Elliptical Time-Density model to estimate wildlife utilization distributions

Jake Wall[*1,2], George Wittemyer[2,3], Valerie LeMay[4], Iain Douglas-Hamilton[2], Brian Klinkenberg[1]

[1]Lab for Advanced Spatial Analysis, Department of Geography, University of British Columbia, 1984 West Mall, Vancouver B.C., Canada, V6T 1Z2

[2]Save the Elephants, P.O. Box 54667, Nairobi, Kenya, 00200

[3]Wittemyer Lab, Department of Fish, Wildlife and Conservation Biology, Colorado State University, 1474 Campus Delivery, Fort Collins CO, USA, 80523

[4]Department of Forest Resources Management, Faculty of Forestry, University of British Columbia, Forest Sciences Centre #2045 - 2424 Main Mall, Vancouver, B.C., Canada, V6T 1Z4

[*]Corresponding author: walljcg@gmail.com

Running Title: Elliptical Time-Density Model

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/2041-210X.12218

Abstract

We present a new animal space-use model (Elliptical Time-Density - ETD) that uses discrete-time tracking data collected in wildlife movement studies. The ETD model provides a trajectory-based, non-parametric approach to estimate the utilization distribution (UD) of an animal, using model parameters derived directly from the movement behavior of the species. The model builds on the theory of ’time-geography’ whereby elliptical constraining regions are established between temporally-adjacent recorded locations. Using a Weibull speed distribution fitted for an animal’s movement data, a time-density value (i.e., time per unit landscape) is determined from the expectation of all elliptical regions equal to, or greater-than, the minimum bounding ellipse for a given landscape point. We tested the ETD model using a tracking dataset for an African elephant (*Loxodonta africana*) and compared the resulting UDs for regularly sampled, frequently recorded locations, as well as irregular random time intervals between locations and also infrequent temporal-sampling regimes, providing insight to the method’s performance with different resolution data. We compared the performance of the ETD model, the Brownian Bridge Movement Model (BBMM), the Time-Geography Density Estimator (TGDE) and the Kernel Density Estimator (KDE) by calculating omission/commission errors from the predicted space-use distribution of each model relative to the true known UD of our elephant test data. The comparison was made for the 10% to 99% percentile UD model areas. The ETD90 model (i.e., ETD model parameterized using the 90% percentile value of the Weibull speed distribution) resulted in the fewest errors of commission and omission with regards to locating the true movement path at the 99% percentile UD area. The ETD model provides an improved approach for estimating animal UDs since: i) parameters are derived directly from the tracking data rather than assumed; ii) parameter values are biologically interpretable; iii) the Weibull speed distribution is adaptable to various temporal-sampling regimes; and iv) the ETD model handles the case of degenerate ellipses thus preserving landscape connectivity in the UD. Software (freeware) for calculating the ETD and a Bayesian framework for estimating the Weibull distribution speed parameters are also introduced in the paper.

Keywords: Animal Movement, Home-range, Utilization Distribution, GPS Tracking, GIS

Introduction

Quantifying the movements of an animal and its use of a landscape and the resources available therein is a fundamental component of wildlife ecology and conservation. Tracking animals with technology such as Global Positioning System (GPS) tracking devices is a widely used methodology that can provide detailed positional information (Kie *et al.*, 2010; Hebblewhite & Haydon, 2010). Although the technology for continuous-time monitoring of animals is fast approaching (e.g., using tri-axial accelerometers (Wilson *et al.*, 2006)), GPS and similar relocation technology (e.g., Argos or Very High Frequency (VHF) collar-based information) only sample at discrete and relatively infrequent times. Models are therefore needed to predict an animal's spatial utilization when not being directly sampled and are generally referred to as animal 'home-range' models (Laver & Kelly, 2008; Fieberg & Börger, 2012).

Home-range models have started to emerge that explicitly incorporate the temporality of sampled positions into the model definition, thereby leveraging the inherent movement structure of the sample to derive a more accurate statistical and biological estimate; e.g., Brownian Bridge Movement Model (BBMM) (Horne *et al.*, 2007), Time-Geographic Density Estimation (TGDE) (Downs, 2010; Downs *et al.*, 2011) dynamic BBMM (dBBMM) (Kranstauber *et al.*, 2012), Potential Path Area (PPA) (Miller, 2005; Long & Nelson, 2012), Time Localized Convex Hull (T-LoCoH) (Lyons *et al.*, 2013). As such, these home-range models are better suited to handle the temporally and spatially autocorrelated nature of high-resolution datasets (Lyons *et al.*, 2013).

The BBMM and dBBMM models both estimate landscape utilization based on mechanistic assumptions (i.e. Brownian motion) of the underlying movement behavior between sampled positions. In contrast, the theory of 'time-geography' (Hägerstrand, 1970; Miller, 2005) uses elliptical areas to demarcate the bounding region that a moving object could have occupied between recorded positions, but does not make assumptions about the form of movement the object took between points (Long & Nelson, 2012). Downs *et al.* (2011) have recently expanded the time-geographic approach to formulate a UD model from geo-ellipses connecting recorded animal locations. Models that produce a utilization distribution (i.e., a spatial model that gives a probability of occupancy, or time spent, for every point in the landscape (Van Winkle, 1975; Jennrich & Turner, 1969; Worton, 1989; Seaman & Powell, 1996; Marzluff *et al.*, 2001; Fieberg & Kochanny, 2005; Keating & Cherry, 2009)), contain more spatial structure and information compared with models that outline potential areas of use (e.g., Minimum Convex Polygons (MCP) and PPA) and, therefore, are generally preferred in analysis of animal space-use (Fieberg & Kochanny, 2005; Fieberg & Börger, 2012).

In this paper, we advance time geographic approaches for space-use estimation by developing a novel approach that quantifies the probable time-density of occupancy within geo-ellipses bounding the potential space use of an individual. Deriving a UD based on the probability of time-spent at a given point in the landscape provides a logical advancement of the TGDE model in line with typical space-use constructs in the wildlife ecology field. Secondly, we present a robust Bayesian framework for modeling the maximum speed parameter critical as input to time-geographic approaches, that provides a construct for interpreting this user selected parametrization of the model. Thirdly, we exemplify application of the model using data collected from an Africana elephant (*Loxodonta africana*) specifically exploring the effects of differential temporal sampling regimes and model parameter inputs on model output. From this example we demonstrate how model parametrization can be conducted in a biologically interpretable manner. Finally, we compare the omission and commission error rates between an absolute movement path (generated) and UD estimates of the ETD model, BBMM, TGDE and the traditional kernel density estimator (KDE), providing a platform by which to contrast our model with commonly used approaches.

Elliptical Time-Density Model Development

Similar to other time-geographic approaches, we begin development of the ETD movement model by considering a pair of sequentially acquired positions from a tracked animal (see Figure 1). We assume the animal was at position $\vec{p}_i$ at time t_i and then at position $\vec{p}_j$ at time t_j where $T_j = t_j - t_i$ and $j = i+1$. Together i and j index a chronologically sorted list of recorded locations. We assume no knowledge of where the animal traveled in between the recorded locations and that it could have taken any complicated, but unknown route with a path length equal to r_j and an average speed of $s_j = \frac{r_j}{T_j}$. The speed s_j puts bounds on the area reachable by the animal during time T_j, also known as the 'Potential Path Area' by Miller (2005); Downs *et al.* (2011); Long & Nelson (2012). It can be shown (see derivation in Supplementary Information Appendix 4) that this region is defined by an ellipse with area $A_j = \pi \frac{r_j}{4} \sqrt{r_j^2 - D_j^2}$ where D_j is the straight line distance between $\vec{p}_i$ and $\vec{p}_j$. Our goal in deriving an animal's UD, is to know where in the landscape the animal is likely to have spent its time over the period it was being monitored. Mathematically, a utilization distribution (UD) is a surface that maps, for every point $\vec{z}$ in $\mathbb{R}$ (defining the extent of all points that could possibly be occupied by the animal), the probability density of finding the animal at a given $\vec{z}$. From our delineation of the elliptical area possibly reached during a time T_j we can calculate a 'time-density' value as $\rho_j = \frac{T_j}{A_j}$ in units of Hr/Km^2.

