Table 2 revealed that a unit increase in wind velocity in January will result in 61.375, 54.101, 57.780 and 0.076 times more likely to stay in gusty-westerly, gusty-easterly, normal-westerly and normal-easterly categories respectively, relative to it staying in the none/light wind category. The gusty easterly wind is highly expected from February to December since a unit increase in the wind during these months results in higher ratios relative to other wind categories. Considering the annual factors, a unit increase in wind velocity resulted in the high gusty westerly wind category in 2004, while high normal easterly wind velocity observed during 2005-2006. In addition, high gusty easterly wind velocity was observed in 2007 while normal westerly wind velocity was prevalent in 2008.

The goodness of fit of the multinomial model for the occurrence probabilities for the various wind velocity categories, ROC curve revealed that unlike the normal easterly and westerly, which show poor classification, the remaining wind categories events did fairly well. ROC curve revealed an area under the ROC curve values of 0.69, 0.38 and 0.46 for wind velocity categories of none/light, normal easterly and normal westerly respectively.

The overall accuracy of the classification on the occurrence and non-occurrence of the various wind categories were 30%, 31%, 28%, 5%, and 4% having a true positive rate of 64%, 51%, 18%, 8%, and 34% with a false positive rate of 15.6%, 21.3%, 33%, 4%, and 3.4% respectively. These indicators depict a reasonable fit of the model to the data.



Note:- Sensitivity/AUC/True positive rate values are in same order of legend
FalsePositiveRate - 15.6%,21.3%,33%,4%,3.4%
Sensitivity - 4%,51%,18%,8%,34%
AUC - 0.685,0.382,0.462,0.599,0.647

Fig. 1. ROC curves for the multinomial model to evaluate wind velocity categories

The logistic regression modelling of the 4-hourly mean wind velocity category of gusty westerly and easterly was fixed with 3 predictors, the 4-hourly period of the day, the bi-monthly, and the year with 17 parameters including the constant term in the models. The parameters were highly significant in the models.

Evaluation of the fitted logistic regression model with the ROC curve (Fig. 2) shows that the model with a 4-hour period, bi-month are reasonable and provide a good description of the main features of the data.

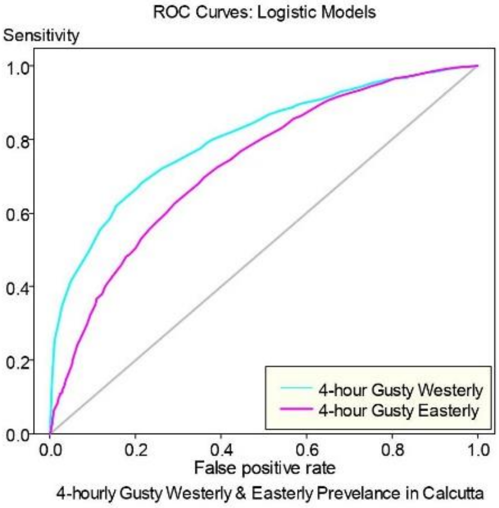


Fig. 2. ROC curves for the logistic regression model to assess the wind velocity category of gusty westerly and easterly

AUC values of 0.60 and 0.65 were observed for the gusty easterly and gusty westerly respectively. These values show a fair classification of the prevalence and non-occurrence of wind velocity categories.

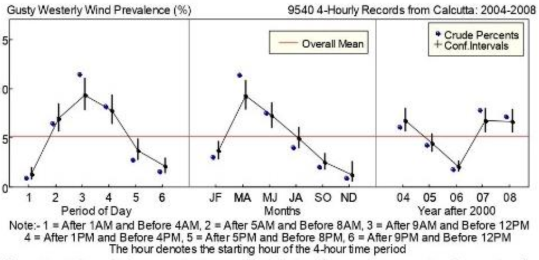
The plot of confidence intervals illustrates the gusty westerly and easterly prevalence percentage respectively for a period of the day, month and year (Figure 3 and 4) and the crude percentage. The horizontal line denotes the overall mean of the gusty westerly and easterly prevalence of approximately 5%. The gusty westerly prevalence patterns for the period of the day were significantly different from the overall mean (Figure 3).

