Vol.2.Issue.3.2014

http://www.bomsr.com

BULLETIN OF MATHEMATICS AND STATISTICS RESEARCH

A Peer Reviewed International Research Journal

ISSN:2348-0580

RESEARCH ARTICLE

RELATIVE PERFORMANCE OF CIRCULAR MODELS IN FITTING MOVEMENTS OF HUNDRED ANTS DATA

A.J.V. RADHIKA[1], S.V.S. GIRIJA[2], A.V. DATTATREYA RAO[3]

[1]Assistant Professor of Mathematics, University College of Engineering and Technology, Acharya Nagarjuna University, Guntur, India

[2]Associate Professor of Mathematics, Hindu College, Guntur, India.

[3]Professor of Statistics, Acharya Nagarjuna University, Guntur, India

Article Info:
Article received :15/09/2014
Revised on:27/09/2014
Accepted on:30/09/2014

ABSTRACT

Hitherto method of wrapping is frequently used in constructing circular models. On those lines Dattatreya Rao et al (2007) derived a good number of wrapped circular models. A new method of generation of circular models by using the Rising Sun function was developed by Girija (2010). Offsetting and inverse stereographic projection are the two important methods of constructing a circular model. Applying these techniques Stereographic circular models (Phani (2013)) and Offset circular models (Radhika (2014)) have been constructed and they provide a rich and very useful class of models for circular as well as *l*-axial data. Movements of hundred ants data (available in Fisher (1993) p.243) is tried to verify whether some of these circular models will be good fits or not, in order to establish their applicability in practical live cases. Here, Offset Pearson Type II, Stereographic Reflected Gamma, Stereographic Logistic, Stereographic Double Exponential and Wrapped Logistic models are found to be good fits for the above data.

In case if more than one model is found to be good fit for a particular data set, as in this case one has to suggest, choosing the best fit in order to select the most appropriate model for that situation and these aspects are discussed here.

INTRODUCTION

The applications of the proposed circular models are based on the fitting of the new circular models to the live data available in Fisher (1993). The fit of various parametric circular models has been compared with proposed models using likelihood based model selection criteria viz., Maximized log-likelihood (*MLL*), Akaike's Information Criteria (*AIC*) and Bayesian Information Criteria (*BIC*). We apply various goodness of fit tests such as **Rayleigh Test, Kuiper's Test, Watson's U^2 - Test, Hodges –Ajne Test, Range Test, Rao's Equal Spacings and Ajne's test** to find whether the data follows the given distribution.

The main focus of this work is to study goodness-of-fit with respect to the parametric models for certain live data sets. Further to choose the most appropriate model by examining among the models which are found to be good fits.

Section 2 is devoted to specify live data set and to describe the given data by means of simple data plot. Section 3 deals with the nonparametric estimation of mean direction and concentration parameter of the parametric models under consideration. Probability plot curves will give visual impressions whether the specified parametric model fits well to the given data or not. However available rigorous test procedures are adopted for verifying goodness-of-fit of the model for the given data. The methodological aspects of the said goodness-of-fit procedures are given in section 4 while the implementation aspects are presented in section 5. Five circular models viz., Stereographic versions of Reflected Gamma, Logistic, Double Exponential, Offset Pearson Type II and Wrapped Logistic models are found to be good fits for movements of ants data set of size n = 100 and these are described in section 6. Obviously, the logical question that arises is to decide which of the circular models mentioned above is to be selected as most appropriate for further data analysis. This resolutional aspect is discussed in section 7.

LIVE DATA SET

To analyze circular data, the following live data set is considered.

Data Set : Movements of Hundred Ants Data [Fisher (1993) p.243]

Directions chosen by 100 ants in response to an evenly illuminated black target placed.