Although the true mean speed of the animal s_j along the true path length r_j is unknown, it is bounded by a lower value of $s_{min} = \frac{|\vec{p}_j - \vec{p}_i|}{T_j}$ (i.e. the animal had to move at least as fast as s_{min} to move along the shortest distance path - a straight line - path from $\vec{p}_i$ to $\vec{p}_j$). It is also bounded by some biologically realizable upper limit s_{max} based on the physiology of the animal being studied. These constraints on speeds translate into constraints in ellipse areas and each possible average speed value from s_{min} to s_{max} in moving from $\vec{p}_i$ to $\vec{p}_j$ corresponds to a unique bounding elliptical area. Remembering that our goal is to determine the amount of time an animal is likely to have spent at a given point in the landscape, we can choose a particular point $\vec{z}$, and note that it can only be reached if the animal moves at an average speed greater than or equal to

$s_{z,j} = \frac{|r_j|}{T_j} = \frac{(|\vec{z}-\vec{p}_i|+|\vec{p}_j-\vec{z}|)}{T_j}$ where $s_{z,j} \leq s_{max}$. If $s_{z,j} > s_{max}$ then point $\vec{z}$ is unreachable by the animal in time T_j and the corresponding UD value will be zero. If point $\vec{z}$ is reachable, then we can estimate the amount of time spent at $\vec{z}$ by integrating the time-density value over a differential area at $\vec{z}$. Time-density values at $\vec{z}$ can vary according to the area of the elliptical bounding region being considered. We proceed by computing the expected time-density value function at a point $\vec{z}$: $\Theta_j(z) = E\left\{\rho_{z,j}\right\}$ by calculating the expectation of time density values from each possible ellipse equal to, or greater-than, the ellipse required to move between $\vec{p}_i$ via a point $\vec{z}$ to reach $\vec{p}_j$ in time T_j:

$$\Theta_j(\vec{z}) = \int_{s_{z,j}}^{s_{max}} f(s)\rho_j(s)ds \tag{1}$$

where $f(s)$ is the probability density function of average speed over time T_j and where the volume integral of the elliptical time-density function $\Theta_j(z)$ over all points $\vec{z}$ across the overall region $\mathbb{R}$ being considered, is equal to the time T_j:

$$\oint_{\mathbb{R}} \Theta_j(\vec{z})dA = T_j \tag{2}$$

In practice, the integral in equation 2 is approximated by discretizing the landscape into a grid where each cell has an area ΔA, and the discrete set of evaluation points $\{\vec{z}_m\}$ are taken as the grid-cell center points. The integral is then approximated by the sum:

$$\sum_{k=1}^{m} \Theta_j(\vec{z}_m)\Delta A = T_j \tag{3}$$

where $k = 1..m$ indexes the set of discrete evaluation points $\{\vec{z}_m\}$ reachable in time T_j . The discrete UD is a probability mass function whereby the UD value of a given grid-cell is the probability of use by an animal within the grid-cell and the probabilities across all grid-cells are normalized to sum to one.

It follows that the animal's UD can be constructed by adding the fractional amounts of time spent per landscape area as determined using the elliptical time-density function Θ_j from each successive point pair within the movement dataset:

$$UD(z) = \frac{\sum_j \Theta_j(\vec{z})}{T_{Total}} \tag{4}$$

where $T_{Total} = \sum_{j=1}^{n-1} T_j$ is the total time-span of the movement dataset.

Weibull Probability Density Function

Selection of the probability density function $f(s)$ in equation 1 should be based on choosing the mathematical form that best-fits the speed distribution of the data, and $f(s)$ can either be a parametric or non-parametric function. The most basic assumption is that $f(s)$ is uniform, in which case each time-density value $\rho_j(s)$ in equation 1 is equally likely. However, from empirical observation we know that the probability distribution of speeds is unlikely to be uniform and we generally expect that faster speeds are less likely than slower ones. The two-parameter Weibull distribution has a wide range of flexibility in representing variations in shapes and has previously been used to model animal movement (Morales *et al.*, 2004). Here we incorporate a two-parameter Weibull distribution in equation 1 to condition the time-density expectation function Θ_j, although a Gamma distribution or other similarly versatile function are equally useful for this approach. Equation 1 then becomes:

$$\Theta_j(\vec{z}) = \int_{s_{z,j}}^{s_{max}} \frac{4k}{\pi \lambda s} \left(\frac{s}{\lambda}\right)^{k-1} \frac{e^{-\left(\frac{s}{\lambda}\right)^k}}{\sqrt{s^2 t_j^2 - D_j^2}} ds \tag{5}$$

where k is the Weibull shape parameter and λ is the Weibull scale parameter. There is no known analytical solution to equation 5 and the integration must be performed using numerical methods.

The shape and scale parameters that define the functional form of the Weibull distribution in equation 5 are generally not known but can be empirically estimated by fitting a Weibull curve to the distribution of straight line speeds from the consecutive pairs of positions $\{\vec{p}_i, \vec{p}_j\}$ in the movement dataset. The likelihood function to be maximized is then:

$$L = \prod_{j=1}^{n-1} \frac{k}{\lambda} \left(\frac{s_j}{\lambda}\right)^{k-1} e^{-\left(\frac{s_j}{\lambda}\right)^k} \tag{6}$$

and the best-fit parameters solved for using standard methods (Clark, 2007; Gelman & Hill, 2007; Bolker, 2008).

Multi-temporal data

In empirically collected movement datasets, the time interval between sampled locations, T_j, may vary considerably owing to missed fixes or studies that purposely choose to vary the temporal sampling regimes (Katajisto & Moilanen, 2006). In the case of datasets with irregular time intervals between locations, we would expect the Weibull distribution to vary its shape and scale parameters as a function of the temporal sampling regime T_j, making it necessary to explicitly model this variation. We can therefore add an additional parametrization to the Weibull model by writing the scale parameter as a function of T_j:

$$\lambda(T_j) = aT_j^b c^{T_j} \tag{7}$$

where a, b, c are three parameters to be estimated in addition to the shape parameter k. The likelihood function can then be written as:

$$L = \prod_{j=1}^{n} \frac{k}{aT_j^b c^{T_j}} \left(\frac{s_j}{aT_j^b c^{T_j}}\right)^{k-1} e^{-\left(\frac{s_j}{aT_j^b c^{T_j}}\right)^k} \tag{8}$$

Maximum Speed Parameter

The maximum speed value s_{max}, effectively limits the size of the area reachable by an animal in moving between successive recorded locations (Downs *et al.*, 2011; Long & Nelson, 2012) and is therefore an important input parameter to the ETD movement model. The maximum speed value can be chosen in several ways: 1) based on a known biological speed limit for a given species given its physiology 2) using the maximum empirically measured speed in the tracking dataset or 3) using a value that corresponds to a percentile of the empirically best-fit two-parameter Weibull cumulative distribution function (CDF) where

$$CDF_{Weibull} = 1 - e^{-\left(\frac{s}{\lambda}\right)^k} \quad (9)$$

A desired maximum speed value for a given percentile is then determined using:

$$s_{max} = \lambda \left(ln\left(|1-\alpha|\right)\right)^{1/k} \quad (10)$$

where α is the percentile value (e.g., 50th percentile).

Using a percentile value of the Weibull distribution has the benefit that the maximum speed value can vary according to the temporal resolution at which the animal is being sampled as in the case of using the multi-temporal parametrization. Alternatively, using the top speed attainable by an animal over short time periods will lead to an over estimation of the area reachable when considering longer time intervals and either the biological or empirically-derived maximum speed values would have to be specified in consideration of the temporal sampling regime. If the s_{min} value for a given positional pair $\{\vec{p}_i, \vec{p}_j\}$ is greater than or equal to the s_{max} value (i.e. $s_{min} \geq s_{max}$) then the corresponding ellipse eccentricity becomes infinite and the model collapses to a straight line connecting the positional pair (also called the degenerate ellipse (Long & Nelson, 2012)).

Depending on the choice of the maximum speed value, and as a result of discretization of the integral in equation 3 into discrete grid cells, there may be instances when no evaluation points are reachable between a position pair even if s_{min} is less than s_{max}. For example, in the case of a degenerate ellipse where the point pair separation represents the maximum speed in the dataset, then no evaluation points will be reachable, unless they fall exactly on the straight line connecting the positional pair. In this situation, we employ the technique of Wall *et al.* (2013) of first joining the point pair with a straight line and calculating a constant time-density value along the line in units of hours/meter, then clipping the line based on the geometry of the underlying grid, such that any grid intersecting the line will have a time-density value equal to $\Theta_j(\vec{z}) = \frac{d_z}{T_j}$ where d_z is the fractional length of the straight line crossing the grid cell associated with $\vec{z}$.

