



## ARTICLE

# Sine-skewed von Mises- and Lindley/Gumbel models as candidates for direction and distance in modelling animal movement

 Gopika Devi Ramkilawon,<sup>\*,1,2</sup>  Johannes T. Ferreira,<sup>1,2</sup> and  Najmeh Nakhaeirad<sup>1,2</sup>

<sup>1</sup>University of Pretoria, Department of Statistics, Pretoria, South Africa.

<sup>2</sup>Centre of Excellence in Mathematical and Statistical Sciences (CoE-MaSS), Johannesburg, South Africa.

\*Corresponding author. Email: gopika.ramkilawon@up.ac.za

(Received: May 19, 2022; Revised: August 1, 2022; Accepted: August 12, 2022; Published: June, 2023)

### Abstract

Movement of animals is often characterised by direction (measured on the circle) and distance (measured on the real line); but traditionally employed models often do not account for potential asymmetric directional movement, or departures from the usual von Mises assumption for direction and gamma/Weibull assumptions for distance. This paper focuses on the modelling of circular data in this animal movement context relying on a previously unconsidered circular distribution (the sine-skewed von Mises) which provides a platform for departures from symmetry. In addition, alternative models to usual distance assumptions are considered, namely the power Lindley (as a mixture of gamma and Weibull distributions) as well as a Gumbel candidate. Computational aspects and investigations of this joint modelling (presented as a consensus model) are highlighted, accompanied by an extensive bootstrap study. A general hidden state Markov model is used to incorporate these essential components when estimating via the use of the EM algorithm, and goodness of fit measures verify the validity and viable future consideration of the newly proposed theoretical models within this practical and computational animal movement environment.

**Keywords:** Angular regression, Asymmetry, Bootstrap, Directional statistical model, EM algorithm, Hidden Markov model, Spatio-temporal time series.

## 1. Introduction

Understanding an animal's movement and behavioural pattern is essential but is often complex. The considerations of several factors such as climate change, pollution, land use and proliferation of other invasive species and diseases are often difficult to quantify (Van Niekerk, 2018). The primary

challenge in animal movement modelling considering these factors and various environmental features identified within a habitat is the validation of a statistical approach since the true and exact natures of the animals are rarely known. To this end, most quantitative and statistical techniques fall short in appropriately accounting for sufficient directional targets influencing the animal's behaviour of movement.

There is no single universally accepted method for the inference of the behavioural states i.e., "resting", "foraging", and "travelling", which generate the recorded geographical location of animals. Various robust statistical techniques and pliable animal movement models are required for the modelling of animal movement (Nicosia *et al.*, 2017). Displacement may be expressed as the direction and distance between two consecutive points with a circular-linear process modelling the movement within a two-dimensional discrete environment. Hidden Markov Models are popular considerations within animal movement models to segment movement pathways into latent behavioural states due to their flexibility and practicality (Goodall *et al.*, 2017). In animal movement modelling, the theoretical models for the directional movement and distances travelled by the animal is of utmost importance. This extends to animal movement models utilising Markov hidden states to classify the latent behavioural states of the animal's movement. The gamma and Weibull distributions are popular candidates for distance modelling in conjunction with the von Mises distribution for directional modelling. The directional model includes the von Mises distribution which is the most popular choice in animal movement modelling (Nicosia *et al.*, 2017). The symmetric, unimodal nature of the von Mises distribution makes it an attractive choice when modelling direction (Jammalamadaka & Sengupta, 2001), specifically when the directional data of the animal shares these symmetric, unimodal characteristics. Furthermore, the simplicity of the form of the pdf (probability density function) makes it an easily implementable choice within an animal movement models (Nicosia *et al.*, 2017).

### 1.1 Model Outline

Consider a data set which consists of the time series:

$$\{\gamma_t, d_t, \mathbf{x}_t, \mathbf{z}_t, t = 0, \dots, T\} \tag{1}$$

where

- $\gamma_t \in [0, 2\pi)$  represents the direction between the animal's location from time step  $t$  to time step  $t + 1$ .
- $d_t \geq 0$  represent the distance between the animal's location from time step  $t$  to time step  $t + 1$ .
- $\mathbf{x}_t = (x_{1t}, \dots, x_{nt})$  are the values of the  $n$  explanatory angular variables which are measured which could be used to predict  $\gamma_t$  or  $d_t$ .
- $\mathbf{z}_t = (z_{1t}, \dots, z_{nt})$  are the values of the  $n$  explanatory real variables which are measured which could be used to predict  $\gamma_t$  or  $d_t$ .

The explanatory variables  $x_{jt}$  and  $z_{jt}$ ,  $j = 1, \dots, n$  are associated to the direction and distances of the animal being observed with respect to the position of the animal at the previous time step. The set of all observed directions and distances are denoted by  $(\gamma_{0:T}, d_{0:T}) = \{(\gamma_t, d_t), t = 0, \dots, T\}$  for simplicity. The observed information for the variables involved, i.e. directions, distances and explanatory variables are denoted by a filtration  $\mathcal{F}_t^0$ . The reason for this denotation is to differentiate the observed data from unobserved data (ibid.).

The hidden process  $S_t$ ,  $t = 1, \dots, T$  is used in this model to represent the behaviour i.e. state in which the animal is in at time  $t$ . This is due to the fact that animals exhibit multiple behavioural patterns over time. These changes in behavioural patterns are not observed over time, so this hidden

state will account for these behavioural changes over time. The set of the possible hidden states are denoted by  $S_{0:T} = \{S_0, \dots, S_T\}$ . The complete data filtration is denoted by  $\mathcal{F}_t^c$  which is the filtration generated by  $\mathcal{F}_t^0$  and the hidden information until time  $t$ . The joint pdf for the complete data is therefore:

$$f(y_{0:T}, d_{0:T}, S_{0:T}) = \prod_{t=1}^T g(S_t | \mathcal{F}_{t-1}^c) k(y_t, d_t | S_t, \mathcal{F}_{t-1}^c) \tag{2}$$

where  $g(\cdot)$  represents the pdf of the hidden data and  $k(\cdot)$  represents the pdf of the observed data. Specific assumptions have been made for the directional-distance proposals to remain theoretically and empirically practical. These assumptions are detailed by *ibid.* and readers are encouraged to view this paper for further details.

### 1.2 Circular-regression model

The direction  $y_t$  is assumed to have its mean direction being dependent on  $y_{t-1}$  together with other explanatory variables and a homogeneous error (*ibid.*). This homogeneous error depended on a fixed concentration parameter  $\kappa$ . A consensus error model was adapted in order to combat multi-modal log-likelihoods which affected the estimation procedure outlined by Nicosia *et al.*, 2016. This method is adapted for the sine skewed von Mises (ssvM) distribution for comparison where this was originally designed by Nicosia *et al.*, 2017 for the von Mises distribution.

