

Bias in estimating animal travel distance: the effect of sampling frequency

J. Marcus Rowcliffe^{1*}, Chris Carbone¹, Roland Kays^{2,3}, Bart Kranstauber^{4,6} and Patrick A. Jansen^{2,5,6}

¹ZSL Institute of Zoology, Regent's Park, London NW 4RY, UK; ²Smithsonian Tropical Research Institute, Balboa, Panamá City, Panamá; ³Nature Research Center, North Carolina Museum of Natural Sciences, 11 W. Jones Street, Raleigh, NC 27601, USA; ⁴Max Planck Institute for Ornithology, Department of Migration and Immuno-Ecology, Radolfzell, Germany; ⁵Wageningen University, Centre for Ecosystem Studies, Wageningen, the Netherlands; and ⁶University of Groningen, Centre for Ecological and Evolutionary Studies, Groningen, the Netherlands

Summary

1. The distance travelled by animals is an important ecological variable that links behaviour, energetics and demography. It is usually measured by summing straight-line distances between intermittently sampled locations along continuous animal movement paths. The extent to which this approach underestimates travel distance remains a rarely addressed and unsolved problem, largely because true movement paths are rarely, if ever, available for comparison. Here, we use simulated movement paths parameterized with empirical movement data to study how estimates of distance travelled are affected by sampling frequency.

2. We used a novel method to obtain fine-scale characteristics of animal movement from camera trap videos for a set of tropical forest mammals and used these characteristics to generate detailed movement paths. We then sampled these paths at different frequencies, simulating telemetry studies, and quantified the accuracy of sampled travel distance estimation.

3. For our focal species, typical telemetry studies would underestimate distances travelled by 67–93%, and extremely high sampling frequencies (several fixes per minute) would be required to get tolerably accurate estimates. The form of the relationship between tortuosity, sample frequency, and distance travelled was such that absolute distance cannot accurately be estimated by the infrequent samples used in typical tracking studies.

4. We conclude that the underestimation of distance travelled is a serious but underappreciated problem. Currently, there is no reliable, widely applicable method to obtain approximately unbiased estimates of distance travelled by animals. Further research on this problem is needed.

Key-words: Barro Colorado Island, camera traps, daily distance, day range, movement models, radiotracking, random walk, telemetry, travel distance, tropical forest

Introduction

The distance that animals must travel to meet their daily nutritional and reproductive needs affects demographic outcomes through the net availability of energy for reproduction and survival and through exposure to predation (Morales *et al.* 2010). Distance travelled thus forms a quantifiable mechanistic link between behaviour and energetics as well as between abundance and demographic processes (see for example Carbone, Teacher & Rowcliffe 2007; Carbone, Pettorelli & Stephens 2011). Measurement of distance travelled inevitably relies on

intermittent location data, usually derived from telemetry or visual tracking, and summing the straight-line distances between locations to give total distance. Except in the unusual case that the locations recorded are resting places with straight-line movement between them, this approach underestimates true distance travelled because actual paths taken between location fixes are not in fact straight. The more tortuous a movement path is (that is, the more an animal turns between fixes), and the lower the frequency of fixes, the greater this underestimation will be. This presents a problem for robust and comparable measurement of distance travelled, as the outcome is a function not only of the actual distance, but also of the animal's turning behaviour and the scale at which we observe its path.

*Correspondence author: E-mail: marcus.rowcliffe@ioz.ac.uk

This underestimation problem has long been recognised. For example, Laundré *et al.* (1987) and Reynolds & Laundré (1990) showed that distances travelled estimated from low sampling frequencies were much shorter than, and uncorrelated with those based on higher sampling frequencies for several mammal species, while Pépin *et al.* (2004) illustrated the dependence of estimated distance travelled by red deer *Cervus elaphus* on sampling frequency. In this study, we build on this work by developing a novel methodological framework for estimating true, scale-independent distance travelled by animals, and quantifying the effect of path sampling on apparent distance moved. We define true or scale-independent distance as that measured at a scale commensurate with scale of the animal's moves (*sensu* Turchin 1998, p129). In practice, this means measuring at roughly the same sampling frequency as the animal's step rate, as in principle there can be no further tortuosity of paths at a scale finer than this.

A great deal of research in movement ecology has been devoted to understanding how the scale of observation interacts with animal movement behaviour to determine observed outcomes (Kareiva & Shigesada 1983; Turchin 1998; Nouvellet, Bacon & Waxman 2009). Key strands of this research include finding the right balance between over- and undersampling to avoid spatial autocorrelation between observations while minimising loss of potentially important spatial information (Turchin 1998); more recent developments in understanding two-dimensional space use that can deal with non-independence and location error in observations, such as Brownian bridge (Horne *et al.* 2007) and state space models (Jonsen, Flemming & Myers 2005; Patterson *et al.* 2008); defining measures of movement tortuosity (also termed sinuosity) that are independent of the scale of observation (Bovet & Benhamou 1988; Benhamou 2004); and the use of random walk models to study animal movement, dispersal and population redistribution (Codling, Plank & Benhamou 2008).

Among these approaches, only the last two touch on the problem of how to measure distance travelled in a scale-independent way. Specifically, this research provides equations linking net straight-line displacement expected for given time between observations and path tortuosity (Bovet & Benhamou 1988; Codling & Hill 2005). However, these equations are largely based on relatively simple models whose assumptions may not be satisfied in most real world situations. Most importantly, they assume that the parameters defining animal velocity are constant over time, whereas in practice, animals typically shift between slower, more tortuous and faster, more direct movement in response to their environment and internal state at multiple temporal and spatial scales (Morales & Elnor 2002; Morales *et al.* 2004; Fryxell *et al.* 2008). This further limits the applicability of current random walk theory to the problem of scale-independent measurement of distance travelled.

The only study that we are aware of that has explicitly attempted to estimate a scale-independent total distance travelled from intermittent fix data is that of Pépin *et al.* (2004), who suggested subsampling positions from telemetry data at a range of lower sampling frequencies, and fitting an asymptotic curve to the relationship between apparent distance and

sampling frequency, with the estimated asymptote assumed to represent true distance travelled. In practice, the hyperbolic curve suggested in this study fits real data very poorly, giving an asymptote substantially lower even than the apparent distance derived from the actual telemetry sampling frequency. In the absence of a generally accepted method for estimating true, scale-independent distance travelled from intermittent fix data, most empirical and comparative studies of distance travelled do not deal with, or even acknowledge, this estimation problem, casting doubt on the value of these estimates.