Direction (in degrees)

330 290 60 200 200 180 280 220 190 180 180 160 280 180 170 190 180 140 150 150 160 200 190 250 180 30 200 180 200 350 200 180 120 200 210 130 30 210 200 230 180 160 210 190 180 230 50 150 210 180 190 210 220 200 60 260 110 180 220 170 10 220 180 210 170 90 160 180 170 200 160 180 120 150 300 190 220 160 70 190 110 270 180 200 180 140 360 150 160 170 140 40 300 80 210 200 170 200 210 190

The data plot is shown in figure 1

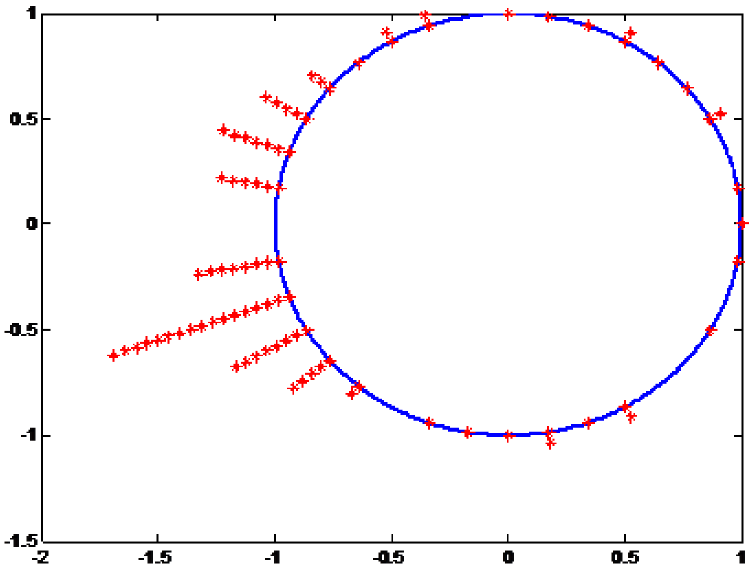

Fig. 1 Graph of 100 ants data plot

ESTIMATION OF PARAMETERS

In order to fit any model for given data, one has to estimate the parameters of the model as a pre processing step. Any circular model is characterized by mean direction and concentration parameter. The pair of parameters can be estimated using several methods available in the literature of statistics. Here we perform the estimation of the parameters using the following nonparametric method.

Let $\theta_1, \theta_2, \theta_3, ..., \theta_n$ be a set of circular observations.

The mean direction is

$$\hat{\mu} = \begin{cases} \tan^{-1}(S/C) & \text{if } S>0,\ C>0 \\ \tan^{-1}(S/C)+\pi & \text{if } C<0 \\ \tan^{-1}(S/C)+2\pi & \text{if } S<0,\ C>0 \end{cases} \qquad (3.1)$$

and the mean resultant length (estimate of the concentration parameter) is $\hat{\sigma} = \sqrt{C^2 + S^2}$,

where $C = \frac{1}{n}\sum \cos\theta_i$ and $S = \frac{1}{n}\sum \sin\theta_i$

To demonstrate the modeling behavior of Stereographic Reflected Gamma and Offset Pearson Type II distributions, the index parameter *c* and other parameter ρ respectively are to be estimated. Their ML Estimates are computed by invoking fmincon, the built-in MATLAB function.

Table 1 Estimates for Data Set

Sample Size	$\hat{\mu}$	$\hat{\sigma}$	$\hat{c}$	$\hat{\rho}$
Data Set(*n* =100)	-3.0868 $\mu \in (-\pi, \pi)$ 3.1964 $\mu \in [0, 2\pi)$	0.6101	0.8042	0.0176

The cut off points for these data sets are computed on the lines of algorithm from Devaraaj (2012) for circular models.

METHODOLOGIES OF GOODNESS –OF- FIT

Because of central role played by circular Uniform distribution, one of the most important hypotheses about a distribution on the circle is that of Uniformity. The important test procedures of Uniformity are reviewed here. Substituting data set in the commutative distribution function (cdf) of the said model, corresponding Uniform $[-\pi, \pi)$ variates denoted by $\theta_1, \theta_2, \theta_3, ...\theta_n$ are obtained. Using these θ_i's , the test statistics of Rayleigh test , Kuiper's test, Watson's U^2 –test , Hodges-Ajne's test, Range's test, Rao's equal spacing test and Ajne test are computed and these tests are furnished in Mardia and Jupp (2000).

