

PAPER • OPEN ACCESS

Review on circular-linear regression models

To cite this article: Hanin Hazwani Mohammad *et al* 2021 *J. Phys.: Conf. Ser.* **1988** 012108

View the [article online](#) for updates and enhancements.

You may also like

- [Large deviations for the Kardar–Parisi–Zhang equation from the Kadomtsev–Petviashvili equation](#)
Pierre Le Doussal
- [A shortcut through the Coulomb gas method for spectral linear statistics on random matrices](#)
Fabio Deelan Cunden, Paolo Facchi and Pierpaolo Vivo
- [A unified fluctuation formula for one-cut - ensembles of random matrices](#)
Fabio Deelan Cunden, Francesco Mezzadri and Pierpaolo Vivo




ECS The Electrochemical Society
Advancing solid state & electrochemical science & technology

242nd ECS Meeting

Oct 9 – 13, 2022 • Atlanta, GA, US

Presenting more than 2,400 technical abstracts in 50 symposia

 **ECS Plenary Lecture** featuring **M. Stanley Whittingham**, Binghamton University Nobel Laureate – 2019 Nobel Prize in Chemistry

 Register now!



Review on circular-linear regression models

Hanin Hazwani Mohammad¹, Siti Zanariah Satari² and Wan Nur Syahidah Wan Yusoff³

^{1,2,3}Centre for Mathematical Sciences, College of Computing and Applied Science, Universiti Malaysia Pahang, Lebuhraya Tun Razak, 26300 Gambang, Kuantan, Pahang, Malaysia.

E-mail: ¹haninhazwani@gmail.com, ²zanariah@ump.edu.my, ³wnsyahidah@ump.edu.my

Abstract. Classical linear statistics method is no longer appropriate when handling circular data since the data is influenced by direction or angle. Considering the possibility of circular data appeared as dependent variable, it has resulted in the remodeling of classic linear regression model into circular-linear regression model over the past few decades. It is important to acknowledge these circular data characteristics as it can affect the descriptive and inference of statistical analysis. With the growing body of literature regarding this issue, this paper will review on circular-linear regression model by highlighting and exploring their benefits and limitations.

1. Introduction

Circular data is a type of data that representing points on the circumference of the unit circle measured in degrees or radian [1]. Over the past century, a great emphasis of circular data is being utilized in various kind of area such as physics, psychology, meteorology, geology, paleontology, biology, astronomy and many more. Circular data can be defined in different kind of distributions such as point, uniform, cardioid, triangular, von Mises (VM), wrapped normal (WN) and wrapped Poisson (WP) distributions. The VM distribution is the most common distribution for circular data as it takes the role held by the normal distribution in standard linear statistics. In fact, it is shaped like the normal distribution, except that its tails are truncated. Unlike linear data, the origin of circular data is undefined because they exist on circumference and spherical surfaces. The difference of data characteristics can influence the descriptive and inference of statistical analysis [2]. Therefore, the classical linear statistics method is no longer appropriate when handling for circular data.

Hence, without disregarding the possibility of circular data to be appeared as dependent variable, for over the past decades, the development of circular-linear regression model has growth rapidly. However, the application of regression models proposed may be differ depending on the data generation process, data distribution, sample size, parameters involved in the link function and presence of outlier in the data set [3]. These aspects must be taken into consideration when choosing the right regression model to be applied. Therefore, the aim of this paper is to review on circular-linear regression model by highlighting and exploring their benefits and limitations. Section 2 in this paper will give an overview on the association of circular and linear variables. Next, in Section 3, this paper will give the overview of the model development for circular-linear regressions based on VM distribution. Later, this paper will discuss the benefits and limitations for other types of circular-linear regression model Section 4.



2. Linear-Circular Association

It is important to understand the difference between regression and correlation analysis. Regression is an expression of the relationship in the form of an equation. Meanwhile, correlation analysis is a method used to determine the existence of association between variables and to measure the strength of the association between them. The first step to develop a regression model is to investigate the association between linear and circular variables. Pearson Product-Moment correlation is a correlation analysis introduced by [4] in 1900. Pearson Product-Moment correlation is used to indicate the linear degree of a scatter diagram particularly for the strength and direction of a linear relationship between variables. Be mindful, it is not the measure of slope or steepness of the regression line. However, correlation analysis by [4] is specifically used for the association between linear variables only.

