



Using time-series similarity measures to compare animal movement trajectories in ecology

Ian R. Cleasby^{1,7} · Ewan D. Wakefield² · Barbara J. Morrissey³ · Thomas W. Bodey^{4,5} · Steven C. Votier⁶ · Stuart Bearhop⁴ · Keith C. Hamer⁷

Received: 22 May 2019 / Revised: 11 October 2019 / Accepted: 17 October 2019 / Published online: 21 November 2019
© The Author(s) 2019

Abstract

Identifying and understanding patterns in movement data are amongst the principal aims of movement ecology. By quantifying the similarity of movement trajectories, inferences can be made about diverse processes, ranging from individual specialisation to the ontogeny of foraging strategies. Movement analysis is not unique to ecology however, and methods for estimating the similarity of movement trajectories have been developed in other fields but are currently under-utilised by ecologists. Here, we introduce five commonly used measures of trajectory similarity: dynamic time warping (DTW), longest common subsequence (LCSS), edit distance for real sequences (EDR), Fréchet distance and nearest neighbour distance (NND), of which only NND is routinely used by ecologists. We investigate the performance of each of these measures by simulating movement trajectories using an Ornstein-Uhlenbeck (OU) model in which we varied the following parameters: (1) the point of attraction, (2) the strength of attraction to this point and (3) the noise or volatility added to the movement process in order to determine which measures were most responsive to such changes. In addition, we demonstrate how these measures can be applied using movement trajectories of breeding northern gannets (*Morus bassanus*) by performing trajectory clustering on a large ecological dataset. Simulations showed that DTW and Fréchet distance were most responsive to changes in movement parameters and were able to distinguish between all the different parameter combinations we trialled. In contrast, NND was the least sensitive measure trialled. When applied to our gannet dataset, the five similarity measures were highly correlated despite differences in their underlying calculation. Clustering of trajectories within and across individuals allowed us to easily visualise and compare patterns of space use over time across a large dataset. Trajectory clusters reflected the bearing on which birds departed the colony and highlighted the use of well-known bathymetric features. As both the volume of movement data and the need to quantify similarity amongst animal trajectories grow, the measures described here and the bridge they provide to other fields of research will become increasingly useful in ecology.

Significance statement

As the use of tracking technology increases, there is a need to develop analytical techniques to process such large volumes of data. One area in which this would be useful is the comparison of individual movement trajectories. In response, a variety of measures

Communicated by L. Z. Garamszegi

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00265-019-2761-1>) contains supplementary material, which is available to authorized users.

✉ Ian R. Cleasby
i.r.cleasby@gmail.com

¹ RSPB Centre for Conservation Science, Inverness, UK

² Institute of Biodiversity, Animal Health and Comparative Medicine, University of Glasgow, Glasgow, UK

³ Rivers and Lochs Institute, Inverness College, University of the Highlands and Islands, Inverness, UK

⁴ Centre for Ecology and Conservation, University of Exeter, Cornwall Campus, Penryn, UK

⁵ School of Biological Sciences, University of Auckland, Auckland, New Zealand

⁶ Environment and Sustainability Institute, University of Exeter, Cornwall Campus, Penryn, UK

⁷ School of Biology, University of Leeds, Leeds, UK

of trajectory similarity have been developed within the information sciences. However, such measures are rarely used by ecologists who may be unaware of them. To remedy this, we apply five common measures of trajectory similarity to both simulated data and real ecological dataset comprising of movement trajectories of breeding northern gannets. Dynamic time warping and Fréchet distance performed best on simulated data. Using trajectory similarity measures on our gannet dataset, we identified distinct foraging clusters centred on different bathymetric features, demonstrating one application of such similarity measures. As new technology and analysis techniques proliferate across ecology and the information sciences, closer ties between these fields promise further innovative analysis of movement data.

Keywords Information science · Movement ecology · Route fidelity · Site fidelity · Tracking data · Trajectory clustering

Introduction

In recent years the widespread adoption of radio- and satellite-based telemetry has led to a marked increase in the volume of animal movement data (Kays et al. 2015). While such rapid technological development has advanced the study of animal ecology, the amount of data obtained presents a challenge to researchers analogous to big data problems seen in other disciplines (Thums et al. 2018). In response, techniques have been developed concerning habitat selection (Aarts et al. 2009), navigation (Freeman et al. 2010) and behavioural classification (Langrock et al. 2012). Nevertheless, an open problem for ecologists remains how best to quantify similarity in animals space use, both within and amongst individuals and groups, using movement data. This is a necessary step both in addressing fundamental research questions and for many wildlife management and conservation tasks (Fieberg and Kochanny 2005; Demšar et al. 2015; Lascelles et al. 2016). For example, at the group level, assessing the similarity of space use underpins studies on territoriality (Bateman et al. 2015), spatial segregation (Wakefield et al. 2013; Cleasby et al. 2015) and temporal changes in distribution (Clapp and Beck 2015). At the individual level, the concept of space use similarity has been used to investigate site and route fidelity (Freeman et al. 2010; Wakefield et al. 2015), habitat specialisation (McIntyre et al. 2017) and the ontogeny of foraging behaviour (Votier et al. 2017).

Typically, animal movement data consist of a discrete time-series of observed locations that collectively form a trajectory. The challenges of analysing complex trajectory data are not unique to animal tracking studies and are encountered by all studies that track moving objects (Demšar et al. 2015). Consequently, analytical approaches developed in other disciplines may be applicable in ecology, including a range of measures developed to calculate the similarity of movement trajectories (Ranacher and Tzavella 2014). However, because most trajectory similarity indices have been developed outside of ecology, they may not be well known amongst ecologists. For example, most trajectory similarity studies in ecology use nearest neighbour distance (NND) (e.g. Freeman et al. 2010; Demšar et al. 2015; Fayet et al. 2017; Votier et al. 2017), but use of other trajectory similarity measures is rare. Moreover,

in a recent review of metrics for ecological movement data (Seidel et al. 2018) none of the similarity measures described here was mentioned yet they are routinely used in other fields. Here, we illustrate how a range of similarity measures can be applied to animal movement data to quantify the similarity of two or more movement trajectories. To do so, we first define similarity conceptually. We then introduce similarity measures commonly used in other fields and assess their performance as we alter different parameters of an Ornstein-Uhlenbeck (OU) movement model. Using an OU model, we alter (1) the point of attraction, (2) the strength of attraction to this point and (3) the noise added to the movement process. Similarity between two trajectories is expected to decline as the strength of attraction to the point of attraction declines or as the noise in the movement process increases. In addition, as the point of attraction differs between trajectories, similarity should decline and when points of attraction are in the opposite direction, similarity should be at its lowest. Ideally, any similarity measure should be able to distinguish between different scenarios in which these parameters are altered, and this should be reflected in the similarity values calculated. As well as simulated data, Keogh and Kaszety (2003) highlight the importance of applying time-series similarity measures to real datasets, which typically have much higher variance than simulated datasets. Therefore, we also demonstrate the application of similarity measures to an example ecological dataset comprising trajectories of breeding northern gannets (*Morus bassanus*). We go on to demonstrate how similarity measures can be used to create distance matrices, enabling similar trajectories to be grouped using clustering methods in order to compare trajectory similarity within and between individuals.

Concerning similarity

The concept of similarity is founded upon three intuitions (Lin 1998): (1) the more features two objects share, the more similar they are; (2) the more differences there are between two objects, the less similar they are; and (3) maximum similarity occurs when two objects are identical. Similarity (or dissimilarity) is generally quantified as the either the cost of transforming one object into another or as the inverse of the

distance between objects (Faloutsos et al. 1997). The general notion of similarity can be conceptualised by mathematical functions called similarity measures that are used to compare objects. The inputs of a similarity measure are two objects and the output is a number. Similarity is related to distance which is generally the inverse of similarity (i.e. as distance between objects increases similarity decreases, see also Supplementary Material, Section S1 Similarity and Distance Metrics) and in practice distance measures are also often used to assess similarity directly.

From an ecological perspective, we are often interested in identifying and explaining patterns underlying movement data. Estimating the similarity of two or more instances of movement, particularly in geographic space and time, provides one means of achieving this (Demšar et al. 2015). When considering movement, one can assess similarity across a hierarchy of movement patterns (see Dodge et al. 2008). For example, movement can be split into separate temporal, spatial and spatio-temporal components and a distinction is often made between primary (e.g. spatial coordinates) and derived (e.g. speed) movement parameters (see Ranacher and Tzavella 2014). One can also distinguish instances in which the focus of analysis rests upon assessing the similarity of movement within an object over time versus instances in which the similarity of movement between different objects is key (Dodge et al. 2008). The focus of the current paper is to look at the similarity of movement in geographic space. However, it should be noted that similarity can also be assessed in other dimensions, such as derived spatial indices (e.g. speed) or environmental dimensions (e.g. wind speed).

In movement ecology, space use is often summarised using a utilisation distribution (UD) (Fieberg and Kochanny 2005). At their simplest, UD's are 2-dimensional probability distributions that represent the probability of encountering an animal at a given location (though UD methods can incorporate a temporal component, Kranstauber et al. 2012). Comparing UD's across individuals or groups have been used to address ecological questions pertaining to site fidelity, territoriality and temporal change in space use amongst others (Wakefield et al. 2015; Abrahms et al. 2018). To this end, UD's are often compared quantitatively using the degree to which they overlap in space, with high spatial overlap indicating UD's are more alike (Fieberg and Kochanny 2005). Recently, methods to assess the similarity of utilisation distributions have been developed and have been used to visualise and compare patterns of similarity in space use over time using clustering techniques (Kranstauber et al. 2017), demonstrating the broader utility of similarity-based approaches within movement ecology. However, as UD's are primarily designed to characterize space use, the measures described here may be better suited to questions relating to trajectory similarity and route fidelity.

Alongside comparison of UD's, an alternative approach to assessing movement similarity is to investigate the similarity

of movement trajectories directly using location data. Assessing the similarity of trajectories provides a formal means of quantifying the relationship between two or more moving objects. Ecological studies that focus upon either route similarity or route fidelity (similar trajectories within a defined group e.g. within the same individual) are typically based on this approach (Freeman et al. 2010; Fayet et al. 2016; Votier et al. 2017). Similarity measures provide a relative measure to compare the similarity of one trajectory to another. Moreover, by comparing the similarity of all trajectory pairs one can create a distance matrix that provides a basis for clustering trajectories into groups with similar properties (Demšar et al. 2015). The most common clustering approaches for time-series data are hierarchical or partitional clustering (Sarda-Espinosa 2017). Clustering can be a useful means of visually summarising large datasets, making it easier to discern movement patterns (Andrienko and Andrienko 2013). Clustering has also been used to identify representative or prototypical paths within groups of trajectories (Lee et al. 2007; Freeman et al. 2010); index and classify trajectories (Vlachos et al. 2002); investigate temporal change in space use (Shoji et al. 2016); and define the location of key activity centres using the centroids of the different trajectory clusters identified (McClintock and Michelot 2017).