The period from dawn (Period 2) to afternoon (Period 4) observed high westerly gust prevalence rates with the maximum gusty westerly prevalence of around 9% observed in the noon (Period 3). The low gusty easterly prevalence rates were observed from midnight to dawn (Period 1), evening (Period 5) to the night (Period 6) with minimum westerly gust prevalence of around 1 % in dawn (Period 1) respectively.

The monthly gusty westerly prevalence patterns revealed a sharp increase in prevalence from January to April where it attained its maximum (9%) above the overall mean then it reduced rapidly from April to December to attain its minimum (1%) below the overall mean.

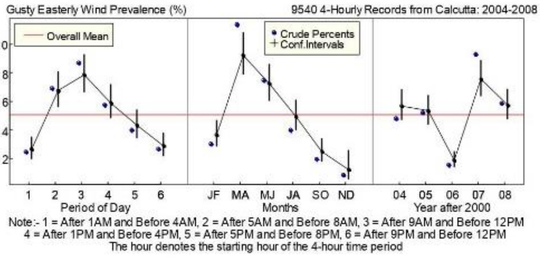
The yearly gusty westerly prevalence patterns revealed an instant decrease in gusty westerly prevalence from 2004 to 2006. It then attained its minimum of 2% below the overall mean after that, it had severe increase to attain its maximum of almost 7% above the overall mean in 2007 and a gradual decrease towards 2008.

The gusty easterly wind gust patterns for the period of the day were also significantly different from the overall mean (Figure 4). During (period 1) at dawn to (period 3) noon illustrated a quick increase of prevalence where it attained its maximum of almost 8% above the overall mean, after that, it decreased rapidly during (period 3) to (period 6) at night aligning with the overall mean. The minimum prevalence was observed in (period 1) at midnight (2%).



Note:- 1 = After 1AM and Before 4AM, 2 = After 5AM and Before 8AM, 3 = After 9AM and Before 12PM
4 = After 1PM and Before 4PM, 5 = After 5PM and Before 8PM, 6 = After 9PM and Before 12PM
The hour denotes the starting hour of the 4-hour time period

Fig. 3. Confidence interval plot of gusty westerly wind prevalence for the logistic model.



Note:- 1 = After 1AM and Before 4AM, 2 = After 5AM and Before 8AM, 3 = After 9AM and Before 12PM
4 = After 1PM and Before 4PM, 5 = After 5PM and Before 8PM, 6 = After 9PM and Before 12PM
The hour denotes the starting hour of the 4-hour time period

Fig. 4. Confidence interval plot of gusty easterly wind prevalence for the logistic model.

The monthly gusty easterly prevalence patterns revealed a sharp increase in prevalence from January to April where it attained its maximum (9%), after that, it decreased briskly from April to December to attain its minimum (1%) below the overall mean.

The yearly gusty easterly prevalence patterns revealed decreasing trends during 2004 and 2006 where it attained its minimum of almost 2% below the overall mean, and after that, it had a rapid increase to attain its maximum of almost 8% above the overall mean in 2007 and a gradual decrease towards 2008.

Conclusion and Discussion

Statistical models were used to investigate the 4-hourly wind velocity categories in India. The categories comprised none/light, normal easterly, normal westerly, gusty westerly, and gusty easterlyof Calcutta during 2004-2008. The multinomial regression method used to describe the wind velocity categories pattern fits the data reasonably well. These patterns of wind velocity categories were predominantly due to discrete wind variability rates surrounding over Indian subcontinent landmass. The prevalence of gusty westerly and gusty easterly patterns was investigated. The gusty westerly and easterly prevalence patterns for the period of the day, month and year illustrated distinct, varied patterns. The overall gusty westerly and easterly prevalence were observed to be almost 5% all through the year.