In this study, we used empirical fine-scale movement patterns from a community of terrestrial mammals as the basis for developing a general framework for understanding the impact of sampling frequency on the apparent distance travelled. To quantify apparent distance travelled for a given species and sampling frequency, we used correlated random walk simulations. We parameterised these simulations using data on the movement speed and turning behaviour of animals recorded by camera traps. On the basis of the patterns found, we elucidate how sampling frequency and path tortuosity interact to determine the degree to which sampling locations along movement paths underestimates true distance travelled. Finally, in the light of these results, we evaluate the potential for a more consistent, scale-independent approach for the measurement of distance travelled in future.

Methods

FIELD METHODS

Empirical data on fine-scale movements of animals were derived from camera traps that recorded videos of terrestrial animals walking through the tropical moist forest of Barro Colorado Island, Panama. Twenty camera traps (Rapidfire RC55; Reconyx Inc., Holmen WI, USA) were deployed for 1 year, from February 2008 to February 2009 (Kays *et al.* 2011; Rowcliffe *et al.* 2011). The cameras were placed at computer-generated random locations and were moved to new random locations approximately every 8 days. The cameras were triggered by passive infrared motion sensors and recorded a sequence of consecutive digital images, using an infrared flash at night. Cameras were mounted approximately 20 cm above-ground on small trees, with the angle of view parallel to the overall slope of the ground and unobstructed by vegetation and topographic relief. For every motion trigger event, cameras recorded 10 low-resolution (1 mega pixel) pictures at a frame rate of just under 1 s⁻¹ and could immediately be triggered again without delay, producing short video clips of animals moving in front of the camera. A total of 6312 trap-nights at 789 locations yielded 17 226 animal detections of 25 different vertebrate species.

To obtain data on the movement patterns of animals, we examined a subset of the image sequences in the field before removing the cameras. We noted the position of the animal in each image relative to nearby landmarks such as trees and rocks and traced this movement path across the field of view with a measuring tape (Supporting information). Estimates of total distance travelled using this method were highly correlated with estimates made remotely through video image analysis (Kays *et al.* 2011), and based on this and our experience of the process, we believe that our estimates of the point locations in movement path were accurate to within a few cm. This is much more

accurate than most GPS or VHF tracking location estimates, which typically have 10–20 m error (Frair *et al.* 2010).

We measured the movement speed of animals from 2140 video sequences by dividing the length of each animal's path during the sequence by the time between first and last images recording the animal. We measured the angle of each change in direction of travel between consecutive images with a compass. We focused this analysis on 10 core species for which we obtained at least 30 sequences (the minimum for reasonably precise estimation) of at least three images (the minimum within which a turn could be observed), a total of 1172 sequences, yielding 635 recorded turns. Sample sizes are summarised in Table 1.

ESTIMATION OF MOVEMENT PARAMETERS

The goal here was to derive measures of movement behaviour that can be used to parameterise a realistic simulation of fine-scale movement (simulation methods below). Parameters to be estimated were average speed and its variance, and three aspects of turning behaviour were the following: (i) the probability of turning at each step; (ii) the probability that successive turns are made to the same side (side persistence); and (iii) the distribution of turn angles.

Speed observations were log-normally distributed, and we therefore used this distribution to simulate moves, requiring us to estimate the mean and standard deviation of log speed. The appropriate measure of average rate when the numerator is held constant is the harmonic mean (Ferguson 1931), which effectively controls for oversampling of higher rates by weighting rates by their inverse. Furthermore, this inverse weighting remains appropriate even if the numerators vary, so long as they vary independently of the rates. In the case of camera trap records, speed is observed within an arena of essentially fixed size (the camera detection zone), and the numerator (distance travelled) is thus indeed independent of the rate (speed). We therefore estimate average log speed as its arithmetic mean weighted by inverse speed and take the standard deviation of log speed to describe its dispersion.

As change in movement behaviour over time is a pervasive characteristic of animal movement (Fryxell *et al.* 2008), with important consequences for path tortuosity, and hence for perceived distance travelled (Turchin 1998), we also needed to characterise this change in behaviour. On the basis that different movement behaviours are likely to be characterised by shifts between slower, more tortuous and faster, more direct movement patterns (Morales & Elner 2002; Morales *et al.* 2004), we looked for correlations between speed and

each of the three elements of turning behaviour listed earlier. In the case of turn probability, we used a binomial generalised linear model (GLM), with dependent variable defined as the number of turns observed given the total possible turns in each sequence (i.e. two less than the number of images in the sequence). Probability of turning to the same side in sequential turns was also modelled using a binomial GLM for pairs of turns within sequences, with the dependent variable defined as the number of turns made in the same direction as the previous turn given the total number of turn comparisons in each sequence. For the purposes of describing simulations below, we denote the probability of turning to the same side predicted by this GLM s .

To quantify turn angle distribution, we fitted von Mises distributions (Mardia & Jupp 2000) to the absolute radian values of turn angles obtained in the field. The choice of distribution was based on a visual comparison of goodness of fit between von Mises, wrapped Cauchy and wrapped normal distributions, which indicated very little difference in functional forms for our data, but a marginal preference for von Mises. Preliminary exploration of turn angles also indicated fewer shallow turns than expected, giving a peak frequency at intermediate turn angles. While it is possible that this is a genuine distribution of angles, with animals tending to avoid shallow turns, our experience suggests that it is more likely due to a tendency for observers to miss shallow turns when attempting to trace paths on the ground while looking at footage from oblique camera angles in the field. To solve this problem, we used a mixture model composed of an underlying half von Mises distribution with location parameter 0 to estimate the underlying distribution of absolute turn angles ($0 < a < \pi$), with an exponentially declining probability of missing turns of increasing angle. The probability of an observed angle $p(a)$ is thus given by:

$$p(a) = \frac{e^{\kappa \cos(a)}}{\pi I_0(\kappa)} (1 - \varepsilon^{-a}) \quad \text{eqn 1}$$

where $I_0(x)$ is the zero-order modified Bessel function (Bowman 1958), ε is a parameter defining the rate of increase in probability of detecting an angle, and κ defines the von Mises concentration parameter for the underlying distribution of angles. High values of κ indicate angles concentrated around 0, while $\kappa = 0$ indicates a uniform distribution of angles between 0° and 180° . The probability density of an observed angle is given by:

$$f(a) = \frac{p(a)}{\int_{a=0}^{\pi} p(a) da} \quad \text{eqn 2}$$

Table 1. Sample sizes for movement parameters estimated for the 10 Panamanian forest mammal species with at least 30 sequences examined for turns. Abbreviated common names in brackets are used hereafter

