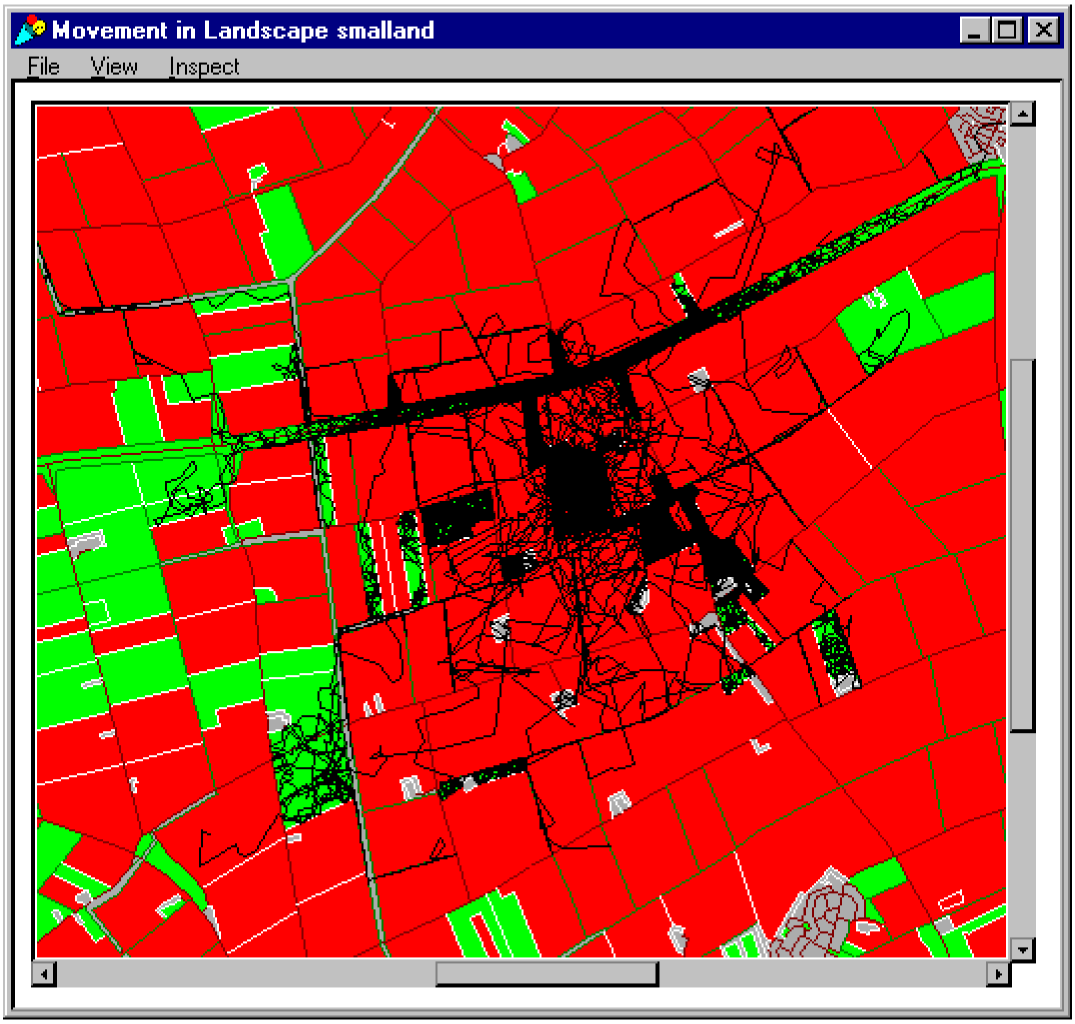
SmallSteps Movement Model



Manual

Version 1.0

ALTERRA

0. Preface

1. SmallSteps - the Model

1.1. Short Overview

SmallSteps is a movement simulation model, in which movement is represented by a correlated random walk, or variants thereof. Movement may occur through a heterogeneous landscape, of arbitrary complexity, consisting of (many) different landscape elements. Movement may be influenced either by the linear elements in the landscape, or the surface-shaped elements, or by both. The landscape itself is represented by data in a vector format, unlike the grid-based approaches applied by Schippers et al. 1996 (Gridwalk), Wiens et al. (1997), With et al. (1997). Comparable approaches to the simulation of movement are applied by Tischendorf and Wissel (1997), Tischendorf et al. (1998), Vermeulen (1995), Haddad (1999ab), and in the Polywalk model.
Output statistics are in terms of parameters of population spread or in terms of arrival probabilities. The first may be useful in answering questions related to the conductivity of the landscape (e.g., how the presence of corridors affects the conductivity); the second estimates are useful in a metapopulation context, where the probability of colonization of isolated habitat patches is a critical parameter.
The landscape is obtained from standard Arc\Info coverages, and can be of arbitrary scale and complexity. SmallSteps has been applied for, among others, Tree frog, Root vole, Reed warblers, Lynx, Pine marten, at different spatial scales. The framework is currently being extended to include fish migration.

1.2. Spatial Representation

The landscape organisms move through is generally complex, made up by many different landscape elements. These elements may be (on the relevant spatial scale) considered either as surfaces or as linear elements. In the vocabulary of vector-based geographical information systems, these are referred to as polygons and arcs. The ease with which linear elements (and therefore the borders of surfaces as well) are represented, is one of the advantages of a vector-based approach, over a grid-based approach.
The vector-data used by SmallSteps, are obtained from standard Arc\Info files, and easily generated from any complete Arc\Info coverage. Both polygon and arc attributes are required to rebuild the right topology within SmallSteps; an additional file (ungenerate) should provide the data on the actual vertices of all the arcs in the landscape.
An underlying grid-representation of the landscape is generated on-demand from within SmallSteps, for the sake of computational performance. In addition, raster-data may be provided in Arc\Info asciigrid format, to define any local property possibly affecting movement parameters (e.g. vegetation cover).

