Appendix C. Modeling movement based on empirical data for use in the agent-based model DeerLandscapeDisease

1. Methods

1.1. Empirical data on movement

We obtained empirical movement data through field studies from 2002-2009 at two different study sites: southern and east-central Illinois (See Fig. 1 and Table 1 in main manuscript). The southern Illinois study site was an exurban setting approximately 4 km southeast of Carbondale, Illinois, USA (37° 42′14′′N, 89° 9′2′′E), whereas the east-central Illinois study site was a predominantly agricultural setting around Lake Shelbyville, Illinois, USA (39° 24' 30" N, 88° 46 '40" W). At both study sites, we captured female white-tailed deer and fitted them with GPS collars (from 2002-2006 in Southern Illinois and from 2006-2009 east-central Illinois). For individual and group movement modeling, we used location data from GPS collars deployed on 41 deer. We identified 2 pairs of deer at the southern Illinois study site and 2 pairs of deer at the east-central Illinois study site as being in the same social groups as their movements were highly correlated (Schauber et al., 2007). When analyzing data for individual movement, we included only the one deer with the most GPS locations from each group. To account for seasonal variations in behavior, we separated location data into 4 seasons pertinent to deer biology: gestation (1 Jan - 14 May), fawning (15 May - 31 Aug), prerut (1 Sep - 31 Oct), and rut (1 Nov - 31 Dec). For all our analyses, we only used locations 2 hours apart in order to compare movement models across seasons. The landscape, capture procedures, collared deer, and deer population in the southern and central Illinois study areas are thoroughly described elsewhere (Kjær et al., 2008; Schauber et al., 2015, 2007; Storm et al., 2007).

1.2. Individual movement

We used semi-hourly GPS location data from individual deer and season (Schauber et al., 2015). For each dataset (deer and season), the basis of our analysis was to predict each turn angle (θ_t) and step length (s_t) at time t as a function of the animal's displacement (d_t) from the home range centroid (Fig. C1). The centroid location was calculated simply as the averages of UTM x and y coordinates for that animal in each season and year (if data were collected from the same deer in the same season in >1 year). Turn angle was modeled with the wrapped Cauchy distribution:

$$f(\theta_t|\mu_t, \rho_t) = \frac{1}{2\pi} \times \frac{1-\rho_t^2}{1-\rho_t^2 - 2\rho_t \cos(\theta_t - \mu_t)} \qquad \text{Eq. 1}$$

where μ_t is the central angle of the distribution and ρ_t is the mean cosine of deviations from μ_t. The parameter ρ_t controls the variance in turn angles, where $\rho_t = 0$ indicates a uniform distribution of turn angles over {- π, π}, and $\rho_t = 1$ indicates zero variance in turn angles. Negative values of ρ_t are possible, and are equivalent to positive ρ_t with the central angle opposite (i.e., $\mu_t \pm \pi$ radians).

For a correlated random walk (CRW) with $\mu_t = 0$, negative ρ_t implies negative temporal correlation of step directions, i.e., a tendency to reverse direction more often than not. We fitted 5 variants of the wrapped Cauchy model (Table C1): simple (μ_t and ρ_t constant for each deer and season), simple return (where $\mu_t = \theta_{ct}$, the turn angle that would take the animal toward the centroid), converging angle (μ_t converges from 0 (CRW) to θ_{ct} (biased random walk, BRW) as d_t increases), a generalization of Frost et al.'s (2009) FOCUS model ($\mu_t = \theta_{ct}$ and ρ_t converges to a greater value as d_t increases), and both converging (μ_t converges from 0 to θ_{ct} and ρ_t converges to a greater value as d_t increases).

We modeled step length (in m) with the Weibull distribution:

$$f(s_t \mid \alpha_t, \beta_t) = \frac{\alpha_t}{\beta_t^{\alpha_t}} s_t^{\alpha_t - 1} \exp(-(\frac{s_t}{\beta_t})^{\alpha_t}) \qquad \text{Eq. 2}$$

where β_t is the scale parameter and α_t is the shape parameter. We fitted 4 variations (Table C1): simple (β_t and α_t constant for each deer and season), scale changing (β_t changing linearly with d_t), shape changing (α_t changing linearly with d_t), and both changing (both changing linearly with d_t).

1.3. Group movement

To model group movement, we quantified correlation of movements from empirical location data to identify pairs of female deer in a group together in each of our study sites (Schauber et al., 2007). For each such within-group pair, we calculated the distance between simultaneous locations at each time (x_t). Because we only had 2 deer pairs per study site, the distance data for both study sites were pooled. We then fitted exponential distributions (Eq. 1) to seasonal within-group distances to estimate λ values for each season.

$$f(x) = \lambda e^{-\lambda x} \qquad \text{Eq. 1}$$

A grouping deer in the model would then move with a random direction and a step length drawn from the above fitted exponential distributions.