Measures of trajectory similarity

Generally, trajectory similarity is quantified using the distance between trajectories. Within the plethora of trajectory similarity measures available, there are several that are used frequently, usually in the analysis of time-series (Ding et al. 2008) or for geometric shape matching (Alt 2009). Popular measures of trajectory similarity include Euclidean distance, Dynamic Time Warping (DTW, Senin 2008), Longest Common Subsequence distance (LCSS, Vlachos et al. 2002) and Edit Distance for Real sequence (EDR, Chen et al. 2005), while popular shape-based measures that try to catch geometric features of the trajectories include the Fréchet distance (Fréchet 1906). Due to their widespread use, functions for calculating these measures have already been implemented in commonly used statistical software, such as R and Python, making them easily accessible to ecologists (Salvador and Chan 2007; Toohey 2015; Sarda-Espinosa 2017; Supplementary Material - Table S1). Within ecology the use of such measures is not commonplace and instead variations of the nearest neighbour distance (NND) tend to be used to calculate trajectory similarity (e.g. Fayet et al. 2016; Votier et al. 2017). The Euclidean distance, sometimes called the L_2 -Norm, is the most commonly encountered similarity measure. However, it is a lock-step measure that requires one-to-one point matching of trajectories of equal length (Ranacher and Tzavella 2014), which may limit its use when analysing animal movement

trajectories as these have typically irregular lengths. Therefore, we focus here on DTW, LCSS, EDR, Fréchet distance and NND. Below we briefly describe each of these measures (see also Table 1). To aid interpretation we also provide example R code for calculating DTW, LCSS, EDR and Fréchet distance (Supplementary Material, S2: R Code For Similarity Measures, see also S3: Visual Examples of Similarity Measures).

In the following section we assume a trajectory T_A that contains (m) a series of m timestamped n dimensional points $a_i = (a_{i,1}, \dots, a_{i,n})$:

$$T_A = ((t_1, a_1), \dots, (t_m, a_m)) \tag{1}$$

where t_i are ordered, discrete timestamps. The length of a trajectory is given by the number of discrete timestamps it contains. Trajectory points are usually recorded in two dimensions (longitude and latitude represented by x and y), but higher dimension trajectories are possible. In the pairwise examples that follow, T_A is compared with another trajectory T_B of length k :

$$T_B = ((t_1, b_1), \dots, (t_k, b_k)) \tag{2}$$

Dynamic time warping

DTW algorithms search recursively through all point combinations between two trajectories to identify the path between those trajectories with the smallest

Table 1 Summary of different similarity measures. Computation time refers to standard calculations, different statistical packages may provide faster computation if using optimized algorithms. Note all measures are non-metric with the exception of Fréchet distance

Similarity Measure	Method	Range	Parameters	Elastic	Sensitivity to outliers	Sensitivity to time shifting	Computation time
Dynamic time warping	Point-based	0–∞ Larger values = lower similarity	Point index spacing—integer of the maximum index difference between trajectory 1 and trajectory 2 allowed in calculation. Can be set as unlimited. Tuning may improve clustering performance.	Yes	High. All points must be matched.	Low	Quadratic, $O(n^2)$
Longest common sub-sequence	Point-based	0–1 Larger values = greater similarity	δ —maximum index (time) difference when comparing points from two trajectories. Can be set as unlimited. Tuning may improve clustering performance. ϵ —maximum distance in each dimension allowed for two points to be considered equivalent	Yes	Low. Only common sub-sequences are matched. Outliers will not be matched. Gaps between matching sub-sequences not considered which can lead to inaccuracy.	Low	Quadratic
Edit distance with real penalty	Point-based	0–1 Larger values = lower similarity	ϵ —maximum distance in each dimension allowed for two points to be considered equivalent	Yes	Medium. Similar to LCSS points are classified as matching or not matching. Unlike LCSS gap elements receive a penalty.	Low	Quadratic
Fréchet distance	Shape matching	0–∞ Larger values = lower similarity	None	Yes	High. Comparison based on shape of trajectories which will be influenced by noise/outliers.	High	Polynomial, $O(nm \log(nm))$
Nearest neighbour distance	Point-based	0–∞ Larger values = lower similarity	None	Yes	High. All points must be matched.	Low	$O(nd + kn)$ d = no. dimensions being compared k = no. of neighbours to identify

distance using dynamic programming. Consequently, DTW is an elastic measure because it can be used with trajectories of different lengths, but is sensitive to noise, because all spatial points are matched, including any outliers. Given two trajectories T_A and T_B of length m and k respectively, DTW aligns these trajectories by creating an m -by- k distance matrix in which the (i th, j th) element is $(a_i - b_j)^2$. Here, the distance between points a_i and b_j represents the cost of aligning them. To find the best alignment between T_A and T_B , a path is drawn through the distance matrix that minimizes the total cumulative distance between them, called the *warping path*. The warping path begins in the bottom-left corner of the matrix and ends at the top-right corner (the first element of both trajectories must be matched and the last element of both trajectories must be matched) and must be contiguous (warping path can only go forward one step at a time) and monotonic (warping path cannot go backwards in time). Typically, the lowest cost warping paths will be close to the diagonal of the created distance matrix. DTW is defined as

$$DTW(T_A, T_B) = \begin{cases} 0 & \text{if } m = k = 0 \\ \infty & \text{if } m = 0 \text{ or } k = 0 \\ d(a_1, b_1) + \min \begin{cases} DTW(\text{Rest}(T_A), \text{Rest}(T_B)), \\ DTW(\text{Rest}(T_A), T_B), \\ DTW(T_A, \text{Rest}(T_B)) \end{cases} & \text{otherwise} \end{cases} \quad (3)$$

$$LCSS(T_A, T_B) = \begin{cases} 0 & \text{if } m = 0 \text{ or } k = 0 \\ LCSS(\text{Rest}(T_A), \text{Rest}(T_B)) + 1 & \text{if } |m-k| \leq \delta \text{ and } |a_{m,x} - b_{k,x}| \leq \varepsilon \text{ and } |a_{k,y} - b_{k,y}| \leq \varepsilon \\ \max \begin{cases} LCSS(\text{Rest}(T_A), T_B) \\ LCSS(T_A, \text{Rest}(T_B)) \end{cases} & \text{otherwise} \end{cases} \quad (4)$$

Here, δ represents the index difference used when comparing points between trajectories and controls how far to search in time to match a point in one trajectory with a point in another. LCSS is unbounded and depends upon the length of the sequences compared; therefore, the length of the shortest trajectory is typically used to normalize this method as an LCSS ratio (LCSS ratio = $LCSS / \min(|T_A|, |T_B|)$) that ranges from 0 (lowest similarity) to 1 (highest similarity). Like DTW, LCSS is an elastic measure, but unlike DTW it does not require that all points in a trajectory are matched, making it less sensitive to outliers and noise (Toohey and Duckham 2015); however, the user must subjectively define ε and δ (Long and Nelson 2013). Within an ecological context, Vlachos et al. (2002) used LCSS to cluster the movement patterns of a variety of marine mammals.

Here, $\text{Rest}(T_A)$ and $\text{Rest}(T_B)$ refer to the trajectories T_A and T_B with the first element (location) removed and $d(a_1, b_1) = |a_1 - b_1|$. DTW measures are non-negative and unbounded, with larger DTW values denoting greater distance (dissimilarity) between time-series. DTW was recently used in an ecological context by McClintock and Michelot (2017) to cluster the trajectories of foraging grey seals (*Halichoerus grypus*) and identify centres of activity.

Longest common subsequence

The LCSS finds the longest subsequence that two trajectories, A and B , have in common (Vlachos et al. 2002). A subsequence is defined as an alignment of elements that occurs in both trajectories given that the order of the remaining elements is preserved. Elements within a subsequence need not occur at the same time within both trajectories, but they must occur in the same order. Traditional LCSS was designed to capture similarities between discrete points where exact matches were possible. For use on movement trajectories Vlachos et al. (2002) extended LCSS to cover elements with real values such as spatial coordinates via a distance-based matching threshold, ε . Points that are $\leq \varepsilon$ units apart are considered matching and the LCSS is the total count of the number of points that are matched. LCSS can be calculated following Vlachos et al. (2002) as:

Edit distance for real sequence

Edit distance measures of similarity are based on counting how many operations are required to turn one string, A , into another string, B . Operations include inserting, deleting or substituting elements in A . The more similar strings A and B are, the fewer operations are required to make them equivalent, resulting in a lower edit distance score. For assessing the similarity of movement trajectories two common edit distance approaches are edit distance with real penalty (ERP, Chen and Ng 2004) and edit distance for real sequence (EDR, Chen et al. 2005). In these approaches, deletion is treated as adding in a gap to the other string, with the added element referred to as a gap element. When matching strings, the elements in A and B are symbols or characters. However, when dealing with time-series and trajectories, elements in T_A and T_B are real

numbers. Therefore, instead of enforcing strict equality between elements in T_A and T_B a distance threshold ϵ (as with LCSS) is used to determine approximate equality. In fact, edit distance is closely related to LCSS, but whereas LCSS counts matching locations, edit distance counts the number of operations required to turn T_A into T_B . One key reason for the development of EDR and ERP was that LCSS does not account for the length of gaps between matching elements when identifying common subsequences. In contrast, both EDR and ERP

assign penalties to such gap elements. While EDR and ERP are related, EDR may be more robust to noise and a more accurate measure of trajectory similarity as demonstrated by Chen et al. (2005); hence, we focus upon EDR here. EDR can be normalized by the length of the longest trajectory being compared (EDR ratio = $EDR / \max(|T_A|, |T_B|)$) to generate a score between 0 (highest similarity) and 1 (lowest similarity). EDR is defined as

$$EDR(T_A, T_B) = \begin{cases} \begin{matrix} m \\ k \end{matrix} & \begin{matrix} \text{if } k = 0 \\ \text{if } m = 0 \end{matrix} \\ \min \left\{ \begin{matrix} EDR(Rest(T_A), Rest(T_B) + \text{subcost}, \\ EDR(Rest(T_A), T_B) + 1, \\ EDR(T_A, Rest(T_B)) + 1 \end{matrix} \right\} & \text{otherwise} \end{cases} \tag{5}$$

A pair of trajectory elements $(a_{i,x}, a_{i,y})$ and $(b_{i,x}, b_{i,y})$ are considered matching if and only if $|a_{i,x} - b_{i,x}| \leq \epsilon$ and $|a_{i,y} - b_{i,y}| \leq \epsilon$. Cost of insertion, deletion or substitution operations is assumed to be 1. Subcost = 0 if (a_i, b_i) matches and 0 otherwise. EDR is not metric as it does not satisfy triangle inequality (Chen and Ng 2004). At present, we are not aware of any ecological study that has compared trajectories using EDR.

Fréchet distance

The Fréchet distance is a distance measure for continuous shapes or curves that considers the location and sequence of points within the compared trajectories (Besse et al. 2015). The Fréchet distance can be described by the analogy to a person walking a dog on an extendable leash (Aronov et al. 2006). The person and dog move along their respective trajectories. Both can control their own speed but are not allowed to backtrack. The Fréchet distance is the minimum length of the leash that ensures that the dog is always connected to its owner.