1.3. Movement in Time and Space

Complementary views on modeling processes in time and space exist, i.e. event-based and a time step-based. Approaches in modeling movement can build on either concept, while the necessary empirical data can be one and the same, but lumped in different ways (by time or space).
With respect to movement, field data from e.g. radio-tracked individuals are usually obtained at fixed time intervals. The data can be applied directly in time step-based simulation models, using the calculated distributions of move-length (or velocity) and angle. Thus, these distributions basically define the distance covered and the direction taken within a standard time interval.
However, a strict representation of a correlated random walk process (Turchin 1998) would require the calculation of an additional probability distribution, that of move-duration. Several observed (fixed-time) steps in fact might constitute a single move. Acknowledging this aspect would lead to an event-based approach, where besides velocity and the turning angle of the move, also its duration is drawn from a probability distribution. In practice, movement data are often too scarce or the environment is too heterogeneous to calculate move-duration per habitat type; the default approach taken in SmallSteps is therefore the fixed-time step one. For extreme spatial configurations (movement in networks) an

event-based implementation is available. This variant was applied among others to simulate recreant path-use in nature reserves.

2. Movement

2.1. Movement Patterns (polygons)

Correlated Random Walk

A basic assumption of many movement models is that movement, at least in a homogeneous environment, is essentially random. The random-walk process is simulated applying a few simple rules:

1. Select a move-angle α from the move-angle probability distribution
2. Select a move-length l from the move-length probability distribution
3. Make the move with selected angle and length
4. Go to 1, etc.

For a pure random walk, a Uniform distribution of angles on the interval 0 to 2π is used. The move length l_i is usually assumed to have a Gamma or Exponential distribution. If the goal is to reproduce observed time-dependent movement patterns, an additional parameter, the motility, may be required, defining the probability of a move to take place (Turchin, 1998).

Clearly, a pure random walk is an overly simplistic view of the movement process. A correlated random walk may be a (slightly) more realistic alternative. With a correlated random walk one assumes that there is a correlation between the current move angle and the next (Kareiva and Shigesada 1983). Not the absolute angle of a move is what matters, but the angle between two consequent moves, the turning angle θ_i. In the rules given above, replace move-angle by turning-angle, to obtain the algorithm for a correlated random walk. The turning angle is obtained from a probability distribution, often a Normal distribution with zero mean (no bias in turning left or right). The resulting walking pattern will become less tortuous (winding, sinuous) as the variance decreases. For a small standard deviation the process may be called a directed random walk (see Tischendorf). The process is given by the equation:

$$\alpha_i = \left(\alpha_{i-1} + \theta_i\right) \operatorname{mod} 2\pi$$

where θ denotes the turning angle. More (random) walk variants are shortly discussed in the chapter Random Walk Variants.

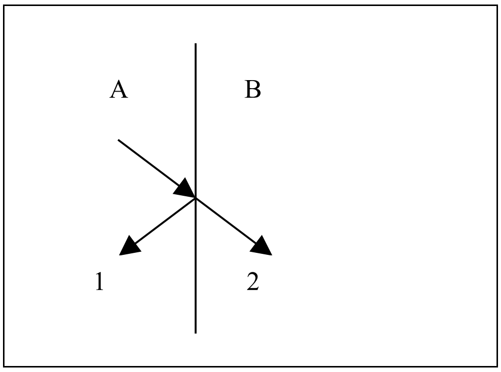
2.2. Heterogeneous Landscapes

SmallSteps is developed to simulate movement through a landscape consisting of more than one type of landscape element. Each type that is encountered in the landscape, needs to have a set of movement parameters assigned to it. This set consists of at least two parameters: to define an Exponential distribution for move-length, a single parameter is required; assuming no bias for turning left or right another parameter sets the variance in turning-angle. For such heterogeneous landscapes, simulation is the only feasible approach to solve for the expected spatial distribution of individuals; even for a relatively simple movement pattern like a correlated random walk an analytical solution is only available assuming a homogenous environment (Kareiva and Shigesada 1983, McCulloch and Cain 1989).

2.3. Boundary Behavior

In a landscape mosaic, moving individuals will frequently encounter edges, the boundaries between elements of different types. It is possible to ignore boundaries in SmallSteps while still taking into account the impact of different landscape elements on movement attributes (see above). Usually though, boundaries are considered to be the locations where a decision is taken: will individuals cross the boundary between two types? And how will movement continue?

The first decision is simulated by defining transition probabilities for every possible combination of *i* types of landscape elements (resulting in an *i*i* matrix). As the figure shows, conditional probabilities are defined taking into account the location of the origin of the move. At the border between type A and B, P_{AB} defines the probability of moving from A into B, and P_{BA} defines the probability of moving from B into A. The probability of turning back into A is $(1-P_{AB})$.



The number of parameters may be reduced by assuming that these probabilities are complementary, that is assuming $P_{AB} + P_{BA} = 1$. This assumption should be checked against empirical data, and consistently be applied (it affects the way transition probabilities are estimated from the data). Conceptually, complementary probabilities mean that one does not believe memory to be important: once arrived at a boundary in the landscape, the subsequent handling of boundary behavior will not be affected by the origin of the move. It is the responsibility of the user to decide whether to make probabilities complementary or not – SmallSteps expects both P_{AB} and P_{BA} to be supplied independently of each other.

If the individual crosses the boundary, it is assumed that its movement will continue on the same vector. The remaining vector-length is scaled by the ratio of the mean move-length in the patch-type entered, to the mean move-length of the patch-type left. Thus,

$$L_2 = \frac{\bar{l}_2}{\bar{l}_1} \cdot \left(L - L_1\right)$$

with L representing the original move-length, L_1 the part effectuated in patch 1, and L_2 the remaining length to be covered in patch 2.

On deciding not to cross a boundary, individuals will bounce off the edge in one of the following ways:

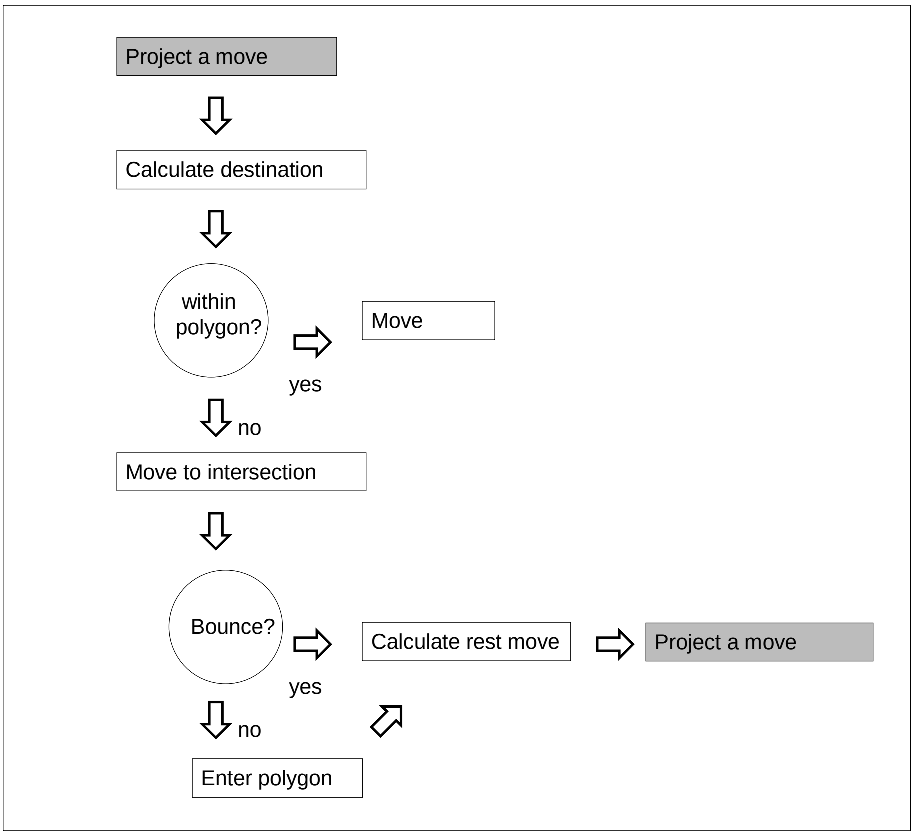
- bouncing, (“billard-ball”) at an angle equaling the incidence angle (set as a default)
- reversing, adding π to the move-angle
- turning left or right, adding +/- ½π to the move-angle, depending on the incidence angle

2.4. Landscape Boundary Conditions

In a finite landscape, moving individuals will eventually encounter the edges of the model landscape. Boundary conditions will affect model outcome, depending on the choice of spatial and temporal scales. Absorbing boundary conditions will lead to an underestimate of the circulation within the landscape (individuals leaving the area are lost); reflecting boundary conditions will lead to an overestimate. Periodic boundary conditions may be interesting for theoretical applications, but hard to implement and probably not very relevant to the complex realistic landscape case. By default, absorbing boundary conditions are assumed.

2.5. Movement Algorithm

The algorithm applied to simulate movement in a polygon world, irrespective of the actual type of random walk (pure random or correlated) is pretty simple. For the more complex walking patterns, e.g. oriented or goal-directed walks, the basic algorithm may be embedded in a hierarchical decision-making process, where internal or external conditions are checked first, to determine the kind of move to make.



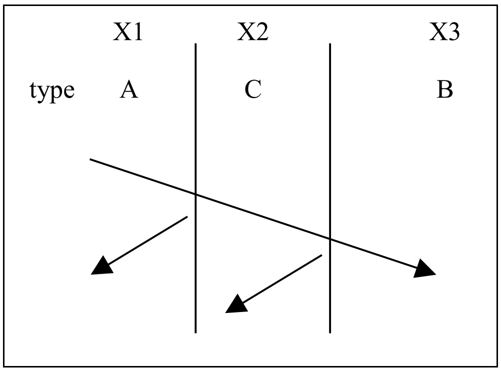
2.6. Movement Patterns (arcs)

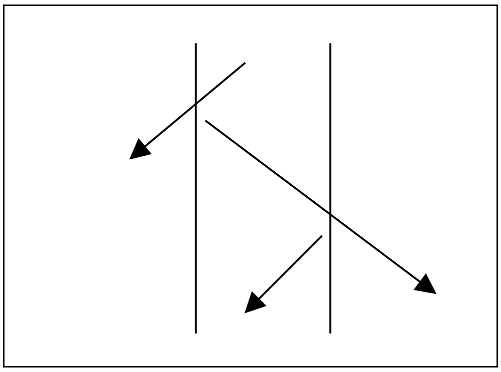
There are (at least) two situations in which the "correlated random walk in polygons" approach described above, appears inadequate. In the first place, individuals may deliberately track the edges in the landscape, e.g. forest-pasture edges. This border-tracking can be considered a movement pattern in its own right, and interwoven with correlated random walk movement patterns. It is not accounted for in SmallSteps, though, but it could be added easily.
In the second place, individuals may have been observed to move preferably through linear landscape elements. These linear landscape elements can be represented as polygons, with the drawback that a zigzag movement path will be simulated with many irrelevant border-collisions. Alternatively, one may represent these elements, like hedgerows, as arcs, and apply an algorithm for movement on arcs, analogue to the correlated random walk in polygons. In the extreme, the movement landscape may consist only of such linear elements, for instance when simulating the redistribution of seeds floating through a network of ditches. In most cases though, one needs to combine the "arc-walking pattern" with the "polygon-walking pattern", as most species can hardly be supposed *never* to leave linear landscape elements when dispersing.

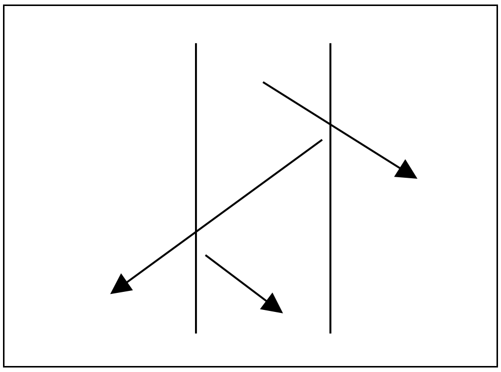
Arc-walking is specified by a move-length probability distribution (just as polygon-walking), and a *turning probability*, determining how frequent the direction of movement is reversed. The handling of transition probabilities is a lot more complex when arc-walking is combined with polygon-walking. By definition, when arc-walking, each move starts on a boundary between two polygons; thus, at the start of each move the chances of a transition to either polygon need to be evaluated.

Transition probabilities between polygons and arcs are defined in the following way. Whenever an individual in X1 hits the polygonborder, the probability to enter X2 (the corridor-type of arc C) is P_{AC};

the probability to turn is 1 – P_{AC}. If it enters the arc, it is assumed that immediately also the possibility to continue into X3 is evaluated.







