A framework for generating and analyzing movement paths on ecological landscapes

Wayne M. Getz[a,b,1] and David Saltz[c]

[a]Department of Environmental Science, Policy, and Management, University of California, Berkeley, CA 94720-3114; [b]Mammal Research Institute, University of Pretoria, Pretoria 0002, South Africa; and [c]Ben-Gurion University of the Negev, Be'er Sheva 84105, Israel

Edited by Ran Nathan, The Hebrew University of Jerusalem, Jerusalem, Israel, and accepted by the Editorial Board June 2, 2008 (received for review February 26, 2008)

The movement paths of individuals over landscapes are basically represented by sequences of points (x_i, y_i) occurring at times t_i. Theoretically, these points can be viewed as being generated by stochastic processes that in the simplest cases are Gaussian random walks on featureless landscapes. Generalizations have been made of walks that (*i*) take place on landscapes with features, (*ii*) have correlated distributions of velocity and direction of movement in each time interval, (*iii*) are Lévy processes in which distance or waiting-time (time-between steps) distributions have infinite moments, or (*iv*) have paths bounded in space and time. We begin by demonstrating that rather mild truncations of fat-tailed step-size distributions have a dramatic effect on dispersion of organisms, where such truncations naturally arise in real walks of organisms bounded by space and, more generally, influenced by the interactions of physiological, behavioral, and ecological factors with landscape features. These generalizations permit not only increased realism and hence greater accuracy in constructing movement pathways, but also provide a biogeographically detailed epistemological framework for interpreting movement patterns in all organisms, whether tossed in the wind or willfully driven. We illustrate the utility of our framework by demonstrating how fission–fusion herding behavior arises among individuals endeavoring to satisfy both nutritional and safety demands in heterogeneous environments. We conclude with a brief discussion of potential methods that can be used to solve the inverse problem of identifying putative causal factors driving movement behavior on known landscapes, leaving details to references in the literature.

fission–fusion | GPS | landscape matrices | random and Lévy walks | dispersal | movement ecology

The movement of all organism, but especially sentient animals, is a complex process that depends on both an individual's ability to perform various tasks and the nature of the landscape through which it moves (1–6). These tasks include the individual's intrinsic ability to move in different ways (e.g., a horse walks, trots, canters, and gallops), the individual's internal state to perform certain activities (e.g., forage, head home, flee, and seek a mate), and the individual's ability to sense its environment, remember landmarks, construct mental maps, and process information (7). Landscape variables that influence movement include topography, abiotic variables (8, 9), location of resources (10), conspecifics by gender and age, and heterospecific competitors and predators.

Emerging digital and communications technologies have refined our ability to measure movement at the resolution of fractions of seconds with concomitant spatial precision (11), while kinematical (e.g., acceleration), physiological (e.g., heart beat and temperature), and behavioral (e.g., vocalizations) information are simultaneously recorded. The internal state driving movement, however, remains largely hidden: in animals, for example, states of hunger, thirst, and fear are either inaccessible or, at best, only indirectly inferable.

The sampling frequency of movement data affects our ability to detect short-duration fundamental movement elements (FME) (6), such as a lunge versus a step taken at normal speed. Such fine-scale events are generally unrecoverable from sampling movement at intervals coarser than the duration of these events (12) [see supporting information (SI) Fig. S1]. Furthermore, activities such as foraging or heading home involve a mix of FMEs such as being stationary, ambling, and walking; and these activities may differ only in the way the FMEs are strung together. If a string of normal steps interspersed with stationary periods, for example, is on the order of minutes for both foraging and heading to a target, then movement paths sampled every 10 min during either of these two activities can be distinguished only if they produce different characteristic "distance moved in each sampling interval" distributions. This suggests that to appropriately characterize movement components of an individual's path over time, we should endeavor to identify canonical activity mode (CAM) distributions that emerge from the mix of FMEs that characterize the activity in question: i.e., CAMS are composites of the FMEs (Fig. S1), and their characteristic step size and direction of heading distributions will depend on the length of sampling intervals (Fig. S2) and scale of analysis (13).

Ideally, if one assumes that FMEs are characterized purely by a fixed speed (because they relate to biomechanical traits of individuals), then one could mechanistically construct a movement path by specifying a sequence of FMEs with a direction of heading according to know distributions of sequence lengths and correlated heading directions for particular activities. Alternatively, at fixed points in time one could specify the next location of an individual by drawing "distance moved" and "heading direction" from empirical step-size and heading distributions (1, 6, 14–17). To date, such distributions are invariably derived from sampling intervals considerably longer than the shortest FME. Thus, with most current data, it is not possible to construct distributions in terms of strings of FMEs, but only in terms of longer-lasting CAMs. Consequently, CAMs are currently the preferred place to start developing a framework for movement analysis, despite the fact that any set of CAMs is unlikely to account for all of an individual's time. In many cases, however, a reasonable tradeoff may exist in defining several CAMS that account for most of an individual's time.

Movement Paths on Featureless Landscapes

Kinds of Data. The most basic set of data that can be collected on the movement path of an individual is a sequence of positions (x_i, y_i) at points t_i, $i = 0, 1, 2, \ldots, n$: that is, the set $\mathcal{D} = \{\mathbf{u}_0, \ldots, \mathbf{u}_n\}$ with $\mathbf{u}_i = (x_i, y_i)$, where $x_i = x(t_i)$ and $y_i = y(t_i)$. For

Author contributions: W.M.G. and D.S. designed research, performed research, contributed new reagents/analytic tools, analyzed data, and wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. R.N. is a guest editor invited by the Editorial Board.

[1]To whom correspondence should be addressed. E-mail: getz@nature.berkeley.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0801732105/DCSupplemental.

convenience, we set $t_0 = 0$ but do not necessarily require all of the time intervals $\tau_i = [t_{i-1}, t_i]$ to be of equal length. In fact, in some analyses, the focus is on the distribution of waiting times τ_i associated with the events of first being in position (x_{i-1}, y_{i-1}) and then next in position (x_i, y_i) (14). These data can be transformed into polar coordinates (6) to obtain a set of vectors $\mathbf{z}_i = (t_i, d_i, \theta_i)$ with

$$d_i = \sqrt{(x_i - x_0)^2 + (y_i - y_0)^2}$$

and

$$\theta_i = \arctan\left(\frac{y_i - y_0}{x_i - x_0}\right).$$

A considerable body of diffusion and stochastic process theory exists to analyze such data in the context of featureless landscapes, including uncorrelated and correlated random walks (14–17) and super- and subdiffusive Lévy walks (18–21) and Lévy modulated correlated walks (22).

Brief Review of Random Walks and Diffusion. Consider the distribution of waiting times τ_i, velocities v_i, and absolute displacements d_i associated with a set of displacement data $\mathcal{D}$. If the distributions of both waiting times and velocities have finite mean and variance, then the stochastic process associated with the data is said to be diffusive. From theory (16–20), this implies that over an ensemble of K sets $\mathcal{D}_k, k = 1, \ldots, K$, where each set is one realization of the same stochastic process, the mean-square displacement (msd) ψ_i of the series d_{ik} averaged over all K sets is asymptotically linear, that is $\psi_i = (1/K)\sum_{k=1}^{K} d_{ik}^2 \sim t$. By contrast, if upon plotting this relationship we find that $\psi_i \sim t^p$ for $p > 1$ or < 1, then the walk is respectively referred to as superdiffusion or subdiffusion. As we demonstrate below, using simulated data drawn from a modified Pareto power law distribution (23) (Fig. S3) and plotted on a log–log scale, such plots can be misleading if time is not sufficiently large. The reason is that, initially, the relationship is affected by the fact that the distance moved in the first sampling interval is controlled by the actual step-size distribution, but from the second step onwards the distance moved from the origin is now also affected by turning angles.

