IOSR Journal of Mathematics (IOSR-JM)
e-ISSN: 2278-5728, p-ISSN: 2319-765X. Volume 12, Issue 2 Ver. II (Mar. - Apr. 2016), PP 56-65
www.iosrjournals.org

Application Of Multiple Circular-Linear Regression Models To Animal Movement Data With Covariates.

Dr Robert Mathenge Mutwiri (Ph.D.)

School of pure and applied sciences, Kiriinyaga University College, Kenya and School of Mathematics, Statistics and Computer science, University of Kwazulu Natal, South Africa, P.O Box 143 – 10300Kerugoya, Kenya.

Abstract: *In a many biological and physical sciences studies, a set of techniques have been developed to analyse the relationship between the circular and linear data derived from the geographical positioning system (GPS) telemetry to describe animal movement. Yet, many of the models used by ecologists do not provide a link between the circular and linear variables. This chapter demonstrates the application of the circular-linear regression in describing such a relationship. We describe numerical methods of obtaining maximum likelihood model parameter estimates. We discuss the technical limitations of the model through simulation and application to real elephant movement data with covariates collected from Kruger national park, South Africa. These results provide a new statistical paradigm for understanding the need to landscape features in elephant and similar animal models and evolutionary forces driving unpredictable.*

Keywords: *Animal movement, circular statistics, von Mises, regression model, GPS tracking, turn angles*

I. Introduction

Advances in GPS tracking technology are revealing new insights regarding animal movements (Ropert-Coudert, and Wilson, 2005). Tracking devices tagged on animals are becoming smaller yet larger in memory capacity re yielding huge datasets (Urbano et al., 2010). GPS metrics (circular or linear) of animal movement can be derived from the relocation data. A huge body of literature exists on the analysis of linear metrics (step length, speed and mean square displacement) (Edwards et al., 2007) but little is known about the circular metrics due to lack of computational softwares. Accurate and objective interpretation of biological data on movement patterns of animals is needed because conclusions may have management implications (Owen and Smith, 2010). For instance, the analysis of turn angles data can reveal animal space use and the ecosystem drivers of movement (Duffy et al., 2011), and efficient foraging strategies (Viswanathan et al., 2005) that can facilitate applications in conservation ecology (Duffy et al., 2011).

Circular data can also be visualized as being distributed on the circumference of a unit circle in the range of 0 to 2π (Mardia and Jupp, 2009). Handling such data creates difficulties due to the restriction of support to the unit circle, the sensitivity of descriptive and inferential statistics to the starting point on the circle. There exists a substantial literature on circular data but broadly, it is confined to descriptive statistics and limited to inference for simple univariate models Batschelet1981. For instance, Cain (1989) reviewed the statistical methods for analysing angular ecological data and found that standard statistical methods were not appropriate for circular data.

Many ecological models have relied on circular statistics to develop complex models. For instant the authors in Siniff et al., (1969) and Tracey et al., (2005) fitted von Mises distribution to turn(move) angle data and come up with a set of non-linear regression models relating animal movement to landscape features orientation. Cain(1985) found that von Mises distribution provides good statistical fit to insect turn angles and clonal plant branching angles. According to behavioural and foraging theory, magnitude of the turn angle and the step lengths depends on fundamental movement elements (Getz and Wayne, 2008) such as habitat type, distance to the river (water point), landscape structure, quality of food patches etc. Elephants locates their home ranges in areas of high food density, water source and low human disturbance (De Beerand Van Aarde, 2008) while (Shannon et al., 2009) argued that the location of artificial water points can have negative impact on the bio-diversity by putting pressure on habitats closure to them. Landscape features are physical barriers such as rivers and roads which constrain animal movement by preventing access to adjacent patches and impeding dispersal (Vanak et al., 2010). Redfern et al., (2003) found correlation between distance to water and elephant distribution across the poor and rich forage habitat patches in KNP. To model the effects of landscape feature on animal turning behaviour a tool that relates the circular and linear variables is needed.

We apply circular statistics, a specific area of methods for analysing data arising from animal movement studies. The primary motivation of our approach arises from the rich data set of elephant turning angle and landscape features covariates data of an individual elephant collected from Kruger National park using GIS telemetry tags. The turn angles of the animal are derived from the location and the covariates extracted. Using a circular-linear regression model Fisher and Lee (1992), we model the turn angles of the elephant movement as a function of landscape and climatic covariates. Using this model, we can determine the mean orientation of the elephant and identify important predictors of animal orientation.

This paper is organized as follows. Section 2, we describe the von Mises distribution and its properties. Section 3 deals with the von Mises regression model, parameter estimation, confidence intervals and model checking methods. In Section 4 we discuss the limitations of the model based on simulated data and real elephant movement data with covariates. In Section 6 we present the results of the fitted model and there description. Finally we end the chapter with discuss and conclusion of the findings.

II. The Von Mises Distribution

One of the most important distributions for circular data is the von Mises distribution, represented by $vm(\mu,\kappa)$, whose probability density function (pdf) is given

$$f(\theta) = \frac{1}{2\pi I_0(\kappa)} exp\{\kappa cos(\theta-\mu)\} \tag{1}$$

where $[2\pi I_0(\kappa)]$ is the normalizing constant, $I_0(\kappa)$ is the modified Bessel function of the first kind and order zero. The parameter μ represents the mean direction while κ is the concentration parameter of the distribution (Abramowitz and Stegun, 1972). These parameters must satisfy $\mu \in [-\pi,\pi)$ and $\kappa \in [-\infty,\infty)$.

