

# Chapter 12

## An Implementation of the Pathway Analysis Through Habitat (PATH) Algorithm Using NetLogo

William W. Hargrove and James D. Westervelt

### 12.1 Background

When developing plans to protect populations of species at risk, we tend to focus on protecting and preserving habitat where the species naturally occurs. Ecologists employ historic sightings of individuals, habitat suitability index (HSI) models, and the expert advice of local naturalists familiar with the environments desired and required. In many cases, the optimal habitats are fragmented to the extent that a species at risk could be extirpated from any single area by storms, drought, disease, or other local insult. Alternately, over time, a small population could become so inbred that it becomes susceptible to disturbances that increase the probability of local extirpation. However, if nearby populations can reach the isolated population across the fragmented space, then the isolated population becomes part of a larger metapopulation. The strength of habitat connections, therefore, can be critical to ensure the viability of at-risk populations over time. Habitat connectivity plays this important role by increasing the effective population size, maintaining gene flow, and facilitating regular migration, dispersal, and recolonization. Each of these processes helps ensure the long-term persistence of a population. A connected landscape is preferable to a fragmented one (Beier and Noss 1998; Bennett 1999). Natural landscapes are generally more connected than landscapes altered or manipulated by humans, so establishing or maintaining corridors is a viable strategy to sustain the natural connectivity (Noss 1987).

---

WW Hargrove (✉)

USDA Forest Service Southern Research Station, Eastern Environmental Threat Assessment Center, 200 W.T. Weaver Blvd., Asheville, NC 28804-3454, USA  
e-mail: whargrove@fs.fed.us; hnw@geobabble.org

J.D. Westervelt

Construction Engineering Research Laboratory, US Army Engineer Research and Development Center Champaign, IL USA  
e-mail: james.d.westervelt@usace.army.mil

*Metapopulation* is the term for a collection of discrete local breeding populations that occupy distinct habitat patches but are connected by migration (Hanski and Gilpin 1997). Population viability analysis (PVA), a method for forecasting the probability that a metapopulation will persist over time, has been automated by numeric modeling applications such as RAMAS (Applied Biomathematics 2003). A key input to a PVA is the probability of migration among all individual populations of the metapopulation. Understanding the viability of an at-risk population or metapopulation depends to a significant extent on identifying the migration routes that connect the discrete constituent populations. Also, without this understanding, land managers may not have enough information to protect important migration paths. That deficiency, in turn, may lead to the loss of more-isolated populations, an increased threat to the survival of the greater metapopulation.

In order to help improve our understanding of species migration routes among separated populations, Hargrove, Hoffman, and Efrogmson (2005) developed the Pathway Analysis Through Habitat (PATH) computer simulation model. Originally developed for implementation on a supercomputer, the purpose of PATH is to help a decision-maker to reliably predict where potential dispersal corridors are likely to exist in real-world landscape maps. This information makes it possible to project which habitats will support population growth (sources) and which will tend to lose population (sinks). The PATH algorithm works by launching *walkers* (i.e., virtual animals) from each habitat patch to simulate the journey of individuals through land cover types in the intervening matrix until arriving at a different habitat patch or dying. Each walker is given a set of user-specified habitat preferences that direct its walking behavior to resemble the animal of interest. As originally implemented, PATH was designed for a massively parallel computing environment in order to analyze the activity of very large numbers of random walkers in large landscapes with many habitat patches. PATH produces three outputs: (1) a map of the most heavily traveled potential migration pathways between patches, (2) a square transfer matrix that quantifies the flow of animals successfully dispersing from each patch to every other patch, and (3) a set of importance values that quantifies, for every habitat patch in the map, the contribution of that patch to successful animal movement across the landscape. The transfer matrix is square and not triangular since the rate of animal movement is likely to be asymmetrical between any two habitat patches. That is, the rate of successful migration from patch 1 to patch 2 will likely not equal the successful migration from 2 to 1.

One problem with the original implementation of PATH is the inherent barrier to its use: the application is designed to perform massive simulations that require a supercomputer to run them in a reasonable amount of time. This project was conceived as a way to make migration pathway analysis more accessible to users of desktop computers by implementing the core PATH algorithm in the NetLogo simulation modeling environment (Wilensky 1999).