Movement Pathways and Diffusion. Over the past decade, analyses of the movement paths of several organisms, including albatrosses (24) and spider monkeys (14), have concluded that the associated movement processes are superdiffusive, although a re-analysis of these data refute this finding (ref. 1, but see ref. 25). A possible source of error in estimating p in the msd relationship $\psi_i \sim t^p$ (Fig. 1) is t must be sufficiently large for the asymptotic value to emerge. Another source of error is that superdiffusion predicts fractal looking movement paths (14, 20, 24). Using one of several different methods to estimate the fractal dimension (26) of such paths, some organisms have been pronounced as superdiffusive (14). In these organisms, however, repeated fractal patterns occur at no more than two or three particular scales and are more a feature of the way resources are distributed across the landscape than of a genuine superdiffusive process. Thus, it is important to understand how animal movement is influenced by landscape features and to assess the extent to which step-size distributions are modified by landscape heterogeneity.

We note here that because all bounded step-size distributions produce Gaussian random walks when turning angles are uncorrelated, critical information in the step-size distribution, such as multimodality arising from mixed distributions of FMEs, is lost when using the statistics of emergent global characteristics such as msd as a function of time. This stresses the importance of knowing the actual step-size distributions when deconstructing movement paths and trying to understand the causal processes creating local path structures.

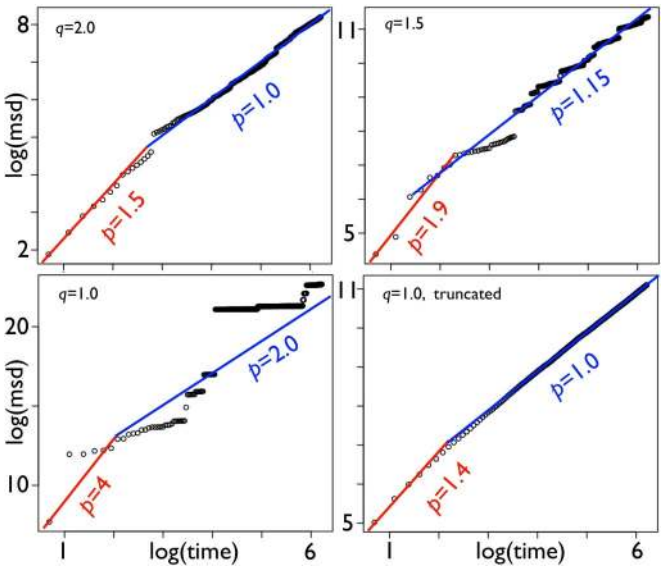

Fig. 1. The logarithm of mean-square displacements (msd) ψ versus the logarithm of time t average over 10,000 simulations are plotted for a random walk with step lengths drawn from modified Pareto distributions (*Upper Left*, $q = 2$; *Upper Right*, $q = 1.5$; *Lower Left*, $q = 1$) and directions for each step completely random. From the lines inserted by ''eye'' (red, small t; blue, large t), *Upper Left* represents diffusion ($P = 1$ for large t), and *Upper Right* and *Lower Left* represent super diffusion (respectively, $P = 1.15 > 1$ and $P = 2$ for large t). In *Lower Right*, for the case $q = 1$, the length of these excursions are truncated at a step size of 100 (biologically, an upper bound is set by the maximum velocity of the organism multiplied by the length of the time interval), which is far out in the tails of the distribution (two orders of magnitude beyond the mode; see Fig. S3). In this case, the noisy superdiffusive behavior is completely tamed even though, initially, it looks superdiffusive ($\psi \sim t^{1.4}$ for $t \leq 2$).

Movement Paths on Structured Landscapes

FMEs and CAMs. The framework presented here is formulated in the context of a group of N known individuals indexed by k. This specificity allows us to account for the following: (*i*) species, gender, and age-specific differences; (*ii*) unique memory and knowledge of landscape; and (*iii*) cues and vectors individuals use to select a new position on the landscape. Further, we assume that each individual has n_{m}^k FMEs (Table 1), each with its own characteristic speed $s_j^k, j = 1, \ldots, n_{\mathrm{m}}^k$, from which its movement track is generated. The set of FMEs constitute the basic motion capacity $\mathbf{\Omega}^k$ (7). Furthermore, in any segment of the movement path (Fig. S1), these FMEs are mixed in various proportions to constitute a set of A_r^k CAMs (Table 1), $r = 1, \ldots, n_{\mathrm{a}}^k$ best characterized by distributions of speeds (equivalently distances) with means and standard deviations $\bar{s}_{\mathrm{r}}$ and σ_{π} for each r (Fig. S2). We note, however, that when the sample intervals $\tau_i^k = \tau^k$ are fixed over all intervals i, for reasons discussed below only the standard deviation (and not the mean) depends on sampling frequency $1/\tau^k$.

As an example, $\mathbf{\Omega}^k$ in horses has been defined in terms of five FMEs or gaits (27) of increasing speeds: stationarity ($s_1 = 0$), walking (a four beat gait with $s_2 \approx 4$ mph), trotting (a two beat gait with $s_3 \approx 8$ mph on average), cantoring (a three beat gait for which $s_4 > s_3$), and galloping (a four beat gait for which $s_5 \approx 25$–30 mph, varying across horses). Beyond these natural gaits, some horses have been bred to implement so-called "ambling gaits" that provide a smoother ride, but not all horses can execute these, thereby

Table 1. Definition of parameters and variables

Frames (indices)	Mathematical objects	Descriptions
Time (i)	t_i τ_i τ	Current time, variable and fixed inter-interval size
Random walks	d_i θ_i ψ_i	Step size, heading, mean-square displacement (msd)
	v_i w_i	Linear and radial velocities at time t_i
Dataset $\mathcal{D}_k$ of individuals (k)	N $\mathbf{u}_i^k = (x_i^k, y_i^k)$	Number, Cartesian location for each k updated each t_i
Fundamental movement elements (FMEs) (j)	n_{m}^k s_j^k Ω^k	Number, characteristic speed, "capacity" for each k
Canonical activity modes (CAMs) (r)	n_{a}^k $A_{\pi r}^k$ $\mathcal{A}_{\pi r}^k$	Number, 1D and 2D distance distributions
Internal state of individual k	$\mathbf{w}_i^k$ w_{ir}^k	Vector and elements ($r = 1,\cdots, n_{\mathrm{a}}^k$) updated each t_i
Landscape grid $C^{\alpha\beta}$ (α rows, β columns)	$\mathrm{c}_i^{\alpha\beta} = (c_{i1}^{\alpha\beta}, \ldots, c_{in_c}^{\alpha\beta})'$	Vector state of grid cell (α, β) updated each t_i
Landscape modifier matrices (LMMs)	L_{ir}^k $\ell_{ir,\alpha\beta}^k$	Modifier matrix and elements on (α, β) updated each t_i
Ideal movement distributions (fixed τ)	$G_{\pi i}^k(s)$ $\mathcal{G}_{\pi i}^k(\alpha, \beta)$	Continuous 1D and binned 2D representations
Realized movement distributions (fixed τ)	$M_{\pi i}^k(\alpha, \beta)$	2D histograms for each k updated each t_i

providing an example of selection at work on the motion capacity component Ω of our guiding conceptual model (7).

For each of the n_{a}^k, CAMs associated with individual k, the distribution of speeds (equivalently distances or "step sizes") associated with a particular CAM is affected by the value of τ: if τ is small an individual will generally be in one movement mode or another in the mix that constitutes the activity, whereas if τ is large an individual will more likely be executing some mix of modes over one interval (Fig. S2). Thus, even though the proportion of different movement mode events used to construct the movement tracks may be quite stable with regard to a specific activity, the directions of heading from one event to the next make the relationship between the length of a track and displacement quite complicated. Only for very simple cases, such as tracks generated from one type of FME, is the relationship between length-of-track and displacement easily cracked by analytical methods. For the rest, the easiest way to generate the distribution of speeds (distances) $A_{\pi r}(s)$ as a function of sample interval size τ is to use Monte Carlo simulation.

The proportions of FMEs in a particular CAM is likely to vary with changes in landscape. For example, the mix of stationary, walking, and running modes used by a foraging African antelope will differ in an open savannah compared with thick bushveld. This problem can be dealt with by assuming that the parameters of a CAM distribution depend on external landscape factors in addition to the frequency of sampling along a movement path.

