Journal of Modern Mathematics and Statistics
MEDWELL PUBLICATIONS SINCE 2000

Analysing the Effects of Landscape Factors on Animal Movement Step Lengths with Stable Law Regression Models

[1]R.M. Mutwiri, [2]H. Mwambi and [3]R. Slotow
[1]*Department of Mathematics Computing and Information Technology, University of Embu, Embu, Kenya*
[2]*School of Mathematics, Statistics and Computer Sciences, University of KwaZulu Natal, Durbar, South Africa*
[3]*School of Life Sciences Sciences, University of KwaZulu Natal, Durbar, South Africa*

Key words: Stable law regression model, elephant movement, habitat types, GPS data, student regression model

Corresponding Author:
R.M. Mutwiri
Department of Mathematics Computing and Information Technology, University of Embu, Embu, Kenya

Page No.: 1-9
Volume: 14, Issue 1, 2020
ISSN: 1994-5388
Journal of Modern Mathematics and Statistics

Abstract: The potential advantage of stable distribution assumption in modelling ecological disturbances is of animal movement is the central theme of this study. Studies relating animal movement paths to structured landscape data are particularly lacking despite the obvious importance of such information to understanding animal movement. Previous studies of elephant movement have shown that speed is heavy tailed and skewed. In this study, we model the heavy tails using the student t regression model, the skewness and the heavy tails with the stable law regression. The new models add substantial flexibility and capabilities including the ability to incorporate multiple variables. We use a likelihood based approach that utilizes the fourier transform technique to evaluate the densities and demonstrate the approach with movement data from five elephant herds (Africana Loxadonta). The proposed methodology can be useful for GPS tracking data that is becoming more common in monitoring of animal movement behaviour. We discuss our results in the context of the current knowledge of animal movement and in particular elephant ecology highlighting potential applications of our approach to the study of wide ranging animals.

INTRODUCTION

Regression analysis is one of the most popular methods in ecology and statistics where most variables of interest such animal movement step lengths are assumed to be normally distributed. However, the normality assumption is not appropriate for many ecological variables, especially, animal movement metrics (speed, step lengths) variables and also in some cases circular metrics (turn angles) Bartumeus[1]. Animal movement linear metric are typically heavy tailed and excessively highly peaked around zero. A stable distribution whose shape is governed by the stability index parameter α represent one such alternative. Thus, such a distribution is better suited to describing such variables; the normal distribution is a special case of the stable distribution. To this end, the four parameter family of stable distribution is more of a generalization of the central limit theorem than an alternative.

The flexibility of stable distribution can be explored in a regression modelling framework to overcome some of the deficiencies of linear regression models when analysing heavy tailed and skewed data. The non-Gaussian stable distributions have heavier tails than the Normal distribution and allow skewness?. Heavy tails and skewness implies that extreme observations are given a greater probability of occurring and are thus given less weighting in maximum likelihood estimation, so that, fitted lines are not biased towards these extreme observations[2]? Therefore, it is a reasonable extension to the regression models to assume a stable distribution as the distribution of the error terms. Alternative models to the stable law regression models are the student's t, skewed student's t and skewed normal regression models? demonstrates that the regression model with student's t errors also suffers from monotone likelihood. The use of the normal distribution to model errors of linear model is under increasing criticism for its inability to model fat or heavier tailed distributions as well as being non-robust. Lange *et al*.[3] generalized the traditional regression model with normal distributed errors to more robust regression models with t distributed errors. It is well known that the t distribution provides a convenient description for regression analysis when the residual term has a density with heavy tails. From, the classical linear model can be modelled as follows.

The stable distribution has found wide applications in financial problems, Biology, genetics, ecology and geology[4] with a few applications in movement ecology[5]. The assumption that animal step lengths or speed follow a stable distribution has far reaching consequences for both foraging ecology and statistical theory[5]. For example, the problem of L'evy flight search patterns is well studied[6-9], robustness to the sampling frequency is studied by Kawai and Petrovskii[10] and for a specific discussion of movement ecology and statistical issues[5]. However, in all these studies, the link between linear metrics and the environmental heterogeneity variables remains unexplored. According to Duffy *et al*.[11], different vegetation cover types have varying impacts on elephant movement. Surface water availability, patch quality, rainfall and distance to the water bodies is known to affect elephant movement. The effects of artificial water points and fences has been investigated Loarie *et al*.[12]. In this study, we examine the effects of vegetation cover type in a stable regression model setup in order to understand elephant movement.

Advances in statistical computation have made it possible to estimate the unconditional stable density as well as incorporate covariates[2]. However, estimates of the stable distribution conditional on a set of explanatory variables in the context of regression framework used by applied researchers poses an overwhelming computational problem?. One of the methods used for evaluating the stable density (the direct numerical integration techniques) is non-trivial and burden-some from a computational perspective[13]. As a consequence, maximum likelihood estimation algorithms based on such approximations are difficult to implement, especially for huge data sets encountered in movement ecology[5]. However, with increasing computational power and efficient algorithms, maximum likelihood estimation and other comparative techniques have been implemented by Nolan and Ojeda-Revah[13]. Due to the above mentioned drawbacks, stable distributions are not well explored in movement ecology.

