

Stochastic Individual-Based Model of Animal Movement

Timothy C. Haas¹ & Sam M. Ferreira²

¹ Lubar School of Business, University of Wisconsin-Milwaukee, Milwaukee, USA

² Scientific Services, SANParks, Skukuza, South Africa

DRAFT: Please do not quote.

Abstract

??

Acknowledgement: Travel for Timothy C. Haas was supported by a SANParks Visiting Scholar grant.

keywords: animal movement modeling, stochastic movement simulator, individual based models

1 Introduction

We describe an individual-based model (IBM) of the spatio-temporal movement through a landscape of either a terrestrial animal such as a rhinoceros, or a semi-aquatic animal such as a crocodile (see “crocodile” (2009)). Such a model is called a Stochastic Movement Simulator (SMS) (see Palmer et al. (2010)).

1.1 Previous Work

1.1.1 A Grid-Based Approach

Aben et al. (2014) models the movement of several species of birds by expanding a stochastic IBM of animal movement due to Palmer et al. (2010). To begin, the authors discretize the birds’ landscape into a grid composed of $5 m$ by $5 m$ cells. At each simulation time step, each bird selects one of the eight cells adjacent to its current cell by drawing a cell randomly from a probability distribution over these eight cells. Cell probabilities are proportional to the inverse of a cell’s “effective cost.” Viewing cost as negative utility, Aben et al. (2014)

define a bird’s utility of moving to a cell to be determined by two variables: *directional persistence*, and *goal bias*. The authors give a bird four characteristics: an ability to move exactly one cell during a simulation time step, knowledge of its home territory location, a *perceptual range*, and a *memory size*. A cell’s directional persistence utility increases if the animal does not have to make a sharp turn to enter that cell. Direction is computed from the locations of cells visited by the animal over a number of previous steps. The animal’s memory size determines this number of previous steps. A cell’s goal bias utility increases if moving to that cell brings the bird closer to its home territory.

A cell’s utility is the harmonic mean of the directional persistence and goal bias values taken over that cell and a number of adjoining cells. The number of these adjoining cells is determined by animal’s perceptual range.

1.1.2 A Continuous-Space Approach

Latombe et al. (2014) use data on the distance that a caribou covers over different time intervals to form an empirical distribution of caribou step length. Data on turning angle of caribou paths is used to form an empirical distribution of turning angle. The simulation is conducted over four-hour time steps. At each time step and for each caribou, 21 step length-turning angle pairs are randomly drawn from the step length and turning angle distributions. A utility “score” is computed for each of these pairs. The pair having the highest score is executed by that caribou. This score is a function of the individual’s current energy level, the amount of food at the potential new location, the new location’s proximity to anthropogenic features such as roads, and the individual’s memory of the most recently-visited locations. Note that this movement model does not rely on a discretized landscape.

This model consists of two submodels: the caribou population and the environment. Energy budgets in both submodels are quite detailed and change as lichens grow and are consumed by caribou. An IBM that lacks a movement submodel but otherwise contains a detailed energy budget submodel such as the rhino IBM of Haas and Ferreira (2015) is easily extended to include an animal movement submodel.

Note that because they are modeling a herbivore, the authors can take advantage of

satellite imagery to map animal food sources.

1.2 Our Approach

The model described below is similar to the caribou model of Latombe et al. (2014). As such, it does not require a grid to be imposed on the region under study. Not only does this allow easy integration of the model's input and output with a vector-based GIS system such as QGIS, but also separates the size of an animal's step from the coarseness of the landscape's GIS layers.

2 Model Description

Let h be an animal's *heading* expressed as the angle from easting, i.e., $h = 90$ is due north. A terrestrial animal such as a rhino, is usually able to move in any direction (heading) it desires. Restrictions on a terrestrial animal's heading are usually due to obstacles such as fences, mountains, or rivers that cannot be forded. Such obstacles may completely close off a region to the animal's access or merely force the animal to enter the desired region via a circuitous path. Although capable of moving overland, a semi-aquatic animal's main mode of movement is swimming along river reaches. But rivers often open into either a delta, a mudflat, or a marsh. Such areas are marked by high soil moisture and hence while in such an area, a semi-aquatic animal has nearly the same range of headings available to it as a terrestrial animal. Therefore, a general model of animal movement should allow either type of animal to select h from the interval $[0, 360)$.

