AN ALTERNATIVE DIAGRAMMATICAL REPRESENTATION OF WIND DATA.

FAKHRULROZI HUSSAIN, YONG ZULINA ZUBAIRI,
ABDUL GHAPOR HUSSIN.
Centre for Foundation Studies in Science,
University of Malaya
50603 Kuala Lumpur,
MALAYSIA

Abstract: -
Rose diagram is often used for diagrammatical representation of circular data. In this study, an alternative method of representation of circular data is described. Using the wind data collected at various stations by the Malaysian Meteorological Services throughout parts of Malaysia to monitor the patterns of the climatic changes in Malaysia, Wind diagram and Spoke plot are developed to visually describe the data as well as calculate the correlation value. This program has great potential in providing an alternative analysis of circular data.

Key-Words: - Diagrammatical representation, Wind data, Circular, Rose

1 Introduction

The wind data is a type of data known as the circular data. The data is measured on a scale which records speed with angular direction [3,4]. Further discussions on analyzing circular data can be found in Fisher [3] and Mardia K.V. [4]. A useful graphical representation of the circular histogram is the Rose diagram, in which the bars of the circular histogram are replaced by sectors [4]. Later on, AXIS software was introduced in 1993 which in addition to the Rose diagram, summary statistics such as circular mean, variance and the correlation between two circular data also calculated using this software [2]. Alternatively, Spoke plot could be used to represent the diagrammatical of correlation [5,6].

In this study, an alternative diagrammatical representation for circular data was developed. As opposed to the Rose diagram that does not exhibit the pattern with respect to the nature of time series data, this research applied MATLAB programming to provide an alternative graphical representation of the wind data.

This study utilizes data from Malaysian Meteorological Service which collects wind data at five different locations.

2 Representation of Circular Data

2.1 Circular Data

Circular data is one which takes values on the circumference of a circle, i.e. they are angles in the range $(0, 2\pi)$ radians or (0^0, 360^0). To analyze this type of data, we must use techniques differing from those appropriate for the usual Euclidean type variables because the circumference is a bounded closed space, for which the concept of origin is arbitrary or undefined [4,5]. Thus, the technique that has been used for continuous linear data cannot be applied for circular data. Furthermore, continuous linear data are realized on the straight line or real line which may be analyzed straightforwardly by usual statistical techniques.

2.2 Diagrammatical Representation for Circular Data in Terms of Time, Direction and Speed.

The normal representation for circular data in terms of time, direction and speed can be obtained from 3-D scatter plot as shown in Fig 1. As we can see in Fig. 1, it was difficult for us to capture pattern to be analyzed or make any conclusions of the plot. The Rose diagram, however, gives us a better visualization on circular data (Fig.2).

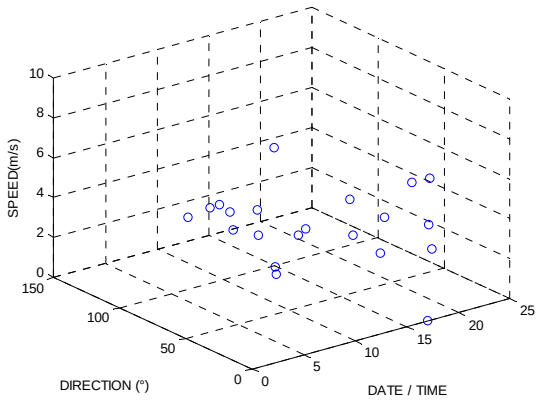

Fig 1. 3-D scatter plot for circular data in term of date, direction and speed.