MATERIALS AND METHODS

Stable distribution: Stable distributions are a four parameter family of probability models which was first introduced by L'evy in a study of normalized sums of independent and identically distributed (i.i.d) terms. A random variable X is said to be stable distributed if for any positive integer n>2, there exist constant $a_n>0$ and $b_n \in \mathbb{R}$ such that $X_1 +, ..., + X_n \overset{d}{=} a_n X + b_n$ where $X_1, ..., X_n$ are independent identically distributed copies of X and "$\overset{d}{=}$" -signifies equality in distribution. The coefficients a_n is necessarily of the form $a_n = n^{1/\alpha}$ for some $\alpha \in (0, 2]$ Feller[14]. The parameter α is called the index of stability (tail index) of the distribution and a random variable X with index α is called α-stable. An α-stable distribution is described by four parameters and will be denoted by s (α, β, γ, μ). Closed form expressions for the probability density function of the α-stable distribution is known to exist only for three special cases (cauchy, normal and L´evy distribution).

The research by Mandelbrot[15] and Fama Fama[16] elicited a lot of interest in using stable distributions to model heavy tailed and skewed phenomena but research has been restricted to theoretical context due to computational complexities involved in calculating the probability densities and the consequently what this has for the maximum likelihood procedures. Notable contributions in this field of study are found in DuMouchel[17, 18], DuMouchel[19], Zolotarev[20], Samoradnitsky and Taqqu[21], Janicki and Weron[22] and more recently Nolan[22, 23], Nolan and Ojeda-Revah[13].

Although, the probability density function of the stable distribution cannot be written in closed form, the characteristic function which can be specified in a closed form for all stable distributions, allows the only opening for practical use of the distributions in real life problems McHale and Laycock[24]. The characteristic function can be expressed in several different forms, each of which has advantages over others, for example, formula simplicity over computational consistency. However, the Zolotarev's form has the advantage of being continuous in all the four parameters and behaves more intuitively than in other forms Nolan and Ojeda-Revah[13]. Lambert and Lindsey[2] discuss complexities in fitting their regression model caused by the sensitivity of the location parameter to the skewness parameter. For numerical purposes, several researchers have recommended the use of Zolotarev's parameterization as the most practical in application to real life data sets Nolan and Ojeda-Revah[13]. The characteristic function of a stable random variable X is given by:

$$E\left[\exp(iuZ)\right] = \begin{cases} \exp(-\gamma)|u|^{\alpha}\left[\begin{array}{l} 1+i\beta(\text{sign } u)\tan\frac{\pi\alpha}{2} \\ \left(|u|^{-u-1}-1\right)+ \\ i\mu u \end{array}\right] \alpha \neq 1 \\ \exp\left(-\gamma|u|\left[1+i\beta(\text{sign } u)\frac{2}{\pi}\log|u|\right]+i\mu u\right) \alpha = 1 \end{cases} \quad (1)$$

The family of stable probability density can be calculated using the fourier transform of the characteristic function given by:

$$S(x,\ \alpha,\ \beta,\ \gamma,\ \mu) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{itx}\phi(x;\ \alpha,\ \beta,\ \gamma,\ \mu)\,dt \quad (2)$$

Statistical software to fit stable distribution and density functions are available in Rmetrics for R Wuertz, stable Nolan and Ojeda-Revah[13] or as standalone program stable Nolan[23]. These resources allow one to evaluate the consequences of replacing the normal assumption with the more general stable distribution. Further, advances in theory and computation will aid the development of new models in the coming years and the use of the stable distribution will become more common.

Stable paretian regression model: In may practical applications in animal ecology, it is known that animal movement rate can be affected by a number of covariates (explanatory variables) such as the nearest distance to the water point, vegetation cover type, distance to tourist roads, soil topology, seasons, amount of rainfall, temperature and many others Duffy *et al*.[11]. However, animal movement data is characterized by skewed and heavy tailed distributions. Thus, a model that provides a good fit to movement data will definitely yield more precise estimates of the quantities of interest. Based on the stable distribution assumption, we propose a linear regression type model linking the response and the explanatory variables $X = (x_1, x_2, ..., x_n)$ as:

$$y_i = \beta_0 + \sum_{i=1}^{n} \beta_{ij} x_i + \epsilon_i, i=1,...,n \quad (3)$$

Where:

