MODELLING WIND PATTERNS WITH IRREGULARLY DISTRIBUTED MEASUREMENT DATA USING- R

D.R.Welikanna*

*Dept. of Surveying and Geodesy, Sabaragamuwa University of Sri Lanka,Belihuloya 70140, Sri Lanka – (drw@geo.sab.ac.lk)

Abstract

Wind behavior in Sri Lanka, exposed to two major monsoon season is complex to model. The acute changes in the topography further make it difficult to map the wind patterns. Irregularly distributed wind measurements are the main source of information available to interpret and model this dynamic Geographic phenomenon. Sri Lanka is in need of an accurate and timely wind maps in order to support wind energy resource planning and to empower weather prediction system's accuracy. Heavy and continuous winds in the southern part of the island encourage the wind harvesting for energy. Wind energy is arguably the most affordable per megawatt hour of the renewable energy source and it is growing nearly as quickly as conventional generation techniques. This study has used 23 diurnal wind measuring stations data in the form of wind speed and direction. These data were analyzed on the basis of morning and evening time periods and was aggregated on weekly basis by using the mean wind speed and the mode of the wind direction. R which is an open source statistical programming environment has been used completely for all the data processing, modelling and analysis stages. The spatial correlation of the wind speed and the direction was determined in the beginning by using the semi-Variograms. Spherical semi-variogram was fitted with the observations. Ordinary Kriging which uses the samples in the local neighborhood avoiding the mean estimation was used to perform the spatial interpolations. For each month of the year 2016, four weekly wind maps were generated, both for the direction and speed. These 48 wind maps were further analyzed for the wind patterns that exist within the island during the year 2016. The changes in the morning and evening wind patterns and the significant wind pattern changes along the coastal areas and the central highlands are extracted and presented in the study. The generated maps could be useful in identifying potential areas for wind farming.

Keywords: Wind Speed, Wind Direction, Variograms, Kriging, R

Corresponding Author: Welikanna D R
ORCID:
Email address:drw@geo.sab.ac.lk
DOI:

1 INTRODUCTION

1.1 Background

Climate can be considered as a sum of atmospheric elements and their variation along time. Those main constituent elements contributing to the climate are solar radiation, temperature, humidity, clouds and precipitation (type, frequency, and amount), atmospheric pressure, and wind (speed and direction). Hence Wind becomes one of the main components in the determination of the Climatic conditions pertaining to certain time frame. Wind is an extremely dynamic Geographic field. They are externally discontinuous. Mapping and predicting wind behavior and the directions is extremely challenging task. Geostatistical approaches provide feasible platform to map and predict Wind patterns using ground based measurement data. Sri Lanka as an island which is positioned within the tropics between 5° 55' to 9° 51' North latitude and between 79° 42' to 81° 53' East longitude, and holds a tropical climate. Due to the tropical position of the island it is not affected by seasonal variations due to the ecliptic of the Sun. Though this is the case considering mainly Temperature variations the islands climate is mainly controlled by the seasonal winds. If closely observed the main element that determines the climate of Sri Lanka, is its Wind. Out of these seasonal wind regimes, two stands out as the Southwest and Northeast monsoons regional scale winds. The Climate experienced during 12 months period in Sri Lanka can be characterized in to 4 climate (Table 1) seasons depending mainly on the seasonal winds experienced. First inter monsoon season (March and April) brings in thunderstorm-type rain, particularly during the afternoons or evenings. The distribution of rainfall during this period shows that the entire South-western sector at the hill country receiving heavy rainfall, with localize area on the South-western slops experiencing rainfall in excess of 700 mm. Southwest monsoon season (May to September), Windy weather during this monsoon eases off the warmth that prevailed during the 1st Inter monsoon season. Southwest monsoon rains are experience at any times of the day and the nights, mainly focused on the southern part of the country. Second inter monsoon season (October to November), brings in thunderstorm type of rain particularly during the afternoon or evening. Under such conditions, the whole country experiences strong winds with wide spread rain, sometimes leading to floods and landslides. Northeast monsoon season (December and February) is characterized by dry and cold wind blowing from the Indian land-mass making cooling effect all over the island. Cloud free skies provide days full of sunshine and pleasant and cool night. During this period, the highest rainfall figures are recorded in the North, Eastern slopes of the hill country (Department of Meteorology, Sri Lanka 2016).