Species	Sample sizes		
	Turn rate	Turn angle	Speed
Mouse unknown species (mouse)	39	15	42
Tome's spiny rat <i>Proechimys semispinosus</i> (rat)	84	67	135
Red-tailed squirrel <i>Sciurus granatensis</i> (squirrel)	39	27	68
Central American agouti <i>Dasyprocta punctata</i> (agouti)	601	313	980
White-nosed coati <i>Nasua narica</i> (coati)	41	21	129
Nine-banded armadillo <i>Dasyurus novemcinctus</i> (armadillo)	30	21	41
Paca <i>Agouti paca</i> (paca)	110	47	196
Ocelot <i>Leopardus pardalis</i> (ocelot)	48	24	93
Central American red brocket deer <i>Mazama temama</i> (brocket)	70	48	184
Collared peccary <i>Tayassu tajacu</i> (peccary)	110	52	272

and parameters κ and ε are estimated by maximising the likelihood:

$$L(\mathbf{a}) = \prod f(a_i) \quad \text{eqn 3}$$

The parameters of both GLMs and the angle distribution model were fitted as functions of speed and species in various combinations and for the angle model either with or without failure to detect shallow angles (defined, respectively, by $0 < \varepsilon < \infty$ and $\varepsilon = \infty$). Models were compared using the Akaike Information Criterion (Akaike 1973) corrected for small sample size (AICc, Burnham & Anderson 2002).

The proportion of turns made that were in fact recorded was then estimated as the area under the probability curve:

$$q = \int_0^{\pi} p(a) da \quad \text{eqn 4}$$

Given the predicted probability of turning derived from the GLM, u , which assumes that all turns are detected, a corrected estimate of true turning probability, w , was given by:

$$w = \frac{u}{q} \quad \text{eqn 5}$$

RANDOM WALK SIMULATION

The speed and turning parameters (probability, angle and side persistence) obtained from camera trapping were used to parameterise a correlated random walk model for each species, in which movement paths were simulated through a series of short straight-line steps. We used a step rate of 0.9 s per step (the average frame rate in the camera data). For each species, we simulated a single continuous series of 88 000 steps, corresponding to 22 h of movement. We use a single long sequence rather than shorter replicate sequences because we are interested in defining the average outcome as precisely as possible, not in variation between replicates, which is defined primarily by the arbitrary length of sequence chosen. Assuming starting x, y co-ordinates of 0,0, the position after the first step was given by:

$$\begin{aligned} x &= 0.9v \sin(\theta) \\ y &= 0.9v \cos(\theta) \end{aligned} \quad \text{eqn 6}$$

where v is speed in m s^{-1} and θ is direction of travel in radians, chosen from a random uniform distribution between 0 and 2π . Positions after subsequent steps were then given by:

$$\begin{aligned} x_{t+1} &= x_t + 0.9v_t \sin(\theta_{t+1}) \\ y_{t+1} &= y_t + 0.9v_t \cos(\theta_{t+1}) \end{aligned} \quad \text{eqn 7}$$

where direction θ_{t+1} is related to direction in the previous time step by:

$$\theta_{t+1} = \theta_t + \begin{cases} 0 & \Phi_t < w \\ [a_{t-}]a_t & \Phi_t > w, \Gamma_t < s \\ -[a_{t-}]a_t & \Phi_t > w, \Gamma_t > s \end{cases} \quad \text{eqn 8}$$

Here, a are random turning angle deviates drawn from a half von Mises distribution with location parameter 0 and concentration parameter κ , and $[a_{t-}]$ is the side of the previous turn, assigned 1 for right and -1 for left. Φ and Γ are standard random uniform deviates defining, respectively, whether a turn is made (with respect to turn probability parameter w), and if so whether it is to the same side as the previous turn (with respect to side persistence parameter s). At each step, v was drawn from an autocorrelated log-normal distribution, with sequential values given by:

$$v_t = \exp\left(\bar{v} + \sigma\sqrt{(1-r^2)}\sum_{s=1}^t z_s r^{t-s} + z_0 r^t\right) \quad \text{eqn 9}$$

where \bar{v} and σ are, respectively, the mean and standard deviation of $\log_e(v)$, z is a set of uncorrelated random normal deviates, and r is the correlation coefficient between successive $\log_e(v)$ values. Autocorrelation was introduced to allow persistence of movement behaviour over time, simulating shifts between periods of faster and slower movement. Through correlation between speed and turning parameters, this also allows for linked shifts in tortuosity of movement.

Because we had no information on patterns of autocorrelation in speed, we explored the sensitivity of results to a range of r values between 0.9 and 0.9999, roughly corresponding to persistence of movement patterns over several seconds to several hours (as judged by how long it takes for the correlation to fall close to 0). While persistence of behaviours may occur over longer timescales than this in some cases (e.g. Fryxell *et al.* 2008), this study focused on a suite of non-migratory species resident within a relatively small, self-contained area, and we would not expect strong persistence over days or longer in this case. Turn angle concentration parameter values, κ , were given by predicted values from the best fitting turn angle model described in the previous section. The probability of making a turn at each step was defined by w , given by predicted values from the best fitting turn probability GLM and corrected for missed turns. The side of the first turn in a sequence was chosen randomly with equal probability of left or right, and thereafter, each turn was to either the same or the opposite side relative to the previous turn, with probability given by predicted values from the best fitting turn side GLM.

QUANTIFYING UNDERESTIMATION OF DISTANCE TRAVELLED

To quantify the effect of sampling frequency on apparent distance travelled for each of the 10 species, we simulated their movement paths, sampled each path at different frequencies (simulating telemetry at different intensities), and calculated the proportional accuracy for each sampling frequency. On the basis of our experience examining sequential images, we believe that the sampling frequency of camera images (on average 0.9 s per image) is close to the pace rates of the species covered and that there was therefore negligible turning at finer resolution than that observed. On the basis of the definition of true, scale-independent travel distance given in the introduction, we therefore assumed that the sum of all steps taken between start and end points of a simulated path at maximum sampling frequency represented the true distance travelled (d_{TRUE}). Apparent distance travelled (d) at a given sampling frequency was then calculated as the sum of straight-line distances between fixes, averaged across all possible starting points on the path. From these, we calculated proportional accuracy, D , quantifying the degree to which apparent distance reflects the true distance travelled:

$$D = \frac{d}{d_{\text{TRUE}}} \quad \text{eqn 10}$$

D takes values between 0 and 1, with $D = 1$ indicating perfect accuracy.