β = A vector of the unknown parameters to be estimated

ϵ_i = The random error term

The notion of Stable Regression Models (SRMs) was developed by McCulloch[25] for symmetric stable distribution and discussed in detail by McHale and Laycock[24]. In SRMs, the error terms $\epsilon_1, \epsilon_2, ..., \epsilon_n$ are assumed to be independent identically distributed stable random variables denoted by $\epsilon_i \sim S(\alpha, \beta, \gamma, \mu)$ Standard methods of approximating such integrals are of unknown accuracy in real settings. Instead, DuMouchel[17] suggested the use of numerical inversion of the First Fourier Transform (FFT) to obtain a closed density and hence the likelihood for stable distributions. In a similar manner, numerical inversion of the first fourier transform can be used to obtain the parameters of the stable Paretian regression model. DuMouchel[17-19], showed that subject to certain conditions, the maximum likelihood estimates of the parameters of an α-stable distribution have the usual asymptotic properties of a maximum likelihood estimator. They are asymptotically normal, unbiased and have an asymptotic covariance matrix $n^{-1} I(\alpha, \beta, \gamma, \mu)^{-1}$ where $I(\alpha, \beta, \gamma, \mu)$ is the fisher information matrix. McCulloch[25] examines the linear regression model in the context of α-stable distribution paying particular attention to the symmetric case. Here, the symmetry constraint is not imposed. If we denote the stable density function by $S(\epsilon_i;\ \alpha, \beta, \gamma, \mu)$ then we may rewrite the density of ϵ_i as:

$$S(x;\ \alpha,\ \beta,\ \gamma,\ \mu) = \frac{1}{\gamma} S\left(\frac{y_i - \sum_{j=1}^{k} x_{ij}\beta_j}{\gamma},\ \beta, 1, 0\right)$$

The likelihood as:

$$L(\epsilon_i;\ \alpha,\ \beta,\ \gamma,\ \mu) = \frac{1}{\gamma}\prod_{i=1}^{n} S\left(\frac{y_i - \sum_{j=1}^{k} x_{ij}\beta_j}{\gamma},\ \beta,\ 1,\ 0\right)$$

Hence, the log-likelihood function for the vector of parameters $\theta = (\alpha, \beta, \gamma, \mu, \beta_0, \beta_1, ..., \beta_p)$ from model (3) has the form:

$$l\left(\epsilon_i;\ \alpha,\beta,\gamma,\mu,\beta_0,\beta_1,...,\beta_p\right) = -n\sum_{i=1}^{n}\log(\gamma) + \sum_{i=1}^{n}\log\left(s\left(\frac{y_i - \sum_{j=1}^{k} x_{ij}\beta_j}{\gamma},\ \beta,\ 1,\ 0\right)\right) \quad (4)$$

The ML estimator $\hat{\theta}$ of the vector of unknown parameters can be calculated by maximizing the log-likelihood (4) to obtain the solution to the equations:

$$\frac{\partial l}{\partial \beta_m} = \sum_{i=1}^{n} -\phi(\epsilon_i) x_{im} = 0, m = 1, ..., k$$

$$\sum_{i=1}^{n} -\frac{\phi(\hat{\epsilon}_i)}{\hat{\epsilon}_i} \hat{\epsilon}_i \ x_{im} = 0, m = 1, ..., k$$

$$\begin{aligned} &\sum_{i=1}^{n} -\frac{\phi(\hat{\epsilon}_i)}{\hat{\epsilon}_i}\left(y_i - \sum_{i=1}^{n} x_{ij}\beta_j\right) x_{im} = 0, m = 1, ..., k \\ &\sum_{i=1}^{n} -\frac{\phi(\hat{\epsilon}_i)}{\hat{\epsilon}_i} y_i x_{im} = -\sum_{i=1}^{n} -\frac{\phi(\hat{\epsilon}_i)}{\hat{\epsilon}_i} \sum_{i=1}^{n} x_{ij}\beta_j, m = 1, ..., k \end{aligned} \quad (5)$$

where, $\epsilon_i = y_i \, x_{im} - \sum_{i=1}^{n} x_i{}'\hat{\beta}$ if we let W be the diagonal matrix given by:

$$w = \begin{bmatrix} \frac{\phi(\hat{\epsilon}_i)}{\hat{\epsilon}_i} & 0... & 0 & \\ 0 & \frac{\phi(\hat{\epsilon}_i)}{\hat{\epsilon}_i} & ... & 0 \\ \vdots & \vdots & ... & \vdots \\ 0 & 0 & ... & \frac{\phi(\hat{\epsilon}_i)}{\hat{\epsilon}_i} \end{bmatrix}$$

Then using the least squares notation, we may write the normal (Eq. 5) as:

$$X'Wy = (X'WX)\hat{\beta} \quad (6)$$

And if X' W X is not singular, the parameter estimates of β are given by:

$$\beta = (XWX)^{-1} X'Wy \quad (7)$$

Nolan[26] showed that the evaluation of the likelihood function is made possible by using efficient non-linear optimizers. Maximum likelihood algorithm used in this research are provided by Nolan and Ojeda-Revah[13] within the R package stable 5.1 which can be obtained commercially from www. Robust Analysis.com. Initial values for α, β, γ, μ, β_0, β_1, ..., β_p can be taken from the fit of the stable distribution model.