2. Data analysis

2.1. Individual movement

We used Akaike's Information Criterion (AIC) and Akaike weights (Anderson et al., 2000) to assess support from each dataset for each alternative model. When the best model was determined for turn angle and step length, we tested for main and interactive effects of study site and season on fitted values of each parameter, using generalized linear mixed models (glmer from lme4 package (Bates et al., 2015)) in R.3.5.2 (R Development Core Team, 2018) with deer ID as the random effect as well as pairwise comparisons (Tukey's HSD) to test for differences in model parameters between study sites and seasons using R.3.5.2 (R Development Core Team, 2018). To approach normality, the parameters β_0, γ_μ, and γ_ρ were log transformed, and we used a Gaussian error distribution and identity link in glmer. Values of ρ_0 potentially ranged from -1 to 1, but resembled a beta distribution when rescaled as $(\rho_0 + 1)/2$ so we used this rescaled beta distribution for errors with a logit link function in glmer to test for site and season effects on ρ_0. We found a fitted value of $\rho_\infty = 1.0$ for many datasets (see Results), so we used mixed-model logistic regression in glmer (binomial error distribution and logit link function; again with deer ID as random effect) to test whether the proportion of cases where $\rho_\infty = 1.0$ differed between study sites and seasons.

To measure how well the resulting BRW models were able to reproduce the large-scale home range patterns of deer, we simulated a 5000-step movement path for each deer and season, using its fitted parameters for the best supported turn angle and step length models. Then we compared deciles (10^{th} percentile, 20^{th} percentile, etc.) of the resulting distribution of displacement values with those from the empirical data for that deer and season, to provide a rough comparison of home range shape. Simulations were conducted in R.3.5.2 (R Development Core Team, 2018).

2.1. Group movement

We used the package MASS (Venables and Ripley, 2002) in R.3.5.2 (R Development Core Team, 2018) to fit seasonal exponential distributions to the distance data.

3. Results

3.1. Individual movement

Over all deer and seasons, the turn-angle model that most frequently had the greatest support (lowest AIC) was “FOCUS”, followed by “Both Converging”; other turn-angle models received much less support (Table C2). Cumulative Akaike weight (summed over datasets) was also slightly greater for the “FOCUS” model than both converging, and much lower for other turn-angle models (Table C2). In both study areas, support for the “Both Converging” model was strongest during the gestation period and weakest during fawning. Given their similarly high levels of support, we present patterns of fitted parameter estimates under the “FOCUS” and “Both Converging” models for all deer and seasons. However, the distributions of turn-angle parameters present in both models (ρ_0, ρ_∞, and γ_ρ) for each study area and season were very similar between the “Both Converging” and “FOCUS” models (Fig. C2), so we present statistical comparisons between sites and seasons for the “Both Converging” model. In DLD, however, we chose the simpler model “FOCUS” to simulate turn angles.

For the “Both Converging” model, estimates of γ_μ were highly variable, but most were >0.01, which implies that mean turn angle converged halfway to θ_{ct} at distances of <70 m (-ln(0.5)/0.01; Fig. C2, Fig. C3). Estimates of γ_μ were similar between study sites ($F_{1,89} = 2.22$, P = 0.14) and seasons ($F_{3,89} = 2.30$, P = 0.083). Fitted values of ρ_0 were predominantly negative (Fig. C2), indicating that deer near their home-range centers tended to reverse direction or move away from the center, and did not appear to differ between study sites ($F_{1,89} = 0.01$, P = 0.93) but differed between seasons ($F_{3,89}$==3.98, P = 0.0099), primarily due to ρ_0 having a higher value for the gestation season compared to the other 3 seasons (Fig. C2). Nearly all fitted values of ρ_∞ were >0.4, and most were at or near 1.0 (Fig. C2), indicating that deer far from their home-range centers had turn angles strongly concentrated around the turn angle leading to the centroid (Fig. C3). The proportion of datasets with $\rho_\infty = 1$ was lower in the east-central Illinois site (18/59 = 0.31) than in southern Illinois (37/75 = 0.49, $F_{1,89} = 4.62$, P = 0.034), but were similar among seasons ($F_{3,89} = 2.19$, P = 0.095). Typical fitted values of γ_ρ were in the range of 0.001 to 0.01 (Fig. C2) implying that ρ converged more slowly than μ, with convergence halfway from ρ_0 to ρ_∞ occurring at displacements between ca. 70 and 700 m, respectively. Estimates of γ_ρ did not differ between study sites ($F_{1,89} = 0.65$, P = 0.42) or seasons ($F_{3,89} = 2.30$, P = 0.083).

Model selection for step lengths also yielded two competitive models: "Scale Changing" and "Both Changing" (Table C3). The "Scale Changing" model had top support in the greatest number of datasets, followed by "Both Changing", whereas the "Simple" and "Shape Changing" models received little support (Table C3). Akaike model weights yielded a similar pattern, with nearly 90% of cumulative weight resting on the "Scale Changing" and "Both Changing" models (Table C3). However, although the "Both Changing" model was competitive in terms of data support, fitted values of the Weibull shape slope (γ_α) in the "Both Changing" model were inconsistent in sign (38% positive) and the mean γ_α over all deer and seasons was not significantly different from zero (t = -0.4659, df = 133, P = 0.642, Fig. C4). Also, distributions of shared parameters (α_0/α, β_0, and γ_β) were generally quite similar between the two models (Fig. C4). Therefore, we present statistical analysis of fitted parameter values from the "Scale Changing" model, which was the step length model used in DLD.

Fitted values of the shape parameter (α) were close to one (Fig. C4), causing the Weibull distribution to resemble an exponential distribution (Fig. C3). We detected significant interactive effects of study site and season on α, with fitted values being elevated during the fawning season for the east-central Illinois study site ($F_{3,89} = 3.23$, $P = 0.026$). Fitted values of the Weibull scale intercept (β_0) were generally in the range of 60 to 150 m (Fig. C4). In analyzing β_0, We found a significant interaction between study site and season ($F_{3,89} = 5.12$, $P = 0.0026$), mostly due to lower values for the fawning and prerut season in the east-central Illinois study site (Fig. C4). Fitted values of the scale slope (γ_β) were nearly all positive (Fig. C4), indicating that step length increased with distance from home range centroid (Fig. C3). Values of γ_β did not differ significantly between seasons ($F_{3,89}= 2.48$, $P = 0.066$), or study sites ($F_{1,89} = 2.62$, $P = 0.11$).