A consensus error model for the direction  $y_t$  with knowledge of the animal being in state  $k$  depends on the vector (*ibid.*)

$$V_t^k = \kappa_0^{(k)} \begin{pmatrix} \cos(y_{t-1}) \\ \sin(y_{t-1}) \end{pmatrix} + \sum_{i=1}^n \kappa_i^{(k)} z_{it} \begin{pmatrix} \cos(x_{it}) \\ \sin(x_{it}) \end{pmatrix}, \quad t = 1, \dots, T, \tag{3}$$

where  $\kappa^{(k)} = (\kappa_0^{(k)}, \dots, \kappa_n^{(k)})$  represent the unknown parameters dependent on state  $k$ . In terms of the ssvM distribution, its mean direction denoted by  $\mu_t^{(k)}$  is the direction of  $V_t^{(k)}$ . The parameters  $\kappa_i^{(k)}$  quantify how target  $i$  influences the animal's directional movement. Further detail of the circular-regression model is outlined in Section 2.1.

### 1.3 Paper outline

The distance models considered in literature include the Gamma and Weibull distributions (Lan-grock *et al.*, 2012; Nicosia *et al.*, 2017; Van Niekerk, 2018) where distance refers to the distance travelled by the animal between consecutive time steps i.e. their step lengths. Both distributions have a shape and scale parameter with appealing pdf characteristics which make them good candidates for the modelling of distances, and need to be able to accurately measure these distances within their respective movement states i.e. hidden Markov states. However, these models are sometimes limited in their flexibility and usage when accounting for irregular movement, and theoretical limitations of the distributions themselves such as lesser degrees of skewness and kurtosis. To this effect, two alternative distance models are considered in this paper, namely the power Lindley distribution and the Gumbel distribution. To further circumvent theoretical underpinnings of the von Mises distribution in terms of skewness and unimodality, the sine-skewed von Mises distribution is introduced as a contender for modelling direction of animal movement - in both cases, are introduced and described in Section 2. The models are comparatively investigated and applied to real caribou data after the necessary description of the Expectation Maximisation (EM) algorithm in Sections 3 and

4 respectively, and a bootstrap study sheds further insight into these computational considerations of new theoretical contenders. Throughout, computational considerations remain at the forefront of this paper’s main propositions to allow an interpretative contribution for the future practitioner. The paper concludes with final thoughts in Section 5.

## 2. Alternate models for direction and distance

In this section, focus is drawn to the ssvM distribution as an alternative in direction modelling and the power Lindley distribution as an alternative in distance modelling in this specific animal movement model context.

### 2.1 Sine-skewed von Mises distribution

The pdf of the ssvM distribution is given by (Abe & Pewsey, 2011):

$$f(y_t) = \frac{\exp(\ell_t \cos(y_t - \eta))}{2\pi I_0(\ell_t)} (1 + \phi \sin r(y_t - \eta)), \quad -\pi \leq y_t < \pi, \tag{4}$$

where  $\ell_t \geq 0$  is the concentration parameter,  $-\pi \leq \eta < \pi$  is the location parameter,  $\phi \in [-1, 1]$  is the skewness parameter and  $I_0(\ell_t)$  is the modified Bessel function of the first kind of order 0 with form (Jammalamadaka & Sengupta, 2001):

$$I_0(\ell_t) = \frac{1}{2\pi} \int_0^{2\pi} \exp(\ell_t \cos(y_t)) dy_t.$$

The pdf of the ssvM is skewed to the left when  $\phi > 0$  and skewed to the right when  $\phi < 0$ . The parameter  $r = 1, 2, 3, \dots$  is a positive integer value which influences the multimodality of the pdf. When  $r = 1$ , the pdf is both unimodal and bimodal, but for  $r > 1$  the pdf is always multimodal. When the skewness parameter  $\phi = 0$ , the model simply reduces to the usual von Mises distribution (Mardia & Jupp, 2009). This parameter enrichment adds value within the practical considerations as an already existing model is encapsulated within this proposed generalised circular consideration.

When  $\ell_t$  increases with  $\phi$  remaining constant, the peak of the ssvM pdf increases and is more centred around its mean. However, as  $\ell_t$  decreases, the flatter the shape of the ssvM pdf becomes and the peak of the pdf decreases. The cdf of the ssvM is defined by (Abe & Pewsey, 2011):

$$F(y_t) = F_0(y_t) + \frac{\phi}{2\pi \ell_t I_0(\ell_t)} (\exp(-\ell_t) - \exp(\ell_t \cos r(y_t - \eta))),$$

$$-\pi \leq y_t < \pi, \ell_t \geq 0, -1 \leq \phi \leq 1, -\pi \leq \eta < \pi$$

where  $F_0(y_t)$  is the cdf of the base von Mises distribution. The p-th trigonometric moment of the ssvM is given by (ibid.):

$$\varphi_p = \exp(ip\eta) A_p(\ell_t) \left( 1 + \frac{ip\phi}{\ell_t} \right)$$

where  $A_p(\ell_t) = \frac{I_p(\ell_t)}{I_0(\ell_t)}$  is the  $p^{\text{th}}$  trigonometric moment of the von Mises distribution. The ability to account for bimodality, skewness and asymmetry is an appealing feature of the ssvM, therefore making it an appropriate consideration for directional movement modelling.

In context of the circular-regression model outlined in Section 1.2, the concentration parameter of the ssvM distribution denoted by  $\ell_t^{(k)}$  is the length of the vector (3) for the consensus error model.

This indicates that the concentration parameter  $\ell_t^{(k)}$  is dependent on the level of agreement between the directional targets under consideration. The pdf of the direction given the observed data and concentration parameter is (Nicosia *et al.*, 2017):

$$f_k(y_t | \mathcal{F}_{t-1}^c; \kappa^{(k)}) = \frac{\exp\{\ell_t^{(k)} \cos(y_t - \mu_t^{(k)})\}}{2\pi I_0(\ell_t^{(k)})} (1 + \Phi_t^{(k)} \sin(y_t - \mu_t^{(k)})), \quad t = 1, \dots, T. \quad (5)$$

Since  $\mu_t^{(k)}$  and  $\ell_t^{(k)}$  are the direction and length of the same vectors, from (3) and (5):

$$f_k(y_t | \mathcal{F}_{t-1}^c; \kappa^{(k)}) = \frac{\exp\{\kappa_0^{(k)} \cos(y_t - \gamma_{t-1}) + \sum_{i=1}^n \kappa_i^{(k)} z_{it} \cos(y_t - x_{it})\}}{2\pi I_0(\ell_t^{(k)})} \times \left( 1 + \Phi_t^{(k)} \left( \sin(y_t - \gamma_{t-1}) + \sum_{i=1}^n \kappa_i^{(k)} z_{it} \sin(y_t - x_{it}) \right) \right), \quad t = 1, \dots, T. \quad (6)$$

This parameterisation is of importance since it results in a numerically stable model with  $\kappa_i^{(k)}$  being the canonical parameters of a distribution which belongs to the exponential family.

