An application of Stable Paretian time series models to animal movement GPS telemetry data

Dr. Robert Mathenge Mutwiri\textsuperscript{1,2}, Prof. Henry Mwambi\textsuperscript{1} and Prof Rob Slotow\textsuperscript{3}

1. School of Mathematics, Statistics and Computer sciences, University of KwaZulu Natal, South Africa,  
2. School of Pure and Applied Sciences, Kirinyaga University College, Kenya,  
3. School of Life Sciences, University of KwaZulu Natal, South Africa

Abstract

Many time series models in movement ecology are often based on the assumption that the random variables follow a Gaussian distribution. It is well known that empirical data sets have infrequently occurring rare events and cannot be modelled with the Gaussian distribution. Such data sets can be described by ARMA-GARCH models with an appropriate underlying distribution that can cater for skewed and heavy tailed distributions. In this study we compare four ARMA-GARCH models as potential candidates for modelling heavy tailed and skewed time series data of animal movement step lengths. Estimation difficulties have however hindered the use of stable Paretian distributions among practitioners in fields such as movement ecology. Four distributional hypotheses are tested: the Gaussian, Student’s t, skewed Student’s t and the Stable Paretian distributions and fitted to time series step lengths of single female elephant herd. We found that the stable Paretian distributed ARMA-GARCH model performs better than the Gaussian and student’s t distribution ARMA-GARCH models.

Keywords: Stable Paretian, Heavy tailed data, Animal movement, biological signals
1 Introduction
Many ecological models of animal movement rely on the assumption that step lengths follow a Normal distribution (Gurarie et al., 2009). However, animal movement data often depart from the Normality assumption, in that their marginal distributions are heavy tailed and skewed. Animal location data are being collected at increasing high resolution (0.1-30 minutes sampling intervals of large mammals are now quite common) over several seasons. For such data, first order autoregressive models that assume normal distribution may miss important properties of the data. The use of stable distributions in movement ecology has been mentioned in a few studies via the works of Viswanathan et al., (1996) and emphasized by Bartumeus et al., (2007). Since then, the power law a simpler variant of the stable distribution has been used to model the step lengths of animal movements (Edwards et al., 2007). Such an approach ignores the temporal correlation potentially leading to spuriously precise estimates of parameters (Edwards et al., 2007) Stable distributions in general have not gained much attention in movement ecology due to lack of closed form density and computational difficulties (Kawai, 2012). Biological signals modelled using these models can provide a useful approach for statistically detecting and characterizing the temporal dependency in animal movement data.

While different heavy tailed distributions such as the Weibull and Log-normal distribution can be used for modelling animal movement metrics, stable distributions are preferred due to the generalized Central Limit Theorem. According to this theorem, the limiting distribution of a sum of i.i.d random variables is stable (Nolan, 1998). A notable feature of the stable distribution is the slow decay in the tails so that large rare steps are captured naturally. All the stable distributed variables have the nice property of additive invariance adding two or more independent realisations yields another variable of this type. The normal distribution is a limiting case of the generalised central limit theorem when $\alpha=2$ (Zolotarev, 1986). These and many other properties make stable distribution more tractable to the analysis of animal movement data. Thus they have gained popularity in modelling heavy tailed data (Curto, 2009) in finance and economics and a few cases in animal movement (Kawai, 2012). Animals typically move in a non-random manner with short clustered steps when foraging or resting and long rare steps when migrating or in search of scarce food resources like water resulting to heavy tailed and skewed step lengths. These data sets are autocorrelated a property which most ecologists tend to eliminate though de-trending or filtering (Cushman, 2005). However, Drayet al., (2010) notes that such autocorrelation is an intrinsic
property of biological data thus eliminating it could reduce the relevance of ecological studies. Traditional analysis often assume uncorrelated or weakly correlated temporal structure in animal movement time series constructed using sequential location data (Dray et. al., 2010).