To predict proportional accuracy for any given sampling frequency, we developed a model to describe the relationship between these variables. A preliminary heuristic examination of simulated data suggested that that D plotted against sampling frequency (F) followed a log-log relationship characterised by an initial linear

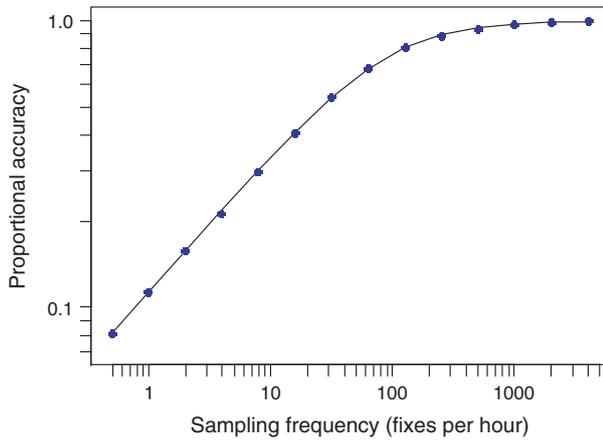


Fig. 1. Example of the relationship between proportional accuracy D (apparent distance relative to the true total distance travelled) and sampling frequency. Points are data generated by a correlated random walk model parameterised for agouti, assuming speed autocorrelation $r = 0.999$. The curve is eqn 12 fitted to these data with parameters: $\delta = 0.428$; $\beta = 1.236$; $\gamma = 4.342$. Axes are logarithmic.

increase, followed by gradual transition to an asymptote (Fig. 1). The rate of change in such a pattern can be characterised by a negative sigmoid function:

$$\frac{d \log(D)}{d \log(F)} = \frac{\gamma}{1 + e^{\beta(\log(F) - \gamma)}} \quad \text{eqn 11}$$

where α defines the slope of the initial increasing arm, β defines the rate of change between increasing and asymptotic arms, and γ defines the midpoint of the transition between these arms on the $\log(F)$ axis. Taking the indefinite integral, we derive:

$$\log(D) = \delta \frac{-\log(1 + e^{\beta(\gamma - \log(F))})}{\beta} \quad \text{eqn 12}$$

which gives an asymptotic proportional accuracy of 1. Figure 1 illustrates this function fitted to data from a random walk simulation of the form described earlier with proportional accuracy calculated for 14 points, demonstrating a close match between data and model.

All simulations and analyses were carried out using R 2.13.1 (R Development Core Team, 2011).

Results

QUANTIFYING MOVEMENT BEHAVIOUR

Modelling of turning angle distribution including failure to detect shallow angles (including finite ε parameter) was strongly supported over models without this process (Table 2). The turning angle concentration parameter κ was most strongly related to speed alone, with essentially no support for additional variation in turning angle between species (Table 2). The best model in the set ($\kappa(\mathbf{v}) \varepsilon(\cdot)$) had very strong support (AIC weight = 0.95), and visual inspection of data distributions and values predicted by this model for different speeds suggests a good fit between model and data (Fig. 2). The relationship between κ and speed was positive, indicating sharper turning angles at higher speeds (Fig. 2). Model $\kappa(\mathbf{v}) \varepsilon(\cdot)$ was used to generate turning angles for the simulations.

Table 2. Summary of support for 10 alternative models of turning angle distribution as a function of species and speed of movement, indicating the number of parameters (K), difference in AIC corrected for small sample size from that of the best model (ΔAICc) and AICc weight (w). κ denotes von Mises dispersion for the underlying distribution of turns, and ε denotes rate of decline in probability of missing shallow turns. Where no ε appears, it is assumed that no turns are missed. (\mathbf{v}) denotes speed as a covariate, (\mathbf{s}) denotes stratification by species, and (\cdot) denotes no covariate or stratification used for the parameter indicated

Model	K	ΔAICc	w
$\kappa(\mathbf{v}) \varepsilon(\cdot)$	3	0	0.952
$\kappa(\mathbf{v}) \varepsilon(\mathbf{s})$	12	6.4	0.039
$\kappa(\mathbf{v} + \mathbf{s}) \varepsilon(\cdot)$	12	9.4	0.009
$\kappa(\mathbf{v} * \mathbf{s}) \varepsilon(\cdot)$	21	21	0
$\kappa(\mathbf{v})$	2	37.3	0
$\kappa(\mathbf{s}) \varepsilon(\mathbf{s})$	20	48.2	0
$\kappa(\mathbf{s}), \varepsilon(\cdot)$	11	48.5	0
$\kappa(\cdot) \varepsilon(\cdot)$	2	52.4	0
$\kappa(\mathbf{s})$	10	84	0
$\kappa(\cdot)$	1	85.8	0

The model for turning probability that included speed, species and their interaction ($\mathbf{v} * \mathbf{s}$) was strongly supported over all others (AIC weight = 0.9; Table 3), indicating that there were trends in turning probability with speed that varied between species (Fig. 3). These relationships were generally positive, indicating higher turning rates at higher speeds; however, this trend was weak or non-existent for most species, and in all cases the expected turning rate per unit distance decreased with speed. The model including speed, species and their interaction ($\mathbf{v} * \mathbf{s}$) was used to generate turning probabilities (u) for the simulations.

Modelling the probability of sequential turns being made on the same side (side persistence), there was no support for variation between species, but speed had some support as a predictor (AIC weight = 0.47). This indicated a weak positive effect of speed on persistence in turning side, with predicted probabilities of maintaining turning side from one turn to the next varying between about 0.38. and 0.5. However, given that support for this model was not overwhelming and that the trend described by it was weak, we used the null model (AIC weight = 0.53) to define turning side persistence in the simulations. The average probability of turning to the same side as the previous turn was 0.41.

QUANTIFYING UNDERESTIMATION OF DISTANCE TRAVELLED

Accuracy of distance travelled was insensitive to the degree of autocorrelation in speed assumed (Fig. S1). The four autocorrelation scenarios explored simulated persistence of movement behaviour over timescales from a few seconds to several hours, but for all species, proportional accuracy of distance travelled varied by a maximum of about 5% between scenarios. For further exploration of patterns, we focus on simulations assuming $r = 0.99$, indicating persistence of movement behaviour over a timescale of a few minutes (Fig. 4).