### 2.2 Power Lindley distribution

If  $V$  has a Lindley distribution,  $D_t = V^{\frac{1}{\alpha}}$  has the power Lindley distribution with the following pdf (Ghitany *et al.*, 2013):

$$f(d_t) = \frac{\alpha\beta^2}{\beta + 1} (1 + d_t^\alpha) d_t^{\alpha-1} \exp(-\beta d_t^\alpha) = p s_1(d_t) + (1 - p) s_2(d_t),$$

where  $p = \frac{\beta}{\beta+1}$ ,  $s_1(d_t) = \alpha\beta d_t^{\alpha-1} \exp(-\beta d_t^\alpha)$  and  $s_2(d_t) = \alpha\beta^2 d_t^{2\alpha-1} \exp(-\beta d_t^\alpha)$ . Thus, the power Lindley distribution is a two-component mixture of a Weibull distribution with shape parameter  $\alpha > 0$  and scale parameter  $\beta > 0$  and a generalised gamma distribution (Bain & Engelhardt, 1987) with shape parameter of 2 and a scale parameter of  $\beta > 0$  with a mixing proportion of  $0 \leq p \leq 1$  as defined above.

The intrigue about the power Lindley distribution lies in its shape characteristics and its mixture distribution with the gamma and Weibull distributions whilst remaining tractable as a contender since these distributions are the most popular distance models in animal movement. The moments maximum likelihood estimators are explained in detail by Ghitany *et al.*, 2013. Ghitany *et al.*, 2013 illustrates that the skewness of the power Lindley distribution may be negative which offers flexibility over the traditional Lindley distribution and other parametric considerations for distance modelling. This flexibility in accounting for skewness, variety of shape characteristics and versatility of the shape characteristics of the pdf make the power Lindley distribution an interesting consideration in the context of animal movement.

### 2.3 Gumbel distribution

The Gumbel distribution is a special case of the generalised extreme value type 1 distribution (Kotz & Nadarajah, 2000) and is unimodal with a pdf (Chakraborty & Chakravarty, 2014)

$$f(d_t) = \frac{1}{\sigma} \exp\left(-\frac{d_t - \mu}{\sigma}\right) \exp\left(-\exp\left(-\frac{d_t - \mu}{\sigma}\right)\right), \quad -\infty < d_t < \infty,$$

where  $-\infty < \mu < \infty$  represents the location parameter and  $\sigma > 0$  represents the scale parameter. The Gumbel distribution is also known as the log-Weibull distribution. The corresponding cdf for the Gumbel distribution is given by (ibid.)

$$F(d_t) = \exp \left( -\exp \left( -\frac{d_t - \mu}{\sigma} \right) \right).$$

In the case where  $\mu = 0$  and  $\sigma = 1$ , this distribution is the standard extreme value distribution. The Gumbel distribution is typically used for modelling the maximum or minimum of a number of samples of various distributions. The shape of the Gumbel distribution pdf is unimodal and is typically skewed to the right. However, if  $\mu$  increases, the pdf will shift to the left. ibid. highlight the moments and moment generating function in further detail. The intrigue surrounding the Gumbel distribution is firstly its relation to the Weibull distribution and secondly, could be a useful consideration when distances tend to have more extreme values or consist of primarily of smaller distances.

### 3. Estimation Aspects And Inferential Procedures

The model framework is largely inspired by the work of Nicosia *et al.*, 2017 with the alternative direction and distance considerations incorporated into this modelling framework. The ssvM, the power Lindley and Gumbel distributions are being evaluated as alternative considerations for direction and distance respectively.

#### 3.1 The Markov Processes For The Hidden States

The hidden process  $S_{0:T}$  is a homogeneous Markov chain (i.e. transition probabilities independent of  $t$ ). Therefore, at any time step  $t$ , the animal is in a state ranging from  $\{1, \dots, K\}$  and  $g(S_t | \mathcal{F}_{t-1}^c) = g(S_t | S_{t-1})$  (ibid.). This allows the description of  $S_{0:T}$  as a discrete-time homogeneous multinomial process. This permits the definition of the sequence  $\{S_t, t = 0, \dots, T\}$  of multinomial vectors  $S_t = (S_{1t}, \dots, S_{Kt})$  where  $S_{it} = 1$  and  $S_{i't} = 0$  for all  $i' \neq i$  whenever  $S_t = i$  for  $i, i' = 1, \dots, K$ .

At the time  $t = 0$ , it is set that  $P(S_{k0} = 1) = (\pi_0)_k$  such that  $(\pi_0)_k \geq 0, k = 1, \dots, K$  and  $\sum_{k=1}^K (\pi_0)_k = 1$ . It should be noted that  $(\pi_0)_k$  presents the initial distribution of the hidden process. It's assumed that this initial distribution of the hidden process is known. The transition probabilities are denoted by  $\pi_{lk} = P(S_{kt} = 1 | S_{l,t-1} = 1)$  for  $l, k = 1, 2, \dots, K$ . The hidden process contributes to the complete data pdf given in (2) as a function of  $S_{0:T}$  and the transition probabilities in the following way (ibid.):

$$\prod_{t=1}^T g(S_t | \mathcal{F}_{t-1}^c) = \prod_{t=1}^T \prod_{l=1}^K \prod_{k=1}^K \pi_{lk}^{S_{l,t-1} S_{kt}}$$

The aforementioned methodology holds for any number of states  $K$ , but the case of  $K = 2$  is the only case considered in this modelling scenario.

#### 3.2 Inferential Procedures And EM Algorithm

In this section, the inferential procedures and EM algorithm are explained. A procedure to estimate  $\theta = (\pi, \kappa, \lambda)$ , where

- $\pi$  represents the transition probabilities in a  $K \times (K - 1)$  vector,
- $\kappa$  represents the  $K \times (p + 1)$  unknown parameters for the angular model and
- $\lambda$  represents the  $2K$  parameters of the distance variable model.

Given that we have the series of observations as before denoted by (1), the likelihood function of the parameters in the model are written as a product of the one-step ahead predictive densities,

$$L(\theta) = \prod_{t=1}^T \left\{ \sum_{k=1}^K f_k(y_t | \mathcal{F}_{t-1}^0, \kappa^{(k)}) h_k(d_t, \lambda^{(k)}) P(S_{tk} = 1 | \mathcal{F}_{t-1}^0, \theta) \right\} \tag{7}$$

where  $P(S_{tk} = 1 | \mathcal{F}_{t-1}^0, \theta)$  for  $t = 1, \dots, T$  represents the predictive probabilities which are recursively calculated using a filtering-smoothing algorithm which permits the evaluation of the likelihood function (7). The EM algorithm predicts the underlying state of the animal at each time point which leads to greater knowledge of their movements and behavioural patterns.