When calculating Fréchet distance, trajectories are approximated as polygonal curves. For example, trajectory T_A could be approximated as the polygonal curve $P_A[0, m]$ with m segments. P_A can be parameterized using the parameter $a \in \mathbb{R}$ so that $P_A(a)$ refers to a given position on the curve, with $P_A(0)$ and $P_A(m)$ referring to the first and last vertices of the curve respectively. Similarly, T_B could be approximated as the curve $Q_B[0, k]$ with $Q_B(b)$ referring to a position on the curve. The Fréchet distance is

$$FD(P_A, Q_B) = \min_{\substack{a: [0, 1] \rightarrow [0, m] \\ b: [0, 1] \rightarrow [0, k]}} \left\{ \max_{t \in [0, 1]} d(P(a(t)), Q(b(t))) \right\} \tag{6}$$

where $a(t)$ and $b(t)$ range over continuous and increasing functions, with $a(0) = 0, a(1) = m, b(0) = 0$ and $b(1) = k$ and d is a distance operator. Fréchet distance is shape-dependent and does not use point matching as the other techniques described here do. Consequently, it does not consider temporal aspects of movement such as speed (e.g. Buchin et al. 2011a) but it can be used on trajectories of different length. It is also sensitive to noise because every point of the two trajectories being compared is used in its calculation. Nevertheless, it can still be useful for trajectory comparison and clustering (Buchin et al. 2010, 2011b). A discretized version of the Fréchet distance is described by Eiter and Mannila (1994) and can be applied to movement data collected by discrete fixes.

Nearest neighbour distance

Unlike the distance measures listed above, the nearest neighbour distance is relatively well known amongst ecologists and perhaps due to its simplicity has been used widely to assess trajectory similarity in order to determine route fidelity (Freeman et al. 2010; Meier et al. 2015; Fayet et al. 2016). NND identifies for each point in trajectory T_A the closest point in trajectory T_B . NND is then the average of these distances. Calculated in this way, the distance between trajectories is not

symmetric, because $d(T_A, T_B) \neq d(T_B, T_A)$. However, the measure can be made symmetric as $NND(T_A, T_B) = (d(T_A, T_B) + d(T_B, T_A))/2$. It should be noted that this formulation of NND lacks certain features of the other measures described. Specifically, it is not monotonic or subject to temporal continuity constraints.

Simulation study of different similarity measures

We used a series of data simulations to assess the performance of the different similarity measures introduced. Our goal was to simulate different movement trajectories in which the similarity between trajectories was controlled by changing different movement parameters. The similarity measures that are most responsive to changes in these parameters were then judged to be the best performing. To do so we simulated movement trajectories using a multi-variate Ornstein-Uhlenbeck (OU) movement process a commonly used movement model (Blackwell 1997). The OU process is given as:

$$dz(t) = \alpha (\beta - z(t))dt + \sigma dW2(t) \quad (7)$$

where the vector β contains the coordinates of the attraction point and the matrix α contains coefficients that describe the strength of attraction to β . σ controls the noise added to the movement and is sometime referred to as volatility; $dW2(t)$ denotes a Wiener process. We used the OU process to simulate different movement trajectories for comparison by varying the parameters, α , β and σ respectively using the R package `adehabitatLT` (Calenge 2006) in order to create different scenarios in which the similarity between trajectories will vary in a predictable manner. Ideally, similarity measures should respond to changes in these parameters. We chose to vary a range of parameters as a given similarity measure may be sensitive to changes in one parameter but be less responsive to changes in another. We address the following scenarios:

1. Assessing the similarity of movement trajectories that started at point [0,0] and were attracted to point [50,000, 50,000] with attraction strength 0.001, but varying σ between low (250), medium (500) and high (1500) values. When σ is lower we expected similarity of trajectories to be higher as they are both directed towards the same point with less noise.
2. Assessing the similarity of movement trajectories that started at point [0, 0] and were attracted to point [50,000, 50,000] with σ of 500, but varying the strength of the attraction, α , between low (0.0001), medium (0.001) and high (0.01) values. When α is lower we expected the similarity of trajectories to be lower due to weaker attraction to the coordinates of the attraction point.

3. Assessing the similarity movement of trajectories in which σ was 500, α was 0.001 and trajectories began at point [0, 0] but differed in their point of attraction. Specifically, we compared trajectories in which:

- (a) Both trajectories are attracted to point [50,000, 50,000].
- (b) T_A was attracted to point [50,000, 50,000] and T_B was attracted to point [- 50,000, 50,000].
- (c) T_A was attracted to point [50,000, 50,000] and T_B was attracted to point [- 50,000, - 50,000].

We expected similarity of trajectories to be greatest when both trajectories were directed to the same point of interest as in (a) and lowest when trajectories were directed to points of interest in the opposite direction as in (c).

For each set of comparisons 500 trajectories were simulated in total (250 simulations of T_A and 250 simulations of T_B) meaning that our similarity measures were calculated across 250 trajectory pairs. All simulated trajectories had a length of 1000 points. A list of the different movement trajectories simulated and compared is described in Table S2 (see also Fig. S6 and Supplementary Material, S5: R script used to simulate different OU trajectories). We calculated DTW, LCSS, EDR and Fréchet distance in \mathbb{R} (R Core Team 2018) using the *SimilarityMeasures* package (Toohey 2015). We calculated NND using the *RANN* package (Arya et al. 2018). We standardized coordinates prior to calculating similarity for all measures except NND. Standardizing involved centring coordinates using the position of the common departure point and then dividing by the standard deviation of that coordinate. This ensures all trajectories start at the same point [0, 0] and standardized trajectories will be invariant to scaling and shifting (see Giao and Ahn 2016). We did not standardize coordinates when calculating NND in order to mirror the approach used in the previous ecological studies that used this measure (Fayet et al. 2016, 2017; Shoji et al. 2016; Votier et al. 2017). When calculating LCSS and DTW we set δ to *unlimited*. Because trajectory data were standardized prior to analysis we set ϵ at 1.0 for LCSS and EDR which roughly corresponded to points within a standard deviation distance in both x and y coordinates being considered matches. Note that similarity values for LCSS and EDR are influenced by the values of ϵ and δ . Although simulated trajectories were the same length it is generally advisable to convert DTW, LCSS and EDR to ratios to ensure that variation in trajectory length does not overly influence the results. Because all simulated trajectories originate from the same location the start point of each trajectory was identical. Therefore, following Toohey and Duckham (2015), we calculated DTW and EDR ratios by dividing the estimated similarity measure by the length of the larger trajectory minus 1 (minus the common starting point) ($\max(|T_A|, |T_B|) - 1$). For LCSS the ratio was calculated

using the minimum trajectory length ($\min(|T_A|, |T_B|) - 1$). Ratios are calculated in this way because each simulated trajectory starts at the same location. Therefore, the first point in T_A and T_B is a perfect match so LCSS is guaranteed to have at least one match and EDR one element with no gap. DTW will also be scored as 0 for the first points across T_A and T_B . This can alter ratio values in a non-linear way (Toohey and Duckham 2015). Hence, we subtract minus one to account for this.

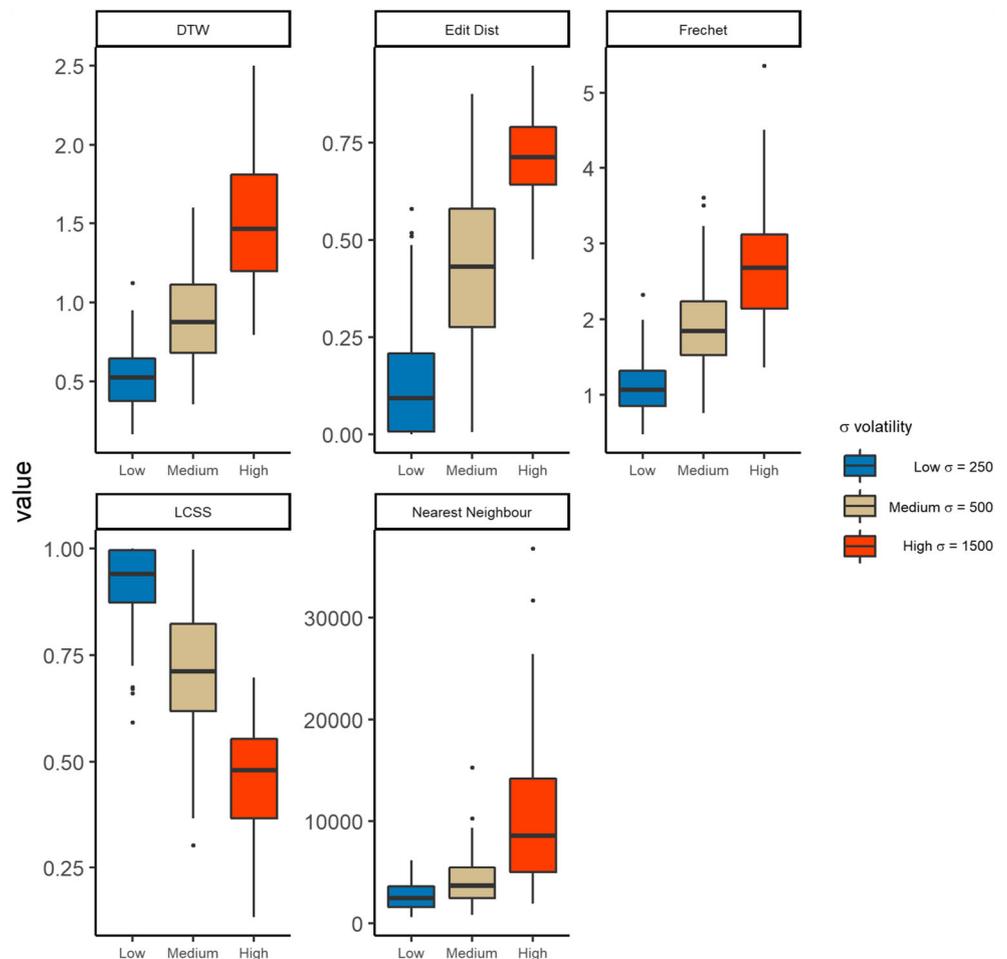
Fréchet distance is less sensitive to variance in trajectory length because it is shape-based rather than a time-series measure, so we did not convert this to a ratio (see also Toohey and Duckham 2015). Similarly, in line with previous ecological studies (Fayet et al. 2016, 2017; Shoji et al. 2016; Votier et al. 2017), we did not convert NND to a ratio.