The period of the day factors revealed that the gusty westerly prevalence was above the overall mean between period 2 and period 4 and almost below the overall mean during the other periods of the day. These patterns of gusty westerly may be due to wind-induced ocean swell (Alves 2006).The monthly maximum gusty westerly prevalence of 9% was observed during January to April. Also, the minimum gusty westerly prevalence occurred between April to December (1%) and may be due to global warming and ocean warming (Mojgan *et al.*, 2017). The yearly gusty westerly wind prevalence was mostly below the overall mean during most periods of 2005 to 2006 but was above the overall mean during 2004 while the maximum of 7% occurred between 2007 and 2008. This pattern was partly because of the monsoon and coastal currents (Rao *et al.*, 2012). In general, the results of the 95% confidence interval plot for the gusty easterly prevalence was similar to that of gusty westerly prevalence. The models have described the patterns of the various wind categories significantly. These methods can be used to describe the patterns of other climate variables such as rainfall and solar radiation.

Acknowledgement

The authors are grateful to Emeritus Prof. Don McNeil for supervising this work. We would like to thank the Higher Education Research Promotion and Thailand's Education Hub for Southern Region of ASEAN Countries Project Office of the Higher Education Commission under grant no. THE/033 and Thanks to Centre of Excellence in Mathematics, Commission on Higher Education, Thailand.

References

Ahmadi, S. O. and Udo, A. O. 2016. Analysis of Trends and Variations of Monthly Mean Wind Speed Data in Nigeria. *IOSR Journal of Applied Physics* (IOSR-JAP). 7(4) : 31-41.

Ailliot, P. and Monbet, V. 2012. Markov-switching autoregressive models for windtime series. *Environmental Modelling and Software*. 30 : 92-101.

Alves, J. H. G. 2006. Numerical modeling of ocean swell contributions to the globalwind-wave climate. *Ocean Modelling*. 11(1-2) : 98-122.

Bardal, L. M. and Saetran, L. R. 2016. Wind gust factors in a coastal wind climate. *Energy Procedia*. 94: 417-424.

Das, K. and Bhattacharjee, S. 2016. The inter-relationship between wind direction andwind speed for 5 meteorological stations across Assam, India: A case study. *International Journal of Approximate Reasoning*. 4 (5) : 591-603.

Daut, I., Irwanto, M., Irwan, Y. M., Gomesh, N. and Rosnazri, N. A. 2011. Potential of solar radiation and wind speed for photovoltaic and wind power hybrid generation in Perlis, Northern Malaysia. In *Power Engineering and Optimization Conference* (PEOCO), 2011 5th International (pp. 148-153). IEEE.

DePaul, F. T. and Sheih, C. M. 1986. Measurements of wind velocities in a streetcanyon. *Atmospheric Environment* (1967). 20 (3) : 455-459.

Dinulica, F., Marcu, V., Borz, S. A., Vasilescu, M. M. and Petritan, I. C. 2016. Wind contribution to yearly silver fir (*Abies alba* Mill.) compression wood development in the Romanian Carpathians. *iForest-Biogeosciences and Forestry*. 9 (6) : 927.

Galdies, C. 2011. The Climate of Malta: statistics, trends and analysis 1951-2010. IMD. 2015. Indian Meteorological Department, Monsoon, Availableat: http://www.imd.gov.in/pages/monsoon_main.php (accessed February 5, 2017).

Ganske, A., Tinz, B., Rosenhagen, G. and Heinrich, H. 2016. Interannual and Multidecadal Changes of Wind Speed and Directions over the North Sea from Climate Model Results. *Meteorologische Zeitschrift*. 25 : 463-478.

Gillings, R.J. 1972. *Mathematics in the time of the pharaohs*, Dover Publications Inc, New York, USA.

Glickman, T.S. and Zenk, W. 2000. *Glossary of Meteorology*, American Meteorological Society, Boston, USA.