Movement paths simulated on the basis of the fitted models for each deer and season produced displacement deciles that were strongly correlated with deciles from observed movement datasets (Fig. C5). Simulated paths had a somewhat greater spread of displacement values than was observed in empirical datasets. The 10^{th} percentile displacement from simulated movement paths exceeded the analogous value from the observed data by an average factor of 1.18, and that mean ratio increased to 1.57 for the 90^{th} percentile (Fig. C6). Therefore, return tendency of our actual deer was a bit stronger than was captured in our fitted models. However, it is important to note that this positive bias was much smaller than the differences in observed decile values among individual deer and seasons (Fig. C5).

3.2. Group movement

The group modeling analysis indicated that distances between group members were greatest in fawning and prerut seasons (Fig. C7), and mean distances from the fitted exponential distributions were 226 m for gestation, 302 m for fawning, 315 m for prerut and 246 m for rut seasons.

4. Fitting movement to DLD

4.1. Individual movement

To allow for stochasticity in DLD, we fitted seasonal- and landscape specific distributions to the individual parameter values obtained from the “FOCUS” (turn angle) and “Scale Changing” models (step length) for each deer within the specific season and landscape. Parameters for the “FOCUS” model were θc_t,ρ_∞,ρ_0 and γ_ρ. Whereas θc_t was just the turn angle leading to the home range center, values of ρ_0 potentially ranged from -1 to 1, but resembled a beta distribution when rescaled as $(\rho_0+1)/2$ so we used this rescaled beta distribution. Most of the fitted ρ_∞ values ranged from >0.4 to 1 and took the value of 1 for many of the data, thus for each season and landscape we calculated the fraction of 1 values. We then used the random engine in Java (Arnold et al., 2005) to create random probabilities, and if these random probabilities were smaller than the calculated fraction, ρ_∞ was set to one. Otherwise ρ_∞ was drawn from a beta distribution with repeated drawings until values were larger than 0.4. γ_ρ and ρ_∞ were negatively correlated, so in order to model γ_ρ, we ran a linear regression of log(γ_ρ) vs. ρ_∞ values over all the seasons and landscapes using R.3.5.2 (R Development Core Team, 2018), resulting in $y=-2.79\rho_\infty - 4.32$, with a MSE of 1.21. In DLD, we then drew a value for ρ_∞ as explained above, calculated y and drew values from a normal distribution with mean y and standard deviation MSE. The drawn value was then back-transformed (from log) to produce γ_ρ. See also Appendix B, Table B3 for all fitted parameters.

For the “Scale Changing” model (step length), the parameters β_0 and α_0 were drawn from log normal distributions, whereas γ_α was drawn from a normal distribution. To avoid unnaturally long or short steps, β_t was restricted to values between 10 m and 5241 m (longest observed movement). See also Appendix B, Table B3 for all fitted parameters.

At each time step, an agent conducting individual movement will move with a step length and turn angle drawn from Weibull and Wrapped Cauchy distributions specified by the above parameter distributions.

4.2. Group movement

At each time step, distances to group leader are randomly drawn from the season-specific distance distributions (Appendix B, Table B3). When conducting group movement, an agent will chose a random direction and then move to a location at a distance from the leader agent drawn from these distance distributions.

Table C1. The 9 correlated random walk models fitted to turn angles and step lengths extracted from GPS-locations from white-tailed deer in southern Illinois (2002-2006) and east-central Illinois (2006-2009).

Base Model	**Model Variant**	**Parameter equations**	**Bounds**
Turn Angle:			
Wrapped Cauchy	Simple	$\mu_t = \mu$ $\rho_t = \rho$	$\mu \in [-\pi, \pi]$ $\rho \in [-1, 1]$
	Simple Return	$\mu_t = \theta c_t$ $\rho_t = \rho$	$\rho \in [-1, 1]$
	Converging Angle	$\mu_t = \theta c_t + (\mu_0 - \theta c_t)*\exp(-\gamma_\mu * d_t))$ $\rho_t = \rho$	$\mu_0 \in [-\pi, \pi]$ $\rho \in [-1, 1]$
	FOCUS	$\mu_t = \theta c_t$ $\rho_t = \rho_\infty + (\rho_0 - \rho_\infty)*\exp(-\gamma_\rho * d_t)$	$\rho_0 \in [-1, 1]$ $\rho_\infty \in [-1, 1]$ $\gamma_\rho \in [0, \infty)$
	Both Converging	$\mu_t = \theta c_t(1-\exp(-\gamma_\mu * d_t))$ $\rho_t = \rho_\infty + (\rho_0 - \rho_\infty)*\exp(-\gamma_\rho * d_t)$	$\gamma_\mu \in [0, \infty)$ $\rho_0 \in [-1, 1]$ $\rho_\infty \in [-1, 1]$ $\gamma_\rho \in [0, \infty)$
Step Length:			
Weibull	Simple	$\alpha_t = \alpha$ $\beta_t = \beta$	$\alpha \in [0, \infty)$ $\beta \in [0, \infty)$
	Scale Changing	$\alpha_t = \alpha$ $\beta_t = \beta_0 + \gamma_\beta * d_t$	$\alpha \in [0, \infty)$ $\beta_0 \in [0, \infty)$ $\gamma_\beta \in (-\infty, \infty)$
	Shape Changing	$\alpha_t = \alpha_0 + \gamma_\alpha * d_t$ $\beta_t = \beta$	$\alpha_0 \in [0, \infty)$ $\gamma_\alpha \in (-\infty, \infty)$

Table C2. Empirical support for each turn-angle model by study site (Illinois, USA) and biological season, measured by the number of datasets for which each model received highest support (lowest AIC) and cumulative Akaike weight. Each dataset consists of semi-hourly GPS collar data from one deer in one season.