CIRCULAR MODELS FOR GOODNESS –OF- FIT

To verify goodness – of - fit, we consider the following five circular models

i) Offset Pearson Type II Model [Radhika et al (2013)]
ii) Stereographic Reflected Gamma Model [Radhika (2014)]
iii) Stereographic Logistic Model [Dattatreya Rao et al. (2011)]
iv) Stereographic Double Exponential Model and [Phani (2013)]
v) Wrapped Logistic Model [Dattatreya Rao et al. (2007)].

For the purpose of verifying goodness of fit the Movements of Ants Data Set is considered.

We consider Offset Pearson Type II, Stereographic Reflected Gamma, Stereographic Logistic , Stereographic Double Exponential and Wrapped Logistic models to verify goodness-of-fit for movements of ants data of size *n* = 100. As this is a **large sample**, we apply Range, Kuiper's and modified Watson's U^2 tests. The statistic of Watson's U^2 test for large sample [Mardia and Jupp (2000), p. 105] is

$$W^2 = \frac{V_n^{\,2}}{\pi^2}$$ where V_n the statistic of Kuiper's described in Mardia and Jupp (2000).

Uniform probability plots are drawn for the given data by implementing the graphical method of goodness-of-fit and are presented below.

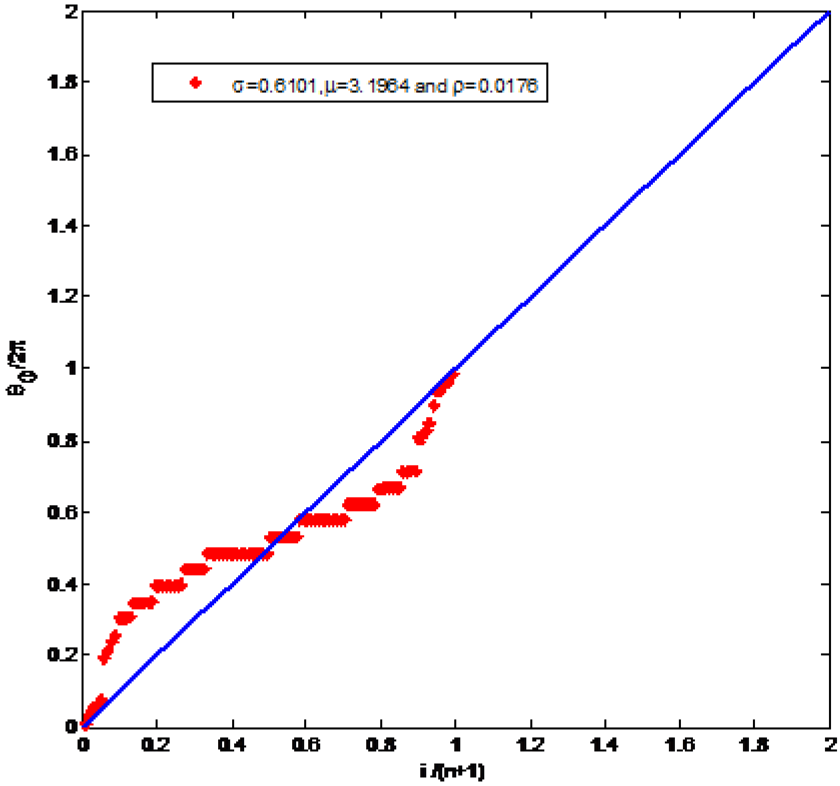

Fig. 2 Graphical Assessment of the Offset Pearson Type II model to movements of ants data for n=100