Elliptical Time-Density Model Application

We applied the ETD model to a movement dataset collected from an African elephant (*Loxodonta africana*) in the Gourma region of Mali (Wall *et al.*, 2013) to highlight use of the ETD model with real data. Additionally, to demonstrate model behaviour as a function of varying temporal sampling regimes typically encountered in applied scenarios, we calculated the ETD model under three different temporal sampling scenarios: 1) the full, high resolution hourly sampled GPS data (dataset 'D1'), 2) by randomly down-sampling D1 to produce a randomly varying temporal resolution dataset with 50% of the original points (dataset 'D2'), and 3) by regularly down-sampling D1 to produce a low temporal resolution data with sampled locations once every 24 hours (dataset 'D3'). For each dataset, the ETD model was calculated using an output grid size of 500 meters.

ETD Model Software

To calculate the ETD model we developed software implemented as part of the ArcMET extension for Esri ArcMap GIS software (Esri, 2013; Wall, 2014) and written using the C# programming language (available at: www.movementecology.net). The algorithm begins by first creating a landscape-wide grid of user-specified cell size that covers the entire extent of the dataset (although, in order to prevent edge effects, an additional expansion option allows the grid to cover an area greater than the locations dataset by a user-specified expansion ratio (default =1.1)).

Once the grid is created, the grid cell center-point coordinates are determined and these become the evaluation points (i.e. the set of points $\{\vec{z}_m\}$) for the time-density function evaluation. Each pair of points in the location's dataset is then considered independently and in parallel, thus speeding the calculation time considerably on machines with multiple logical cores. For each point pair, the straight-line speed (s_{min}) value between points is established. The speed required to reach each landscape evaluation point (s_z) is then calculated and used as the lower bound for the integral in equation 5. If an evaluation point is unreachable without moving faster than s_{max} then it is not considered further. The time-density function is evaluated for each of the points $\vec{z}_m$ reachable within time T_j . The integral in equation 5 is approximated using a trapezoidal Riemann sum with a user specified differential speed unit (default =0.001 Km/Hr). A user defined cut-off parameter allows a minimum output UD probability value to be set (default value = 1E-15), below which, the particular grid cell will be assigned a value of zero, and the rest of the grid cells will have their values adjusted accordingly so that the total sum of values equals one (equation 4).

Speed Parameter Selection

We modeled the speed distributions of datasets D1,D2, D3 using a two-parameter Weibull distribution using a Bayesian framework. For datasets D1 and D3, we started with non-informative uniform prior distributions for both shape and scale parameters and used Markov Chain Monte Carlo (MCMC) with Gibbs Sampling using WinBUGS software called from within R (R Development Core Team, 2011) using the R2WINBUGS library (Gelman & Hill, 2007), to determine the best-fit parameters for the Weibull distribution (λ and k from Equation 5). We ran three MCMC chains each with 100,000 iterations (first 50,000 discarded to allow for burn-in) and confirmed chain convergence using a potential scale reduction factor value of 1.1 as a cut-off (Gelman & Hill, 2007). A complete listing of the procedure and parameter estimates can be found in Supplementary Information Appendix 1 and Appendix 3. A similar procedure was followed for dataset D2 except that parametrization of the scale factor using Equation 7 allowed for variation in the scale parameter as a function of the temporal separation between recorded points

(Supplementary Information Appendix 2).

We explored how the choice of maximum speed value affected the ETD model output by computing the ETD model for dataset D1 and a variety of maximum speed parameter values. We calculated the ETD model using: 1) the maximum speed value in the dataset (Model: D1-MaxSpeed) 2) using the 99% CDF value (Model: D1-99) and 3) using the 50% CDF value (Model: D1-50). For each of these three models, we calculated the core hotspots (defined as the 50% probability contour), the home range (defined as the 95% probability contour (Jennrich & Turner, 1969; Anderson, 1982)), and the total use area based on the 99% probability contour area.

ETD Model Accuracy Assessment

We tested the accuracy of the ETD model by comparing UD percentile contours to true known UD percentile contours for a given dataset as described previously. To establish a true UD, we used a 15-minute temporal resolution dataset collected from an African elephant over a two-week period ($n = 1522$ positions). Using this 'true' path, we established a fine-resolution (10 meter) graticule and calculated the amount of time spent within each grid cell as a percentage of the total time.

We down-sampled the true, 15-minute dataset at hourly intervals and calculated ETD models using 0%, 30%, 50%, 70%, 90%, 99% percentile maximum speed values (note that the 0% model does not allow any off-path movement and effectively connects data points with straight lines; i.e, infinitely eccentric ellipses) using Weibull shape and scale parameters derived from the distribution of hourly speeds: $shape = 0.8638$; $scale = 0.2906$. For additional comparison, we also calculated the BBMM model (animal mobility variance = 893.5 m^2; telemetry standard deviation = 28.85 m^2), a TGDE model parametrized using the maximum dataset speed (2.55 Km/Hr) and a linear decay function, and a fixed bivariate Gaussian kernel (KDE) model (smoothing factor $h_{ref} = 875$) using ArcMET software (Wall, 2014). For each of the nine modeled UDs (*ETD0*, *ETD30, ETD50*, *ETD70*, *ETD90*, *ETD99*, *TGDE, BBMM*, *KDE*) and the true UD (*TrueUD*), we calculated UD areas at 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 99% percentile levels.

We calculated errors of omission and commission to assess the accuracy in capturing space-use along the true movement path of each model. An error of commission was defined to be the number of pixels that fell within a model UD percentile contour that were not within the corresponding *TrueUD* percentile contour (pixels marked as being used that should not have been — Figure 7-A) expressed as a percentage of the total number of unused pixels in the *TrueUD* for a given percentile level. An error of omission was defined as the number of pixels not contained within a model UD percentile contour but that fell within the corresponding *TrueUD* percentile contour (pixels not marked as being used by a model contour but that should have been — Figure 7-B) expressed as a percentage of the total number of used pixels in the *TrueUD* for a given percentile level. We defined a total error metric to be $Total\,Error = \sqrt{omission^2 + commission^2}$ to summarize both commission and omission errors (Figure 7-C).

Results

Conceptually, the morphology of the time-density function kernel can be visualized in Figure 2 where two hypothetical ETD kernels have been generated for differing positional separations for a pair of points separated by one hour of time using the same Weibull parametrization and a maximum speed parameter value of 6 Km/Hr. On the left is a positional pair that are close together with an s_{min} value of 0.3 Km/Hr compared with the right side where the positional pair are located far apart and have an s_{min} value of 5 Km/Hr. The morphology shifts from a peaked circular function, to an elongated elliptical form as the separation of the positional pair increases, and eventually collapses to a straight line as s_{min} approaches s_{max}.

Understanding the performance of a movement model under differing temporal sampling strategies is an important component in model testing and in the applicability of the movement model within generalized animal tracking scenarios. We contrasted the ETD model calculation under the three selected temporal sampling regimes D1, D2 and D3 (Figure 3 d-f). The maximum speed parameters derived from the Weibull fit were D1: 6.44 Km/Hr, D2: 5.82 Km/Hr, D3: 1.70

Km/Hr. Areal values for core areas, home ranges and total use areas across the three datasets demonstrate similar trends depending on temporal sampling regime (Figure 4). As temporal sampling frequency decreases, the ETD model morphology erodes and becomes more spatially diffuse and less peaked (Figure 4). For example, the area of the UD 50% percentile core area inflates by a factor of 4.7 when moving from hourly to 24-hour sampling, the UD 95% percentile area inflates by a factor of 3.6, and the 99% percentile UD total use area inflates by a factor of 3.2 times when moving from an hourly sampling regime to 24-hour sampling for this dataset.

The maximum speed parameter value influences the calculation of the ETD model (Figure 5). Using a maximum speed value equivalent to the empirically-derived maximum speed value of a given dataset led to the UD with the largest 99% percentile total use area, and 95% percentile home range area (Figure 6). In contrast, 50% percentile core areas were generally equivalent across all three ETD models. Using successively lower maximum speed values, i.e the 99% and 50% percentile Weibull speed distribution values, had the effect of concentrating the three space-use levels into smaller, probabilistically dense regions. We would expect this to be the case given that, for a lower maximum speed parameter, a greater number of point pairs are necessarily connected by straight lines, thus limiting reachable areas.

There was a general exponential increase in commission errors with increasing percentile levels with the KDE model committing the greatest number of commission errors, followed by the TGDE, ETD99, BBMM, ETD90, ETD70, ETD50, ETD30 models and the ETD0 model committing the least (Figure 7-A). This trend shows that as models progressed towards greater restriction (e.g., ETD0 & ETD30) where fewer off-linear movements are allowed, they committed fewer commission errors at all percentile levels. All models showed a peak in commission errors at high percentile levels.