The general modified Bessel function of the first kind and order p is defined by

$$I_O(\kappa) = \frac{1}{2\pi}\int_0^{2\pi} exp[\kappa cos(\phi-\mu)]d\phi$$

Note that $vm(\mu,\kappa)$, and $vm(\mu+\pi,-\kappa)$, have the same distribution. For this model, the values of κ are set to be non-negative, and the ranges of θ and μ are $[-\pi,\pi]$. The von Mises distribution is unimodal with two parameters κ and μ, and is symmetrical about mean direction $\theta=\mu$. The larger the value of concentration parameter κ, the denser the clustering around the mean direction μ. For κ=0, the von Mises distribution tends to the uniform distribution. As $\kappa\to\infty$ it becomes concentrated at the point $\theta=\mu_o$.

From a random sample $\theta_1,\theta_2,\ldots,\theta_n$, we can calculate C and S according to equation (1) so that the maximum likelihood estimate of κ is given by the solution from

$$A(\kappa) = (C^2+S^2)^{1/2}$$

where the function $A(k)=I_1(k)/I_0(k)$ is defined with $I_1(k)$, the modified Bessel function of the first kind (order 1; Abramowitz and Stegun, 1972). Meaning that we need

$$\kappa = A^{-1}(C^2+S^2)^{1/2}$$

This is sufficiently unwieldy that tabular lookup was used historically and we are now advantaged by software (Gill and Hangartner 2010).

Unfortunately, the MLE here is also biased for finite sample (Upton, 1986). Defining

$$R^2 = [nI_1(\kappa)/I_0(\kappa)]^2$$

Schou(1978) and Batschelet (1981) tabulated unbiased values of $\hat{k}$, which is the solution for $A(\kappa)=RA(R\kappa)/n$ when $R^2 \ge n$. These authors recommends that $\hat{k}$=0 when $\hat{R}$<n. A very useful set of approximations is given by Fisher (1993) given by 1

$$\kappa = \begin{cases} 2(R/n)+(R/n)^3+5(R/n)/6 & \text{for } R<0.53n, \\ -4+1.39(R/n)+0.43(1-R/n) & \text{for } 0.53n \le R \prec 0.85n, \\ 1/((R/n)^3-4(R/n)+3(R/n) & \text{for } R \ge 0.85n \end{cases}$$

which needs to be adjusted for small sample size and small R/n (Upton, 1986):

$$\hat{\kappa} = \begin{cases} \max\left(\hat{\kappa}-\frac{2}{n\kappa}\right) & \text{for } \hat{\kappa}<2, \\ \frac{(n-1)^3\hat{\kappa}}{n^3+n} & \text{for } \hat{\kappa}\ge 2 \end{cases} \tag{2}$$

The MLE for the mean direction is the value m̂ that satisfies the equation $\hat{\mu}$=C/R, $\hat{\mu}$=S/R, with C, S and R defined as above. Upton (1973) gives a likelihood ratio test statistic for $H_0:\mu=\mu_0$ verses $H_0:\mu\ne\mu_0$, provided that α is not large, that rejects H_0 if

$$R^2 > X^2 + \frac{(2n^2 - X^2)}{4n} Z_\alpha, X^2 = [Rcos(\mu_0 - \hat{\mu})]^2$$

for an a significance level α. Upton (1986) extends his hypothesis test to derive two confidence intervals for µ:

$$\hat{\mu} \pm cos^{-1}\left[\left(\frac{4nR^2 - 4nR^2 Z_\alpha}{4nR^2 - R^2 Z_\alpha}\right)^{1/2}\right] \quad \text{for } R \le 0.9n;$$

$$\hat{\mu} \pm cos^{-1}\left[\left(\frac{\mathrm{n}^2 - (\mathrm{n}^2 - \mathrm{R}^2)\exp(\mathrm{Z}_\alpha/\mathrm{n})}{\mathrm{R}^2}\right)^{1/2}\right] \quad \text{for R>0.9n}$$

Under the normal approximation, this reduces to the simpler form

$$\mu \pm (Z_\alpha / (R\hat{\kappa})^{1/2}$$

Finally, note that the R package circular (Agostinelli et al., 2011) provides many of these calculations using the von Mises distribution.

III. Circular orientation model

In order to incorporate the effects of landscape features and environmental covariates in animal movement, we assume that the angular displacement of the animal follows a von Mises distribution with mean m_i and concentration parameter κ given by the equation

$$f(y_i; \mu_i, \kappa) = [2\pi I_O(\kappa)]^{-1} exp\{\kappa cos(y_i - \mu_i)\}, \quad 0 \le y, 2\pi \le 2\pi, \kappa_i > 0 \tag{3}$$

where $I_p(.)$ is the modified Bessel function of the first kind and order p, p=0,1,…. The positive parameter κ measures the concentration of the distribution: as κ=0, the Von Mises distribution converges to uniform distribution around the circumference, whereas for κ>0, the distribution tends to the point distribution concentrated in mean direction.