While 'arc-walking' on a corridor-type of arc, at the start of each move, an individual's decision where to go is affected by what is on the left- and the right-side. We may look at it in the following way. First, the individual looks at one side and determines whether to cross this boundary or not. If it chooses to cross, no further decisions have to be made. If it doesn't, it will look at the other side and decide whether to cross this boundary or not. If not, it will stay within the corridor(-arc). If we assume that there is no bias in whether an individual looks first to the right or to the left, we can summarize the outcome of these decisions in net probabilities of moving into left- or right-poly, or staying in the corridor(-arc). Writing out the probability of moving from the arc (X2) into polygon X1, thus yields:

$$P(x2x1) = 0.5 * Pca + 0.5 * (1 – Pcb) * Pca = Pca * (1 – 0.5 * Pcb).$$

See table below for all other net probabilities. Keep in mind that all these probabilities are derived from the pair of transition/preference parameters defined for each combination of landscape element types. This default decision-taking model builds on the assumptions: 1) "individual in polygon decides whether or not to enter arc; if it does it immediately thereafter decides whether to leave it on the other side" and 2) "individual in arc decides whether to leave arc on one side, if it doesn't it immediately thereafter decides whether to leave arc on the other side". Alternative models may be plausible as well, e.g. models that take into account Pab and Pba (in the example above) for individuals that can be assumed to evaluate left- and right-poly types simultaneously when approaching the arc.

from	*to*	*probability*	*outcome*
X2	X1	Pca * (1 – 0.5 * Pcb)	From arc into left-poly
X2	X3	Pcb * (1 – 0.5 * Pca)	From arc into right-poly
X2	X2	(1 – Pcb) * (1 – Pca)	Stay within arc
X1	X1	1 – Pac	Stay within left-poly
X1	X2	Pac * (1 – Pcb)	From left-poly into arc
X1	X3	Pac * Pcb	From left-poly into right-poly
X3	X3	1 – Pcb	Stay within right-poly
X3	X2	Pbc * (1 – Pca)	From right-poly into arc

X3	X1	Pbc * Pca	From right-poly into left-poly

In addition, when arc-walking, the projected move length may exceed the remaining length of an arc. Arcs begin and end in nodes (GIS-terminology) – so in the implementation one needs to determine what will happen in a node. The following rules are applied, handling node-behavior:

1) if (corridor-)arcs of the same type are present (meeting in the same node), movement continues on one of these (randomly selected) arcs
2) if not, the decision where to go is made, based on an evaluation of all polygons and corridor-arcs bordering or meeting in the node. This is done in the same way as for an individual moving on an arc (evaluating right- and left-polygon); the number of alternatives may however be much larger. If none of the alternatives is selected, the individual will turn and continue on the arc it was already on.

As an illustration, we calculate the probabilities of moving from a node into each of a set of polygons or (corridor)arcs that border this node. For performance reasons, these values are cached in the landscape objects. The results are shown for the cases of two, three and four potential destinations.
TX refers to the type of the considered destination element X; T1 etc refer to the type of alternative destination elements 1 (if there are three potential destinations, we have two alternatives; if there are four potential destinations, we will have three alternatives). C is the type of the (corridor-)arc, in which the individual is assumed to be. We are only interested in probabilities to move from this corridor-arc into either of the destination elements. NB note that technically speaking an individual is not in the node, but in the corridor-arc that ends or starts in the node. Therefore we need to calculate these net probabilities for each corridor-arc connected to the node.

For two potential destinations (identical to the "arc with two bordering polygons case"):

$$P_{inX} = \tfrac{1}{2} \cdot P_{CTX} \cdot \left(1 + \overline{P_{CT1}}\right)$$

For three:

$$P_{inX} = \tfrac{1}{6} \cdot P_{CTX} \cdot \left[2 + \left(\overline{P_{CT1}} + \overline{P_{CT2}}\right) + 2\left(\overline{P_{CT1}} \cdot \overline{P_{CT2}}\right)\right]$$

For four:

$$P_{inX} = \tfrac{1}{24} \cdot P_{CTX} \cdot \left[6 + 2\left(\overline{P_{CT1}} + \overline{P_{CT2}} + \overline{P_{CT3}}\right) + 2\left(\overline{P_{CT1}} \cdot \overline{P_{CT2}} + \overline{P_{CT1}} \cdot \overline{P_{CT3}} + \overline{P_{CT2}} \cdot \overline{P_{CT3}}\right)\right]$$

N.B. the bar denotes complementary probabilities, $\overline{P} = 1 - P$

2.7. Landscape Elements as Barriers

By manipulating transition probabilities, a barrier effect exerted by (polygon) landscape elements is easily incorporated. Absolute barriers will never allow individuals to enter; relative barriers will have transition probabilities approaching zero.
Linear elements acting as barriers (e.g. roads) are more easily accounted for by representing them as arcs (see above). Comparable to polygons acting as barriers, arcs acting as barriers can be simulated implicitly by treating them in the same way as arcs acting as corridors (tag = C), but assigning to these elements very small transition probabilities, approaching zero.
Alternatively, a special barrier-type of arc can be defined (tag = B), invoking more elaborate 'barrier-crossing' boundary behavior. These barrier-arcs do not require movement parameters like move-length, as individuals are only supposed to cross barrier-arcs. The behavior at such a barrier is handled in three steps, as a hierarchical decision-making process, involving two additional parameters. Firstly, the

motivation to move to the other side of the barrier is determined, depending on the *transition probabilities* between the elements on both sides of the barrier (optional). This is handled in the same way as for a plain arc, separating two polygons. Secondly, a willingness to cross a barrier is determined, depending on the *crossing-probability*. Thirdly, mortality at a barrier is effectuated, depending on the *crossing-mortality*. The latter two parameters are type-specific; they may differ for e.g., roads with different traffic intensities.

2.8. More (Complex) Movement

Oriented Walk

A consistent bias in the selected angle can be the result of individual choices (orientation on earth magnetic field, sun) and preferences (climatic factors), or simply some kind of drifting caused by the forces of wind or currents. In any case, the 'preferred' angle does not change with the changing position of the individual: each step is chosen relative to a fixed compass direction. Such an oriented walk (Marsh and Jones 1988) can be defined as:

$$\alpha_i = \left(\alpha^* + \theta_i\right) \bmod 2\pi$$

or more directly as a Wrapped Normal distribution with mean α^*. The random variable is simply the random variable obtained from a Normal distribution modulo 2π. Note that the variance ρ of a Wrapped Normal distribution relates to the variance σ^2 of the Normal distribution it is derived from, as:

$$\rho = e^{-\frac{1}{2}\sigma^2}$$

Other circular probability distributions, like the von Mises distribution, are described in e.g., Fisher (1993).