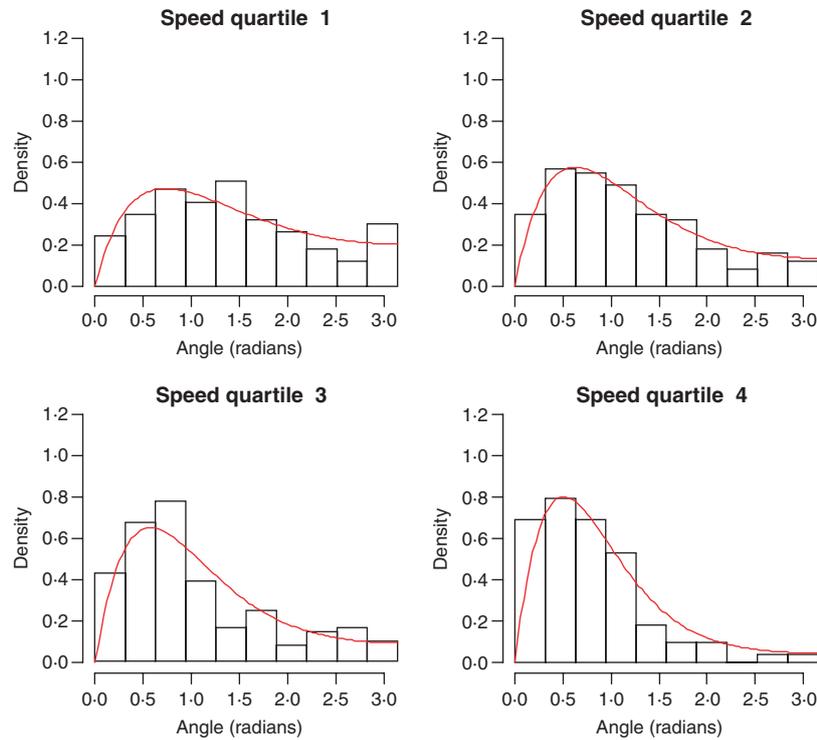


Fig. 2. Turning angle distributions observed (bars) and predicted by model $\kappa(\mathbf{v}) \alpha(\cdot)$ (curves) for different speeds. Speed quartiles used to split observations are from slow (quartile 1) to fast (quartile 4), and the fitted curve on each panel is for mid-quartile speed.

Table 3. Summary of support for alternative models of turn probability (probability of turning at each step, left-hand panel) and turn side persistence (probability of successive turns being on the same side, right-hand panel) in relation to species and speed of movement. Columns contain the number of parameters including intercept (K), difference in AIC corrected for small sample size from that of the best model (ΔAICc) and AICc weight (w) for each model. \mathbf{v} denotes speed as a covariate, \mathbf{s} denotes stratification by species, and **NULL** denotes no covariate or stratification used

Turn probability				Turn side persistence probability			
Model	K	ΔAICc	w	Model	K	ΔAICc	w
$\mathbf{v}*\mathbf{s}$	20	0	0.902	NULL	1	0.0	0.528
$\mathbf{v} + \mathbf{s}$	11	4.4	0.098	\mathbf{v}	2	0.2	0.466
\mathbf{v}	2	28.9	0	$\mathbf{v} + \mathbf{s}$	11	11.0	0.002
\mathbf{s}	10	71.3	0	$\mathbf{v}*\mathbf{s}$	20	11.0	0.002
NULL	1	109.9	0	\mathbf{s}	10	11.6	0.002

For all species, proportional accuracy was sensitive to sampling frequency below a threshold of about one fix per minute. The degree of underestimation of distance travelled at sampling frequencies that are typically used in field studies was large. At sampling frequencies of 0.5 to 6 fixes h^{-1} (which are typical of the more intensive radiotracking studies that report distance travelled to date), proportional accuracy was in the region of 0.07–0.33. This implies that distances travelled would be underestimated by a factor of at least 3, and as much as 14, given a typical radiotracking study of these species. Even at an

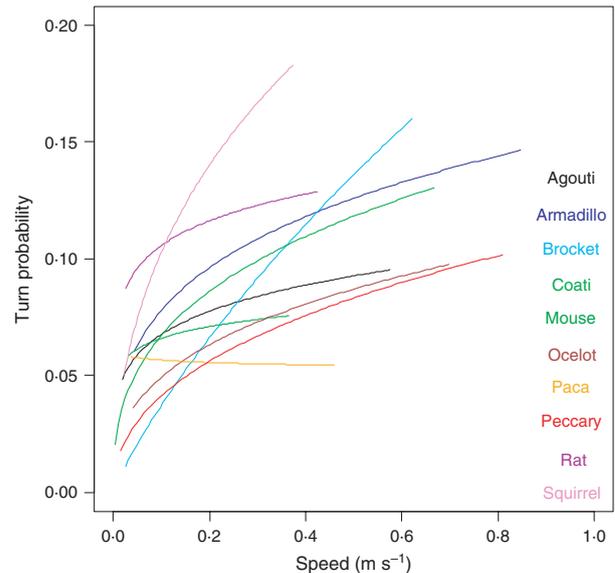


Fig. 3. Predicted turn probability values derived from a generalised linear model with speed, species and their interaction term as predictors. The range of speeds plotted covers the central 95% quantile for each species.

exceptionally high sampling frequency of 60 fixes h^{-1} , distance travelled would still be underestimated by 25–45%. A frequency of at least 400 fixes h^{-1} (one fix every 9 s) would be necessary to estimate distance travelled to within 5% of the true value for the group of species that we studied. The degree

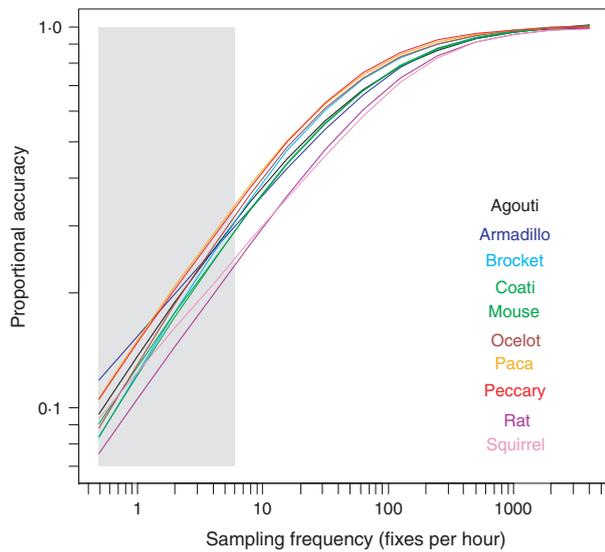


Fig. 4. Proportional accuracy (ratio of apparent to true distance travelled), plotted against sampling frequency for 10 terrestrial mammal species on Barro Colorado Island, Panama. Autocorrelation of speed in simulations used to generate curves (r) was 0.99 for all curves. Typical sampling frequencies in radiotracking studies are indicated by the shaded region. The parameters δ , β and γ in eqn 12 were estimated by fitting curves to data generated from simulations of the species-specific random walk models.

of underestimation at a given sampling frequency tended to be lower for the larger species (such as paca and peccary) than for the smaller species (such as squirrel and rat).