Hewston, R. and Dorling, S. R. 2011. Analysis of observed daily maximum wind gusts in the UK. *Journal of Wind Engineering and Industrial Aerodynamics*. 99 (8): 845-856.

Keevallik, S. and Soomere, T. 2008. Shifts in early spring wind regime in North-East Europe (1955–2007). *Climate of The Past*. 4 (3) : 147-152.

Kious, W.J. and Tilling, R.I. 1996. This dynamic Earth: the story of plate tectonics, DIANE Publishing, Darby, USA.

Mojgan, G. M., Mehdi, M. M. and Reza, B. M. 2017. The Trend of Changes in SurfaceWind in the Indian Ocean, in the Period from 1981 to 2015, Using Reanalysis Data, NCEP/NCAR. *Open Journal of Marine Science*. 7 (04): 445.

Obermann, A., Bastin, S., Belamari, S., Conte, D., Gaertner, M. A., Li, L. and Ahrens, B. 2018. Mistral and tramontane wind speed and wind direction patterns in regional climates imulations. *Climate Dynamics*. 51 (3) : 1059-1076.

Pérez Foguet, A. 2014. Characterization of local wind patterns in complex mountain valleys. *International Journal of Climatology*. 34 (6) : 1741-1759.

Prabhu, J., George, T., Vijayakumar, B. and Ravi, P. M. 2016. Extreme value statistical analysis of meteorological parameters observed at Kudankulam site during 2004–2014. *Radiation Protection and Environment*. 39 (2) : 107.

R Development Core Team. 2008. A Language and environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria.

Rao, A. D., Mahapatra, D. K., Babu, S. V. and Pandey, S. 2012. Spatial-temporal variation and mechanism associated with mini-cold pool off the southern tip of India during summer and winter monsoon season. *Water Quality Research Journal*. 47 (3-4) : 333-352.

Schindler, D., Bauhus, J. and Mayer, H. 2012. Wind effects on trees. *European Journal of Forest Research*. 131 (1): 159-163.

Simley, E., Angelou, N., Mikkelsen, T., Sjöholm, M., Mann, J. and Pao, L. Y. 2016. Characterization of wind velocities in the upstream induction zone of a wind turbine using scanning continuous-wave lidars. *Journal of Renewable and Sustainable Energy*. 8 (1) : 013301.

Suomi, I. 2017. Wind gusts in the atmospheric boundary layer. Schindler, D., Bauhus, J. and Mayer, H. 2012. Wind effects on trees. *European Journal of Forest Research*. 131 (1) : 159-163.

Tongkumchum, P. and McNeil, D. 2009. Confidence intervals using contrasts for regression model. *Songklanakarin Journal of Science and Technology*. 31(2): 151-156.

Udo, K., Kuriyama, Y. and Jackson, D. W. 2008. Observations of wind blown sandunder various meteorological conditions at a beach. *Journal of Geophysical Research: Earth Surface*. 113 (F4).

UNDP. 2006. Hazard profile of Indian districts, Available at https://web.archive.org/web/20060519100611/http://www.undp.org.in/dmweb/hazardprofile.pdf [August 23, 2006].

Wang, C. H., Wang, X. and Khoo, Y. B. 2013. Extreme wind gust hazard in Australia and its sensitivity to climate change. *Natural Hazards*. 67 (2) : 549-567.

Wiegand, P. 2004. *Oxford International Student Atlas*, Oxford university press, Chennai, India.

Xi, W., Peet, R. K., Decoster, J. K. and Urban, D. L. 2008. Tree damage risk factors associated with large, infrequent wind disturbances of Carolina forests. *Forestry*. 81 (3) : 317-334.

APPENDIX D

Conference Proceeding

Episode based graphs, time series graphs, signal plus noise models for analyzing and predicting wind velocity of Indian weather stations

Prashanth Gururaja; Donald Roy McNeil

ABSTRACT

Extreme weather events such as strong winds, high temperatures, and heavy rainfall can cause major devastation. Evaluation of wind performance is still in its initial stages. An informative graphical display of improvised hourly wind velocity measurement helps in understanding how this costly event can be modeled and predicted. Uninterrupted wind for 6 hours is called **Episodes**.