Goal Directed Walk

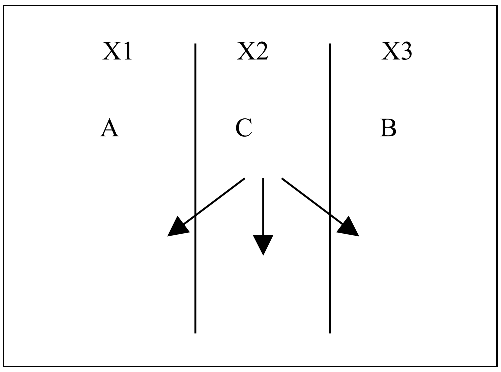
Goal directed walks imply movement in the direction of a (relatively nearby) target: each move will significantly change the angle of the line between current position and this target (point) location (except of course when the individual is already heading for the target). Social attraction (homing) in Tree frogs may lead to a goal directed walk. Attraction strength, the stimulus to turn into the direction of the target location, is likely to be a function of the distance. For Tree frogs, one may assume a negative exponential relationship with distance (perceptibility of the sound of calling males decreases exponentially). For other species, depending on visual clues, a threshold (detection distance) may be a simpler option. A mixture of correlated random walk and goal directed steps results when interpreting the attraction function $f_{(D)}$ as the probability of a goal directed step:

$$f_{(D)} = e^{-\delta \cdot D}$$

$$\alpha_i = \begin{cases} \alpha_{i-1} + \theta_i & x_i > f_{(D)} \\ \gamma_i + \theta_i & x_i \le f_{(D)} \end{cases}$$

with γ_i denoting the angle of the vector from current location to target location, x_i a random number on the interval 0-1, and θ the turning angle.

Another definition of arc-polygon transition probabilities



An alternative way to estimate the chances of moving into left- or right polygon, is the following. When inside a corridor-type of arc, at the start of a move, the chance to move from X2 to X1 amounts to 0.5 * P_{CA}, the chance to move from X2 into X3 amounts to 0.5* P_{CB}, and the chance to remain in X2 amounts to 1 – (0.5 * P_{CA} + 0.5 * P_{CB}). This means that we assume that individuals from X2 are shared between X1 and X3 proportionally to the preferences (transition probabilities), and that no one side will receive more than 50 % of the individuals present. Compared to the default approach, here one takes into account a single decision related to one side only (and assumes that there is no bias for either side). As a consequence, the following situation may occur: when at a single border, individuals would always move into X1 (P_{CA} appr. 1), the probability to select X1in this case can not be more than 0.5, even if X3 is not attractive at all (P_{CB} appr. 0). Clearly, this algorithm needs to be applied carefully. The resulting net exchange probabilities between X1 and X2 and X3 are given in the table below.

from	*to*	*probability*	*outcome*
X2	X1	0.5 * Pca	From arc into left-poly
X2	X3	0.5 * Pcb	From arc into right-poly
X2	X2	1 – 0.5 * (Pcb + Pca)	Stay within arc

The same simplified approach can be taken for the transition probabilities in nodes. The artifact described above can be much more extreme: when one of the potential destinations is very attractive, the probability to select it will never be larger than 1 divided by the number of alternatives.

2.9. Additional Demographic Processes

Mortality

When dispersal is simulated over a long time-span, mortality can hardly be ignored, especially when the net chance of arrival is studied. In a time-step based approach, mortality is easily incorporated by assuming a fixed individual mortality probability μ, per time-step. In a complex, heterogeneous landscape the risks of accidental death are clearly not spread out evenly: some landscape element types present a higher risk (predation, traffic, drowning, etc.) than others. In SmallSteps, thus, landscape element specific mortality probabilities can be applied. A complicating factor is the composite character of movement steps that hit element edges (see above). Within a single time-step, the individual resides in several landscape elements, and mortality probabilities are likely to be some weighted average of the time spent in each of *n* elements:

$$\mu_t = \sum_{i=1}^{n} \Delta_i \mu_i$$

with Δ_i referring to the fraction of the time-step spent in element *i*. Although a simple average, it still seems contrived, given the complete absence of empirical data. Therefore, the default approach is to relate mortality risk to the type of element containing the individual at the onset of each move, ignoring the possibility of actually visiting more than one element type within the timestep.
In many cases, little will be known about mortality risk as an element-type specific parameter, with the single exception of risk related to crossing barriers like roads. Instead of applying an overall mortality risk, equal for all types, mortality may be ignored in the simulations and the results corrected

afterwards, for a constant mortality rate. The reason: effectuating (any) mortality, decreases the sample size (the number of completed movement paths) and therefore requires a larger number of simulated trajectories to obtain a sufficiently large sample size. The sample-size problem can be circumvented by defining dispersers as 'super-individuals': each one representing a population (an internal number) whose size decreases by mortality, as described in Scheffer et al. (1993).

Reproduction, immigration and emigration

By incorporating both death and birth, the dispersal model is converted into a complete population dynamics model. Although beyond the scope of dispersal modeling, it may be useful to list the extensions that are minimally required.
(1) individuals usually do not reproduce everywhere in the landscape; thus some (types of) landscape elements are identified as 'habitat'. (2) individuals usually do not reproduce continuously; thus certain time-intervals should be defined as 'reproductive periods'. (3) in a population dynamics model, one needs density-dependence to keep populations from growing exponentially; thus, in selected parts of the landscape (probably in habitat areas) and during parts of the simulation time, one needs to keep track of local population densities, and make either reproduction or mortality rates dependent on density.
In addition, assumptions should be made with regard to the relationships between local population density and emigration or immigration (density-independent, density-dependent or inversely density-dependent).

3. Output & Analysis

3.1. Output Statistics

Basic Statistics

Basic statistics on movement of individuals and the redistribution of the whole population, depend on the ability of SmallSteps to keep track of the position of each individual (in terms of coordinates and containing spatial element) through time. In addition we have the ability to store the complete path covered by any individual (in terms of coordinates and/or spatial elements visited), and to keep track of the flow of individuals through each spatial element.
The focus of an analysis may be on the statistics of population spread, of arrival or of residence (see below).