Internal States and Goals. One of the goals in an analysis of empirical data is to see how cleanly a set of CAM distribution of step-size FMEs can be extracted from the movement track to explain the actual activity producing different segments of the movement track. Some segments may reflect a pure activity (e.g., foraging) while others are a mix of activities (e.g., foraging and resting) or even a compromise between competing activities such as an individual foraging as it heads to a known water source or home. To be able to account for such mixed activities, as well as assess factors that may lead to an individuals switching from one CAM to another, we must infer that the individual has a multidimensional internal state that drives the behavior (7). The current state of this driver can in turn be associated with a goal emerging from an individual's internal state, which in general will vary with time. The relationship between an individual's internal state (i.e., the vector $\mathbf{w}$—see Fig. S4) and its current goal state can be treated in various ways. One way is to construct a mapping of a continuous n-dimensional vector space to a discrete space of g goals. Another is to consider the internal state as a weighting vector

$$\mathrm{w}_i^k = (w_{i1}^k, w_{i2}^k, \cdots, w_{n_{\mathrm{a}}^k}^k)$$

that produces a goal-modified "ideal" distribution $G_{\pi i}^k(s)$ of speeds (distances) s for individual k at time $t_i = i\tau$ with the weighted sum of its n_{a}^k CAMs: i.e.,

$$G_{\pi i}^k(s) = \sum_{r=1}^{n_{\mathrm{a}}^k} w_{ir}^k A_{\pi r}^k(s).$$

In the case of organisms that have no internal mechanism for generating goals (e.g., plants), the internal state may, for example, represent elements controlling dehiscence of seed dispersal structures.

Spatial Explication. The spatial information embedded in the distribution $G_{\pi i}^k(s)$ needs to be made explicit before other relevant spatially explicit geographical and biological information can be incorporated into the movement process. The distribution $G_{\pi i}^k(s)$ is defined on $s \in [0, s_{\max}]$. On a flat structureless landscape—that is, in the absence of all cues and other directional biasing factors beyond the context of a correlated walk—an individual is equally likely to move in any direction. Thus, the ideal (i.e., featureless landscape) distribution $G_{\pi i}^k(s)$ can be given an explicit spatial dimension by rotating it around the point [0, 0] on a plane parameterized by coordinates (α, β) to obtain a radially symmetric distribution $\mathcal{G}_{\pi i}^k(\alpha, \beta)$ that has a top-half-of-a-donut-like structure (Fig. S5). The distribution can now be relocated so that its center of symmetry is the current point of location $\mathbf{u}_i^k = (x_i^k, y_i^k)$ of organism k.

Landscape Raster. We incorporate landscape features into each individual's goal as follows. First, we cover the landscape with a raster of rectangular cells $C^{\alpha\beta}$. Second, we locate the position $\mathbf{u}_i^k = (x_i^k, y_i^k)$ of each individual within the cell containing this point: in general, this allows several individuals of one or more types to be located in each cell. Also, we associate an n_{c}-dimensional external state to account for all factors relevant to the movement of each of the N individuals on the landscape, including other individuals. Third, the one-dimensional CAM distributions $A_{\pi r}^k(s)$ are used, as described above, to generate two-dimensional radially symmetric, but discretized, distributions $\mathcal{A}_{\pi r}^k(\alpha, \beta) \geq 0$ with $\sum_{(\alpha,\beta)} \mathcal{A}_{\pi r}^k(\alpha, \beta) = 1$ for all k and $r = 1, \ldots, n_{\mathrm{a}}^k$. Fourth, we incorporate the landscape effects through a set of n_{a}^k landscape modifier matrices (LMM) L_{ir}^k (i.e., specific to individual k, their activity r, and time $i\tau$) with elements $\ell_{ir,\alpha\beta}^k$ constructed from those elements of the external state vector that are applicable for the activity in question. These elements $\ell_{ir,\alpha\beta}^k$ are used to modify the CAM distributions to reflect the preference that each individual has for each of the landscape cells while involved in one of its CAMs (e.g., one cell may be the most desirable from a foraging point of view, whereas another is desirable as a target when heading for water). These LMMs are used to modify the movement distribution $\mathcal{G}_{\pi i}^k(\alpha, \beta)$ to account for

the state of the landscape, as well as individual specific information that *inter alia* relates to the location of places [remembered, sensed, and inferred through landscape cues, including the earth's magnetic field (28)] associated with heading activities, the distribution of resources across the landscape, and the location of other individuals on the landscape; and hence the elements of these matrices must be updated each time step.

The simplest way to incorporate this landscape information into the movement process is to use the activity-specific LMM elements $\ell^k_{ir,\alpha\beta}$ (Table 1) as a way to weight the terms in the idealized goal determined movement distribution $G^k_{\pi i}(s)$ to produce a composite realized movement distribution, with spatial matrix elements

$$M^k_{\pi i}(\alpha, \beta) = \frac{1}{M^k_{\pi i}} \sum_{r=1}^{n^k_a} w^k_{ir}\, \ell^k_{ir,\alpha\beta} \mathcal{A}^k_{\pi r}(\alpha, \beta)$$

for all α, β, i, and $k = 1$, where

$$M^k_{\pi i} = \sum_{\alpha,\beta} \sum_{r=1}^{n^k_a} w^k_{ir}\, \ell^k_{ir,\alpha\beta} \mathcal{A}^k_{\pi r}(\alpha, \beta)$$

normalizes the discrete elements of the probability distribution over the cells $C^{\alpha\beta}$ for each individual k and all time i.

Decision Mechanism. The final component of the movement process is how the organism selects the particular cell (α, β) that will determine its next position $\mathbf{u}^k_{i+1} = (x^k_{i+1}, y^k_{i+1})$ at time $t = (i + 1)\tau$. The simplest rule is for an organism to move to cell (α, β) with largest value $M^k_{\pi i}(\alpha, \beta)$, effectively without making any interim decision on its way to the new cell. For the case of an organism driven purely by stochastic landscape processes (such as wind), movement to the next cell can be regarded as random with the probability of selecting a particular cell (α, β) equal to the values $M^k_{\pi i}(\alpha, \beta)$. A third possibility is to select cells with a probability that is proportional to some power of $M^k_{\pi i}(\alpha, \beta)$, a solution that is intermediate between the first two if this power is >1. If this power is <1, the solution is intermediate between the second mechanism and a purely random solution that gives no weight to the relative values $M^k_{\pi i}(\alpha, \beta)$.

Food, Safety, and Fission–Fusion Dynamics

The following example illustrates the application of our framework to simulating the movement of a herd of social ungulates foraging on a heterogeneous landscape with conflicting needs to both assuage hunger and remain safe by staying close to other individuals.

In social ungulates, the existence of groups is presumably the result of safety offered by the group (29). However, membership in a group comes at a cost, namely competition for limited resources (usually food). These conflicting needs presumably drive the observed fission–fusion dynamics typical of many social ungulates (30). Thus, the goal for each individual is feeding while remaining near conspecifics, and the main internal states for this goal are levels of hunger and safety, with each individual seeking to assuage hunger while remaining safe. The CAM in this example is pure foraging, and the FMEs allow either moving between patches constrained by a maximum distance traversed in each time set, or remaining in the current patch while feeding (see *SI Text*). Navigation in our example is relatively simple: individuals move directly using visual information to a cell selected within a fixed radius interpreted as the observable range. The distribution of food patches and the location of other herd members represent the external states that a given herd-member responds to in its choice of direction. The external states

