Noname manuscript No.
(will be inserted by the editor)

Emergence of the wrapped Cauchy distribution in mixed directional data

Joseph D. Bailey · Edward A. Codling

Received: date / Accepted: date

Abstract Inferring the most appropriate distribution (or distributions) to describe observed directional data is important in many applications of circular statistics. In particular, animal movement paths are typically analysed and modelled by considering the distribution of step lengths and turning (or absolute) angles. Here we demonstrate that a single wrapped Cauchy distribution can appear to fit directional data mixed from two different underlying wrapped normal distributions. We derive mathematical expressions to calculate the parameter space for which this occurs and illustrate the result by analysing an example data set of the movements of African bull elephants (*Loxodonta Africana*). We conclude that the presence of a wrapped Cauchy distribution in observed directional data can, in certain cases, be explained by data coming from two distinct underlying distributions. We discuss how this may relate to the presence of multiple movement modes within an observed path when analysing animal movement data.

Keywords Circular distributions · wrapped normal · wrapped Cauchy · animal movement · directional data

Mathematics Subject Classification (2010) MSC 62P10 · MSC 92B05

J. D. Bailey
Department of Mathematical Sciences, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, U.K.
Tel.: +441206 874113
E-mail: jbailef@essex.ac.uk

E. A. Codling
Department of Mathematical Sciences, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, U.K.

1 Introduction

The analysis and applications of circular statistics to directional data plays a significant role in the study of many environmental processes from plant phenology (Morellato et al (2010)) and tree growth (Aradottir et al (1997)) to wind direction (Masseran et al (2013)) and the general movement patterns of animals and cells (Rivest et al (2016); Landler et al (2018)). Ascertaining the distribution which most closely describes circular data is important as characteristics of circular distributions, such as sharper peaks and slow decaying tails, have significant effects on the qualitative and quantitative results of descriptive and predictive models.

The most common distributions used to describe angular data are the wrapped normal (WN), von Mises (vM) (or circular normal) and the wrapped Cauchy (WC) (Jammalamadaka and SenGupta (2001); Mardia and Jupp (2000); McClintock et al (2012); McClintock and Michelot (2018)). These are defined by a probability density function (PDF) on the unit circle, and in the case of the WN and WC distributions, can be formed by 'wrapping' the equivalent one dimensional distributions on the real line around the unit circle (Stephens (1963); Jammalamadaka and SenGupta (2001); Mardia and Jupp (2000); Abe and Shimatani (2018)). Although the von Mises and wrapped normal distributions have differing PDFs, the two approximate each other very closely and produce similar qualitative results (Stephens (1963); Collett and Lewis (1981); Jammalamadaka and SenGupta (2001); Mardia and Jupp (2000); Codling et al (2010)). The WC distribution qualitatively differs from the WN and vM as it has a taller peak around the mean value and heavier tails which decay more slowly. Hence, many analyses classify angular data as being either having a sharp peaks with slow decaying tails (and therefore similar to wrapped Cauchy) or near normal (and thus either a von Mises or wrapped normal).

In particular, when modelling movement by random walk (RW) or step-turn processes it is often necessary to understand the distribution of turning angles and movement directions (Kareiva and Shigesada (1983); Bartumeus et al (2008); Codling et al (2008, 2010); Parton and Blackwell (2017)). Methods to determine the distribution which best describes observed directional data typically involves finding MLE parameters for the model distributions and choosing between them by the use of a likelihood or distance measure (Nilsen et al (2013); Li and Bolker (2017)).

Evidence that a WC distribution is the 'best-fit' for the distribution of turning angles or global orientations in a movement path has been found across a wide range of animals from insects and beetles, such as pea aphids *Acyrthosiphon pisum* (Nilsen et al (2013)) and the Baltimore checkerspot butterfly *Euphydryas phaeton* (Brown and Crone (2016)) to larger animals such as common brushtail possums *Trichosurus Vulpecula* (Postlethwaite and Dennis (2013)), cow elk, *Cervus elaphus*, (Morales et al (2004)), free-range cattle *Nothofagus Antarctica* (Seoane (2015)), Florida panthers, *Puma concolor coryi*, (van de Kerk et al (2015); Li and Bolker (2017)), California sea lions *Zalophus californianus* (Breed et al (2012)), Giant tortoises *Testudinidae* (Blake et al (2013)), american lobster *Homarus americanus* (Bowlby et al (2007)) and seals *Erignathus barbatus* and *Monachus schauinslandi* (McClintock et al (2015)). The studies mentioned above range from high frequency data having locations given every second to large scale movement with data sent every 24hr.

In comparison the WN or vM distribution has been reported across a similarly wide range of animals, from E. coli bacterium (Taylor-King et al, 2015), Fender's blue butterfly *Icaricia icarioides* fender (Schultz and Crone (2001)) and bog fritillary butterfly *Proclossiana Eunomia* (Schtickzelle et al (2007)) to larger animals such as red-cockaded woodpecker *Picoides borealis* (McKellar et al (2014)), lesser black-backed gull *L. fuscus* (Taylor-King et al (2015)), king penguin *Aptenodytes patagonicus* (Pistorius et al (2017)), reindeer *Rangifer tarandus* (Langrock et al (2014)) and southern elephant seal *Mirounga leonine* (Michelot et al (2017)).

A basic search on Google Scholar reveals that whilst both types of distribution have been frequently used and reported in recent animal movement data analyses, there has been a marked increase in the prevalence of the WC distribution over the last 10-15 years. A key word search of "animal movement" and "wrapped Cauchy" returns only 22 articles published before 2008, compared to the equivalent for either "von Mises" or "wrapped normal" returning 81, a four-fold difference. However, since 2008 this ratio has halved with 200 articles mentioning WC and 398 for WN demonstrating a marked increase in the prevalence of the WC distribution.

One biological interpretation of the presence of a WC distribution is that the individual mainly travels on a near constant bearing, with the majority of turns occurring within small deviations from 0 whereas medium to large turns happen only occasionally but with a similar frequency. This would indicate the animal has a tendency for sudden large changes in direction of movement, rather than a gradual change in orientation over a course of a series of larger turns which would be expected from a normal or Gaussian distribution. Various RW movement models have shown that observably different qualitative and quantitative results are produced depending upon whether a WC or vM distribution has been used (Bartumeus et al (2008); Codling et al (2010)), demonstrating the importance of accurately determining the underlying distributions.

If the way in which data is collected, analysed or processed can affect how well a candidate distribution fits the observed data then this needs to be well understood and acknowledged. For example it has been shown that errors in GPS data locations can give rise to spurious large 180° turns, which would enhance the slow decaying tailed nature of recorded turning angles (Jerde and Visscher (2005); Hurford (2009)). Other reported issues with data collection which artificially increased the number of large turns include the effect of recording data in a restricted area, where edge effects can cause sudden large turns as the animal encounters a wall. Young et al (2013) found that flour beetles, *Tribolium confusum* took smaller steps with larger turn angles closer to the border of the experimental setup, which resulted in a distribution with slow decaying tails. Similar results relating the experimental setup to artificially enhancing the presence of distributions with slower decaying tails (such as a WC) has been recorded in other species such as parasitic wasps, *Encarsia formosa* (Drost et al (2000)).