The conditional distribution of step lengths is assumed to be Normal in Autoregressive models. However, this model specification is not proper for many animal movement time series because of the leptokurtic nature of the data (Bartumeus et. al., 2007). Therefore, distributions such as the student’s t, Log-normal and the Laplace have been suggested (Meerschaert et. al., 2009). Previous studies have shown that movement step length data set is heavy tailed and skewed but little has been done to evaluate time series models that capture both the skewness and heavy tails properties. It is this gap that this study seeks to contribute by proposing an ARMA-GARCH model with stable Pareto distributed innovations to analyse animal movement data sets.

Recent methodological developments have improved the ability of statistical models to handle the biological complexity of animal movement data (Nathan, 2008). This paper examines the statistical properties of the step lengths, using a set of symmetrical and asymmetrical time series. These tools are applied to high-frequency telemetry data within the framework of movement ecology.

The remainder of this paper is organized as follows. The next section discusses ARMA-GARCH models with Normal, student’s t, skewed student’s t and stable innovations. In section 2, we explain the statistical properties of animal movement step lengths data used in this paper and presents the initial findings. We also discuss parameter estimation and compare the goodness-of-fit of the fitted models. Section 4 summarizes the findings and gives the concluding remarks.

2 Statistical models
In this section, we describe time series models that can be adopted to analyse animal movement step length time series data in ecology. We also suggest an extension to a new class of time series models based on the stable distribution whose density has the capacity to capture both skewness and heavy tailed nature of animal movement data.

Let $y_t$ be an ARMA(p,q)process of order p and q given mathematically by the equation:

$$y_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i y_{t-i} + \sum_{j=1}^{q} \beta_j \mu_{t-j} + \epsilon_t$$

(1)
where $a_i$ and $b_j$ are the model parameters and $i = 0, \ldots, p; j = 0, \ldots, q$ are the orders of the autoregressive and moving average process respectively, $u_t$ is the error term and $y_t$ is the step length of an individual animal at time $t$.

The generalized autoregressive heteroscedasticity (GARCH) models can be used to model animal movement data by expressing the conditional variance as a linear function of past information, allowing the conditional heteroscedasticity of step lengths. Animals alternates between fast movement while migrating and slow movement while foraging or resting which can be visualized as small clustered steps with long rare steps whose outcome is heavy tailed and skewed data. The assumption that the residuals $u_t$ in equation (1) are i.i.d is typically not valid in ecological time series. The width of the residuals is clustered and depends on the time itself. This property is called volatility clustering. Engle and Bollerslev introduced a class of GARCH(r,s) models which covers this property (Bollerslev, 1986). The volatility clustering $u_t$ can be modelled by expressing the residual terms as follows:

$$u_t = \gamma_t \epsilon_t$$

with the recursive relation

$$\gamma_t^2 = \omega + \sum_{m=1}^{r} \psi_m u_{t-i}^2 + \sum_{n=1}^{s} \phi_n \gamma_{t-j}^2, \quad u_t | U_{t-1} \sim N(0, \gamma_t^2) \quad (2)$$

where $\omega > 0$ is the constant coefficient of the variance equation, $\epsilon_t$ is a sequence of i.i.d random variables with mean zero and variance 1, $\psi_m > 0$ and $\phi_n \geq 0$ are the ARCH and GARCH parameters respectively and $\sum_{m=1}^{r} \psi_m + \sum_{n=1}^{s} \phi_n < 1$, $m=1, \ldots, r$ and $n=1, \ldots, s$ are the lags of the ARCH and GARCH effects and $u_t | U_{t-1}$ is the history of the movement. Here it is understood that $\alpha_i = 0$ for $m > r$ and $b_j = 0$ for $n > s$. The latter constraint on $\alpha_i + \beta_j$ implies that the conditional variance of $u_t$ is finite, whereas its conditional variance $\gamma_t^2$ evolves overtime. Thus, volatility clustering depends on its previous values and on the squared residuals $\epsilon_t^2$.