Omission errors showed very different trends. Above the 20% percentile UD contour, the ETD0 model committed the greatest number of omission errors as would be expected because of the restrictiveness to straight-line paths between points. KDE and TGDE caused the highest number of omission errors at the 10% and 20% percentile levels but these converged to zero by the 99% percentile level (Figure 7-B). The ETD99 model, BBMM, and ETD90 demonstrated the best performance in terms of low omission error rates across all percentile levels. Overall, the ETD99 model had the lowest root mean square error (Total Error) when considering both errors of omission and commission across all percentile levels although the ETD90 model had the least total error at the 90% and 99% percentiles (Figure 7-C). The ETD99 and BBMM showed similarly low total error values between the 30% and 80% percentiles.

Discussion

The ETD movement model provides a powerful framework for estimating animal space-use based on discrete time tracking data. Elliptical based modeling of movement, rooted in the theory of time-geography, is a unique approach to estimate animal space-use, yet development in this direction has been relatively recent (Downs, 2010; Downs *et al.*, 2011; Long & Nelson, 2012). The ETD model builds on this foundation and, by doing so, offers several distinct properties relative to other methods of animal UD estimation including biologically-based parametrization that avoids assumptions regarding the underlying movement process. The lack of biological realism or interpretability in model parametrization has been raised as a weakness in previous space-use and movement modeling approaches.

Trajectory-based model

An important characteristic of the ETD model is that it is developed based on consideration of the trajectory of an animal and explicitly incorporates the temporality of recorded positions by considering pairs of observed locations in their temporal sequence in the parametrization of the geo-ellipses used to estimate space use. Use of the multi-temporal parametrization of the Weibull scale function in equation 7 lets the ETD model adapt to changing temporal lengths between positional pairs and makes the model flexible in order to handle any temporal sampling regime with biological realism. Explicit treatment of the serial structure within trajectory-based models obviates consideration of independence issues that typically arise with other non-trajectory based spatial-use estimators (Swihart & Slade, 1985, 1997; Horne *et al.*, 2007). The recent development of trajectory-based space-use models including the BBMM (Horne *et al.*, 2007), the dBBMM (Kranstauber *et al.*, 2012), the TGDE model (Downs *et al.*, 2011) and the PPA model (Long & Nelson, 2012) has been a departure from methods such as KDE (Worton, 1989) that focus on the point pattern of recorded locations but don't explicitly model temporality or movement between points. Trajectory-based models are critical when interested in the connectivity or directionality of movements within a landscape (Horne *et al.*, 2007).

The ETD model introduced here has several strengths relative to other, recently developed trajectory-based methods. Our method requires no assumptions about the form of movement used by the animal in contrast to both the BBMM and dBBMM models which assume Brownian motion despite its known short-comings for modeling the movements of most species (Horne *et al.*, 2007). Rather, the ETD model simply establishes constraints on movement based on the empirically observed data or, alternatively, a user's knowledge of the biology of the species. Our comparison of omission-commission error demonstrates the strength of our approach in that the ETD99 model provided the best estimate of space used (lowest error rate) relative to BBMM, TGDE or KDE (Figure 7). This analysis also provides a useful framework for determining the strengths and weaknesses of different UD model structures.

In contrast to the ellipse based approach of Long & Nelson (2012), our approach is a probabilistic model leading to a UD, whereas the PPA model establishes overall bounds of where an animal might have traveled between locations. This outer bounding output makes the PPA more similar to a minimum convex polygon (MCP) approach which delineates the outer bounds of movement when considering independent fixes (White & Garrott, 1990). As a result of this fundamental difference, we do not directly compare our method to this approach.

Finally, Downs *et al.* (2011) used two probability density functions - uniform and linearly decreasing - to calculate probability density values at landscape points within the maximum ellipse area. Our approach is fundamentally different in that by calculating the expected time-density value (of all elliptical areas corresponding to speeds equal-to, or greater-than, the minimum speed necessary to reach a landscape point) using the probability distribution of speed (equation 1), we address the question of how long an animal is likely to have spent at a given point in the landscape.

Parametrization

Space-use models have generally been classified as being either parametric or non-parametric (White & Garrott, 1990; Kernohan *et al.*, 2001). We consider the ETD to be non-parametric given that the model is developed solely on empirical data and doesn't assume a distributional form as with some models (e.g., the bivariate-normal class of models (Kernohan *et al.*, 2001)). Although we selected the probability density function $f(s)$ in equation 1 to be the two-parameter Weibull distribution because of its close fit to the empirically observed speed distribution. The ETD is not limited to this distribution, and parameters can be drawn from other distributional forms (e.g., Gamma) or by using the empirical distribution itself. The maximum speed parameter is the primary user defined parameter affecting model output. But unlike many models, this parameter can be selected based on biologically relevant emergent properties of the underlying data. Within the BBMM model for example, Long & Nelson (2012) have noted that the animal mobility variance parameter, which controls the spatial extent of each Brownian bridge density function connecting positional pairs, and is a critical model parameter for BBMM, is difficult to interpret. Similarly, the smoothing parameter used in the popular KDE model (Worton, 1989) is also difficult to link to the underlying biology of the species and its initial selection can even be subjective (Katajisto & Moilanen, 2006; Horne *et al.*, 2007; Kie *et al.*, 2010). Even should a user choose to parametrize ETD subjectively, the parametrization and model output can be interpreted biologically, a strength lacking from other approaches.

The ability to vary the maximum speed parameter, either as a percentage of the fitted Weibull distribution, or biologically, based on prior knowledge of an animal's movement characteristics, is important in constraining the extent of the ETD model output. Use of higher maximum speed parameter values are least prone to errors of omission but lead to the largest area of use in any given scenario (Figure 7). However, it is biologically unrealistic for most terrestrial animals to frequently move at their top speeds (although this may in fact be true for certain aquatic or avian species). As such, using a high maximum speed value (e.g., the 99% value) may inherently over estimate the areas used by the animal leading to errors of commission (Figure 7). Selection of maximum speed parameter may also be driven by the goals of the intended output of the space-use model and be guided by the particular questions the analyst is asking. The model user must also decide what the consequences are for a given modeling scenario in committing omission errors by selecting a low maximum speed value versus commission errors when selecting a high value. Interestingly, for the elephants sampled here, the 99% percentile maximum speed value model (ETD99) provides the least overall error when considering all UD percentile contour areas and was nearly matched by the BBMM model between the 30% and 80% percent contour areas. For resource selection in a use-availability framework, the ETD output for the biologically maximum speed provides a realization of availability that provides a sound estimate of availability (based on our omission analysis) given the sampling of the elephant's movement path in line with that provided using conditional logistic regression (Boyce, 2006).

Choice of a maximum speed value of anything less than the empirically-derived maximum speed will result in the generation of straight line segments between positional pairs, rather than elliptical regions (i.e., infinitely eccentric ellipses with zero area). The ETD model handles this situation by calculating the time-density value along the straight line segment and therefore every positional pair contributes towards the overall UD. Skipping the positional pair if the maximum speed parameter is below the speed necessary to move between positions, as conducted by Downs *et al.* (2011), biases the UD. The connectivity property of the ETD model ensures that connectivity of the sequentially recorded dataset is preserved in the output UD, a particularly important quality when assessing animal movement corridors (Horne *et al.*, 2007).

The effects of the maximum speed parameter are also linked to the user-selected output UD grid cell size. If the selected grid cell size is too coarse, then many landscape evaluation points (grid cell center points) will become unreachable, especially as the minimum straight line speed value s_{min} approaches the maximum speed parameter value. A finer grid cell size will therefore result in a more precise output UD. A consequence of choosing a finer output grid is increased computation time and system resource requirements, thus this decision will likely be a balance among these trade-offs.

Determining the best-fit parameters for the two-parameter Weibull distribution from empirical data is based on using straight line segments to connect positional pairs in the tracking dataset. An argument against this approach might follow the lines that since the animal rarely moves in straight lines between points, especially as the temporal separation between observed locations increases, then the Weibull parametrization is not representative of the true movement speed distribution of the animal and is likely to be an under estimate of the speeds capable by the animal. However, we are modeling probable space-use between points, and therefore deriving the top speed from the animal's movements at the sampling interval is a biologically sensible solution. Further refinement of the method could attempt to incorporate biologically relevant structure in the parametrization of the Weibull model. For instance, circadian, movement state specific, landscape related or seasonal patterns in movement could be modeled directly and used to parametrize ETD for relevant time periods. Such refinement would likely offer more accurate parametrization of movement properties without increasing computational requirements should such structure be incorporated post hoc. In addition, the error structure of the location estimates (GPS points or otherwise) can be modeled to incorporate additional uncertainty to the output UD, though the utility of incorporation of location error relates to the data (high versus low accuracy), underlying grid parametrization (coarse versus fine), and movement characteristics of the animal.