Dynamic or Steady-State

Analysis may focus on dynamic or steady-state properties. The steady-state distribution is easily observed in case of a closed universe, with reflecting (or periodic) landscape boundary conditions, by letting the simulation evolve over a long enough time span. With absorbing boundary conditions (individuals leaving the study area), only the relative abundance can be studied, as the number of individuals will continuously decline. Focusing on steady-state properties has the advantage of independence of initial conditions (by definition). However, it is unlikely that data are available to corroborate the simulation results. Also the relevance can be doubted. The spatial scale of most of our problems will be such that an equilibrium in the spatial distribution of a species will not be reached by dispersal alone - other demographic processes like mortality and reproduction will play an important role.

Landscape- or Grid/Vector-Based

When the scales of movement and landscape heterogeneity permit it, we can perform an analysis completely based on the actual landscape elements. For instance, one element may represent the release location of dispersers, and we may observe how often and when other elements are visited. In some cases, such an approach may not be feasible, e.g. when the whole landscape consists of only a very few number of elements. In that case, one may construct an underlying grid and keep track of grid cells being visited/traversed, or draw circles are the release location, representing distances for which we want to record first passage times and/or angles.

3.2. Statistics of Population Spread

Dynamic properties relate in the first place to aspects of population spread, as in models of animal range expansion. Mean velocity of range expansion can be calculated from the positions of all individuals, each timestep. This velocity can be interpreted as a measure of average landscape conductivity or resistance to movement.

Population Distributions

Basic data on spread include the **Population Distance Distribution**, a frequency distribution showing the distances covered at a certain time, or put reversely, the **Population Time Distribution**, a frequency distribution showing the times required to cover a certain distance. In practice, these distributions are summarized, using population average or median values.

Net Squared Displacement

When taking the mean of all squared displacements, we obtain the **net squared displacement** (NSD). It is expected to be a linear function against time, at least for a correlated random walk in a homogeneous environment (Turchin 1998). I suppose that in a heterogeneous environment this linearity is preserved, if the landscape structure is relatively random and fine-grained. The coefficient of the observed NSD can be used as a measure of the average <u>conductivity</u> of a landscape (or the reverse of the <u>resistance</u> of the landscape to movement).

First Passage Distributions

When looking at a fixed distance, another way to characterize the conductivity of the landscape is by recording **first passage distributions** of times and angles: the times when and locations where individuals covered this distance for the first time (imaging drawing a circle around the start location).
First Passage Time Distribution: the (cumulative) fraction of individuals that has covered a certain distance, mapped against time. This relationship can be summarized in the median first passage time (the time needed for 50 % of the population to cover this distance).
First Passage Angle Distribution: the (cumulative) occurrence of all first passage angles. This statistic (if it departs from a Uniform distribution) may be useful to decide whether the landscape is structured and guides movement in such a way, that at certain angles more individuals cross the circle than at others. A next step in the analysis may be to relate such biases to the presence and position of specific landscape elements.
Recording time and location of each circle-crossing in the simulation, allows us to create both distributions simultaneously. The result can be stored in a two-dimensional matrix, and e.g. for calibration purposes, compared to observed data.

3.3. Statistics of Arrival

In metapopulation studies, we usually need to know what the chances are for an individual when leaving location (habitat patch) A, to arrive at location (habitat patch) B. In SmallSteps, a simulation experiment to calculate these probabilities, is easily set up, by defining polygons or arcs in the landscape as belonging to a habitat cluster, and simulating moving individuals that become immobilized at their first encounter with an arrival cluster. For each release location the number of individuals that arrive at each cluster is recorded. After dividing by the number of dispersers released, the result is a matrix of probabilities, with rows representing the release locations, and columns representing the arrival clusters. Of course, probabilities will to some extend depend on the simulation length; this should be checked in exploratory runs.

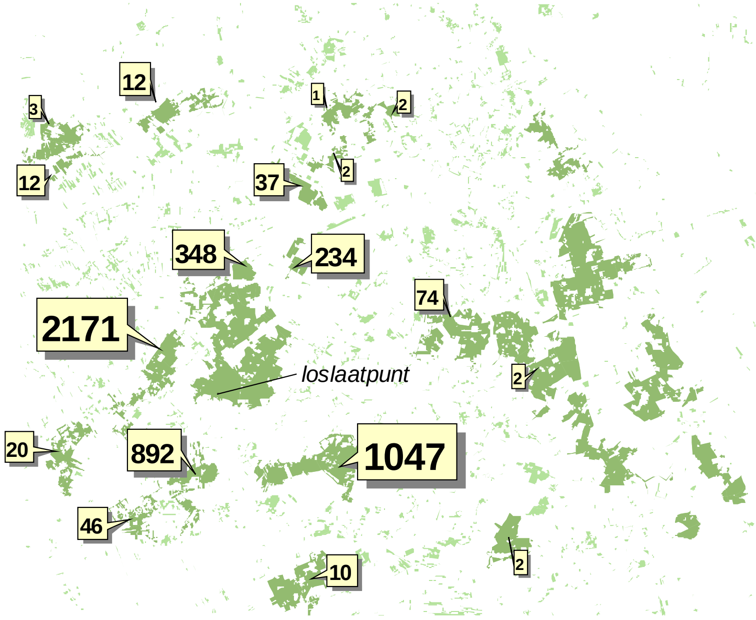


Figure. Results of recording arrivals in several clusters of woodland habitat.

3.4. Statistics of Residence and Density

During a movement simulation the density of individuals in each spatial elements changes. At any time, SmallSteps may report the density of individuals, as their number divided by polygon-area or arc-length. Default output is a table with all polygons; for large landscapes it should be a map with different colors or shading representing different density categories.
Data on densities per element can be lumped into data per element-type, as the landscape-averaged density in each element type (total number of individuals in type i divided by total area type i).
If interest arises, SmallSteps may be instructed to keep track of entrance- and exit-events of individuals for each spatial element, or to keep track if the complete history of all spatial elements visited by each individual – in order to calculate variables like residence time.