When considering movement behaviour it is known that animals can exhibit different movement modes when travelling (Schtickzelle et al (2007); Gurarie et al (2016); Cagnacci et al (2016); Nicosia et al (2017)), perhaps due to switching from a foraging/exploration phase to an encamped/feeding phase which can lead to periods of small turns followed by periods of larger turns (McClintock et al (2015); Torres et al (2017); Nicosia et al (2017)). Similarly, changes in the terrain or climate could alter the movement behaviour (Patterson et al (2009); Dahmen et al (2017); Pérez-Barbería et al (2015)). The qualitative behaviour of each movement phase will be best described by a specific model and set of parameters and if these strategies are not known a priori the movement data could be analysed under the assumption of a single movement strategy resulting in the mixing of the data from the individual behavioural states. The simplest example of such multiple movement behaviour is a two state movement model where one phase is described by large variability in turning angles between steps relating to highly tortuous movement perhaps indicative of foraging or encamped behaviour, and another phase with more directed, straighter movement where the deviation in turning angles from the mean is smaller, akin to purposeful goal based movement or flight behaviour (Patterson et al (2010); Langrock et al (2012); Jonsen et al (2013); Nams (2014); Parton and Blackwell (2017); Nicosia et al (2017); McClintock and Michelot (2018)).

Here we demonstrate that a single wrapped Cauchy distribution can appear to fit directional data mixed from two different underlying wrapped normal distributions. We derive analytical expressions used to calculate the parameter space for which this occurs. Our results show that, in general, when the two WN distributions forming the mixed distribution have a large difference in their respective concentration parameters (≥ 0.5) a WC is the best fitting single distribution, indicating that a mixed distribution can enhance the appearance of a slow decaying tails in the distribution of turning angles when interpreted as a single distribution.

2 Background: Circular Statistics and Distributions

A symmetric wrapped stable (SWS) distribution has the density function given by:

$$f_{\mathrm{sws}}(\theta;\rho,\mu)=\frac{1}{2\pi}\left(1+2\sum_{n=1}^{\infty}\rho^{n^{a}}\cos n(\theta-\mu)\right),\ n\in\mathbb{N}, \tag{1}$$

where $\rho \in [0,1)$ is the concentration parameter, $\mu \in [-\pi,\pi)$ is the location parameter around which the distribution is symmetric and $a \in (0,2]$, with $\theta \in [-\pi,\pi)$.

In the specific case for $a = 1$ the SWS distribution returns the WC distribution and for $a = 2$ we get the WN distribution (Jammalamadaka and SenGupta (2001)).

It is well known that for any given WN distribution a vM can be found as an accurate approximation (Stephens (1963); Collett and Lewis (1981); Jammalamadaka and SenGupta (2001)). Hence, both give qualitatively similar results when used in random walk (RW) models (Codling et al (2010)). Therefore, we consider only a WN distribution as it allows for easier algebraic manipulation.

If we let the SWS be centred around 0 or $\pm\pi$, ($\mu = 0, \pm\pi$), then $\rho^{n^a} = \alpha_n$, where α_n is the nth cosine moment of f_{sws}. Note that in this case we need only consider the cosine moments as the sine moments are 0 (Mardia and Jupp (2000); Jammalamadaka and SenGupta (2001)).

2.1 Mixed wrapped distributions

One can also consider distributions formed by mixing two circular distributions, f_{md}, where random variables are drawn from one of the two distributions according to a certain probability, ω. Defined as,

$$f_{\text{md}}(\theta; \omega) = \omega f_1(\theta) + (1 - \omega) f_2(\theta),$$

where f_1 and f_2 are SWS distributions and $\omega \in [0, 1]$ is defined as the mixing parameter of the two initial distributions. Note that the trivial cases for $\omega = 0, 1$ are equivalent to $f_{\text{md}} = f_2$ and $f_{\text{md}} = f_1$ respectively.

Lemma 1 *Let $f_{md}(\theta; \omega)$ be a distribution formed by mixing two SWS distributions centred around zero ($\mu = 0$), then f_{md} itself is a wrapped distribution with cosine moments, $\alpha_n^{\{md\}}$, given by*

$$\alpha_n^{\{md\}} = \omega \alpha_n^{\{1\}} + (1 - \omega) \alpha_n^{\{2\}},$$

where $\alpha_n^{\{1\}}$, $\alpha_n^{\{2\}}$ are the nth cosine trigonometric moments of f_1 and f_2 respectively.

Proof Lemma 1 follows directly from the definition of f_{md} and f_{sws}.

As both f_1 and f_2 are SWS distributions with $\mu = 0$, we have

$$\begin{aligned} f_{\text{md}}(\theta) =& \omega \frac{1}{2\pi} \left(1 + 2 \sum_{n=1}^{\infty} \alpha_n^{\{1\}} \cos(n\theta) \right) + (1 - \omega) \frac{1}{2\pi} \left(1 + 2 \sum_{n=1}^{\infty} \alpha_n^{\{2\}} \cos(n\theta) \right) \\ =& \frac{1}{2\pi} \left(1 + 2 \sum_{n=1}^{\infty} \left[\omega \alpha_n^{\{1\}} + (1 - \omega) \alpha_n^{\{2\}} \right] \cos(n\theta) \right). \end{aligned}$$

This is the form of a general SWS given in (1) centred around 0 with trigonometric moments $\omega \alpha_n^{\{1\}} + (1 - \omega) \alpha_n^{\{2\}}$ as required. Note, as f_{md} is a wrapped distribution, symmetric around $\mu = 0$ the trigonometric sine moments are all 0 and hence the trigonometric moments of f_{md} are purely the cosine moments.

2.2 Determining between best-fit circular distributions

When calculating a measure of the distance between two given PDFs, one can consider many statistical measures (Gibbs and Su (2002)). However one of the simplest is to consider the sum of the squares of the differences between the distributions across their domain; equivalent to finding the L^2-distance. Hence, we will consider the distance, $d(\cdot,\cdot)$ between two continuous probability functions, f and g, over the finite domain X as

$$d(f,g) = \int_{x\in X} \left[f(x) - g(x)\right]^2 \, \mathrm{d}x \, .$$

We use the L^2 distance throughout this study as it allows for simple algebraic manipulation, however Online Resource 1 (Figs. S1-S4) detail the results of using other metrics, found using simulations and show they are qualitatively similar to those found using the L^2 distance (Figure 1). Therefore, when comparing between multiple PDFs we infer the closest fitting distribution as the one which minimises, d, the L^2 distance.

3 Fitting individual wrapped distributions to a mixed distribution

3.1 Statement of main claim

Here we consider a mixed distribution formed from two WN distributions and demonstrate the parameter space for which it is best described by either a single WN or single WC distribution. By considering the L^2 distance we derive expressions for calculating the best-fitting WN and WC distributions as functions of the parameters of the mixed distribution. Subsequently we determine the parameter space for whether a WN or WC best describes the mixed distribution, by selecting the distribution with the smallest L^2 distance.

Proposition 1 *Let f_{md} be a mixed wrapped distribution formed by mixing two wrapped normal distributions, defined as*

$$f_{md}(\theta;\mu_0,\rho_1,\rho_2,\omega) = \omega f_{wn}(\theta;\mu_0,\rho_1) + (1-\omega) f_{wn}(\theta;\mu_0,\rho_2), \tag{2}$$

for $\theta \in [-\pi,\pi)$, with $\mu_0 \in [-\pi,\pi)$, $\rho_1,\rho_2 \in [0,1)$ and $\omega \in [0,1]$. Define

$$\begin{aligned}\Delta_{wn}(\rho,\mu) =& d\left(f_{wn}(\theta;\rho,\mu), f_{md}(\theta;\mu_0,\rho_1,\rho_2,\omega)\right),\\ \Delta_{wc}(\rho,\mu) =& d\left(f_{wc}(\theta;\rho,\mu), f_{md}(\theta;\mu_0,\rho_1,\rho_2,\omega)\right).\end{aligned}$$

Let ρ_{wn}, μ_{wn} minimise Δ_{wn} and ρ_{wc}, μ_{wc} minimise Δ_{wc}. Then there always exists ρ_1,ρ_2,ω:

$$\Delta_{wc}(\rho_{wc},\mu_{wc}) < \Delta_{wn}(\rho_{wn},\mu_{wn}).$$

As Δ_{wn} and Δ_{wc} give the values of the L^2 distance for each distribution compared to f_{md}, the smaller value of Δ_{wn} and Δ_{wc} indicates the closer fitting distribution.