		# Datasets as Top Model					Akaike Weight				
Study Site	Season	Simple	Simple Return	Converg. Angle	FOCUS	Both Converg.	Simple	Simple Return	Converg. Angle	FOCUS	Both Converg.
Southern	Pre-rut	0	0	1	13	3	0.02	0.12	0.67	9.08	7.11
	Rut	0	1	0	13	6	0.00	0.65	0.43	9.87	9.06
	Gestation	0	0	1	7	15	0.00	0.25	0.68	6.27	15.81
	Fawning	0	0	0	13	2	0.00	0.09	0.23	9.35	5.34
East-central	Pre-rut	0	0	1	8	5	0.89	0.61	1.86	27.37	20.26
	Rut	0	0	1	10	3	0.00	0.28	0.60	6.39	4.73
	Gestation	0	0	1	7	8	0.00	0.11	0.76	6.51	6.61
	Fawning	1	0	0	14	0	0.89	0.00	0.00	8.41	3.70
Both Sites	Total (%)	0.9	0.9	4.3	73.3	20.7	1.09	1.27	3.17	50.45	44.02

Table C3. Empirical support for each step-length model by study site (Illinois, USA) and biological season, measured by the number of datasets for which each model received highest support (lowest AIC) and cumulative Akaike weight. Each dataset consists of semi-hourly GPS collar data from one deer in one season.

		# Datasets as Top Model				Akaike Weight			
Study Site	Season	Simple	Scale Changing	Shape Changing	Both Changing	Simple	Scale Changing	Shape Changing	Both Changing
Southern	Pre-rut	3	9	1	4	1.98	7.59	1.39	6.05
	Rut	0	8	2	10	0.17	6.79	1.99	11.05
	Gestation	1	15	2	5	0.72	11.18	1.59	9.51
	Fawning	1	2	3	9	1.02	3.69	1.71	8.59
East-central	Pre-rut	0	7	1	6	5.18	39.00	8.90	46.92
	Rut	1	6	2	5	0.26	4.95	0.73	6.07
	Gestation	0	10	0	6	0.53	4.76	1.02	5.70
	Fawning	1	7	0	7	0.00	6.78	0.16	7.06
Both Sites	Total (%)	7	64	11	52	4.13	40.43	7.10	48.34

Figure C1. Quantities used in modeling deer movement paths as correlated random walks. Locations at times *t*-1, *t*, and *t*+1 are indicated by black dots, connected by arrows indicating direction and distance of movement. The step length (s_t) and displacement from the home range centroid (d_t) are measured from the animal's location at time *t*. Turn angle (θ_t) is calculated relative to the bearing of the last step (from time *t*-1 to *t*), as is the turn angle that would take the animal directly toward the home range centroid (θc_t).

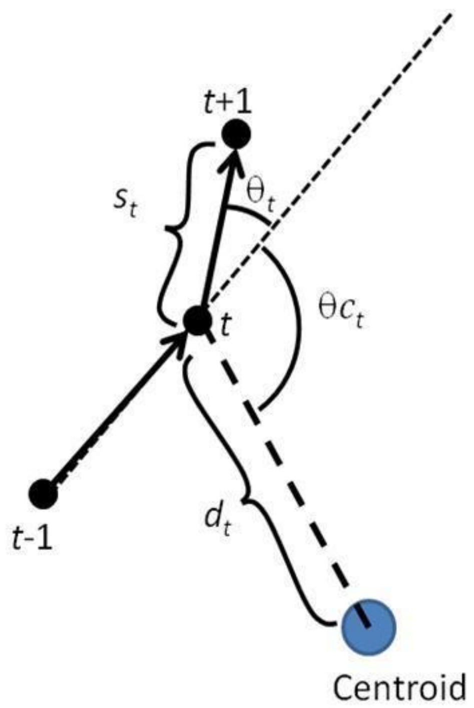


Figure C2. Box plots of fitted parameter values for the "Both Converging" (first column) and "FOCUS" model for turn angle described in the text. A) the southern Illinois study site and B) the east-central Illinois study site. ρ_0 = mean cosine of turn angles at the center of home range, ρ_∞ = mean cosine far from center of home range, γ_μ = parameter controlling rate of convergence between mean turn angle and turn angle that would take the animal directly toward the home range centroid, γ_ρ = parameter controlling rate of convergence between ρ_0 and ρ_∞. γ_μ and γ_ρ are depicted on a log scale.