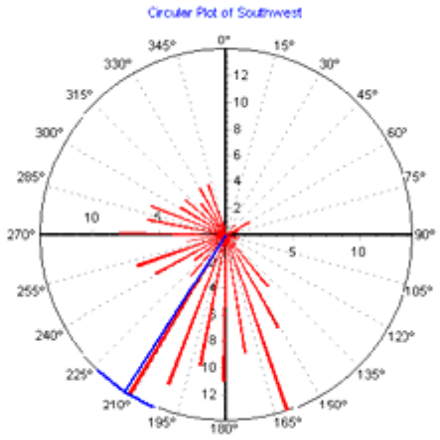

Fig 2. Rose diagram by AXIS software for Malaysian Southwest Wind Direction Data (June 2004 to September 2004), Subang Airport, 19.2 m above ground.

However, as we could see, Rose diagram in Fig. 2 represents a directional statistic in terms of frequency along that period. Each segment on the histogram is centered on its midpoint in its grouping interval. The grouping interval is determined by the bin size on the working data tab [2]. We could develop an alternative diagram that can represent circular data in terms of date, direction and speed.

2.3 Correlation between Circular Data.

An important statistic in comparing two circular data is the correlation value. Given n pairs of circular data $(\theta_1,\varphi_1),...,(\theta_n,\varphi_n)$, where $0 \le \theta_i,\varphi_i < 2\pi$ of circular variables θ and φ with mean direction μ and ν respectively. The circular correlation coefficient is defined by :

$$\hat{\rho}_T = \frac{\sum_{1\le i\le j\le n} \sin(\theta i-\theta j)\sin(\varphi i-\varphi j)}{\sqrt{\sum_{1\le i\le j\le n}\sin^2(\theta i-\theta j)\sum_{1\le i\le j\le n}\sin^2(\varphi i-\varphi j)}} \quad (1)$$

[3]

For example, correlation value between wind direction at Bayan Lepas and wind direction at KLIA. for 31 days, August 2005. (Time=1200, Pressure=500) is 0.0979 which implies a positive and weak correlation between the two locations. Spoke plot is also often be used to represent the correlation plot.

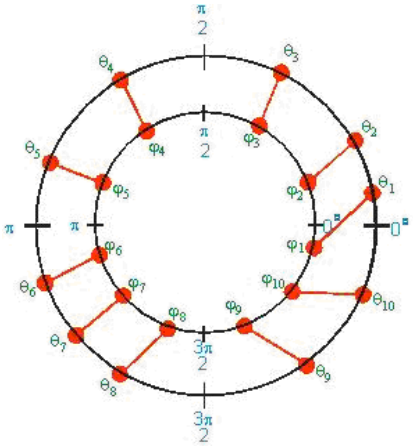

Fig 3. Example of Spoke plot for positive correlation for hypothetical data.

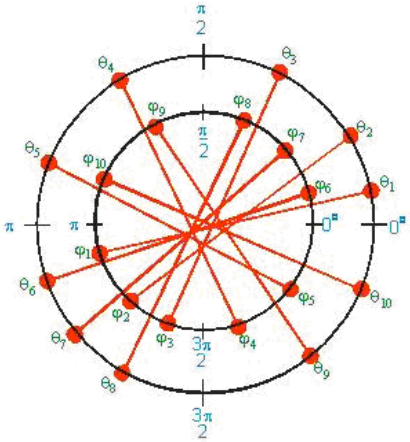

Fig 4. Example of Spoke plot for negative correlation for hypothetical data.

From Fig 3 and Fig.4, one can easily determine the sign of correlation. However, these plots have to be plotted manually and the exact correlation value is either calculated manually or using any circular software such as AXIS, Microsoff Excel and MATLAB. In this study, we further develop the Spoke plot to display both the diagrammatical

representation as well as the numerical value of the correlation.