Conclusion

The ETD model is a probabilistic, trajectory-based model for estimating the UD of an animal from discrete-time positional data and can be used with both regularly and irregularly sampled datasets of varying temporal frequency. The ETD movement model builds on the concept of elliptical constraining regions and time-geography by introducing a new time-density function which determines the most likely time-density value (time spent per unit landscape) at a particular landscape point and is an important conceptual departure from other methods. The model is non-parametric and makes no assumption about the mechanistic movement behavior of the observed animal, providing an unbiased estimate of the animal's temporal space-use. Time information is implicitly encoded in the model formulation, freeing the output UD estimate from statistical issues of auto-correlation. The ETD model development presented in this paper and Supplementary Information, along with freely available software for calculating the ETD model, provides a framework for incorporation of the ETD model within generalized animal tracking studies.

Acknowledgments

We'd like to thank Save the Elephants, Kenya for provision of elephant tracking data used in the analyses. We also thank the Esri Conservation Program for provision of the ArcMap software used in this research. The research was supported financially by a Canadian National Science and Research Council (NSERC) - PGSD3 (348450) award and by Save the Elephants, Kenya. We also thank Dr. Lang Wu, Dr. Cindy Greenwood and Ed Kroc at the University of British Columbia and especially three anonymous reviewers for contribution of their ideas to the paper.

Data Accessibility

Tracking datasets used in this study are available through the *MoveBank* data repository (www.movebank.org):

1. 'Salif' tracking dataset: DOI.xxx

2. Model accuracy assessment tracking dataset: DOI.xxx

References

Anderson, D.J. (1982) The Home Range: A new nonparameteric estimation technique. *Ecology*, **63**, 103–112.

Bolker, B.M. (2008) *Ecological Models and Data in R*. Princeton University Press.

Boyce, M.S. (2006) Scale for resource selection functions. *Diversity and Distributions*, **12**, 269–276.

Clark, J.S. (2007) *Models for Ecological Data: An Introduction*. Princeton University Press.

Downs, J.A. (2010) Time-geographic density estimation for moving point objects. *Lecture Notes in Computer Science*, **6292**, 16–26.

Downs, J.A., Horner, M.W. & Tucker, A.D. (2011) Time-geographic density estimation for home range analysis. *Annals of GIS*, **17**, 163–171.

Esri (2013) *ArcGIS Desktop: Release 10.2.1*. Environmental Systems Research Institute, Redlands, CA.

Fieberg, J. & Börger, L. (2012) Could you please phrase "home range" as a question? *Journal of Mammalogy*, **93**, 890–902.

Fieberg, J. & Kochanny, C.O. (2005) Quantifying home-range overlap: the importance of the utilization distribution. *Journal of Wildlife Management*, **69**, 1346–1359.

Gelman, A. & Hill, J.C. (2007) *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press, Cambridge ; New York.

Hägerstrand, T. (1970) What about people in regional science? *Papers in Regional Science*, **24**, 7–21.

Hebblewhite, M. & Haydon, D.T. (2010) Distinguishing technology from biology: a critical review of the use of GPS telemetry data in ecology. *Philosophical Transactions of the Royal Society of London - Series B: Biological Sciences*, **365**, 2303–2312.

Horne, J.S., Garton, E.O., Krone, S.M. & Lewis, J.S. (2007) Analyzing animal movements using Brownian bridges. *Ecology*, **88**, 2354–2363.

Jennrich, R.I. & Turner, F.B. (1969) Measurement of Non-circular Home Range. *Journal of Theoretical Biology*, **22**, 227–237.

Katajisto, J. & Moilanen, A. (2006) Kernel-based home range method for data with irregular sampling intervals. *Ecological Modelling*, **194**, 405–413.

Keating, K. & Cherry, S. (2009) Modeling utilization distributions in space and time. *Ecology*, **90**, 1971–1980.

Kernohan, B.J., Gitzen, R.A. & Millspaugh, J.J. (2001) Analysis of Animal Space Use and Movements. J.J. Millspaugh & J.M. Marzluff, eds., *Radio Tracking and Animal Populations*, chapter 5, pp. 125–166. Academic Press, San Diego, California, USA.

Kie, J.G., Matthiopoulos, J., Fieberg, J., Powell, R.A., Cagnacci, F., Mitchell, M.S., Gaillard, J.M. & Moorcroft, P.R. (2010) The home-range concept: are traditional estimators still relevant with modern telemetry technology? *Philosophical Transactions of the Royal Society of London - Series B: Biological Sciences*, **365**, 2221–2231.

Kranstauber, B., Kays, R., Lapoint, S.D., Wikelski, M. & Safi, K. (2012) A dynamic Brownian bridge movement model to estimate utilization distributions for heterogeneous animal movement. *Journal of Animal Ecology*, **81**, 738–746.

Laver, P.N. & Kelly, M.J. (2008) A Critical Review of Home Range Studies. *Journal of Wildlife Management*, **72**, 290–298.

Long, J.A. & Nelson, T.A. (2012) Time geography and wildlife home range delineation. *The Journal of Wildlife Management*, **76**, 407–413.

Lyons, A., Turner, W. & Getz, W. (2013) Home range plus: a space-time characterization of movement over real landscapes. *Movement Ecology*, **1**, 2.

Marzluff, J.M., Knick, S.T. & Millspaugh, J.J. (2001) High-Tech Behavioral Ecology: Modeling the Distribution of Animal Activities to Better Understand Wildlife Space Use and Resource Selection. J.J. Millspaugh & J.M. Marzluff, eds., *Radio Tracking and Animal Populations*, chapter 12, pp. 309–326. Academic Press, San Diego, California, USA.

Miller, H.J. (2005) A Measurement Theory for Time Geography. *Geographical Analysis*, **37**, 17–45.

Morales, J.M., Haydon, D.T., Frair, J., Holsiner, K.E. & Fryxell, J.M. (2004) Extracting more out of relocation data: Building movement models as mixtures of random walks. *Ecology*, **85**, 2436–2445.

R Development Core Team, R. (2011) R: A Language and Environment for Statistical Computing.

Seaman, D.E. & Powell, R.A. (1996) An evaluation of the accuracy of kernel density estimators for home range analysis. *Ecology*, **77**, 2075–2085.

Swihart, R.K. & Slade, N. (1985) Testing for independence of observations in animal movements. *Ecology*, **66**, 1176–1184.

Swihart, R.K. & Slade, N. (1997) On testing for independenceof animal movements. *Journal of Agriculture, Biological, and Environmental Statistics*, **2**, 48–63.

Van Winkle, W. (1975) Comparison of several probabilistic home-range models. *The Journal of Wildlife Management*, **39**, 118–123.

Wall, J. (2014) *Movement Ecology Tools for ArcGIS (ArcMET) Version 10.2.1*. URL: www.movementecology.net.

Wall, J., Wittemyer, G., Klinkenberg, B., LeMay, V. & Douglas-Hamilton, I. (2013) Characterizing properties and drivers of long distance movements by elephants (Loxodonta africana) in the Gourma, Mali. *Biological Conservation*, **157**, 60–68.

White, G.C. & Garrott, R.A. (1990) *Analysis of Wildlife Radio-Tracking Data*. Academic Press, Inc., San Diego, CA.

Wilson, R.P., White, C.R., Quintana, F., Halsey, L.G., Liebsch, N., Martin, G.R. & Butler, P.J. (2006) Moving towards acceleration for estimates of activity-specific metabolic rate in free-living animals: the case of the cormorant. *Journal of Animal Ecology*, **75**, 1081–1090.

Worton, B.J. (1989) Kernel Methods for Estimating the Utilization Distribution in Home-Range Studies. *Ecology*, **70**, 164–168.

List of Figures

Figure 1:

Path and ellipse geometry.

Figure 2:

An example of two different elliptical time-density function morphologies for a pair of points acquired close together (left) and spread apart (right) and their relative locations along the Weibull speed distribution curve (bottom). Vertical lines indicate the locations of the pairs of sequentially acquired animal locations and shading illustrate the relative probability of where the animal might have been found in between the recorded locations.