For our purpose, we assume the response y_i has a von Mises distribution with mean μ_i and the concentration parameter κ. We denote the vector of turn angles $Y_i = (y_1, \ldots, y_n)'$ which corresponds to the mean orientation μ_i. We note that each y_i is a circular random variable while the corresponding vector of covariates X_i are not circular i.e., are continuous or categorical as with standard regression approaches. Hence in order to study the relationship between Y's and X's, we use a so-called circular-linear regression model (Gould, 1969). This model assumes a monotonic link function that maps the explanatory variables to a circle. Though a variety of choices of link function can be used as discussed in Fisher and Lee (1992) or Fisher(1993), a generalized linear model (GLM) for the mean turn angle $\mu_i = E[Y_i X_i]$ may be formulated as follows:

$$\mu_i = \mu_0 + 2arctan(\beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k) \tag{4}$$

where μ_0 is an offset mean parameter representing the origin. If $Y_i^* = Y_i - \mu_0$ is taken as a surrogate response, then the corresponding mean turn angle is $\mu_i = \mu_i - \mu_0 = 2arctan(\eta_i)$ with origin of 0^O. This implies that $tan(\mu_i^* / 2) = \eta_i = X_i\beta$.

We note that equation (4) can incorporate covariates measured at both individual and herd levels. However, we focus solely on the individual level covariates as our model results in a single mean orientation model. Individual level predictors in our model would output a vector of mean turn angle for an individual, which though statistically correct, would not be easily interpretable in our setting. This challenge is analogous to covariates in a cross sectional data. Our goal is to find individual level characteristics that may be associated with animal turning pattern. We note that some of the individual level covariates can be a summary statistics, i.e., average distance to landscape features, environmental drivers throughout the mean orientations when the animal is moving, seasons, rainfall (Duffy et al., 2011) and distance to roads (Tracey et al., 2005).

One of the crucial assumptions of the von Mises distribution is that the turn angle data of the animal movement is unimodal (Goodwin and Fahrig, 2002). We note that no conclusive proof in ecology regarding whether unimodality is a valid statistical assumption, as any evidence of multi-modality based mostly on the evidence from very small samples of the data with no formal attempt to determine analytically whether more than one mode can be detected relative to the amount of variability in the data. SenGupta and Ugwuowo (2006) derived recent advances in the analyses of directional data in ecological and environmental sciences. Otieno and

Cook, (2006) found Measures of preferred direction for environmental and ecological circular data. Tracey et al., (2005) studied the effects of landscape features on rattle snake move angles and found that distance to the roads reduced the move angle of the rattle-snakes. One limitation of there model is that it cannot accommodate more than one covariate. Tracey et al., (2011) used neural network approach and extended their single covariate model to a population model with several covariates in a semi-parametric approach. However, their model is computationally challenging. Fisher and Lee(1992) studied the orientation of birds and discusses the challenges of the model.

Computing the Mean Parameter Estimates

Our goal here is to estimate the regression coefficient $\beta = (\beta_0, \ldots, \beta_k)'$ in equation (4) and derive consistent variance estimates of the parameter estimates. Our data is composed of turn angles of an individual elephants herd. The parameter estimates of μ, κ and β of equation (4) are obtained by maximizing the log-likelihood:

$$logL = -NlogI_O(\kappa) + \kappa \sum_{i=1}^{N} cos(\theta - \mu - 2arctan(x_i\beta)). \quad (5)$$

We have

$$\frac{\partial L}{\partial \mu} = \kappa \sum_{i=1}^{N} \sin\{\theta_i - \mu - g(\beta x_i)\}$$

$$= \kappa \left[\cos\mu \sum_{i=1}^{N} \sin\{\theta_i - g(\beta x_i)\} - \sin\mu \sum_{i=1}^{N} \cos\{\theta_i - g(\beta x_i)\} \right] \quad (6)$$

$$\frac{\partial L}{\partial \beta} = \kappa \sum_{i=1}^{N} \sin\{\theta_i - g(\beta x_i)\} \frac{2x_i}{1 + \beta^2 x_i^2} \quad (7)$$

where $u_i = sin\{\theta_i - \mu - g(\beta x_i)\}$, $H = diag(g'(x_1\beta), \ldots, g'(x_n\beta))$, $g'(\beta x_i) = 2x_i / (1 + \beta x_i^2)$, and

$$\frac{\partial L}{\partial \kappa} = -N\frac{I_0'(\kappa)}{I_0(\kappa)} + \sum_{i=1}^{N} \cos\{\theta_i - \mu - g(\beta x_i)\} \quad (8)$$

$$= -NR + C\cos\mu + S\sin\mu.$$

From which the estimates are the solutions of the following equations:

$$X'Hu = 0 \quad (9)$$

where $S = \sum_{i=1}^{n} sin(\theta_i - g(x_i\beta)) / n$, $C = \sum_{i=1}^{n} cos(\theta_i - g(x_i\beta)) / n$ $R = (S^2 + C^2)^{1/2}$ and $A(\kappa) = \frac{d}{d\kappa} logI_0(\kappa) = \frac{I_1(\kappa)}{I_0(\kappa)}$.