The function of the EM algorithm is to maximise the likelihood in the presence of missing or unobserved data. To evaluate the EM algorithm, the complete data log-likelihood function should be computed which follows from (2) and (7):

$$\begin{aligned} \log L_{\text{complete}}(\theta; \mathcal{F}_{t-1}^0) &= \sum_{t=1}^T \sum_{r=1}^K \sum_{k=1}^K S_{r,t-1} S_{kt} \log(\pi_{rk}(n_r, q_r)) \\ &+ \sum_{t=1}^T \sum_{k=1}^K S_{kt} \log(f_k(y_t | \mathcal{F}_{t-1}^0, \kappa^{(k)})) \\ &+ \sum_{t=1}^T \sum_{k=1}^K S_{kt} \log(h_k(d_t | \phi^{(k)})) \end{aligned} \tag{8}$$

Let  $\hat{\theta}_n$  denote the value of the estimate of  $\theta$  as previously defined after the  $n^{\text{th}}$  iteration of the EM algorithm. The  $(n + 1)^{\text{th}}$  iteration begins with one of the iteration of the E-step which will evaluate the expected value of the complete log-likelihood with respect to the conditional distribution of the missing data given the observed data which is expressed as follows:

$$\begin{aligned} Q(\theta | \hat{\theta}_n) &= E_{S_{0:T}} [\log L_{\text{complete}}(\theta; \mathcal{F}_T^0 | \mathcal{F}_T^0, \hat{\theta}_n)] \\ &= \sum_{t=1}^T \sum_{r=1}^K \sum_{k=1}^K E(S_{r,t-1} S_{kt} | \mathcal{F}_T^0, \hat{\theta}_n) \log(\pi_{rk}(n_r, q_r)) \\ &+ \sum_{t=1}^T \sum_{k=1}^K E(S_{kt} | \mathcal{F}_T^0, \hat{\theta}_n) \log(f_k(y_t | \mathcal{F}_{t-1}^0, \kappa^{(k)})) \\ &+ \sum_{t=1}^T \sum_{k=1}^K E(S_{kt} | \mathcal{F}_T^0, \hat{\theta}_n) \log(h_k(d_t | \lambda^{(k)})) \end{aligned} \tag{9}$$

Hereafter, the value of  $\hat{\theta}_{n+1}$  is calculated in the M-step as the value of  $\theta$  which will maximise  $Q(\theta | \hat{\theta}_n)$ .

**The E-Step**

1. Compute the filtering probabilities for  $E(S_{r,t-1} S_{kt} | \mathcal{F}_T^0, \hat{\theta}_n)$  and  $E(S_{kt} | \mathcal{F}_T^0, \hat{\theta}_n)$  from (9) using the filtering-smoothing algorithm as defined by (ibid.).

2. Compute the smoothing probabilities using Bayes theorem for both  $E(S_{r,t-1}S_{k,t}|\mathcal{F}_T^0, \hat{\theta}_n)$  and  $E(S_{kt}|\mathcal{F}_T^0, \hat{\theta}_n)$  using the filtering-smoothing algorithm.

**The M-step**

1. Maximise each individual element in (9) since they each depend on different parameter sets.
2. Maximise the hidden Markov process by:

$$\hat{\pi}_{rk} = \frac{\sum_{t=1}^T E(S_{r,t-1}S_{k,t}|\mathcal{F}_T^0, \hat{\theta}_n)}{\sum_{t=1}^T E(S_{r,t-1}|\mathcal{F}_T^0, \hat{\theta}_n)}$$

using the estimates obtained from the E-Step.

3. The M-step depends on  $h_k(d_t|\lambda^{(k)})$  which represents the distance component of the model. The estimates for this pdf is calculated using numerical optimisation such as the Newton-Raphson algorithm of the weighted log-likelihood.

To combat the issue of convergence to spurious or local maxima of the likelihood function (7) and maintain consistency in the algorithm, the EM algorithm is run using many random initial starting values for a few iterations and perform a check for spurious and local maxima. Parameter values are chosen based on which values yield the highest likelihood values for use as the starting values of the EM algorithm which is run until convergence (Biernacki *et al.*, 2003).

Two hidden hidden states are assumed based on previous literature and studies performed by various authors in this research field (Nicosia *et al.*, 2017). The selection of two states for the hidden process enables interpretability within the animal movement model environment. The directional targets and environmental features to be considered were selected using the AIC, BIC and Wald’s test criteria (with the results presented in Table 1). These criteria together with the log-likelihood values were used to determine which combination of alternative direction and distance models could be useful considerations when modelling animal movement.

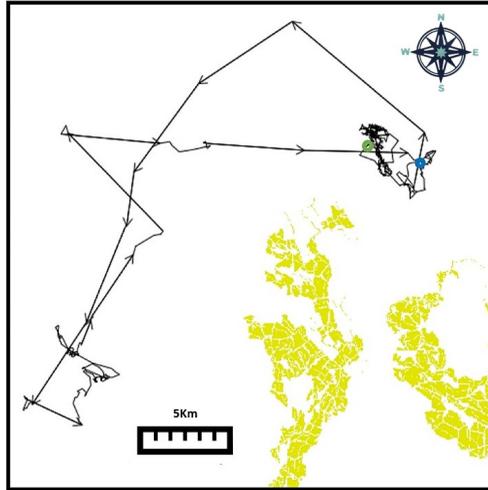
The various combinations of direction and distance models were determined using these criteria and the log-likelihood values for animal movement modelling.

**4. Data Application And Discussion**

In this section, the results from the aforementioned fitted model and their respective alternate parametric considerations are presented. This model was applied to data about the movement from forest caribou in Canada. This data involved 23 animals with various home ranges. Locations were measured every 4 hours during the 2006-2007 winter period with a total number of 617 observations. The data consists of four variables, namely:

- $\gamma_t$ : the direction of the animal at time point  $t$ .
- $x_{cut}$ : the direction to the closest cut (which is a forest stand which has been cut within a period ranging from 5 to 20 years ago).
- $x_{center}$ : the direction pointing towards the centroid of a cluster of recently visited locations.
- $d_t$ : the distance between the animal’s location from time step  $t$  to time step  $t + 1$  in kilometres.

Descriptive statistics revealed that majority of the values of  $d_t$  lie between 0 and 1, indicating that the animals tended to travel small distances within the four hour intervals. A map indicating the animal trajectory is presented in Figure 1 (ibid.).



**Figure 1.** The trajectories of the caribou in he observed time period where the blue circle represents the starting point and the green circle represents the end of the observed trajectory. The area highlighted in yellow represent the areas in the environment with regenerating cuts (ibid.).

The locations which the caribou visits between times 0 and time  $t - 1$  at time  $t$  are grouped into clusters. For example, cluster 1 refers to the set of locations visited by the animal between times 0 and  $t_1$ . The centroids (i.e. middle-points of the clusters) are calculated for these clusters and the cluster whose centroid is closest to the animal’s current position is used in order to compute the  $x_{center}$ .

### 4.1 Original Model Fit

A two-state model with  $K = 2$  which was proposed and is fitted to the caribou data. For this model, state 1 represents the "travelling" state where larger distances are travelled whereas state 2 represents the "encamped" state where minimal movement is noticed. The explanatory variables considered in this study are directional persistence ,  $x_{cut}$  and  $x_{center}$ . These explanatory variables are represented by the parameters  $\kappa_{persist}^{(k)}$ ,  $\kappa_{cut}^{(k)}$  and  $\kappa_{center}^{(k)}$  in the model respectively where  $k = 1, 2$  represents the current state of the model. In the model,  $n_k$  and  $q_k$  denote the size and probability of a negative binomial distribution of the waiting time of state  $k$ . Both  $q_1$  and  $q_2$  form the transition probability matrix for the hidden Markov chain denoted as:

$$P = \begin{bmatrix} 1 - q_1 & q_1 \\ q_2 & 1 - q_2 \end{bmatrix}$$

For each of the parameters in the model, the parameter estimates and corresponding standard errors (S.E) were computed. Both directional models, namely, the von Mises and ssvM distributions are presented within the same table together with the relevant distance model considerations. The first distance model being fitted is the Gamma distribution. No skewness was found to be present in the directional data so an arbitrarily small value of  $\phi = 0.1$  is selected for the skewness parameter of the ssvM. Note that when  $\phi = 0$ , the ssvM simplifies to the usual von Mises distribution.