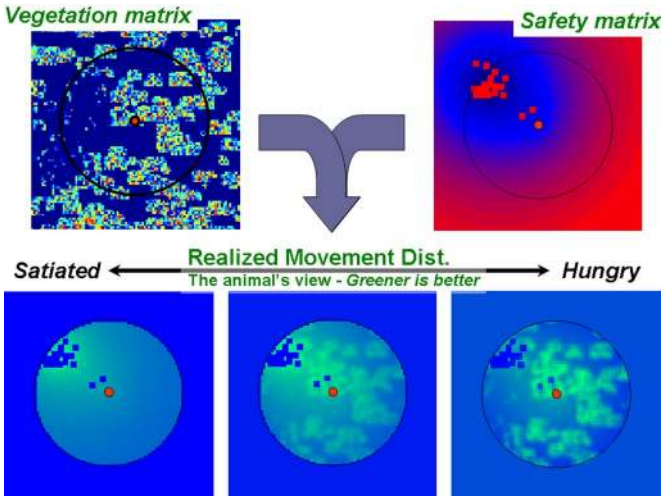

Fig. 2. Each of the five panels is an extract from a much larger mapping of the values of elements in the landscape modifier matrices (LMMs) for the vegetation and safety landscapes of the realized discretized movement distributions constructed from these matrices. The focal individual is represented by the small red squares in each of the five panels, with the positions of its conspecifics represented by other small squares in *Upper Left* and *Lower*. In the distributions represented in *Lower*, the most attractive areas are the lighter areas (but ignoring the small dark blue squares, which are just conspecific position markers for reference). The relative weighting of safety over resources ranges from safety being the only consideration (*Lower Left*), safety and resources being equally important (*Lower Center*), and resources being the only consideration (*Lower Right*). Imposed upon *Lower* is a circle representing the maximum possible movement displacement in one time step.

are combined by using LMMs to obtain the realized movement distribution from which an individual selects its next target.

Our model is spatially explicit with each grid cell, in this particular case roughly the size of an individual. For each time step an animal may perform one of three activities: (*i*) moving toward a selected patch (one cell per time step), (*ii*) feeding in a patch (consuming one unit per time step), or (*iii*) resting in a patch when it has a full gut. There are two matrices, each representing one of the external states: a food (vegetation) matrix and a safety matrix (Fig. 2 *Upper*). The vegetation matrix is static and is generated at the start of each simulation by using a combination of random procedures to create a patchy heterogeneous landscape (see *Materials and Methods* and *SI Text*). The safety matrix is a function of the spatial location of all group members (Fig. 2), risk does not vary with landscape features, and there are no visible predators. In each time step, a safety score is calculated to each cell as the sum of the inverse of the distance of all herd members to that cell (see *Materials and Methods*). The LMM is a weighted sum of the vegetation and safety matrices, where the weightings α and $(1 - \alpha)$ depend on the internal driver (state) $\alpha \in [0, 1]$ reflecting the current priorities to the individual in trading safety against hunger. When an animal stays within a patch to feed, its value of α increases and when it moves its value of α decreases so that the relative importance of food to safety oscillates up and down as the animal approaches satiation or becomes increasingly hungry (see *Materials and Methods* and *SI Text*).

By playing around with simulation parameters, it is possible to test how various movement-related questions—such as fission–fusion dynamics, subgroup structure, and trajectory patterns—are affected by patch size, gut size, and feeding strategies. For example, we explored how patch density affects fission–fusion and subgroup structure as follows: 20 animals with equal gut sizes were initially spread randomly over a 20 × 20 cell section of a 220 × 1,100 cell grid, with movement rules detailed in *SI Text*. We assumed that the vegetation in each cell is not

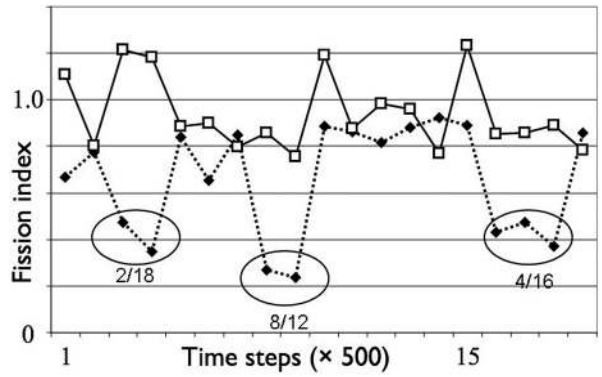

Fig. 3. Simulation of fission–fusion behavior as a function of vegetation quality. Open squares, high quality; filled diamonds, low quality. Population defined to be in a two-herd (one herd) state when the fission index is <0.5 (>0.5). See main text and *SI Text* for details.

renewable, so that depleted patches act as repulsive regions. The narrow landscape channels the herd into a directional movement along the long axis, but the short axis is sufficiently wide to enable fission events (Fig. 3).

The simulation was repeated with the same starting locations with two different vegetation matrices, each with a mean patch size of 38 cells. The first, representing "good" habitat, produced a landscape where ≈75% of the cells were positive (Fig. 3), whereas the second, representing "poor" habitat, produced a landscape with ≈65% positive cells. We ran each simulation for 10,000 time steps. Group structure was analyzed every 500 steps, starting with the 1,000th step, using cluster analysis with a two-group restriction (see *SI Text*). We then calculated the mean individual distance for the entire herd and for each subgroup. We considered a ratio of 0.5, between the sum of the two subgroup mean distances to the mean distance of the entire herd, as an indicator of fission. Simulating the herd's movement by using the "poor" habitat matrix produced three fission events compared with none for the "good" habitat (Fig. 3). All fissions were followed by fusion after 1,000–1,500 time steps. We note that the herd was generally cohesive, with fission and fusion events emerging rather than explicitly constructed, despite that fact that the movements of the individuals themselves were completely deterministic. Our approach highlights the importance of quantifying the internal drivers in understanding animal movements. In the context of our specific simulation, empirical data on giving-up densities (31) appear to be a good tool for quantifying internal drives and understanding movement patterns.

Conclusion

The framework presented here provides a mathematically detailed exposition of the conceptual model developed by Nathan *et al.* (7) that can be used for both the construction and deconstruction of the pathway of organisms on structured landscape, the former through simulations and the latter through state-space methods (32) for fitting parameters. We have not dealt in any detail with the problem of movement mode identification that is the key to the deconstruction component other than stressing that the frequency with which data are collected limits our ability to identify CAMs and their underlying FMEs (12). To undertake such an analysis is not a trivial problem: it requires computationally complex methods that can only be successfully applied to high-resolution data, but a start has been made (33–35).

Before automated GPS data collection, VHF telemetry position data were typically collected too infrequently and were also not sufficiently accurate to be useful for reconstructing pathways; but these data were suitable for constructing home ranges or types of utilization distributions at a seasonal scale, using both parametric (36–38) and nonparametric (39, 40) kernel methods as the preferred methodology. Since the mid-1990s (24) movement data have been fitted to Lévy models to evaluate the extent to which this movement is superdiffusive with, as we have discussed, mixed success (1) and also to assess the degree to which path characteristics can be fitted by correlated random walks (17) or mixtures of random walks (4), but these analyses do not explicitly incorporate landscape structure or factors.

In terms of general methods for deconstructing movement pathways, various time series and frequency domain techniques can be brought to bear on the problem under the rubric of exploratory data analysis (EDA) (41). Furthermore, stochastic differential equation methods (42) can be used to construct vector fields from data on the contemporaneous movement of many individuals, and then thin plate splines can be used to fit potential fields to these vector fields to identify regions on the landscape that are either repealing or attracting the individuals at a particular time of day (2). Recently, techniques new to the field of movement ecology, such as wavelet analysis (33) and artificial neural networks (34), are being applied to obtain insights into the effects that the internal and environmental states of a system have on movement paths. Beyond these, as the resolution of movement and landscape data improves dramatically over the next decade, we should expect to see the application of state-space estimation methods (32) that can take advantage of formulations such as ours, because our formulation permits the inclusion of detailed landscape information.

Materials and Methods

Details are elaborated on in *SI Text*.

The Vegetation Matrix. This matrix is the only component in the model that has stochastic elements. It consists of a series of patches set up using Monte Carlo methods, with a parameter controlling patch density and a beta distribution controlling patch size. The quality of resources in the center of each patch was then assigned a number at random between 1 and 10 with values declining to the edge of the patch. The resource value of each cell was reduced by a set amount in each time step for which the patch was occupied by a feeding individual. Fig. 2 *Upper Left* depicts the result of one such construction.