## 12.2 Objective

The objective of this modeling project was to develop a computationally efficient PATH-based tool for desktop computers that can identify important species migration corridors between habitats based on expert information about inter-habitat patch lethality, the energy cost to cross, and the energy available to animals to move outside of core habitat.

## 12.3 Model Description

### 12.3.1 Purpose

This implementation of PATH is intended to illuminate the essential mechanisms that help to identify animal migration corridors.<sup>1</sup> A premise of the design was to avoid including any data or processes not directly illuminating the successful migration of species between two or more separated habitat patches. This model converts expert knowledge about habitat patch locations, traversal cost, and probability of mortality through the interstitial landscape into information about the relative connectivity of all pairs of habitat patches and the impact of interstitial lands on successful migrations. In future versions, certain additional aspects of “realism,” such as seasonal effects and time steps, could be added in instances where model results might be improved. As you read further you will note that, unlike the other models in this book, this model does not use time steps.

### 12.3.2 State Variables and Scales

The scale and extent of the area are determined by the user, with the practical limit depending on the processing power of the user's computer. The demonstration data set used in this model supports a simulation space encompassing more than 1,000 by 1,000 cells (i.e., one million cells). The primary state variable that changes over simulation time is the number of successful migrations supported by each patch in the simulation space.

---

<sup>1</sup>An operational copy of this model is available through <http://extras.springer.com>.

### ***12.3.3 Process Overview and Scheduling***

The model is initialized by reading data from a set of three location-specific maps that characterize (1) the location of habitat patches, (2) the energy cost to cross the space between patches, and (3) the lethality associated with migrating through non-habitat. During initialization, first the habitat map loads, then the model aggregates, outlines, and numbers all contiguous habitat patches (i.e., “habitats”). A user’s chosen number of travelers is initialized. The simulation runs in discrete steps, but these steps do not represent time because time has no bearing on an individual’s energy level, the energy cost of migrating, or patch lethality. At each step in the simulation, walkers are randomly distributed along the edges of a habitat based on area size. Walkers are faced away from the habitat interior and started on a walk that is partially directed with an adjustable level of randomness. The amount of energy consumed during one step is based on the data provided by the energy cost input map, and travelers randomly die according to the patch-specific probability of mortality read from the patch lethality input map. As a traveler moves through the interstitial space between habitats, it remembers its course. If the traveler succeeds in arriving alive at a habitat patch different from where it started, it communicates to every patch along the successful path that the patch supported a successful crossing. Each patch then updates a habitat-to-habitat crossing array that tracks the number of successful crossings among all patch combinations.

### ***12.3.4 Design Concepts***

#### **12.3.4.1 Emergence**

PATH reveals the value of every patch in supporting habitat-to-habitat migration through the emergent behavior of the walkers, as an increasing number of them successfully complete migrations and the most favorable paths become more evident.

#### **12.3.4.2 Stochasticity**

Walkers begin their attempted migrations by facing away, randomly, from the interior of their beginning habitats, and then proceed to move through the space separating habitat patches. The travel direction of walkers may be set to be fully random, partially random, or fully deterministic depending on the value selected for a user-set variable. The movement of a walker at each simulation step may be characterized as a “wobble,” turning to the left between 0 and  $X$  angular degrees (where  $X$  is a user-selected value); then turning to the right between 0 and  $X$  degrees. If the user sets  $X$  to 0, then the travel path is straight; if the angle is set to 360, then every step is fully random. The user will assign a value to  $X$  based on what is known of the subject species’ tendency to maintain a direction; movement tendencies vary among species.

#### **12.3.4.3 Observation**

In this version of the model, walkers observe nothing about their surroundings because that information is largely extraneous to path selection on a collective level. However, this model could be modified for experimentation purposes to support species-specific observation and evaluation of travel-direction options by individual walkers.

