



Swansea University
Prifysgol Abertawe



Cronfa - Swansea University Open Access Repository

This is an author produced version of a paper published in:
Methods in Ecology and Evolution

Cronfa URL for this paper:
<http://cronfa.swan.ac.uk/Record/cronfa48296>

Paper:

Potts, J., Börger, L., Scantlebury, D., Bennett, N., Alagaili, A. & Wilson, R. (2018). Finding turning-points in ultra-high-resolution animal movement data. *Methods in Ecology and Evolution*, 9(10), 2091-2101.
<http://dx.doi.org/10.1111/2041-210X.13056>

This item is brought to you by Swansea University. Any person downloading material is agreeing to abide by the terms of the repository licence. Copies of full text items may be used or reproduced in any format or medium, without prior permission for personal research or study, educational or non-commercial purposes only. The copyright for any work remains with the original author unless otherwise specified. The full-text must not be sold in any format or medium without the formal permission of the copyright holder.

Permission for multiple reproductions should be obtained from the original author.

Authors are personally responsible for adhering to copyright and publisher restrictions when uploading content to the repository.

<http://www.swansea.ac.uk/library/researchsupport/ris-support/>

1 **Finding turning-points in ultra-high-resolution animal**
2 **movement data**

3 Jonathan R. Potts^{1,a}, Luca Börger², D. Michael Scantlebury³, Nigel C. Bennett^{4,5}, Abdulaziz Alagaili⁵, Rory
4 P. Wilson²

5 **Short title:** Turning-point algorithm

6 **Key words:** Animal movement, Behaviour, Biologging, Change point, High-resolution data,
7 Movement ecology, Turning point

8 **1** School of Mathematics and Statistics, University of Sheffield, Hicks Building, Hounsfield
9 Road, Sheffield, UK, S3 7RH

10 **a** E-mail: j.potts@sheffield.ac.uk

11 **2** Department of Biosciences, College of Science, Swansea University, Singleton Park, Swansea,
12 SA2 8PP, UK

13 **3** School of Biological Sciences, Institute for Global Food Security, Queens University Belfast,
14 Belfast BT9 7BL, UK

15 **4** Department of Zoology and Entomology, University of Pretoria, Pretoria 0002, South Africa

16 **5** Zoology Department, King Saud University, P.O Box 2455, Riyadh 11451, Saudi Arabia

17 Abstract

18 1. Recent advances in biologging have resulted in animal location data at unprecedentedly high
19 temporal resolutions, sometimes many times per second. However, many current methods for
20 analysing animal movement (e.g. step selection analysis or state-space modelling) were devel-
21 oped with lower-resolution data in mind. To make such methods usable with high-resolution
22 data, we require techniques to identify features within the trajectory where movement deviates
23 from a straight line.

24

25 2. We propose that the intricacies of movement paths, and particularly turns, reflect deci-
26 sions made by animals so that turn points are particularly relevant for behavioural ecologists.
27 As such, we introduce a fast, accurate algorithm for inferring turning-points in high-resolution
28 data. For analysing big data, speed and scalability are vitally important. We test our algo-
29 rithm on simulated data, where varying amounts of noise were added to paths of straight-line
30 segments interspersed with turns. We also demonstrate our algorithm on data of free-ranging
31 oryx (*Oryx leucoryx*). We compare our algorithm to existing statistical techniques for break-
32 point inference.

33

34 3. The algorithm scales linearly and can analyse several hundred-thousand data-points in
35 a few seconds on a mid-range desktop computer. It identified turnpoints in simulated data
36 with complete accuracy when the noise in the headings had a standard deviation of $\pm 8^\circ$, well
37 within the tolerance of many modern biologgers. It has comparable accuracy to the existing
38 algorithms tested, and is up to three orders of magnitude faster.

39

40 4. Our algorithm, freely available in R and Python, serves as an initial step in processing
41 ultra high-resolution animal movement data, resulting in a rarefied path that can be used as
42 an input into many existing step-and-turn methods of analysis. The resulting path consists of
43 points where the animal makes a clear turn, and thereby provides valuable data on decisions

44 underlying movement patterns. As such, it provides an important breakthrough required as a
45 starting point for analysing sub-second resolution data.

46 **1 Introduction**

47 Animal movement is a key process underlying many ecological systems (Nathan et al., 2008;
48 Kays et al., 2015; Hays et al., 2016). Until recently, our understanding of the drivers of move-
49 ment was limited by the resolution of data obtainable by technologies such as global positioning
50 systems (GPS) and Argos telemetry (Johnson et al., 2002; Jerde and Visscher, 2005; Hurford,
51 2009; McClintock et al., 2015). However, technological advances, particularly regarding ac-
52 celerometers and magnetometers, have enabled tracks to be constructed at extremely high
53 resolutions (Wilson et al., 2008; Brown et al., 2013; Noda et al., 2014; Walker et al., 2015;
54 Bidder et al., 2015; Williams et al., 2017). Indeed, often the time interval between consecutive
55 locations is shorter than the time it takes for an animal to travel a distance equal to its body
56 length, so that the movement data is, for all practical purposes, continuous (Wilmers et al.,
57 2015).

58 Whilst such data open up a wealth of opportunity for better understanding of animal
59 movement, many of the existing mathematical and statistical techniques for analysing location
60 data were developed with older, coarser data in mind. As such, they often fail to scale-up to
61 the new world of big, high-resolution data: techniques that work well on 1,000s of data points
62 gathered at hourly intervals may be very different to those required to analyse 1,000,000s of
63 data points at a resolution of 10Hz.

64 For example, many highly successful techniques, such as state-space modelling (Morales
65 et al., 2004; Jonsen et al., 2005; Patterson et al., 2008) and step selection analysis (Fortin
66 et al., 2005; Rhodes et al., 2005; Forester et al., 2009; Avgar et al., 2016), were formulated for
67 data where there is reasonable chance of finding interesting behavioural information in each
68 ‘step’ between successive data points. Yet, if the datapoints are only a fraction of a second
69 apart, the resulting information is minimal, and it is necessary to find points that correspond to

70 the animal doing something more interesting than simply carrying-on in the same straight-line
71 trajectory.

72 Consequently, for high-resolution data, we need techniques that can infer when the animal
73 is making a turn. The idea of examining animal paths as ‘steps and turns’ has been at the heart
74 of movement ecology for several decades (Kareiva and Shigesada, 1983; Bovee and Benhamou,
75 1988; Turchin et al., 1991; Turchin, 1998) in various forms, including step selection analysis,
76 biased correlated random walks (Codling et al., 2008), many state-space models, and even
77 continuous-time models (Parton et al., 2016). So to apply such ‘step and turn’ modelling
78 techniques to modern, ultra-high resolution data, we need a way of rapidly and accurately
79 inferring the turning points in the data-stream.

80 The aim of this paper is to provide such a technique. The idea is to view the path as
81 a stream of *headings* of the animal, rather than locations. We then look for switches in the
82 heading by sliding a small window across the path and observing where the standard deviation
83 across the window spikes (Fig. 1). This indicates a turn. If required, we can further post-
84 process the data by removing ‘small’ turning angles, that are not deemed to be indicative of
85 an actual behavioural decision. The resulting algorithm scales linearly with the length of the
86 data stream, and can process 100,000s of data points in a few seconds on a mid-range desktop
87 (Intel i7 2.5GHz processor).

88 This algorithm leads to a description of the animal’s movement in terms of straight-line
89 segments interspersed with turns, giving a biologically meaningful summary of the animal’s
90 movement behaviour whereby the turns are likely to represent actual decisions of the animal.
91 This contrasts with many studies involving lower-resolution data, where the turns are implicitly
92 assumed to occur precisely at the points where the locations were measured (Morales et al.,
93 2004; Fortin et al., 2005; Avgar et al., 2016) [but see Turchin et al. (1991); Codling and Plank
94 (2011), mentioned in more detail below]. Therefore, combined with high-resolution data, our
95 algorithm opens the door to more biologically accurate application of popular techniques for
96 analysing ‘move and turn’ data, such as step selection analysis and state space models.

97 To test the efficacy of our algorithm, we use a combination of simulated and real data.
98 For the simulations, we construct paths of straight-line segments joined together by sharp
99 turns, then add varying amounts of noise, to reflect both the error inherent in data-gathering
100 technologies and the noise arising from animal locomotion (e.g. small ‘rocking’ movements of
101 the sensor due to the animal’s gait, or avoidance of small obstacles like rocks). This enables us
102 to analyse the accuracy of our algorithm in inferring the correct turning points. We then use
103 data on free ranging oryx to demonstrate how to apply the technique to a real-world scenario.