The Safety Matrix. For each of the cells containing an individual (i.e., focal individual), a matrix of values for all of the remaining cells was constructed. The values associated with each of these remaining cells is based on the sum of the inverse distances of all of the remaining organisms to these cells. Thus, cells of highest values are those closest in an integrative sense to the organisms as a group that excludes the focal individual (see Fig. 2 *Upper Right*).

The Navigation Matrix. All points within a fixed distance of an individual (and only these points) were regarded as selectable targets to move to next (circles in Fig. 2). An individual then moves toward the cell that has the highest value of a weighted sum of the vegetation and safety matrix values for that cell.

Feeding and Energetics. Each animal has an energy bank that determines its level of hunger and, hence, its weighting of vegetation and safety matrix values. During each time step, an animal can either feed and increase its energy bank or move and decrease its energy bank. Once an individual has consumed the vegetation locally, it moves to a patch within navigation range that maximizes its current tradeoff for resources versus safety.

ACKNOWLEDGMENTS. We thank the members of the Institute of Advanced Studies movement ecology group for ideas that we have absorbed into our text during the course of our weekly meetings over many months, and Leo Polansky for comments on the manuscript. This work was supported by the Institute of Advanced Studies at the Hebrew University of Jerusalem and by a James S. McDonnell 21st Century Science Initiative Award.

1. Edwards AM, *et al.* (2007) Revisiting Lévy flight search patterns of wandering albatrosses, bumblebees and deer. *Nature* 449:1044–1048.

2. Preisler HK, Ager AA, Johnson BK, Kie JG (2004) Modeling animal movements using stochastic differential equations. *Econometrics* 15:643–657.

3. Zhang X, Johnson SN, Crawford JW, Gregory PJ, Young IM (2007) A general random walk model for the leptokurtic distribution of organism movement: Theory and application. *Ecol Modell* 200:79–88.
4. Morales JM, Haydon DT, Frair JL, Holsinger KE, Fryxell JM (2004) Extracting more out of relocation data: Building movement models as mixtures of random walks. *Ecology* 85:2436–2445.
5. Morales JM, Fortin D, Frair JL, Merrill EH (2005) Adaptive models for large herbivore movements in heterogeneous landscapes. *Landsc Ecol* 20:301–316.
6. Root RB, Kareiva PM (1984) The search for resources by cabbage butterflies (pieris-rapae)—Ecological consequences and adaptive significance of Markovian movements in a patchy environment. *Ecology* 65:147–165.
7. Nathan R, *et al.* (2008) A movement ecology paradigm for unifying organismal movement research. *Proc Natl Acad Sci USA* 105:19052–19059.
8. Mandel JT, Bildstein KL, Bohrer G, Winkler DW (2008) The movement ecology of migration in turkey vultures. *Proc Natl Acad Sci USA* 105:19102–19107.
9. Revilla E, Wiegand T (2008) Individual movement behavior, matrix heterogeneity, and the dynamics of spatially structured populations. *Proc Natl Acad Sci USA* 105:19120–19125.
10. Wright SJ, *et al.* (2008) Understanding strategies for seed dispersal by wind under contrasting atmospheric conditions. *Proc Natl Acad Sci USA* 105:19084–19089.
11. Cooke SJ, *et al.* (2004) Biotelemetry: A mechanistic approach to ecology. *Trends Ecol Evol* 19:334–343.
12. Jerri AJ (1977) The Shannon sampling theorem—Its various extensions and applications: A tutorial review. *Proc IEEE* 65:1565–1596.
13. Fryxell JM, *et al.* (2008) Multiple movement modes by large herbivores at multiple spatiotemporal scales. *Proc Natl Acad Sci USA* 105:19114–19119.
14. Ramos-Fernández G, Mateos JL, Miramontes O, Cocho G (2004) Lévy walk patterns in the foraging movements of spider monkeys (Ateles geoffroyi). *Behav Ecol Sociobiol* 55:223–230.
15. McCulloch CE, Cain ML (1989) Analyzing discrete movement data as a correlated random walk. *Ecology* 70:383–388.
16. Bartumeus F, Da Luz MGE, Viswanathan GM, Catalan J (2005) Animal search strategies: A quantitative random-walk analysis. *Ecology* 86:3078–3087.
17. Dai X, Shannon G, Slotow R, Page B, Duffy KJ (2007) Short-duration daytime movements of a cow herd of African elephants. *J Mammal* 88:151–157.
18. Shlesinger MF, Klafter J, West BJ (1986) Lévy walks with applications to turbulence and chaos. *Phys A* 140:212–218.
19. Klafter J, Zumofen G, Shlesinger MF (1993) Lévy walks in dynamical systems. *Phys A* 200:222–230.
20. Metzler R, Klafter J (2004) The restaurant at the end of the random walk: Recent developments in the description of anomalous transport by fractional dynamics. *J Phys A Math Gen* 37:R161–R208.
21. Bartumeus F, Catalan J, Fulco UL, Lyra ML, Viswanathan GM (2002) Optimizing the encounter rate in biological interactions: Lévy versus Brownian strategies. *Phys Rev Lett* 88:097901.
22. Bartumeus F, Levin SA (2008) Fractal reorientation clocks: Linking animal behavior to statistical patterns of search. *Proc Natl Acad Sci USA* 105:19072–19077.
23. Reed WJ (2001) The Pareto, Zipf and other power laws. *Econ Lett* 74:15–19.
24. Viswanathan GM, *et al.* (1996) Lévy flight search patterns of wandering albatrosses. *Nature* 381:413–415.
25. Sims DW, *et al.* (2008) Scaling laws of marine predator search behavior. *Nature* 451:1098–1102.
26. Nams VO (1996) The VFractal: A new estimator for fractal dimension of animal movement paths. *Landscape Ecol* 11:289–297.
27. Harris SE (1993) *Horse Gaits, Balance and Movement* (Howell Book House, New York), p 178.
28. Lohmann KJ, Putman NF, Lohmann MF (2008) Geomagnetic imprinting: A unifying hypothesis of long-distance natal homing in salmon and sea turtles. *Proc Natl Acad Sci USA* 105:19096–19101.
29. Roberts G (1996) Why individual vigilance declines as group size increases. *Anim Behav* 51:1077–1086.
30. Rubenstein DI (1986) Ecology and sociality in horses and zebra. *Ecological Aspects of Social Evolution: Birds and Mammals*, eds Rubenstein DI, Wrangham RW (Princeton Univ Press, Princeton), pp 282–302.
31. Kotler BP, Brown JS, Bouskila A (2004) Apprehension and time allocation in gerbils: The effects of predatory risk and energetic state. *Ecology* 85:917–922.
32. Patterson TA, Thomas L, Wilcox C, Ovaskainen O, Matthiopoulos J (2008) State-space models of individual animal movement. *Trends Ecol Evol* 23:87–94.
33. Wittemyer G, Polansky L, Douglas-Hamilton I, Getz WM (2008) Disentangling the effects of forage, social rank, and risk on movement autocorrelation of elephants using Fourier and wavelet analyses. *Proc Natl Acad Sci USA* 105:19108–19113.
34. Dalziel BD, Morales JM, Fryxell JM (2008) Fitting probability distributions to animal movement trajectories: Dynamic models linking distance, resources, and memory. *Am Nat*, in press.
35. Forester JD, *et al.* (2008) State-space models link elk movement patterns to landscape characteristics in Yellowstone National Park. *Ecol Monogr* 77:285–299.
36. Worton BJ (1989) Kernel methods for estimating the utilization distribution in home-range studies. *Ecology* 70:164–168.
37. Seaman DE, Powell RA (1996) An evaluation of the accuracy of kernel density estimators for home range analysis. *Ecology* 77:2075–2085.
38. Horne JS, Garton EO (2006) Likelihood cross-validation versus least squares cross-validation for choosing the smoothing parameter in kernel home-range analysis. *J Wildl Manage* 7:641–648.
39. Getz WM, Wilmers CC (2004) A local nearest-neighbor convex-hull construction of home ranges and utilization distributions. *Ecography* 27:489–505.
40. Getz WM, *et al.* (2007) LoCoH: Nonparametric kernel methods for constructing home ranges and utilization distributions. *PLoS ONE* 2:e207.
41. Brillinger DR, Preisler HK, Ager AA, Kie JG (2004) An exploratory data analysis (EDA) of the paths of moving animals. *J Stat Plann Infer* 122:43–63.
42. Brillinger DR, Preisler HK, Ager AA, Kie JG, Stewart BS (2002) Employing stochastic differential equations to model wildlife motion. *Bull Brazil Math Soc* 33:93–116.