#### **12.3.4.4 Time**

This model deals with time in a manner very different from any other model in this book, which all include an idea of time passing as the model executes. In this PATH model, each step involves releasing a new batch of walkers, which expend energy and risk death as they move. One may imagine that time passes as they move, but their “time” is not associated with any other walker’s “time” in any model step. Regardless of the number of ticks, the total set of walkers active during the simulation can be assumed to begin their walking at the same time. Also, the state of the landscape never changes as it often does in other models, which leaves the landscape essentially timeless.

### ***12.3.5 Initialization***

Patches are initialized with data read from the habitat-location, energy-cost, and patch-lethality maps. Habitat quality is represented by a binary variable, either 0 for non-habitat or 1 for habitat. The walkers’ energy store is set by the user before migration attempts begin, and the amount of energy lost to walkers as they cross patches is accounted for by data from the model’s energy-cost map. Finally, the probability of mortality while crossing a patch is represented by a value ranging from 0 to 100 as determined by the model’s lethality map. Also during initialization, as briefly noted above, patches are clumped together to form contiguous habitats, and each habitat is given a unique identification number.

### ***12.3.6 Input***

Input is provided by the three raster maps described above. The habitat-location, energy-cost, and patch-lethality maps are prepared within a raster-based geographic information system (GIS) and provided to the model as EsriASCII grid files. PATH requires maps that are location- and species-specific. The habitat map could be produced using the results of a Habitat Suitability Index study. It also might be based on a regression analysis that correlates known habitat and non-habitat patches

with factors in other GIS maps such as slope, land cover, soil type, land use, canopy cover, and elevation. In any case, the end product must be a map of habitat for the target species—places where the species can establish a home range and survive. The lethality map represents the probability (0–100%) that an individual will die while crossing any patch. A patch lethality value is based on exposure to predators, the probability of being caught in an inescapable situation, and the species' ability to deal with the land cover. GIS data involved in developing this information might include land cover, land use, slope, and aspect. The energy-cost map is similar to the lethality map, but instead of encountering immediate death in specific patches, the walker loses energy based on the energy-cost value for the patch being crossed. For example, crossing through dense woods, a swamp, or a pond may require more energy than crossing an open short-grass field. Energy cost, of course, varies with the species. Like the lethality map, the energy-cost map will likely be based on an ecologist's analysis of land cover, slope, and aspect data.

### 12.3.7 Submodels

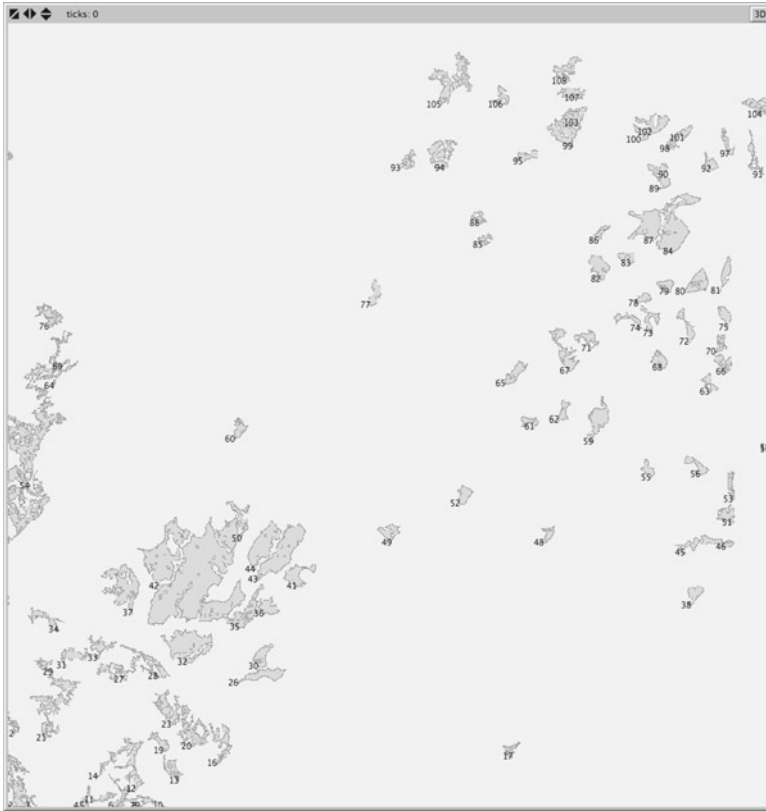
Our NetLogo implementation of PATH sets aside all ecosystem and species functions that are not essential to the modeling of animal migration path formation. It is a very simple model that operates on data provided by three raster maps as initialized with species-specific values set by the user for a small number of variables.