We solve equation (9) using iteratively re-weighted least squares (IRLS) algorithm of Green and Williams (1984) to obtain the parameter estimates μ, κ and β. In order to obtain the starting values of the IRLS algorithm, we assume the data are uncorrelated. We fit the von Mises distribution to the data and obtain the maximum likelihood estimates of μ and κ as described by Fisher and Lee (1992). We choose the MLE of κ as its starting value for our algorithm and the MLE of μ is not used further. From equations (7), the values of β that maximizes the log-likelihood, assuming independent observations, is equal to the value that maximizes the log-likelihood equation. So, we begin with an initial value for $\hat{\beta}$, then calculate S, C and R and hence $\hat{\mu}$ and $\hat{k}$ using (9). These estimates are then used to solve (9) for an updated value $\hat{\beta}$. This procedure is repeated until convergence is achieved. The updating equations for $\hat{\beta}$ in the IRLS algorithm is

$$X'H^2X(\hat{\beta}^+ - \hat{\beta}) = X'H^2r, \quad (10)$$

From the theory of maximum likelihood, the asymptotic variance covariance matrix for the $\hat{\beta}$ is given by

$$var(\hat{\beta}) = \frac{1}{\kappa A(\kappa)} \left((X'H^2X)^{-1} + \frac{(X'H^2X)^{-1}X'hh'X(X'H^2X)^{-1}}{N - h'X(X'H^2X)^{-1}X'h} \right), \quad (11)$$

where h is a vector whose elements are the diagonal elements of H. The asymptotic variance of $\hat{\kappa}$ is equal to $(nA'(\kappa))^{-1}$ and the asymptotic circular variance of μ_0 is $(n(n-p)kA(k))^{-1}$ where $A'(\kappa) = \frac{\partial A(\kappa)}{\partial \kappa} = 1 - \frac{A(\kappa)}{\kappa} - A^2(\kappa)$.

Tests and confidence intervals

Statistical inference on random variables comprises estimation and testing procedures that allow to characterize the underlying distribution regardless of the variables nature and or/dimensions. Tests and confidence intervals can be derived utilizing the standard normal distribution. The confidence interval of β is given by

$$\hat{\beta} \pm z_{N-1,\alpha/2}\sqrt{var\left(\hat{\beta}\right)};$$

$$\hat{\beta} \pm z_{N-1,\alpha/2}\sqrt{\frac{1}{\kappa A(\kappa)}\left((X'H^2X)^{-1} + \frac{(X'H^2X)^{-1}X'hh'X(X'H^2X)^{-1}}{N-h'X(X'H^2X)^{-1}X'h}\right)}$$

Similarly, a test of $H_0 : \beta = 0$ versus $H_1 : \beta > 0$ would be to reject the H_0 if

$$\hat{\beta} / \sqrt{\frac{1}{\kappa A(\kappa)}((X'H^2X)^{-1} + \frac{(X'H^2X)^{-1}X'hh'X(X'H^2X)^{-1}}{N-h'X(X'H^2X)^{-1}X'h})} > t_{N-1,\alpha/2}.$$

Tests and confidence intervals for $\hat{\delta}^2$ can be derived in the same manner. A confidence interval for s^2 can be calculated as

$$\left(\frac{N\hat{\sigma}^2}{\chi^2_{N-2,1-\alpha/2}}, \frac{N\hat{\sigma}^2}{\chi^2_{N-2,\alpha/2}}\right)$$

Model checking and diagnostics testing

Residuals are used to identify discrepancies between models and data. It is useful to establish residuals as contributions made by individual observations on the goodness of fit measures. One of the most useful in GLMs is the deviance $D(y,\hat{\mu}) = \sum_{i=1}^{n} d_i^2$ where

$$d_i = \pm\sqrt{2}(\ell_i(y_i;\tilde{\mu}_i,\kappa) - \ell_i(y_i;\hat{\mu}_i,\kappa))^{1/2}$$

and $\ell_i(y_i;.)$ is the contribution of y_i for the total log-likelihood, $\tilde{m}_i$ is the maximum likelihood estimate of m_i based only on y_i, $\tilde{m}_i$ is the maximum likelihood estimate of m_i based on the whole sample and the sign of d_i is the same as that of $y_i - \hat{\mu}_i$. In the case of the von Mises distribution, $VM(\mu_i, k)$, the deviance residual is defined by $d_i = d(y_i;\hat{\mu}_i,k) = \pm\sqrt{k}(1-cos(y_i-\hat{\mu}_i))^{1/2}$. By using relations between the trigonometric functions we can rewrite d_i in a more convenient form,

$$d(y_i;\hat{\mu}_i,k) = \pm\sqrt{k}cos\frac{1}{2}(y_i - \hat{\mu}_i) \quad (12)$$

that is useful in the development of the approximations.

IV. Applications

An illustration using simulated data

To investigate the behaviour of the parameters of the fitted circular linear model, turn angles of animal movement are simulated from a Von Mises distribution $vm(\mu_i, k)$ with $\mu_i = \mu + arctan(x_i\beta)$ and x values from a uniform distribution on (0,50). Such model checking is required since wrong choice of initial parameter values in model specification can lead to non-identifiability and non-convergence of the model (Gould 1969).