Supporting Information

Getz and Saltz 10.1073/pnas.0801732105

SI Text

This text contains details of the simulation methods used to generate results pertaining to Figs. 2 and 3.

The Vegetation Matrix. The vegetation matrix was set up to provide a high level of flexibility in its design so habitat structure could be easily changed for future analyses. It is the only component in the model that has stochastic elements. However, the matrix is formed at the onset of the model and, except for depletion due to consumption by the animals, it remains unchanged throughout a given model run (i.e., there is no plant regrowth). The vegetation matrix consists of a series of patches and its attributes include patch density, patch size, patch quality, edge effect (i.e., reduced quality of cells at the edges of the patch), and some level of randomization in cell quality within the patch.

Patch location was probabilistic based on a predetermined density. Each cell could become a center of a patch using a Monte Carlo draw. In our case, the probability of a cell becoming a center of a patch was 0.02 (i.e., in a matrix of 1,000 cells, the expected number of patches would be 50) in the "good quality" habitat as opposed to 0.017 in the "poor quality" habitat. Patches were square with the side of the square ranging from 1 to 10 cells drawn from a beta distribution with parameters $\alpha = 2.3$ and $\beta = 1.5$ on the interval [0, 10]. In this manner, patches with 7- to 8-unit sides were the most common (≈35%), followed by sides of 3–6 units (≈25%). Small (1–2 units on the side) and large (9–10 units on the side) patches were rare (accounting for a total of ≈15%).

Patch quality was determined randomly assigning a value ranging from 1 to 10 (drawn from uniform distribution) to the center cell. These values reflect units of energy that could be consumed by the feeding animals. The quality of the remaining cells around the center cell declined as a function of the distance from it. Specifically, this was calculated as one minus the proportion of the distance from the maximum possible distance in a specific patch with an exponent of 0.3. Thus, the decline in quality of the cells relative to the center of the patch was sharp at the edges whereas the more central cells were relatively similar. To generate within-patch variability, the actual quality assigned to each cell was then multiplied by a random number generated from a beta distribution with $\alpha = 100$ and $\beta = 1$, so that most cells received values very near their original one (>0.95) but on rare occasion could drop by as much as 15%.

Patches could overlap, and cells falling within more than one patch received the values generated for the last patch being constructed in the code. The vegetation grid was surrounded by a 40-cell-wide band with zero values. In this manner, individuals did not bounce off of the edge but rather simply avoided it because of lack of resources.

The Time Step. Time in this model is not explicitly defined and can be viewed as a sampling time step (τ). The time step is short and represents points at which the individual may make decisions. In each time step, an animal could: move, feed, or remain standing. The decision the animal makes between these three behaviors depends on the animal's energetic status, its location, and the location of other group members (see below). However, in this specific model, after a target is selected and the animal decides to move there, the decision remains unchanged until the animal arrives at that target—i.e., the animal can make only one choice and then continue moving in the same direction until arrival at the selected target. However, a decision between several choices in each movement step easily can be incorporated (e.g., reassessing the target after other herd members have changed their position).

Feeding and Energetics. Each animal has an energy bank that has an upper limit (E_s). This upper limit may vary between individuals, although in the our example all animals were assigned the same upper limit of 100 units. For each time step, an animal loses a given percentage of its energy reserves (i.e., a basal cost) set in our case to $0.006E_s$. When moving, there is an additional energetic cost of $E_m = 0.003E_s$ per time step. When feeding, uptake is constant at a rate of one unit of energy per time step.

Animals consider the value of the vegetation in patches of 3×3 cells. An animal arriving at a cell in a given patch remains in that cell and in each time unit consumes one vegetation unit from the total number of units in the 3×3 cell matrix surrounding it. If its energy bank is full, the animal remains in place but does not feed. Each time step a given proportion ($0.006E_s$) of the energy bank is emptied. Thus, an animal in a patch with a full energy bank (gut) will alternate between feeding and resting in consecutive time steps.

Movement Rules. Once the nine cells around a given individual are consumed, the animal checks its surroundings within a given perception radius and selects a new target to move to. This is done by the animal evaluating both the vegetation and the safety matrices (the latter, as described in the main text, is a score for each cell calculated as the sum of the inverse of the distance of all herd members to that cell—see definition of S_p below) provided by each cell within the radius, based on the animals' energetic status (i.e., the realized movement distribution—see Fig. 2). In this model, each cell of the realized movement matrix receives a score reflecting the vegetation content (i.e., energetic value) of the 3×3 cell patch around it (E_p), the level of safety (S_p) offered by the cell in terms of its spatial location relative to other herd members, and the cost of moving to the new cell based on the distance to it (D_p) and the cost of movement (E_m). S_p is given in terms of the distances D_i is the distance from the cell to herd member i (i covers all members of the herd excluding the one currently evaluating the patch) by the formula

$$S_P = \sum_{i=1}^{\#\,\mathrm{ind}} \left(\frac{1}{D_i}\right)^x,$$

where the exponent x modifies the shape of the safety curve as a function of distance from an individual from a sharp decline for $x \geq 1$, to a more gradual decline for $x < 1$. In our specific model, this value was fixed to 1 but can be used as a function of the level and type of risk. Only one animal can feed in a cell at any one time, and the cell and the eight cells surrounding it are not accessible to other members of the group.

We then weighted the vegetation quality and safety values for each cell by the internal drivers to produce the actual scores for the realized movement distribution (see below). In our case, there are two drivers, hunger and safety. Hunger (α) is that proportion of the bank of energy reserves that is empty. The drive for safety is determined as $1 - \alpha$.

The score of each cell of the realized movement distribution within the perception range of individual i is calculated as

$$\text{Score}=(E_{\text{P}}-E_{\text{m}}\times D_{\text{P}})\times\alpha+S_{\text{P}}\times(1-\alpha).$$

Thus, a satiated animal has a zero hunger drive and is interested only in safety. As the amount of reserves decline, the animal is driven more by hunger and less by safety.

Once a target is selected, the animal moves toward it on a straight line one cell each time step traveling along the imaginary line from the center of the cell the animal currently occupies to the center of the selected target. During movement, the animal does not feed and does not consider new targets.

We note that to keep our simulation simple, we have not made explicit use of step size distribution $A^k(s)$ for the two activities of foraging within a patch ($k = 1$) and foraging between patches ($k = 2$). Rather, the distributions $A^k(s)$ are implicitly embedded in our movement rules that keep individuals within a 3×3 block of cells (a patch) until the patch is depleted or must be exited according to the rules specified, and then an individual must move to a new patch that may be as close as contiguous with the current patch or as far as the edge of circle defined by the perceptual radius centered on middle cell of the current patch. In essence, we have not specified the landscape matrices and movement distributions separately, but we have combined the quantities

$$\mathcal{L}^k_{ri,\alpha\beta}\mathcal{A}^k_{\pi r}(\alpha,\beta)$$

through a single definition, and some computational effort, not required for our simulations, is needed to separate movement distributions from the landscape matrices.

Cluster Analysis. The (x, y) locations at the given time step were run through cluster analysis limited to two groups. Clusters were identified by using the "clusdata procedure" in MATLAB with the default setting of Euclidean distance among points. Once points had been assigned to a cluster, the mean distance among points in a cluster was calculated as well as the mean distance among all 20-group members of the entire herd. The sum of the two means for the subgroups divided by the mean for the entire herd was then calculated and plotted as the ratio (y-axis) in Fig. 3.