While a GARCH model can describe volatility clustering one still needs an adequate model to estimate the conditional mean of the time series. If the prediction of the conditional mean is not
reasonable, the construction of the conditional volatility is not possible. We combine the equations (1) and (2) to form an ARMA-GARCH(p,q,r,s) models defined by equation:

\[ y_t = a_0 + \sum_{i=1}^{p} a_i y_{t-i} + \sum_{j=1}^{q} b_j u_{t-j} + \epsilon_t \gamma_t \]  

where \( \gamma_t \) is given by equation (2) and \( \epsilon_t \) is assumed to have a mean zero and variance 1. For \( \epsilon_t \), we can assume various probability density functions to test different types of ARMA-GARCH models. For example, the normal, the Student’s t, the skewed Student’s t and the asymmetric stable distributions can be used. In the Gaussian case, \( \epsilon_t \sim N(0,1) \) and in the student’s t case, \( \epsilon_t \sim t(v) \) where \( v \) is the degrees of freedom scale parameter equal to one. We denote the Gaussian model as Normal-ARMA-GARCH, the student’s t model as t-ARMA-GARCH and the skewed student’s t model as st-ARMA-GARCH model. We refer to these models as the conventional ARMA-GARCH models. An important property that makes ARMA-GARCH process attractive to models of animal movement ecology data is that the its tail is heavier than that of a normal distribution which allow a time varying volatility clustering of animal movement step lengths. The clustering of animal movement step lengths is known as Lévy flight pattern in ecology (Kawa, 2012). It is expected that the skewed student’s t ARMA-GARCH model which has slightly heavier tails will be able to capture rare large steps better than the student’s t and the normal ARMA-GARCH process. However, models that can better capture both heavy tails and skewness are still of interest.

It has been argued in the literature by Bartumeus et. al., (2007) that a model that captures both the heavy tails and skewness inherent in ecological data is the four parametric family of stable distributions. However, the stable density for a random variable \( X \) does not have a simple mathematical description, so it is instead represented using its characteristic function given by

\[ E[\exp(itx)] = \begin{cases} 
\exp(-(\gamma|t|^\alpha[1 + i\beta \text{sgn}(t)\tan^{\alpha/\gamma}\frac{|t|^{1-\alpha}}{\alpha}(1-\alpha - 1)]) + i\mu t) & \alpha \neq 1 \\
\exp(-\gamma|t|[1 + i\beta \text{sgn}(t)\frac{\ln|t|}{\alpha} + i\mu t]) & \alpha = 1
\end{cases} \]  

Where \( \alpha \in (1, 2) \) is the index of stability, \( \beta \in [-1, 1] \) is the skewness parameter, \( \gamma > 0 \) is the scale parameter and \( \mu > 0 \) is the location. The distribution is symmetric about the location \( \mu \) if \( \beta = 0 \) and the characteristic exponent \( \alpha \) determines the thickness of the tails. When \( \alpha=2 \), the underlying stable Paretian distribution is Normal distribution: \( N(0, 2) \), with finite moments of all orders. As \( \alpha \) tends to 0 from 2, the tail of the stable distribution becomes increasingly fatter than the Normal case and thus, the scale parameter \( \gamma \) is used as the analog of the variance. In this paper we note that the \( k^{th} \)
moments of the stable distribution exist if \( k<\alpha \) and we assume that \( E[U]<\alpha \). In other words, we restrict \( \alpha \) to be in the interval \((1, 2]\) throughout.

Assuming that step lengths of animal movement follow a stable Paretoian distribution, we define the volatility model of the stable ARMA-GARCH process \( y_t \) denoted by \( S-ARMA-GARCH \) and given by the equation

\[
y_t = \psi_0 + \sum_{i=1}^{p} \psi_i |u_{t-i}| + \sum_{j=1}^{q} \phi_j y_{t-j}
\]

where \( u_t = \gamma_t^{1/\alpha} e_t \) and \( e_t \) is specified to have a stable density.

The \( S-ARMA-GARCH(p,q,r,s) \) process defined by equations (1) and (5) with \( \alpha \in (1,2] \) has a unique strictly stationary solution if \( \alpha_i > 0, i=0,1,...,r, \beta_j > 0, j=1,...,s \) and the measure of volatility persistence, \( V_s = \lambda_{(\alpha,\beta,\gamma,\lambda)} \sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j \), satisfies \( V_s \leq 1 \) (Curto, 2009), where