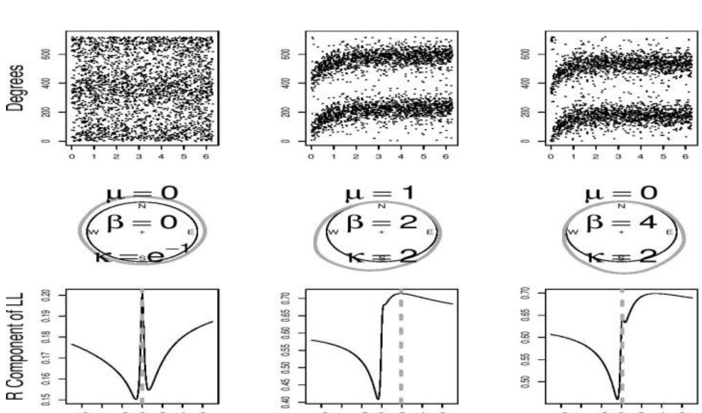

Figure 1: Simulation summaries for the von Mises distribution

Figure (1) shows the scatter plots of the raw data (first row), the densities (second row) and the corresponding likelihood as a function of the estimated β coefficients where $\hat{\beta}$ coefficients are estimated using the IRLS algorithm (third row). The first row visualizes possible patterns in the plot of (θ, x) and the additional points $(\theta + \pi, x)$ in the cartesian coordinates. For the two strongly non-uniform cases, we see the effect of rolling past zero point as the mode determines a well defined point cloud. The second row shows that these are not strongly modal forms despite the patterns in the first row. The third row informative because it demonstrates the difficulty in naively applying a mode finding algorithm, and this is why we recommend always using multiple starting points to find the global maxima (dashed line)

The first column of Figure (1) shows that if the true β is close to or near zero, the log-likelihood function not only has peaks near $\hat{\beta} = 0$, but also asymptotes out to $\pm\infty$ as β gets big in absolute value. In such a case, the only practical estimate of the mode is the peak near zero, which is the solution produced by the probability density function of a von Mises distribution. Despite this seemingly arbitrary choice, the parameter $\hat{\beta}(\text{mod}2\pi)$ is fully identified in the mathematical sense, as opposed to the ecological sense. This difference in identification definitions is a direct result of wrapping around the circle.

If the true value of $|\beta| > 0$, as illustrated in the second and third column of figure (1), the mean resultant length (R) of the log-likelihood function is not globally concave and local maximum exist quite close to the global maximum. This makes the maximization of the likelihood function difficult. The inspection of R as a function of β usually produces good starting values for the maximization of the likelihood. If computationally feasible, a grid of starting values can be used in subsequent runs of the IRLS algorithm. However, due to lack of global log-concavity, a proper exploration of the likelihood plot is always highly advisable when using the maximum likelihood approach in the von Mises regression model (Fisher and Lee, 1992).

Elephant movement data

In this section, we apply the discussed regression model to a real elephant movement data of orientation with covariates. We analyse the data that depicts the reorientation of the Elephants collected using GPS telemetry radio tracking in Kruger National Park (KNP). The data set consists of 63,265 observations recorded at an interval of 30 minutes from April 17, 2006 to May 16, 2009. However, due to computational challenges we only use 4,000 observations in our analysis. The turning angle (θ) is defined as the change in the direction of movement made by an individual from one location to another. The turning angle is a right-hand turn that ranged from -π to π. To illustrate the utility of the reviewed methods, we analysed movement data from foraging African elephant in KNP. Spatial locations of the movement path of an individual foraging herd of female African elephant (Loxadonta) were recorded using GPS radio-telemetry device during 2006-2009. The GPS locations provided data every 30 minutes during an entire day with an accuracy of the locations within 50 meters. This information was sent via cellular phone (GSM) network to a website from where the information was downloaded.

GPS tags provided by African wildlife tracking (http://www.awt.co.za) transmitted location data through the GSM (cellular phone) network to a website, from where these records were received. Collars bearing GPS tags were placed on female elephants representing the movements of the breeding herds with which they were associated. Animal capture was undertaken using chemical immobilisers by South African National parks staff, following standard ethical procedures. An individual elephant representing a heard of eight elephants in the southwestern (Pretorious Kop) region of Kruger National Park received collars in May 2006. If the animal was outside of the cell phone reception, the location data were stored and then downloaded once the animal came into a reception area (Vanak et al., 2010). Movement turn angles were calculated from GPS locations recorded every 30 minutes apart using the methods given by (Tracey et al., 2005).

In the South Africa, elephant habitat is rapidly being destroyed, fragmented and degraded by urban and agricultural development. An objective of the radio-telemetry study was to assess the effects of landscape features on the movement of the elephant. Location data and other associated information were collected by radio-tracking the elephant (Birkett et al., 2012). Elephant herd locations were obtained by tagging a small transmitter on the elephant's body and then locating the transmitter by a receiver attached to a directional antenna. Spatial coordinates for the elephant location were obtained using global positioning system(GPS) receivers and differential correction techniques were applied to improve their accuracy. Data for relocation intervals during which the animal did not move were excluded from this analysis (Vanak et al., 2010). The set of closed line segments, where the from-vertex of the first segment was identical to the to-vertex of the last segment, form a polygon representing the habitat patch. The resulting ArcView shapefile was exported as a tab delimited text file for input to R (*R Development Core Team 2008*) environment in which we perform the remaining analysis. An R function cm developed by extending the source code of the circular R package was used to generate the results (Lund et al., 2007).

V. Results

In Figure (4), we show a wind rose diagram representing the elephant movement turn angles data. Unlike linear data, failure to account for the cyclic nature of the circular data in graphics deceives viewers because it appears there are explicit endpoints. For example, consider the animal movement data. The turn angles derived from the GPS tracking data recorded half hourly certainly affect how animals move. Figure (2) shows the turn angles in a linear histogram. It shows that the turn angles are unimodal and bell-shaped distributed. Conversely, Figure (3) is a linear histogram of the elephant turn angles, it accurately depicts the orientation pattern of elephant movement during foraging or moving to a target.