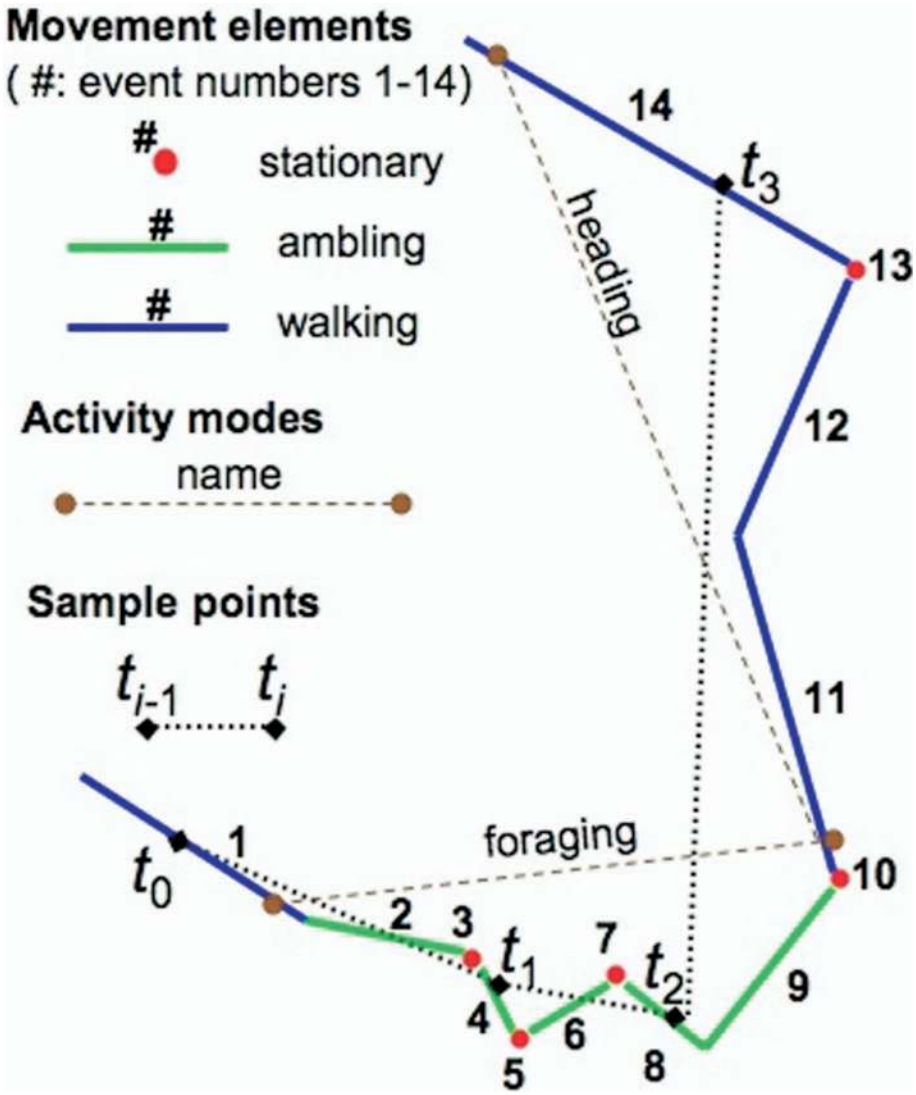

Fig. S1. An idealized movement track consists of an ordered sequence of events selected from a set of fundamental movement elements (FMEs). A canonical activity mode (CAM) is a mixture of FEMs. Sample points t_i, $i = 0, 1, 2, \ldots$, are typically independent of event start and stop times and intervals $\tau_{i-1} = [t_{i-1}, t_i]$ may mix parts of two or more CAMs.

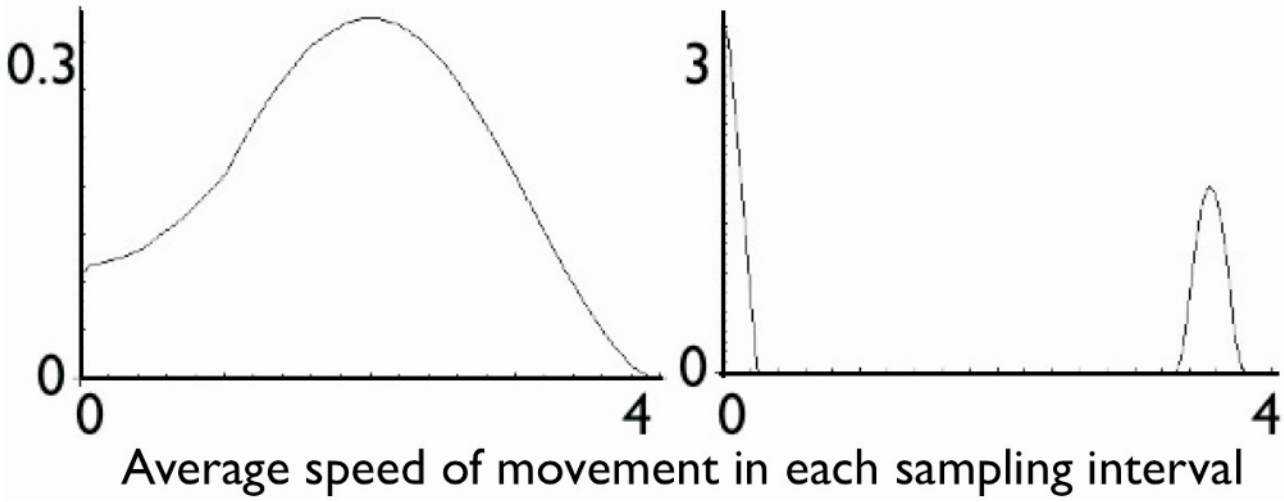

Fig. S2. These distribution illustrate how the same canonical activity mode distribution made up of two fundamental movement elements, one stationary element $s_1 = 0$ and one with a characteristic speed $s_2 = 4$, changes when the sampling interval is several times longer (*Left*) or an order of a magnitude shorter (*Right*) than the characteristic lengths of individual movement mode events. (Note the different vertical scales because the area below both curves is 1.)

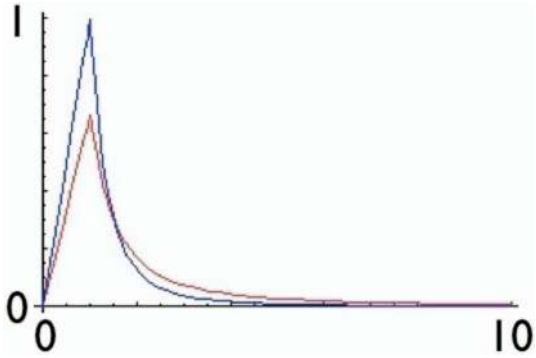

Fig. S3. The modified Pareto power-law distributions we used: $f(x) = \begin{cases} \left(\frac{2q}{2+q}x\right) & on\ [0,\ 1] \\ \left(\frac{2q}{2+q}x^{-q-1}\right) & on\ [1,\ \infty) \end{cases}$ with $q = 2$ (more peaked: blue) and $q = 1$ (less peaked: red).