Note, as both distributions forming the mixed distribution, f_{md}, are from the same family of distributions (2) without loss of generality we can consider the distribution with concentration parameter ρ_1 to be the distribution which has the smaller probability of being chosen and, therefore, by symmetry we need only consider $\omega \in [0, 1/2]$.

We only consider the distributions within the mixed distribution f_{md} to be WN rather than WC as the latter leads to the mixed distribution always being classified as a single WC and our main concern is determining when a slowly decaying tailed distribution fits data from distributions with normal type tails (see Online Resource 2 for a complete discussion of this along with the results of having a WC and a WN as the initial mixed distributions). More generalised cases, such as when one distribution is centred at $\pm\pi$ and having a mixed distribution formed of multiple underlying distributions, are commented upon in the Discussion (sect. 6) as well as in Online Resources 4-5.

To demonstrate this proposition we give an analytical method for calculating the specific parameter values which minimise Δ_{wn} and Δ_{wc} for fixed ρ_1, ρ_2, ω. By directly comparing these minimised values we show the parameter space for which the WC distribution (or WN) is the closest fitting distribution when the L^2 distance metric is used.

3.2 Demonstration of main claim

First we note that if we assume both underlying WN distributions are centred around the same value, here $\mu = 0$, then clearly f_{md} is centred around the same value also, $\mu_{\mathrm{md}} = 0$, and we must have $\mu_{\mathrm{wn}} = \mu_{\mathrm{wc}} = \mu_{\mathrm{md}} = 0$ (Mardia and Sutton (1975)).

Using Lemma 1 we can write f_{md} from Prop. 1 as

$$f_{\mathrm{md}} = \frac{1}{2\pi}\left(1 + 2\sum_{n=1}^{\infty} \alpha_n^{\{\mathrm{md}\}} \cos(n\theta)\right),$$

with

$$\alpha_n^{\{\mathrm{md}\}} = \omega\alpha_n^{\{1\}} + (1-\omega)\alpha_n^{\{2\}},$$

where $\alpha_n^{\{1\}}$ and $\alpha_n^{\{2\}}$ are the nth cosine moments of the WN distributions with concentration parameters ρ_1 and ρ_2 respectively.

Recalling that $\alpha_n^{\mathrm{wn}} = \rho_{\mathrm{wn}}^{n^2}$, we have

$$\alpha_n^{\{\mathrm{md}\}} = \omega\rho_1^{n^2} + (1-\omega)\rho_2^{n^2}.$$

We now show that when considering the L^2 distance between two zero centred SWS distributions it suffices to calculate the sum of the squares of the differences between their respective cosine moments.

Lemma 2 *Let $f_1(\theta)$ and $f_2(\theta)$ be SWS distributions centred around 0 with cosine moments $\alpha_n^{\{1\}}$ and $\alpha_n^{\{2\}}$ respectively, then*

$$d(f_1, f_2) = \frac{1}{\pi} \sum_{n=1}^{\infty} \left(\alpha_n^{\{1\}} - \alpha_n^{\{2\}} \right)^2 .$$

Proof As $f_1(\theta)$ and $f_2(\theta)$ are zero-centred SWS distributions, the square of the difference between the distributions at any given value of $\theta \in [-\pi, \pi)$ is given by

$$\begin{aligned} [f_1(\theta) - f_2(\theta)]^2 &= \left[\frac{1}{2\pi} \left(1 + 2 \sum_{n=1}^{\infty} \alpha_n^{\{1\}} \cos(n\theta) \right) - \frac{1}{2\pi} \left(1 + 2 \sum_{n=1}^{\infty} \alpha_n^{\{2\}} \cos(n\theta) \right) \right]^2 \\ &= \left[\frac{1}{\pi} \sum_{n=1}^{\infty} \left(\alpha_n^{\{1\}} - \alpha_n^{\{2\}} \right) \cos(n\theta) \right]^2 , \end{aligned}$$

integrating over $[-\pi, \pi)$ with respect to θ gives

$$\int_{-\pi}^{\pi} [f_1(\theta) - f_2(\theta)]^2 \, \mathrm{d}\theta = \int_{-\pi}^{\pi} \left[\frac{1}{\pi} \sum_{n=1}^{\infty} \left(\alpha_n^{\{1\}} - \alpha_n^{\{2\}} \right) \cos(n\theta) \right]^2 \, \mathrm{d}\theta .$$

Expanding the right hand side gives

$$\begin{aligned} = \frac{1}{\pi^2} \int_{-\pi}^{\pi} \Bigg[& \sum_{n=1}^{\infty} \left(\alpha_n^{\{1\}} - \alpha_n^{\{2\}} \right)^2 \cos^2(n\theta) \\ & + 2 \sum_{i=1}^{\infty} \sum_{j=i+1}^{\infty} \left(\alpha_i^{\{1\}} - \alpha_i^{\{2\}} \right) \left(\alpha_j^{\{1\}} - \alpha_j^{\{2\}} \right) \cos(i\theta) \cos(j\theta) \Bigg] \mathrm{d}\theta , \end{aligned}$$

$$\begin{aligned} = \frac{1}{\pi^2} \Bigg[& \sum_{n=1}^{\infty} \left(\alpha_n^{\{1\}} - \alpha_n^{\{2\}} \right)^2 \int_{-\pi}^{\pi} \cos^2(n\theta) \, \mathrm{d}\theta \\ & + 2 \sum_{i=1}^{\infty} \sum_{j=i+1}^{\infty} \left(\alpha_i^{\{1\}} - \alpha_i^{\{2\}} \right) \left(\alpha_j^{\{1\}} - \alpha_j^{\{2\}} \right) \int_{-\pi}^{\pi} \cos(i\theta) \cos(j\theta) \, \mathrm{d}\theta \Bigg] , \end{aligned}$$

noting that the integral in the first term yields π and the integral in the second gives 0, this expression reduces to

$$\frac{1}{\pi}\sum_{n=1}^{\infty}\left(\alpha_n^{\{1\}}-\alpha_n^{\{2\}}\right)^2.$$

For a more complete derivation including the intermediary steps see Online Resource 3

Therefore, we can re-write Δ_{wc} and Δ_{wn} as

$$\Delta_{\mathrm{wc}}(\rho,0)=\frac{1}{\pi}\sum_{n=1}^{\infty}\left(\alpha_n^{\{\mathrm{wc}\}}-\alpha_n^{\{\mathrm{md}\}}\right)^2=\frac{1}{\pi}\sum_{n=1}^{\infty}\left[\rho^n-\left(\omega\rho_1^{n^2}+(1-\omega)\rho_2^{n^2}\right)\right]^2, \tag{3}$$

$$\Delta_{\mathrm{wn}}(\rho,0)=\frac{1}{\pi}\sum_{n=1}^{\infty}\left(\alpha_n^{\{\mathrm{wn}\}}-\alpha_n^{\{\mathrm{md}\}}\right)^2=\frac{1}{\pi}\sum_{n=1}^{\infty}\left[\rho^{n^2}-\left(\omega\rho_1^{n^2}+(1-\omega)\rho_2^{n^2}\right)\right]^2. \tag{4}$$

As ρ_{wc} and ρ_{wn} are the values which minimise Δ_{wc} and Δ_{wn} respectively, they can be found by differentiating (3) & (4) with respect to ρ and equating for 0.