We model both types of animals as able to arrive at utility values for locations (point B) located some distance, d away from their current location (point A). These animals compute d to be the shortest *traversable* path from point A to point B. For a terrestrial animal, a traversable path is one that does not cross obstacles. Semi-aquatic animals have more restrictions than a terrestrial one on where they can step: the main one being sufficient soil moisture. Here, soil moisture is defined broadly to include dry soil, moist soil, mud, standing water puddles, shallow water, and flowing water. One way to quantify soil moisture is with the Topographic Wetness Index (TWI) (see Jablonicky (2013)). In our

approach, a traversable path for a semi-aquatic animal is one whose minimum TWI across the path is above the animal’s minimum wetness threshold.

2.1 Movement Algorithm

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 should be a stochastic function of utility rather than simply the heading at which the animal’s utility function is maximized. 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. Next, 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 (covariates) that influence (drive) the animal’s utility function. Let β contain the associated parameters. Say that animal i is presently located at \mathbf{l}_0 . Then that particular 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) $f(h) = u(h) / \int_0^{360} u(y)dy$. Call this the *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.

Our SMS functions as follows. First, n animals are assigned starting locations. Then, as mentioned above, at simulation time t , each animal takes a step of size s along a heading, h_q found by randomly drawing a value from its 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) formed by numerically integrating the heading distribution’s PDF (see above) over the interval $(0, h)$. An adaptive Newton-Cotes nine-point scheme (Hultquist 1988) is used to perform the numerical integration, and Brent’s method (Brent 1973) is used to find the root of $g(\cdot)$.

Each member of the simulated animal population continues to take such steps until the

simulation’s end-time is reached.

Because the mean of this uniform distribution is π , 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.

2.2 Rhino Movement

Once a mother rhino has moved, all of that mother’s calves move to that location also.

2.3 Crocodile Movement

Crocodiles like to eat and mate. Let $twi(\mathbf{l}_0, t)$ be the TWI value at location \mathbf{l}_0 and time t . Let $d_f(\mathbf{l}_0, h, t)$ be the distance to the closest food (prey) from location \mathbf{l}_0 along heading h at time of year t . Likewise, let $d_m(\mathbf{l}_0, h, t)$ be the closest potential mate from location \mathbf{l}_0 along heading h at time of year t . Let \mathbf{c}_i be the demographics of crocodile i (gender, age, or given to wearing funny magnet hats). Then $\mathbf{x}' = (twi(\mathbf{l}_0, t), d_f(\mathbf{l}_0, h, t), d_m(\mathbf{l}_0, h, t), \mathbf{c}'_i)$ and $u(h) = \mathbf{x}'\boldsymbol{\beta}$.

3 Parameter Estimation

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. Here, a *path* is an ordered list of vertices, $P = \{v_1, \dots, v_n\}$. Different paths may have different numbers of vertices. 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. These two animals have the same demographic values. Run this simulation m times.

Because each animal’s utility function may be changing as the animal moves to different locations in the landscape, the path likelihood function may be difficult to obtain analytically. In such a case, Maximum Simulated Likelihood (MSL) can be used to produce parameter estimates that share the same desirable properties as Maximum Likelihood estimates (see for example, Lai and Tsay (2016)).

Say that animal movement over the observed time interval has been simulated m times.

The path's PDF at the j^{th} observed path, p_j may be approximated with a v nearest-neighbor, nonparametric density estimator (see Thompson and Tapia (1990)) by setting it equal to the inverse of the (scaled) dissimilarity between p_j and the v^{th} most-similar simulated path. Let $\tilde{f}(p_j)$ denote this approximate PDF. The value of v equals $\alpha \times m$ where α 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 $\delta(u, v)$. Our measure of path dissimilarity, due to Martí et al. (2009), is:

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

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 .

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.