Consequently, the need for development of correlation analysis between linear and circular variables has caught the attention of researchers. In this section, there are several types of correlation analysis will be discussed. The first correlation coefficient for linear and circular variables is proposed by [5] in 1976 with the assumption that the association is approximately sinusoidal. The scatter plot of linear variable and circular variable should be distributed around a sine wave [5,6]. The correlation coefficient between linear and circular variables, $R_{x\theta}^2$ is defined as

$$R_{x\theta}^2 = \frac{r_{xc}^2 + r_{xs}^2 - 2r_{xc}r_{xs}r_{cs}}{1 - r_{cs}^2} \tag{1}$$

where

$$\begin{aligned} r_{xc} &= r((x_1, \dots, x_n), (\cos \theta_1, \dots, \cos \theta_n)), \\ r_{xs} &= r((x_1, \dots, x_n), (\sin \theta_1, \dots, \sin \theta_n)), \quad \text{and} \\ r_{cs} &= r((\cos \theta_1, \dots, \cos \theta_n), (\sin \theta_1, \dots, \sin \theta_n)) \end{aligned}$$

The $R_{x\theta}^2$ value is within the range of $[0, 1]$ and the higher value of $R_{x\theta}^2$ implies a stronger association between linear variable, X and circular variable, Θ . Zero value of $R_{x\theta}^2$ indicates that X is independent of Θ . The value of $R_{x\theta}^2$ is invariant under a change of scale and origin of X and change of zero or sense of direction for Θ . If X and Θ are independent and X is normally distributed, hence the population has an F distribution with 2 and $n-3$ degrees of freedom. Next, for independence test based on null hypothesis, $R_{x\theta}^2$ is used as test statistic for randomization test. Smaller p -value as compared with the n th significance value, α will emphatically reject the independency, thus there is an association between X and Θ .

In addition, correlation coefficient between linear and circular variables, $R_{x\theta}^2$ based on a rank-based analogue is also another correlation analysis developed by [5]. This rank correlation coefficient follows the Spearman's rank correlation coefficient logic by replacing the linear value, x_i and associated circular, θ_i with their ranks and uniform scores respectively. This results a less restrictive assumption for linear-circular association which there is only one maximum and one minimum value. Firstly, the data vectors, $(x_i, \theta_i), \dots, (x_n, \theta_n)$ is arranged in order of $x_1 \leq \dots \leq x_n$. Then, the uniform scores of their associated angels, θ_i are calculated. The linear rank of the angle θ_i paired with x_i that is denoted by r_i has a uniform score of $\frac{2\pi r_i}{n}$. Their ranks are replaced by x_i . The rank correlation coefficient is

based on the value of $nR_{x\theta}^2$ calculated for $\left(1, \frac{2\pi r_1}{n}\right), \dots, \left(1, \frac{2\pi r_n}{n}\right)$. Furthermore, for independence test of this correlation analysis, U_n is used and it is described as

$$U_n = \frac{24(T_c^2 + T_s^2)}{n^2(n+1)} \tag{2}$$

where

$$T_c = \sum_{j=1}^n j \cos\left(\frac{2\pi r_j}{n}\right) \text{ and } T_s = \sum_{j=1}^n j \sin\left(\frac{2\pi r_j}{n}\right)$$

If the value of U_n is larger than the percentiles of its sampling distribution under independence as in Appendix 2.17 of [1], hence U_n rejects independence. A change of origin for the linear variable or under rotation of the circular will not affect the test.

In summary, both correlation analysis introduced by [5] do not have the restriction of the nature of circular variable either it is independent or dependent variable. Nevertheless, [7] suggested that there must be a restriction for circular variable as independent variable and later presented in [8]. After all, it will not be discussed in this paper as it is not in the context of this paper. Henceforth, it is crucial to analyze the association between linear and circular variables before developing the regression model and the correlation analysis should take the nature of regression model being contemplated into account either the circular variable is independent or dependent variable.

3. Circular-Linear Regression Models based on Von Mises (VM) Distribution

The development of circular-linear regression model can be varies depending on the process of data generation which is the data distribution for circular variable and the association between linear and circular variables. There are a few approaches that can be applied for generating circular distribution such as perturbation, wrapping, transformation and projected or offsets distributions. Table 1 give the overview of some basic circular-linear regression models that using circular variable with von Mises (VM) distribution along with their benefits and limitations.