104 Ours is not the first algorithm to segment data into straight lines and turns. Turchin et al.
105 (1991) developed a method that has proved popular in movement ecology for a number of
106 years. An alternative method was later put forward by Codling and Plank (2011). However,
107 these methods were both designed for the sort of low-resolution data that has historically been
108 available, and we show here that they do not perform so well with higher resolution data. Away
109 from movement ecology, several sophisticated and general techniques have been developed to
110 segment data streams, mainly concerned with studying DNA sequences [e.g. Picard et al.
111 (2005); Erdman and Emerson (2008); Franke et al. (2012); Rivera and Walther (2013)]. These
112 have the advantage of being well-grounded in statistical theory. The best-performing technique
113 [pruned dynamic programming, according to Hocking et al. (2013)] has also been written into a
114 flexible and convenient R package, called `Segmentor3IsBack` (Cleynen et al., 2014). However,
115 when we applied this algorithm to data, it was typically 10^3 times slower than ours, which
116 could cause it to be prohibitively slow for very long data streams.

117 In summary, we describe here a fast, light-weight algorithm for inferring turning points in
118 high-resolution animal movement data. We hope that this will enable more sophisticated use
119 of step-and-turn analysis techniques, where the turns are more closely related to the underlying
120 behavioural decisions of the animal (Wilson et al., 2013).

121 2 Methods

122 2.1 The turning-point algorithm

123 We describe an algorithm to be used on data of animal *headings*. This contrasts with many
 124 animal movement studies which focus on *locational* (or positional) data. The reasons for this
 125 are that (a) high resolution data tends to arrive from magnetometers that record headings
 126 rather than locations, and (b) headings are the natural parameter for determining turning
 127 points (TPs).

128 The essence of the algorithm is contained in the following two steps. First, we slide a window
 129 across a time-series of headings and looking for places where the squared circular standard
 130 deviation (SCSD) across that window spikes. The SCSD is a measure of the ‘spread’ of angles,
 131 and is used in place of the variance to account for the circular nature of angular distributions.
 132 Note that we choose not to use the term ‘circular variance’ as, in circular statistics, this is
 133 not consistently defined and may not be the square of the circular standard deviation [see e.g.
 134 Berens (2009)]. The spikes in SCSD indicate that the animal has turned. Second, we refine
 135 the set of candidate TPs by rejecting those that are below some threshold value.

136 To describe the algorithm precisely, we need to introduce some notation. Let h_1, \dots, h_N
 137 be a time series of headings for the animal, collected at evenly-spaced time-points t_1, \dots, t_N .
 138 Consider a small time window, W , about each data point and calculate the SCSD, s_i , of the
 139 heading across this window. In other words, s_i is the SCSD of $\{h_{i-W/2}, \dots, h_{i+W/2}\}$ for each
 140 time-point t_i where $i \in \{1 + W/2, \dots, N - W/2\}$ (W must be an even number). The SCSD is
 141 given by the following formula

$$142 \quad s_i = \ln \left(\frac{1}{\bar{R}_i^2} \right), \quad (1)$$

144 where $\bar{R}_i = \sqrt{\overline{\sin^2(h_i)} + \overline{\cos^2(h_i)}}$, $\overline{\sin(h_i)}$ is the average of $\{\sin(h_{i-W/2}), \dots, \sin(h_{i+W/2})\}$, and
 145 $\overline{\cos(h_i)}$ is the average of $\{\cos(h_{i-W/2}), \dots, \cos(h_{i+W/2})\}$. The SCSD copes with the fact that

$0^\circ = 360^\circ$, which stops us from using the usual definition of ‘standard deviation’.

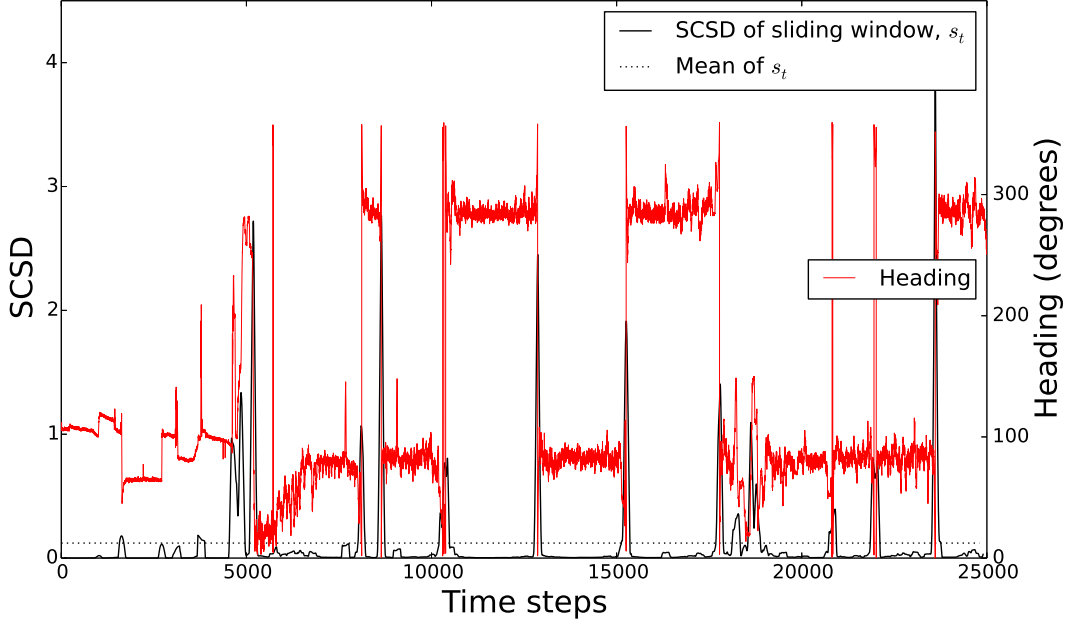


Fig. 1. Method for inferring turning points from circular standard deviation. The red curve gives the recorded headings of an example oryx, measured at 40Hz resolution. The squared circular standard deviation (SCSD) of the heading across a sliding window of size $W = 200$ data points is given by the black curve. Where this spikes above the mean SCSD, we infer that a turn might have taken place. We then do a second check to see that the heading has changed by more than a given threshold angle θ_{thresh} (see Main Text for more details). For example, this removes the points misidentified as turns at around timesteps 21,000 and 22,000.

146

147 The value of s_i will ‘spike’ when the animal turns sharply. We use these spikes to infer
 148 changes in the direction of the animal’s movement (Figure 1). More precisely, a *spike* in
 149 the time series, $\{s_{1+W/2}, \dots, s_{N-W/2}\}$, of SCSDs is defined to be a contiguous set of points,
 150 $\{s_i, \dots, s_{i+k}\}$, each of which is greater than the mean, μ , of the set $\{s_{1+W/2}, \dots, s_{N-W/2}\}$. The
 151 mid-point of each spike is collected, to form a subset of $\mathcal{T} = \{t_1, \dots, t_N\}$ of candidate TPs.
 152 This set is reduced further by removing any candidate TPs for which the turning angle is below
 153 a certain threshold, θ_{thresh} (see Supplementary Appendix A for details). This procedure results
 154 in a set $\{t_{c_1}, \dots, t_{c_n}\} \subset \mathcal{T}$ of *inferred TPs*.

155 Code for the complete algorithm is given in the Supplementary Information, as the R script
 156 `find_turnpoints.R` and the Python script `find_turnpoints.py`. Supplementary Appendix B
 157 explains how to modify and run the R code. Both programmes have the same function and the
 158 user can choose whichever language is more convenient.

159 2.2 Simulated data

160 To test the efficacy of this algorithm at picking out turning angles, we construct a collection of
 161 simulated trajectories. Each trajectory consists of 72,000 data points, which can be viewed, for
 162 example, as a 30 minute path collected at a resolution of 40Hz, or a four-hour path collected
 163 at 5Hz. The actual heading of the animal at time t is denoted by μ_t .