In Table 1,  $\lambda_1^{(k)}$  and  $\lambda_2^{(k)}$  represent the shape and scale parameters of the Gamma and Weibull distributions, the shape and scale parameters for the for the power Lindley distribution and the

location and scale parameter for the Gumbel distribution respectively for state  $k$ . In Table 1, the "\*" indicates that these parameters are insignificant at a 10% level of significance.

**Table 1.** The estimation of the model with the von Mises and ssvM distributions for the directional component and the Gamma, Weibull, power Lindley and Gumbel distributions as the distance component

	Gamma Distribution				Weibull Distribution				power Lindley Distribution				Gumbel Distribution				
	von Mises		ssvM		von Mises		ssvM		von Mises		ssvM		von Mises		ssvM		
	Estimate	S.E	Estimate	S.E	Estimate	S.E	Estimate	S.E	Estimate	S.E	Estimate	S.E	Estimate	S.E	Estimate	S.E	
State 1	$\eta_1$	0.2788	0.0891	0.2788	0.0891	0.2716	0.0859	0.2716	0.0859	0.0734	0.0409	0.2942	0.1128	0.9484	0.1879	0.8344	0.2177
	$\mu_1$	1	-	1	-	1	-	1	-	1	-	1	-	1	-	1	-
	$\kappa_{\text{persist}}^{(1)}$	1.2666	0.3328	1.2666	0.3328	1.2225	0.3326	1.2225	0.3326	0.1872	0.0707	3.4148	1.5498	0.9161	0.4563	0.5617	0.3463
	$\kappa_{\text{cut}}^{(1)}$	0.1601*	0.3013	0.1601*	0.3013	0.1582*	0.2929	0.1582	0.2929	-0.0852	0.0883	-4.0876	1.8133	-0.8856	0.4668	-0.5757	0.3541
	$\kappa_{\text{center}}^{(1)}$	0.3733*	0.2994	0.3733*	0.2994	0.3298*	0.2895	0.3298	0.2895	0.1868	0.0743	1.014	0.8774	0.9493	0.4168	0.7412	0.3215
	$\lambda_1^{(1)}$	0.6478	0.1373	0.6478	0.1373	0.7179	0.0954	0.7179	0.0954	3.3596	0.1758	3.1673	0.3069	0.2099	0.0123	0.2134	0.0106
	$\lambda_2^{(1)}$	3.0445	0.8451	3.0445	0.8451	1.425	0.5103	1.425	0.5103	0.7012	0.0287	0.7712	0.0907	0.1407	-	0.1416	-
State 2	$\eta_2$	0.0231	0.0231	0.0231	0.0231	0.0258	0.0258	0.0258	0.0258	0.2709	0.2709	0.0114	0.0114	0.2824	0.2824	0.3723	0.3723
	$\mu_2$	1	-	1	-	1	-	1	-	1	-	1	-	1	-	1	-
	$\kappa_{\text{persist}}^{(2)}$	0.0274*	0.0618	0.0274*	0.0618	0.0183*	0.0644	0.0183	0.0644	-0.7817*	0.4073	0.0591*	0.0585	-0.0034*	0.1054	0.0021*	0.1446
	$\kappa_{\text{cut}}^{(2)}$	0.1454	0.0698	0.1454	0.0698	0.146	0.0712	0.146	0.0712	4.5847	1.5542	0.2478	0.0654	0.3707	0.1118	0.4198	0.1481
	$\kappa_{\text{center}}^{(2)}$	0.259	0.0694	0.259	0.0694	0.2612	0.0705	0.2612	0.0705	3.9761*	1.2426	0.3326*	0.0682	0.1774*	0.1066	0.1467*	0.1485
	$\lambda_1^{(2)}$	1.2626	0.0774	1.2626	0.0774	1.1143	0.0478	1.1143	0.0478	4.1603	0.3915	3.5031	0.1794	0.2109	0.0066	0.2095	0.0069
	$\lambda_2^{(2)}$	0.1351	0.012	0.1351	0.012	0.1757	0.0093	0.1757	0.0093	0.7602	0.0264	0.7075	0.0288	0.1388	-	0.1382	-
Likelihood		-794.91	-	-795.433	-	-798.301	-	-798.825	-	-1005.93	-	-1012.17	-	-523.568	-	-524.586	-
AIC		1613.82	-	1614.867	-	1620.603	-	1621.65	-	2035.856	-	2048.35	-	1071.135	-	1073.171	-
BIC		1666.918	-	1667.965	-	1673.701	-	1674.748	-	2088.954	-	2101.448	-	1124.234	-	1126.27	-
Model Run Time		76.01s	-	78.39s	-	92.69s	-	96.23s	-	65.65s	-	67.19s	-	72.53s	-	74.96s	-

Across all four models, minor increases in the AIC and BIC values are noted in the ssvM case in comparison to the usual von Mises case. The estimates also do not significantly differ across all models between the von Mises and ssvM distribution. Based on the results presented in Table 1, the Gumbel distribution performs the best when modelling animal movement distances based on the BIC values across both directional cases, however, estimated parameters do not differ significantly from zero which yields this models position in modeling as poor. The Gumbel distribution should be used cautiously as the pdf could shift to the left. It is interesting that although the power Lindley distribution is a mixture of the Gamma and Weibull distributions, it had the worst performance in terms of the BIC values. The performance of the power Lindley distribution will be further investigated in the bootstrap simulation study.

The power Lindley distribution had the fastest run time for the modelling procedure, followed by the Gumbel distribution, Gamma distribution and finally, the Weibull distribution from Table 1. The run times were computed using the `tic` package available in R (Izrailev, 2021). This illustrates that the alternative parametric distance considerations are more computationally efficient for distance in this BCRW model. An interesting distinction in the results is the significance of certain parameters in the model. Across all models,  $\kappa_{\text{persist}}^{(2)}$  was insignificant at a 10% level of significance. In both the Gamma and Weibull distributions,  $\kappa_{\text{cut}}^{(1)}$  and  $\kappa_{\text{center}}^{(1)}$  were both found to be insignificant as at a 10% level of significance whereas this was not the case for the power Lindley and Gumbel distributions. This difference could be due to the manner in which the model classifies distances into the two movement states. Therefore, the amount of information available in the particular movement state used for parameter estimation could differ between the models, influencing the p-value. These differences will be highlighted by the stationary distributions in Table 2.