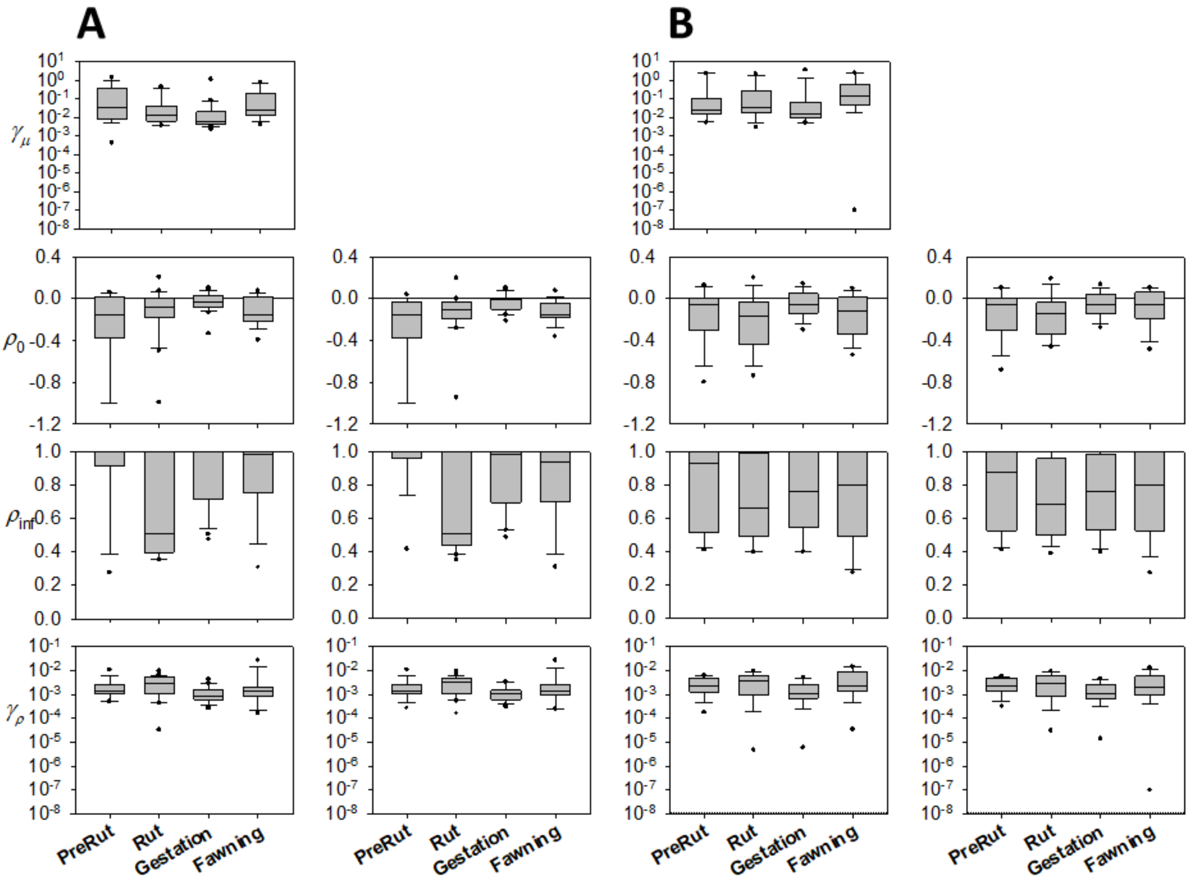


Figure C3. Examples of empirically fitted models of turn angles ("Both converging") and step lengths ("Scale changing", Table C1) depending on displacement from the home range centroid. Parameters based on maximum likelihood fitting to GPS location data from an individual white-tailed deer from east-central Illinois during the rut season. A) Central turn angle (μ_t) for wrapped Cauchy model of turn angles converging exponentially (at rate $\gamma_\mu = 0.0199$) from 0 (continue straight) to θ_{ct} (turn angle that would point animal directly toward centroid). B) Wrapped Cauchy probability density functions for turn angle deviations from μ_t at displacements ranging from 0 (deer is at the centroid) to 1000 m, with mean cosine converging exponentially (at rate $\gamma_\rho =$ 0.0047) from $\rho_0 = -0.206$ to $\rho_\infty = 0.712$. C) Weibull probability density functions for step lengths, with a
constant shape parameter ($\alpha = 0.967$) and scale parameter increasing linearly with displacement from $\beta_0 = 100.3$ with slope $\gamma_\beta = 0.331$.