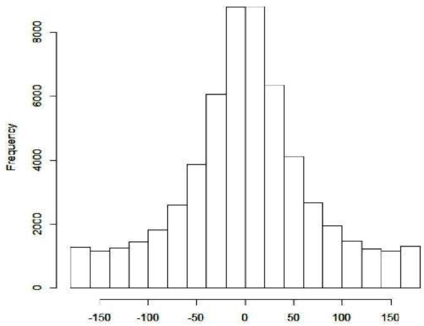

Figure 2: Histogram of elephant turn angle data.

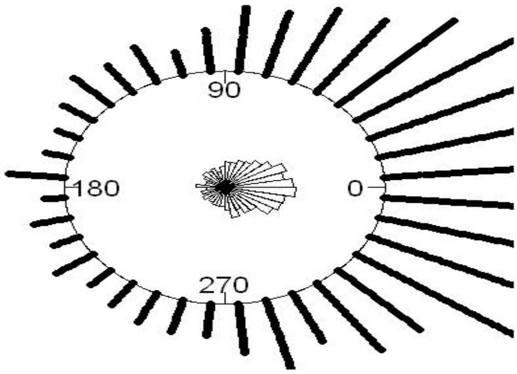

Figure 3: Rose diagram of elephant turn angle data.

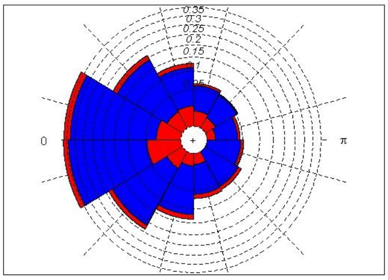

Figure 4: Displays of wind rose diagram of elephant turn angle data.

We model the relationship between the turn angles and the step lengths. To make the IRLS algorithm converge to the global maxima, we carefully select the starting values based on visual inspection of the likelihood plot in Figure (1) of the simulated data. Table (2) gives the output from the circular regression model for these data, which is fit using our version of the Fisher and Lee (1992) algorithm described in section 3. To ensure that the IRLS algorithm converged to the global maxima, the algorithm in section 3 was run from several different starting values since the likelihood function is not guaranteed to be unimodal, making the model somewhat fragile as seen in Figure (2). The fitted model provides a substantial improvement over the null (μ only) model since the difference in deviance is (3669.977) is far into the tail of the chi-squared distribution with 15 degrees of freedom.

Table 1: ML Estimates of directionality of turn angles of elephant movement

Parameters	coef	std.err	|z|-value	p value
Mean direction (□)	2.182	0.3978	5.4839	0.0000
□	0.0101	0.0068	1.4835	0.01379
□	0.1126	0.0448	2.5109	0.012042

Applying the iterative process (10) to solve 9, we found the maximum likelihood estimates $\mu_0 = 2.182(SE = 0.3978)$, The step length parameter $\hat{\beta}$ =0.0101(0.00068) is significantly positive which implies that the length of a step an animal makes increases the magnitude of the turn angle. The concentration parameter of the elephant turn angles is statistically significant $\hat{k}$ =0.1126(0.0448) which implies that the turn angles are not uniformly distributed.

Table (2) presents the parameter estimates of the circular regression of the animal movement and so as expected there is a significant and positive relationship between the turning behaviour of elephants and the landscape features. The coefficient of step length is positively and significantly associated with the turning angle of the animal (β=0.0101, $p<0.05$). This implies that the turning behaviour of elephants is positively influenced by the length of distance it moves within a given interval of time in search of food resources (Duffy et al., 2011). The distance to the river is negatively and significantly associated with the turn angle the animal makes (β=-0.0012, $p<0.05$). This finding supports the earlier argument by resources (Duffy et al., 2011) who noted that elephants forage in habitats closer to the river than habitats far away from the river. The distance to the water points is negatively associated with the turn angle but not significant. The distance to the roads was positively and significantly related with the elephant turning pattern. This implies that an elephant adjust there turning angles to avoid approaching the roads, or when crossing the roads, thereby avoiding risks involved in moving through the human-dominated landscape elements. The distance to the water points and distance to the road were also negatively associated with the turn angles with a coefficient of -0.004 but not significant ($p<0.05$).

Table 2: ML Estimates of landscape features on the mean direction of turn angles on Elephant movement

Parameters	coefficients	Std.err	z-value	p-value
□	-2.6640	0.3010	8.8494	0.0000
Step Length	0.0101	0.0043	2.3456	0.0190
Distance to roads	0.0012	0.0005	2.5869	0.0097
Distance to water point	-0.0004	0.0003	1.4263	0.1538
Distance to river	-0.0004	0.0002	1.9097	0.0562
Temperature	0.0700	0.0827	0.8462	0.3974
Rainfall	-0.0008	0.0046	0.1774	0.8592
□	0.0744	0.3010	0.2470	0.8049

This result indicates that elephants decreases there turn angles as the distance to the water points or river increases. However, the mean annual rainfall had negative but not statistically significant effects on elephant turning (β=-0.0008, $p>0.05$) which indicates that elephants turned less as rainfall increased.