1.2 R Open Source

R is a free software environment for statistical computing and graphics also is an integrated suite of software facilities for data manipulation, calculation and graphical display. It includes an effective data handling and storage facility, a suite of operators for calculations on arrays, in particular matrices, a large, coherent, integrated collection of intermediate tools for data analysis, graphical facilities for data analysis and display either on-screen or on hardcopy, and a well-developed, simple and effective programming language which includes conditionals, loops, user-defined recursive functions and input and output facilities. While it is a GNU project, R provides a wide variety of statistical (linear and nonlinear modelling, classical statistical tests, time-series analysis, classification, clustering, Remote Sensing data processing, GIS, variogram modelling, spatial data handling and graphical techniques (www.r-project.org). R Packages provide more power in data handling; they are collections of functions and data sets developed by the community. They increase the power of R by improving existing base R functionalities, or by adding new ones. A set of R Packages used in this study are, libraray(rgeos), libraray(maptools), libraray(rgdal), libraray(sp), libraray(rgoogleMaps), libraray(ggmap), libraray(ggplot2), libraray('latticeExtra), library(ggrepel), library(gstat), library(dplyr), library, library(scales) and library(magritr). These libraries were chosen and implemented depending on different need in the spatial data processing stages, in addition most base packages in R were used for data input and analysis.

2 STUDY AREA AND THE DATA PREPERATION

Although the climate of Sri lanka is controlled strongly by its wind behavior, Sri lanka does not have comprehensive local wind maps to understand the patterns and the changes. Wind measuring stations are located as shown in Figure 1 below. The irregular distribution of the wind measuring stations not covering the island in a grid based system is concern in wind predictions and the mapping process. In this study weekly wind measurement values for the twenty three stations shown, recorded by the Sri Lanka Meteorological Department for the year 2015 has been used. Weekly wind measurement values were averaged for the four main wind seasons based on the wind values in the morning and the evening hours further for the analysis.

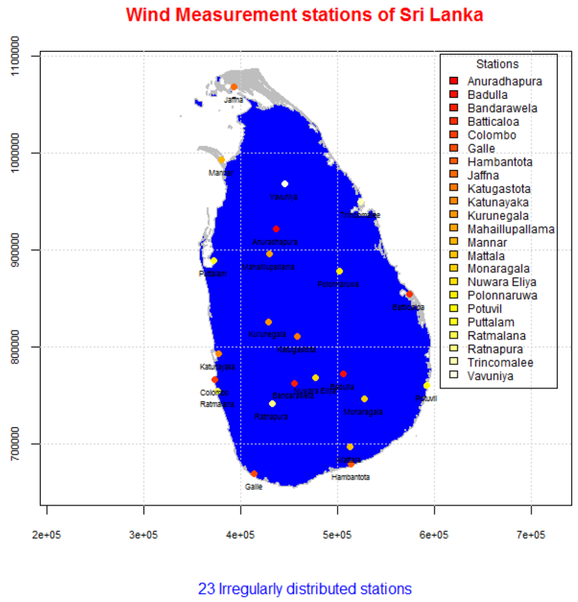

Figure 1: Wind measuring stations of Sri Lanka

Table 1: Summery statistics of the four wind regimes for the mornings and evening, year 2015

1st Inter monsoon March-April	**South West Monsoon May- September**
Morning Wind stat Min. 1st Qu. Median Mean **0.853 1.608 2.585 3.004** 3rd Qu. Max. **3.328 7.907** Evening Wind stat Min. 1st Qu. Median Mean **2.196 3.566 4.859 5.461** 3rd Qu. Max. **6.575 12.675**	Morning Wind speed stat Min. 1st Qu. Median Mean **1.207 3.101 5.404 6.062** 3rd Qu. Max. **8.030 15.653** Evening Wind stat Min. 1st Qu. Median Mean **2.952 5.328 6.019 7.282** 3rd Qu. Max. **8.119 17.872**
2nd Inter monsoon October-November	**North East Monsoon December-February**
Morning Wind stat Min. 1st Qu. Median Mean **0.654 1.375 2.449 2.822** 3rd Qu. Max. **3.1726 7.7060** Evening Wind stat **0.915 2.181 3.463 3.832** Min. 1st Qu. Median Mean 3rd Qu. Max. **4.734 10.339**	Morning Wind stat Min. 1st Qu. Median Mean **0.493 1.618 3.242 3.814** 3rd Qu. Max. **4.056 11.587** Evening Wind stat Min. 1st Qu. Median Mean **2.017 3.238 4.863 5.861** 3rd Qu. Max. **7.211 14.513**