Figure 3:

The ETD movement model. UD probability values have been color coded to range from low chance of occurrence (green) through to a high chance of occurrence (red), while blank (white) areas had less than 1E-15 probability of occurrence and were subsequently set to zero probability. A) Dataset D1 (1 hour resolution) B) Dataset D2 (50% randomly deleted points resulting in variable resolution up to 20 hours data-gap) C) Dataset D3 (24 hour re-sampled dataset). D) ETD model of D1 using max speed value and parameters from Supplementary Information Appendix 1, Table A1 E) ETD model of D2 using max speed and the multi-temporal parametrization of the Weibull with parameters found in Supplementary Information Appendix 2, Table A2 F) ETD model of dataset D3 using Weibull parameters found in Supplementary Information Appendix 3, Table A3.

Figure 4:

Comparison of the 50%, 95% and 99% percentile areas calculated from ETD models based on the datasets D1, D2 and D3. The maximum speed value found within each dataset has been used as the maximum speed value in the ETD model.

Figure 5:

ETD UDs calculated from dataset D1: A) ETD calculated using the maximum speed (6.43 Km/Hr) b) ETD calculated using the 99% empirical Weibull speed value (3.01 Km/Hr) c) ETD calculated using 50% empirical Weibull speed value (0.29 Km/Hr)

Figure 6:

Comparison of the 50%, 90% and 99% ETD model percentile contour areas calculated for the elephant 'Salif'. The ETD has been parametrized in three ways: using a maximum speed value of 50% and 99% of the empirical Weibull speed distribution and by using a constant speed equivalent to the maximum empirically-derived speed (6.44 Km/Hr) found in the tracking dataset.

Figure 7

Accuracy assessment of the ETD, BBMM and KDE models using 1-hour locations sampled from a 15-min elephant GPS tracking dataset ($n = 1522$ positions) over a two week period. Models were evaluated using a 10 meter grid resolution and compared with the true UD calculated from the 15-min data. A) Errors of commission are defined to be the number of pixels that fell within a model UD percentile contour that were not within the corresponding *TrueUD* percentile contour expressed as a percentage of the total number of unused pixels in the *TrueUD* for a given percentile level. B) Errors of omission are defined as the number of pixels not contained within a model UD percentile contour but that fell within the corresponding *TrueUD* percentile contour expressed as a percentage of the total number of used pixels in the *TrueUD* for a given percentile level. C) The total error taking into account both omission and commission errors is defined as $Total\ Error = \sqrt{omission^2 + commission^2}$.