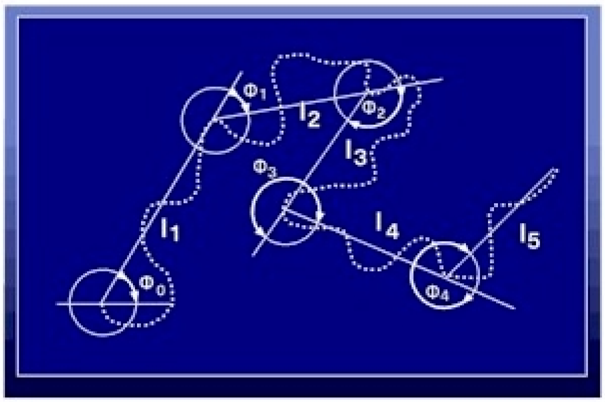
4. Model Parameters

CRW Parameters

To simulate a correlated random walk, we need for each type of landscape element a

- **move-length distribution**. Assuming that move-length is exponentially distributed, there is only a single parameter required (α, with mean = $1/\alpha$).
- **turning-angle distribution**. Assuming that turning-angle is normally distributed, and that there no bias for left nor right (mean is zero), there is only a single parameter required, the standard deviation.

The calculation of CRW parameters from field data is a relatively straightforward process, described in detail in Turchin (1998).



Transition Probabilities

For each combination of spatial element types (all polygon types plus all the corridor-arc types) a transition probability needs to be defined. For n such types, the result will be a n*n matrix of transition probabilities. The estimation of transition probabilities is a complicated process, since besides actual border-crossings we also need to establish when individuals could have crossed a boundary, but didn't.

5. Model Calibration

In calibration, model parameters are tuned in such a way to obtain the closest fit between the model output and an observation data set. In general, the parameters that can not be directly estimated from the data-set are the candidates to be varied in a calibration procedure. In movement models, the correlated random walk parameters are the more reliable ones, while transition probabilities are often "best-guesses".

5.1. Maximum Likelihood approach

The method of maximum likelihood (ML) is a standard way of estimating parameters of statistical models. It involves choosing as estimates the parameter values that make the probability of obtaining the observed data (the **likelihood function**) as large as possible (Manly 1990). The likelihood function is a function of any unknown parameters. An example taken from Manly (1990), concerning count data: "For example, suppose that an insect population is distributed randomly over an area that is divided into a large number of 1 m^2 quadrats. The counts in the quadrats are then expected to have a Poisson distribution with mean μ. If a random sample of *n* quadrats is selected, and gives insect counts of r_1, r_2, ….r_n, then the probability of observing the *i*th count is

$$P_{(r_i)} = \frac{\mu^{r_i} e^{-\mu}}{r_i\,!},$$

and the probability of observing the whole sample is the product of probabilities

$$L(\mu) = \prod_{i=1}^{n} P(r_i) = \prod_{i=1}^{n} \frac{\mu^{r_i} e^{-\mu}}{r_i\,!}$$

This is a function of the Poisson mean μ. If this mean is unknown then the ML estimate μ* of μ is the value that makes L(μ) as large as possible. It can be shown that this value is the mean $\sum r_i/n$ of the observed counts." To deal with very small values of the likelihood function, usually one calculates the log-likelihood function (simplifying the formula: summation instead of multiplication).

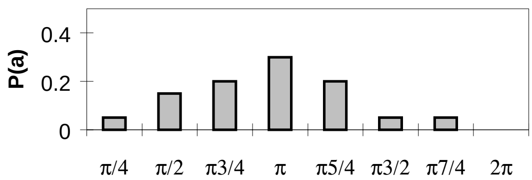
Taking this approach over to the area of movement models, the procedure to follow is: run the simulation model for different parameter sets, calculate from the spatial distribution produced by the model the probabilities for individuals to be at a certain moment at a certain location, and use these probabilities to calculate a likelihood value for the set of observations. The parameter set producing the highest value of the likelihood function, is the one best fit to the data.

This procedure can be applied irrespective of how locations are defined in the observations (and the simulations): landscape elements as they are present in the landscape, imaginary circles drawn around release coordinates, or an underlying grid structure, all may do well depending on the circumstances.

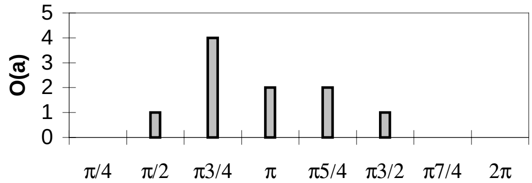
The maximum likelihood approach is not only useful in calibration, it is also useful in model selection. Different movement patterns may be possible (see the sections above). In the Tree frog case, comparing the values of the likelihood functions, we checked whether taking homing behavior into account would improve the fit between model results and movement data.

5.2. Example Calibration

Imagine a calibration based on recorded first-passages of a circle with a certain radius. Time-angle pairs are collected from a large number of simulated movement paths (10^4 - 10^5). Leaving out the time-aspect, a histogram of (first passage) angle frequencies is constructed (with the circle divided into 8 sections of π/4 each):



The first passage observations from the data-set (*n*=10), are collected in the same 8 classes:



As the probability of observing each first passage observation is $P(a_i)$, the likelihood function in this case is

$$L = \prod_{i=1}^{n} P(a_i)$$

and the log-likelihood function:

$$\log L = \sum_{i=1}^{n} \log\left(P_{(a_i)}\right) = \sum_{a=1}^{8} O_{(a)} \log\left(P_{(a)}\right)$$

6. Examples

7. References

Literature

Batschelet, E. (1981) Circular statistics in biology. Academic Press, London.

Fisher, N.I. (1993) Statistical analysis of circular data. Cambridge University Press, pp. 277.

Haddad, N.M. 1999a. Corridor and distance effects on interpatch movements: a landscape experiment with butterflies. Ecological Applications 9(2): 612-622.

Haddad, N.M. 1999b. Corridor use predicted from behaviors at habitat boundaries. American Naturalist 153(2): 215-227.

Kareiva, P.M. & Shigesada, N. (1983) Analyzing insect movement as a correlated random walk. Oecologia 56: 234-238.

Manly, B.F.J. (1990) Stage-structured populations; sampling, analysis and simulation. Chapman and Hall, London, pp.187.

Marsh, L.M. & Jones, R.E. (1988) The form and consequences of random walk movement models. J. Theor. Biol. 133: 113-131.

