A Stochastic Movement Simulator

Timothy C. Haas\textsuperscript{1}  Sam M. Ferreira\textsuperscript{2}

\textsuperscript{1}Lubar School of Business, University of Wisconsin-Milwaukee, United States, haas@uwm.edu

\textsuperscript{2}Scientific Services, SANParks, Skukuza, South Africa
This talk will cover the following.

1. Define a stochastic movement simulator
2. Estimate its parameters
3. Tips for working with wildlife agencies
Overview of animal movement modeling

One of the earliest animal movement models was a 1994 paper on bison foraging behavior bison in Yellowstone.

Many such models can be categorized as mainly concerned with modeling animals’ habitat selection (steady state movement dynamics) or a one-time movement across the landscape known as dispersal.
Two Movement Ecology papers


1. Models for which the landscape is partitioned into a grid of cells are easy to program and easy to fit to available computational resources through adjustments to the model’s cell size.

2. But continuous-space models are realistic and can be directly related to GIS input and output once coordinate systems are aligned.
Observed animal paths → Read-in paths and GIS layers → GIS layers

$m = m + 1$

Initialize start time, stop time, step duration, step length, and the number of simulation replications ($M$)

time = time + step duration

Locate animal at his/her first observed location

Compute this animal’s heading distribution constant

Randomly sample a heading

Is a step along this heading traversable?

Yes

Take a step along this heading

Have all simulated animals taken a step?

Yes

time = stop time?

Yes

$m = M$?

Yes

End

Yes

Yes

Yes

Yes
Say that an animal moves a step of size $s$ along a heading that the animal thinks will maximize its utility. There is some noise in the animal’s decision so that the heading ultimately chosen is a stochastic function of utility.

This stochasticity is modeled by first computing a utility function in terms of $h$; denoted $u(h)$, and then tabulating a probability distribution that is proportional to this utility.
A deviate is drawn from this heading distribution by evaluating the associated quantile function at a value that is randomly drawn from the unit-interval uniform distribution. Finally, the animal moves one step along this randomly-drawn heading.

Let $\mathbf{x}$ be the vector containing all variables that influence the animal’s utility function. Let $\beta$ contain the associated parameters. Say that animal $i$ is presently located at $\mathbf{l}_0$. 
For instance, rhinos like to eat and mate. Let $twi(l_0, t)$ be the *topographic wetness index* value at location $l_0$ and time $t$.

Let $d_f(l_0, h, t)$ be the distance to the closest food from location $l_0$ along heading $h$ at time of year $t$. 
Likewise, let $d_m(l_0, h, t)$ be the closest potential mate from location $l_0$ along heading $h$ at time of year $t$. Let $c_i$ be the demographics of rhino $i$, e.g. gender, and age.

Then $x' = (twi(l_0, t), d_f(l_0, h, t), d_m(l_0, h, t), c_i')$ and $u(h) = x'\beta$. 
The animal will take a step of size \( s \) in the direction \( h \). We assume that the animal has drawn (in-effect) the value \( h \) from a distribution defined by the probability density function (PDF) \( f(h) = \frac{u(h)}{\int_0^{360} u(y)dy} \). Call this the \textit{heading distribution}.

A heading is repeatedly drawn until the path defined by that heading and the pre-determined step size is traversable by the animal.
SMS algorithm

First, $n$ animals are assigned starting locations. Then, at time $t$, each animal takes a step of size $s$ along a heading, $h_q$ found by randomly drawing from the animal’s heading distribution.

This random draw is accomplished by finding the root of the function $g(h) = p - F(h)$ where $p$ is the deviate drawn from the uniform distribution over the unit interval, and $F(h)$ is the Cumulative Distribution Function (CDF).
Example of Three Observed Paths on a TWI Surface
Numerical integration

This CDF is formed by numerically integrating the heading distribution’s PDF over the interval \((0, h)\).

An adaptive Newton-Cotes nine-point scheme is used to perform the numerical integration, and Brent’s method is used to find the root of \(g(.)\).

Each member of the simulated animal population continues to take such steps until the simulation’s end-time is reached.
Westing correction

Because the mean of this uniform distribution is $\pi$, a bias towards a westing heading is possible.