The concept of time is not essential to understanding the establishment of migration paths, so each NetLogo tick (i.e., step) occurs independently of time, and represents only one discrete action that involves walkers attempting to migrate from a home habitat to another habitat. At each NetLogo tick, a user-chosen number of travelers make a crossing attempt. That number is not critical, but the number of total attempts is. The number of travelers leaving each habitat is proportional to the size of the habitat, based on the assumption that larger habitats have larger populations and therefore will send forth more emigrants. Walkers depart home habitats from a random edge location.

As the simulation proceeds, successful migration paths are traced onto the output map in black, and each path becomes denser as more individuals use the same path. These paths accrete into increasingly stable shaded areas that reveal the corridors between habitats that are most successfully used in migration. The darkest traces within these grayscale areas indicate the most successfully used corridors.

## 12.4 Simulation Experiments

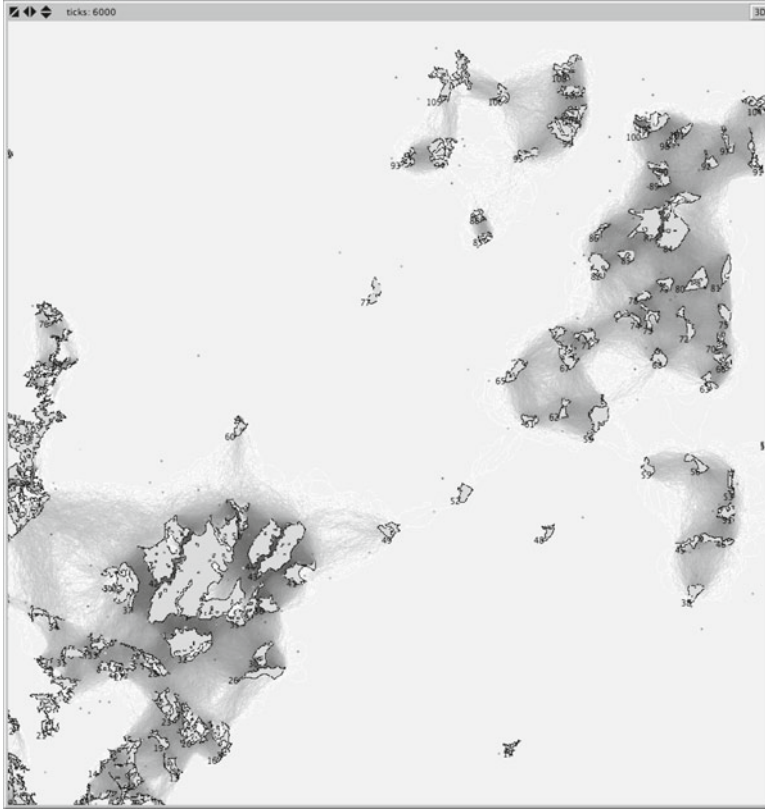
We used this PATH application to identify the corridors within Fort Benning, GA, that might be used by gopher tortoises (*Gopherus polyphemus*) migrating among tortoise habitat fragments. This animal, an at-risk species that has been nominated



**Fig. 12.1** The study area after being initialized

for Federal Threatened status, is carefully managed on Fort Benning. The three GIS maps required for the model were derived from National Land Cover Data (NLCD) maps. Areas with land cover suitable for tortoises were selected, and contiguous areas greater than 18 ha were identified as habitat. Each of the NLCD land cover types was then associated with a transit-energy cost and a probability of mortality (0–100%), and these parameters were then used as the basis for reclassifying the NLCD map into the other two input maps.