McCulloch, C.E. & Cain, M.L. (1989) Analyzing discrete movement data as a correlated random walk. Ecology 70(2): 383-388.

Scheffer, M. et al(199?). Super individuals etc.

Schippers, P., Verboom, J. Knaapen, J. P., and Van Apeldoorn, R. C. 1996. Dispersal and habitat connectivity in complex heterogeneous landscapes: an analysis with a GIS-based random walk model. Ecography 19: 97-106.

Tischendorf, L., and Wissel, C. 1997. Corridors as conduits for small animals: attainable distances depending on movement pattern, boundary reaction and corridor width. Oikos 79: 603-611.

Tischendorf, L., Irmler, U. and Hingst, R. 1998. A simulation experiment on the potential of hedgerows as movement corridors for forest carabids. Ecological Modeling 106: 107-118.

Turchin, P. 1998. Quantitative analysis of movement; measuring and modelling population redistribution in animals and plants. Sunderland, Massachusetts: Sinauer Associates, Inc.

Vermeulen, H. J. W. 1995. Road-side verges: habitat and corridor for carbid beetles of poor sandy and open areas. PhD dissertation, Wageningen, The Netherlands.

Vos, C.C. 1999. A frog's-eye view of the landscape. PhD thesis, Wageningen University.

Wiens, J. A., Schooley, R. L., and Weeks, Jr., R. D. 1997. Patchy landscapes and animal movements: do beetles percolate? Oikos 78: 257-264.

With, K. A., Gardner, R. H., and Turner, M. G. 1997. Landscape connectivity and population distribution in heterogeneous environments. Oikos 78:151-169.

Student Reports

(for internal use only)

Jansen, V. 2000. Dispersal of the Great Reed Warbler (Acrocephalus arundinaceus) and the influence of landscape configuration. Graduation report Alterra/Wageningen University.

Martens, A.J. 1999. Dispersiegedrag van de Boommarter. Een simulatie van looppatronen. Graduation report Alterra/Wageningen University/RUU.

Talsma, J. 2000. De weg van de minste weerstand. De weerstand van het landschap tegen de dispersie van grondgebonden soorten. Stageverslag Alterra/RUL.

Other resources

The SmallSteps documentation is maintained by Hans Baveco. For comments, suggestions, or additions, he can be reached at j.m.baveco@alterra.wag-ur.nl.

Up-to-date information on the application of SmallSteps in current research projects, can be found on the SmallSteps coweb, hosted at http://ibn08006:8080/smallsteps (intranet-DLO accessible only).

8. Samenvatting

SmallSteps is de verzamelnaam voor een serie modellen, waarin bewegingen van individuen door een complex, heterogeen, landschap gesimuleerd worden (dispersie-modellen). Een gemeenschappelijk kenmerk van deze modellen is dat de landschaps-representatie vector-gebaseerd is, en opgebouwd wordt uit informatie in Arc\Info bestanden, in een standaard format. Rasterbestanden kunnen echter naar behoefte gebruikt worden om lokale eigenschappen (vegetatie, landschapstype) van invloed te laten zijn op het bewegingspatroon ter plekke. Voor het dispersie-proces zelf kan een tijdstap- of een event-gebaseerde weergave gekozen worden, afhankelijk van welke weergave het beste past bij de bewegingsdata waarop gecalibreerd en/of gevalideerd wordt.

Het grondpatroon voor individuele verplaatsingen is een gecorreleerde random-walk (CRW) of een variant daarop (georienteerde bewegingen, etc.). In principe is voor ieder type landschapselement dat in het dispersie-landschap onderscheiden wordt, een set parameters (2) nodig die deze CRW beschrijft (kansverdelingen voor stapgrootte en draaihoek). Naast dit grondpatroon van verplaatsingen binnen vlakken (polygonen) is het, in een heterogeen landschap, noodzakelijk grensgedrag te definieren. Grensgedrag wordt bepaald door een matrix van overgangskansen tussen alle onderscheiden (n) landschapselement-typen (dus n^2); wanneer geheugen (de voorgeschiedenis) geen rol speelt, zijn overgangskansen complementair (en daarmee het aantal overgangskansen gehalveerd). Elementen kunnen door hun overgangskansen fungeren als relatieve of zelfs absolute barrieres.

Een voor het opschalen belangrijk aspect is dat lineaire landschaps-elementen waar individuen door heen bewegen, zowel als lijnen (arcs) als als vlakken (polygonen) weergegeven kunnen worden, met een consistente afhandeling van verplaatsingen en overgangen. Voor de ene toepassing(sschaal) past weergave van bijvoorbeeld een houtwal als vlak beter; voor de andere (schaal) als lijn.

SmallSteps is ontwikkeld aan de hand van twee toepassingen op een kleine ruimtelijke schaal (enkele km's). Doel van deze toepassingen was het schatten (calibreren) van bewegingsparameters en overgangskansen, m.b.v. telemetrische data, en het ontwikkelen van een methodiek hiervoor. In deze kleinschalige landschappen worden een vijftal verschillende landschapselement typen onderscheiden. Zonder verdere aanpassingen worden deze modellen momenteel toegepast op schaal van enige tientallen km's, waarbij de dispersie-landschappen gedestilleerd worden uit topografische kaart bestanden (1:10000 tot 1:50000).

Toepassingen op nog grotere ruimtelijke schaal zijn alleen mogelijk bij verdergaande vereenvoudiging van het dispersie-landschap en de bijbehorende landschaps-specifieke bewegingspatronen. Voor toepassing op bijna-landelijke schaal wordt een methode ontwikkeld waarbij in het dispersie-landschap enkel grofschalige patronen worden weergegeven, en waarbij binnen deze patronen de CRW-parameters (deels) bepaald worden door informatie uit begroeingstype-bestanden of LKN-bestand (rasterbestanden). De noodzakelijke onderbouwing hiervoor dient echter grotendeels nog plaats te vinden.

Een andere benadering wordt ontwikkeld voor soorten die in hun dispersie sterk gestuurd worden door lineaire landschapselementen, bijvoorbeeld aquatische organismen. Het dispersie-landschap wordt vereenvoudigd tot een dispersie-netwerk. Het hiervoor ontwikkelde model wordt momenteel toegepast binnen een andere context (recreatie-onderzoek), waarbij het gedrag van wandelaars en fietsers op een paden-netwerk wordt onderzocht.