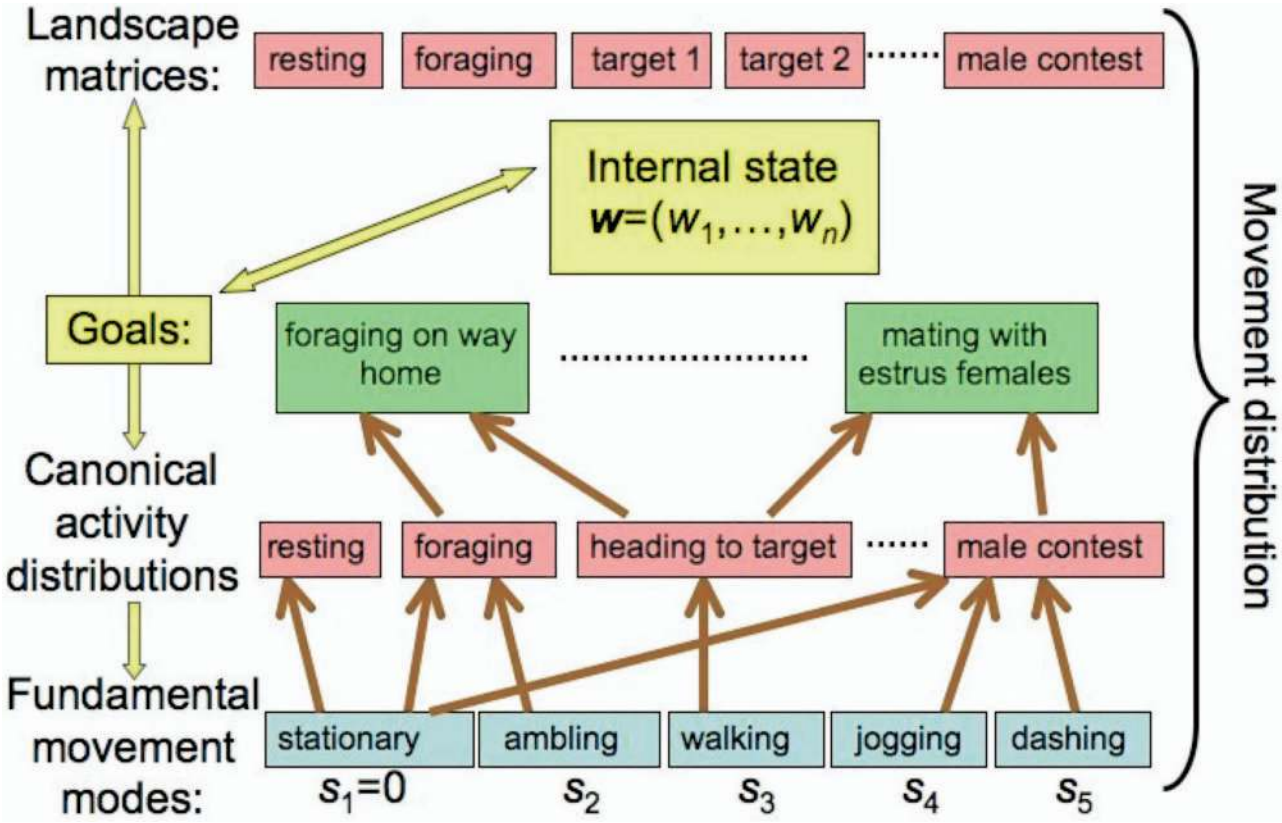

Fig. S4. An illustrative example of the different layers and flows of information needed to construct goal-modified movement distributions and analyze movement data. The yellow arrows relate to movement path constructions, beginning with an identification of goal states based on the internal state **w** of the individual and the landscape matrices incorporating the external state **r** and the navigation capacity Ψ discussed in ref. 1. These goal states, in turn, determine a mix of CAM distributions that are modified (weighted) using the values in the associated landscape matrix (representing a distribution of landscape values across a covering of cells). The CAM distributions are themselves constructed from FMEs—constituting the motion capacity Ω introduced in the general conceptual model (1)—in a way that depends on the simulation time step (Fig. S2). In terms of data analyses, state space (2), exploratory data (3), and other suitable methods may be used to identify segments of movement tracks that are generated under different CAMs and mixes of CAMs.

1. Nathan R, *et al.* (2008) A movement ecology paradigm for unifying organismal movement research. *Proc Natl Acad Sci USA* 105:19052–19059.
2. Patterson TA, Thomas L, Wilcox C, Ovaskainen O, Matthiopoulos J (2008) State-space models of individual animal movement. *Trends Ecol Evol* 23:87–94.
3. Brillinger DR, Preisler HK, Ager AA, Kie JG (2004) An exploratory data analysis (EDA) of the paths of moving animals. *J Stat Plann Infer* 122:43–63.

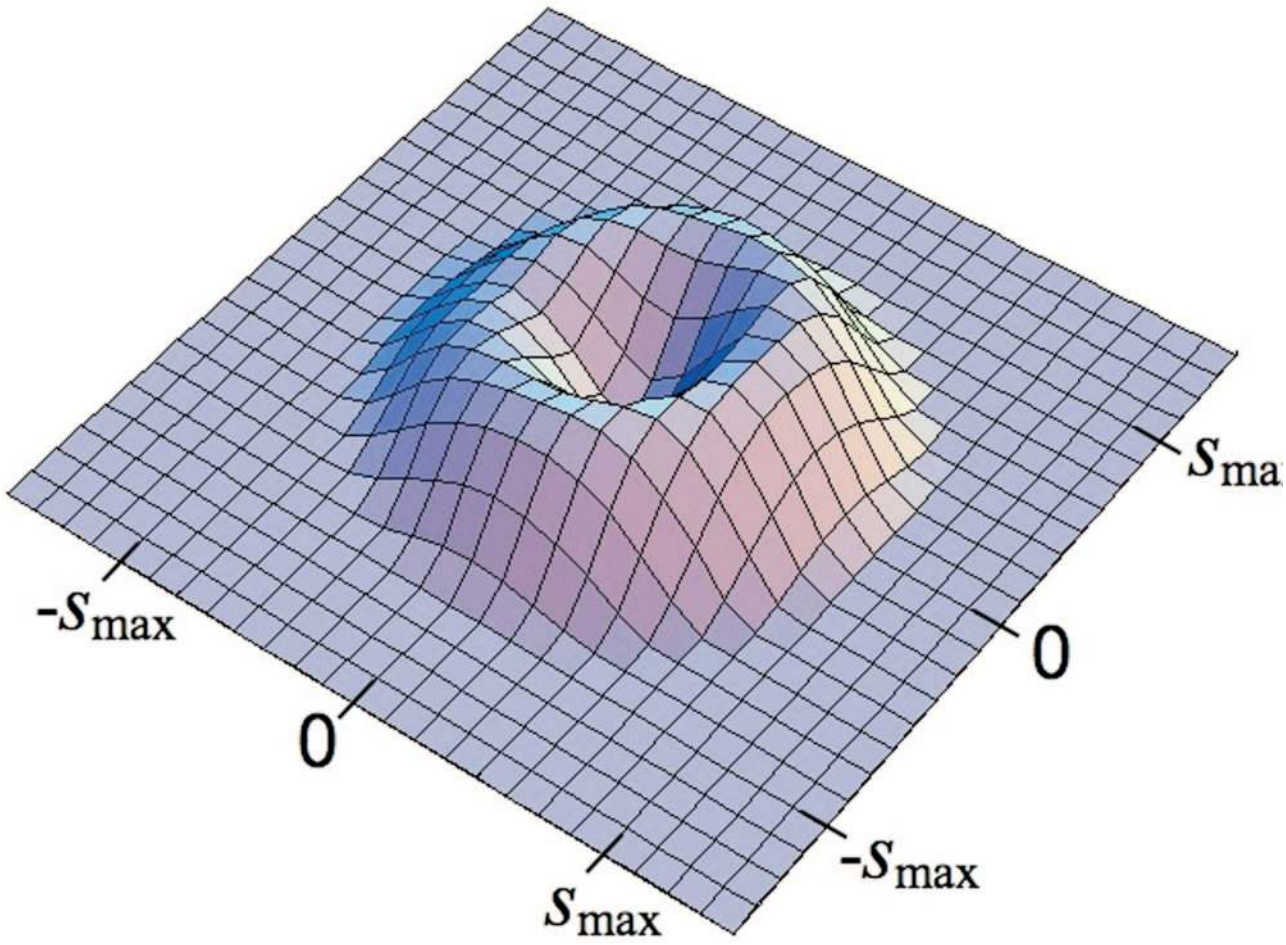

Fig. S5. A unimodel one-dimensional step-size distribution $G_{ri}^{k}(s)$ on [0, s_{max}] is rotated through 360° to convert it to a spatially explicit two-dimensional distribution $G_{ri}^{k}(\alpha, \beta)$ defined over a featureless landscape grid $C^{\alpha\beta}$ [i.e., all grid cells (*a*, *b*) have the same neutral state in terms of $\mathbf{c}_i^{\alpha\beta}$ being defined to reflect a completely neutral landscape structure].