\[
\lambda_{(\alpha,\beta,\gamma,\lambda)} = E[\varepsilon_t] = \frac{2}{\pi} \Gamma \left( 1 - \frac{1}{\alpha} \right) \left( 1 + \frac{2}{\alpha} \right)^{1/2} \cos \left( \frac{1}{\alpha} \arctan \left( \frac{\beta_{\alpha,\beta,\gamma,\lambda}}{\alpha} \right) \right)
\]

and \( \tau_{\alpha,\beta,\gamma,\lambda} = \beta_{\tan} \frac{\pi \alpha}{2} \). If \( V_s \) is strictly less than one, this implies a conditional volatility equation where the impact of long rare steps dies out over time. In practice, the estimated volatility persistence \( V_s \) tends to be quite close to one for highly volatile series which would offer a reasonable description of the animal movement data.

### 3 Parameter estimation

The parameters were estimated using the maximum likelihood method and it was assumed that the innovations \( \varepsilon_t \) are Normal, student’s t, skew student’s t and stable Paretoian and that \( \gamma_t \) satisfies ARMA-GARCH recursions (3) and (5), respectively. We follow the algorithm of Paolella and Mittnik (1999) who approximates the stable Paretoian density function through the Fast Fourier Transform (FFT) of the characteristic function.

To compare the fitted models, we employ three likelihood based goodness of fit criterion proposed by Mittnik and Paolella(2003). The first is the maximum log-likelihood value obtained from ML estimation (MLE). This value allows us to judge which model is more likely to have generated the data. The second method is the bias-corrected Akaike (Akaike, 1978) information
criterias (AICC) and the third method is the Schwarz Bayesian criteria (SBC, Schwarz, 1978) given by:

\[
AICC = -2\ln L(\hat{\theta}) + \frac{2n(k+1)}{n-k+2}, \quad SBC = -2\ln L(\hat{\theta}) + \frac{2\ln(n)}{n},
\]

where \(\ln L(\hat{\theta})\) is the maximum log-likelihood value, \(n\) is the number of observations and \(k\) is the number of parameters. The model with a lower value for these information criteria is judged to be preferable. These criteria are also recommended by Curto et al., (2009).

4 Application to elephant movement data

To illustrate the utility of the reviewed methods, we analyse a time series data set of an individual female elephant herd (Africana loxadonta) collected via radio-telemetry (Birkett et al., 2012 for details on data collection) from May 2006 to April 2009. Animal capture was undertaken using chemical immobilisers by South African National parks staff, following approved ethical procedures (University of KwaZulu Natal, Ref. 009/10/Animal). These elephants inhabited Kruger national park isolated by paved roads and rivers in South Africa. Africana loxadonta is an elephant species mostly found in Southern of Africa.

One of the objectives of the radio-telemetry study was to assess the effects of autocorrelation in the presence of heavy tails and skewness in the step lengths of elephant movement. Spatial coordinates of the elephant location were obtained using global positioning system (GPS) receivers and differential correction techniques were applied to improve their accuracy. These animals were typically located at 30 minutes intervals between May 2006 and April 2009 yielding 35272 observations for that herd. Step lengths (linear metrics) of animal movement were derived using the tracking tool in ArcGis 10.0 and 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.

To summarize the statistical properties of animal movement step lengths, firstly we carry a visual inspection of the time series graphs of step lengths of the herd. Secondly, relevant descriptive statistics are computed (including the Jarque-Bera normality test). Thirdly, we carry out two unit root tests: the augmented Dickey-Fuller (ADF) and the KPSS tests to decide whether the series is stationary. Fourthly, we discuss the autocorrelation of step length and we test the presence of nonlinearities in the movement data.
5 Results of data analysis

Figure 1 gives the time series plot of step lengths of herd AM108 and shows the long run movement pattern of elephant spanning a period of 3 years. The plots show relatively large number of high spikes in the step Lengths which shows the presence of clustering of steps and the possibility that the step length innovations of the herd are non-Gaussian.

Table (1) summarizes the basic descriptive statistical properties of the data. The mean step length is 213 meters and the median is 145 meters signifying the presence of outliers in the data. The step lengths appear to be somewhat asymmetric as reflected by the positive skewness estimate which implies the presence of rare large steps in the series. The kurtosis is positive and higher than that of a standard Gaussian distribution which is 3, showing the presence of heavy tails for these empirical distributions.