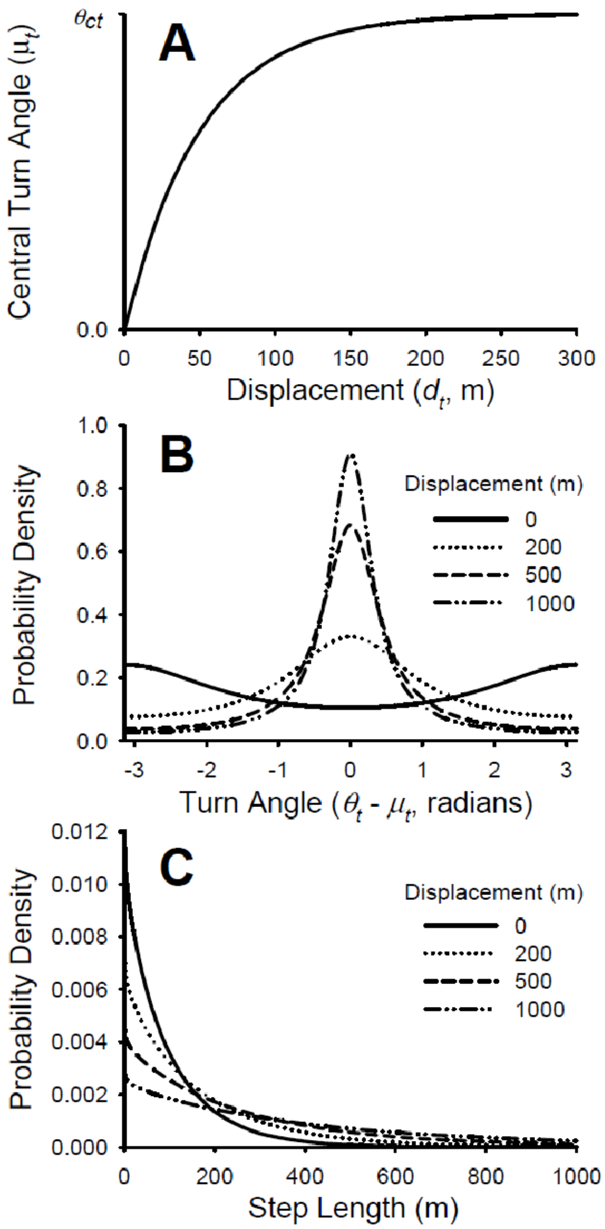


Figure C4. Box plots of fitted parameter values for the "Both changing" and "Scale changing" model for step length described in the text. A) the southern Illinois study site and B) the east-central Illinois study site. β_0 = scale intercept, γ_β = scale slope, α = shape of Weibull distribution, γ_α = shape slope.

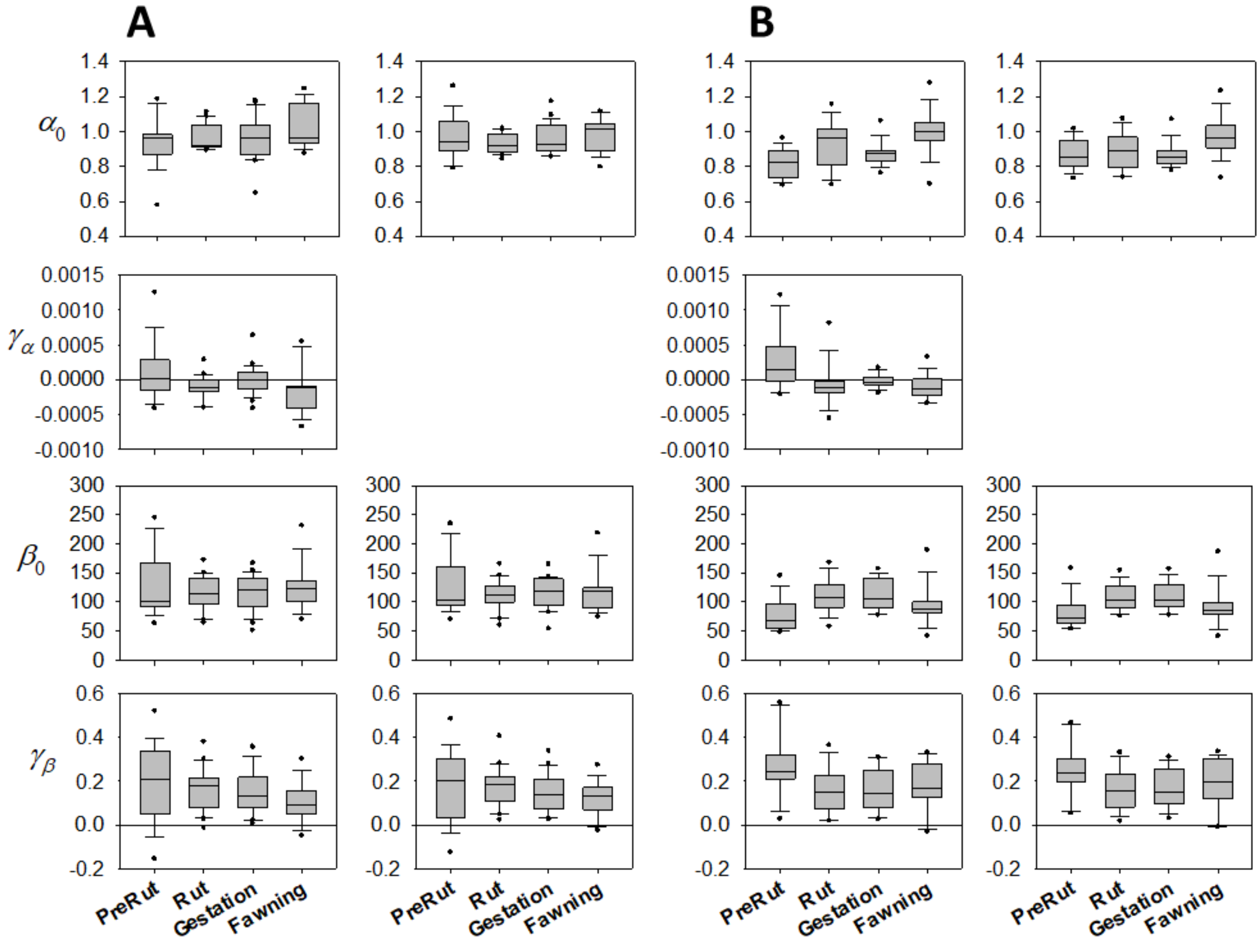


Figure C5. Relationship of percentile displacement distances (from home-range center) between observed movement paths and simulated paths based on empirically fitted random walk models with distance-dependent return tendency. Each symbol represents 10th or 90th percentile displacement distances for a single deer and season.