3 Alternative Diagrammatical Representation of Circular Variables.

3.1 Wind Diagram

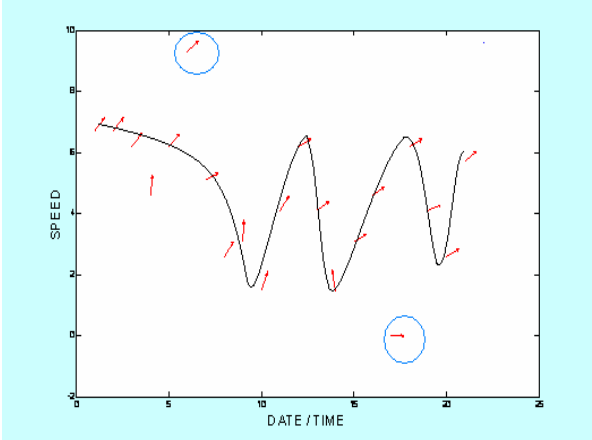

Fig.5 Wind diagram for wind direction and speed for 21 days in Bayan Lepas, January 2005. (Pressure=1000hpa, Time=0000)

Fig.5 shows a diagrammatical representation that was developed by MATLAB software for circular data in terms of time, direction and speed. Wind diagram could be an alternative method to analyze the pattern of wind direction.

3.2 Spoke Plot

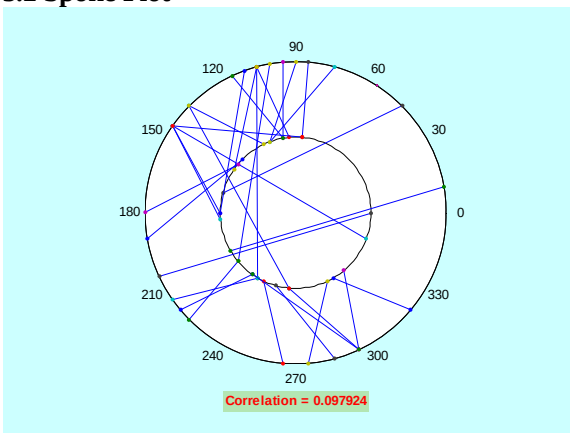

Fig. 6 Spoke plot of Correlation data for 31 days, August 2005. (Time=1200, Pressure=500)
θ_n(°) = wind direction at Bayan Lepas (inner circle).
φ_n(°) = wind direction at KLIA(outer circle).

Fig. 6 shows a diagrammatical representation for correlation between two circular data that was developed using MATLAB software. From this figure, we could see some of the lines were crossing the inner circle (implying negative correlation) and a lot of the lines were crossing in between of inner circle to the outer circle (implying positive correlation). Therefore, comparing the frequencies of the lines, it can be concluded that the correlation value is positive. To confirm the findings from the plot, the calculated correlation value of 0.0979 is shown below the diagram.

4 Conclusion

Using available software, this study provides an alternative diagrammatical representation of wind data. Wind diagram exhibits the plot for wind data in terms of time, direction and speed. Spoke plot provides a diagrammatical correlation between two circular data and calculated the value of correlation between two circular variables. This research could be improved by further developing Graphical User Interface software for ease of application by user as well as integrating into the existing AXIS software.

References:

[1] Fakhrulrozi Hussain, Yong Zulina Zubairi, Abdul Ghapor Hussin. Development of Diagrammatical Representation for Circular Variables. *Poster for Research, Invention and Innovation Exposition, University of Malaya,* 2007

[2] Abdul Ghapor Hussin, Jedzry Fadzlin Jalaludin, Ibrahim Mohamed, Analysis of Malaysian Wind Direction Data using AXIS, *Journal of Applied Science Research,* Vol.2, No. 11, 2006, pp. 1019-1021

[3] Fisher, N. I.. *Statistical analysis of circular data*. Cambridge University Press, 1993

[4] Mardia, K. V. and Jupp, P. E. *Directional Statistics*. Academic Press Inc. (London), 1999

[5] Mardia, K. V.. *Statistics of directional data*. Academic Press Inc. (London), 1972

[6] Jammalamadaka, S. Rao. *Topics in circular statistics*. World Scientific Publishing Co. Pte. Ltd., 2001