3.1 Goodness-of-Fit

A graphical assessment of a movement model's agreement with one observed path of an animal of type j can be made with a plot of 100 simulated paths all traced in one color together with the observed path in a different color.

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. The first two indices used herein are measures of a path's *tortuosity* (see Almeida et al. (2010)). These are the path's *straightness*, and *sinuosity*. Straightness is defined as $ST = dE/L$ where dE is the Euclidean distance between the beginning and end of the path; and L is the path's length. Note that if the SMS employs a pre-determined, fixed number of steps, L is constant for all animals and all simulated paths. Sinuosity is defined as

$$SI = 2\sqrt{p \left\{ \frac{1 - c^2 - s^2}{(1 - c)^2 + s^2} + b^2 \right\}} \quad (2)$$

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. Again, if simulated steps all of the same size, $b = 0$. For a semi-aquatic animal who is largely river bound, the only possible angles are zero and 180 degrees and hence sine is always zero and cosine is always either -1 or 1. Therefore, for a semi-aquatic animal, sinuosity reduces to

$$SI = 2\sqrt{p \left\{ \frac{1 - c^2}{(1 - c)^2} + b^2 \right\}}. \quad (3)$$

The third index, R is the path's overall spatial range – a measure of the animal's search effort. The fourth and fifth indices, MND and SND are the average, and standard deviation of the path's *net displacement* defined to be the distance between the path's starting location and a subsequent location on the path, respectively. These two indices can be used to assess the animal's migratory state (Quinn et al. 2012). A resident, e.g. an animal engaged in browsing has low values of these two indices. A transient animal has a high MND value and a low SND value.

These five indices, $\{ST, SI, R, MND, SND\}$ may exhibit pairwise correlation but are linearly independent of each other. One way to portray realizations of a multivariate random vector is with a *heatmap* (see Ivanisevic et al. (2014)). In such a plot, each path is assigned a row, and each variable, a column in a matrix. The value of a variable for a particular path is indicated by a color in that box that ranges from blue (low) to red (high). Here, for each demographically unique animal, m simulated paths are displayed in a heatmap along with their observed paths (Figure 1). The row position of a path simulated or observed is determined by its density value computed with the fitted SMS model. Goodness-of-fit of the SMS to the sample of observed paths improves as the rows denoting observed paths are more evenly mixed-in with those simulated.

3.2 Effect Size

Once the goodness of fit of the model has been ascertained, effect size of driver variables may be investigated. Following recommendations in Johnson (1999, 2005), confidence intervals are used in lieu of hypothesis tests to assess effect sizes of SMS driver variables. Here, this is accomplished by computing jackknife simultaneous confidence intervals of the model's parameters (Haas 2011). These confidence intervals are used to assess the effect size of

each driver variable.

Heatmaps for each demographically-unique observed crocodile (Figure 2) show ??.