The Jarque-Bera normality tests, as shown in Table (1), are far beyond their critical value at 5%, indicating that the normal distribution hypothesis for the empirical step lengths should clearly be rejected. We computed the ADF and the KPSS tests as shown in Table (2) to determine the stationarity of elephant movement data.

The ADF unit root hypothesis is rejected at 5% level of significance and the KPSS stationarity hypothesis is rejected at 10% level of significance, we conclude that the elephant movement is nonstationary. For the purpose of this work, and according to the statistics obtained for the step length time series data of herd AM108, there is evidence for higher order lags but in this study we shall confine the analysis to the autoregressive models of order 1 as in Gurarie et al., (2009). The first order autocorrelation is positive which suggests that the elephants have some memory although it is very short and there is a portion of movement step lengths that is predictable although it is very small. So the Lévy flight foraging hypothesis does not hold strictly as stipulated in Bartumeus et. al., (2007). Owen-Smith et al., (2010) suggest that this behaviour could be due to environmental heterogeneities.

Table 1: Descriptive analysis of two female elephant movement step length in Kruger National Park collected every 30 minutes

<table>
<thead>
<tr>
<th>Observations</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>35272</td>
<td>5.039</td>
<td>2616</td>
<td>213</td>
<td>145</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Stats_Stddev</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Jarque-Bera</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.226</td>
<td>2.25</td>
<td>7.39</td>
<td>164028.9</td>
<td>0.000</td>
</tr>
</tbody>
</table>
In order to determine our mean equation, we note that the lag value which has the minimum AIC value corresponding to the ARMA terms was chosen to model the step lengths, thus supporting the choice of our model.

Table 2: KPSS tests and Augmented Dickey-Fuller

<table>
<thead>
<tr>
<th>Test</th>
<th>Statistic</th>
<th>lag</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>KPSS</td>
<td>0.3567</td>
<td>43</td>
<td>0.096</td>
</tr>
<tr>
<td>ADF</td>
<td>-31.378</td>
<td>32</td>
<td>0.010</td>
</tr>
</tbody>
</table>

Even though the series of step lengths seems to be correlated over time, the Ljung-Box statistic for up to 20th order serial correlation is highly significant at any level for the animal movement linear metrics, suggesting the presence of strong nonlinear dependence in the data. From the time series plots (2), volatility clustering of steps are clearly visible suggesting the presence of conditional heteroscedasticity which implies that nonlinearities must enter though the variance of the processes (Mittnik and Paolella, 2003). Thus the nonlinear dependencies can be explained by the presence of conditional heteroscedasticity and one might represent this behaviour using a model in which successive conditional variances of step lengths are positively autocorrelated and such specification is consistent with the optimal foraging hypothesis of animal movement.

Lagrange multiplier (LM) test can be used to test formally the presence of conditional heteroscedasticity and evidence of ARCH effects. The LM test for a first order linear effect as shown in Table (2) suggests that the herd exhibit ARCH effects as the null hypothesis is rejected. In the next section we use ARMA-GARCH models to describe the conditional distribution of step lengths.

Table 3: Step Length autocorrelation

<table>
<thead>
<tr>
<th>Test</th>
<th>Statistic</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ljung-Box</td>
<td>98.02398</td>
<td>0.000</td>
</tr>
<tr>
<td>LM Arch</td>
<td>133.6647</td>
<td>0.000</td>
</tr>
</tbody>
</table>
Figure 1: Time series plots of the step lengths of Herd AM108 of GPS-derived telemetry data collected from Kruger National Park South Africa (May 2006 - April 2009)

Figure 2: Histogram plot of the step lengths of Herd AM108 GPS-derived telemetry data collected from Kruger National Park South Africa (May 2006 - April 2009)
Figure 3: QQ plots of standardized residuals from an ARMA-GARCH (1, 1, 1, 1) to step lengths of elephant movement.