The embarking of circular-linear regression model study development are started in 1969, as [9] invented a regression model to predict circular response variable based on linear predictors. The circular data in [9] is based on von Mises distribution. Hence the model by [9] is invented particularly for biological field, specifically for the analysis of vectorcardiogram projections in which the circular variable is the time of the the interval elapsed from the start of the QRS loop to that point and the duration of the QRS loop were supplied is observed, and the linear variable is the duration of the heartbeat measured in beats per minute. The circular-linear regression model with link function, $g(x) = \beta'x$ is written as

$$\mu = \mu_0 + \sum \beta_i x_i, \tag{3}$$

where μ is the circular response variable, μ_0 is the mean level, β_i is the regression coefficient and x_i is the linear independent variable. [9] provided a coordinated method for calculating the maximum likelihood estimations (MLEs) of the model parameters which is also equivalent to least squares and offered some approximate methods of inference for large value of concentration parameter, κ [10]. Nonetheless, [11] indicated that the maximum likelihood function in [9] has infinitely many equally high peaks which can lead to ambiguously defined maximum likelihood estimations (MLEs). The infinite equally high peaks will be occurred because of the circular mean response, $\bar{\mu}$ that conditional

on linear variable, $X = x$ is actually a curve winding in an infinite number of spirals up the surface of an infinite cylinder. This is one of the various forms of the so-called ‘barber’s pole’ model that used in [9] regression model. As the weakness of the finding is also addresses by [9], the author believes that the employment of additional information such as the magnitude of the vector, hemodynamic data and the position of any particular vector on the surface of sphere as a function of the age of the individual for further research will give more elaborated and realistic models with more precise and meaningful data description. Subsequently, [11] extended [9] model for a circular-linear regression in a case of single linear predictor. The regression model is applied in meteorological field in which the circular variable is the wind direction than influenced by the wind speed as the linear variable. This proposed model is developed by using maximum entropy conditional distribution, which is a completely specified cumulative distribution function, $g(x) = 2\pi F(x)$ as the link function. The regression model is outlined as

$$f_{\Theta,R}(\theta, \tau) = 2\pi g(2\pi(F_{\Theta}(\theta) + F_R(\tau))) f_{\Theta}(\theta) f_R(\tau) \tag{4}$$

The $f_{\Theta}(\theta)$ is the density function of Θ , $F_{\Theta}(\theta)$ is the cumulative distribution function (CDF) of $f_{\Theta}(\theta)$, $f_R(\tau)$ is the density function of R and $F_R(\tau)$ is the CDF of $f_R(\tau)$. Despite of the fact that the circular-linear regression model with conditional centering is not a multiplicative model in the parameters, the response, θ completes just a single spiral as x becomes larger through its range as praised by [12]. The single spiral of the response variable is known as another different class of the so-called “barber’s pole” as compared to [9] model is a curve winding in an infinite number of spirals. However, [11] model is only covering for the case of a single linear predictor [13]. Accordingly, [12] generalized [9] model by allowing the response variable to have a concentration parameter for the case of multiple linear predictors using link function mapping the real line to the unit circle [13,14]. [12] model is used to determine the distances moved by a type of small marine snails, small blueperiwinkles, *Nodilittorina unifasciata*, after they had been transplanted down shore from the height at which they normally live. The circular-linear regression model is given as

$$\mu_i = \mu + g(\beta^T x_i) \tag{5}$$

where μ_i is the circular response variable, μ is the mean level, β^T is the k -vector of regression coefficient and x_i is the linear independent variable. The link function, $g(x) = 2 \tan^{-1}(\text{sgn}(\beta^T x) |\beta^T x|^2)$ can avoid the non-identifiability of MLEs in [9] model. [12] is taking the account of the dependence of both mean direction and dispersion on distanced moved into their model. Anyhow, the multi-modality with very narrow and sharp modes resulted by the proposed model causes a difficulty in optimizing the likelihood function [15].