Example of Three Observed Paths on a TWI Surface

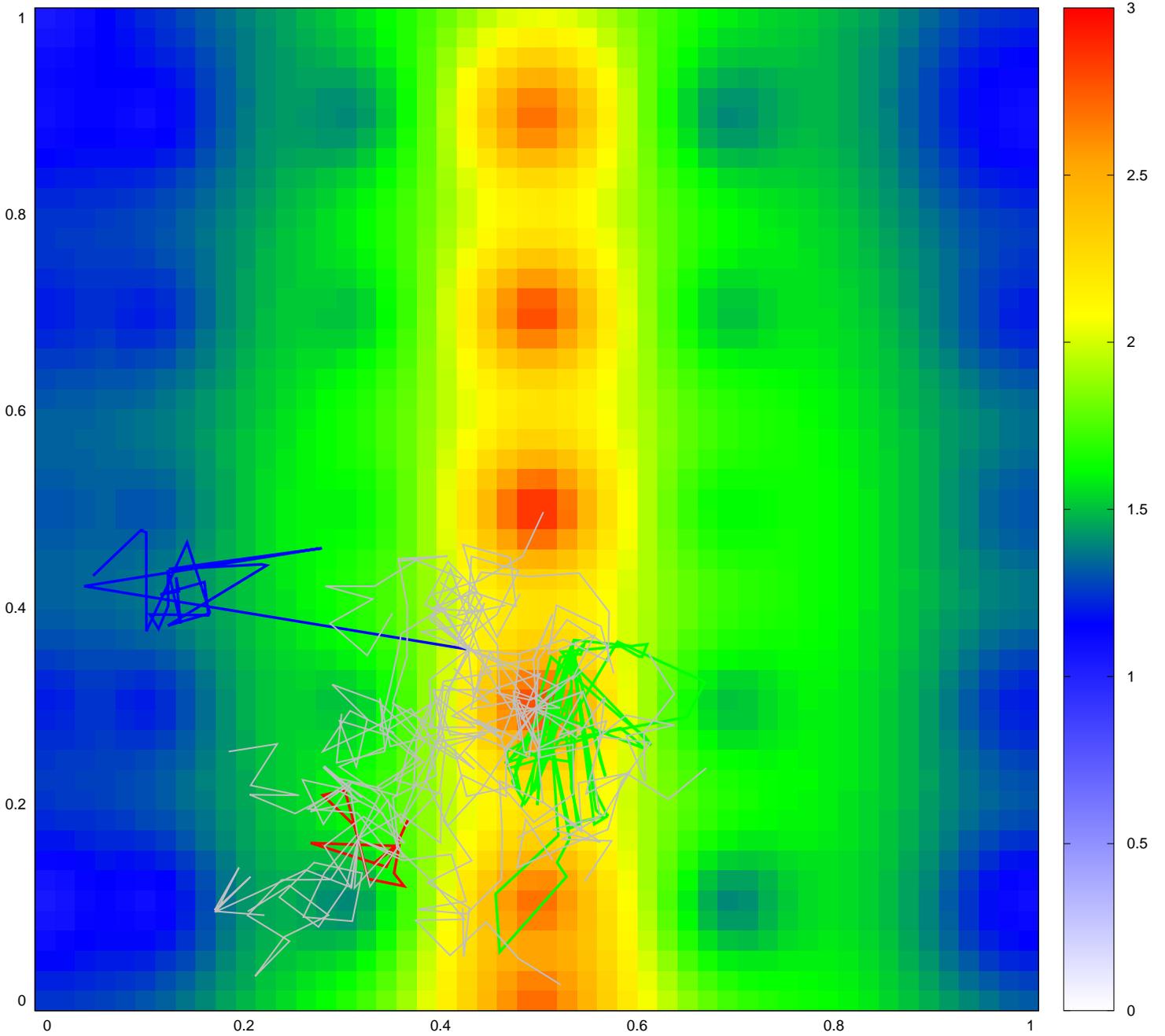


Figure 1: Paths of three observed animals along with 21 simulated paths generated by the SMS model. TWI is the plotted spatial variable.