**For the first state of the model,** large estimates were obtained across all models for the average distance travelled by the animal i.e.  $\hat{\lambda}_1^{(k)}\hat{\lambda}_2^{(k)}$  with the exception of the Gumbel distribution. An interesting highlight is that the power Lindley obtained similar distances to the Gamma distribution in state one, reinforcing the idea to further investigate its performance as a distance model consideration in this animal movement modelling context. The caribou exhibit positive directional

persistence in their movement illustrated by  $\kappa_{\text{persist}}^{(1)}$  across the models. For both the Weibull and Gamma models,  $\kappa_{\text{cut}}^{(1)}$  and  $\kappa_{\text{center}}^{(1)}$  had positive estimates with fairly large standard errors. In contrast, the power Lindley and Gumbel models noted a negative estimate for  $\kappa_{\text{cut}}^{(1)}$  and a positive estimate for  $\kappa_{\text{center}}^{(1)}$  with fairly large standard errors. These large standard errors could be due to the fact that there is limited amounts of data which is available for state 1. In order to obtain the average speed at which the animal moves in this state,  $\hat{\lambda}_1^{(1)}\hat{\lambda}_2^{(1)}/4$  is computed. The Gamma distribution reported a distance of approximately 494m per hour, with the power Lindley reporting 611m per hour, the Weibull reporting 255m per hour and lastly, the Gumbel reporting 7m per hour. The Gumbel distribution clearly doesn't provide a reasonable estimation in this scenario as previously highlighted by the average distances obtained.

**For the second state of the model**, the animal is in its "encamped" state in which it is primarily stationary. Smaller estimates were found for the average speeds with the Gamma reporting 44m per hour and the Weibull reporting 49m per hour. The power Lindley and Gumbel distributions reported 357m per hour and 7.4m per hour respectively which is a stark contrast to the Gamma and Weibull results. The directional persistence  $\kappa_{\text{persist}}^{(2)}$  was found to be insignificant in this state across all models which is a sensible result given the stationary movement state of the caribou. Both of the environmental factors considered in the model were found to be significant. It is interesting to note that  $\kappa_{\text{cut}}^{(2)}$  was significant and positive for this state across all models when the caribou never reaches this area as illustrated by Figure 1. This may be a fortuitous relationship given the north-south movement trajectory of the caribou and the southern location of the regenerating tree cuts.

The stationary distribution of the latent fitted Markov chain is in the first state with the corresponding probability  $\frac{q_2}{q_2+q_1}$ . This probability can also show the number of sightings the caribou was observed travelling. Using the von Mises directional results, the stationary distributions are summarised in Table 2 below:

**Table 2.** Stationary distribution of the latent fitted Markov chain for state 1

	Probability	No. of travelling States
Gamma	0.0765	47
power Lindley	0.1126	69
Weibull	0.0868	54
Gumbel	0.2294	142

The number of travelling states observed for the caribou expressed in the table above provides an explanation to the low precision and high standard errors found across all models in the first state. The Gumbel distribution found the largest amount of travelling states, yet, provided the smallest estimate for distance travelled. In contrast, the power Lindley distribution found a larger amount of travelling states than the Gamma and Weibull counterparts and provided a larger estimate for distance travelled in state 1.

Using the stationary distributions in Table 2 and parameter estimates obtained in Table 1, Figure 2 provides an illustration of how the various distributions fitted the distances for state 1. Figure 2 illustrates that the power Lindley and Gamma distributions fit the distances relatively well with the Weibull distribution providing a slightly worse fit, especially when modelling the longer distances. The worst is clearly the Gumbel distribution, where the shape of the pdf hardly fits the distances

being modelled, particularly in the beginning where the larger distances are observed.

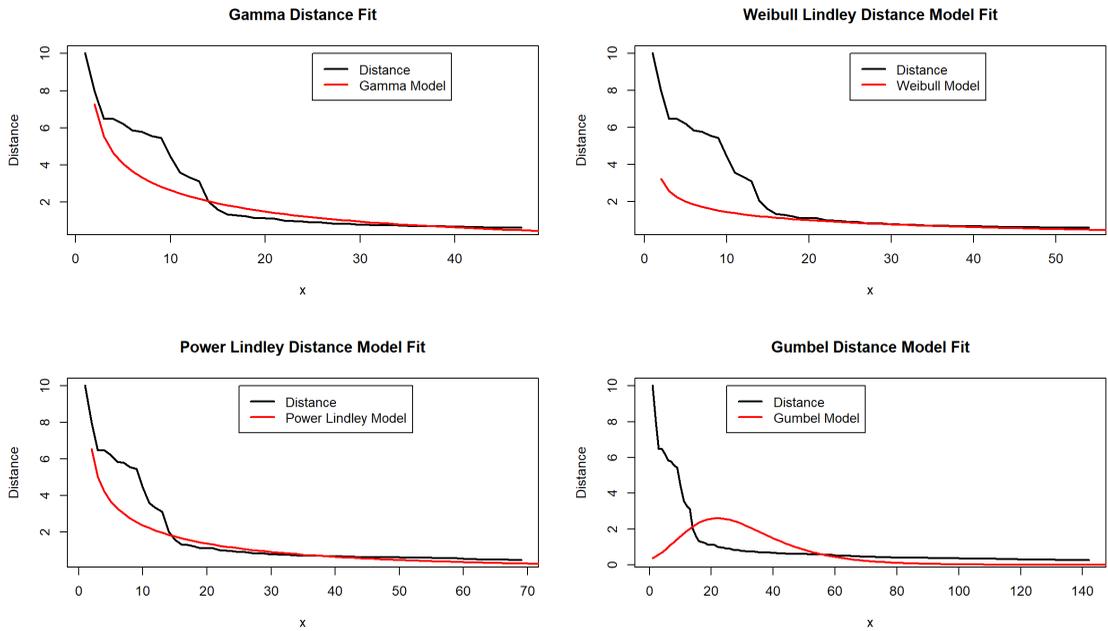


Figure 2. The various distance component fits for the model for state 1.

## 4.2 A deeper insight: bootstrap study

In the previous section, the power Lindley distribution was considered as an alternative for the distances in animal movement modelling. Upon closer inspection of the obtained estimates, the power Lindley distribution proved to be a unique consideration in the context of animal movement modelling in light of its parametric performance and interpretability. The motivation for further investigation into this model is the construction of the power Lindley distribution, which consists of a mixture between the Gamma and Weibull distributions as well as the appeal of increased computational efficiency as illustrated by the lower model run times in Table 1. In this section, a bootstrap procedure is performed using the von Mises distribution for the directional component to obtain the bootstrap distribution for the parameters of the power Lindley distribution from the fitted model. It should be noted that both the directional and distance components were jointly modelled, but only the distance component fits are illustrated. This bootstrap study aims to "mimic" real data for illustrative purposes which will be able to highlight the potential capabilities of the power Lindley distribution.