164 We assume that the times between successive TPs are drawn from an exponential distri-
 165 bution with mean η . For our simulations, $\eta = 1,200$ time-steps. Turning angles are drawn
 166 randomly and uniformly from the set $[-\pi, -\phi_{\text{thresh}}) \cup (\phi_{\text{thresh}}, \pi)$, so that ‘turns’ are always
 167 greater than a threshold value, ϕ_{thresh} . The set of times at which the simulated animal makes
 168 a turn is denoted by $\{T_1, \dots, T_m\}$. Our choice of $\eta = 1,200$ and a trajectory of length 72,000
 169 roughly mimics 30 minutes of oryx data, collected at 40Hz, turning on average every 30 seconds.

170 Because real data contains noise, we do not record the actual headings μ_t . Rather we
 171 simulate ‘observed’ headings, h_t , drawn from a von Mises distribution with mean μ_t and con-
 172 centration parameter κ . Therefore the resulting path is a sequence of straight-line segments
 173 with noise added, interspersed with turns of greater than ϕ_{thresh} . For our analysis, we set
 174 $\phi_{\text{thresh}} = 30^\circ$.

175 We construct simulated trajectories for a variety of values of κ and run each simulated
 176 trajectory through our inference algorithm for a variety of values of W and θ_{thresh} . For each
 177 trajectory, to determine how close the inferred set of TPs, $\{t_{c_1}, \dots, t_{c_n}\}$, is to the actual set,
 178 $\{T_1, \dots, T_m\}$, we calculate a True Positive Rate (TPR) and a False Positive Rate (FPR) using
 179 the following procedure. We split the path into windows of size W . If a window contains a
 180 true TP (i.e. one of T_1, \dots, T_m) then this window is considered a *True Condition*, otherwise it

181 is a *False Condition*. If the window corresponding to a True Condition (resp. False Condition)
 182 contains one of the values $\{t_{c_1}, \dots, t_{c_n}\}$ then it is a *True Positive* (resp. *False Positive*). Then
 183 the TPR (resp. FPR) is the number of true positives (resp. false positives) divided by the
 184 number of true conditions (resp. false conditions). Using the window in this way means that
 185 we accept as “True Positives” inferred TPs that are very close to real TPs (i.e. within W time-
 186 steps), but they do not have to be exactly the same points. Calculating (TPR,FPR) pairs for
 187 a variety of values of W and θ_{thresh} enables us to construct a receiver operating characteristic
 188 (ROC) curve for each value of κ (Brown and Davis, 2006).

189 We compare each ROC curve to the corresponding curve obtained by applying a previous
 190 turning-point algorithm, introduced by Turchin et al. (1991), to each path. Turchin’s algorithm
 191 determines a turning point by iterating through a time series of *locations* (rather than headings),
 192 x_0, \dots, x_K . If the locations x_0, \dots, x_{k-1} are all within a distance of ϵ from the straight line
 193 between x_0 and x_{k-1} , but some of the locations x_0, \dots, x_k are at a distance *greater* than ϵ from
 194 the straight line between x_0 and x_k , then the algorithm says that x_{k-1} is the first TP. Other
 195 TPs are constructed iteratively (for full details, see Turchin et al. (1991); Turchin (1998)).

196 For a given simulated trajectory, to compute the associated ROC curve for Turchin’s al-
 197 gorithm, we vary ϵ and calculate each TPR- and FPR-value. One would expect low ϵ to give
 198 many TPs, so produce high values for both FPR and TPR. On the other hand, a high value of
 199 ϵ might be expected to produce low FPR and TPR. We also compare our method to a more
 200 recent method of Codling and Plank (2011) (see Supplementary Appendix D).

201 **2.3 Case study on oryx movement**

202 To demonstrate the efficacy of our algorithm on a real dataset, we use high-resolution magne-
 203 tometer data of oryx living in Mahazat as-Sayd, a protected area located in west-central Saudi
 204 Arabia (28°15 N, 41°40E). The area consists of open steppe desert and is characterised by arid
 205 climate with hot summers, mild winters and low rainfall (Ostrowski et al., 2003). Vegetation is
 206 sparse, and predominated by perennial grasses and sporadically distributed small Acacia trees

207 (Mandaville, 1990).

208 The six oryx used in this study were captured during February 2015. After capture, they
 209 were fitted with loggers containing tri-axial accelerometers and tri-axial magnetometers (Daily
 210 Diary units, Wildbyte Technologies Ltd., Swansea, UK) which were set to record for 10 days
 211 at 40Hz in each channel. Each oryx was fitted with two daily diary units. One unit was glued
 212 to the head of the animal, behind the horns, using quick-set epoxy resin, and the other was
 213 fixed using cable ties and adhesive tape to a collar around the animals neck. Data for this
 214 study were taken from the magnetometer fitted to the neck, so that the headings represent the
 215 trajectory of the animal, rather than the direction it is facing.

216 Following logger deployment, animals were allowed to recover in an outside enclosure ($25 \times$
 217 25m) for approximately eight hours after which they were released into the larger enclosure
 218 ($2 \times 1\text{km}$). For this study, we examine a sample path in the larger enclosure for each of the six
 219 oryx. These paths have varying lengths, the shortest is 1 hour ($\sim 140,000$ headings) and the
 220 longest is 2 hours 15 minutes ($\sim 320,000$ headings; see Supplementary Table ST1).

221 For real trajectories, analysis of TPR and FPR is not possible, since we do not have knowl-
 222 edge of the ‘true’ turning-points. Instead, to assess how good a proposed set of TPs is, we
 223 construct a path of straight-line segments between each pair of consecutive TPs. We com-
 224 pare this piecewise-linear model with the path given by the data, assuming that the animal
 225 is moving at constant speed, v (chosen arbitrarily). This latter path is given by a collection
 226 of points $\mathbf{x}(t_1), \dots, \mathbf{x}(t_N)$, where $\mathbf{x}(t_i) = (x(t_i), y(t_i))$ is a 2D vector for each $i \in \{1, \dots, N\}$,
 227 $\mathbf{x}(t_1) = (0, 0)$, and

$$\begin{aligned}
 228 \quad x(t_i) &= \sum_{j=1}^{i-1} v(t_{j+1} - t_j) \cos(h_j), \\
 229 \quad y(t_i) &= \sum_{j=1}^{i-1} v(t_{j+1} - t_j) \sin(h_j). \tag{2} \\
 230
 \end{aligned}$$

231 For a given set of inferred TPs, $\{t_{c_1}, \dots, t_{c_n}\}$, the model path is

$$232 \quad \mathbf{m}(t_i) = \mathbf{x}(t_{c_k}) + \frac{t_i - t_{c_k}}{t_{c_{k+1}} - t_{c_k}} [\mathbf{x}(t_{c_{k+1}}) - \mathbf{x}(t_{c_k})] + \xi_t, \quad \text{for } t_i \in [t_{c_k}, t_{c_{k+1}}], \quad (3)$$

233

234 where $\xi_t \sim N(\mathbf{0}, \sigma^2)$ is a random variable from a Normal distribution with zero mean and
 235 covariance matrix $\sigma^2 I$, where I is the two-dimensional identity matrix.

236 Note that, if we were to have a data set containing the speed at any point in time, the
 237 assumption that v is constant could be dropped and the actual speeds used instead. Further-
 238 more, if we were to have locational data rather than just headings, we could use the measured
 239 locations as $(x(t_i), y(t_i))$ rather than constructing them using Equation (2). For the oryx data
 240 we do not have such information. However, if any future study contains locational (or po-
 241 sitional) information at a high frequency – e.g. from dead-reckoning (Wilson et al., 2008) –
 242 then these locations should be used directly to construct a piecewise-linear model similar to
 243 Equation (3).

244 An estimate for the standard deviation, σ , can be calculated empirically, as follows. Let
 245 $\mathbf{d}(t_i) = \mathbf{x}(t_i) - \langle \mathbf{m}(t_i) \rangle$ for each $i \in \{1, \dots, N\}$, where $\langle \mathbf{m}(t_i) \rangle$ is the mean of $\mathbf{m}(t_i)$. Since $\langle \mathbf{m}(t_i) \rangle$
 246 is the point on a straight line between inferred turning points corresponding to $\mathbf{x}(t_i)$, $\mathbf{d}(t_i)$ is
 247 the deviation of the measured location from the corresponding location on this straight line (i.e.
 248 the residual). Then σ is estimated to be the standard deviation of the set $\{\mathbf{d}(t_1), \dots, \mathbf{d}(t_N)\}$. A
 249 smaller σ indicates a better fit of the model to the data. Since σ is dependent on our (arbitrary)
 250 choice of animal speed, we define a *normalised standard deviation*, $\bar{\sigma} = \sigma / \langle l_i \rangle$, where $\langle l_i \rangle$ is
 251 the mean of the step lengths $l_i = |\mathbf{x}(t_{c_k}) - \mathbf{x}(t_{c_{k-1}})|$. Then $\bar{\sigma}$ is a dimensionless quantity,
 252 independent of v .