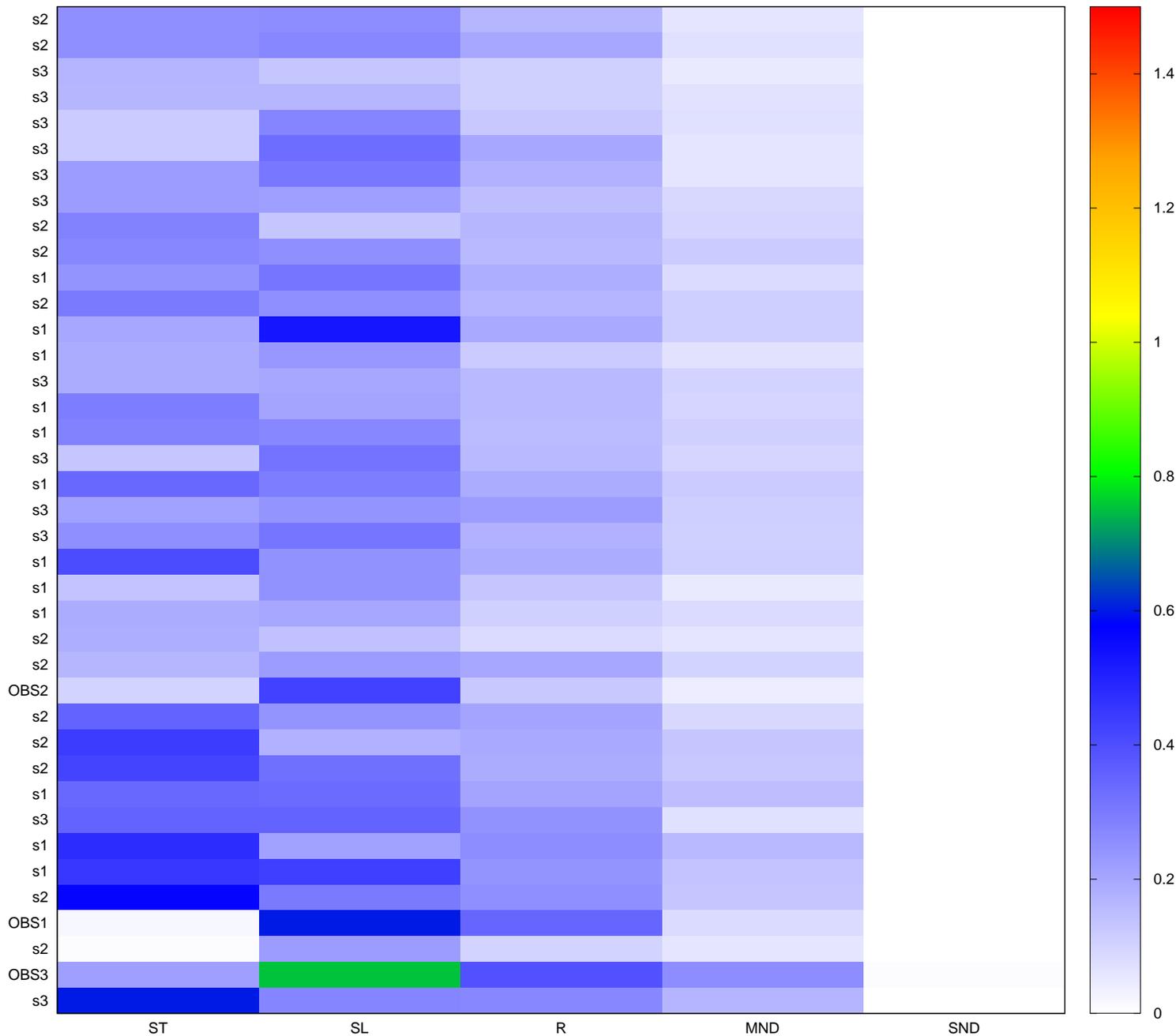


Figure 2: Heatmap of movement behavior indices (ST , SI , R , MND , SND) of the observed paths along with those simulated with the fitted model. The path label “OBS i ” indicates the path of the i^{th} observed animal, and “ s_i ” indicates a simulated path of this animal. Paths are ordered by their density computed at the fitted model. The observed paths have the lowest densities indicating this model’s lack of fit.