The algorithm for the bootstrap of the power Lindley distribution is outlined below:

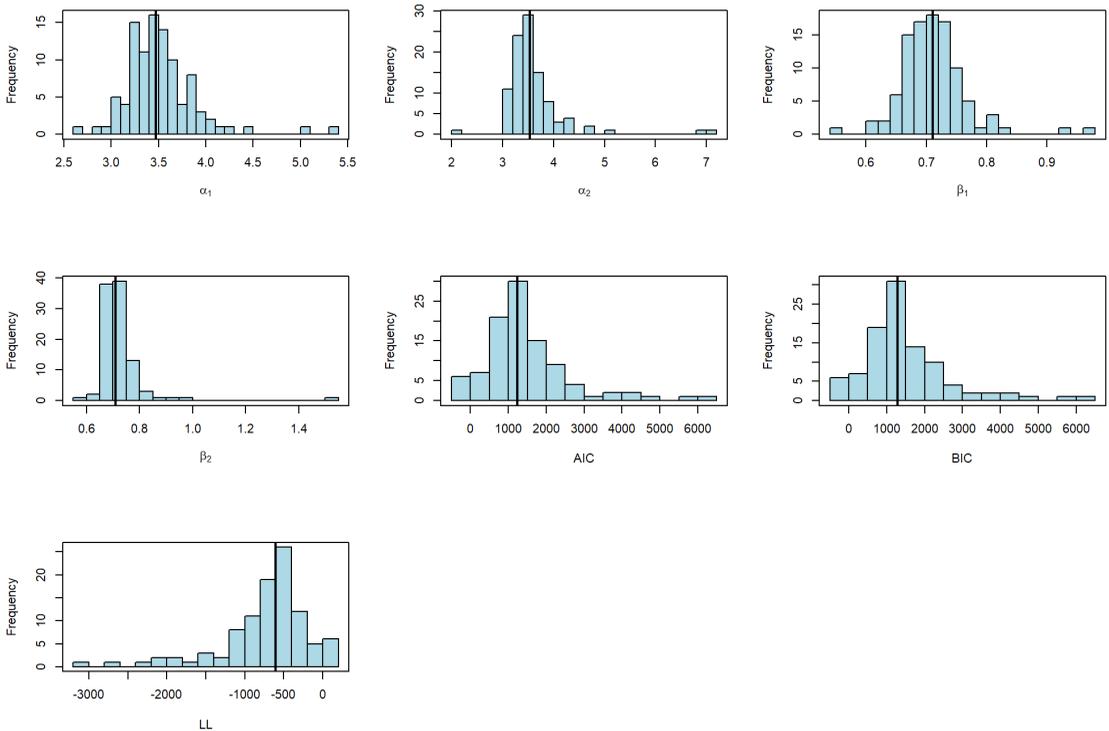
1. Generate  $m = 100,300$  samples of size  $n = 617$  i.e. the size of the dataset with replacement from the original dataset.
2. Calculate and store the estimated scale and shape parameters of the power Lindley obtained from fitting the BCRW model i.e  $\lambda_1^{(k)}$  and  $\lambda_2^{(k)}$  defined as in the previous section for states  $k = 1, 2$  as well as the AIC, BIC and log-likelihood values for each resample.
3. Calculate the bootstrap distribution and summary statistics for each of the parameters.

4. Plot the bootstrap distribution (histogram) with a vertical reference line for the median.

Across all the histograms, the solid black vertical reference line represent the median value of the sample statistic. The median is considered to be a more robust measure of location than the mean due to the skewed nature of some of the sampling distributions. The results for the summary statistics and bootstrap distributions are shown in Table 3 for  $m = 100$ :

**Table 3.** Summary statistics of the bootstrap distributions obtained for  $m = 100$

	Min	1st Quartile	Median	Mean	3rd Quartile	Max
$\lambda_1^{(1)}$	2.64	3.29	3.47	3.53	3.68	5.38
$\lambda_1^{(2)}$	2.01	3.32	3.54	3.62	3.70	7.16
$\lambda_2^{(1)}$	0.54	0.68	0.71	0.71	0.74	0.97
$\lambda_2^{(2)}$	0.58	0.68	0.71	0.73	0.74	1.53
AIC	-363.29	848.22	1231.16	1461.97	1833.92	6078.44
BIC	-310.20	901.32	1284.26	1515.07	1887.02	6131.54
LL	-3027.22	-904.96	-603.58	-718.98	-412.11	193.65



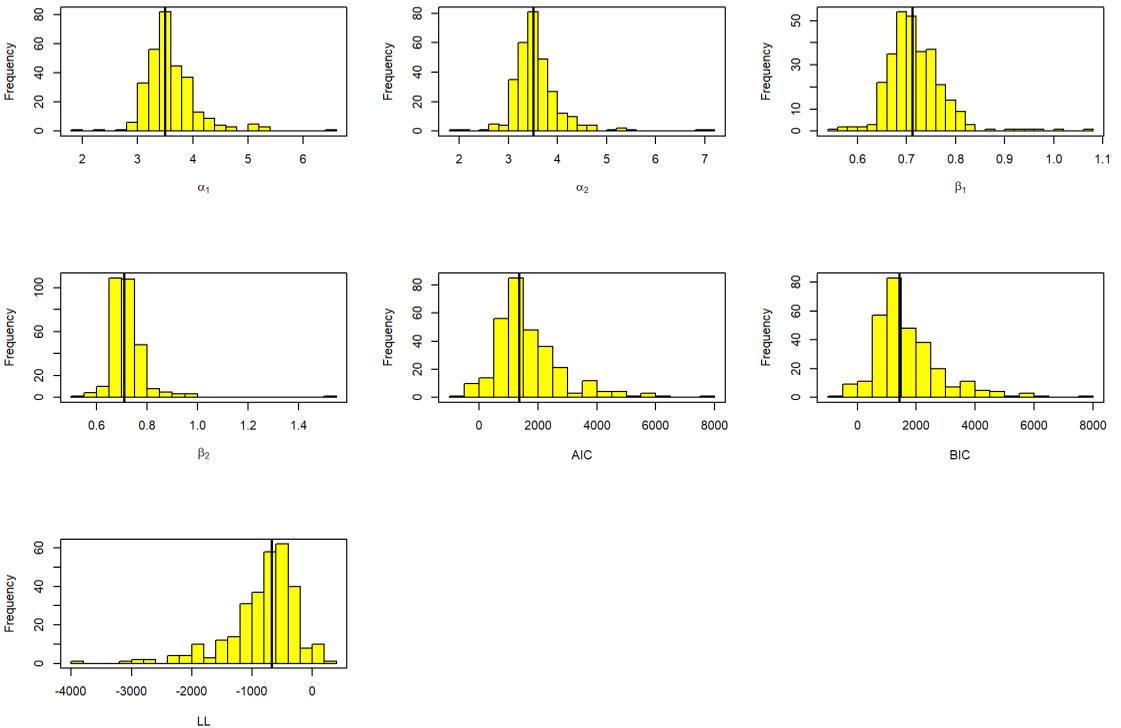
**Figure 3.** Bootstrap histograms for  $m = 100$  where the black vertical line represents the median of the sample statistic.

These values obtained when  $m = 100$  resemble the results found in the previous section, however, a large spread of values is once again noted for the scale and shape parameter estimates i.e.  $\lambda_1^{(k)}$  and  $\lambda_2^{(k)}$  for states  $k=1,2$ . It is interesting to note that the AIC and BIC values have decreased in

comparison to the results shown in Table 1. Furthermore, the bootstrap distributions are right-skewed with the exception of the likelihood bootstrap distribution. Large amounts of variability are exhibited by the bootstrap results as illustrated by the spread of values obtained in Table 3. The results for the summary statistics and bootstrap distributions are tabulated in Table 4 for  $m = 300$ .