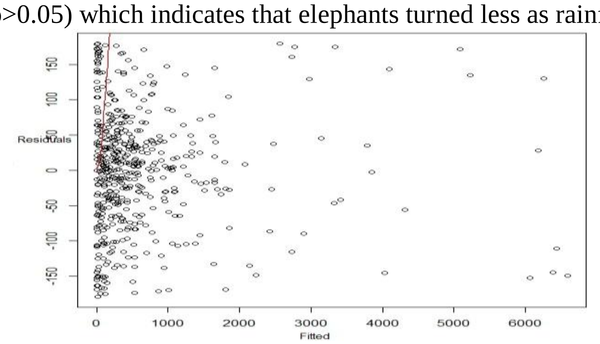

Figure 5: Diagnostic plot of the fitted von Mises regression model of elephant movement.

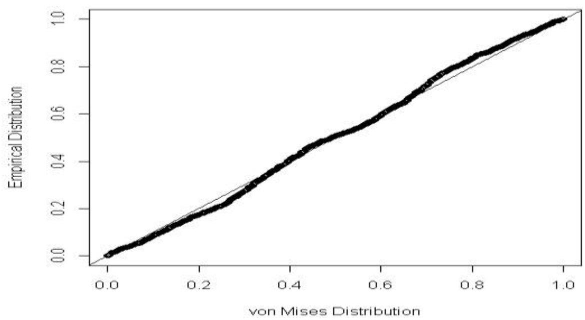

Figure 6: Probability plot of elephant movement turn angle model.

The temperature is positively associated with turn angle but not significant (β=0.070, $p>0.05$). The result indicates that the elephant did not alter the turn angle as the temperature increased. Figure (6) presents the QQ-plot of the turn angles data. The figure indicates that elephant turn angles fits elephant movement well. Figure (5) is a diagnostic plot of the residuals verses the fitted values of the model.

VI. Discussion

Our interest in this chapter was to apply the circular linear regression methodology to animal movement data. In particular, we show that the models presented here can be an alternative to modelling animal movement data with circular responses and linear or categorical covariate. Despite the substantial literature involving information from several species, there is no clear application of circular statistical methods in ecology and especially on animal movement GPS tracking data.

The application of circular statistical methods to optimal foraging theory across in situ systems has been hindered by the inability to account for all critical variables. The regression models provide a link between the circular metrics (turn angles) and the environmental drivers of animal movement. This allows broader inspection of the influence of ecological fluctuations on movement behaviour and the behavioural mechanisms adopted by species to cope with ecological constraints they face. An important characteristic of animal movement pattern is the clustering of the visited locations which can be modelled through the concentration parameter of the von Mises model (Fisher and Lee, 1992). Further, more complex models the model both the effects of landscape features on the concentration parameter and the mean orientation are also possible. However, a major challenges is in implementing such a model occasioned by is the lack statistical softwares and computational capacity. In particular, the turning angle models enabled us to evaluate the effects of internal states of the animal on movement orientation such as moving to water and foraging separately. The findings of this study support the descriptive analysis of resources (Duffy et al., 2011) who argued that animals turned less from large permanent water bodies and more from semi-permanent water bodies like seasonal rivers. Elephants used a direct movement strategy-turning less-when needing to get a destination more quickly (e.g. toward water or mates), rather than significantly increasing their speed. Further, it supports the hypothesis that the elephant optimizes energy efficiency while still varying their foraging approach and search intensity. Movement ecology will advance in parallel to developments in circular statistics, and the development of circular statistic will be promoted by the practical demands made from movement ecology.

Although estimating the parameters of the von Mises distribution was straightforward, in that estimates of μ and κ were always obtained, the angular regression models does not tend to a sensible solution immediately. Thus, careful note of the contour of the likelihood surfaces is required. However, these problems will be in part data-specific particularly (we postulate) if the value of κ is small. More recent methodological work on circular regression which refers to some early animal movement examples concerned with the study of landscape features on animal movement orientation is found in Tracey et al., (2005)

References

[1]. Agostinelli, C. and Lund, U. (2011). R package ’circular’: Circular statistics (version 0.4-3). URL https://rforge.r-project. org/projects/circular.

[2]. Batschelet, E. (1981). Circular Statistics in Biology. Mathematics in Biology. Academic Press.

[3]. Birkett, P. J., Vanak, A. T., Muggeo, V. M. R., Ferreira, S. M., and Slotow, R. (2012). Animal perception of seasonal thresholds: Changes in elephant movement in relation to rainfall patterns. PLoS ONE, 7(6):e38363.

[4]. Cain, M. L. (1985). Random search by herbivorous insects: A simulation model. Ecology, 66(3):pp. 876–888.

[5]. Cain, M. L. (1989). The analysis of angular data in ecological field studies. Ecology, 70(5):pp. 1540–1543.

[6]. De Beer, Y. and Van Aarde, R. (2008). Do landscape heterogeneity and water distribution explain aspects of elephant home range in southern africa’s arid savannas? Journal of Arid Environments, 72(11):2017–2025.

[7]. Downs, T. D. and Mardia, K. V. (2002). Circular regression. Biometrika, 89(3):683–698.

[8]. Duffy, K. J., Dai, X., Shannon, G., Slotow, R., and Page, B. (2011). Movement patterns of african elephants (Loxodonta africana) in different habitat types. South African Journal of Wildlife Research, 41(1):21–28.

[9]. Edwards, A. M., Freeman, M. P., Breed, G. A., and Jonsen, I. D. (2012). Incorrect likelihood methods were used to infer scaling laws of marine predator search behaviour. PloS one, 7(10):e45174.