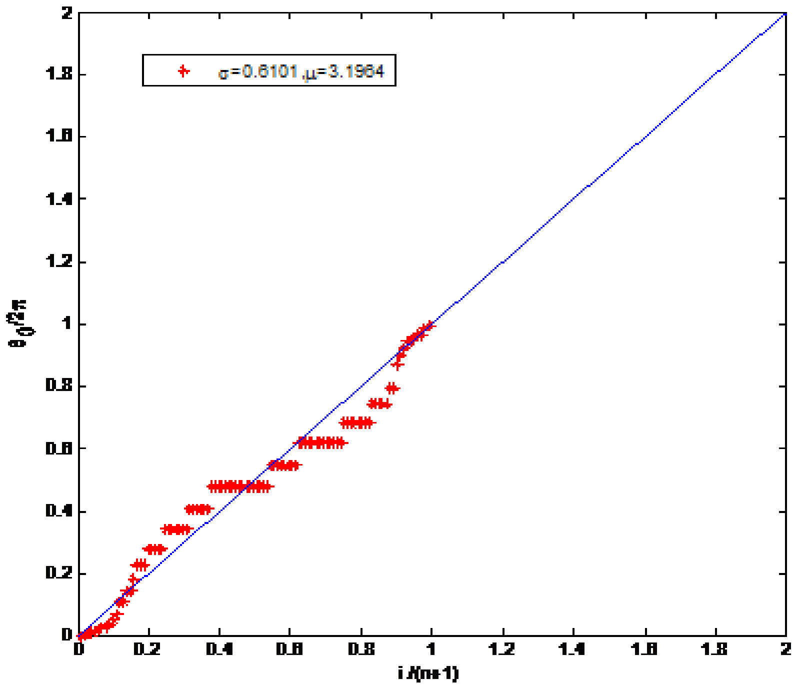

Fig. 3 Graphical Assessment of the Wrapped Logistic model to movements of ants data for n=100

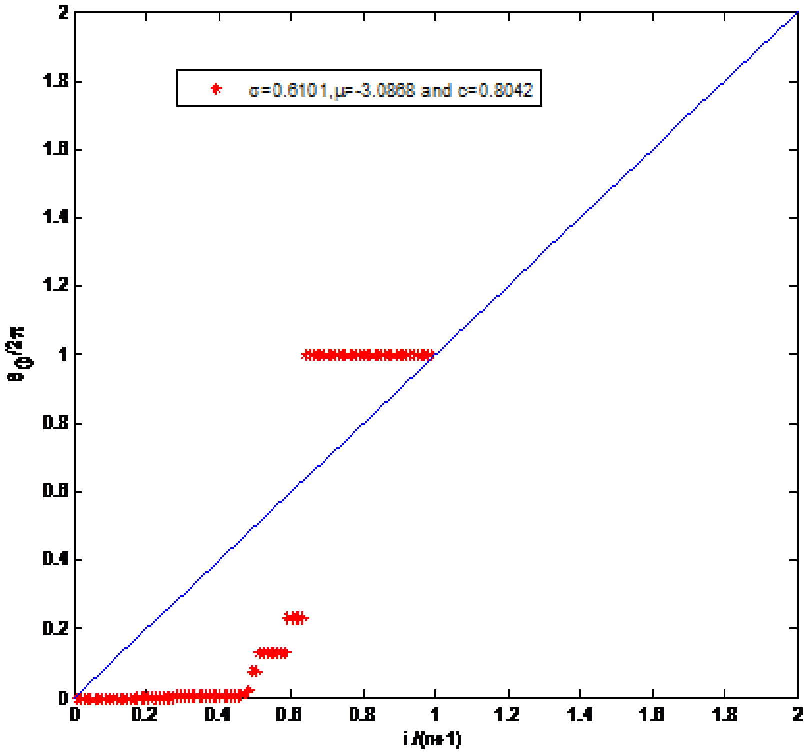

Fig. 4 Graphical Assessment of the Stereographic Reflected Gamma model to movements of ants data for n=100