Regression model with t errors: We consider the univariate nonlinear regression model where the observations $y = (y_1, ..., y_n)$ are independent, y_i having a student t distribution with location parameter μ_i, scale parameter σ and v degrees of freedom. The density of y_i, for each i = 1, ..., n is therefore, given by:

$$f(\beta, \delta, \nu; y, X) = \frac{\Gamma\frac{(\nu+1)^n}{2}}{\Gamma\left(\frac{\nu}{2}\right)^n (\pi\nu)^{n/2} \delta n} \begin{bmatrix} 1+ \\ \frac{1}{\nu}\left(\frac{y_i - x_i{}'\beta}{\delta}\right)^2 \end{bmatrix}^{-(\nu+1)/2} \quad (8)$$

where $\sigma > 0$ and $\nu > 1$ are both unknown. We define a linear regression for y by:

$$y_i = \beta_0 + \sum_{i=1}^{n} \beta_{ij} x_i + \epsilon_i, i = 1, ..., n \quad (9)$$

where, $\epsilon = (\epsilon_1, ..., \epsilon_n)$ is the error vector where the components are independent and identically distributed according to the student t distribution with location zero and scale δ and degrees of freedom v Lange *et al.*[3]. $X = (x_1, ..., x_n)$ is the n×k matrix of explanatory variables. The parameter space is given by $\theta = (\beta, \delta, \nu)$. The likelihood is given by:

$$L(\beta, \delta, \nu; y, X) = \frac{\Gamma\frac{(\nu+1)^n}{2}}{\Gamma\left(\frac{\nu}{2}\right)^n (\pi\nu)^{n/2} \delta n} \prod_{i=1}^{n} \begin{bmatrix} 1+ \\ \frac{1}{\nu}\left(\frac{y_i - x_i{}'\beta}{\delta}\right)^2 \end{bmatrix}^{-(\nu+1)/2} \quad (10)$$

The parameter estimates θ are obtained by maximizing the log-likelihood equation:

$$\begin{aligned} &\text{Log}\, L = \text{In}\,\Gamma\left(\frac{\nu+1}{2}\right)^n - \text{In}\,\Gamma\left(\frac{\nu}{2}\right)^n (\pi\nu)^{n/2\delta^n} - \\ &\frac{(\nu+1)}{2} \sum_{i=1}^{n} \log\left[1 + \frac{1}{\nu}\left(\frac{y_i - x_i{}'\beta}{\delta}\right)^2\right] \end{aligned} \quad (11)$$

The least squares estimator of β is:

$$\beta = (X'X)^{-1} X'Y$$

The variance-covariance matrix for $\hat{\beta}$ is:

$$\text{var}(\hat{\beta}) = E\left[(\hat{\beta}-\beta)'(\hat{\beta}-\beta)'\right] = \frac{\nu\sigma^2}{\nu-2}(X'X)^{-1}$$

Lange *et al*. (1989) noted that this is also the maximum likelihood estimate of β. provided the following stimate of the degrees of freedom parameter:

$$\hat{\nu} = \frac{2(\hat{\alpha}-3)}{\hat{\alpha}-3}$$

Where:

$$\hat{\alpha} = \frac{\frac{1}{n}\sum_{i=1}^{n}(y_i - x'_i\hat{\beta})4^2}{\frac{1}{n}\sum_{i=1}^{n}(y_i - x'_i\hat{\beta})}$$

The maximum likelihood estimator of σ^2 is:

$$\sigma^2 = \frac{1}{n}(y_i - x\hat{\beta})'(y_i - x\hat{\beta})$$

as in the normal case. For $\nu>2$ $E\left(\hat{\sigma}^2\right)=\frac{(n\text{-}p)}{n}\sigma_u^2$ where $\sigma_u^2=\nu\sigma^2/\nu\text{-}2$ is the common variance of the elements of $\in$. Thus, $\hat{\in}/\hat{\in}/n\text{-}p$ is an unbiased estimator for σ_u^2 while:

$$\sigma^2=\frac{\upsilon\text{-}2}{\upsilon(n\text{-}p)}\hat{\in}'\hat{\in} \tag{12}$$

Is unbiased estimator for σ^2. In the class of estimator $q\hat{\in}'\hat{\in}$ with q being a positive scalar, the minimal mean squared error estimator for σ^2 is:

$$\sigma^2=\frac{\upsilon\text{-}4}{\upsilon(n\text{-}p+2)}\hat{\in}'\hat{\in} \tag{13}$$

While the minimal mean squared error estimator for σ^2 in this class is $(\nu\text{-}4)\hat{\in}'\hat{\in}(\nu\text{-}2)(n\text{-}p+2)$. The variances of the unbiased and the minimal mean squared error estimators of σ^2 are:

$$\text{var}\left(\hat{\sigma}^2\right)=\frac{2\sigma^4}{(n\text{-}p)}\frac{n\text{-}p+\upsilon\text{-}2}{\upsilon\text{-}p} \tag{14}$$

Maximum likelihood algorithm used in this research are provided by Osorio and Galea[27] in R statistical package 'heavy'.