Wind episodes are defined as periods of uninterrupted wind separated by occasions of zero wind speed. Graphs provide useful supplements to standard frequency-based displays providing an appropriate basis for data enhancement, fitting statistical models that can be used for measuring and forecasting. We illustrate these concepts using sample data from Indian weather stations.

The objectives of the study are:

- To analyze wind speeds, wind direction, develop Episode based graphs and time series graphs to measure hourly wind velocity.
- Develop a statistical model for predicting wind velocity and evaluate extreme weather events.

We fit linear regression model for analyzing wind velocity. The Measured variables Wind speed, wind direction with continuities, discontinuities, is regressed on predictor variables wind direction, days, months, and 5 day periods. Model error

analysis is performed by extracting the residuals from the model. Quantile graphs of the model, auto correlation function graphs, and partial auto correlation function graphs on the model residuals are plotted. The coefficients, standard errors, crude mean, adjusted mean are calculated. The graphs of the wind direction, hours, and 5 day periods on the mean of the wind speed are plotted. Graphs plotted illustrate there were confounding factors of zero and nonzero wind speed, when we observe the crude mean, adjusted mean on the mean of the wind speed. We fit logistic regression model to analyze zero, non zero wind speed, Evaluate wind speed on hourly, and 5 day periods. There was confounding effect of zero and nonzero wind speed.

We developed a novel graphical method of displaying wind velocity as an episode based time series using hourly observations. A statistical model was fitted for wind speed. The result illustrates wind velocity can be analyzed on mean wind speed, wind direction, hours and 5 day periods. The model had an adjusted r-squared of 27% but the errors contained substantial serial correlation, some departure from normality. Further research will involve using **arima model for** extensive hourly wind velocity measurements and forecasting.

*** 5 days = 1 period., 365 days = 73 periods (5*73)

Renewable
Energy India Expo
A UBM Event

23-25 SEPTEMBER, 2015
India Expo Centre, Greater Noida, India

UBM

Certificate of Appreciation

Certificate of Appreciation is hereby presented to

Mr. Prashant Gururaja

For submitting your Research in the Wind Sector
At

9th RENEWABLE ENERGY INDIA
EXPO 2015

Mr. Rajneesh Khattar
Group Director
UBM India Pvt. Ltd.

VITAE

Name	Mr. Prashanth Gururaja
Student ID	5720330001

Educational Attainment

Degree	**Name of the Institution**	**Graduated on**
BSc Computer Science	Newport University, Bangalore, India.	1999
MBA. Computer Resource and Information Management	Webster University, USA	2003

Scholarship Awards during Enrolment

- Thailand's Education Hub for Southern Region of ASEAN Countries (TEH-AC) Scholarship.
- Thesis support grant from Graduate School, Prince of Songkla University, Thailand

Work-Position and Address:

Position: Program Manager

Address: Capgemini India Pvt. Ltd, Priteck Park, Block B, (Gr. Floor), Primal Project Pvt Ltd, Ltd SEZ, Bellandur, Village, Varthur Hobli, Bangalore East Taluk, Bangalore Urban District, Bangalore - 560103.

List of Publications and Proceedings

Publications

Gururaja, P., McNeil, N. and Owusu, B.E. 2019a. Statistical Modelling of 4-hourly Wind Patterns in Calcutta, India. Nature Environment and Pollution Technology, 18(1),.73-80.

Gururaja, P., McNeil, N. and Owusu, B. E. 2019b. Statistical modelling of wind velocity in Calcutta, India. Ecology, Environment and Conservation, 25 (3): 1088-1095.

Proceedings

Prashanth Gururaja and Donald Roy McNeil. 2015. Episode based graphs, time series graphs, signal plus noise models for analyzing and predicting wind velocity of Indian weather stations. 9th Renewable Energy Expo, India.