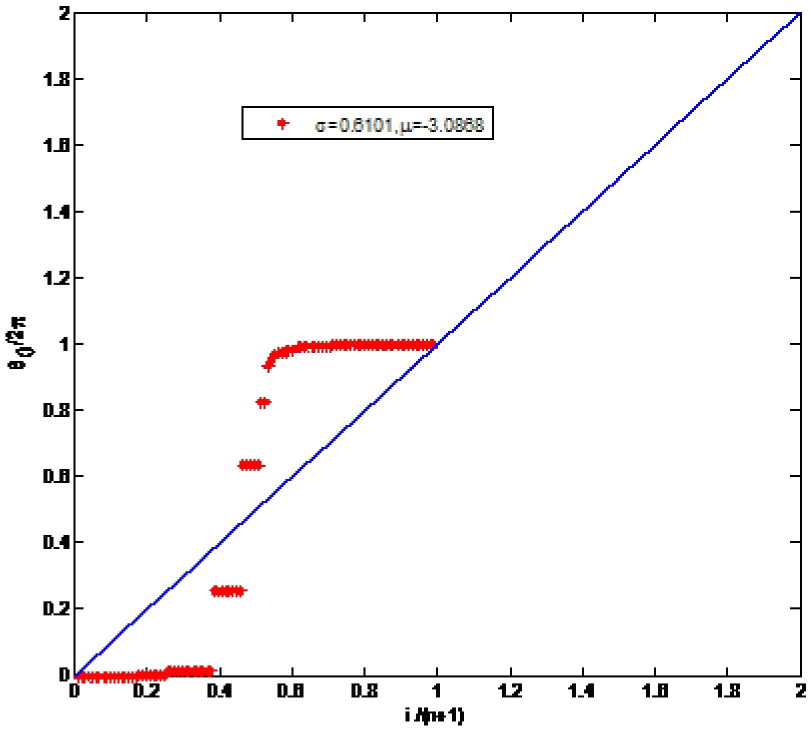

Fig. 5 Graphical Assessment of the Stereographic Logistic model to movements of ants data for n=100

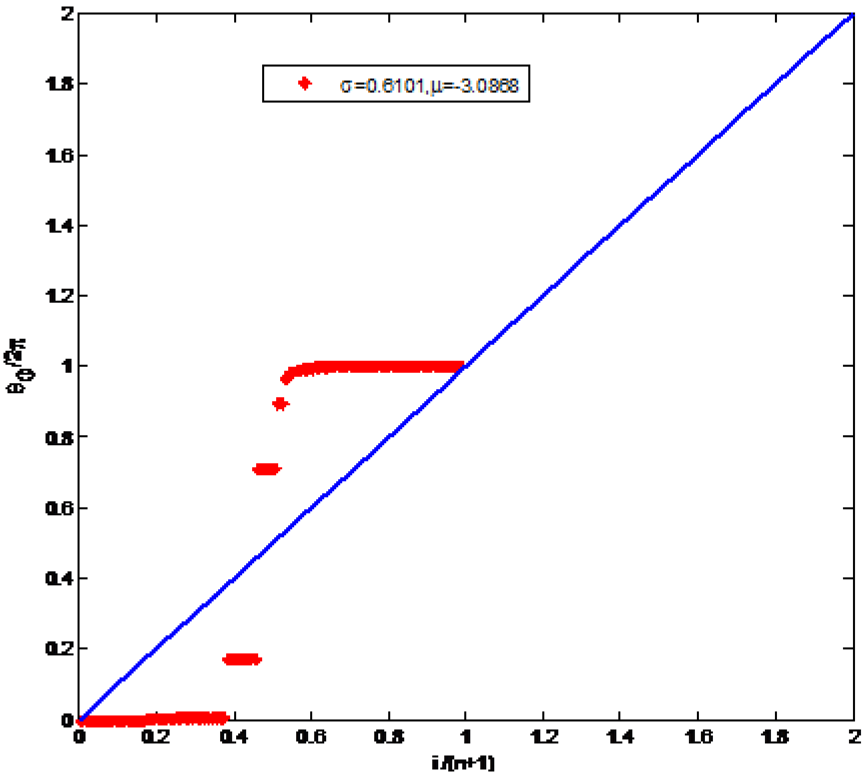

Fig. 6 Graphical Assessment of the Stereographic Double Exponential model to movements of ants data for n=100