The form of the relationship between proportional accuracy and sampling frequency varied among species primarily in the position of the curve on the sampling frequency axis (parameter γ in eqn 12; Fig. 4). Other parameters describing the curve (the slope of the increasing arm of the curve, δ , and the rate of transition from increase to asymptote, β) were similar among species.

Discussion

Distance travelled is an important determinant of space use and energy expenditure that is traditionally estimated by summing the linear distances between a limited number of location fixes across the day. That this method underestimates the actual distance travelled owing to unobserved tortuosity of movement paths is well known and has been extensively addressed through analysis of random walk models (e.g. Bovet & Benhamou 1988; Benhamou 2004; Codling & Hill 2005). However, the implications of this phenomenon for the estimation of distance travelled by animals in the field are less widely acknowledged or understood (Hebblewhite & Haydon 2010). Using detailed movement paths simulated with speed and tortuosity parameters estimated from camera trap footage, we were able to quantify the magnitude of this bias for a range of terrestrial vertebrates in a neotropical forest. We found a marked effect of sampling frequency on estimates of distance travelled. For example, with the traditional approach to

day-range calculation, a typical sampling frequency of two fixes h^{-1} could underestimate the true distance travelled by over 80%, with marked uncertainty over the magnitude of underestimation.

Given these results, we expect previous estimates of distance travelled - based on intermittent fixes below the frequency we found to be reliable for estimating day range - to be biased, with worrying implications for the accuracy of some conclusions. For example, Kowalczyk, Zalewski & Jedrzejewska (2006) present an interesting comparison of daily movement distances across European badger populations, showing 6-fold variation between the extremes. However, the sampling frequencies used to derive these estimates varied between 1 and 12 fixes h^{-1} . Assuming tortuosity of movement similar to that in our study species, our results suggest that the true variation in distance travelled between populations is probably closer to 14-fold, more than twice that suggested by the uncorrected comparison. This elucidates the importance of accounting for sampling frequency in calculations of distance travelled. We therefore urge future studies of travel distance based on intermittent fixes to report the sampling frequency used and preferably also to archive the raw location data for future reanalysis (see for example <http://www.movebank.org>).

We also found significant variation among species in turning behaviour, with smaller species in our sample tending to show more tortuous movement paths, hence greater underestimation of distance travelled at a given sampling frequency. We expect that this is likely to be a general result, arising because smaller species have shorter stride lengths, and thus use space at a finer spatial scale (Haskell, Ritchie & Olf 2002). Underestimation is therefore likely to scale with body mass, thus potentially confounding the scaling relationship found between body mass and telemetry-based estimates of distance travelled (Garland 1983; Carbone *et al.* 2005). Systematic bias in estimates of distance travelled across species may have important implications for our understanding of animal space use (Jetz *et al.* 2004; Carbone *et al.* 2005; Carbone, Teacher & Rowcliffe 2007) and requires further attention.

Although the problem of underestimation of distance travelled from tracking data is well known, few studies have attempted to quantify or correct for the magnitude of underestimation. Two examples from Poland (Zalewski, Jedrzejewski & Jedrzejewska 1995; Musiani, Okarma & Jedrzejewska 1998) compared radiotracking data with high-resolution movement data obtained by tracing foot prints that radiocollared individuals left in snow and found that radiolocations with a frequency of 4 h^{-1} produced apparent distances 55% and 21% lower than the actual distance (respectively, pine marten *Martes martes*, Zalewski, Jedrzejewski & Jedrzejewska 1995; and wolf *Canis lupus*, Musiani, Okarma & Jedrzejewska 1998), which are smaller biases than those we found. This is likely due to less tortuous movement in these species. In particular, Musiani, Okarma & Jedrzejewska (1998) pointed out that the wolves frequently followed relatively linear features, such as roads and rivers, and that more tortuous movements during social interactions or around kills were probably

underrepresented in their data. In contrast, our field data represent a random sample of all forms of movement behaviour and show very little evidence of travel along linear features. These comparisons emphasise the need for a better understanding of variation between species, habitats and environmental conditions in tortuosity of movement.

Measuring distance travelled through visual tracking of focal animals, as widely applied to primates (Wrangham, Gittleman & Chapman 1993) and occasionally to other species (e.g. van Gysegem 1984; Schradin 2006), has the potential to provide sufficiently fine-scale information on movement paths. Here, however, great care is needed to ensure that every small-scale movement is traced. Without such rigour, estimated distance becomes sensitive to the scale at which movements are effectively recorded in the field, and this will be hard to standardise. Without such standardisation, comparisons between distance travelled estimates are likely to reflect differences in the approach to measurement as much as any true underlying differences. This potential pitfall appears not to be widely appreciated among authors publishing results based on these methods.

Several authors have suggested that underestimation of distance travelled in telemetry studies could be avoided by increasing fix frequency (Mills, Patterson & Murray 2006; Johnson & Ganskopp 2008; Loneragan, Fedak & McConnell 2009), an approach made possible, in principle, by the advent of GPS tracking (Hebblewhite & Haydon 2010). However, our results suggest that many fixes per minute would be necessary to achieve reasonable insensitivity to sampling frequency, at least for species showing similar turning patterns to those investigated here. Such high frequencies are probably unrealistic at present, as they bring their own problems. First, high power usage leading to short tag lifespan (Ganskopp & Johnson 2007; Hurford 2009; but see Brown *et al.* in press). Second, typical GPS accuracy for intermittent location estimates is currently within 10–20 m, and at high sampling frequencies this degree of error leads to substantial upward bias in both apparent turning angles (Hurford 2009) and apparent distance travelled (Frair *et al.* 2010). While it is possible that more frequent GPS fixes might improve location accuracy (R. Kays, unpublished data), thus reducing this bias to some degree, this possibility has not been tested in the field, and methods to deal with this problem will clearly be needed if extremely high sampling frequencies are to provide reliable estimates of distance travelled in the future.