253 For analysis of the oryx data, we choose values for W and θ_{thresh} by examining (a) those that
 254 perform well on simulated data (i.e. low FPR and high TPF), (b) those that are biologically
 255 justifiable (i.e. expert opinion). We also construct videos of the trajectories, with the turnpoints
 256 super-imposed, so one can visually inspect whether the estimations of TPs look reasonable to
 257 the human eye. This aids in determining whether our choice of W and θ_{thresh} give the correct

258 information.

259 We use the oryx data to compare our algorithm with the output of the `Segmentor3isBack`
 260 package. `Segmentor3isBack` is a general-use programme that segments data-streams into K
 261 segments, where K is fixed (and user-defined). The package can also find the theoretical optimal
 262 value of K for fitting a trajectory of straight-line segments interspersed with break-points. We
 263 compare the time it takes to run the respective algorithms, as well as the resulting $\bar{\sigma}$ values,
 264 when K is set to be equal to the number of segments given by our algorithm. For this, we fix
 265 $W = 40$ and $\theta_{\text{thresh}} = 30^\circ$.

266 We also compare our algorithm to Turchin’s algorithm, where ϵ is set so that resulting
 267 number of segments is equal to that given by our algorithm. Turchin’s algorithm has to be
 268 applied to the reconstructed path (Equation 2) rather than the raw headings. It also defines
 269 turns in such a way as to minimise the distance between the model path (Equation 3) and the
 270 reconstructed one (Equation 2). It can thus be viewed as providing an rough estimate of the
 271 minimum $\bar{\sigma}$ that may be attainable. We also examined the effect of pre-processing our data
 272 by smoothing-out possible noise arising from the animal’s gait or minor obstacle avoidance,
 273 before running it through our algorithm. This provides a prior smoothing before the implicit
 274 smoothing given by choice of window size, W . Details are given in Supplementary Appendix
 275 C.

276 3 Results

277 3.1 Simulated data: comparison with previous approaches

278 Fig. 2 shows three simulated trajectories, with an increasing level of noise from left to right.
 279 The ROC curves (Panels j-l) indicate that the method proposed here finds the TPs with
 280 significantly better accuracy than the method of Turchin et al. (1991). Indeed, for the example
 281 where the SD in the error of the heading measurements is only $\pm 8.1^\circ$, our method had a TPR
 282 of 1 and FPR of 0, meaning it caught all of the true TPs and did not falsely identify any

283 TPs. This suggests that if an animal really is moving in straight-lines separated by distinct
 284 turns, then our method will be extremely accurate at picking these up, as long as the SD in
 285 the heading measurements is not too great (i.e. of the order of $< 10^\circ$). The method of Codling
 286 and Plank (2011) performed worst of the three and is probably only suitable where data is
 287 much lower resolution (see Supplementary Appendix D).

288 3.2 Oryx data

289 To identify turns in the oryx data, we found that a window size of $W = 40$ data points
 290 and $\theta_{\text{thresh}} = 30^\circ$ gave accurate results for determining both broad- and fine-scale movement
 291 decisions. Since data were taken 40 times per second, this means that we are only integrating-
 292 out behavioural features that occur on a subsecond resolution, which are likely to be minimal-to-
 293 nonexistent. Fig. 3a shows an example of an oryx path with these inferred turns superimposed.
 294 At first glance, it appears as if there are a number of places where turns are identified where
 295 they do not appear to be present. However, by zooming-in, we observe that the algorithm is
 296 actually correctly identifying very fine-scale movements correctly (see inset of Fig. 3a).

297 Depending on the biological question being sought, a user may not be interested in very
 298 fine-scale movements, so may wish to smooth out behaviour over a longer time-interval. For
 299 example, we also used a window size of $W = 200$, corresponding to five seconds of movement,
 300 to analyse the same oryx path as in Fig. 3b. Here, the very fine-scale movements are integrated
 301 out, leaving a much smaller set of TPs (155 compared with 498). Videos of the trajectory of
 302 Oryx 1 with the inferred turning points for $W = 40$ and $W = 200$ are given in Supplemen-
 303 tary Videos SV1 (cpsv1.mp4) and SV2 (cpsv2.mp4), respectively. Pre-processing data using
 304 subsampling or smoothing had almost no effect on the inference (Supplementary Appendix C).

305 Table 1 summarises the number of TPs inferred for each of the six oryx paths, using
 306 $\theta_{\text{thresh}} = 30^\circ$ and $W \in \{40, 200\}$, together with the normalised standard deviation, $\bar{\sigma}$ of the
 307 data from the piecewise-linear model given by Equation (3). Observe that this normalised
 308 standard deviation is similar for both $W = 40$ and $W = 200$. The reason for this is that,

Table 1. Application to oryx data. The first column is the oryx identifier. The second (resp. fourth) gives the normalised standard deviation of the model from the data for a window size of $W = 40$ (resp. $W = 200$) data points, representing 1 second (resp. 5 seconds) of movement. The third (resp. fifth) gives the number of inferred turning points (TPs) for a window size of $W = 40$ (resp. $W = 200$) data points.

Oryx ID	$\bar{\sigma}$ ($W = 40$)	No. TPs ($W = 40$)	$\bar{\sigma}$ ($W = 200$)	No. TPs ($W = 200$)
1	0.0601	498	0.0612	155
2	0.0783	892	0.0792	261
3	0.381	557	0.508	153
4	0.477	863	0.180	217
5	0.176	639	0.187	183
6	0.247	929	0.198	237

309 although the step lengths are longer for the piecewise-linear models with fewer TPs ($W = 200$),
 310 small-scale turns are treated as noise rather than signal, thus increasing the amount of error
 311 proportionately.

312 Comparing our algorithm with `Segmentor3isBack`, we see that the latter tends to be about
 313 10^3 times slower (e.g. 46 minutes compared to 3.6 seconds; see Supplementary Table ST1 for
 314 precise figures). The resulting path of straight lines and turns is a marginally better fit in five
 315 of the six cases (Supplementary Table ST1), which is to be expected, since `Segmentor3isBack`
 316 is designed to find the theoretical best-fit path. However, the difference tends to be minor,
 317 both by comparing $\bar{\sigma}$ -values and by visually inspecting the paths (Supplementary Figure SF4).

318 Comparison with Turchin’s algorithm, applied to the reconstructed path (Equation 2),
 319 reveals that Turchin’s algorithm is 1-2 orders of magnitude slower (e.g. 220 seconds compared
 320 to 3.6 seconds; see Supplementary Table ST2 for precise figures). Turchin’s algorithm generally
 321 results in a lower $\bar{\sigma}$. This is to be expected as Turchin’s method defines turns as places where
 322 there is a deviation from a straight-line of more than a fixed value, so implicitly seeks to
 323 minimise $\bar{\sigma}$. However, our simulation analysis reveals that this is not such an accurate method
 324 for determining where turns have actually occurred, as it is more likely to misdetect noise as
 325 signal than our approach (Fig. 2). Therefore the resulting inferred set of turning points is not
 326 as reliable as our algorithm, even though the constructed piecewise linear path may turn out

327 to have a slightly better fit.

328 4 Discussion

329 We have described a fast, accurate algorithm for detecting turning-points in animal movement
330 data, particularly tailored for use with very high-resolution data. Given a path of straight-
331 moves and turns, where headings have been measured to within an accuracy of within $\pm 8.1^\circ$
332 standard deviation, the algorithm succeeds in detecting all turning points, without falsely
333 detecting any (Fig. 2). If the accuracy is only $\pm 19.6^\circ$, the algorithm was still able to identify
334 56 of 59 turnings points, whilst misclassifying 8 non-turns as turning points. Since many
335 modern measuring devices, such as magnetometers, have an accuracy within $\pm 5^\circ$ (Li et al.,
336 2006), this suggests our algorithm is well-suited to identifying turning-points in such data (as
337 long as the tag is attached well and does not shift location on the animal significantly).