**Table 4.** Summary statistics of the bootstrap distributions obtained for  $m = 300$

	Min	1st Quartile	Median	Mean	3rd Quartile	Max
$\lambda_1^{(1)}$	1.97	3.32	3.50	3.60	3.80	6.60
$\lambda_1^{(2)}$	1.99	3.29	3.51	3.57	3.73	7.16
$\lambda_2^{(1)}$	0.54	0.69	0.71	0.72	0.75	1.07
$\lambda_2^{(2)}$	0.52	0.68	0.71	0.72	0.75	1.53
AIC	-633.49	937.37	1371.99	1652.03	2090.86	7865.78
BIC	-580.39	990.47	1425.09	1705.13	2143.96	7918.88
LL	-3920.89	-1033.43	-674.00	-814.02	-456.68	328.74



**Figure 4.** Bootstrap histograms for  $m = 300$  where the black vertical line represents the median of the sample statistic.

The parameter values of  $\lambda_1^{(k)}$  and  $\lambda_2^{(k)}$  for  $k = 1, 2$  are similar to the values obtained by Table 1. As the sample size increased from  $m = 100$  to  $m = 300$ , the median AIC and BIC values increased but remain lower than the Gamma distribution and Weibull distribution results in Table 1. This indicates that for certain subsets of data, the power Lindley distributions performs on par with the other models and in some instances, outperforms these models whilst retaining interpretability and

computational efficiency. This is an important discovery which highlights the potential of the power Lindley distribution as a candidate for distance modelling. The performance of the power Lindley for modelling distance in animal movement illustrates better performance than the Weibull and Gamma distributions, whilst still remaining true to the movement nature of the caribou as previously highlighted in the discussion of the parameter estimates obtained.

## 5. Final thoughts

This paper proposes alternative parametric considerations for direction and distance in the context of animal movement modelling. The two alternative parametric considerations introduced in this paper are the power Lindley- and Gumbel distributions to model the animal's distances in the model. Furthermore, the ssvM distribution is a more flexible directional modelling consideration in comparison to the traditional von Mises distribution due to its ability to account for both multimodality and asymmetry without the loss of interpretability or logic. The gamma distribution outperformed the power Lindley and Weibull distributions when modelling the distances in the original fit model. However, the power Lindley distribution had a similar performance to the Gamma distribution based on parameter estimates and calculated measures obtained whilst remaining interpretable and more computationally efficient. This warranted further investigation via an insightful bootstrap study to further investigate the performance of the power Lindley distribution as a consideration for the modelling of distances within the animal movement. The bootstrap study revealed that the power Lindley distribution is a powerful consideration in distance modelling of animal movement. The power Lindley distribution may be suited to certain animal movement datasets (or subsets of this data), which emphasises its inclusion into the animal movement modelling context. No perfect model exists that will have the best performance across many datasets or studies. The performance of models and model selection is a data-driven process and will differ across studies. This study has highlighted certain distributions may outperform others within contexts and certain distributions are more well suited for this modelling context than others. Future work may include other potential candidates for the direction- and distance marginals (for example, Bekker *et al.*, 2022 and Ferreira & van der Merwe, 2022), as the study and implementation of this type of joint model with angular- and linear data remains theoretically challenging and computationally heavy.

## Acknowledgements

The support of the DSI-NRF Centre of Excellence in Mathematical and Statistical Sciences (CoE-MaSS), Johannesburg, South Africa towards this research is hereby acknowledged. Funding for this study was obtained from the NRF (grant RA201125576565, nr 145681) as well as grant RDP296/2021 based at the University of Pretoria. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to the CoE-MASS nor the NRF.

## Conflicts of Interest

The authors declare no conflict of interest.

## References

1. Abe, T. & Pewsey, A. Sine-skewed circular distributions. *Statistical Papers* 52, 683–707 (2011).
2. Bain, L. J. & Engelhardt, M. *Introduction to probability and mathematical statistics* (Brooks/Cole, 1987).
3. Bekker, A., Nakhaei Rad, N., Arashi, M. & Ley, C. in *Directional Statistics for Innovative Applications* 161–186 (Springer, 2022).

4. Biernacki, C., Celeux, G. & Govaert, G. Choosing starting values for the EM algorithm for getting the highest likelihood in multivariate Gaussian mixture models. *Computational Statistics and Data Analysis* **41**, 561–575 (Jan. 2003).
5. Chakraborty, S. & Chakravarty, D. A discrete Gumbel Distribution. *arXiv:1410.7568 [math.ST]* (Oct. 2014).
6. Ferreira, J. & van der Merwe, A. A Noncentral Lindley Construction Illustrated in an INAR(1) Environment. *Stats* **5**, 70–88 (2022).
7. Ghitany, M., Al-Mutairi, D., Balakrishnan, N. & Al-Enezi, L. Power Lindley distribution and associated inference. *Computational Statistics and Data Analysis* **64**, 20–33 (2013).
8. Goodall, V., Fatti, P. & Owen-Smith, N. Animal movement modelling: independent or dependent models? *South African Statistical Journal* **51**, 295–315 (Dec. 2017).
9. Izrailev, S. *tictoc: Functions for Timing R Scripts, as Well as Implementations of Stack and List Structures* R package version 1.0.1 (2021). <https://CRAN.R-project.org/package=tictoc>.
10. Jammalamadaka, S. R. & Sengupta, A. *Topics in circular statistics* ISBN: 9789810237783 (World Scientific, Jan. 2001).
11. Kotz, S. & Nadarajah, S. *Extreme value distributions: theory and applications* (World Scientific, 2000).
12. Langrock, R., King, R., Matthiopoulos, J., Thomas, L., Fortin, D. & Morales, J. M. Flexible and practical modeling of animal telemetry data: Hidden Markov models and extensions. *Ecology* **93**, 2336–2342. eprint: <https://esajournals.onlinelibrary.wiley.com/doi/pdf/10.1890/11-2241.1>. <https://esajournals.onlinelibrary.wiley.com/doi/abs/10.1890/11-2241.1> (2012).
13. Mardia, K. V. & Jupp, P. E. *Directional statistics* (eds Barnett, V., Creasie, N. A. C. & Fisher, N. I.) (John Wiley and Sons, 2009).
14. Nicosia, A., Duchesne, T., Rivest, L.-P. & Fortin, D. A general angular regression model for the analysis of data on animal movement in ecology. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* **65**, 445–463 (2016).
15. Nicosia, A., Duchesne, T., Rivest, L.-P. & Fortin, D. A general hidden state random walk model for animal movement. *Computational Statistics and Data Analysis* **105**, 76–95. ISSN: 0167-9473. <http://www.sciencedirect.com/science/article/pii/S0167947316301736> (2017).
16. Van Niekerk, B. *Application of Hidden Markov Models and their Extensions to Animal Movement Data* MA thesis (Nelson Mandela University, 2018).