6 Parameter estimates and interpretation
Table 3 shows the estimation results for the stable Paretian, normal, student’s t and skewed student’s t distribution ARMA-GARCH (1, 1, 1,1) models for herd AM108. The estimate of the intercept which is denoted by $a_0$ is statistically significant for all the models except for the stable ARMA-GARCH model. The autoregressive parameter denoted by $a_1$ is positive and statistically significant for all the four models implying that there is a positive autocorrelation in the data. This indicates that the AR (1) model is able to model the conditional mean of animal movement linear metrics as suggested by Gurarie et. al., (2009).

Both $\varphi_1$ and $\phi_1$ are highly significant which implies that there is persistent volatility clustering in the step length time series data of the herd. These results indicate that as the animals increase their step lengths during the search for scarce food resources or migration so does volatility. Since the step length is the displacement between two positions covered by the animal within a given period of time is assumed to be a proxy measure of animal movement pattern or unobserved information flow. We conjecture that as the food resources becomes scarcer or decreases, the volatility also increases. This clearly makes sense because the animal is in the search mode state.
From the stable distribution parameterisation, the estimates of the shape parameter \( \alpha \) is statistically significant at 5% and less than 2 which means that the step length time series of the herd have heavy tailed patterns and thus rejecting a Gaussian hypothesis. The estimates of the skewness parameters \( \beta \) are positive and significant which means that the step length time series are skewed. These findings agree with the results of the descriptive analysis which indicates that the data set is highly skewed and leptokurtic as shown in Table (1).

One important aspect of these estimates relates to the connection between the skewness and kurtosis of a normal distribution, the asymmetry and the tail index parameter of the stable distribution (Mittnik and Paolella, 2003). Recalling the sample skewness and kurtosis obtained in Table (1), one should expect that a fit with stable distribution will probably give stable skewness with the same sign and with tail index less than 2. Indeed this happens in Table (1). We note that there is no quantitative connection between the sample kurtosis and the population kurtosis as it is undefined under the stable distribution. The real difficulty when analysing the ecological data sets is that we do not know what the exact distribution of the underlying population is. Sampling variations make this problem even more complicated (Kawai, 2012).

Bases on the log-likelihood, AICC and the SBC results of Table (3), we find that the ARMA-GARCH(1,1,1,1) model with stable distributed innovations performed better than the normal followed by the skewed student’s t and student’s t ARMA-GARCH models respectively. We also note that the skewed student’s t performed better that the student’s t and the normal ARMA-GARCH models respectively. This implies that the asymmetric models are better suited for analysing heavy tailed and skewed data of animal movement like this one. This finding indicates that asymmetric models perform better with heavy tailed and skewed data. The results further confirm the advantage of using the stable distributions that can take into account the non-normal nature of the time series data examined. However, we point out that despite the computational challenges highlighted by Kawai (2012), stable distribution through the index of stability and skewness parameter can capture more properties of animal movement data and provide more biological insights than the conventional normal models.

7 Conclusion
We investigated the empirical 30 minutes time series data of elephant movement step lengths collected from Kruger national park using both asymmetric and symmetric distribution in
ARMA-GARCH models. We note that models that appropriately capture the heavy tails and skewness properties of animal movement and other ecological applications remain largely unexplored. Stable Pareti an distribution has rich properties that can appropriately capture the skewness and heavy tailed nature of animal movement data better than the conventional normal distribution. Using asymmetric models, we are able to extract biological insights from a rich statistically challenging empirical GPS animal location data which could otherwise be lost if the data was be subjected to treatment to remove the autocorrelation. These models provide an objective way of modelling animal movement data sets without loss of biological information due to data treatment as proposed by Cushman et. al., (2005). By incorporating the Biological signals through stability index and skewness parameter of the stable distribution, the model provides an approach of statistically detecting and characterizing the temporal dependency in animal movement facilitating model development.

However, stable Paretian time series models are computationally challenging and no well-developed softwares have been implemented. We note that the skewed student’s t distribution can be a better alternative as it comes second in goodness of fit to the stable Paretian model. Models of higher order lags can be of interest as a future research. We also note that our models can be extended through the random coefficient autoregressive models to investigate the effects of several herds. We leave this and other considerations for future investigation.

References
24 Abi Vanak, Maria Thaker, and Rob Slotow. Do fences create an edge-effect on the movement