338 This accuracy compares well with previous methods. Perhaps the most oft-used in move-
339 ment ecology has been that of Turchin et al. (1991), which ours markedly improves upon (Fig.
340 2). A more sophisticated method, imported from literature on statistics and DNA segmen-
341 tation (Cleyngen et al., 2014), does a reasonable job on real data (Supplementary Table ST1,
342 Supplementary Figure SF4) but is around three orders of magnitude slower than our method
343 (Supplementary Table ST1). Indeed, the speed of our algorithm is a very important feature.
344 Datasets are becoming ever larger, so having fast algorithms without significant scaling prob-
345 lems is very important. Ours will analyse hundreds of thousands of data points in a few seconds
346 on an ordinary desktop and scales linearly. Therefore, we expect that even tracks of a billion
347 locations (40Hz for a year) would be analysable in only a few hours.

348 Our method complements existing research in analysis of behavioural change-points in an-
349 imal paths, recently reviewed by Edelhoff et al. (2016). These methods look at movement
350 paths at a broader scale, segmenting them into sections corresponding to different behavioural
351 modes. Edelhoff et al. (2016) explained how this analysis can be broken down into four stages
352 [see Fig. 1 from Edelhoff et al. (2016)], with the third stage ostensibly very similar to the

353 sort of turning-point analysis described here. However, a detailed look reveals that the papers
354 that are referenced regarding this third stage are, in fact, seeking answers to issues that are
355 somewhat different to the aims of this paper, which we explain in the next two paragraphs.
356 Our contention will be that the method presented here is a sub-step prior to Edelhoff *et al.*'s
357 third step, required when data is very high resolution (a case not considered in Edelhoff *et al.*).

358 Several methods for behavioural changepoint analysis (BCPA) have been proposed in the
359 literature. Many of them begin with a description of movement in terms of summary statistics.
360 For example, Gurarie et al. (2009) gives an algorithm for determining significant changes in
361 persistence velocity and turning velocity. Similar ideas were given a more general and the-
362 oretical treatment by Buchin et al. (2011). Nams (2014) generalises BCPA by developing a
363 technique for detecting behavioural changepoints that can make use of a wide variety of sum-
364 mary statistics, and also clusters the resulting path-segments into distinct behavioural states.
365 Postlethwaite et al. (2013) proposes a ‘straightness index’ for rapid inference of behavioural
366 states. Gurarie et al. (2016) summarises and compares a variety of methods for detecting
367 behavioural changes.

368 However, all such behavioural changepoint techniques require that the path be already
369 described using some sort of summary statistic (e.g. velocity, tortuosity, turning angle distri-
370 bution etc.). Our paper provides a method to infer specific summary statistics (i.e. step lengths
371 and turning angles) from big, high-resolution datasets, thus enabling existing behavioural
372 changepoint analysis techniques to be used with high-resolution data. We thus anticipate
373 that the output of our algorithm could be effectively used as an input to BCPA and similar
374 methods.

375 Our method, based on the circular statistics of headings, has some mathematical similarities
376 with certain methods of deriving tortuosity in movement paths (Benhamou, 2004). If the un-
377 derlying distribution of headings comes from a wrapped normal distribution then the SCSD is
378 an unbiased estimator of the variance of the underlying (unwrapped) normal distribution (Mar-
379 dia, 2014). The mean of the cosine of a wrapped normal distribution is then $c = \exp(-\text{SCSD}/2)$

380 (Mardia and Jupp, 2009). The quantity c has been applied to turning angles of animal paths
 381 to measure the tortuosity of such paths, since it interpolates between 0 for an uncorrelated
 382 random walk to 1 for ballistic movement (Bovet and Benhamou, 1988). Indeed, it has been
 383 used, combined with a sliding window, to detect changes in the behavioural mode of animals
 384 (Benhamou, 2004). So there are some strong similarities between this approach and ours. The
 385 main differences are that the existing studies using c have been concerned with behavioural
 386 changepoints rather than (smaller scale) turning points, and generally applied to turning angles
 387 rather than headings (since the underlying questions are different). Also, the wrapped normal
 388 assumption that links the two is not so easy to justify when applied to our scenario, especially
 389 near turning points.

390 The study of Byrne et al. (2009) also examines changes in behaviour, but this time by
 391 explicitly looking for a change in direction. The aim was to identify the points at which
 392 an animal decides to move towards a particular location. The method compares the sum of
 393 the lengths of two straight line segments $|\mathbf{x}_{n-1} - \mathbf{x}_n| + |\mathbf{x}_n - \mathbf{x}_{n+1}|$ with the resultant length
 394 $|\mathbf{x}_{n-1} - \mathbf{x}_{n+1}|$ to infer a change in direction if the latter is much smaller than the former.

395 Whilst this method asks a similar question to the one examined here, in fact it is not
 396 designed to pick up every turn, but just those that indicate a decision to move to a specific
 397 location. Indeed, it quite deliberately ignores small, temporary changes in direction, as Fig.
 398 4 from Byrne et al. (2009) demonstrates. Our algorithm, on the other hand, does attempt to
 399 detect every change in direction, however temporary it is. However, it is possible for the user
 400 to factor-out temporary changes by choosing a large window size, W . Ultimately, the choice of
 401 whether it is best to use our algorithm or the one from Byrne et al. (2009) will depend on the
 402 specific biological question, and the summary statistics desired to answer it (notwithstanding
 403 additional issues regarding computational speed for big data sets).

404 In general, the choice of both W and θ_{thresh} depends on various factors and a combination
 405 of statistical tests and expert knowledge will be required in order for this be set appropriately.
 406 Our simulation analysis indicates that there is an optimal W and θ_{thresh} for a given simulation

407 scenario, defined by the point at the upper-left-most extreme of the ROC curve (see Fig.
 408 2j-1). However, for real data it is not possible to construct such a ROC curve and find the
 409 actual optimum. Instead we recommend calculating the variation of the real trajectory from a
 410 piecewise-linear model trajectory with turns at the inferred turning points. Such a trajectory
 411 is described in Equation (3). Furthermore, we give a dimensionless quantity, denoted by $\bar{\sigma}$,
 412 for testing this fit. To choose W and θ_{thresh} , we recommend, in the first instance, running our
 413 algorithm over a range of values and calculating $\bar{\sigma}$ for each.

414 Although lower values of $\bar{\sigma}$ indicate a better fit, the correct choice of W and θ_{thresh} also
 415 depends upon the biological properties of the study species and the underlying scientific ques-
 416 tions. This is where expert opinion becomes important, and blindly picking the W and θ_{thresh}
 417 that minimise $\bar{\sigma}$ may not always be the best option. In particular, the turn radii of the species
 418 is an important quantity. The minimum turn radius of an animal depends partly on its move-
 419 ment speed, with faster moving individuals (or species) tending to have greater turn radii
 420 (Alexander, 2002a), with this condition generally holding whether the animal in question is
 421 aerial (Thomas, 1996), terrestrial (Alexander, 2002b) or aquatic (but see Blake et al. (1995)
 422 and references therein), although values differ in the different media. In particular, it is worth
 423 noting that terrestrial mammals, such as the oryx used in this work, may turn through 90° in
 424 less than 1s whereas, because a flying bird has a turn radius that is proportional to the flight
 425 speed squared (Thomas, 1996), a similar 90° turn by a large gliding bird such as a condor
 426 (*Vultur gryphus*) may take several seconds during which time the bird may have travelled 50m
 427 (McGahan, 1973).

428 In addition to this, there are extrinsic factors that may mean an animal moves in a curve
 429 rather than a straight line between successive decisions to change direction. For example,
 430 topography could affect a terrestrial animal, and water (resp. air) currents will affect aquatic
 431 (resp. airborne) animals. Therefore, when finding turning-points in such data, it is necessary
 432 to factor-out such extrinsic effects. (Note that the oryx studied here are unlikely to be largely
 433 affected by such factors, as they roam on relatively flat and open terrain.) If, once all these

434 factors are accounted for, an animal’s path is curvilinear, rather than consisting of straight-
435 lines and turns, then our algorithm is simply inappropriate for analysing the path and should
436 not be used.