Another possible approach, suggested by Pépin *et al.* (2004), is to correct bias in distance estimation by subsampling positions from standard telemetry data at a range of lower sampling frequencies and fit an asymptotic curve to the relationship between apparent distance and sampling frequency. The estimated asymptote is then assumed to represent true distance travelled. Unfortunately, the hyperbolic curve suggested by Pépin *et al.* (2004) fits real data very poorly, giving an asymptote substantially lower even than the apparent distance derived from the actual sampling frequency used. Our Equation 11 describes a possible alternative curve that fits trends in apparent distance with sampling frequency

extremely well (Fig. 1), but also shows why the principle of attempting to find an asymptote is problematic. On log-transformed data, the relationship at low sampling frequencies is initially linear, giving a relationship on untransformed axes that is decelerating but not asymptotic. As a result, if data are available only within this range of sampling frequencies, as is typical of telemetry studies, they contain no information about the true distance travelled. Although the relationship eventually does approach an asymptote, the sampling frequency needed to identify the asymptote would be more than one per minute for the species studied here. This extrapolation approach could be more realistic for animals that have substantially less tortuous movement paths (likely to be true for larger species, and those living in more open habitats), but we cannot currently speculate on how reliable it would be in any given case. Further analysis is needed, based on a broader understanding of the variation in tortuosity that occurs in animal movement.

Finally, while we used camera traps in this study to obtain data with which to highlight the problem of bias in distance measurement, they are also a potential solution to the problem themselves, as the simulations illustrated in Fig. 4 provide correction factors with which to adjust apparent distances travelled derived from telemetry data. For example, given proportional accuracy of 0.16 for ocelot at a sampling frequency of 1 fix per hour, we could assume that the true distance travelled per unit time would be around six times greater than the apparent distance derived from a telemetry study of this intensity. This is likely to be most reliable where simulation parameters are estimated for the same population from which the telemetry data come; however, it may also be valuable as a way of approximately correcting telemetry-derived distances across populations.

In either case, a degree of caution is needed. First, there may be relevant movement processes that we have not been able to model in our simulations because they take place at a scale that we cannot observe using camera traps. In particular, movement within a home range is a fundamental process in many species, with important implications for the approach developed here. Specifically, if ranging behaviour is governed by some form of biased or memory-based random walk (Börger, Dalziel & Fryxell 2008) rather than the unbounded correlated random walk assumed here, underestimation of distance travelled is likely to be greater than our results suggest and increasingly so as sampling frequency decreases. Second, if applying a correction factor derived from one population to telemetry data from another, it must be recognised that tortuosity of movement and home range usage patterns are likely to vary between populations, potentially confounding the application. In practice, it may be that variation in movement behaviour between populations of the same species living in similar habitats will be minimal, in which case application of the approach across populations should be reasonably reliable. However, we are not aware of any existing studies that shed light on the degree of variability in relevant movement parameters between populations, highlighting a need for synthesis in this area.

Acknowledgements

We thank Nadia Sitas, Anthony Turner, Daniel Rasmussen and Lennart Suselbeek for assistance in the field, as well as the Smithsonian Tropical Research Institute, and especially Oris Acevedo, for logistical support. We also thank four reviewers for their extremely diligent and constructive comments. Required permits for the work described in this article were obtained from the Autoridad Nacional del Ambiente, Panama, and the Smithsonian Tropical Research Institute. The work was funded by NSF (DEB 0717071) and the British Ecological Society. PAJ was funded by the Netherlands Foundation of Scientific Research (NWO-ALW 863-07-008).