In the meantime, over the years [9], [11] and [12] have become the main references for new authors to generate new method for statistical analysis such as regression models, outlier detection methods and correlation analysis. As discussed earlier, the extension of [11] and [12] models are based on the model introduced by [9] and model suggested by [11] has been extend by [12]. Next, [13] altered the model based on [9] to be applied for semiparametric Bayesian model for circular-linear regression. Furthermore, extension of [9] and [11] has produced new model based on nonnegative trigonometric sums by [16]. Meanwhile, the new models for asymmetric circular-linear regression model for environmental data by [14], inverse circular-linear and linear-circular regression by [18], application of Bayesian hidden Markov modelling by [19] are the extension based on [9], [11] and [12]. Besides, [15] presented the projected multivariate for circular-linear regression model based on [12]. Next, [9] model is used by [20] and [21] to propose new outlier detection methods such as DMCE and COVRATIO. In addition, COVRATIO statistic method also is used by [22] based on [9] and [12] models. [23] presented mean circular error (MCE) as outlier detection method based on [9] model. Later, the models by [9],

[11] and [12] are used by [24] to apply COVRATIO and [25] used DMCE. [26] developed COVRATIOs statistic for multiple circular regression model based on [9] model. Other than that, DMCE circular statistic is applied by [27] using [9] model. Lastly, [11] model is used for correlation analysis by [28].

Ultimately, [9] is the main contributor for the development of circular-linear regression model with the assumption of circular data follows the VM distribution. Eventually, with the same distribution assumption for circular variable, the development is continued by [11] and [12] with their own uniqueness modification of the link function to encounter the limitations for specific problem. To sum up, the establishment of the link function is influenced by the distribution of circular data and consequently produced multiple different approaches in developing circular-linear regression models. Moreover, the link function also can be differed from each other depending to a case of circular response with one linear predictor or with a set of linear covariates. Consequently, the model developed by [9], [10] and [12] are extended into further works for statistical analysis such as generates new models, outlier detection methods and correlation analysis.

Table 1. Review on Circular-Linear Regression Model based on Von Mises Distribution.

Data	Benefits	Limitations	Analysis Tools	References	Further Work
<p><i>Circular variable:</i> The time the particular vector is observed (the interval elapsed from the start of the QRS loop to that point and the duration of the QRS loop were supplied).</p> <p><i>Linear variable:</i> The duration of the heartbeat (in beats per minute).</p> <p><i>Circular variable:</i> Wind direction</p> <p><i>Linear variable:</i> Ozone concentration (Level of pollutant)</p>	<p>This model give an interactive method for calculating the maximum likelihood estimates (MLEs) of the model parameters and that has considered various forms of the so-called 'barber's pole' model which provided some approximate methods of inference for the case when κ can be considered large.</p> <p>The circular-linear regression model with conditional centering is not a multiplicative model in the parameters and this model causes the response completes just a single spiral as x increases through its range.</p>	<p>The maximum likelihood function has indefinitely many equally high peaks which can lead to ambiguously defined maximum likelihood estimates (MLEs).</p> <p>This model is only for the case of a single linear predictor.</p>	<p>Not mention</p> <p>Not mention</p>	<p>[9]</p> <p>[11]</p>	<p>[11,12,13,14,16,18,19] proposed new model</p> <p>[20,21,22,23,24,25,26] introduced new method for outlier detection.</p> <p>[12,14,15,16,18,19] proposed new model</p> <p>[24,25] introduced new method for outlier detection.</p> <p>[28] developed new method for correlation analysis.</p>
<p><i>Circular variable:</i> Direction of small blue periwinkles moved.</p> <p><i>Linear variable:</i> Distance of small blue periwinkle travelled.</p>	<p>Extended [9] model to the case of multiple linear predictors using link function mapping the real line to the unit circle.</p> <p>The likelihood function of the proposed model is very difficult to optimize because multi-modality is very narrow and has sharp modes and unidentifiability of parameters.</p>	<p>The likelihood function of the proposed model is very difficult to optimize because multi-modality is very narrow and has sharp modes and unidentifiability of parameters.</p>	<p>Not mention</p>	<p>[12]</p>	<p>[14,15,18,19] proposed new model.</p> <p>[22,24,25,27] introduced new method for outlier detection.</p>

4. Other Types of Circular-Linear Regression Model

Formerly, VM distribution has been known as the foundation and has been widely used for the circular-linear regression model development as originally introduced by [9]. Notwithstanding of the concept, the controversy of VM distribution is inappropriate to analyze circular data with multimodality or skewness has been raised by [8]. On top of that, the cumulative distribution function in VM distribution does not have a closed form expression. As a matter of fact, [14] brought out a concerning reality that in practice the circular distribution encountered are usually asymmetric. In that event, the development of circular-linear regression model has been extended into several new approaches such as non-parametric smoothing, multimodal distributions, asymmetric and projected or offsets distributions by other authors which is discussed in Table 2.