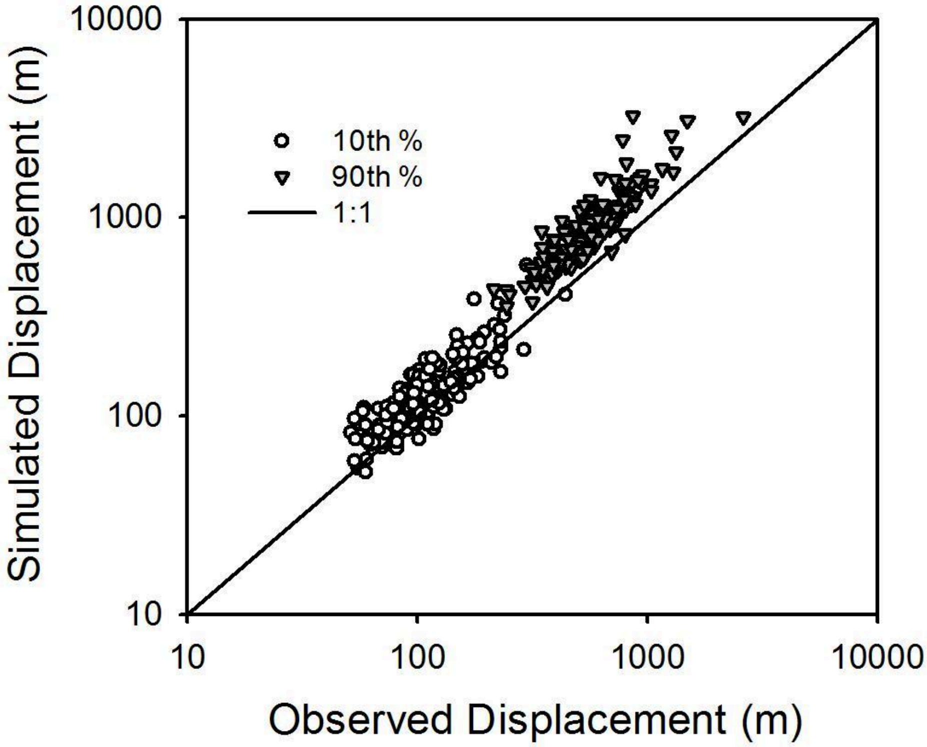


Figure C6. Ratio of 10th-90th percentile distance from home range center for observed and simulated movement data for A) the southern Illinois study site and B) the east-central Illinois study site. For the simulated data, the "Both Converging" model for turn angle and Scale changing model for step length were used. Data were pooled over seasons.