From the above graphs, it is observed that the data probably follows the five circular models viz., Offset Pearson Type II, Wrapped Logistic, Stereographic Reflected Gamma, Stereographic Logistic and Stereographic Double Exponential models with estimates of mean direction, concentration parameter and index parameter $\widehat{\mu}= -3.0868\ (3.1964),\ \widehat{\sigma}=0.6101\ \hat{c}=0.8042$ .and $\hat{\rho}=0.0176.$

The statistics of the three goodness-of-fit tests viz., Range, Kuiper's and Watson's U^2 tests are computed for the said circular models and are tabulated.

Table 2 Statistics of the Range, Kuiper's and Watson's U^2 tests of goodness-of-fit tests

Sample Size *n* =100	Offset Pearson model	Wrapped Logistic model	Stereographic Reflected Gamma model	Stereographic Logistic model	Stereographic Double Exponential model
Kuiper's Test	4.7050	2.5887	8.4098	7.9811	8.2945
Watson's U^2 Test	0.0612	0.0066	0.0692	0.0623	0.0673
Range Test	5.4168	5.7878 1,5	1.4895	3.9094	1.8187

On the lines of algorithm in Devaraaj (2012) the cut off points for the data set of sample size *n* = 100 are computed using MATLAB techniques.

Table 3 The cut of points for the data set of sample size n = 100

LOS / Tests	1%	5%	10%
Rayleigh Test	0.0106 - 10.3886	0.0464 - 7.4802	0.1035 - 5.9778
Kuiper's Test	0.6987 - 2.1225	0.7924 - 1.8348	0.8474 - 1.7437
Watson's U^2 - Test	0.0163 - 0.2976	0.0231 - 0.2231	0.0270 - 0.1888
Hodges –Ajne Test	3.8000 - 7	4.2000 - 6.6000	4.4000 - 6.4000
Range Test	5.7003 - 6.0818	5.7839 - 6.0616	5.8193 - 6.0515
Rao's Equal Spacing Test	1.9084 - 2.7230	2.0024 - 2.6024	2.0436 - 2.5601
Ajne Test	0.0264 - 1.0959	0.0406 - 0.7920	0.0490 - 0.6564

From the above two tables it is identified that the Data Set follows all the five circular models at different levels of significance by some of the tests of Uniformity.

CHOOSING THE BEST MODEL

For a given data, in a case more than one parametric circular model fits well, then choosing the appropriate model which is the best fit, can be decided by the criteria of MLL, AIC and BIC.

Now, the task from data analysis point of view is to decide the most appropriate circular model. Here this issue is resolved by applying the criteria of AIC, BIC and MLL. The computations of measures of relative performances are presented in table 4.

Table 4 Measures of Relative Performance for Goodness–of-fit at $\widehat{\mu} = -3.0868\ (3.1964),\ \widehat{\sigma} = 0.6101$ $\hat{c} = 0.8042$ **.and** $\hat{\rho} = 0.0176$. **for Data Set**

Distribution	Parameters	*MLL*	*AIC*	*BIC*
Offset Pearson Type II distribution	$\widehat{\mu} = 3.1964$ $\widehat{\sigma} = 0.6101$ $\hat{\rho} = 0.0176$	-183.7642	187.7642	376.7387
Wrapped Logistic distribution	$\widehat{\mu} = 3.1964$ $\widehat{\sigma} = 0.6101$	**-143.9360**	**147.9360**	**297.0823**
Stereographic Reflected Gamma	$\widehat{\mu} = -3.0868,$ $\widehat{\sigma} = 0.6101$ $\hat{c} = 0.8042$	-220.3863	224.3863	449.9829
Stereographic Logistic distribution	$\widehat{\mu} = -3.0868,$ $\widehat{\sigma} = 0.6101$	-246.2138	250.2138	501.6379
Stereographic Double Exponential model	$\widehat{\mu} = -3.0868,$ $\widehat{\sigma} = 0.6101$	-224.2926	228.2926	457.7955