Application: elephant movement data

Data description: The telemetry data employed in this study was collected by the South African National parks (SAN- PARKS). In May 2006, 18 African elephants were fitted with GPS-argos telemetry collars (Telenics). Capturing and handling was done according to the University of Kwa Zulu Natal animal care regulations. GPS locations were recorded every 30 min during the first 3 years after collaring and transmitted to SANPARKS via. an Argos satellite uplink every day when the elephant was within network range[28]. Telemetry points collected within the first 24 h after capturing and those with obvious errors were excluded from the analysis. Overall, the telemetry data set was composed of >50,000 GPS points, taken over a period of three years, across a 19, 485 km^2 area Birkett *et al*.[28], Vanak *et al*.[29].

Vegetation cover types: To determine the effects of various habitat types in the pattern of elephant movement, we extracted the vegetation cover types data of before and after the breakpoints. Land cover types and distances to different landscape features within a spatial resolution of 25 m pixels were obtained from the Kruger national park Land cover database. This database is based on the Thematic mapper sensor on landsat Earth-resource satellites using data frames recorded between 2006 and 2009 (spectral analysis Inc. 2009). Dummy variables of

Table 1: Dummy variables of vegetation cover type under investigation

Types	Variables
y_i	Speed of the animal
x_{11}	Nearest distance to the river
x_{12}	Comb
x_{13}	Thicket
x_{14}	Mixed combretum/terminalia sericea woodland
x_{15}	Combretum/mopane woodland of Timbavati
x_{16}	Acacia welwitschii thickets on Karoo sediments
x_{17}	Kumana sandveld
x_{18}	Punda maria sandveld on cave sandstone
x_{19}	Sclerocarya birrea subspecies caffra/Acacia nigrescens savanna
x_{20}	Dwarf acacia nigrescens savanna
x_{21}	Bangu rugged veld
x_{22}	Combretum/acacia nigrescens rugged veld
x_{23}	Lebombo South

vegetation cover types were created and fitted to a regression model assuming stable distributed error terms. The land cover of Kruger National Park (KNP) consist of fourteen vegetation cover types.

Model formulation: The observations of the response y_1, y_2,..., y_n variable represent the movement rate of five elephant herds derived before and after breakpoint home ranges[28]. The covariate vector x_i is the dummy variables representing the vegetation cover type created from the habitat variable. Due to computational complexity of the stable regression model and lack of rich data set with covariates of elephant herds, we shall demonstrate the results of habitat cover types only in this study. The dummy variables created from vegetation cover type are presented in Table 1. Now, we present the results by fitting the model:

$$y_i=\beta_0+\beta_1 x11+\beta_2 x12+,...,+\beta_p\, x_{23} \tag{15}$$

Where the dependent variable y_i speed of elephants follows the stable law distribution or the student's t distribution for i = 1, ..., 200. The dependent variable y_i is the speed of elephant before and after breakpoint home ranges obtained as described by Birkett *et al*.[28] and Vanak *et al*.[29]. The MLEs of the model parameters are calculated using the procedure nlm in R statistical software. Iterative maximization of the logarithm of the likelihood function of the stable law regression starts with some initial values for the θ taken from the linear regression model.

RESULTS AND DISCUSSION

Table 2 lists the MLEs of the parameters for the SRMs and HTRMs Models fitted to the current data. The SRMs Model involves four extra parameters which gives it more flexibility to fit the elephant movement data. Due to lack of rich data set of animal movement with covariates we investigate only the effects of habitat types

Table 2: Summary of heavy tailed -t distribution and stable law regression model

	Heavy tailed model			Stable law regression model		
Parameters	Estimate	SE	p-values	Estimate	SE	p-values
α				1.309	0.051	1.000
β				0.857	0.081	1.000
γ				0.091	0.003	1.000
intercept	0.376	0.025	1.0000	0.340	0.011	0.000
X_{12}	-0.089	0.054	0.0000	-0.086	0.024	0.000
X_{13}	-0.114	0.047	0.0000	-0.144	0.020	0.000
X_{14}	-0.068	0.034	0.0000	-0.053	0.015	0.000
X_{15}	-0.001	0.044	0.0000	0.008	0.019	0.663
X_{16}	0.067	0.037	0.9999	0.033	0.016	0.978
X_{17}	0.228	0.058	1.0000	0.165	0.026	1.000
X_{18}	-0.047	0.045	0.0202	-0.080	0.020	0.000
X_{19}	0.008	0.034	0.6772	0.008	0.015	0.500
X_{20}	0.064	0.039	0.9999	0.043	0.017	0.994
X_{21}	0.006	0.042	0.6103	-0.006	0.019	0.382
X_{22}	0.021	0.041	0.8508	0.014	0.018	0.773
X_{23}	-0.038	0.047	0.0559	-0.048	0.020	0.009
Log like	88.00481			84.16449		
AIC	-170.4644			-162.7838		

as dummy variables. Most of the environmental variables considered, here, were selected as drivers of movement rates before and after break point analysis of home ranges[29]. The fitted SRM indicates that the dummy variables X_{12}, X_{13}, X_{14} X_{18} and X_{23} are significant at 5% level of significance. The linear regression intercept was however, significantly <1 indicating the ability of our models to predict the movement of the elephant at moderate speed.