Hence, we require

$$0=\frac{\mathrm{d}}{\mathrm{d}\rho}\Delta_{\mathrm{wc}}=\frac{1}{\pi}\sum_{n=1}^{\infty}2n\rho^{n-1}\left[\rho^n-(\omega\rho_1^{n^2}+(1-\omega)\rho_2^{n^2})\right], \tag{5}$$

$$0=\frac{\mathrm{d}}{\mathrm{d}\rho}\Delta_{\mathrm{wn}}=\frac{1}{\pi}\sum_{n=1}^{\infty}2n^2\rho^{n^2-1}\left[\rho^{n^2}-(\omega\rho_1^{n^2}+(1-\omega)\rho_2^{n^2})\right]. \tag{6}$$

The precise values of ρ which satisfy (5) & (6) will therefore be the values for ρ_{wc} and ρ_{wn} respectively, and can be found via numerical methods. Substituting ρ_{wc} and ρ_{wn} back into the expressions for Δ_{wc} and Δ_{wn} in (3) & (4) yield the respective minimum values. We can now determine the parameter space of ρ_1, ρ_2, ω for which a WC distribution is favoured over a WN distribution when compared to f_{md} by considering when $\Delta_{\mathrm{wc}}(\rho_{\mathrm{wc}}, \mu_{\mathrm{wc}}) < \Delta_{\mathrm{wn}}(\rho_{\mathrm{wn}}, \mu_{\mathrm{wn}})$. Therefore, calculating

$$D_{\Delta}=\Delta_{\mathrm{wc}}(\rho_{\mathrm{wc}},0)-\Delta_{wn}(\rho_{\mathrm{wn}},0),$$

will give us an indication, not only of which distribution is favoured (negative in the case of a WC and positive for a WN), but also the relative 'strength'; that is the larger the absolute value, the larger the difference in the L^2 distance between the distributions, and thus the closer the preferred distribution is to the mixed distribution.

%%Figure 1 here%%

4 Results

Fig. 1 shows the results of plotting D_Δ across the parameter space of ρ_1, ρ_2, ω with the areas in yellow (areas bounded by the dashed line) representing combinations for which the WN distribution is considered closer to the mixed distribution f_{md} and blue areas showing where a WC is considered closer; the darker the colour the stronger the preference.

In the simplest case where the mixed distribution is formed by mixing the two underlying distributions equally ($\omega = 0.5$; Fig 1K) the plot is symmetric about the lead diagonal (corresponding to $\rho_1 = \rho_2$) as expected, with the WN favoured whenever $|\rho_1 - \rho_2|$ is small. The areas for which the WC is favoured occur predominantly when the difference between ρ_1, ρ_2 is large ($|\rho_1 - \rho_2| \geq 0.5$). However, in the case when both concentration parameters are greater than 0.5, the area for which a WC is favoured is much smaller occurring now only if $5\rho_1 - \rho_2 \geq 4$ (or $5\rho_2 - \rho_1 \geq 4$).

In general the plots remain almost unchanged across $0.35 \leq \omega \leq 0.5$ (Figs 1H-J). In particular, the areas of the plots above the lead diagonal (corresponding to $\rho_1 > \rho_2$) remain remarkably unchanged for $0.1 \leq \omega \leq 0.5$. However, as $\omega \to 0$ the area favouring the WC (blue) begins to vanish (demonstrated in Fig 1B with $\omega = 0.05$) and disappears entirely when $\omega = 0$ (Fig 1A) due to the the definition of the mixed distribution, f_{md} (2) ($\omega = 0$ is equivalent to $f_{\mathrm{md}} = f_2$ and since f_2 was chosen to be WN it will never be best classified as WC).

Considering now the areas of the plots below the lead diagonal (corresponding to $\rho_2 > \rho_1$), as ω decreases below 0.3 (Fig. 1A-G) the area favouring the WC shrinks and only exists for large values of ρ_2 (> 0.8). And for $\omega \leq 0.15$, corresponding to distributions where the majority of angles are drawn from the distribution with parameter ρ_2, the plots indicate that there is no combination of parameters for which the WC will be favoured when $\rho_2 > \rho_1$ (Figs 1A-C).

5 Example: analysis of elephant movement

As an example of data which is well-fitted by a single WC and after a simple analysis appears to be better fitted by a mixed distribution, we use tracking data from bull African elephants Loxodonta africana previously published in Wall et al (2014b)). Here, location data were recorded for two elephants, we consider the data for the elephant id: *Habiba* which had locations recorded every 15 minutes for a period of over 4 days giving 1522 data points (data from Movebank data repository; Wall et al (2014a)). Turning angles were found by calculating the difference in bearings for subsequent pairs of locations. Visual inspection of the movement path (Fig. 2A) appears to show segments of high tortuosity where the movement path includes large variations in turning angles, along with periods of more straight-line behaviour with mainly small deviations in direction and fewer larger turns. Simply pooling the turning angles across the entire path gives the distribution shown in Fig. 2B. Using the standard practice of best fitting a WN and WC distribution using the packages in R (in this case *CircStats*) reveals that a WN (black dotted, $\rho_{\mathrm{wn}} = 0.3558$) is a poor fit, whereas a WC (black dashed, $\rho_{\mathrm{wc}} = 0.4385$) is a close fit (Fig 2C). If instead, we assume turns are drawn from two distinct distributions, we can consider the observed data to be a mixed distribution (as in (2)). Further assuming that the underlying distributions are WN, the best fitting mixed distribution can be found by simply comparing the density distribution of the observed data (calculated using the circular package in R (R Core Team (2018)) with all possible mixed distributions formed with parameters ρ_1, ρ_2, ω at 0.01 intervals, selecting the specific combination of parameters which minimises the L^2 distance. In this case we find that the best fitting mixed distribution is one with $\rho_1 = 0.20, \rho_2 = 0.91, \omega = 0.71$ (Fig 2D). In calculating the continuous density curve for the discrete observed data, the bandwidth used was the automatic selection from R as would be the case for a simple initial analysis, however, fixing this width at other values did not change the qualitative results.

When comparing this mixed distribution with the best fitting WC we see that both are close matches (Fig 2C & 2D), however, as the visual inspection of the movement path indicated more than one movement behaviour, then one could conclude the mixed distribution is the better for describing the movement as it implies that the turning angles across the elephant's path came from two distinct distributions. With the value of ω indicating that 29% of angles were drawn from a highly peaked distribution and 71% from a flatter, more uniform distribution.

The possible presence of a mixed distribution could indicate two distinct movement behaviours over the path, with one behaviour admitting turning angles drawn from a distribution tightly peaked around 0 and the other behaviour with angles taken from a flatter distribution. However, it should be noted that one cannot use this analysis as a method of predicting such multiple state behaviour, as it provides no information of the movement state any given part of the path is likely to be in, neither does it provide a 'switching' parameter which determines the likelihood of switching between states; as is expected in behavioural state analyses although ω acts as a proxy for this (Johnson et al (2008); Patterson et al (2009); Parton and Blackwell (2017); McClintock and Michelot (2018)). Similarly, it does not consider any other covariates or parameters of the movement path typically used in CRW movement models, such as step-length or bout distribution, nor any correlation between these parameters (i.e. having smaller step lengths when the variation in turning angle is large and vice-versa). All of which can be considered by the use of cylindrical distributions (Abe and Ley (2017); Imoto et al (2019)).

A method of analysis which does consider switching parameters and other covariates to predict behavioural states is the *momentuHMM* package in R which was introduced in McClintock and Michelot (2018) and used to analyse the companion elephant dataset from Wall et al (2014a). The results found by applying this analysis to these data gives the best-fitting mixed distribution formed from WN distributions to be one with concentration parameters $\rho_1 = 0.11, \rho_2 = 0.80$ and a mixing parameter of $\omega = 0.56$ (The package requires using von Mises rather than WN distributions, however, as has been discussed these distributions are known to be similar).