On the basis of AIC, BIC and MLL we identify that **Wrapped Logistic model** is the superior fit than Offset Pearson Type II, Stereographic Reflected Gamma, Stereographic Logistic and Stereographic Double Exponential models for $\widehat{\mu} = 3.1964$, $\hat{\sigma} = 0.6101$. Also it is observed that the statistic of Watson's U^2 test for Wrapped Logistic model at $\widehat{\mu} = 3.1964$, $\hat{\sigma} = 0.6101$ is the smallest.

REFERENCES

[1]. Dattatreya Rao, A.V., Ramabhadra Sarma, I. and Girija, S.V.S. (2007). *On Wrapped Version of Some Life Testing Models*, Comm Statist., - Theor.Meth. 36, 11, pp.2027-2035.

[2]. Dattatreya Rao, A.V., Girija, S.V.S. and V.J. Devaraaj, (2009). *Tests for Circular Uniformity*, Journal contemplated by ANU, 151-164.

[3]. Dattatreya Rao A.V., Girija S.V.S. and Phani, Y. (2011a). On Stereographic Logistic Model, Proceedings of NCAMES, AU Engineering College, Visakhapatnam, pp. 139 – 141.

[4]. Dattatreya Rao A.V., Girija S.V.S. and Phani, Y. (2011b). Differential Approach to Cardioid Distribution, Computer Engineering and Intelligen Systems, ISSN 2222-1719, Vol 2, No.8, pp.1-6.

[5]. Dattatreya Rao, A.V., Ramabhadra Sarma, I., and Girija, S.V.S., (2012). *New Wrapped Distributions – Goodness of Fit*, Proceedings of STATPARCONF, Kolkata.

[6]. Devaraaj V. J. (2012), Some Contributions to Circular Statistics, Thesis submitted to Acharya Nagarjuna University for the award of Ph. D.

[7]. Fisher, N. I. (1993). Statistical Analysis of Circular Data. Cambridge University Press, Cambridge.

[8]. Girija, S.V.S. (2010).Construction of New Circular Models, VDM VERLAG, Germany. ISBN 978-3-639-27939-9.

[9]. Girija S.V.S., Phani, Y. and Dattatreya Rao A.V., (Sep. 2013). *On Stereographic Lognormal Distribution*, International Journal of Advances in Applied Sciences (IJAAS), Vol 2 No. 3.

[10]. Jammalamadaka, S.R. and Sen Gupta, A. (2001). Topics in Circular Statistics, *World Scientific Press, Singapore.*

[11]. Mardia, K.V. and Jupp, P.E. (2000). Directional Statistics, John Wiley, Chic ester.

[12]. Phani, Y., Girija S.V.S. and Dattatreya Rao A.V., (2012). *Circular Model Induced by Inverse Stereographic Projection On Extreme-Value Distribution,* IRACST – Engineering Science and Technology: An International Journal (ESTIJ), ISSN: 2250-3498,Vol.2, No. 5, pp. 881 – 888.

[13]. Phani, Y., Girija S.V.S. and Dattatreya Rao A.V., (2013). *On Construction of Stereographic Semicircular Models*, Journal of Applied Probability and Statistics, Vol. 8, no. 1, pp. 75-90.

[14]. Phani, Y. (2013). On Stereographic Circular and Semicircular Models, Thesis submitted to Acharya Nagarjuna University for the award of Ph.D.

[15]. Radhika A J V, "Mathematical Tools in the Construction of New Circular Models", unpublished thesis submitted to Acharya Nagarjuna University for the award of Ph.D.

[16]. Van Rooij, A.C.M. and Schikof, W.H. (1982). A Second Course on Real Functions, *Cambridge University Press, Cambridge*.