Summary of the wind measurement statics are shown above in the Table 1. Slightly higher wind speeds with respect to the mornings could be seen for all the four seasons (Figure 2). Maximum wind speeds could be observed during the South West monsoon season, while almost steady or windless conditions could be seen during the 1st and 2nd Inter monsoon seasons. Large variation in the wind speeds as well as incoming rain could be expected during the South West monsoon season. According to the Figure 2 the evening wind speeds are slightly higher than the morning wind speeds, generally across the island. These changes in the average wind speed measure and the variance in the wind speeds depending on the morning and the evening add the wind modelling several more dimensions.

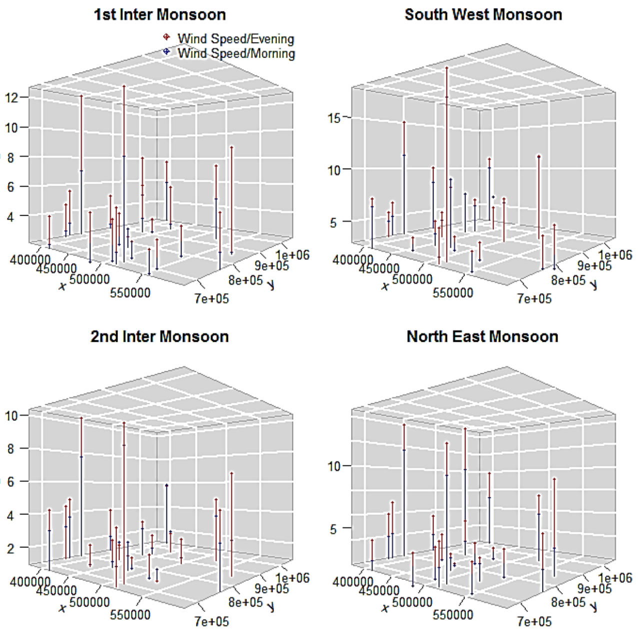

Figure 2: Wind speeds during the morning and the evening hours of the four wind seasons for the wid stations, year 2015

As the topography of the island suggests the coastal belt representing Jaffana, Katunayaka, Puttalam and Trincolmaless receives very high wind effects with their values reaching large, while Colombo, Galle, Hambanthota and Baticaloa eventually are coastal as well reaches moderately high yet steady wind effects (Figure 3). Most of the central part of the country especially Badulla, Anuradhapura, Kurunagala, Nuwara Eliya and Ratnapura receives low wind effects.