Simulation results

Scenario 1—varying σ

As the volatility in the OU movement process increased the similarity of movement trajectories decreased. This

Fig. 1 Similarity measures for trajectories simulated using an OU process (Eq. 7) in which noise, σ , varied, while point of attraction, β , and strength of attraction, α , were held constant. In each group the number of simulated trajectories compared was 250. DTW, LCSS and EDR expressed as ratios (see methods). Boxplots show the median, inter-quartile range, whiskers and outliers. The upper whisker extends from the hinge to the largest value no further than $1.5 * IQR$ from the hinge (where IQR is the inter-quartile range, or distance between the first and third quartiles). The lower whisker extends from the hinge to the smallest value at most $1.5 * IQR$ of the hinge

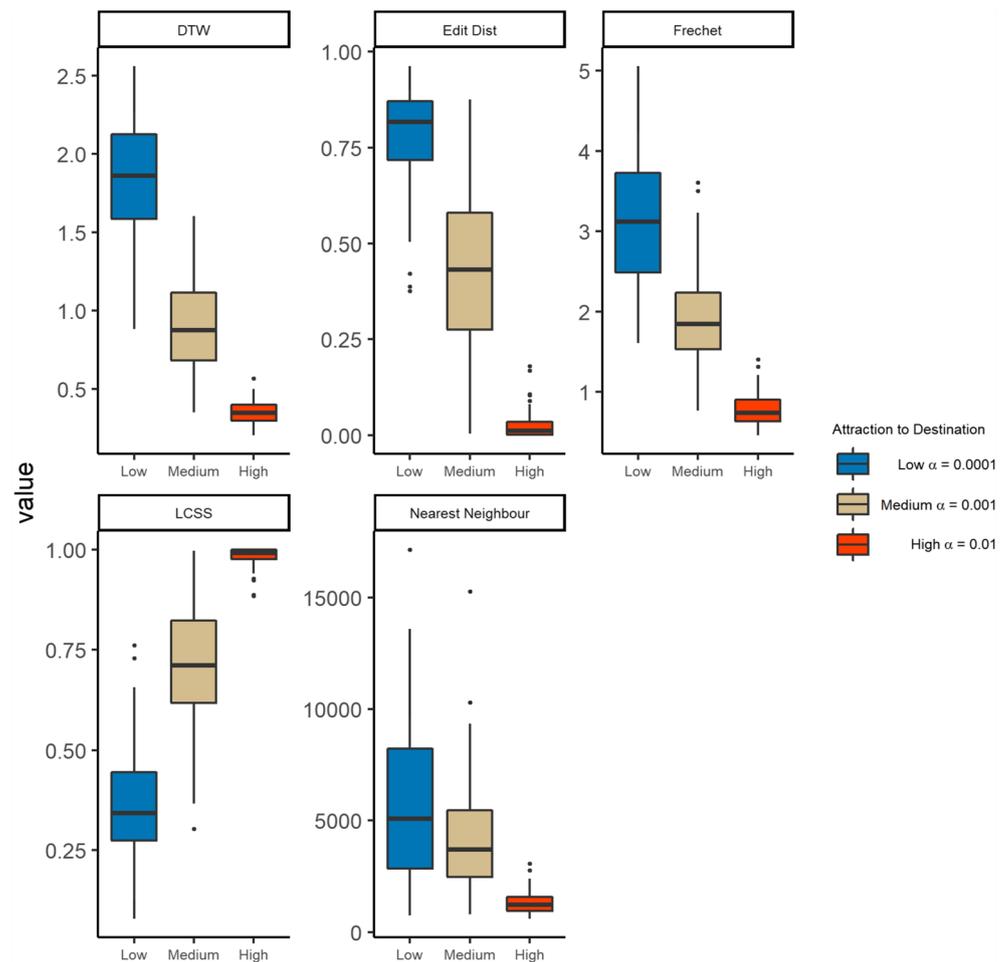


decrease in similarity occurred even though trajectories were directed towards the same point of interest with the same strength of attraction. Across all similarity measures there was a clear separation between comparisons based on low, medium or high values of σ based on inter-quartile ranges of presented box plots (Fig. 1) suggesting all measures were responsive to changes in movement volatility.

Scenario 2—varying α

As the strength of attraction, α , increased, the similarity of compared movement trajectories increased as expected. When α was set at high values the spread of similarity values was also narrower across all similarity measures tested. In general, the similarity measures tested were able to distinguish between simulations when α values were low, medium or high (Fig. 2). The exception was NND in which calculated values showed a high degree of overlap when α was low or medium suggesting NND was less sensitive to changes in α .

Fig. 2 Similarity measures for trajectories simulated using an OU process (Eq. 7) in which α varied, while β and σ were held constant. Boxplots show the median, inter-quartile range, whiskers and outliers



Scenario 3—varying point of attraction

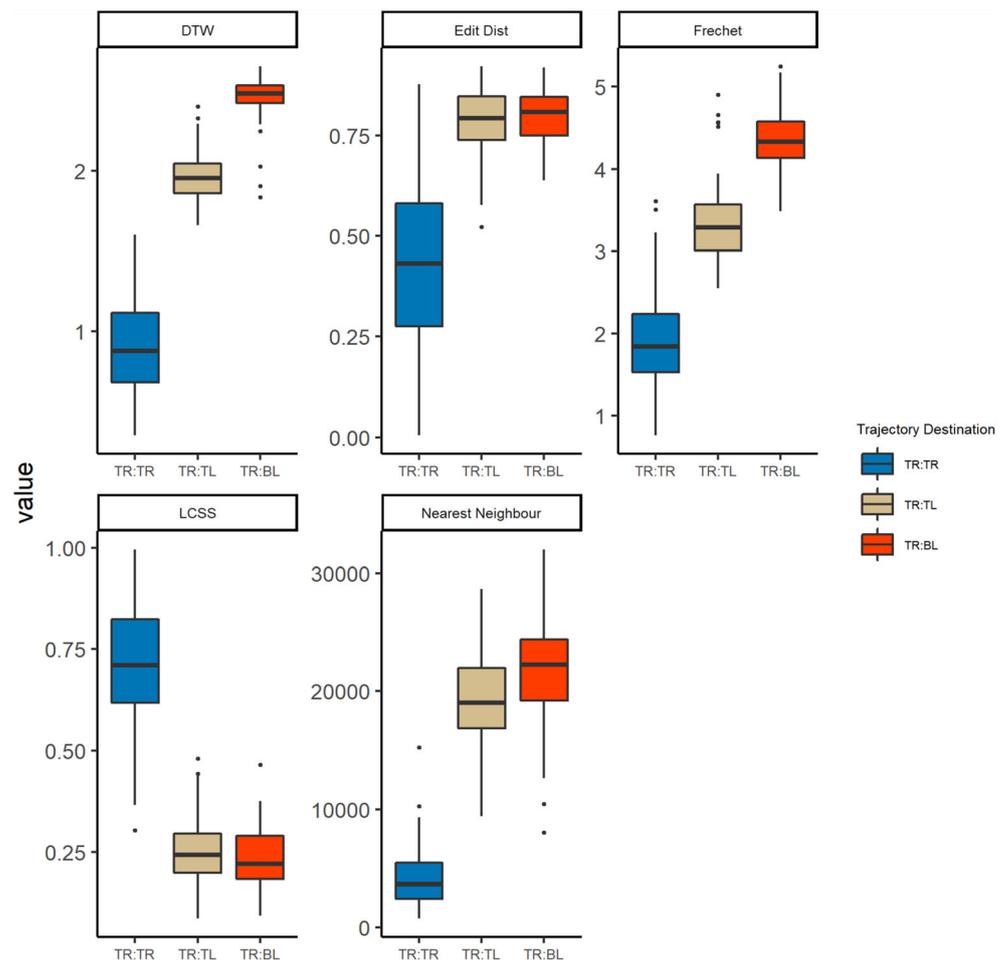
When both trajectories were directed towards the same point of attraction the similarity of trajectories was highest across similarity measures (Fig. 3). In addition, both DTW and Fréchet measures distinguished between simulations in which trajectory T_B was directed towards point $[-50,000, 50,000]$ and those in which trajectory T_B was directed towards $[-50,000, -50,000]$. In contrast, LCSS, EDR and NND did not distinguish between these two scenarios. In the case of EDR and LCSS because values are bounded between 0 and 1 may make discerning similar trajectories more difficult, although adopting a different value for ϵ may improve performance. The reasons for the poor performance of NDD are less clear but may be related to the lack of monotonicity and continuity constraints when using this metric which allows matches between points that would not be allowed when using the other methods where the ordering of points in time is taken into consideration. For example, if we assume that points a_{10} and b_5 of trajectories T_A and T_B are matched then monotonicity

and continuity constraints would mean that subsequently point a_{11} could only be matched with points $b_{\geq 5}$. However, if using NND, matches between a_{11} and $b_{<5}$ remain possible.

Case study: estimating the similarity between northern gannet trajectories

In the following case study, we provide an example of how different trajectory similarity measures can be applied to an ecological dataset. First, we calculate each of the five different similarity measures listed above and examine the correlation between them. Because the similarity measures introduced are related, results should be correlated. Previous empirical studies support this assertion (e.g. Toohey and Duckham 2015) but were based upon traffic data where vehicles move along a road network. This limits the potential routes an object can travel and contrasts with animal movement data which is typically not as constrained. Secondly, we use DTW to cluster trajectories within an individual to demonstrate how such measures could be used to examine trajectory similarity over

Fig. 3 Similarity measures for trajectories simulated using an OU process (Eq. 7) in which β varied, while α and σ were held constant. TR:TR indicates both trajectories were attracted towards point [50,000, 50,000] in the top right. TR:TL indicates trajectory A was attracted to point [50,000, 50,000] and trajectory B was attracted to point [-50,000, 50,000] in the top left. TR:BL indicates trajectory A was attracted to point [50,000, 50,000] and trajectory B was attracted to point [-50,000, -50,000] in the bottom left. Boxplots show the median, inter-quartile range, whiskers and outliers



time. Thirdly, we use DTW to cluster trajectories across all individuals tracked during one year of study to identify distinct movement clusters and highlight important destinations. While such measures have been used to cluster movement trajectories in the past (e.g. McClintock and Michelot 2017) the example provided here is one of the first large-scale (> 150 trajectories) examples of trajectory clustering in ecology.

Study species

Northern gannets (hereafter gannets) are medium-ranging pelagic seabirds. Whilst breeding, they act as central-place foragers, and are capable of travelling > 1500 km on a single round-trip to obtain food for themselves and their offspring (Hamer et al. 2009). Here, we analyse GPS tracking data collected from gannets breeding at Bass Rock, UK (56.0773° N, 2.6408°W, Fig. S7) from 2010 to 2012. Full details of field procedures are provided by Wakefield et al. (2015). In brief, adults attending young chicks were caught at the nest, equipped with a tail-mounted Igotu GT-200 or GT-600 GPS logger (Mobile Action Technology Inc., Taipei, Taiwan, 37 g/± 2% body mass), programmed to record locations every 2 min.

Birds were recaptured after 1–3 weeks, the loggers were removed, and the data downloaded. It was not possible to record data blind because our study involved focal animals in the field. Here, we treat a foraging trip as a complete trajectory, starting and ending at the colony. Individuals were tracked for multiple trips within years and a subset of birds were tracked across multiple years (Wakefield et al. 2015). Previous work has shown that individual gannets exhibit a relatively high degree of foraging site and route fidelity and often follow highly directed paths from their colonies when departing on foraging trips (Wakefield et al. 2015; Votier et al. 2017).

Calculation of trajectory similarity measures

In order to examine the correlation between the different trajectory similarity measures, we randomly selected a subset of 50 pairs of trajectories from gannets tracked in 2012 (see also Toohey and Duckham 2015) and for each pair of trajectories we calculated DTW, LCSS, EDR, Fréchet distance and NND between them. A sample size of 50 was chosen as this gave us 80% power of detecting a correlation of 0.4 with a two-tailed type I error significance level of 0.05 (our a priori expectation

was that measures should be strongly correlated based on Toohey and Duckham (2015)). As we were satisfied with the statistical power, we did not increase sample size further in order to avoid additional increases in computation time associated with the calculation of Fréchet distance (Table 1, Table S3).