437 If there is noise in the data arising from specific known artifacts, such as effects of rocky
438 terrain or animal gait, then it may be beneficial to pre-process the data prior to analysis so as
439 to smooth-out this noise. Some possible pre-processing methods are given in Supplementary
440 Appendix C. We recommend users test for such noise, ideally by examining short paths where
441 the animal has been directly observed to go in a straight line. If this is not possible, attaching
442 the magnetometer to a human in the same terrain where the animal resides can give an idea
443 (albeit imperfect) of the noise due to an uneven terrain. Similarly, we recommend that users
444 obtain an idea of the noise inherent in the magnetometer by leaving it immobile at a fixed
445 heading for some time. If there any of these types of noise are either large or autocorrelated,
446 then it may be beneficial to examine the effect of pre-processing the data. For the oryx exam-
447 ined here, however, such pre-processing had almost no effect on the inference (Supplementary
448 Appendix C).

449 Our method makes an implicit choice to define a candidate turn as a point at which the
450 SCSD goes above the global mean. Whilst this choice appears to work adequately in the
451 situations studied here, it is not the only possible way to define a turn. For example, one could
452 examine the *cumulative* SCSD and look for sharp changes in the resulting time series, using
453 the methods described by Knell and Codling (2012). There, the authors examined how the
454 cumulative sum (CUSUM) of any summary statistic (not necessarily SCSD) will change sharply
455 over time when the behaviour changes (in the context of our study, this ‘behavioural change’
456 would be between straight-moving and turning). However, the CUSUM method also relies on
457 an arbitrary choice of a parameter [labelled ε by Knell and Codling (2012)] to determine where
458 such sharp changes occur in the time series. Although the authors demonstrate a method for
459 calculating an optimal ε in certain circumstances, it is not clear whether it would always be
460 possible to derive such an optimum in any situation. Therefore, whilst a CUSUM approach

461 to SCSD may sometimes be a useful option for the user to bear in mind, it may also end up
462 simply replacing one arbitrary choice with another.

463 In summary, our algorithm is a quick and accurate method for splitting up long streams
464 of ultra high resolution animal movement data into straight-line segments and turns. The
465 output of such segmentation can then be used to detect behavioural features using the myriad
466 techniques that require step-and-turn descriptions, such as step selection analysis, behavioural
467 changepoint analysis, state space models, and more. In particular, step selection analysis (SSA)
468 would greatly benefit from an approach whereby the ‘steps’, which typically mean a move-
469 ment from one measured location to the next, are replaced with the more behaviourally-driven
470 ‘moves’ from one turning point to the next. SSA seeks to understand whether a movement
471 along one straight-line path is preferable to another and how that is correlated to environmen-
472 tal covariates. Evidently, this inference will be improved if the animal’s actual movement from
473 one point to the next well-approximates a straight line. So replacing ‘steps’ with ‘moves’ seems
474 to be the correct way forward.

475 With some of the other aforementioned techniques, such as behavioural changepoint analy-
476 sis, it is less clear whether ‘steps’ or ‘moves’ would be more appropriate [to borrow terminology
477 from Turchin (1998), Section 5.2]. If a technique uses turning angles as a summary statistic
478 for analysing behaviour, it would seem more appropriate to use ‘moves’ as the angles would
479 correspond to actual turns by the animal, which may be energetically costly (Wilson et al.,
480 2013). For example, the calculations of persistence velocity and turning velocity in Gurarie
481 et al. (2009) would be improved by using ‘moves’. However, if the analysis relies upon regular
482 sampling of animal locations then one may have to use ‘steps’. For example, techniques such
483 as Morales et al. (2004); Beyer et al. (2013) rely on step-length distributions between locations
484 gathered (roughly) regularly in time. These would need to be carefully adapted before use with
485 a sequence of ‘moves’ of different time-periods. That said, if such an adaption can be made,
486 a switch from ‘step length distribution’ to ‘move length distribution’ would be possible within
487 these frameworks, and may make them more behaviourally-grounded. In conclusion, given its

488 potential for use to improve a broad range of existing techniques, our algorithm should serve
489 as an important tool for making sense of the type big data increasingly available to movement
490 ecologists.

491 **Acknowledgements**

492 The data-gathering part of this project was funded by the National Plan for Science, Technology
493 and Innovation (MAARIFAH), King Abdulaziz City for Science and Technology, Kingdom of
494 Saudi Arabia, Award Number (11-ENV1918-02) and the Deanship of Scientific Research at
495 the King Saud University through Vice Deanship of Research Chairs. Ethical clearance for
496 the data-gathering was obtained from the University of the Witwatersrand Animal Ethics
497 Committee (clearance certificate number 2014/53/D). Permission to work in the field was
498 granted by the President of the Saudi Wildlife Authority. JRP acknowledges support from the
499 National Environmental Research Council (NERC) grant NE/R001669/1. The authors thank
500 two anonymous reviewers and an associate editor for comments that have helped improve the
501 manuscript.

502 **Author Contributions**

503 JRP, LB, RPW conceived and designed the research; JRP, LB performed the research; LB,
504 DMS, NCB, AA, RPW provided data; JRP led the writing of the manuscript. All authors
505 contributed critically to the drafts and gave final approval for publication.

506 **Data Accessibility**

507 Data used in this manuscript will be archived on FigShare after the manuscript has been
508 accepted for publication.

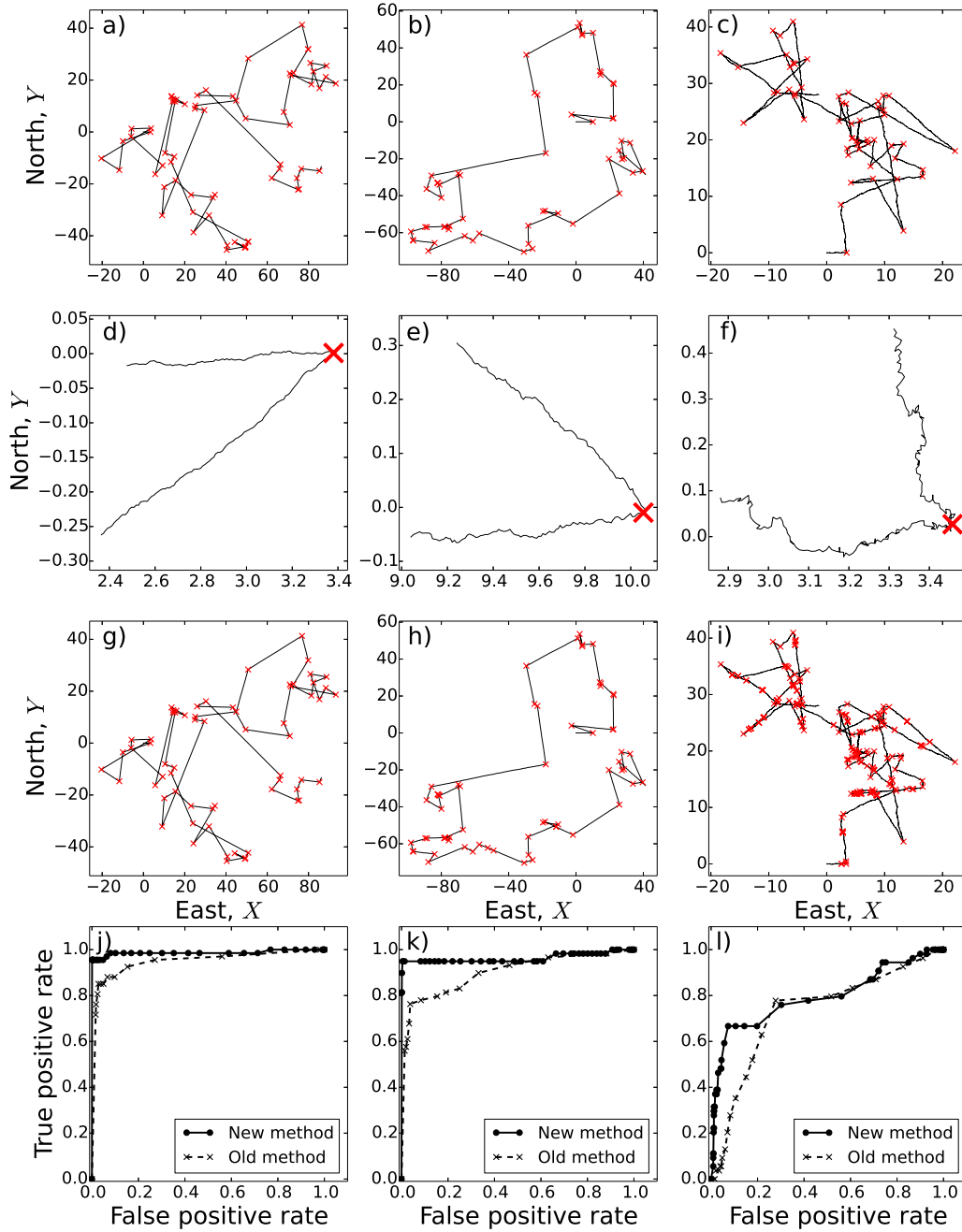


Fig. 2. Simulated paths. The top panels (a-c) give three simulated trajectories with increasing amounts of noise from left to right. Specifically, panel (a) has $\kappa = 50$ so the standard deviation (SD) in the heading is 8.1° ; panel (b) has $\kappa = 10$, corresponding to an SD of 18.6° ; panel (c) has $\kappa = 1$ so $SD = 72.6^\circ$. The actual TPs are superimposed on the trajectories in panels (a-c) as crosses. Panels (d-f) zoom in on panels (a-c) respectively around the first turning point, giving a better visual impression of the noise in the data. Panels (g-i) show the same trajectories as (a-c) respectively, but this time the red crosses show inferred TPs using the inference method described in the Main Text. Panels (j-l) show ROC curves, corresponding to the trajectories in (a-c) respectively, for both the method introduced here (dots and solid curves) and an older method due to Turchin et al. (1991) (crosses and dashed curves).