Whilst the results found considering a mixed distribution and those found using the *momentuHMM* package are qualitatively similar, they are not equivalent since the HMM method of McClintock and Michelot (2018) specifically attempts to identify periods of distinct behaviour taking into account various aspects of the movement path, whereas, our results simply looked for the distribution which best described the distribution of turning angles. The observation that the outcomes are similar indicates that this analysis on the distribution of turning angles can give credible results for the underlying distributions of turning angles and identifying the presence of multiple movement behaviours. Also, in the specific case where HMM techniques wish to be used to analyse movement behaviour, using the mixed analysis approach to provide an initial parameter selection for the concentration and switching parameters could be beneficial due to HMMs sensitivity to the set of initial conditions.

%%Figure 2 here%%

6 Discussion

Accurately identifying parameters of movement models is clearly crucial when analysing, predicting and understanding animal behaviour. Identifying the most accurate distribution in turning angles is important as differing distributions can result in noticeably different predictive outcomes (Bartumeus et al (2008); Codling et al (2010)). In movement data analyses it is often assumed that angles are drawn from a single underlying distribution, here we have demonstrated the parameter space for when a mixed distribution can be best described by a single distribution with either a normal type, WN, or a sharply peaked and slowly decaying tailed distribution, WC; two distributions commonly associated with the analysis of directional movement.

Our results indicate that a mixed distribution formed from two WN distributions will, in general, be best fitted by a WC distribution when the difference between the concentration parameters of the underlying initial WN distributions is large ($|\rho_1 - \rho_2| \geq 0.5$). This has been reported when analysing and classifying animal movement behaviour into two movement states, such as "foraging" and "exploratory" (Langrock et al (2014); Nicosia et al (2017); McClintock and Michelot (2018)) The characteristic distributions found in such movement include a flat almost uniform distribution attributed to the "foraging" stage, and would be equivalent to a low concentration parameter in a SWS distribution, along with another much more peaked distribution for the "exploratory" phase, given by a distribution with a concentration parameter close to 1. Evidence of such results after model fitting have been observed in a range of animals including American lobster, *Homarus americanus*, (Bowlby et al (2007)), African elephants (McClintock and Michelot

(2018)), Cataglyphis desert ants (Dahmen et al (2017)), caribou (Nicosia et al (2017)) and elk, *C. elaphus*, (Parton and Blackwell (2017)). Specifically, Langrock et al (2014), found that reindeer in a 2 state model exhibited angular distributions described by a von Mises distribution with $\kappa = 0.246$ (approximately equivalent to a WN with $\rho = 0.1218$) for the "foraging" behaviour and $\kappa = 3.517$ (approximately equivalent to a WN with $\rho = 0.8389$) for the "exploratory" behaviour.

Here we assumed that the mixed distribution was formed from underlying distributions both centred at 0, however, it has been observed that the underlying distributions may be centred at 0 and $\pm\pi$. This has been found in the movement of female wolves (Mastrantonio et al (2019)), muskox (Pohle et al (2017)), American bison (Langrock et al (2012)) and elk (Patterson et al (2016)). It can be shown that, similar to the the results found here, both the WN and WC distribution are the single favoured distribution in certain cases (see Online Resource 4). However, in general, detection of such mixed behaviour is more readily observable as the resulting mixed distribution would appear bimodal, with a peak (usually higher) at 0 and another (generally lower) at $\pm\pi$. This bimodal behaviour would be the indicator that a mixed distribution, rather than one single distribution, should be considered. In contrast, when both distributions are centred at the same value, determining the presence of multiple underlying distributions is not usually clear. Similarly, the case for when one of the underlying distributions is centred at a point other than 0 or $\pm\pi$, would be expected to demonstrate bimodal behaviour. This case is not considered here and as in terms of animal movement it would imply an animal had a preference for consistent turning regardless of the direction faced, giving a helical movement path. However, this is not uncommon, and has been observed in movement data including; sunfish (Drucker and Lauder (2001)), thrushes (Da Silveira et al (2016)), elk (Fryxell et al (2008)) and wolves (Mastrantonio et al (2019)).

That this relatively straight forward approach of analysing movement data revealed results consistent with those using more complex methods is interesting as it relies solely on the angular data. However, as it gives no information as to which distribution any particular part of the movement path belongs, it cannot be used as an indicator of states of behaviour. Discovering when a period of movement comes from a particular state with prescribed model and parameter set is an active area of research, and as such there has been much work on behavioural change point analysis (BCPA) utilising a range of methods from hidden Markov models (HMMs) (Michelot et al (2016); Jonsen (2016); McClintock and Michelot (2018)), Markov chain Monte Carlo processes (McClintock et al (2012); Parton and Blackwell (2017)) and switching Markov models (Nicosia et al (2017)), to wavelet analysis (Polansky et al (2010)) and time series CUSUM techniques (Knell and Codling (2012)) (see Gurarie et al (2016), for a more complete list). Currently the analysis described here may be used to predict the values of the initial distributions required in HMM techniques (Michelot et al (2016); Jonsen (2016); McClintock and Michelot (2018)), but could also be extended to predict breaks in behaviour by including additional 'smoothing' techniques in order to ascertain when a change from using one distribution to another has occurred, most likely utilising a time-series break point analysis such as that used in Knell and Codling (2012). For this to be the case, many improvements would be needed in the method for finding the ρ_1, ρ_2, ω parameters, for example a more efficient search algorithm, (for example, the Nelder-Mead algorithm (Nelder and Mead (1965))) could be used rather than the slow parameter sweep used in section 5.

These results indicate that when techniques such as HMMs (Michelot et al (2016); McClintock and Michelot (2018)), switching Markov models (Nicosia et al (2017)) or state-space models (Patterson et al (2010)) are used to ascertain the number of movement states from telemetry data, it would be necessary to include more states when the assumed distributions for angular data are of normal-type (WN or vM) compared to using WC distributions. In such a case the resulting best-fitting model will be easier to interpret, with each single distribution more likely to describe a single movement behaviour. For example Pohle et al (2017) determined that, when considering von Mises distributions, 5 states gave the most parsimonious model in the analysis of muskox movement. Had a wrapped Cauchy been used instead of a von Mises, this work indicates that the expected number of states would have been lower, as the WC distribution can better capture distributions with slower decaying tails or bimodal behaviour.

Although only 2 distributions were considered in forming the mixed distribution here, the results generalise when including multiple underlying distributions (Online Resource 5). There are many other potential avenues for enhancing and extending the work shown here, for example, this method could be extended to mixing more than two normal distributions by simply including more ρ_i and mixing ratio terms in the summation for the mixed distribution and editing the subsequent calculations appropriately. However, interpreting the results obviously increases in difficulty due to the increasing dimension of the required parameter space. It should be noted that Jammalamadaka and Kozubowski (2017) have shown that a WC distribution can in fact be recovered precisely when one considers mixing an infinite number of WN distributions and therefore taking the mixture distribution as a continuous function across all possible concentration parameter values in $[0, 1]$ for the initial WN distributions.

Whilst we chose to focus on two particular distributions, other wrapped distributions such as wrapped Gompertz and the wrapped exponential have also been used to describe animal movement (Roy and Adnan (2012); Ravindran and Ghosh (2011)) and could be included in a more complete analysis. Similarly, other distributions on the unit circle such as the Jones-Pewsey (Jones and Pewsey (2005)), Kato-Jones (Kato et al (2013)) and wrapped t (Pewsey et al (2007)) could have been considered, as could cylindrical distributions (see Abe and Shimatani (2018) for a discussion of common cylindrical distributions). However, these are all multi-parameter distributions and as such can prove computationally harder to fit to actual data. Since our main aim here is to illustrate a possible mechanism for how distributions, such as the wrapped

Cauchy, may emerge in observed directional data, a full and and complete classification of mixed circular distributions along with their combinations, is beyond the scope of this work. Instead, this work highlights the importance the choice between the most common distributions used in directional data, can have to practitioners.