Since, we have demonstrated that the residuals are non-Gaussian, we will now compare the stable estimates with those obtained from the heavy tailed regression model with student's t distributed disturbances. The results of student's t regression model in Table 2 indicates that the vegetation covers combo, thicket, mixed combretum, punda maria sandveld and lebombo South significantly reduced the movement rates of elephants while mopane woodland, acacia welwisitchii, Kumana sandveld, Sclerocarya birrea subspecies, dwarf acacia savanna, Bangu rugged and Combretum acacia rugged increased the movement rates though not significant. The the results of stable law regression model indicates that the vegetation covers combo, thicket, mixed combretum, punda maria sandveld and lebombo South significantly reduced the movement rates of elephants while mopane woodland, acacia welwisitchii, Kumana sandveld, *Sclerocarya birrea subspecies*, dwarf acacia savanna, Bangu rugged and combretum acacia rugged increased the movement rates though not significant. We note that the movement of elephants in the resource poor patches are positive and significant indicating that elephants increased there movement speed when moving from search of food and water while in resource rich patches the move at a slower speed as they forage. The stability index parameter estimated is 1.31 which is <2 with a standard error of 0.0511 indicating that the data is heavy tailed. Clearly we can reject the null hypothesis that the random disturbance follows a Gaussian distribution (the hypothesis $\alpha = 2$) in favour of the alternative that the disturbance follows a non-Gaussian stable distribution with infinite variance. Figure 1, further supports the findings of the fitted model with residuals of the stable distribution plotted along the empirical density of the data. The density plot shows that the empirical distribution has heavier tails and a higher more concentrated peak compared to the Gaussian distribution. These attributes convey the ecological importance of the tails with appropriate statistical assumption. We used the Akaike information criterion to compare the student's t and the stable law regression models. The results of Table 2 indicates that the student t regression model better fits the data than the stable law regression model with an AIC of -170.46 and -162.78, respectively.

Biological implications and applications: The empirical analysis shows that the effect of vegetation cover types is to reduce the movement rates of elephant in abundant food patches and increased the movement rate in poor resource areas. This finding is consistent with the descriptive analysis by Duffy *et al*.[11] who hypothesized that the quality and availability of forage suppresses movement rates of elephants. The findings also support the argument by Hopcraft etc who found that resource rich vegetation cover reduced the movement rates of animals in Serengeti game reserve, Kenya.

The stable regression model estimated in this study sheds more light on the earlier results by identifying how the underlying ecological processes result to differential habitat use. African elephants have large effects on vegetation and high numbers can lead to extensive habitat

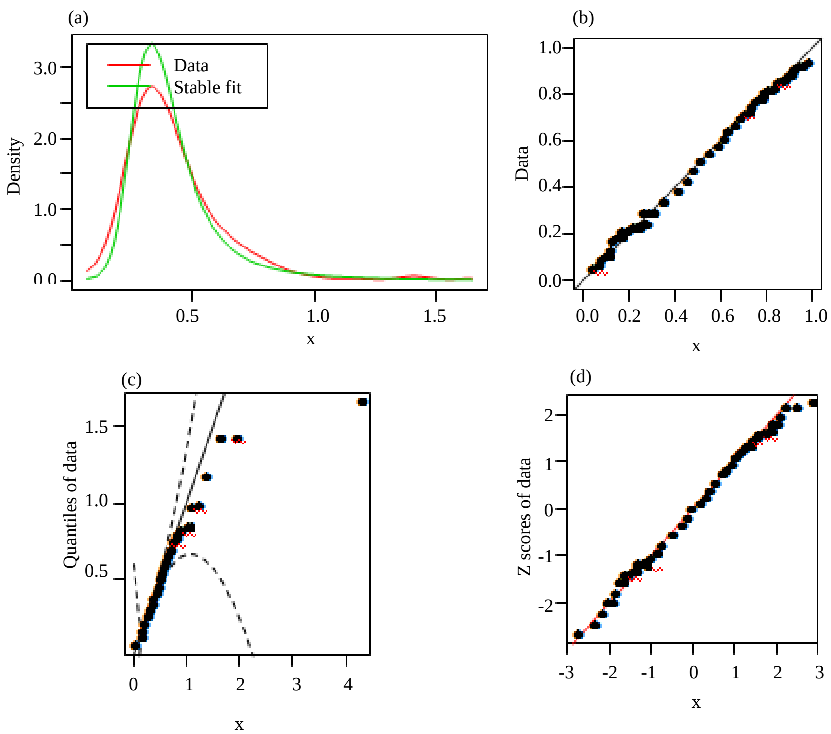

Fig. 1(a-d): Diagnostic analysis of elephant movement data, (b) Stable distribution with theta = (1.346, 0.927, 0.084, 0.351), (c) Stable distribution with theta = (1.346, 0.927, 0.084, 0.351) and (d) z scores of stable distribution with θ = (1.346, 0.927, 0.084, 0.35)