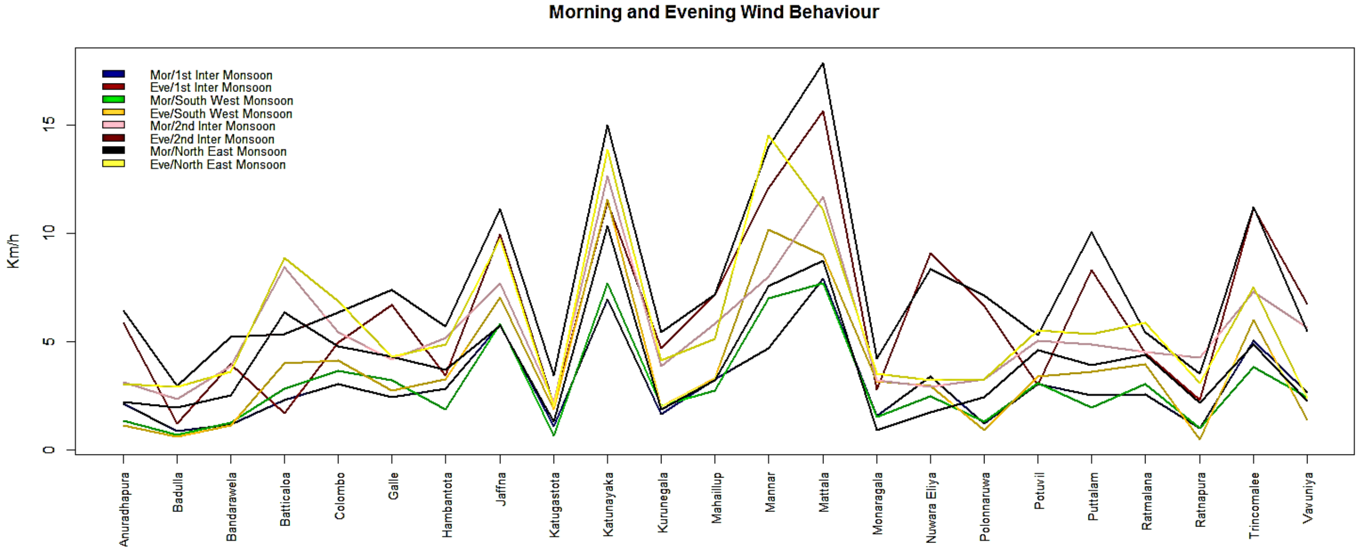

Figure 2: Time series seasonal Wind speeds four each of the 23 measuring stations, 2015

3 GEOSTATITICAL METHOD: SEMIVARIANCE AND KRIGING INTERPOLATION

With Geo-statistical methods modelling data with significant changes from normal distribution and with non-stationary mean and variance is still a challenge. Spatial structure which describes the spatial dependence in distance and direction with respect to the data is an important parameter in the process of estimation. Semivariance can be seen as one of the main parameter determining the Auto-Covariance. Semivariance also is a measure of dissimilarity between observations while the covariance and the correlation both measure the similarity. The Semivariance function shown below quantifies the spatial correlation between measurement points.

$$\gamma(h) = \frac{1}{2m(h)} \sum_{i=1}^{m(h)} [Z(x_i - x_j)] \quad (1)$$

Here $m(h)$ is the number of point pairs separated by vector h.The Semivariance can be a function of both distance as well as direction, making way to account for direction-dependent variability. This change in autocorrelation structure with respect to the direction could be reflected by using anisotropic variogram. From the Empirical variogram computed by the Eq. 1 we derive the variogram model which expresses the Semivariance as function of the separation vector. The model allows us to infer the characteristics of the underlying process from functional form and its parameters. This computed Semivariance between any points pair is used in optimal interpolation such as Kriging to make predictions at unsampled locations. Due to theoretical and mathematical limitations several functional forms can be used to model the variogram, these models are also called the authorized models. In this study three of the authorized models as shown below were used to determine the weights for the nearby supporting data to compute the interpolated values.

The exponential model:

$$\gamma(h) = c\left\{1 - e^{-\frac{h}{a}}\right\} \quad (2)$$

Where h is the lag distance, sill is c and effective range is $3a$. The Spherical model (Figure 4)

$$\gamma(h) = \begin{cases} c\left\{\frac{3h}{2a} - \frac{1}{2}\left(\frac{h}{a}\right)^3\right\} & h < a \\ c & h \geq a \end{cases} \quad (3)$$

And the Gaussian model:

$$\gamma(h) = c\left\{1 - e^{-3\frac{h^2}{a^2}}\right\} \quad (4)$$

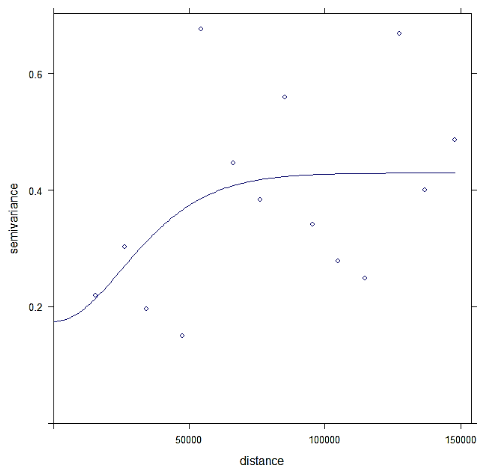

Figure 4: A variogram with a Spherical model fitting

Kriging (Krige, 1966) is a stochastic technique similar to IDW, in that it uses a linear combination of weights at known points to estimate the value at an unknown point. In contrast with deterministic methods, kriging provides a solution to the problem of estimation of the surface by taking account of the spatial correlation which is quantified by the Semivariance function. Varieties of kriging have been developed such as ordinary, universal, simple and indicator but only the ordinary and the universal kriging were used in this study. In this paper ordinary kriging experiments are presented. Ordinary kriging which assumes the observation mean is unknown focuses on the spatial component and uses only the samples in the local neighborhood for the estimation. Universal kriging is similar to that of ordinary kriging, but assumes the presence of a trend in average values across the study area, further details of kriging could be found in (Isaaks and Srivastava, 1989; Cressie, 1993; Webster and Oliver, 2001). In ordinary kriging (OK) , we model the value of variable Z at location x_i as the sum $Z(x_i) = m + e(x_i)$, where m is the regional mean and $e(x_i)$ the spatially correlated random component. The regional mean m is estimated from the sample, but not as the simple average, because there is spatial dependence. It is implicit in the OK system. This mean is constant across the field that is the expected value is the same and unknown; this is the "Ordinary"situation. The spatially-correlated random component $e(x_i)$, is estimated from the spatial covariance structure as revealed by the variogram model. The estimated value $\hat{Z}$ at a point x_0 is predicted as the weighted average of the values at all sample points x_i:

$$\hat{Z}\,(x_0) = \sum_{i=1}^{N} \lambda_i Z(x_i) \qquad (5)$$

Where $\sum_{i=1}^{N} \lambda_i = 1$ so that the estimation is unbiased, as $E[\hat{Z}\,(x_0) - Z(x_0)] = 0$.

4 Experimental results

The variogram parameters for the fitted anisotropic variogram models for the four seasons were calculated and modelled. Two such models for the south west monsoon morning hours and the north east monsoon evening hours are shown below in the Figure 5 and Figure 6 respectively. Exponential variogram model has been the widest choice in fitting the empirical variograms produced for each wind season. Estimating these functions in the case of irregular distributed data could be extremely challenging, mainly due to the need to pool the data pairs into larger bins (Bargaoui and Chebbi, 2009). Larger lag spacing and the tolerance increase the number of data pairs for estimation but reduces the amount of details in the Semivariogram. The spanning of the measuring stations in longitudinal direction is about 219 Km, while in latitudes it is about 398Km. The bin sizes for the variogram was estimated using the bounding box of the data points and it was set to 10Km.

Experimental Semivariogram of average wind speed during the morning hours of the South West monsoon season has been fitted by a spherical model in four different azimuthal directions. The Semivariogram in the direction N 90°E was used for the weight determination for the interpolation. The Semivariogram levels off to the sill almost at a distance of 5 Km. Partial sill is about 0.75.

For the second case Experimental Semivariogram of average wind speed during the evening hours of the South West monsoon season has been fitted by an exponential model in four different azimuthal directions. As for the morning the Semivariogram in the direction N 90°E was used for the weight determination for the interpolation. The Semivariogram levels off to the sill almost at a distance of 4 Km. Partial sill is about 0.07.

In the third case Experimental Semivariogram of average wind speed during the morning hours of the North East monsoon season has been fitted by an exponential model in four different azimuthal directions. As for the morning the Semivariogram in the direction N 90°E was used for the weight determination for the interpolation. The Semivariogram levels off to the sill almost at a distance of 4 Km. Partial sill is about 1.5.

Finally four the 4th monsoon, experimental Semivariogram of average wind speed during the evening hours of the North East monsoon season has been fitted by an exponential model in four different azimuthal directions. As for the morning the Semivariogram in the direction N 135°E was used for the weight determination for the interpolation. The Semivariogram levels off to the sill almost at a distance of 5 Km. Partial sill is about 0.6. The resulting wind speed maps for the 4 seasons are shown in Figure 7.