[10]. Edwards, A. M., Phillips, R. A., Watkins, N. W., Freeman, M. P., Murphy, E. J., Afanasyev, V., Buldyrev, S. V., da Luz, M. G. E., Raposo, E. P., Stanley, H. E., and Viswanathan, G. M. (2007). Revisiting L´evy flight search patterns of wandering albatrosses, bumblebees and deer. Nature, 449(7165):1044–1048.

[11]. Fisher, N. and Lee, A. (1992). Regression models for an angular response. Biometrics, pages 665–677.

[12]. Fisher, N., Lewis, T., and Embleton, B. (1993). Statistical analysis of spherical data. Cambridge University Press.

[13]. Getz, W. M. and Saltz, D. (2008). A framework for generating and analyzing movement paths on ecological landscapes. Proceedings of the National Academy of Sciences, 105(49):19066–19071.

[14]. Gill, J. and Hangartner, D. (2010). Circular data in political science and how to handle it. Political Analysis, 18(3):316–336.

[15]. Goodwin, B. J. and Fahrig, L. (2002). Effect of landscape structure on the movement behaviour of a specialized goldenrod beetle, trirhabda borealis. Canadian Journal of Zoology, 80(1):24–35.

[16]. Gould, A. L. (1969). A regression technique for angular variates. Biometrics, pages 683–700.

[17]. Green, P. J. (1984). Iteratively reweighted least squares for maximum likelihood estimation, and some robust and resistant alternatives. Journal of the Royal Statistical Society. Series B (Methodological), pages 149–192.

[18]. 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.

[19]. Lund, U. and Agostinelli, C. (2009). The circstats package. version 0.2-4.

[20]. Lund, U., Agostinelli, C., Agostinelli, M. C., and SenGupta, A. (2007). The circular package.

[21]. Mardia, K. and Jupp, P. (2009). Directional statistics, volume 494. Wiley.

[22]. Milne-Thomson, L., Abramowitz, M., and Stegun, I. (1972). Handbook of mathematical functions. Handbook of Mathematical Functions.

[23]. Otieno, B. S. and Anderson-Cook, C. M. (2006). Measures of preferred direction for environmental and ecological circular data. Environmental and Ecological Statistics, 13(3):311–324.

[24]. Owen-Smith, N., Fryxell, J., and Merrill, E. (2010). Foraging theory upscaled: the behavioural ecology of herbivore movement. Philosophical Transactions of the Royal Society B: Biological Sciences, 365(1550):2267–2278.

[25]. Owen-Smith, N., Goodall, V., and Fatti, P. (2012). Applying mixture models to derive activity states of large herbivores from movement rates obtained using gps telemetry. Wildlife Research, 39(5):452–462.
[26]. Redfern, J. V., Grant, R., Biggs, H., and Getz,W. M. (2003). Surface-water constraints on herbivore foraging in the kruger national park, south africa. Ecology, 84(8):pp. 2092–2107.
[27]. Ropert-Coudert, Y. and Wilson, R. P. (2005). Trends and perspectives in animal-attached remote sensing. Frontiers in Ecology and the Environment, 3(8):437–444.
[28]. Schou, G. (1978). Estimation of the concentration parameter in von mises–fisher distributions. Biometrika, 65(2):369–377.
[29]. SenGupta, A. and Ugwuowo, F. I. (2006). Asymmetric circular-linear multivariate regression models with applications to environmental data. Environmental and Ecological Statistics, 13(3):299–309.
[30]. Shannon, G., Matthews, W. S., Page, B. R., Parker, G. E., and Smith, R. J. (2009). The affects of artificial water availability on large herbivore ranging patterns in savanna habitats: a new approach based on modelling elephant path distributions. Diversity and Distributions, 15(5):776–783.
[31]. Siniff, D. B. and Jessen, C. (1969). A simulation model of animal movement patterns. volume 6 of Advances in Ecological Research, pages 185–219. Academic Press.
[32]. Tracey, J. A., Zhu, J., and Crooks, K. (2005). A set of nonlinear regression models for animal movement in response to a single landscape feature. Journal of Agricultural, Biological, and Environmental Statistics, 10(1):1–18.
[33]. Tracey, J. A., Zhu, J., and Crooks, K. R. (2011). Modeling and inference of animal movement using artificial neural networks. Environmental and Ecological Statistics, 18(3):393–410.
[34]. Upton, G. J. (1973). Single-sample tests for the von mises distribution. Biometrika, 60(1):87–99.
[35]. Upton, G. J. (1986). Distance and directional analyses of settlement patterns. Economic Geography, pages 167–179.
[36]. Urbano, F., Cagnacci, F., Calenge, C., Dettki, H., Cameron, A., and Neteler, M. (2010). Wildlife tracking data management: a new vision. Philosophical Transactions of the Royal Society B: Biological Sciences, 365(1550):2177–2185.
[37]. Vanak, A., Thaker, M., and Slotow, R. (2010). Do fences create an edge-effect on the movement patterns of a highly mobile mega-herbivore? Biological Conservation, pages 2631–2637.
[38]. Viswanathan, G. M., Raposo, E. P., Bartumeus, F., Catalan, J., and da Luz, M. G. E. (2005). Necessary criterion for distinguishing true superdiffusion from correlated random walk processes. Physical Review. E,72:011111.