modifications. Driven by the need to manage these impacts several models have been developed to better understand the interaction between elephants and trees. Therefore this understanding can be used for both management and habitat conservation Dai *et al*.[30]. Another implication of the stable regression analysis is that the distribution of movement rates-even when conditioning on the vegetation attributes-has infinite variance. This means that the point predictions are useless because they lack precision especial when the stable parameters α and β are at the boundaries. Confirming the foraging success and measuring the impact of environmental drivers is one of the challenges facing ecologists today. Thus, the finding of this study provides a direct link of inferring the effects of vegetation cover types on elephant movement speed. However, the stable Paretian model does not permit the conditional distribution of movement rates to be quantified and it can be used to make probability statements that may be useful in practice, for example, optimal foraging theory. A potentially important practical application of the stable analysis is movement strategies analysis. A strategy that includes both the L'evy stable walks and the L'evy flights are thought to optimize foraging. Kawai and Petrovskii[10] show in movement ecology applications that stable models-because they capture both skewness and heavy tails in movement rates-perform considerably well than models based on power law distribution or the empirical distribution. Further, due to the analytical tractability of the stable distributions, it is possible to use the stable models to construct optimal search strategies for animals within the framework of movement ecology. In animal movement studies where rare steps in the upper tail of the distribution drive search optimality, it appears promising to use the stable regression models developed above as an input into constructing an optimal search strategies for animals that help understand the relationship between elephant herds and their habitats.

CONCLUSION

We have described the theoretical justification for the use of stable law regression models and t regression models in analysing animal movement data. To be useful in practice, a statistical model of the speed of animal movement should capture asymmetry, the heavy tails implied by the importance of extreme events and allow the speed to be conditioned on a vector of explanatory variables. Recent advances in the statistical theory of non-symmetric density functions and their estimation make it feasible to estimate statistical models based on the stable law and the student's t distribution. It is also possible to estimate t regression models using standard maximum likelihood techniques.

Despite several studies detailing analogous statistical approaches, application of such models to GPS tracking is limited due to computational difficulties Kawai[5] and lack of adequate data rich in covariates in ecology. The t regression model is particularly appealing in ecology where the data are characterized by heavy tails and where we are interested in conditional distributions. Unlike some other distributions in the L´evy stable family- t models does not account for infinite variance and is not in the domain of attraction of sum of independent and identically distributed random variables. However, the t model is intuitively appealing in that it extends the normal distribution model by permitting tails to be heavy and symmetric. Also, the t model is computationally straightforward and estimable using standard statistical softwares.

We have identified several key areas to be pursued. Some of these areas are straightforward such as increasing the number of explanatory variables, allowing the parameters of stable distribution to vary with the explanatory variables in a Generalized Linear Model (GLM) framework. Diagnostic testing and model checking tools need to be developed to check the adequacy of the fitted models. Similarly, the random effects can be included in the model to explain herd variability between herds. While each is an extension of the simple models demonstrated they entail estimation of many more parameters.

Our empirical application demonstrates the importance of modeling explicitly the asymmetries and heavy tails that characterize animal movement linear metrics (step length or speed) if one is to make the accurate probability statements required to manage the environmental fluctuations. Typically, elephant movement is not predictable as it is difficult to determine analytically when a step starts and ends. However, quantifying the distribution of the movement rate conditional on specific-environmental variables is one way to describe the effects of the drivers on the elephant movement. The stable regression models appears to be a useful tool for quantifying this relationship and it may have an important and practical application in assessing the value of artificial incentives in wildlife management, especially on private game ranches in South Africa.