Figures

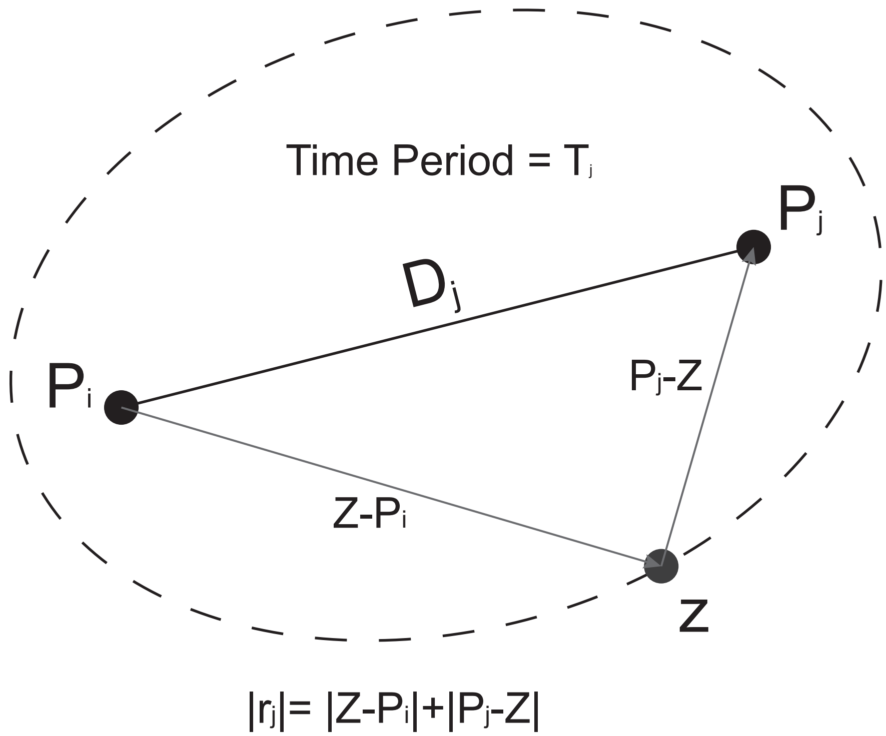


Figure 1

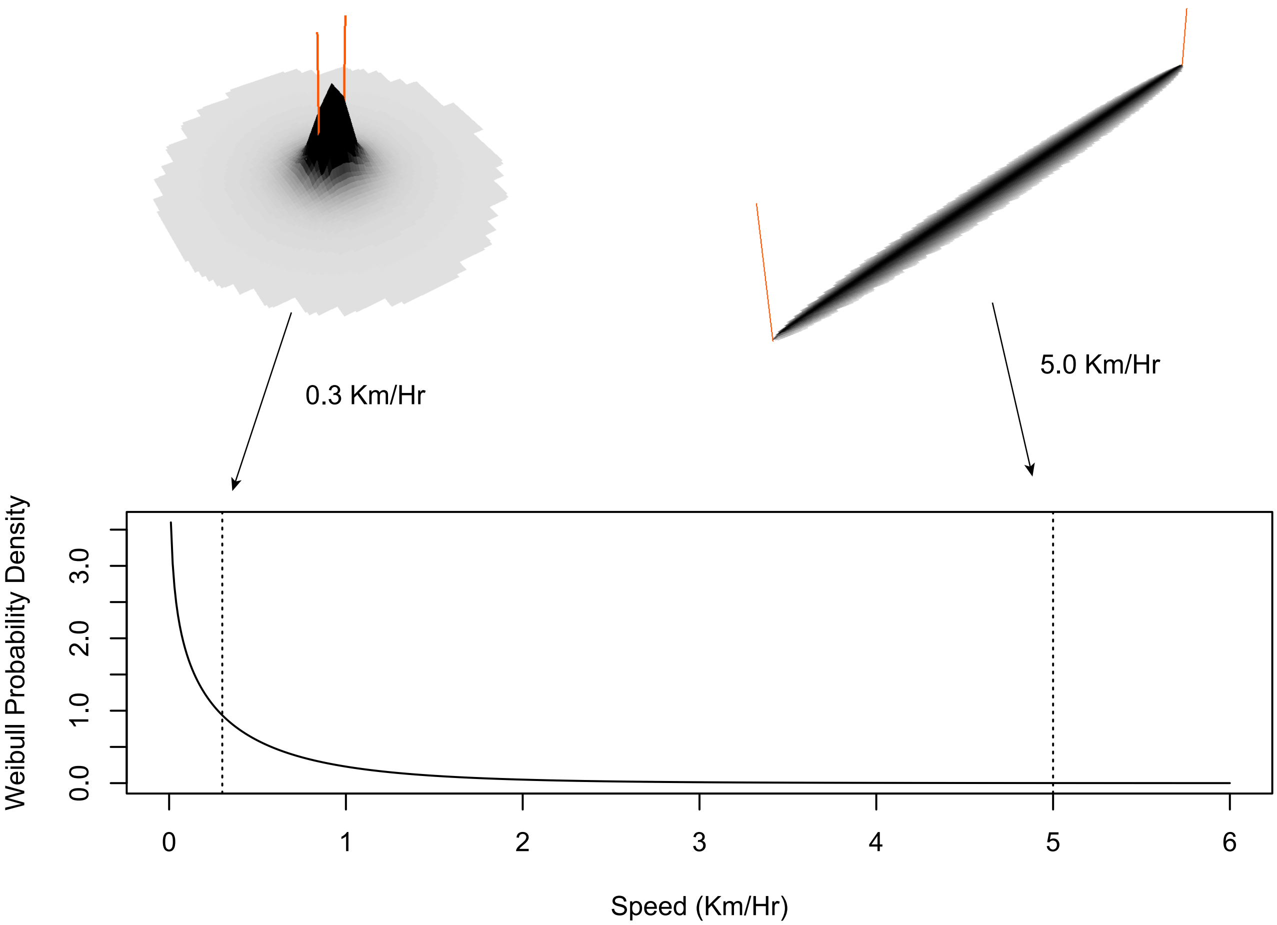

0.3 Km/Hr
5.0 Km/Hr
Weibull Probability Density
0.0
1.0
2.0
3.0
0
1
2
3
4
5
6
Speed (Km/Hr)

Figure 2

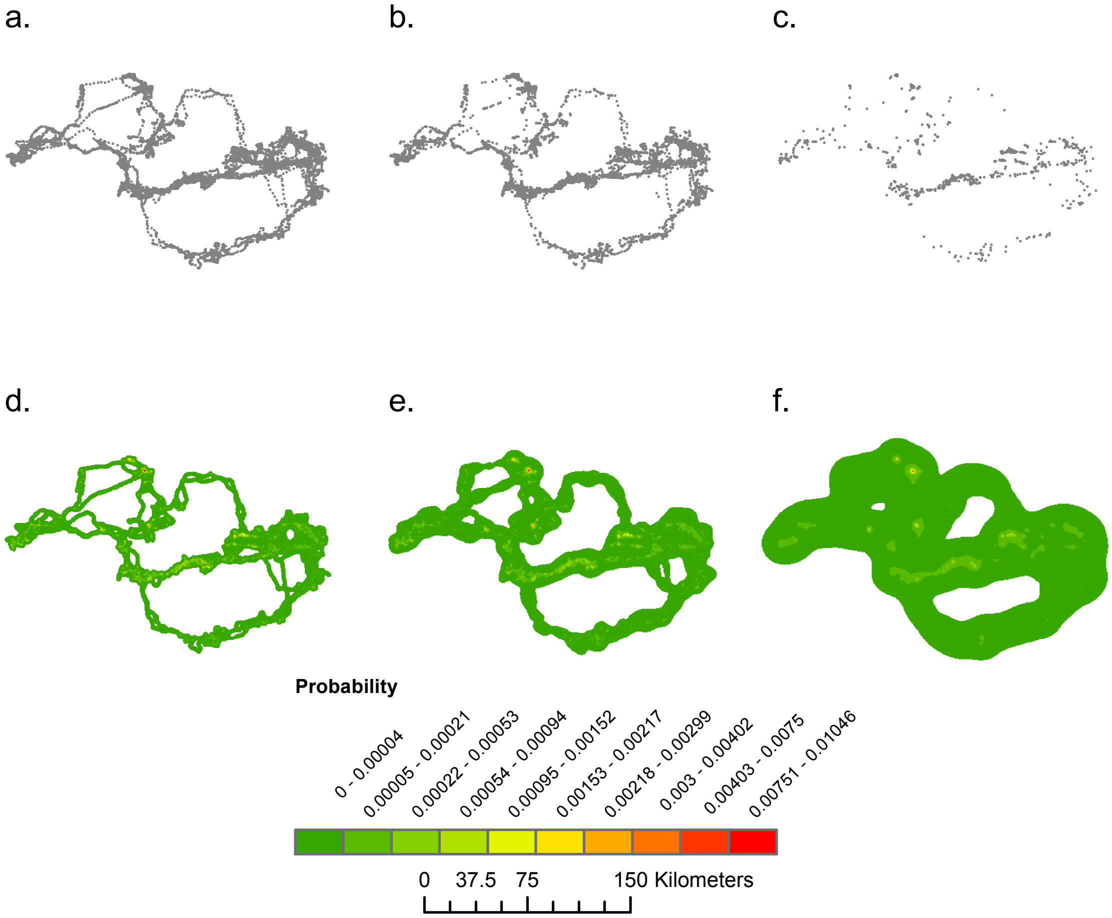
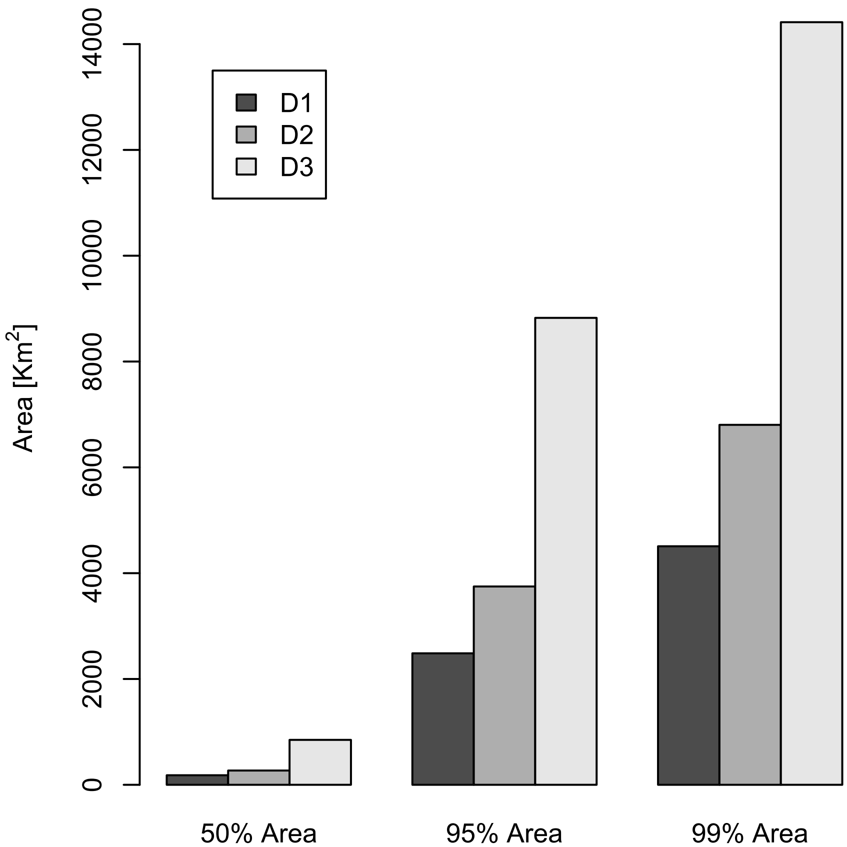

a.
b.
c.
d.
e.
f.
Probability
0 - 0.00004
0.00005 - 0.00021
0.00022 - 0.00053
0.00054 - 0.00094
0.00095 - 0.00152
0.00153 - 0.00217
0.00218 - 0.00299
0.003 - 0.00402
0.00403 - 0.0075
0.00751 - 0.01046
0 37.5 75 150 Kilometers

Figure 3



Area [Km^2]
0
2000
4000
6000
8000
10000
12000
14000
D1
D2
D3
50% Area
95% Area
99% Area