4 References

- Aben, J., Strubbe, D., Adriaensen, F., Palmer, S. C. F., Travis, J. M. J., Lens, L., and Matthysen, E. (2014), “Simple individual-based models effectively represent Afrotropical forest bird movement in complex landscapes,” *Journal of Applied Ecology*, 51: 693-702.
- Almeida, P. J. A. L., Vieira, M. V., Kajin, M., Forero-Medina, G., and Cerqueira, R. (2010), “Indices of movement behavior: conceptual background, effects of scale and location errors,” *Zoologia*, 27(5): 674-680.
- Brent, R. P. (1973), *Algorithms for Minimization Without Derivatives*, Englewood Cliffs, N.J.: Prentice-Hall, pp. 61-80.
- “crocodile.” The Oxford Pocket Dictionary of Current English (2009), Retrieved April 15, 2016 from Encyclopedia.com:
<http://www.encyclopedia.com/doc/10999-crocodile.html>
- Haas, T. C. (2011), *Improving Natural Resource Management: Ecological and Political Models*, Chichester, U.K.: Wiley, pp. 190-193.
- Haas, T. C. and Ferreira, S. M. (2015), “Conservation Risks: When will rhinos be extinct?” to appear in the special issue of *IEEE Transactions on Cybernetics* on Risk Analysis in Big Data Era, online at
<http://ieeexplore.ieee.org/xpl/tocresult.jsp?isnumber=6352949>.
- Hultquist, P. F. (1988), *Numerical Methods for Engineers and Computer Scientists*, Menlo Park, CA: The Benjamin/Cummings Publishing Company, pp. 230-236.
- Ivanisevic, J., Benton, H. P., Rinehart, D., Epstein, A., Kurczy, M. E., Boska, M. D., Gendelman, H. E., and Siuzdak, G. (2014), “An interactive cluster heat map to visualize and explore multidimensional metabolomic data,” *Metabolomics*, DOI 10.1007/s11306-014-0759-2.
- Jablonicky, C. A. (2013), *Spatial distribution of the Nile crocodile (Crocodylus niloticus) in the Mariarano River system*, Master of Science Thesis, University of Southern California, Los Angeles, California. Retrieved April 15, 2016 from
<http://digitallibrary.usc.edu/cdm/ref/collection/p15799coll13/id/130996>

- Johnson, D. H. (2005), “What hypothesis tests are not: a response to Colgrave and Ruxton,” *Behavioral Ecology*, Jan./Feb., 16(1): 323-324.
- Johnson, D. H. (1999), “The insignificance of statistical significance testing,” *Journal of Wildlife Management*, 63: 763-772.
- Lai, H. and Tsay W. (2016). “Maximum simulated likelihood estimation of the panel sample selection model,” *Econometric Reviews* (to appear).
- Latombe, G., Parrott, L., Basille, M., and Fortin, D. (2014), “Uniting statistical and individual-based approaches for animal movement modelling,” *PLOS ONE*, 9(6): e99938. doi:10.1371/journal.pone.0099938
- Martí, R., González, J. L., and Duarte, A. (2009). “Heuristics for the bi-objective path dissimilarity problem,” *Computers and Operations Research*, 36: 2905-2912.
- Palmer, S. C. F., Coulon, A., and Travis, J. M. J. (2010), “Introducing a ‘stochastic movement simulator’ for estimating habitat connectivity,” *Methods in Ecology and Evolution*, doi: 10.1111/j.2014-210X.2010.00073.x
- Quinn, A. C. D., Williams, D. M., and Porter, W. F. (2012), “Postcapture movement rates can inform data-censoring protocols for GPS-collared animals,” *Journal of Mammalogy*, 93(2): 456-463.
- Thompson, J. R. and Tapia, R. A. (1990), *Nonparametric Function Estimation, Modeling and Simulation*, Philadelphia, PA: Society for Industrial and Applied Mathematics.
- Waikawait, W.E. (1989), “Present knowledge on the west African slender-snouted crocodile, and the west african dwarf crocodile,” (in) *Crocodiles: Their Ecology, Management, and Conservation*, Gland, Switzerland: IUCN Publication New Series, p. 265. (shows some river morphological effects on habitat selection)

Additional References

- Turner, M., Wu, YI, Wallace, L., Romme, W., and Brenker, A. (1994). “Simulating winter interactions among ungulates, vegetation, and fire in Northern Yellowstone Park,” *Ecological Applications*, 4: 472-496.
- Allen, A. M. and Singh, N. J. (2016). “Linking movement ecology with wildlife management and conservation,” *Frontiers in Ecology and Evolution*, 3,

doi: 10.3389/fevo.2015.00155

Kevin Buchin et al. (2015), “Deriving movement properties and the effect of the environment from the Brownian bridge movement model in monkeys and birds,” *Movement Ecology*.

Wiebke Neumann et al. (2015), “Opportunities for the application of advanced remotely-sensed data in ecological studies of terrestrial animal movement,” *Movement Ecology*.

Gil Bohrer et al. (2014), “Elephant movement closely tracks precipitation driven vegetation dynamics in a Kenyan forest-savanna landscape,” *Movement Ecology*.