This potential bias is avoided by rotating the heading distribution coordinate system through a random angle each time a step is simulated.
Statistical estimation of SMS parameters

Classical maximum likelihood needs a mathematical form of the likelihood of each and every possible sample.

Simulation models rarely possess such a function.

If the simulation model is *simulable*, the unknown likelihood function may be approximately maximized with Maximum Simulated Likelihood Estimation (MSLE) as it replaces the actual likelihood with a simulated approximation to it.
Say that $n$ animals have been observed moving through the landscape over a time interval.

Let $p_j$ be the path of the $j^{th}$ observed animal. A *path* is an ordered list of vertices, $P = \{v_1, \ldots, v_n\}$.

Let $r_{ijk}$ be the dissimilarity between the path taken by the $i^{th}$ simulated animal during the $k^{th}$ simulation run and the $j^{th}$ observed animal that shares the same demographic values.
Our measure of path dissimilarity, due to Martí et al. (2009), is:

\[ r = \frac{1}{2} \left[ \frac{\sum_{v_i \in P_1} \delta(v_i, P_2)}{|P_1|} + \frac{\sum_{u_j \in P_2} \delta(u_j, P_1)}{|P_2|} \right] \]

where \( \delta(v, P_1) = \min_{v_j \in P_1} \delta(v, v_j) \) and \( |P_i| \) is the number of vertices in path \( P_i \).
Say that animal movement over the observed time interval has been simulated $m$ times.

The path’s PDF at the $j^{\text{th}}$ observed path, $p_j$, may be approximated with a $v$ nearest-neighbor, nonparametric density estimator by setting it equal to the inverse of the (scaled) dissimilarity between $p_j$ and the $v^{\text{th}}$ most-similar simulated path.
Let $\tilde{f}(p_j)$ denote this approximate PDF. The value of $v$ equals $\alpha \times m$ where $\alpha$ is a small value between 0 and 1 (typically about 0.05).

The idea here is that the PDF at path $p_j$ becomes small as the dissimilarity between $p_j$ and the $v^{th}$ most-similar simulated path increases. Denote the distance between two points in the landscape with $\delta(u, v)$. 
The log-likelihood is approximated with $\sum_{j=1}^{n} \log \tilde{f}(p_j)$. Parameters are adjusted such that this approximate log-likelihood is maximized.
Movement behavior indices

A multivariate approach to goodness-of-fit is to define several movement behavior indices, model them as a multivariate random vector, and create diagnostic plots to assess model-data agreement.
To this end, quantify a path’s tortuosity with its straightness, and sinuosity.

\[ ST = \frac{dE}{L} \] where \( dE \) is the Euclidean distance between the beginning and end of the path; and \( L \) is the path’s length.
Sinuosity

\[ Sl = 2 \sqrt{p \left\{ \frac{1 - c^2 - s^2}{(1 - c)^2 + s^2} + b^2 \right\}} \]

where \( p \) is the mean step length, \( c \) is the mean cosine of the path’s turning angles, \( s \) is the mean sine of the path’s turning angles, and \( b \) is the coefficient of variation of step length.
A path’s spatial range, $R$ (the path’s overall spatial range) is a measure of the animal’s search effort.

Let *net displacement* be the distance between the path’s starting location and a subsequent path location.
**MND** and **SND**

**MND**: average net displacement

**SND**: standard deviation of net displacement

A resident has low values of both **MND** and **SND**. A transient has a high **MND** value and a low **SND** value.
Heatmap

Realizations of a multivariate random vector may be portrayed with a heatmap.

Each path is assigned a row, and each variable, a column in a matrix. The value of a path variable is indicated by a color in the associated box.

A path’s row is determined by its density value computed under the fitted SMS model.
The following is a heatmap of the movement behavior indices ( $ST$, $SI$, $R$, $MND$, $SND$) of the observed paths along with those simulated with the fitted model.

The path label “OBSi” indicates the path of the $i^{th}$ observed rhino, and “si” indicates a simulated path of this animal.

Goodness of fit improves as the observed path rows are more evenly mixed-in with those of simulated paths.
Working with a wildlife agency

1. Do some free work for them and then have a chat.
2. Propose a formal but unfunded project.
3. Now, jointly approach a conservation organization for funding.
4. Be patient but persistent about acquiring data.
5. Move with them.