Prior to analysis, we projected trajectories in Lambert Azimuthal equal area (LAEA) coordinates. Latency in the GPS loggers' acquisition of GPS signals resulted in a small amount of variability in the interval between GPS fixes. We therefore re-sampled trajectories using the *adehabitatLT* R package (Calenge 2006) to ensure regular time steps every 120 s using linear interpolation. Regular time steps are not required for the similarity measures discussed but made it easier for us to sample different time-intervals when calculating Fréchet distance for which computation time can be considerable. Ultimately, when calculating Fréchet distance, we re-sampled trajectories every 20 min to reduce computation time. We standardized coordinates prior to calculating similarity for all measures except NND to account for differences in the spatial scale of different trajectories (Toohey and Duckham 2015). Coordinates were standardized by subtracting colony coordinates and dividing by their standard deviation ensuring that all trajectories start at the same location. Trajectory length varied considerably (trip duration: $\bar{x} = 22.58$ h, $\sigma = 13.68$, range = 2.18–95.11 h; total distance travelled per trip: $\bar{x} = 508.0$ km, $\sigma = 283.7$, range = 51.3–1336.4 km) so when calculating LCSS and DTW we set δ to *unlimited as in* Toohey and Duckham (2015). For LCSS and edit distance, we set ϵ at 5 km. This value was set to reflect the fine-scale foraging behaviour of gannets identified using first passage time analysis during a previous study that found foraging bouts occurring at a scale of 1.5 km nested within 9 km search areas (Hamer et al. 2009). To ensure that variation in trajectory length did not dominate the results, we converted DTW, LCSS and EDR similarity values to ratios. We calculated DTW and EDR ratios by dividing the estimated similarity measure by the length of the larger trajectory minus 2 ($\max(|T_A|, |T_B|) - 2$). For LCSS the ratio was calculated using the minimum trajectory length ($\min(|T_A|, |T_B|) - 2$). Ratios are calculated using a value of minus two here because gannets are central place foragers and therefore the first and last points of any trajectories being compared will be identical.

Using DTW to cluster trajectories

Our simulation studies highlighted that DTW performed well when trying to distinguish different models of movement. In addition, while the other similarity measures we describe can also be used to create distance matrices, they currently lack dedicated R packages that combine calculation of similarity measures and downstream clustering for multi-variate time-series such as movement trajectories (Table 1). Therefore, we created a distance matrix

based upon DTW values to cluster trajectories using the R package *dtwclust* (Sarda-Espinosa 2017). Here, we present the results of two different clustering analyses using DTW: (1) Clustering all trips made by one bird tracked in each study year (Individual: 1459907, 3 foraging trips in 2010, 5 trips in 2011, 5 trips in 2012); and (2) clustering all foraging trips recorded within 2012 (166 foraging trips from 33 different individuals). We used hierarchical clustering, with the average linkage method, implemented by the *hclust()* function (R Core Team 2018, see also Tables S3, S4). Hierarchical clustering requires that the number of clusters be selected based upon the clustering output. When clustering data from all trips recorded in 2012, we varied the number of clusters from 2 up to 20 and used silhouette analysis to determine the optimal number of clusters (see Supplementary Material, S7 Silhouette Coefficients and Clustering). Silhouette analysis suggested that the optimal number of clusters was either 4 or 11. For simplicity, we report results based on 4 clusters. When analysing trips from bird 1459907 across multiple years, silhouette analysis suggested the optimal number of clusters was 2.

Results

Comparison of trajectory similarity measures

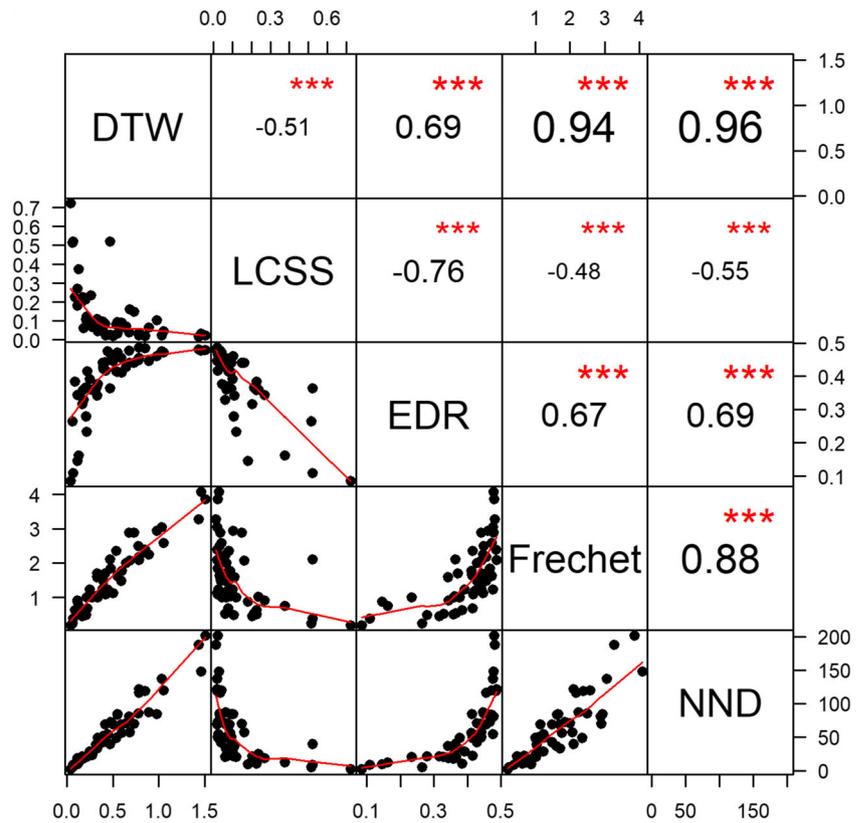
There were significant correlations between all the trajectory similarity measures trialled (Fig. 4). DTW, Fréchet distance and NND showed strong correlations even though DTW and NND are based on point matching, whereas Fréchet distance is shape-based. LCSS was negatively correlated with the other distance measures as higher LCSS values denote greater similarity, whereas for the other measures higher values denote lower similarity. Relationships between distance measures were not always linear. In part, this may arise because DTW, Fréchet distance and NND are unbounded, whereas LCSS and EDR are bounded between 0 and 1. Selected trajectories that show high, moderate and low similarity are displayed for visualisation in the supplementary material (Fig. S8)

Using similarity measures to cluster trajectories

1. Clustering all trips within an individual bird

Our analysis of foraging trips made by bird 1459907 in the 2010, 2011 and 2012 breeding seasons suggest that foraging trips grouped into two clusters. All three trajectories recorded in 2010 were placed within cluster

Fig. 4 Correlations amongst measures of trajectory similarity/distance across 50 randomly selected pairs of gannet foraging trips. Note DTW, EDR and LCSS refer to the ratios here rather than absolute values. Scatterplots shown in bottom left triangle. Spearman's rank correlations in upper right triangle, *** $p < 0.001$

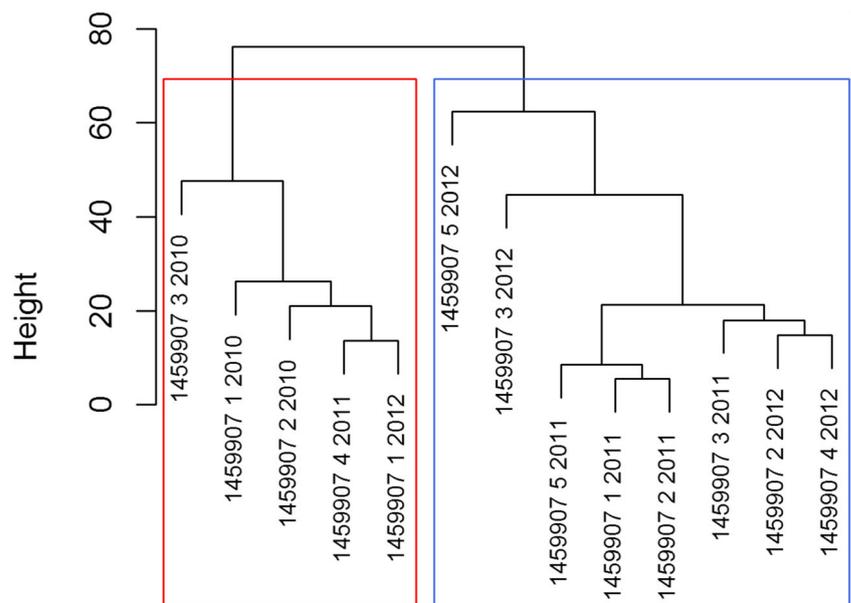


1 (Fig. 5). In contrast, most trajectories recorded in 2011 and 2012 were placed in cluster 2, with just one trajectory from each of these years being placed in cluster 1. The clusters appeared to split trajectories based upon the direction from the colony in which the bird foraged: Cluster 1 comprised trips to the south-east and cluster 2 trips to the east (Fig. 6).

2. Clustering foraging trips recorded within 2012

Hierarchical clustering split the foraging trips made by individuals in 2012 into four clusters (Fig. S9), roughly corresponding to differences in the directions these birds travelled to forage from the colony. For ease of visualisation, we

Fig. 5 Hierarchical clustering of GPS trajectories of gannet 1459907 during the 2010, 2011 and 2012 breeding seasons. Coloured rectangles denote cluster identity. Dendrogram labels report bird identity, trip number within a given year and the year in which the trajectory was recorded



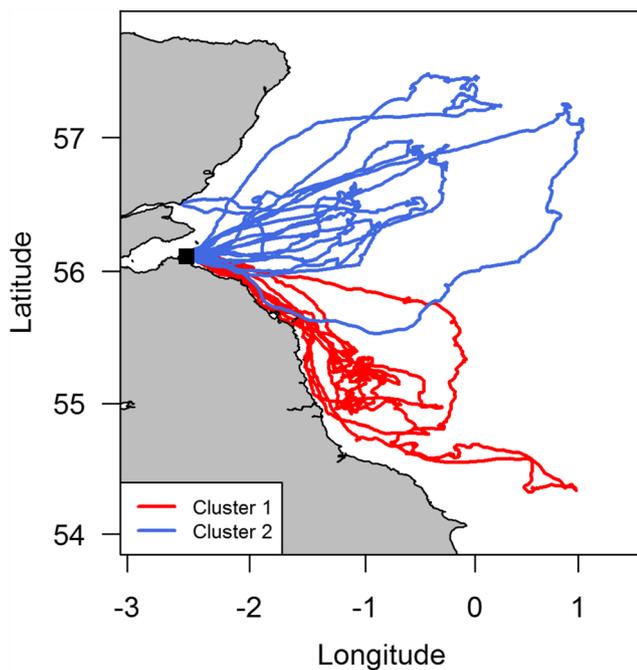


Fig. 6 Foraging trajectories recorded for bird 1459907 coloured by cluster. Trajectories in cluster one were typically to the South-East of Bass Rock, whereas trajectories in cluster two were generally to the East. Location of Bass Rock shown as a black square

summarise the distribution of space use in each cluster with utilisation distributions (UDs) using biased random bridges (Supplementary material, [S10](#) Calculating Utilisation Distributions). Cluster 1 was the smallest, containing 15 trajectories. These characteristically travelled to the east of the colony and UD suggested a key destination for such trajectories was the Devil's Hole, a group of submarine trenches (Fig. [7a](#)). Cluster 2 contained 28 trajectories, which typically went far to the northeast of the colony, with usage concentrated around the Witch Ground basin (Fig. [7b](#); see also Grecian et al. [2018](#)). Cluster 2 birds also used areas around the Montrose and Scalp banks—shallow areas which give rise to tidal mixing fronts. Cluster 3 was the largest cluster, containing 103 trips. Unlike the other clusters, no one pattern clearly stereotyped this cluster. However, trajectories in this cluster tended to be shorter and usage was concentrated, in part, on the Firth of Forth sand banks complex, relatively close to Bass Rock, largely inshore of the Montrose and Scalp banks (Fig. [7c](#)). Increasing the number of clusters to 11 (as identified by silhouette values) largely resulted in a greater partitioning of trajectories within this cluster. Finally, cluster 4 contained 20 trajectories. Trajectories in cluster 4 typically followed routes to the southeast of Bass Rock, with usage concentrated over the Farn Deeps (Fig. [7d](#)). Across individuals with more than one recorded foraging trip, 4/31 (13%) birds performed trips that were classified entirely within only one cluster (all cluster 3), 20/31 birds (65%) performed trips that spanned across two different clusters, 5/31 (16%) birds performed trips

that spanned across three different clusters and 2/31 (6%) birds performed trips that spanned each of the four clusters identified.