References

Abe T, Ley C (2017) A tractable, parsimonious and flexible model for cylindrical data, with applications. Econometrics and statistics 4:91–104

Abe T, Shimatani IK (2018) Cylindrical distributions and their applications to biological data. Applied Directional Statistics: Modern Methods and Case Studies p 86

Aradottir AL, Robertson A, Moore E (1997) Circular statistical analysis of birch colonization and the directional growth response of birch and black cottonwood in south iceland. Agricultural and Forest Meteorology 84(1-2):179–186

Bartumeus F, Catalan J, Viswanathan G, Raposo E, Da Luz M (2008) The influence of turning angles on the success of non-oriented animal searches. Journal of Theoretical Biology 252(1):43–55

Blake S, Yackulic CB, Cabrera F, Tapia W, Gibbs JP, Kümmeth F, Wikelski M (2013) Vegetation dynamics drive segregation by body size in galapagos tortoises migrating across altitudinal gradients. Journal of Animal Ecology 82(2):310–321

Bowlby HD, Hanson JM, Hutchings JA (2007) Resident and dispersal behavior among individuals within a population of american lobster homarus americanus. Marine Ecology Progress Series 331:207–218

Breed GA, Costa DP, Jonsen ID, Robinson PW, Mills-Flemming J (2012) State-space methods for more completely capturing behavioral dynamics from animal tracks. Ecological Modelling 235:49–58

Brown LM, Crone EE (2016) Individual variation changes dispersal distance and area requirements of a checkerspot butterfly. Ecology 97(1):106–115

Cagnacci F, Focardi S, Ghisla A, Van Moorter B, Merrill EH, Gurarie E, Heurich M, Mysterud A, Linnell J, Panzacchi M, et al (2016) How many routes lead to migration? comparison of methods to assess and characterize migratory movements. Journal of Animal Ecology 85(1):54–68

Codling EA, Plank MJ, Benhamou S (2008) Random walk models in biology. Journal of the Royal society interface 5(25):813–834

Codling EA, Bearon RN, Thorn GJ (2010) Diffusion about the mean drift location in a biased random walk. Ecology 91(10):3106–3113

Collett D, Lewis T (1981) Discriminating between the von mises and wrapped normal distributions. Australian Journal of Statistics 23(1):73–79

Da Silveira NS, Niebuhr BBS, de Lara Muylaert R, Ribeiro MC, Pizo MA (2016) Effects of land cover on the movement of frugivorous birds in a heterogeneous landscape. PloS one 11(6)

Dahmen H, Wahl VL, Pfeffer SE, Mallot HA, Wittlinger M (2017) Naturalistic path integration of cataglyphis desert ants on an air-cushioned lightweight spherical treadmill. Journal of Experimental Biology 220(4):634–644

Drost Y, Qiu Y, Posthuma-Doodeman C, Van Lenteren J (2000) Comparison of searching strategies of five parasitoid species of bemisia argentifolii bellows and perring (hom., aleyrodidae). Journal of Applied Entomology 124(2):105–112

Drucker E, Lauder G (2001) Wake dynamics and fluid forces of turning maneuvers in sunfish. Journal of Experimental Biology 204(3):431–442

Fryxell JM, Hazell M, Börger L, Dalziel BD, Haydon DT, Morales JM, McIntosh T, Rosatte RC (2008) Multiple movement modes by large herbivores at multiple spatiotemporal scales. Proceedings of the National academy of Sciences 105(49):19,114–19,119

Gibbs AL, Su FE (2002) On choosing and bounding probability metrics. International statistical review 70(3):419–435

Gurarie E, Bracis C, Delgado M, Meckley TD, Kojola I, Wagner CM (2016) What is the animal doing? tools for exploring behavioural structure in animal movements. Journal of Animal Ecology 85(1):69–84

Hurford A (2009) Gps measurement error gives rise to spurious 180 turning angles and strong directional biases in animal movement data. PLoS One 4(5):e5632

Imoto T, Shimizu K, Abe T (2019) A cylindrical distribution with heavy-tailed linear part. Japanese Journal of Statistics and Data Science 2(1):129–154

Jammalamadaka SR, Kozubowski TJ (2017) A general approach for obtaining wrapped circular distributions via mixtures. Sankhya A 79(1):133–157

Jammalamadaka SR, SenGupta A (2001) Topics in circular statistics, vol 5. world scientific

Jerde CL, Visscher DR (2005) Gps measurement error influences on movement model parameterization. Ecological Applications 15(3):806–810

Johnson DS, London JM, Lea MA, Durban JW (2008) Continuous-time correlated random walk model for animal telemetry data. Ecology 89(5):1208–1215

Jones M, Pewsey A (2005) A family of symmetric distributions on the circle. Journal of the American Statistical Association 100(472):1422–1428

Jonsen I (2016) Joint estimation over multiple individuals improves behavioural state inference from animal movement data. Scientific reports 6:20,625

Jonsen I, Basson M, Bestley S, Bravington M, Patterson T, Pedersen MW, Thomson R, Thygesen UH, Wotherspoon S (2013) State-space models for bio-loggers: A methodological road map. Deep Sea Research Part II: Topical Studies in Oceanography 88:34–46

Kareiva P, Shigesada N (1983) Analyzing insect movement as a correlated random walk. Oecologia 56(2-3):234–238

Kato S, Jones M, et al (2013) An extended family of circular distributions related to wrapped cauchy distributions via brownian motion. Bernoulli 19(1):154–171 van de Kerk M, Onorato DP, Criffield MA, Bolker BM, Augustine BC, McKinley SA, Oli MK (2015) Hidden semi-markov models reveal multiphasic movement of the endangered florida panther. Journal of Animal Ecology 84(2):576–585

Knell AS, Codling EA (2012) Classifying area-restricted search (ars) using a partial sum approach. Theoretical Ecology 5(3):325–339

Landler L, Ruxton GD, Malkemper EP (2018) Circular data in biology: advice for effectively implementing statistical procedures. Behavioral ecology and sociobiology 72(8):128

Langrock R, King R, Matthiopoulos J, Thomas L, Fortin D, Morales JM (2012) Flexible and practical modeling of animal telemetry data: hidden markov models and extensions. Ecology 93(11):2336–2342

Langrock R, Hopcraft JGC, Blackwell PG, Goodall V, King R, Niu M, Patterson TA, Pedersen MW, Skarin A, Schick RS (2014) Modelling group dynamic animal movement. Methods in Ecology and Evolution 5(2):190–199

Li M, Bolker BM (2017) Incorporating periodic variability in hidden markov models for animal movement. Movement ecology 5(1):1

Mardia K, Sutton T (1975) On the modes of a mixture of two von mises distributions. Biometrika pp 699–701

Mardia KV, Jupp PE (2000) Directional statistics, vol 494. John Wiley & Sons

Masseran N, Razali AM, Ibrahim K, Latif MT (2013) Fitting a mixture of von mises distributions in order to model data on wind direction in peninsular malaysia. Energy Conversion and Management 72:94–102

Mastrantonio G, Grazian C, Mancinelli S, Bibbona E, et al (2019) New formulation of the logistic-gaussian process to analyze trajectory tracking data. The Annals of Applied Statistics 13(4):2483–2508

McClintock BT, Michelot T (2018) momentuhmm: R package for generalized hidden markov models of animal movement. Methods in Ecology and Evolution 9(6):1518–1530

McClintock BT, King R, Thomas L, Matthiopoulos J, McConnell BJ, Morales JM (2012) A general discrete-time modeling framework for animal movement using multistate random walks. Ecological Monographs 82(3):335–349

McClintock BT, London JM, Cameron MF, Boveng PL (2015) Modelling animal movement using the argos satellite telemetry location error ellipse. Methods in Ecology and Evolution 6(3):266–277