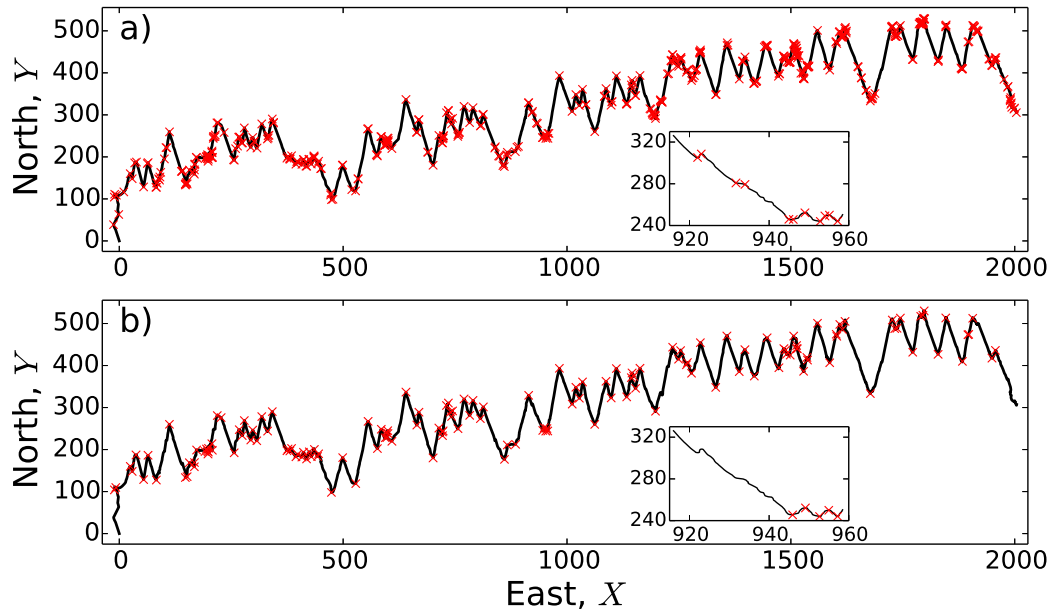


Fig. 3. Example oryx path. Both panels show the path of a single oryx (ID=1 in Table 1). In Panel (a) the red crosses denote the inferred turning points (TPs) using a window size $W = 40$ (corresponding to 1 second of movement) and a turning threshold angle of $\theta_{\text{thresh}} = 30^\circ$. The inset zooms in on a fragment of the path, to demonstrate the very small-scale turns that are revealed by this analysis, alongside broader-scale turning decisions. In Panel (b), the TPs are inferred using a window size of $W = 200$ (5 seconds of movement). The very fine-scale turns are now smoothed-out (inset, Panel b), which allows the user to focus on broader-scale patterns.

References

- 509
- 510 Alexander, R. M. 2002*a*. The Merits and Implications of Travel by Swimming, Flight and
511 Running for Animals of Different Sizes. *Integrative and Comparative Biology*, **42**:1060–1064.
- 512 Alexander, R. M. 2002*b*. Stability and manoeuvrability of terrestrial vertebrates. *Integrative*
513 *and Comparative Biology*, **42**:158–164.
- 514 Avgar, T., J. R. Potts, M. A. Lewis, and M. S. Boyce. 2016. Integrated step selection analysis:
515 bridging the gap between resource selection and animal movement. *Methods in Ecology and*
516 *Evolution*, **7**:619–630.
- 517 Benhamou, S. 2004. How to reliably estimate the tortuosity of an animal's path:: straightness,
518 sinuosity, or fractal dimension? *Journal of Theoretical Biology*, **229**:209–220.
- 519 Berens, P. 2009. CircStat: a MATLAB toolbox for circular statistics. *Journal of Statistical*
520 *Software*, **31**:1–21.
- 521 Beyer, H. L., J. M. Morales, D. Murray, and M.-J. Fortin. 2013. The effectiveness of Bayesian
522 state-space models for estimating behavioural states from movement paths. *Methods in*
523 *Ecology and Evolution*, **4**:433–441.
- 524 Bidder, O., J. Walker, M. Jones, M. Holton, P. Urge, D. Scantlebury, N. Marks, E. Magowan,
525 I. Maguire, and R. Wilson. 2015. Step by step: reconstruction of terrestrial animal movement
526 paths by dead-reckoning. *Movement Ecology*, **3**:23.
- 527 Blake, R., L. Chatters, and P. Domenici. 1995. Turning radius of yellowfin tuna (*Thunnus*
528 *albacares*) in unsteady swimming manoeuvres. *Journal of fish biology*, **46**:536–538.
- 529 Bovet, P. and S. Benhamou. 1988. Spatial analysis of animals' movements using a correlated
530 random walk model. *Journal of Theoretical Biology*, **131**:419–433.
- 531 Brown, C. D. and H. T. Davis. 2006. Receiver operating characteristics curves and related
532 decision measures: A tutorial. *Chemometrics and Intelligent Laboratory Systems*, **80**:24–38.

- 533 Brown, D. D., R. Kays, M. Wikelski, R. Wilson, and A. P. Klimley. 2013. Observing the
534 unwatchable through acceleration logging of animal behavior. *Animal Biotelemetry*, **1**:20.
- 535 Buchin, M., A. Driemel, M. Van Kreveld, and V. Sacristán. 2011. Segmenting trajectories:
536 A framework and algorithms using spatiotemporal criteria. *Journal of Spatial Information*
537 *Science*, **2011**:33–63.
- 538 Byrne, R., R. Noser, L. Bates, and P. Jupp. 2009. How did they get here from there? Detecting
539 changes of direction in terrestrial ranging. *Animal Behaviour*, **77**:619–631.
- 540 Cleynen, A., M. Koskas, E. Lebarbier, G. Rigaille, and S. Robin. 2014. Segmentor3IsBack: an R
541 package for the fast and exact segmentation of Seq-data. *Algorithms for Molecular Biology*,
542 **9**:6.
- 543 Codling, E. A. and M. J. Plank. 2011. Turn designation, sampling rate and the misidentification
544 of power laws in movement path data using maximum likelihood estimates. *Theoretical*
545 *Ecology*, **4**:397–406.
- 546 Codling, E. A., M. J. Plank, and S. Benhamou. 2008. Random walk models in biology. *Journal*
547 *of the Royal Society Interface*, **5**:813–834.
- 548 Edelhoff, H., J. Signer, and N. Balkenhol. 2016. Path segmentation for beginners: an overview
549 of current methods for detecting changes in animal movement patterns. *Movement Ecology*,
550 **4**:21.
- 551 Erdman, C. and J. W. Emerson. 2008. A fast Bayesian change point analysis for the segmen-
552 tation of microarray data. *Bioinformatics*, **24**:2143–2148.
- 553 Forester, J., H. Im, and P. Rathouz. 2009. Accounting for animal movement in estimation of
554 resource selection functions: sampling and data analysis. *Ecology*, **90**:3554–3565.
- 555 Fortin, D., H. Beyer, M. Boyce, D. Smith, T. Duchesne, and J. Mao. 2005. Wolves influence