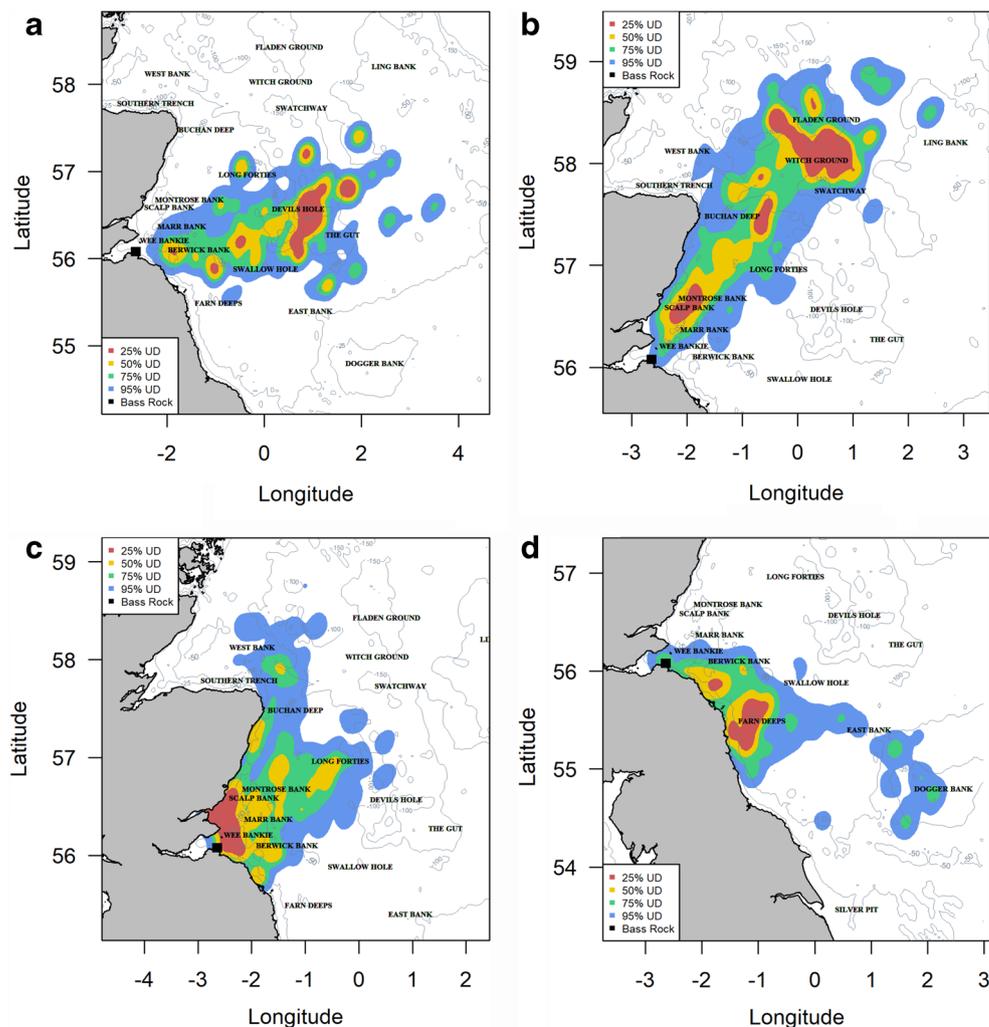
Discussion

As the volume of movement data that ecological studies collect increases, there is a growing need to develop techniques to process and analyses such data efficiently (Demšar et al. [2015](#)). For instance, an increasing number of studies seek to assess route fidelity and route similarity, within and between individuals as well as across different species (Hansen et al. [2013](#); Fayet et al. [2016](#); Votier et al. [2017](#)). Therefore, effective measures of the similarity between different movement trajectories are crucial. Here, we provide a summary and a comparison of some of the more commonly encountered similarity measures. Such similarity measures have a relatively long history outside of ecology (Ranacher and Tzavella [2014](#)) but have only recently begun to be applied in animal movement studies (Demšar et al. [2015](#); McClintock and Michelot [2017](#)). We anticipate that they will be used much more widely in this context in the future. Moreover, clustering trajectories using similarity-based distance matrices makes it possible to detect patterns of similarity amongst animal trajectories.

Overall, DTW and Fréchet distance performed best across our simulations in terms of being able to distinguish between the different parameter combinations we trialled. EDR and LCSS were also able to distinguish between all situations when we varied σ or α , but not the point of attraction. It should be borne in mind that the exact behaviour of LCSS and EDR could be altered if we set the maximum distance threshold for points to be considered equivalent at a different value. This highlights the importance of setting appropriate values for the distance threshold when using LCSS or EDR, but also the difficulty of choosing such a value without trialling different values (Supplementary Material, [S11](#). Trialling different values of ϵ for LCSS and EDR). NND is probably the most common measure of trajectory similarity used in ecology, but it did not distinguish situations where the strength of attraction varied from low to medium or when the point of attraction was shifted. Thus, it appeared to be the least sensitive measure of the five trialled to changes in the underlying movement process we simulated.

One challenging aspect of simulating movement trajectories in the manner done here lies in specifying a movement model and understanding how its results can be generalized across studies. The OU process is a relatively standard model of movement, but there are a variety of other movement models which may better describe the movement behaviour of a given species (e.g. Fleming et al. [2017](#); Gurarie et al. [2017](#)). Therefore, after first identifying an appropriate movement model, researchers could then simulate trajectories to

Fig. 7 Utilisation distributions for trajectories within each cluster identified. The names of marine features and fishing grounds within the study area are also displayed along with depth contours. **a** cluster 1, Devil's Hole cluster; **b** cluster 2, Witch Ground cluster; **c** cluster 3, Firth of Forth sand banks cluster; **d** cluster 4, Farn Deeps cluster



investigate the performance of the different similarity measures using movement models based upon their own data. Such simulations may also allow one to test different threshold distance values if using LCSS or EDR. The simulations we present here are also loosely based upon the behaviour of a central place forager (reflecting our case study) in as much as all trajectories start at the same location. However, not all movement studies are focussed upon central place foraging and this may be another aspect that affects the relative performance of the different similarity measures.

When applied to a real dataset, the five similarity measures we examined were strongly correlated, in line with results from other studies (Toohey and Duckham 2015). However, there were also specific instances in which such measures differed. For example, the unbounded nature of DTW and Fréchet distance allowed them to emphasize differences between trajectories that were less apparent from relative LCSS and EDR scores, which are bounded between 0 and 1. In our case study the correlation between DTW and NND was particularly strong. A similar pattern was observed in our simulation study, but it is worth noting that the strength of this

correlation varied across scenarios. For example, when both simulated trajectories were directed towards a common destination NND and DTW were strongly correlated ($r = 0.83$, 95% CI (0.77–0.87)). However, this correlation was slightly weaker when simulated trajectories were directed towards different destinations ($r = 0.60$, 95% CI (0.46–0.71)). Thus, one should not assume that these two measures will always be highly correlated.

Providing guidance about which is the ‘best’ similarity measure is difficult as each have their own strengths and weaknesses (Table 1). That said, all measures covered here appear to perform better on high sample rate trajectories and may therefore be less suitable for lower resolution tracking data. For example, Wang et al. (2013) compared the same trajectory sampled at both the original and lower resolution rates and showed that as sample rate was reduced the similarity between the original and the resampled trajectory declined. Defining what is a high or low sampling rate is difficult and likely to depend on the scale at which an animal moves. As a rule-of-thumb, Zheng et al. (2012) suggest that sampling rates greater than > 2 min represent a low sampling rate though this

was based upon traffic data in which data points had to be matched to a road network. More broadly, to examine the influence of sampling rate one could sample high resolution data at lower resolutions as in Wang et al. (2013). If such data is lacking, then movement simulation studies in which sample rate varies could be used to gauge the effect of sampling rate when assessing similarity. In addition, irregular sampling rates between trajectories can also be problematic, making it difficult to compare between densely and sparsely sampled trajectories (Li et al. 2018). Such an issue can also occur within the same device if sampling rate is scheduled such that it changes during deployment and this should be borne in mind when comparing trajectories.

The close association between NND and the other similarity measures examined in our case study suggest they show close agreement when assessing the similarity of gannet trajectories. However, unlike the other measures, NND is not constrained to be monotonic or contiguous which are desirable features when comparing time-series. Therefore, we would recommend using other or additional similarity measures when comparing trajectories. Of the other similarity measures, DTW has generally been proven to work well for time-series classification purposes (Wang et al. 2017) and has the advantage that it is incorporated within multiple R packages that provide researchers with a wide range of analytical resources for downstream clustering. DTW was also quicker to calculate than LCSS, EDR or Fréchet distance. Thus, comparing trajectories using DTW may often be a useful and relatively easy first step. However, LCSS and EDR are particularly useful when it is suspected that trajectory data are noisy, as these methods are more robust to noise. Here, we used biological insights from a previous study to set ϵ at a spatial scale reflecting that at which GPS-tracked gannets typically forage, but one could set ϵ to reflect known location error from ones tracking device for example. For DTW, LCSS and EDR one must also set a δ value and while many studies adopt an unconstrained approach as we do here adjusting this parameter can sometimes improve clustering performance (Dau et al. 2018). We also note that the similarity measures covered here represent only a subset of available trajectory similarity measures (Ranacher and Tzavella 2014) and that rather than having to choose between similarity measures it may be possible to use them as an ensemble for machine learning methods of time-series classification purposes (Lines and Bagnall 2015). In addition, it would also be possible to cluster trajectories by combining variables such as trajectory duration, distance travelled, longitude and latitude of furthest point from the colony etc. although some information of the entire trajectory would be lost as any analysis would be based on a summarised version of a movement trajectory rather than the trajectory itself. For example, individuals may visit the same end location but travel there via different routes. Consequently, research focus would shift from comparing the similarity of detailed

movement paths to broader movement syndromes such as territorial versus nomadic (Abrahms et al. 2017).