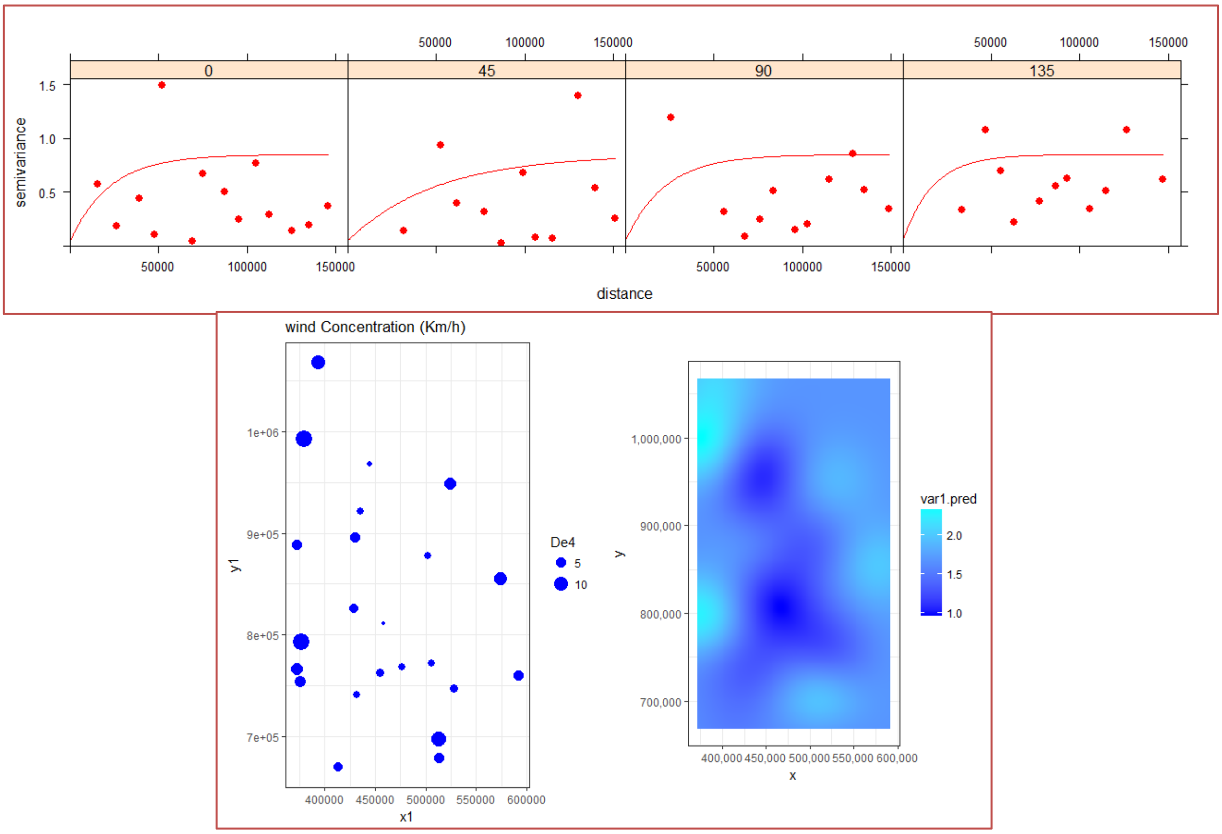

Figure 5: South West Monsoon, Morning hours, Anisotropic Variograms and the wind predictions