[13] implemented the semiparametric Bayesian approach to give flexibility for the model distributional assumptions and increases the robustness to prior misspecifications. In the light of this advantage, it is very beneficial especially when the likelihood surface is ill behaved. On contrary, the priori knowledge in the selection of starting values for the optimization algorithm is needed to locate the global maximum, however such information is not available in this case. Likewise, [16] extended [12] and [11] models for multimodal distribution by using nonnegative trigonometric sums which is known by its specialty to model data that represent multimodality or skewness in the joint density. Thus, it is applicable to a broad range of dependence patterns including from independence to a very complex structures of dependence. Other than that, [18] proposed an inverse regression which is inversely predicting the corresponding value of the independent variable after observing a value of the dependent variable. The inverse circular-linear regression model for asymmetric distribution is based on likelihood method by using the asymmetric generalized von Mises (AGVM) errors in the inverse circular linear regression problem is recommended when one targets a small mean squared errors (MSE) in the result. Lastly, [19] applied the Bayesian framework that able to overcome identifiability issues and computational problems in the classical setting of combining projected normal (PN) and Gaussian distributions. However, the models proposed by [14], [18] and [19] are only for one linear independent variable.

The model introduced by [13] has been extended into developing new model using Bayesian approach by [29] for semiparametric linear-circular regression model, [30] for basic circular-linear regression model and [31] for asymmetric multi modal regression. New outlier detection method is proposed by [32] based on [13] and [16] models. Later, [16] model are extended by [33] for non-parametric copula using wind direction data for circular-linear and linear-circular regression models, [34] for semi-wrapped gaussian and [35] for environment contour using direct sampling method for [34] circular-linear regression model. Next, [36] introduced meta-analysis for linear-circular regression model and [37] used iteration order manipulation to develop circular-linear regression model based on [18] model. Moreover, hidden Markov in [19] model is applied in [38] model and [39] model using wind direction data.

Obviously circular variable can be found in several form of distributions other than VM distribution, hence the development of circular-linear regression model is continuously has been studied and extended to statistical analysis possibilities such as developing new models and outlier detections methods. Despite of the fact that certain approaches may outstanding the other, the behavior and pattern of the circular data can affect the appropriateness of the regression model.

Table 2. Other Types of Circular-Linear Regression Model.

Data	Model Distribution	Benefits	Limitations	Analysis Tools	References	Further Work
<i>Circular variable:</i> Time of peak ozone concentration <i>Linear variable:</i> Standardized virtual temperature	Nonparametric distribution	The semiparametric Bayesian approach to give flexibility for the model distributional assumptions and increases the robustness to prior misspecifications. It is very beneficial especially when the likelihood surface is ill behaved.	The priori knowledge in the selection of starting values for the optimization algorithm is needed to locate the global maximum, however such information is not available in this case.	SAS, Maple 6, CODA	[13]	[27,28,29] proposed new model. [30] introduced new method for outlier detection.
<i>Circular variable:</i> Wind direction <i>Linear variable:</i> Ground level ozone concentration	Multimodal distribution	The nonnegative trigonometric sums are applicable to a broad range of dependence patterns including from independence to a very complex structures of dependence.	The model is for one linear independent variable only.	Not mention	[16]	[33,34,35] proposed new model. [32] introduced new method for outlier detection.
<i>Circular variable:</i> The rotational angle of flywheel <i>Linear variable:</i> The magnitude of the correction weight of flywheel	Asymmetric distribution	Likelihood method using the asymmetric generalized von Mises (AGVM) errors in the inverse circular linear regression problem is recommended when one targets a small mean squared errors (MSE) in the result.	The model is for one linear independent variable only.	Not mention	[18]	[36,37] proposed new model.
<i>Circular variable:</i> Wind direction <i>Linear variable:</i> Wind speed	Projected normal (PN) distribution	The Bayesian framework able to overcome identifiability issues and computational problems in the classical setting of combining projected normal (PN) and Gaussian distributions.	The model is for one linear independent variable only.	Not mention	[19]	[38,39] proposed new model.