- 556 elk movements: Behavior shapes a trophic cascade in Yellowstone National Park. *Ecology*,
557 **86**:1320–1330.
- 558 Franke, J., C. Kirch, and J. T. Kamgaing. 2012. Changepoints in times series of counts. *Journal*
559 *of Time Series Analysis*, **33**:757–770.
- 560 Gurarie, E., R. D. Andrews, and K. L. Laidre. 2009. A novel method for identifying behavioural
561 changes in animal movement data. *Ecology Letters*, **12**:395–408.
- 562 Gurarie, E., C. Bracis, M. Delgado, T. D. Meckley, I. Kojola, and C. M. Wagner. 2016. What is
563 the animal doing? Tools for exploring behavioural structure in animal movements. *Journal*
564 *of Animal Ecology*, **85**:69–84.
- 565 Hays, G. C., L. C. Ferreira, A. M. Sequeira, M. G. Meekan, C. M. Duarte, H. Bailey, F. Bailleul,
566 W. D. Bowen, M. J. Caley, D. P. Costa, et al. 2016. Key questions in marine megafauna
567 movement ecology. *Trends in Ecology & Evolution*, **31**:463–475.
- 568 Hocking, T. D., G. Schleiermacher, I. Janoueix-Lerosey, V. Boeva, J. Cappo, O. Delattre,
569 F. Bach, and J.-P. Vert. 2013. Learning smoothing models of copy number profiles using
570 breakpoint annotations. *BMC Bioinformatics*, **14**:164.
- 571 Hurford, A. 2009. GPS measurement error gives rise to spurious 180 turning angles and strong
572 directional biases in animal movement data. *PLoS One*, **4**:e5632.
- 573 Jerde, C. L. and D. R. Visscher. 2005. GPS measurement error influences on movement model
574 parameterization. *Ecological Applications*, **15**:806–810.
- 575 Johnson, C. J., D. C. Heard, and K. L. Parker. 2002. Expectations and realities of GPS animal
576 location collars: results of three years in the field. *Wildlife Biology*, **8**:153–159.
- 577 Jonsen, I. D., J. M. Flemming, and R. A. Myers. 2005. Robust state–space modeling of animal
578 movement data. *Ecology*, **86**:2874–2880.

- 579 Kareiva, P. and N. Shigesada. 1983. Analyzing insect movement as a correlated random walk.
580 *Oecologia*, **56**:234–238.
- 581 Kays, R., M. C. Crofoot, W. Jetz, and M. Wikelski. 2015. Terrestrial animal tracking as an
582 eye on life and planet. *Science*, **348**:aaa2478.
- 583 Knell, A. S. and E. A. Codling. 2012. Classifying area-restricted search (ARS) using a partial
584 sum approach. *Theoretical Ecology*, **5**:325–339.
- 585 Li, Y., A. Dempster, B. Li, J. Wang, and C. Rizos. 2006. A low-cost attitude heading reference
586 system by combination of GPS and magnetometers and MEMS inertial sensors for mobile
587 applications. *Journal of Global Positioning Systems*, **1**:0.
- 588 Mandaville, J. P. 1990. *Flora of Eastern Saudi Arabia*. London: Kegan Paul Int.
- 589 Mardia, K. V. 2014. *Statistics of directional data*. Academic press.
- 590 Mardia, K. V. and P. E. Jupp. 2009. *Directional statistics*, volume 494. John Wiley & Sons.
- 591 McClintock, B. T., J. M. London, M. F. Cameron, and P. L. Boveng. 2015. Modelling animal
592 movement using the Argos satellite telemetry location error ellipse. *Methods in Ecology and*
593 *Evolution*, **6**:266–277.
- 594 McGahan, J. 1973. Gliding flight of the Andean condor in nature. *Journal of Experimental*
595 *Biology*, **58**:225–237.
- 596 Morales, J. M., D. T. Haydon, J. Frair, K. E. Holsinger, and J. M. Fryxell. 2004. Extract-
597 ing more out of relocation data: building movement models as mixtures of random walks.
598 *Ecology*, **85**:2436–2445.
- 599 Nams, V. O. 2014. Combining animal movements and behavioural data to detect behavioural
600 states. *Ecology Letters*, **17**:1228–1237.

- 601 Nathan, R., W. M. Getz, E. Revilla, M. Holyoak, R. Kadmon, D. Saltz, and P. E. Smouse. 2008.
602 A movement ecology paradigm for unifying organismal movement research. *Proceedings of*
603 *the National Academy of Sciences*, **105**:19052–19059.
- 604 Noda, T., Y. Kawabata, N. Arai, H. Mitamura, and S. Watanabe. 2014. Animal-mounted
605 gyroscope/accelerometer/magnetometer: In situ measurement of the movement performance
606 of fast-start behaviour in fish. *Journal of Experimental Marine Biology and Ecology*, **451**:55–
607 68.
- 608 Ostrowski, S., J. B. Williams, and K. Ismael. 2003. Heterothermy and the water economy of
609 free-living Arabian oryx (*Oryx leucoryx*). *Journal of Experimental Biology*, **206**:1471–1478.
- 610 Parton, A., P. G. Blackwell, and A. Skarin. 2016. Bayesian inference for continuous time animal
611 movement based on steps and turns. In *International Conference on Bayesian Statistics in*
612 *Action*, pages 223–230. Springer.
- 613 Patterson, T. A., L. Thomas, C. Wilcox, O. Ovaskainen, and J. Matthiopoulos. 2008. State-
614 space models of individual animal movement. *Trends in Ecology & Evolution*, **23**:87–94.
- 615 Picard, F., S. Robin, M. Lavielle, C. Vaisse, and J.-J. Daudin. 2005. A statistical approach for
616 array CGH data analysis. *BMC Bioinformatics*, **6**:27.
- 617 Postlethwaite, C. M., P. Brown, and T. E. Dennis. 2013. A new multi-scale measure for
618 analysing animal movement data. *Journal of Theoretical Biology*, **317**:175–185.
- 619 Rhodes, J., C. McAlpine, D. Lunney, and H. Possingham. 2005. A spatially explicit habitat
620 selection model incorporating home range behavior. *Ecology*, **86**:1199–1205.
- 621 Rivera, C. and G. Walther. 2013. Optimal detection of a jump in the intensity of a Poisson
622 process or in a density with likelihood ratio statistics. *Scandinavian Journal of Statistics*,
623 **40**:752–769.

- 624 Thomas, A. L. 1996. The flight of birds that have wings and a tail: variable geometry expands
625 the envelope of flight performance. *Journal of Theoretical Biology*, **183**:237–245.
- 626 Turchin, P. 1998. *Quantitative analysis of movement: measuring and modeling population re-*
627 *distribution in animals and plants, volume 1.* Sinauer Associates Sunderland, Massachusetts,
628 USA.
- 629 Turchin, P., F. Odendaal, and M. Rausher. 1991. Quantifying insect movement in the field.
630 *Environmental Entomology*, **20**:955–963.
- 631 Walker, J. S., M. W. Jones, R. S. Laramée, M. D. Holton, E. L. Shepard, H. J. Williams,
632 D. M. Scantlebury, J. M. Nikki, E. A. Magowan, I. E. Maguire, et al. 2015. Prying into the
633 intimate secrets of animal lives; software beyond hardware for comprehensive annotation in
634 Daily Diary tags. *Movement Ecology*, **3**:29.
- 635 Williams, H. J., M. D. Holton, E. L. Shepard, N. Largey, B. Norman, P. G. Ryan, O. Duriez,
636 M. Scantlebury, F. Quintana, E. A. Magowan, et al. 2017. Identification of animal movement
637 patterns using tri-axial magnetometry. *Movement Ecology*, **5**:6.
- 638 Wilmers, C. C., B. Nickel, C. M. Bryce, J. A. Smith, R. E. Wheat, and V. Yovovich. 2015. The
639 golden age of bio-logging: how animal-borne sensors are advancing the frontiers of ecology.
640 *Ecology*, **96**:1741–1753.
- 641 Wilson, R. P., I. W. Griffiths, P. A. Legg, M. I. Friswell, O. R. Bidder, L. G. Halsey, S. A.
642 Lambertucci, and E. L. C. Shepard. 2013. Turn costs change the value of animal search
643 paths. *Ecol. Lett.*, **16**:1145–1150.
- 644 Wilson, R. P., E. Shepard, and N. Liebsch. 2008. Prying into the intimate details of animal
645 lives: use of a daily diary on animals. *Endangered Species Research*, **4**:123–137.