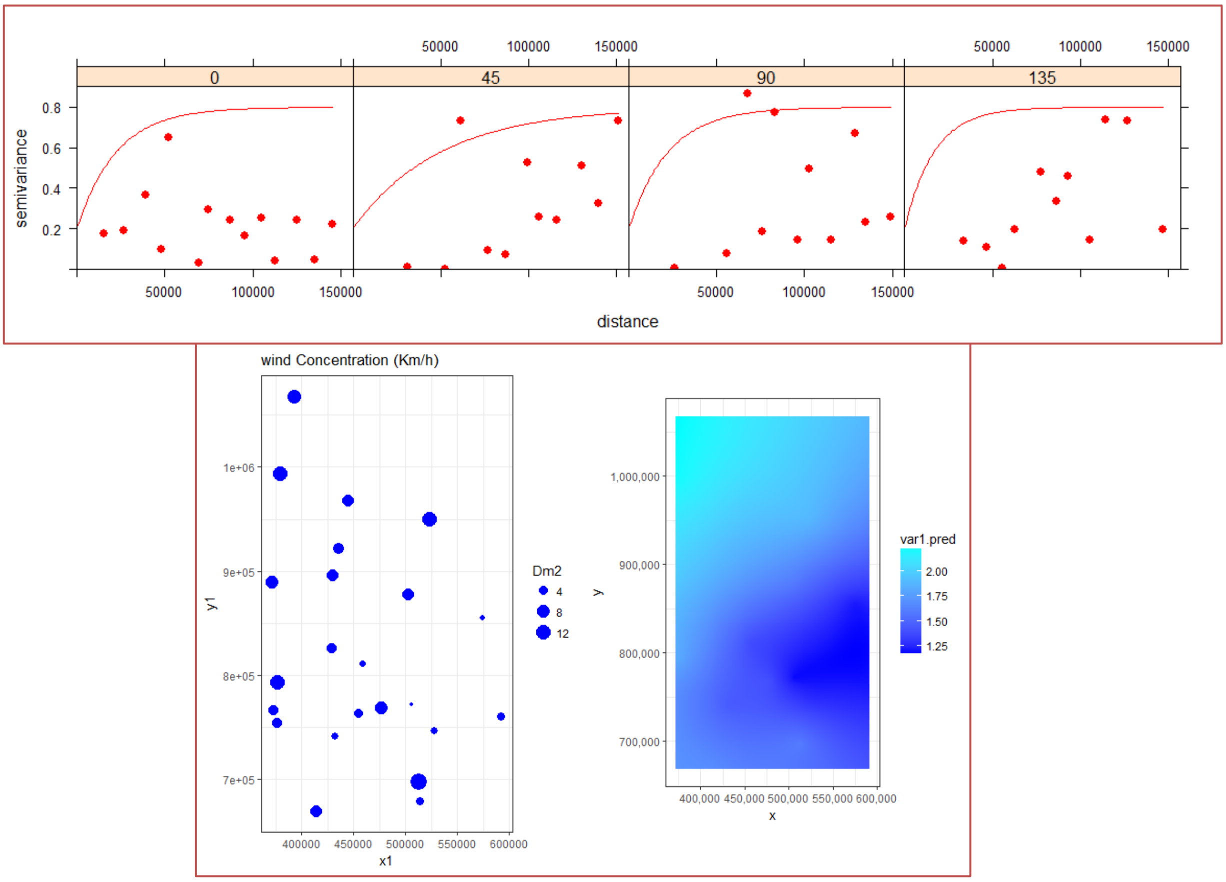

Figure 6: North East Monsoon, Evening hours, Anisotropic Variograms and the wind predictions