5. Conclusion

In a conclusion, the regression model for circular-linear regression model has been frequently transformed over time. The statistical model development is different from one another based on the process of generation of data for circular variable, the sample size, the link function, the parameters involved in the link function development and the presence of outlier in the dataset. The most important aspect before developing the regression model is to determine the association between linear and circular variables to produce an appropriate model. The VM distribution is the most commonly distribution used for predicting circular response as it is firstly introduced by [9] into his model. As discussed, the occurrence of infinite equally high peaks is because of the curve winding in an infinite number of spirals up the surface of an infinite cylinder and the employment of additional information such as the magnitude of the vector, hemodynamic data and the position of any particular vector on the surface of sphere as a function of the age of the individual for further research will give more elaborated and realistic models with more precise and meaningful data description. [11] applied another form of the so-called “barber’s pole” which the response completes only single spiral as independent variable increases. However, this model is only suitable for one predictor only [12] introduced a regression model with multiple number of predictors and the link function can avoid the non-identifiability of MLEs in [9] model. Nevertheless, the multi-modality with very narrow and sharp modes resulted by the proposed model causes a difficulty in optimizing the likelihood function. Gradually, with the exposure of multiple possibilities that circular data in practice can be asymmetric, circular-linear regression model development has expanded to the adaptation of other type of distributions such as nonparametric, asymmetric, multimodal and normal projected (NP) or offset distributions. SAS, Maple 6 and CODA are the most popular statistical analysis tools that have been used by researchers for analysis that includes circular data. R, Python and Matlab are the other example of statistical tools that can be explored and frequently used in industrial field. For further works, there are numerous circular-linear regression models developed by authors can be extended and manipulated to generate new model based on different distribution types and approaches, apply or introduce outlier detection methods in established circular-linear regression model such as DMCE, COVRATIO and MCE and propose new correlation analysis that take the nature of circular data involved into consideration either the data is dependent or independent variables. To summarize, [11] model is the most suitable and simplest model that can be applied when the focus of the research is about other statistical inferences such as confidence intervals and outlier detection other than developing new regression model.

Acknowledgments

Authors would like to thank all the associate editors and referees for their assistance and encouragement for the improvement of this paper and acknowledged the financial support from Universiti Malaysia Pahang (UMP Internal Grant:RDU190363) and Ministry of Higher Education Malaysia (FRGS/1/2019/S7G06/UMP/0216) to conduct this study.

References

- [1] Mardia K V and Jupp P E 1999 *Directional Statistic* John Wiley Chichester
- [2] Gatto R. and Jammalamadaka S R 2015 On two-sample tests for circular data based on spacing-frequencies *Geometry Driven Statistics Wiley Series in Probability and Statistics* **121** pp 129-145
- [3] Sikaroudi A E and Park C 2019 A mixture of linear-linear regression models for a linear-circular regression *Statistical Modelling*
- [4] Pearson K 1900 I. Mathematical Contributions to the Theory of Evolution.—VII. On the Correlation of Characters not Quantitatively Measurable *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character* **195**(262-273) pp 1-47.