McKellar AE, Langrock R, Walters JR, Kesler DC (2014) Using mixed hidden markov models to examine behavioral states in a cooperatively breeding bird. Behavioral Ecology 26(1):148–157

Michelot T, Langrock R, Patterson TA (2016) movehmm: an r package for the statistical modelling of animal movement data using hidden markov models. Methods in Ecology and Evolution 7(11):1308–1315

Michelot T, Langrock R, Bestley S, Jonsen ID, Photopoulou T, Patterson TA (2017) Estimation and simulation of foraging trips in land-based marine predators. Ecology 98(7):1932–1944

Morales JM, Haydon DT, Frair J, Holsinger KE, Fryxell JM (2004) Extracting more out of relocation data: building movement models as mixtures of random walks. Ecology 85(9):2436–2445

Morellato LPC, Alberti L, Hudson IL (2010) Applications of Circular Statistics in Plant Phenology: a Case Studies Approach, Springer Netherlands, pp 339–359

Nams VO (2014) Combining animal movements and behavioural data to detect behavioural states. Ecology Letters 17(10):1228–1237

Nelder JA, Mead R (1965) A Simplex Method for Function Minimization. The Computer Journal 7(4):308–313, DOI 10.1093/comjnl/7.4.308, URL https://doi.org/10.1093/comjnl/7.4.308, http://oup.prod.sis.lan/comjnl/article-pdf/7/4/308/1013182/7-4-308.pdf

Nicosia A, Duchesne T, Rivest LP, Fortin D (2017) A general hidden state random walk model for animal movement. Computational Statistics & Data Analysis 105:76–95

Nilsen C, Paige J, Warner O, Mayhew B, Sutley R, Lam M, Bernoff AJ, Topaz CM (2013) Social aggregation in pea aphids: Experiment and random walk modeling. PLOS ONE 8(12):1–11

Parton A, Blackwell PG (2017) Bayesian inference for multistate 'step and turn' animal movement in continuous time. Journal of Agricultural, Biological and Environmental Statistics 22(3):373–392

Patterson TA, Basson M, Bravington MV, Gunn JS (2009) Classifying movement behaviour in relation to environmental conditions using hidden markov models. Journal of Animal Ecology 78(6):1113–1123

Patterson TA, McConnell BJ, Fedak MA, Bravington MV, Hindell MA (2010) Using gps data to evaluate the accuracy of state–space methods for correction of argos satellite telemetry error. Ecology 91(1):273–285

Patterson TA, Parton A, Langrock R, Blackwell PG, Thomas L, King R (2016) Statistical modelling of animal movement: a myopic review and a discussion of good practice. arXiv preprint arXiv:160307511

Pérez-Barbería FJ, Small M, Hooper RJ, Aldezabal A, Soriguer-Escofet R, Bakken GS, Gordon IJ (2015) State-space modelling of the drivers of movement behaviour in sympatric species. PLOS ONE 10(11):1–21

Pewsey A, Lewis T, Jones MC (2007) The wrapped t family of circular distributions. Australian & New Zealand Journal of Statistics 49(1):79–91

Pistorius P, Hindell M, Crawford R, Makhado A, Dyer B, Reisinger R (2017) At-sea distribution and habitat use in king penguins at sub-antarctic marion island. Ecology and Evolution 7(11):3894–3903

Pohle J, Langrock R, van Beest FM, Schmidt NM (2017) Selecting the number of states in hidden markov models: pragmatic solutions illustrated using animal movement. Journal of Agricultural, Biological and Environmental Statistics 22(3):270–293

Polansky L, Wittemyer G, Cross PC, Tambling CJ, Getz WM (2010) From moonlight to movement and synchronized randomness: Fourier and wavelet analyses of animal location time series data. Ecology 91(5):1506–1518

Postlethwaite CM, Dennis TE (2013) Effects of temporal resolution on an inferential model of animal movement. PLOS ONE 8(5):1–11

R Core Team (2018) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, URL http://www.R-project.org/

Ravindran P, Ghosh SK (2011) Bayesian analysis of circular data using wrapped distributions. Journal of Statistical Theory and Practice 5(4):547–561

Rivest LP, Duchesne T, Nicosia A, Fortin D (2016) 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(3):445–463

Roy S, Adnan MAS (2012) Wrapped generalized gompertz distribution: an application to ornithology. Journal of Biometrics & Biostatistics 3(6)

Schtickzelle N, Joiris A, Dyck HV, Baguette M (2007) Quantitative analysis of changes in movement behaviour within and outside habitat in a specialist butterfly. BMC Evolutionary Biology 7(4)

Schultz CB, Crone EE (2001) Edge-mediated dispersal behavior in a prairie butterfly. Ecology 82(7):1879–1892

Seoane N (2015) Modelling free-range cattle movements in forests using multistate random walks. Journal of Biological Systems 23(supp01):S43–S54

Stephens MA (1963) Random walk on a circle. Biometrika 50(3/4):385–390

Taylor-King JP, van Loon EE, Rosser G, Chapman SJ (2015) From birds to bacteria: Generalised velocity jump processes with resting states. Bulletin of Mathematical Biology 77(7):1213–1236

Torres LG, Orben RA, Tolkova I, Thompson DR (2017) Classification of animal movement behavior through residence in space and time. PLOS ONE 12(1):1–18

Wall J, Wittemyer G, LeMay V, Douglas-Hamilton I, Klinkenberg B (2014a) Data: Elliptical time-density model to estimate wildlife utilization distributions. DOI doi:10.5441/001/1.f321pf80/1, URL http://dx.doi.org/10.5441/001/1.f321pf80/1

Wall J, Wittemyer G, LeMay V, Douglas-Hamilton I, Klinkenberg B (2014b) Elliptical time-density model to estimate wildlife utilization distributions. Methods in Ecology and Evolution 5(8):780–790