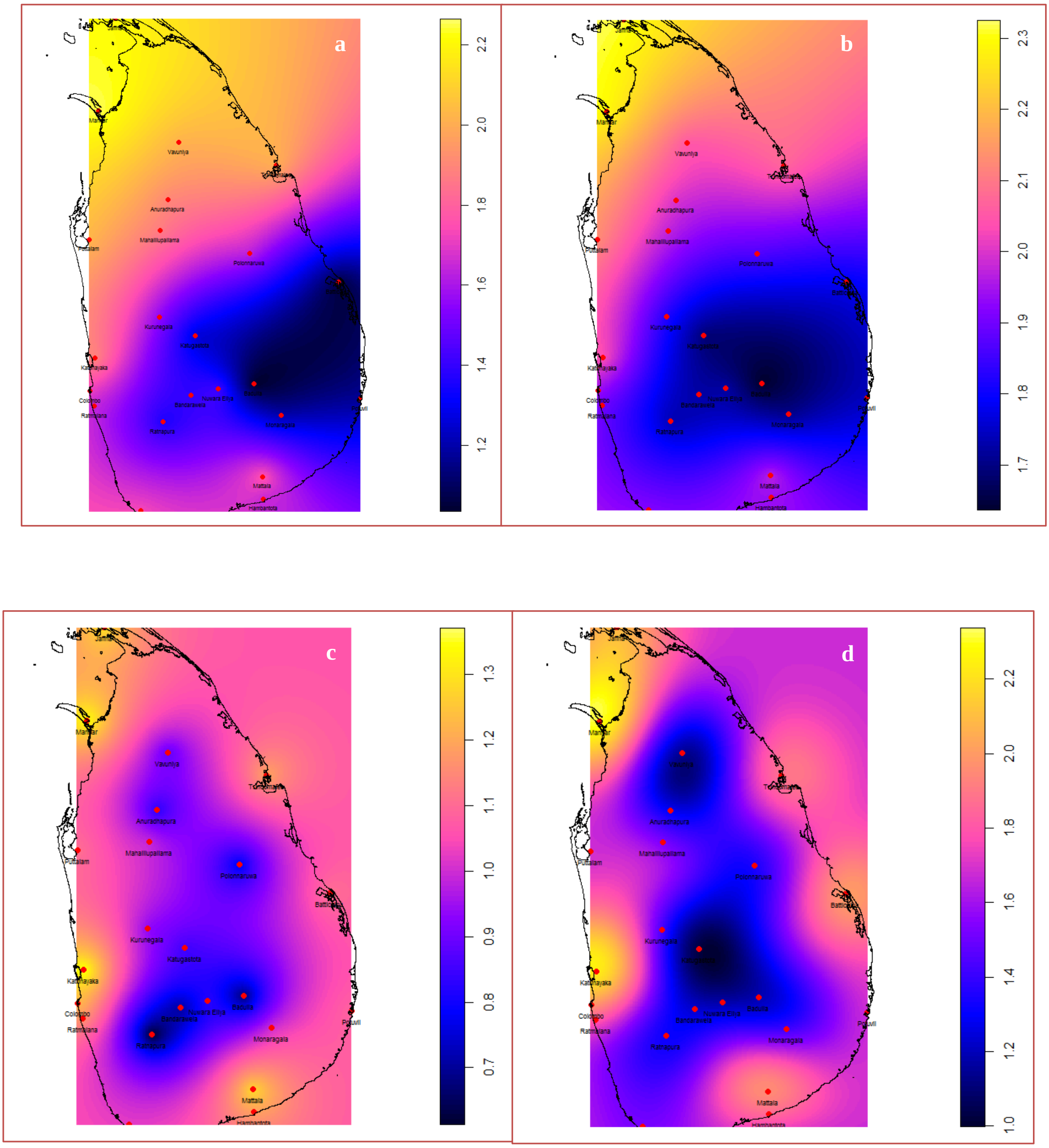

Figure 7: Ordinary Kriging based Interpolated wind speed surfaces (a) Morning of the South West Monsoon (b) Evening of the South West Monsoon (c) Morning of the North East Monsoon and (d) Evening of the North East Monsoon seasons respectively

5 Conclusions and Recommendation

The results of this study indicate the use of variograms and the ordinary kriging interpolation produce simple yet robust wind speed predictions. The whole statistical modelling were done using the R 3.5.0 environment, the possibility to handle spatial data with different projection coordinate systems within R has made the kriging based interpolation more robust (Table 2). The irregularly distributed wind measuring stations brings in difficulties in the variogram modelling. The paper reports only the predictions for the two main wind seasons for Sri Lanka. A comparison with other interpolation methods such as the universal and co-kriging could be an insight for a complete prediction mechanism for the scares data as was used in this study. Comprehensive validation of the prediction results is necessary in order to provide strong justification of the methods adopted. The study will extend in the direction of producing mobile wind measuring data in the regions where the measurements have not been made in order to see the performance of the interpolations. Possibilities of using external drifts such as altitude will also be one further step of this study.

Table 2 .R Packages additional to basic packages employed in the study

R Package	Usage
rgeos	Interface to Geometry Engine
maptools	Manipulating and reading geographic data, in particular 'ESRI Shapefiles'
rgdal	Access to projection/transformation operations
sp	Classes and Methods for Spatial Data
rgoogleMaps	Use the google map as a background image
ggmap	Spatial Visualization with ggplot2
ggplot2	System for 'declaratively' creating graphics
latticeExtra	Powerful and elegant high-level data visualization system
ggrepel	Automatically Position Non-Overlapping Text Labels
gstat	Variogram modelling; simple, ordinary and universal point or block (co)kriging; spatio-temporal kriging; sequential Gaussian or indicator (co)simulation; variogram and variogram map plotting utility functions
dplyr	Consistent tool for working with data frame like object
scales	Graphical scales map data to aesthetics
magritr	Provides a mechanism for chaining commands

6 REFERENCES

E.H. Isaaks and R.M. Srivastava, (1989). *An Introduction to Applied Geostatistics*, Oxford University Press, 561 pages.

Krige, D. G., (1966), Two-dimensional weighted moving average trend surfaces for ore valuation,*in* Symposium on mathematical statistics and computer applications in ore valuation: Johannesberg, South Africa, 1966. *Special issue of J. South African Inst. Min. Metal., p. 13–79*.

Meteorological Department of Sri Lanka, http://www.meteo.gov.lk/index.php?lang=en

P. Goovaerts, 1997. *Geostatistics for Natural Resources Evaluation*, Oxford University Press, 483 pages

R. **Webster**, Margaret A. **Oliver**.,(2009), *Geostatistics for environmental scientists*, John Wiley & Sons

A comprehensive R Archive network, https://cran.r-project.org

W. Luo, M. C. Taylor and S. R. Parker, 2008. *A comparison of spatial interpolation methods to estimate continuous wind speed surfaces using irregularly distributed data from England and Wales*, International Journal of climatology, 28, pp. 947-959

Zoubeida Kebaili Bargaoui., & Afef Chebbi., (2009). *Comparison of two kriging interpolation method applied to spatiotemporal rainfall*, Journal of Hydrology, 365, pp. 56-73