Similarity measures provide a means of quantitatively assessing the distance between two trajectories. Such assessment might be all that is required if the aim is to compare specific individuals (e.g. members of a breeding pair). However, we imagine researchers will often wish to compare multiple trajectories to investigate trajectory similarity both within- and between-individuals. If so, the similarity measures introduced here could be used as response variables in statistical models. For example, Votier et al. (2017) previously used NND between trajectories as a response variable in linear models to investigate route fidelity across different groups of individuals, but the other measures introduced here could be used for this purpose. Likewise, measures of trajectory similarity can be used as inputs to randomization tests to examine whether similarity differs between groups or individuals just as measures of overlap are currently used in such randomization tests when comparing utilisation distributions across different groups (e.g. Wakefield et al. 2015). In addition, we investigated space use of an individual over time by calculating DTW similarity measures for all binary comparisons of foraging trips recorded for a focal individual over three consecutive breeding seasons. The creation of a dendrogram from the resulting distance matrix, and plotting of the resulting clusters, allowed us to easily visualise and compare patterns of space use over time (Kranstauber et al. 2017). In our example, clustering split foraging trips for individual 1459907 into two main groups based primarily on the direction the individual departed the colony.

We also demonstrate how clustering can be used to assess the similarity of movement between individuals across all foraging trips recorded during a breeding season. As before, clustering tended to split foraging trips partly on the direction birds travelled from the colony and the resulting clusters appeared to be focussed towards well-known bathymetric features. Overall, clustering provided a useful tool to visualise patterns in our movement data (Andrienko and Andrienko 2013) and identify potential foraging strategies related to geographic locations. Such tools are particularly important when working on large tracking datasets as discerning patterns from raw data becomes more difficult. For example, clusters 2 and 3 identify patterns in the data that the human eye alone may find difficult to differentiate. In addition, the potential importance of the offshore areas identified in clusters 1, 2 and 4 would be less apparent if we estimated a single population-level UD across all tracked birds. However, these clusters align with features that are visited by gannets in other years (Grecian et al. 2018) and overlap with designated Marine Protected Areas (MPAs) such as the Central Fladen MPA or the Farnes East MPA. The results from trajectory clustering can also be used as the foundation for more in-depth analysis of foraging behaviour as they permit identification of destinations of

interest (McClintock and Michelot 2017). From a conservation perspective the environmental conditions and anthropogenic impacts individuals will experience are likely to vary across clusters. For example, individuals that forage predominantly in a certain cluster may be more affected by localized risk factors, such as wind farms or oil spills, than other individuals from the same colony who utilise a different foraging area (Bodey et al. 2018), with potential downstream effects on population dynamics (Tuck et al. 2015).

In addition, rather than focussing upon whole trajectories it is also possible to compute the similarity and cluster movement sub-trajectories. In human geography, such approaches are often used to discover areas of high traffic density (Liu et al. 2014; Gui et al. 2016). One advantage to such an approach is that we can identify important, highly travelled routes used by an animal population (Whittington et al. 2005; Meier et al. 2015). Such movement corridors or bottlenecks may be particularly sensitive to human activity, such as windfarms or other human infrastructure (Barrios and Rodriguez 2004). Recently, Scharf et al. (2018) reported that habitat suitability models may perform poorly when trying to identify animal movement corridors and suggested instead that movement data be used directly to identify corridors and sub-trajectory clustering is one means of achieving this. Although outside the scope of this work, there are a variety of algorithms available for sub-trajectory clustering (e.g. Lee et al. 2007; Eerland et al. 2017)

Throughout, we have focussed upon clustering movement data based purely upon the geographic locations, which is typical of most trajectory similarity studies. However, trajectories can be compared using a variety of movement parameters other than, or in conjunction with, geographic location. For example, we could estimate the similarity between trajectories using geographic position and the time-stamp at which positions were recorded (absolute time). Similar trajectories would then have to occupy similar positions at similar times, which is sometimes termed movement coincidence (Dodge et al. 2008). Movement in environmental space can also be considered. For example, Buchin et al. (2014) computed the similarity between trajectories of foraging Galapagos albatrosses (*Phoebastria irrorata*) using wind speed, as well as positional location. Dodge et al. (2012) demonstrate how trajectories can be clustered using movement parameters such as speed, turning angle and azimuth to identify concurrence in movement parameters over time, without including positional information. In this case trajectories will be similar if objects move in a similar way over time even if they are not found at the same location. This could be an effective way of identifying the same stereotyped behaviours, such as search patterns and prey capture techniques, or territorial/reproductive displays, which can be expressed in different locations. In addition, the similarity

measures described here could be applied to other time-series (e.g. Marques et al. 2018). For example, temperature-depth recorders (TDR) are often used to record the behaviour of diving animals (seabirds, pinnipeds, etc.). The techniques we describe could be applied to depth time-series data from these devices to quantify individual consistency in dive behaviour or objectively identify typical dive profiles.

Conclusions

Many useful tools to visualise and analyse movement data efficiently have been developed across the information sciences. Thus, as the number of ecological studies concerned with estimating route similarity increases, it is natural to consider the broader range of trajectory similarity measures that are available (Ranacher and Tzavella 2014). We provide an overview of five commonly used trajectory similarity measures (DTW, LCSS, EDR, Fréchet distance and NND) and apply them to both simulated data and an ecological dataset. Simulations suggested that DTW and Fréchet distance performed best on the examples we generated. In contrast, NND was generally the worst performing measure and lacks some of the stricter time-series constraints that characterize the other measures. Therefore, although NND has been most widely used in ecology we recommend considering additional similarity measures including those discussed here. These measures can be calculated via existing software and have the advantage that their mathematical properties and performance have been well studied (Ratanamahatana and Keogh 2004; Wang et al. 2013), though not on ecological data. We also demonstrate how distance matrices can be created to cluster GPS trajectories both within and amongst individuals. Importantly, these measures are not limited solely to comparing spatial positions but can also utilise parameters such as speed and environmental conditions. Thus, these similarity measures represent both a useful tool for ecologists in an area of growing interest, and an introduction into the wider world of movement analysis beyond ecology (Demšar et al. 2015; Miller et al. 2019). As new technology and analysis techniques proliferate across ecology and the information sciences, closer ties between these fields promises further innovative analysis of movement data.

Acknowledgements We thank Sir Hew Hamilton-Dalrymple for access to the Bass Rock and Maggie Sheddan and the Scottish Seabird Centre for logistical support. Fieldwork was carried out with approval from the British Trust for Ornithology and Scottish Natural Heritage. Additional thanks go to Robert J. M. Cleasby and Tony Inchpractice for advice on earlier drafts. We thank Rocio Joo and one anonymous reviewer for comments which improved the manuscript.

Data availability statement All tracking data used in this manuscript are available at <http://seabirdtracking.org>.

Funding information Gannet data was collected as part of a Natural Environment Research Council (Standard Grant NE/H007466/1) to KCH, SB, and SCV and Independent Research Fellowship NE/M017990/1) to EW.

Compliance with ethical standards

Ethical approval Fieldwork was carried out with approval from the British Trust for Ornithology and Scottish Natural Heritage, including permits and special methods endorsements for attachment of rings and data loggers. All work followed international, national and institutional guidelines on the use of animals, including ethical approval from the UK Home Office (licence PPL40/3408).

Conflict of interests The authors declare they have no conflict of interest

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Aarts G, MacKenzie M, McConnell B, Fedak M, Matthiopoulos J (2009) Estimating space-use and habitat preference from wildlife telemetry data. *Ecography* 31:140–160
- Abrahms B, Seidel DP, Dougherty E, Hazen EL, Bograd SJ, Wilson AM, Weldon McNutt J, Costa DP, Blake S, Brashares JS, Getz WM (2017) Suite of simple metrics reveals common movement syndromes across vertebrate taxa. *Mov Ecol* 5:12
- Abrahms B, Hazen EL, Bograd SJ, Brashares JS, Robinson PW, Scales KL, Crocker DE, Costa DP (2018) Climate mediates the success of migration strategies in a marine predator. *Ecol Lett* 21:63–71
- Alt H (2009) The computational geometry of comparing shapes. In: Albers S, Alt H, Näher S (eds) *Efficient Algorithms*. Springer, Berlin, pp 235–248
- Andrienko N, Andrienko G (2013) Visual analytics of movement: An overview of methods, tools and procedures. *Inform Visual* 12:3–24
- Aronov B, Har-Peled S, Knauer C, Wang Y, Wenk C (2006) Fréchet distance for curves, revisited. In: *European Symposium on Algorithms*. Springer, Berlin, pp 52–63
- Arya S, Mount D, Kemp SE, Jefferis G (2018) RANN: fast nearest neighbour search (Wraps ANN Library) Using L2. R package version 2:6. <https://CRAN.R-project.org/package=RANN>
- Barrios L, Rodríguez A (2004) Behavioural and environmental correlates of soaring-bird mortality at on-shore wind turbines. *J Appl Ecol* 41: 72–81
- Bateman AW, Lewis MA, Gall G, Manser MB, Clutton-Brock TH (2015) Territoriality and home-range dynamics in meerkats, *Suricata suricatta*: a mechanistic modelling approach. *J Anim Ecol* 84:260–271
- Besse P, Guillouet B, Loubes JM, François R (2015) Review and perspective for distance based trajectory clustering. arXiv 1508.04904
- Blackwell PG (1997) Random diffusion models for animal movement. *Ecol Model* 100:87–102
- Bodey TW, Cleasby IR, Votier SC, Hamer KC, Newton J, Patrick SC, Wakefield ED, Bearhop S (2018) Frequency and consequences of individual dietary specialisation in a wide-ranging marine predator, the northern gannet. *Mar Ecol Prog Ser* 604:251–262
- Buchin K, Buchin M, Gudmundsson J (2010) Constrained free space diagrams: a tool for trajectory analysis. *Int J Geogr Inf Sci* 24: 1101–1125
- Buchin K, Buchin M, Van Kreveld M, Luo J (2011a) Finding long and similar parts of trajectories. *Comput Geom* 44:465–476
- Buchin K, Buchin M, Gudmundsson J, Löffler M, Luo J (2011b) Detecting commuting patterns by clustering subtrajectories. *Int J Comput Geom Appl* 21:253–282
- Buchin M, Dodge S, Speckmann B (2014) Similarity of trajectories taking into account geographic context. *J Spat Inf Sci* 15:101–124
- Calenge C (2006) The package “adehabitat” for the R software: a tool for the analysis of space and habitat use by animals. *Ecol Model* 197: 516–519
- Chen L, Ng R (2004) On the marriage of lp-norms and edit distance. In: *Proceedings of the Thirtieth international conference on Very large data bases-Volume 30. VLDB Endowment*, pp 792–803
- Chen L, Özsu MT, Oria V (2005) Robust and fast similarity search for moving object trajectories. In: *Proceedings of the 2005 ACM SIGMOD international conference on Management of data*. ACM, pp 491–502
- Clapp JG, Beck JL (2015) Evaluating distributional shifts in home range estimates. *Ecol Evol* 5:3869–3878
- Cleasby IR, Wakefield ED, Bodey TW, Davies RD, Patrick SC, Newton J, Votier SC, Bearhop S, Hamer KC (2015) Sexual segregation in a wide-ranging marine predator is a consequence of habitat selection. *Mar Ecol Prog Ser* 518:1–12
- Dau HA, Silva DF, Petitjean F, Forestier G, Bagnall A, Mueen A, Keogh E (2018) Optimizing dynamic time warping’s window width for time series data mining applications. *Data Min Knowl Disc* 32: 1074–1120
- Demšar U, Buchin K, Cagnacci F, Safi K, Speckmann B, Van de Weghe N, Weiskopf D, Weibel R (2015) Analysis and visualisation of movement: an interdisciplinary review. *Mov Ecol* 3:5
- Ding H, Trajcevski G, Scheuermann P, Wang X, Keogh E (2008) Querying and mining of time series data: experimental comparison of representations and distance measures. *Proc VLDB Endowment* 1:1542–1552
- Dodge S, Weibel R, Lautenschütz AK (2008) Towards a taxonomy of movement patterns. *Inform Visual* 7:240–252
- Dodge S, Laube P, Weibel R (2012) Movement similarity assessment using symbolic representation of trajectories. *Int J Geogr Inf Sci* 26:1563–1588
- Eerland W, Box S, Fangohr H, Söbester A (2017) Teetool—a probabilistic trajectory analysis tool. *J Open Res Softw* 5
- Eiter T, Mannila H. (1994) Computing discrete Fréchet distance. *Tech. Report CD-TR 94/64*. Information Systems Department, Technical University of Vienna, Vienna
- Faloutsos C, Jagadish HV, Mendelzon AO, Milo T (1997) A signature technique for similarity-based queries. In: *Proceedings Compression and Complexity of SEQUENCES*. IEEE. IEEE Computer Society, Los Alamitos, pp 2–20
- Fayet AL, Freeman R, Shoji A, Boyle D, Kirk HL, Dean BJ, Perrins CM, Guilford T (2016) Drivers and fitness consequences of dispersive migration in a pelagic seabird. *Behav Ecol* 27:1061–1072
- Fayet AL, Shoji A, Freeman R, Perrins CM, Guilford T (2017) Within-pair similarity in migration route and female winter foraging effort predict pair breeding performance in a monogamous seabird. *Mar Ecol Prog Ser* 569:243–252
- Fieberg J, Kochanny CO (2005) Quantifying home-range overlap: the importance of the utilization distribution. *J Wildlife Manage* 69: 1346–1359