- [5] Mardia K V 1976 Linear-Circular Correlation Coefficients and Rhythmetry *Biometrika* **63** pp 403-405
- [6] Johnson R A and Wehrly T E 1977 Measures and Models for Angular Correlation and Angular-Linear Correlation *Journal of the Royal Statistical Society: Series B (Methodological)* **39**(2) pp 222-229
- [7] Fisher N I and Lee A J 1981 Nonparametric Measures of Angular-Linear Association *Biometrika* **68** pp 629-639
- [8] Fisher N I 1993 *Statistical Analysis of Circular Data* Cambridge University Press Cambridge
- [9] Gould A L 1969 A Regression Technique for Angular Variates *Biometrics* pp 683-700
- [10] Laylock P J 1975 Optimal Design: Regression Models for Directions *Biometrika* **62** pp 305-3011
- [11] Johnson R A and Wehrly T E 1978 Some Angular-Linear Distributions and Related Regression Models *Journal of the American Statistical Association* **73**(363) pp 602-606
- [12] Fisher N I and Lee A J 1992 Regression Models for an Angular Response *Biometrics* **48** pp 665-677
- [13] George B J and Ghosh K 2006 A semiparametric Bayesian model for circular-linear regression *Communications in Statistics-Simulation and Computation* **35**(4) pp 911-923
- [14] SenGupta A and Ugwuowo F I 2006 Asymmetric Circular-Linear Multivariate Regression Models with Applications to Environmental Data *Environmental and Ecological Statistics* **13**(3) pp 299-309
- [15] Presnell B Morrison S P and Little R C 1998 Projected Multivariate Linear Models for Directional Data *Journal of the American Statistical Association* **93**(443) pp 1068-1077
- [16] Fernández-Durán J J 2007 Models for circular-linear and circular-circular data constructed from circular distributions based on nonnegative trigonometric sums *Biometrics* **63**(2) pp 579-585
- [17] Wehrly T E and Johnson R A 1980 Bivariate models for dependence of angular observations and a related Markov process *Biometrika* **67**(1) pp 255-256
- [18] Kim S and SenGupta A 2015 Inverse circular-linear/linear-circular regression *Communications in Statistics-Theory and Methods* **44**(22) pp 4772-4782
- [19] Mastrantonio G Maruotti A and Jona-Lasinio G 2015 Bayesian hidden Markov modelling using circular-linear general projected normal distribution *Environmetrics* **26**(2) pp 145-158
- [20] Rambli A Mohamed I Abuzaid A H and Hussin A G 2010 Identification of influential observations in circular regression model *Proceedings of the Regional Conference on Statistical Sciences (RCSS'10)* pp 195-203
- [21] Abuzaid A Hussin A G and Rambli A 2011 COVRATIO statistic for simple circular regression model
- [22] Ibrahim S Rambli A Hussin A G and Mohamed I 2013 Outlier detection in a circular regression model using COVRATIO statistic *Communications in Statistics-Simulation and Computation* **42**(10) pp 2272-2280
- [23] Abuzaid A H Hussin A G and Mohamed I B 2013 Detection of outliers in simple circular regression models using the mean circular error statistic *Journal of Statistical Computation and Simulation* **83**(2) pp 269-277
- [24] Rambli A Yunus R M Mohamed I and Hussin A G 2015 Outlier detection in a circular regression model. *Sains Malaysiana* **44**(7) pp 1027-1032
- [25] Rambli A Abuzaid A H Mohamed I B and Hussin A G 2016 Procedure for detecting outliers in a circular regression model *PLoS One* **11**(4)
- [26] Alkasadi N A Ibrahim S Ramli M F and Yusoff M I 2016 A comparative study of outlier detection procedures in multiple circular regression *AIP Conference Proceedings* **1775**(1) p 30032
- [27] Ibrahim S 2013 Some outlier problems in a circular regression model *Doctoral dissertation University of Malaya*

- [28] Jupp P E and Mardia K V 1980 A general correlation coefficient for directional data and related regression problems *Biometrika* **67**(1) pp 163-173
- [29] Bhattacharya S and Sengupta A 2009 Bayesian analysis of semiparametric linear-circular models *Journal of agricultural, biological, and environmental statistics* **14**(1) 33-65
- [30] Gill J and Hangartner D 2009 Bayesian circular-linear regression *Seminar on Bayesian Inference in Econometrics and Statistics*
- [31] Kılıç M B 2015 Bayesian modelling for asymmetric multi-modal circular data
- [32] Nurul Hidayah S 2018 Outlier detection in cylindrical data *Doctoral dissertation University of Malaya*
- [33] Carnicero J A Ausín M C and Wiper M P 2013 Non-parametric copulas for circular-linear and circular-circular data: an application to wind directions *Stochastic Environmental Research and Risk Assessment* **27**(8) pp 1991-2002
- [34] Roy A Parui S K and Roy U 2016 SWGMM: a semi-wrapped Gaussian mixture model for clustering of circular-linear data *Pattern Analysis and Applications* **19**(3) pp 631-645
- [35] Vanem E Hafver A and Nalvarte G 2020 Environmental contours for circular-linear variables based on the direct sampling method *Wind Energy* **23**(3) pp 563-574
- [36] Kim S and Peiris T B 2019 Meta analysis of regression: a review and new approach with application to linear-circular regression model *Communications in Statistics-Theory and Methods* pp 1-9
- [37] Nurhab M I Nurhab B Purwaningsih T and Teng M F 2017 Circular (2)-linear regression analysis with iteration order manipulation *International Journal of Advances in Intelligent Informatics* **3**(2) pp 107-116
- [38] Lagona F Picone M and Maruotti A 2015 A hidden Markov model for the analysis of cylindrical time series. *Environmetrics* **26**(8) 534-544
- [39] Mastrantonio G and Calise G 2016 Hidden Markov model for discrete circular-linear wind data time series *Journal of Statistical Computation and Simulation* **86**(13) pp 2611-2624