Figure 4

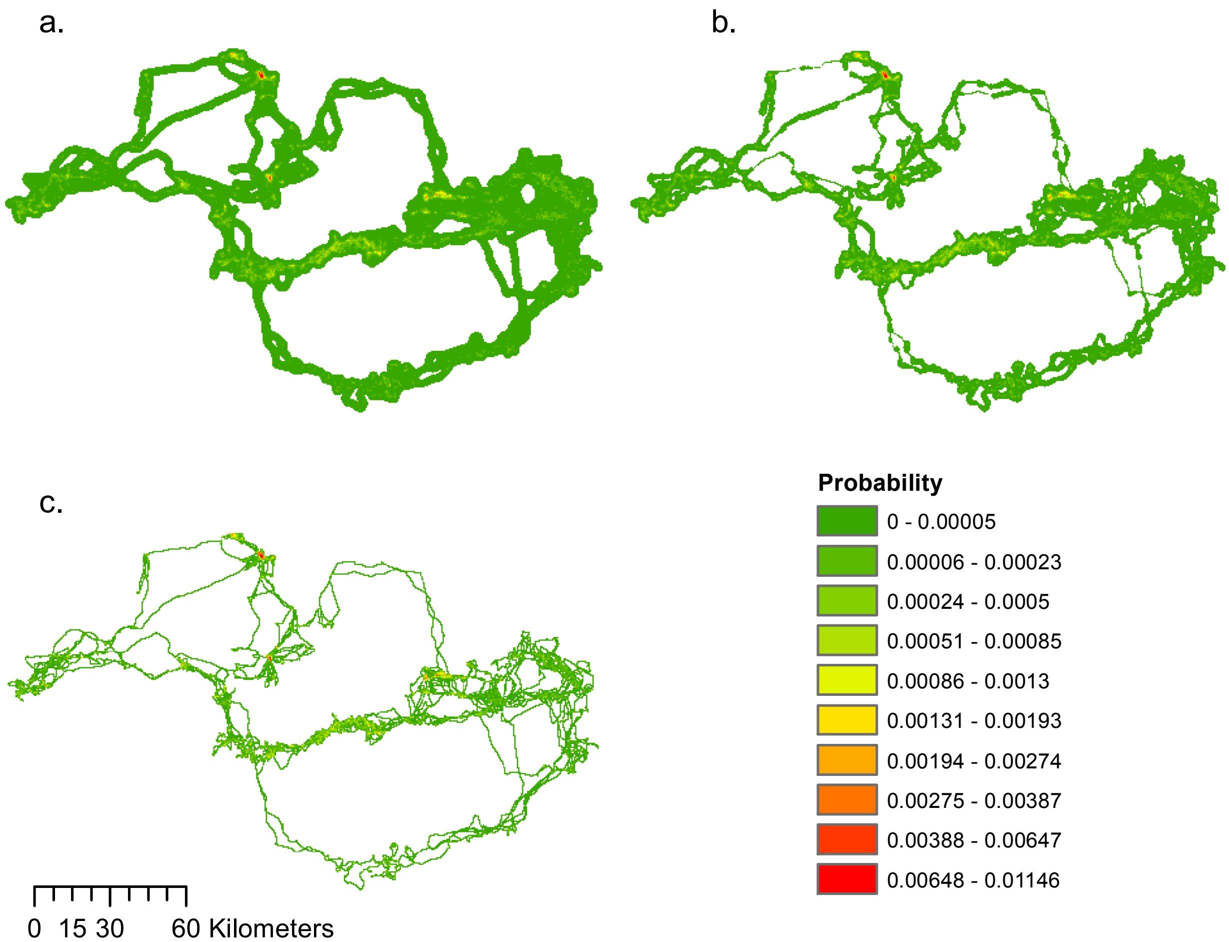

a.
b.
c.
0 15 30 60 Kilometers
Probability
0 - 0.00005
0.00006 - 0.00023
0.00024 - 0.0005
0.00051 - 0.00085
0.00086 - 0.0013
0.00131 - 0.00193
0.00194 - 0.00274
0.00275 - 0.00387
0.00388 - 0.00647
0.00648 - 0.01146

Figure 5

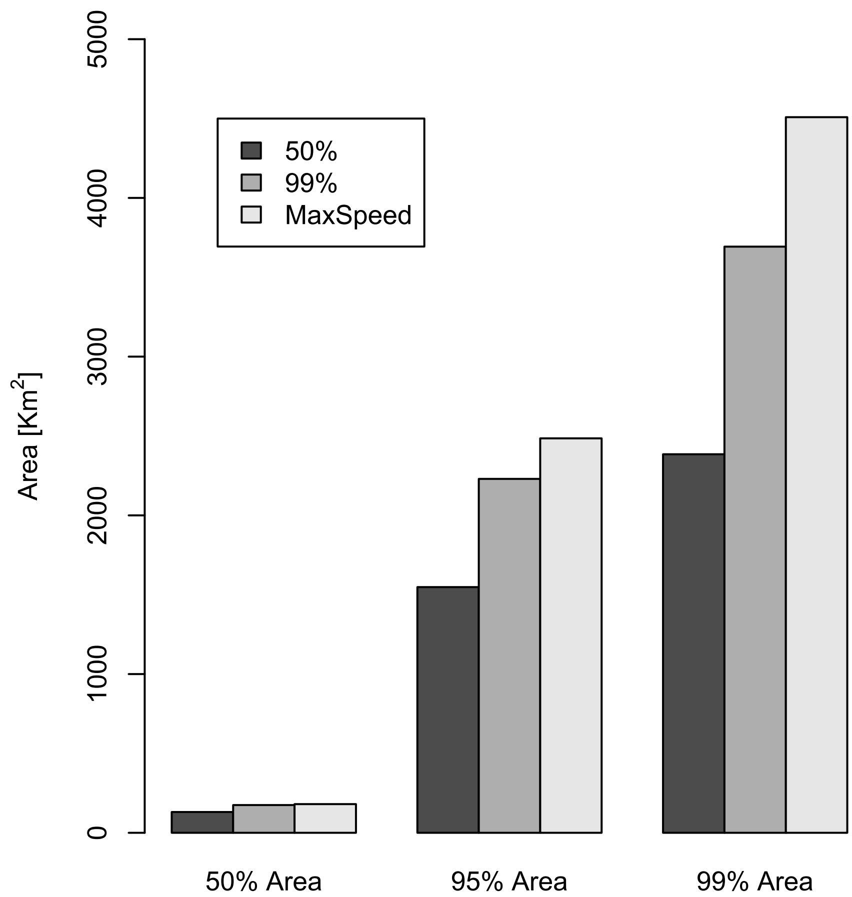

5000
4000
3000
2000
1000
0
Area [Km2]
50%
99%
MaxSpeed
50% Area
95% Area
99% Area

Figure 6

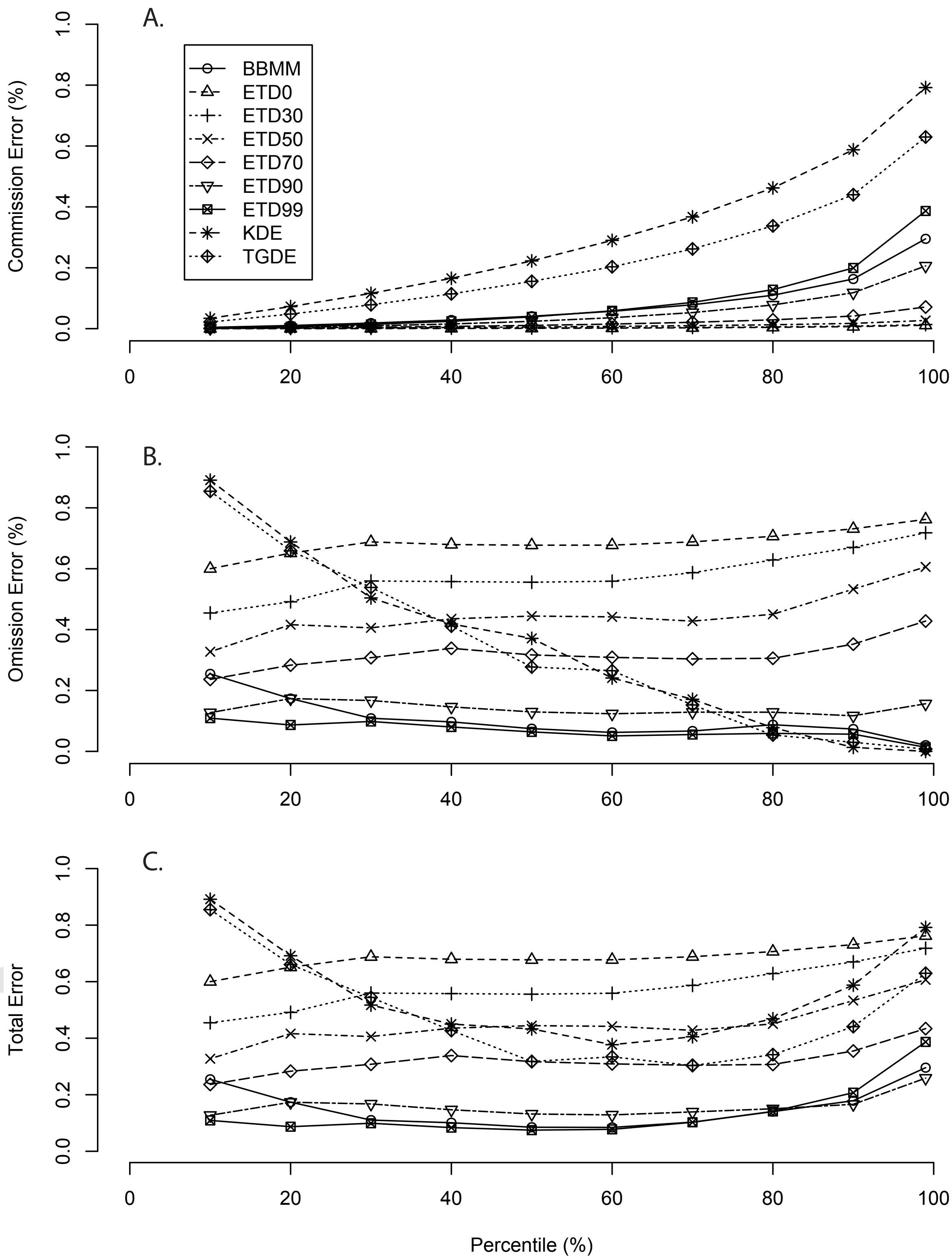

A.
Commission Error (%)
BBMM
ETD0
ETD30
ETD50
ETD70
ETD90
ETD99
KDE
TGDE
B.
Omission Error (%)
C.
Total Error
0.0
0.2
0.4
0.6
0.8
1.0
0
20
40
60
80
100
Percentile (%)

Figure 7