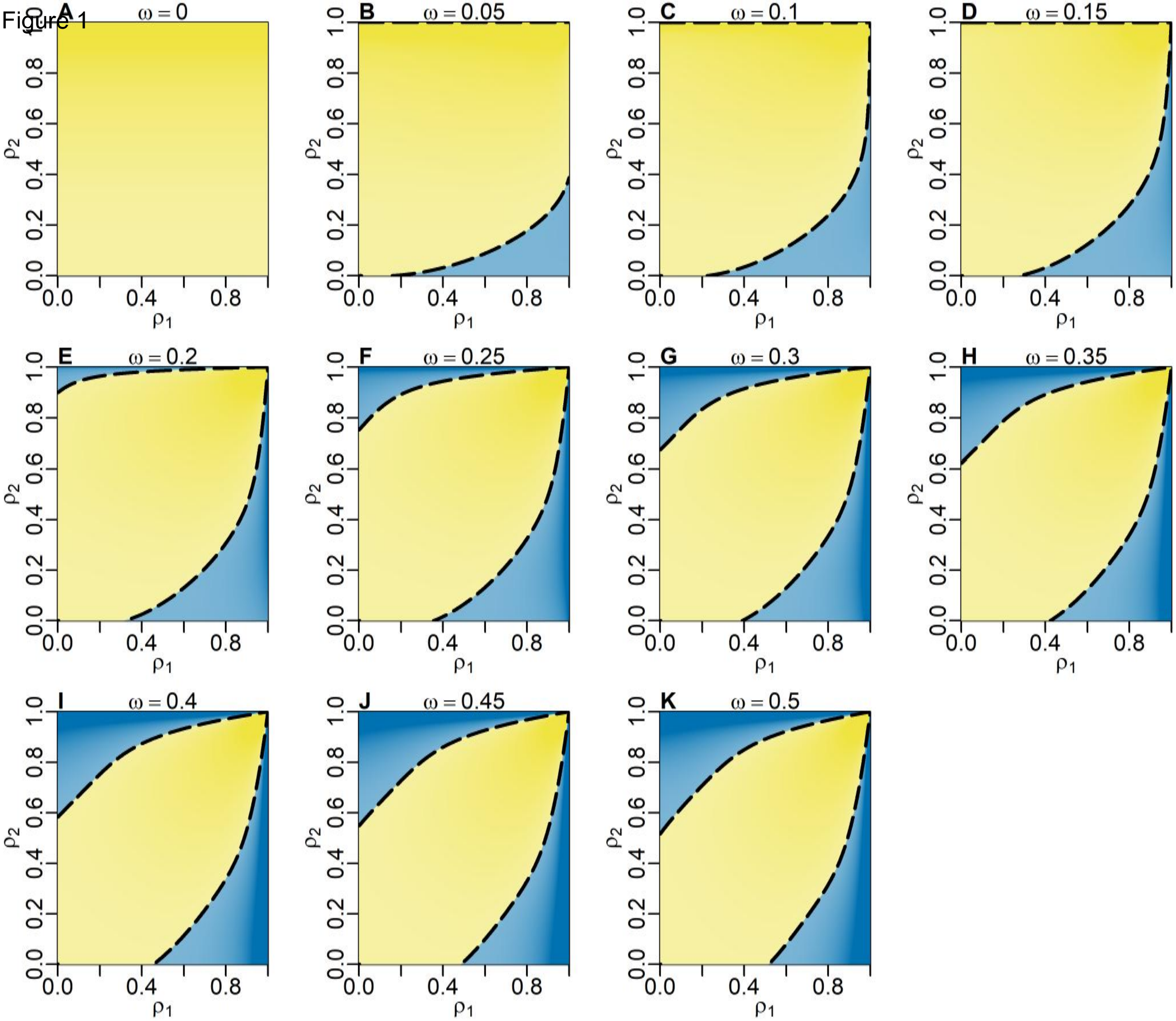
Young HC, Reid TG, Randall LA, Lachowsky LE, Foster DJ, Pengelly CJ, Latty T, , Reid ML (2013) Influences of movement behavior on animal distributions at edges of homogeneous patches. International Journal of Zoology 2013

Figure Legends

Figure 1: Plots demonstrating the parameter space for where a single WC or WN is the favoured distribution for $\rho_1, \rho_2 \in [0,1]$ calculated at 0.001 intervals with the mixing ratio $\omega \in [0,0.5]$ at 0.05 intervals (due to symmetry the results for $\omega > 0.5$ are not displayed). Areas in blue (dark grey – printed version) indicate parameter combinations for which the WC was favoured whereas areas in yellow (light grey – printed version) show combinations for which the WN was favoured. The dashed black line indicates where the transition from preferred distribution occurs.

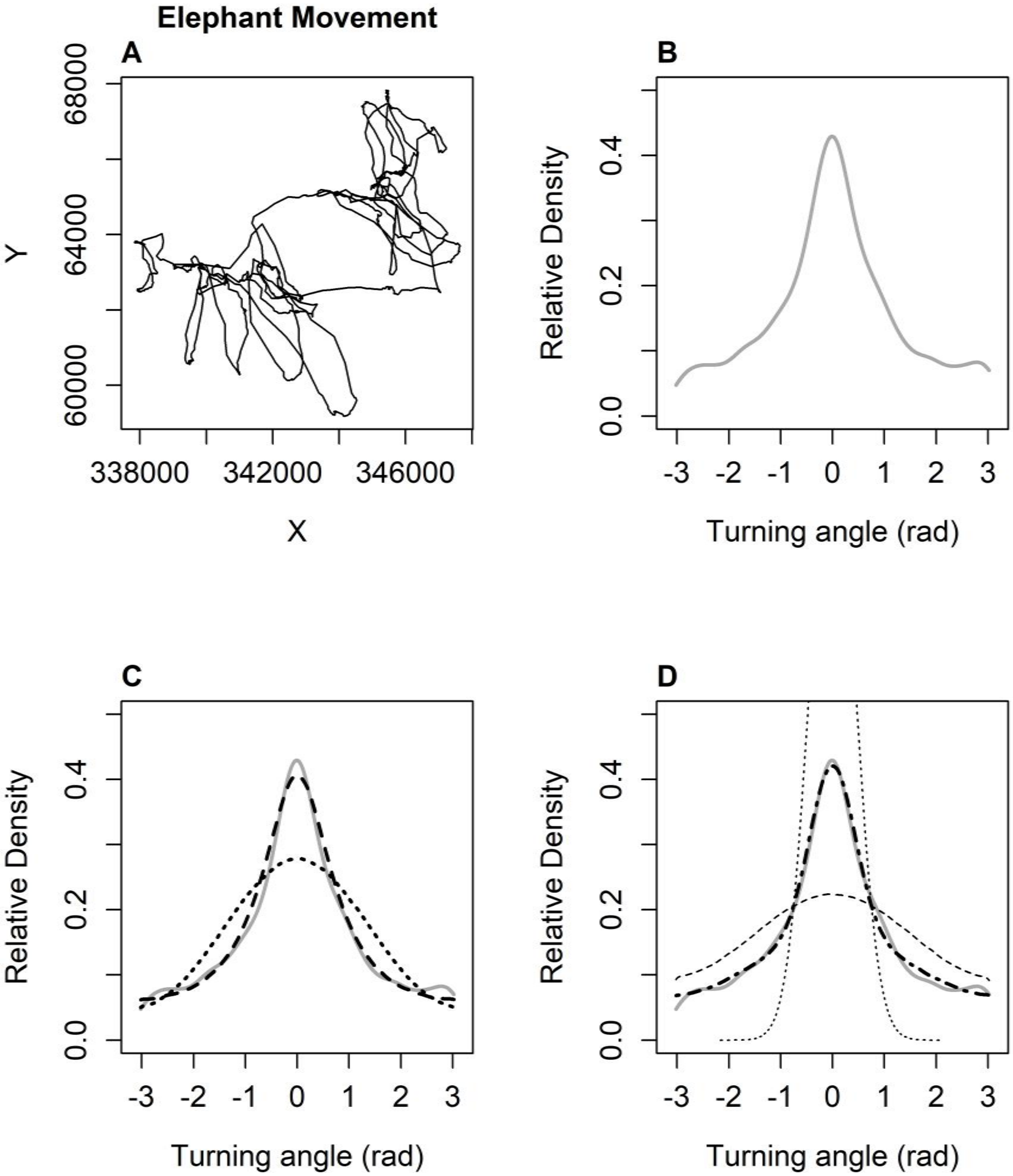
Figure 2: Movement data analysis of African elephant (ID: *Habiba*) from Wall et al (2014a). (A) shows the recorded movement path; (B) the corresponding turning angle distribution (grey); (C) the best fitting WC distribution (black-dashed; $\rho_{wc} = 0.4385$) and best fitting WN (black dotted; $\rho_{wn} = 0.3558$); (D) the best fitting mixed distribution (black – dashed and dotted) determined by numerical simulations with parameters $\rho_1 = 0.20, \rho_2 = 0.91, \omega = 0.71$. Thin black dashed line corresponds to a WN with concentration parameter ρ_1 and the black thin dotted line corresponds to a WN with concentration parameter ρ_2.

Figure 1



A
ω = 0
B
ω = 0.05
C
ω = 0.1
D
ω = 0.15
E
ω = 0.2
F
ω = 0.25
G
ω = 0.3
H
ω = 0.35
I
ω = 0.4
J
ω = 0.45
K
ω = 0.5
ρ1
ρ2
0.0
0.4
0.8
0.2
0.6
1.0

Figure 2


Elephant Movement
A
B
C
D
68000
64000
60000
Y
338000
342000
346000
X
Relative Density
0.0
0.2
0.4
-3
-2
-1
0
1
2
3
Turning angle (rad)