REFERENCES

01. Bartumeus, F., 2007. Levy processes in animal movement: An evolutionary hypothesis. Fractals, 15: 151-162.
02. Lambert, P. and J.K. Lindsey, 1999. Analysing financial returns by using regression models based on non-symmetric stable distributions. J. Royal Stat. Soc., 48: 409-424.
03. Lange, K.L., R.J.A. Little and J.M.G. Taylor, 1989. Robust statistical modeling using the t distribution. J. Am. Stat. Assoc., 84: 881-896.
04. Zolotarev, V.M. and V.V. Uchaikin, 1999. Chance and Stability, Stable Distributions and Their Applications. Walter de Gruyter, Berlin, Germany, ISBN: 978-9067643016.
05. Kawai, R., 2012. Continuous-time modeling of random searches: statistical properties and inference. J. Phys. Math. Theor., Vol. 45, 10.1088/1751-8113/45/23/235004
06. Edwards, A.M., 2008. Using likelihood to test for Levy flight search patterns and for general power-law distributions in nature. J. Anim. Ecol., 77: 1212-1222.
07. Edwards, A.M., M.P. Freeman, G.A. Breed and I.D. Jonsen, 2012. Incorrect likelihood methods were used to infer scaling laws of marine predator search behaviour. PloS One, Vol. 7, No. 10. 10.1371/journal.pone.0045174
08. Edwards, A.M., R.A. Phillips, N.W. Watkins, M.P. Freeman and E.J. Murphy, 2007. Revisiting Levy flight search patterns of wandering albatrosses, bumblebees and deer. Nature, 449: 1044-1048.
09. Edwards, A.M., 2011. Overturning conclusions of Levy flight movement patterns by fishing boats and foraging animals. Ecology, 92: 1247-1257.
10. Kawai, R. and S. Petrovskii, 2012. Multi-scale properties of random walk models of animal movement: Lessons from statistical inference. Proc. Royal Soc. Math. Phys. Eng. Sci., 468: 1428-1451.
11. Duffy, K.J., X. Dai, G. Shannon, R. Slotow and B. Page, 2011. Movement patterns of African elephants (Loxodonta Africana) in different habitat types. Afr. J. Wildl. Res., 41: 21-28.
12. Loarie, S.R., R.J.V. Aarde and S.L. Pimm, 2009. Fences and artificial water affect African savannah elephant movement patterns. Biol. Conserv., 142: 3086-3098.
13. Nolan, J.P. and D. Ojeda-Revah, 2013. Linear and nonlinear regression with stable errors. J. Econometrics., 172: 186-194.
14. Feller, W., 2008. An Introduction to Probability Theory and its Applications. Vol. 2, John Wiley & Sons, New York.
15. Mandelbrot, B., 1963. The stable Paretian income distribution when the apparent exponent is near two. Int. Econ. Rev., 4: 111-115.
16. Fama, E.F., 1963. Mandelbrot and the stable Paretian hypothesis. J. Bus., 36: 420-429.
17. DuMouchel, W.H., 1971. Stable distributions in statistical inference. Ph.D. Thesis, Department of Statistics, Yale University, New Haven, Connecticut.

18. DuMouchel, W.H., 1975. Stable distributions in statistical inference: 2. Information from stably distributed samples. J. Am. Stat. Assoc., 70: 386-393.
19. DuMouchel, W.H., 1973. On the Asymptotic normality of the maximum-likelihood estimate when sampling from a stable distribution. Ann. Statist., 1: 948-957.
20. Zolotarev, V.M., 1986. One-Dimensional Stable Distributions. Vol. 65, American Mathematical Socity, USA., ISBN: 9780821898154, Pages: 284.
21. Samoradnitsky, G. and M.S. Taqqu, 1994. Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance. Vol. 1, CRC Press, Boca Raton, Florida, USA., ISBN: 9780412051715, Pages: 632.
22. Janicki, A. and A. Weron, 1994. Simulation and Chaotic Behavior of Stable Stochastic Processes. Marcel Dekker, New York, USA., ISBN: 9780824788827, Pages: 376.
23. Nolan, J.P., 1998. Parameterizations and modes of stable distributions. Stat. Probab. Lett., 38: 187-195.
24. McHale, I.G. and P.J. Laycock, 2006. Applications of a general stable law regression model. J. Applied Stat., 33: 1075-1084.
25. McCulloch, J.H., 1998. Linear Regression with Stable Disturbances. In: A Practical Guide to Heavy Tails: Statistical Techniques and Applications, Adler, R.J., R.E. Feldman and M.S. Taqqu (Eds.). Springer, Berlin, Germany, ISBN: 9780817639518, pp: 359-378.
26. Nolan, J.P., 2001. Maximum Likelihood Estimation of Stable Parameters. In: Levy Processes: Theory and Applications, Barndorff-Nielsen O.E., S.I. Resnick and T. Mikosch (Eds.). Birkhauser Verlag, Boston, Massachusetts, USA., ISBN: 978-1-4612-6657-0, pp: 379-400.
27. Osorio, F. and M. Galea, 2006. Detection of a change-point in student-t linear regression models. Stat. Pap., 47: 31-48.
28. Birkett, P.J., A.T. Vanak, V.M. Muggeo, S.M. Ferreira and R. Slotow, 2012. Animal perception of seasonal thresholds: Changes in elephant movement in relation to rainfall patterns. PloS One, Vol. 7, No. 6. 10.1371/journal.pone.0038363.
29. Vanak, A.T., M. Thaker and R. Slotow, 2010. Do fences create an edge-effect on the movement patterns of a highly mobile mega-herbivore?. Biol. Conserv., 143: 2631-2637.
30. Dai, X., G. Shannon, R. Slotow, B. Page and K.J. Duffy, 2007. Short-duration daytime movements of a cow herd of African elephants. J. Mammalogy, 88: 151-157.