- Fleming CH, Sheldon D, Gurarie E, Fagan WF, LaPoint S, Calabrese JM (2017) Kálmán filters for continuous-time movement models. *Ecol Inform* 40:8–21
- Fréchet MM (1906) Sur quelques points du calcul fonctionnel. *Rend Circ Mat Palermo* 22(1884-1940):1–72
- Freeman R, Mann R, Guilford T, Biro D (2010) Group decisions and individual differences: route fidelity predicts flight leadership in homing pigeons (*Columba livia*). *Biol Lett* 7:63–66
- Giao BC, Anh DT (2016) Similarity search for numerous patterns over multiple time series streams under dynamic time warping which supports data normalization. *Vietnam J Comp Sci* 3:181–196
- Grecian WJ, Lane JV, Michelot T, Wade HM, Hamer KC (2018) Understanding the ontogeny of foraging behaviour: insights from combining marine predator bio-logging with satellite-derived oceanography in hidden Markov models. *J R Soc Int* 15:20180084
- Gui Z, Yu H, Tang Y (2016) Locating traffic hot routes from massive taxi tracks in clusters. *J Inf Sci Eng* 32:113–131
- Gurarie E., Fleming CH, Fagan WF, Laidre KL, Hernández-Pliego J, Ovaskainen O (2017) Correlated velocity models as a fundamental unit of animal movement: synthesis and applications. *Move Ecol* 5: 13
- Hamer KC, Humphreys EM, Magalhaes MC, Garthe S, Hennenck J, Peters G, Grémillet D, Skov H, Wanless S (2009) Fine-scale foraging behaviour of a medium-ranging marine predator. *J Anim Ecol* 78:880–889
- Hansen IJ, Johnson CJ, Cluff HD (2013) Synchronicity of movement paths of barren-ground caribou and tundra wolves. *Polar Biol* 36: 1363–1371
- Kays R, Crofoot MC, Jetz W, Wikelski M (2015) Terrestrial animal tracking as an eye on life and planet. *Science* 348:aaa2478
- Keogh E, Kasetty S (2003) On the need for time series data mining benchmarks: a survey and empirical demonstration. *Data Min Knowl Disc* 7:349–371
- Kranstauber B, Kays R, LaPoint SD, Wikelski M, Safi K (2012) A dynamic Brownian bridge movement model to estimate utilization distributions for heterogeneous animal movement. *J Anim Ecol* 81:738–746
- Kranstauber B, Smolla M, Safi K (2017) Similarity in spatial utilization distributions measured by the earth mover's distance. *Methods Ecol Evol* 8:155–160
- 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* 93:2336–2342
- Lascelles BG, Taylor PR, Miller MG et al (2016) Applying global criteria to tracking data to define important areas for marine conservation. *Divers Distrib* 22:422–431
- Lee JG, Han J, Whang KY (2007) Trajectory clustering: a partition-and-group framework. In: *Proceedings of the 2007 ACM SIGMOD international conference on Management of data*. ACM, pp 593–604
- Li X, Zhao K, Cong G, Jensen CS, Wei W (2018) Deep representation learning for trajectory similarity computation. In: *2018 IEEE 34th International Conference on Data Engineering*. IEEE, pp 617–628
- Lin D (1998) An information-theoretic definition of similarity. In: *Proceeding ICML '98 Proceedings of the Fifteenth International Conference on Machine Learning*. Morgan Kaufmann Publishers Inc, San Francisco, pp 296–304
- Lines J, Bagnall A (2015) Time series classification with ensembles of elastic distance measures. *Data Min Knowl Disc* 29:565–592
- Liu B, de Souza EN, Matwin S, Sydow M (2014) Knowledge-based clustering of ship trajectories using density-based approach. In: *2014 IEEE International Conference on Big Data (Big Data)*. IEEE, pp 603–608
- Long JA, Nelson TA (2013) A review of quantitative methods for movement data. *Int J Geogr Inf Sci* 27:292–318
- Marques AR, Forde H, Revie CW (2018) Time-series clustering of cage-level sea lice data. *PLoS ONE* 13:e0204319
- McClintock BT, Michelot T (2017) momentuHMM: R package for analysis of telemetry data using generalized multivariate hidden Markov models of animal movement. <https://rdrr.io/cran/momentuHMM/f/inst/doc/momentuHMM.pdf>
- McIntyre T, Bester MN, Bornemann H, Tosh CA, de Bruyn PN (2017) Slow to change? Individual fidelity to three-dimensional foraging habitats in southern elephant seals, *Mirounga leonina*. *Anim Behav* 127:91–99
- Meier RE, Wynn RB, Votier SC et al (2015) Consistent foraging areas and commuting corridors of the critically endangered Balearic shearwater *Puffinus mauretanicus* in the northwestern Mediterranean. *Biol Conserv* 190:87–97
- Miller HJ, Dodge S, Miller J, Bohrer G (2019) Towards an integrated science of movement: converging research on animal movement ecology and human mobility science. *Int J Geogr Inf Sci* 16:1–22
- R Core Team (2018) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna <https://www.R-project.org/>
- Ranacher P, Tzavella K (2014) How to compare movement? A review of physical movement similarity measures in geographic information science and beyond. *Cartogr Geogr Inf Sci* 41:286–307
- Ratanamahatana CA, Keogh E (2004) Everything you know about dynamic time warping is wrong. In: *Third workshop on mining temporal and sequential data*. Citeseer, University Park
- Salvador S, Chan P (2007) Toward accurate dynamic time warping in linear time and space. *Intell Data Anal* 11:561–580
- Sarda-Espinosa A (2017) dtwclust: Time series clustering along with optimizations for the dynamic time warping distance. R package version 3.1.2., <https://cran.r-project.org/web/packages/dtwclust/index.html>
- Scharf AK, Belant JL, Beyer D, Wikelski M, Safi K (2018) Habitat suitability does not capture the essence of animal-defined corridors. *Move Ecol* 6:18
- Seidel DP, Dougherty E, Carlson C, Getz WM (2018) Ecological metrics and methods for GPS movement data. *Int J Geo Inf Sci* 32:2272–2293
- Senin P (2008) Dynamic time warping algorithm review. Information and Computer Science Department. University of Hawaii at Manoa, Honolulu
- Shoji A, Aris-Brosou S, Owen E et al (2016) Foraging flexibility and search patterns are unlinked during breeding in a free-ranging seabird. *Mar Biol* 163:72
- Thums M, Fernández-Gracia J, Sequeira AM, Eguíluz VM, Duarte CM, Meekan MG (2018) How big data fast tracked human mobility research and the lessons for animal movement ecology. *Front Mar Sci* 13:21
- Toohey K (2015) SimilarityMeasures: trajectory similarity measures. R package version 1.4, <https://CRAN.R-project.org/package=SimilarityMeasures>
- Toohey K, Duckham M (2015) Trajectory similarity measures. *Sigspatial Special* 7:43–50
- Tuck GN, Thomson RB, Barbraud C, Delord K, Louzao M, Herrera M, Weimerskirch H (2015) An integrated assessment model of seabird population dynamics: can individual heterogeneity in susceptibility to fishing explain abundance trends in Crozet wandering albatross? *J Appl Ecol* 52:950–959
- Vlachos M, Gunopoulos D, Kollios G (2002) Discovering similar multi-dimensional trajectories. In: *Proceeding ICDE '02 Proceedings of the 18th International Conference on Data Engineering*. IEEE Computer Society, Washington
- Votier SC, Fayet AL, Bearhop S, Bodey TW, Clark BL, Grecian J, Guilford T, Hamer KC, Jeglinski JWE, Morgan G, Wakefield E, Patrick SC (2017). Effects of age and reproductive status on individual foraging site fidelity in a long-lived marine predator. *Proc R Soc B* 284:20171068

- Wakefield ED, Bodey TW, Bearhop S, Blackburn J, Colhoun K, Davies R, Dwyer RG, Green JA, Grémillet D, Jackson AL, Jessopp MJ, Kane A, Langston RH, Lescroël A, Murray S, le Nuz M, Patrick SC, Péron C, Soanes LM, Wanless S, Votier SC, Hamer KC (2013) Space partitioning without territoriality in gannets. *Science* 341: 68–70
- Wakefield ED, Cleasby IR, Bearhop S, Bodey TW, Davies RD, Miller PI, Newton J, Votier SC, Hamer KC (2015) Long-term individual foraging site fidelity—why some gannets don't change their spots. *Ecology* 96:3058–3074
- Wang H, Su H, Zheng K, Sadiq S, Zhou X (2013) An effectiveness study on trajectory similarity measures. In: *Proceedings of the Twenty-Fourth Australasian Database Conference*, vol 137. Australian Computer Society Inc., Darlinghurst, pp 13–22
- Wang Z, Yan W, Oates T (2017) Time series classification from scratch with deep neural networks: a strong baseline. In: *2017 International Joint Conference on Neural Networks (IJCNN)*. IEEE, pp 1578–1585
- Whittington J, St. Clair CC, Mercer G (2005) Spatial responses of wolves to roads and trails in mountain valleys. *Ecol Appl* 15:543–553
- Zheng K, Zheng Y, Xie X, Zhou X (2012). Reducing uncertainty of low-sampling-rate trajectories. In: *2012 IEEE 28th International Conference on Data Engineering*. IEEE, pp 1144–1155

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.