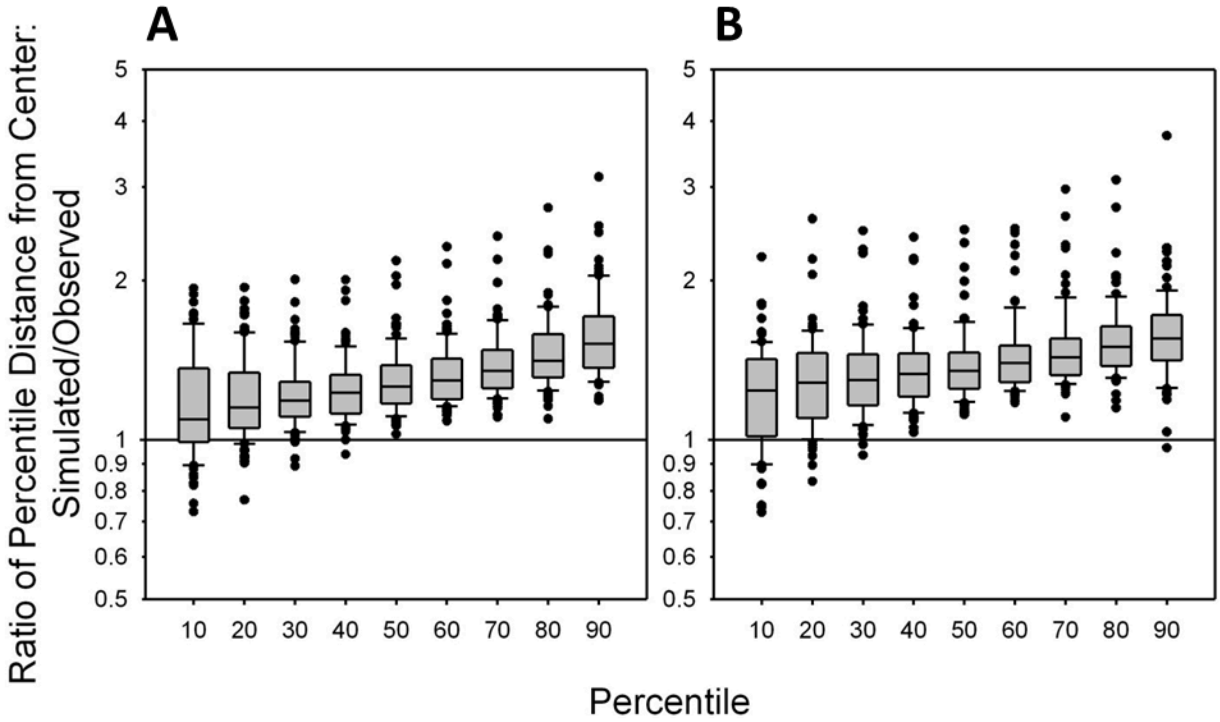
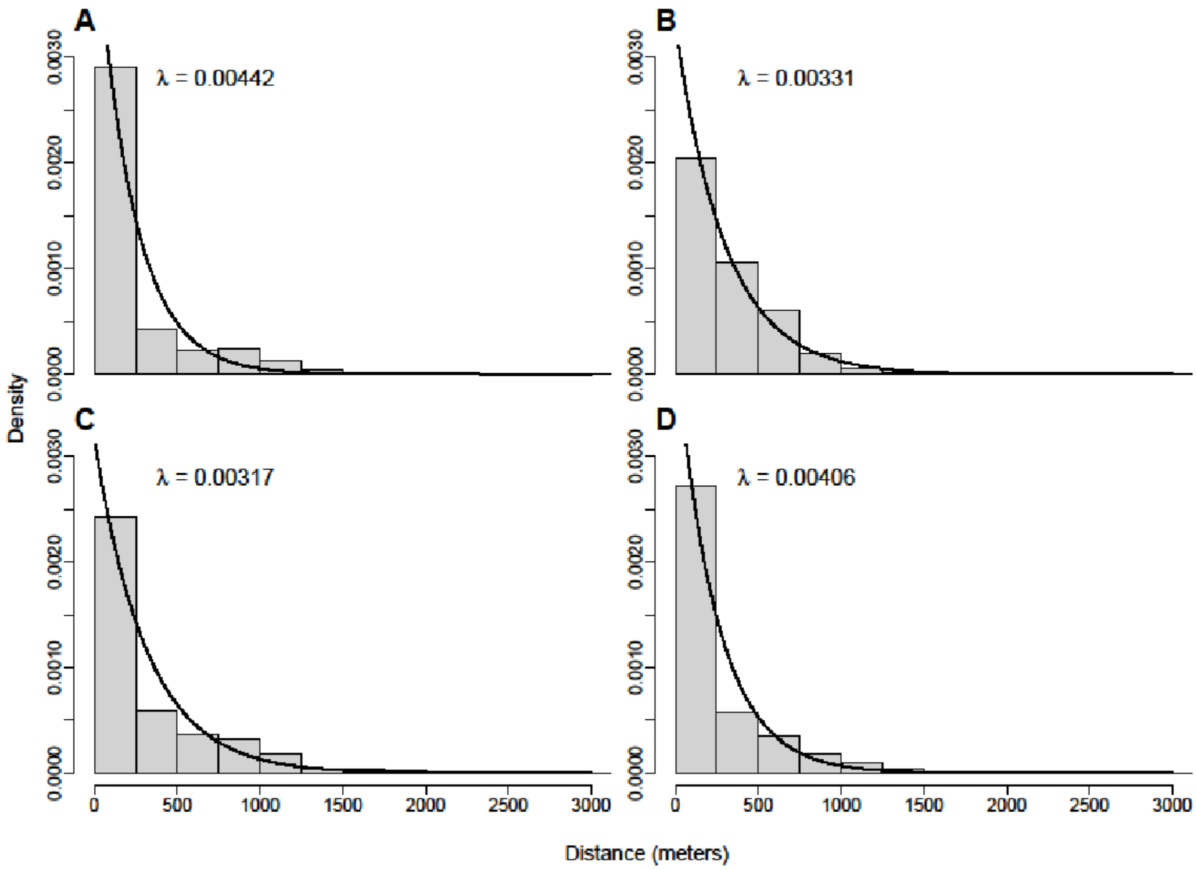


Figure C7. Exponential probability density functions fitted to group member distances from the southern Illinois and east-central Illinois study sites. A) gestation (n=7806 simultaneous location pairs), B) fawning (n=5143), C) prerut (n=2790) and D) rut (n=3978).



References

Anderson, D.R., Burnham, K.P., Thompson, W.L., 2000. Null Hypothesis Testing: Problems, Prevalence, and an Alternative. J. Wildl. Manage. 64, 912. https://doi.org/10.2307/3803199

Arnold, K., Gosling, J., Holmes, D., 2005. The Java programming language. Addison Wesley Professional.

Bates, D., Mächler, M., Bolker, B., Walker, S., 2015. Fitting Linear Mixed-Effects Models Using {lme4}. J. Stat. Softw. 67, 1–48. https://doi.org/10.18637/jss.v067.i01

Frost, C.J., Hygnstrom, S.E., Tyre, A.J., Eskridge, K.M., Baasch, D.M., Boner, J.R., Clements, G.M., Gilsdorf, J.M., Kinsell, T.C., Vercauteren, K.C., 2009. Probabilistic movement model with emigration simulates movements of deer in Nebraska, 1990–2006. Ecol. Modell. 220, 2481–2490. https://doi.org/10.1016/j.ecolmodel.2009.06.028

R Development Core Team, 2018. R: A Language and Environment for Statistical Computing. R Found. Stat. Comput.

Schauber, E.M., Storm, D.J., Nielsen, C.K., 2007. Effects of Joint Space Use and Group Membership on Contact Rates Among White-Tailed Deer. J. Wildl. Manage. 71, 155–163. https://doi.org/10.2193/2005-546

Venables, W.N., Ripley, B.D., 2002. Modern Applied Statistics with S, Fourth. ed. Springer, New York.