References

- Akaike, H. (1973) Information theory as an extension of the maximum likelihood principle. *Second International Symposium on Information Theory* (eds B.N. Petrov & F. Csaki), pp. 267–281. Akademiai Kiado, Budapest.
- Benhamou, S. (2004) How to reliably estimate the tortuosity of an animal's path: straightness, sinuosity, or fractal dimension? *Journal of Theoretical Biology*, **229**, 209–220.
- Börger, L., Dalziel, B.D. & Fryxell, J.M. (2008) Are there general mechanisms of animal home range behaviour? A review and prospects for future research. *Ecology Letters*, **11**, 637–650.
- Bovet, P. & Benhamou, S. (1988) Spatial analysis of animals' movements using a correlated random walk model. *Journal of Theoretical Biology*, **131**, 419–433.
- Bowman, F. (1958) *Introduction to Bessel Functions*. Dover, New York.
- Brown, D.D., Lapoint, S.D., Kays, R.W., Heidrich, W., Kummeth, F. & Wikelski, M. (in press) Accelerometer-informed GPS telemetry: reducing the tradeoff between resolution and longevity. *Wildlife Society Bulletin*. doi: 10.1002/wsb.111.
- Burnham, K.P. & Anderson, D.R. (2002) *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*, 2nd edn. Springer, New York.
- Carbone, C., Pettorelli, N. & Stephens, P.A. (2011) The bigger they come, the harder they fall: body size and prey abundance influence predator–prey ratios. *Biology Letters*, **7**, 312–315.
- Carbone, C., Teacher, A. & Rowcliffe, J.M. (2007) The costs of carnivory. *PLoS Biology*, **5**, e22.
- Carbone, C., Cowlshaw, G., Isaac, N.J.B. & Rowcliffe, J.M. (2005) How far do animals go? Determinants of day range in mammals. *American Naturalist*, **165**, 290–297.
- Codling, E.A. & Hill, N.A. (2005) Sampling rate effects on measurements of correlated and biased random walks. *Journal of Theoretical Biology*, **233**, 573–588.
- Codling, E.A., Plank, M.J. & Benhamou, S. (2008) Random walk models in biology. *Journal of the Royal Society Interface*, **5**, 813–834.
- Ferger, W.F. (1931) The nature and use of the harmonic mean. *Journal of the American Statistical Association*, **26**, 36–40.
- Frair, J.L., Fieberg, J., Hebblewhite, M., Cagnacci, F., J., D.N. & Pedrotti, L. (2010) Resolving issues of imprecise and habitat-biased locations in ecological analyses using GPS telemetry data. *Philosophical Transactions of the Royal Society B-Biological Sciences*, **365**, 2187–2200.
- Fryxell, J.M., Hazell, M., Börger, L., Dalziel, B.D., Haydon, D.T., Morales, J.M., McIntosh, T. & Rosatte, R.C. (2008) Multiple movement modes by large herbivores at multiple spatiotemporal scales. *Proceedings of the National Academy of Sciences*, **105**, 19114–19119.
- Ganskopp, D.C. & Johnson, D.D. (2007) GPS error in studies addressing animal movements and activities. *Rangeland Ecology & Management*, **60**, 350–358.
- Garland, T. (1983) Scaling the ecological cost of transport to body mass in terrestrial mammals. *American Naturalist*, **121**, 571–587.
- van Gysegem, R. (1984) Observations on the ecology and behaviour of the northern white rhinoceros (*Ceratotherium simum cottoni*). *Zeitschrift für Saugetierkunde*, **49**, 348–358.
- Haskell, J.P., Ritchie, M.E. & Olff, H. (2002) Fractal geometry predicts varying body size scaling relationships for mammal and bird home ranges. *Nature*, **418**, 527–530.
- 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 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.
- Hurford, A. (2009) GPS measurement error gives rise to spurious 180° turning angles and strong directional biases in animal movement data. *PLoS One*, **4**, e5632.
- Jetz, W., Carbone, C., Fulford, J. & Brown, J.H. (2004) The scaling of space use in animals. *Science*, **306**, 266–268.
- Johnson, D.D. & Ganskopp, D.C. (2008) GPS collar sampling frequency: Effects on measures of resource use. *Rangeland Ecology & Management*, **61**, 226–231.
- Jonsen, I.D., Flemming, J.M. & Myers, R.A. (2005) Robust state-space modeling of animal movement data. *Ecology*, **86**, 2874–2880.
- Kareiva, P.M. & Shigesada, N. (1983) Analyzing insect movement as a correlated random walk. *Oecologia*, **56**, 234–238.
- Kays, R., Tilak, S., Kranstauber, B., Jansen, P.A., Carbone, C., Rowcliffe, J.M., Fountain, T., Eggert, J. & He, Z. (2011) Monitoring wild animal communities with arrays of motion sensitive camera traps. *International Journal of Research and Reviews in Wireless Sensor Networks*, **1**, 19–29.
- Kowalczyk, R., Zalewski, A. & Jedrzejewska, B. (2006) Daily movement and territory use by badgers *Meles meles* in Białowieża Primeval Forest, Poland. *Wildlife Biology*, **12**, 385–391.
- Laundré, J.W., Reynolds, T.D., Knick, S.T. & Ball, I.J. (1987) Accuracy of daily point relocations in assessing real movement of radio-marked animals. *Journal of Wildlife Management*, **51**, 937–940.
- Loneragan, M., Fedak, M. & McConnell, B. (2009) The effects of interpolation error and location quality on animal track reconstruction. *Marine Mammal Science*, **25**, 275–282.
- Mardia, K.V. & Jupp, P.E. (2000) *Directional Statistics*. Wiley, Chichester.
- Mills, K.J., Patterson, B.R. & Murray, D.L. (2006) Effects of variable sampling frequencies on GPS transmitter efficiency and estimated wolf home range size and movement distance. *Wildlife Society Bulletin*, **34**, 1463–1469.
- Morales, J.M. & Elnor, S.P. (2002) Scaling up animal movements in heterogeneous landscapes: the importance of behavior. *Ecology*, **83**, 2240–2247.
- Morales, J.M., Haydon, D.T., Frair, J., Holsinger, K.E. & Fryxell, J.M. (2004) Extracting more out of relocation data: Building movement models as mixtures of random walks. *Ecology*, **85**, 2436–2445.
- Morales, J.M., Moorcroft, P.R., Matthiopoulos, J., Frair, J.L., Kie, J.G., Powell, R.A., Merrill, E.H. & Haydon, D.T. (2010) Building the bridge between animal movement and population dynamics. *Philosophical Transactions of the Royal Society B-Biological Sciences*, **365**, 2289–23201.
- Musiani, M., Okarma, H. & Jedrzejewska, W. (1998) Speed and actual distances travelled by radiocollared wolves in Białowieża Primeval Forest (Poland). *Acta Theriologica*, **43**, 409–416.
- Nouvellet, P., Bacon, J.P. & Waxman, D. (2009) Fundamental insights into the random movement of animals from a single distance-related statistic. *American Naturalist*, **174**, 506–514.
- Patterson, T.A., Thomas, L., Wilcox, C., Ovaskainen, O. & Matthiopoulos, J. (2008) State-space models of individual animal movement. *Trends in Ecology & Evolution*, **23**, 87–94.
- Pépin, D., Adrados, C., Mann, C. & Janeau, G. (2004) Assessing real daily distance traveled by ungulates using differential GPS locations. *Journal of Mammalogy*, **85**, 774–780.
- R Development Core Team (2011) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Reynolds, T.D. & Laundré, J.W. (1990) Time intervals for estimating pronghorn and coyote home ranges and daily movements. *Journal of Wildlife Management*, **54**, 316–322.
- Rowcliffe, J.M., Carbone, C., Jansen, P.A., Kays, R. & Kranstauber, B. (2011) Quantifying the sensitivity of camera traps: an adapted distance sampling approach. *Methods in Ecology and Evolution*, **2**, 464–476.
- Schradin, C. (2006) Whole-day follows of striped mice (*Rhabdomys pumilio*), a diurnal murid rodent. *Journal of Ethology*, **24**, 37–43.
- Turchin, P. (1998) *Quantitative Analysis of Movement: Measuring and Modeling Population Redistribution in Animals and Plants*. Sinauer, Sunderland, MA.
- Wrangham, R.W., Gittleman, J.L. & Chapman, C.A. (1993) Constraints on group-size in primates and carnivores - population-density and day-range as assays of exploitation competition. *Behavioral Ecology and Sociobiology*, **32**, 199–209.
- Zalewski, A., Jedrzejewski, W. & Jedrzejewska, B. (1995) Pine martens home ranges, numbers and predation on vertebrates in a deciduous forest Białowieża National Park, Poland). *Annales Zoologici Fennici*, **32**, 131–144.

Received 20 June 2011; accepted 14 February 2012
Handling Editor: Emmanuel Paradis

Supporting Information

Additional Supporting Information may be found in the online version of this article.

Fig. S1. Proportional underestimation (ratio of apparent to true distance travelled), plotted against sampling frequency for 10 terrestrial mammal species on Barro Colorado Island, Panama, with varying degrees of autocorrelation in speed of movement (correlation coefficient denoted r , colour coded as indicated). The parameters α , β and γ

in eqn 11 were estimated by fitting curves to data generated from simulations of the species-specific random walk models.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.