Model initialization, which involves reading the three maps as well as identifying and labeling contiguous habitat patches into habitats, takes several minutes and produces the image in Fig. 12.1. The 108 identified habitats are shown in medium gray, with edges outlined in darker gray and uniquely labeled. For this experiment, each walker was given 4,000 units of energy before its crossing attempt and was assigned a maximum turn angle of  $20^\circ$  after each step. Five hundred walkers were arbitrarily initialized at the same time, started at a habitat edge and randomly faced away from the habitat. The total leaving any habitat was based on that habitat's relative size. Each walker moved ahead until it ran out of energy or successfully reached

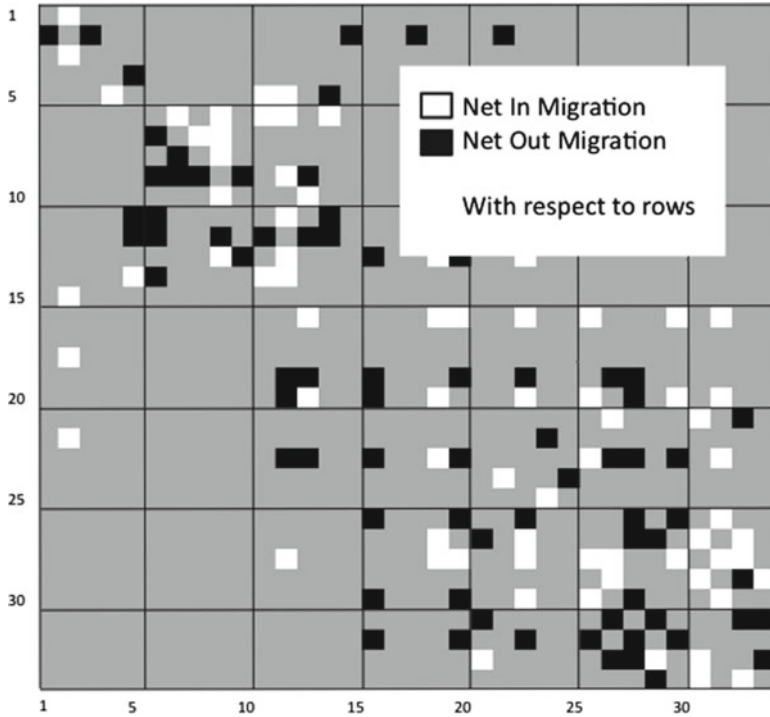


**Fig. 12.2** Successful migration density after three million attempts

a new habitat. Once these walkers were finished, a new set of 500 was initialized and the process continued over and over. In 1 min of “wall-clock time,” the model generated 1.25 million walker attempts, with about 10% successfully migrating from 1 of the 108 habitat patches to another. After three million habitat-crossing attempts were simulated, the patches were assigned a color along a log-adjusted gray scale based on the relative number of successful crossings in which they were used (Fig. 12.2). The specks in the image identify locations at which travelers ran out of energy in their final attempt. The maximum number of successful crossings supported by any one patch was 2,932 and is represented in black. Numerous patches participated in crossing. When the simulation is viewed on a computer screen, these patches are visible as white sinuous lines against the darker background.

Just 357,328 of the 3 million crossing attempts were successful. This value was derived from the inter-habitat migration success table, which can be displayed by selecting NetLogo’s Interface tab. For this simulation, four strong metapopulations emerged, with three of them being tenuously linked (inspect Fig. 12.2). With lower initial energy levels, the number of successful migrations decreased, the number of





**Fig. 12.3** Net in- and out-migration among habitats 1 through 34

separated metapopulations increased, and the number of habitats unconnected to other habitats increased. In terms of species conservation and survival, higher numbers of metapopulations and disconnected habitats are less desirable; and higher numbers of individuals and larger metapopulations are more desirable.

One output of this application is a table that shows the number of successful migrations between all pairs of habitats. By subtracting out-migrations from in-migrations, the user can discover which habitats in each pair are net population sources or sinks. Figure 12.3 shows the net migration from every habitat to every other habitat for habitats 1–34 (out of 108 total). Note that habitat 2 is a net source to habitats 1, 3, 15, 18, and 21 while habitat 16 is a net sink for 7 other habitats.

## 12.5 Discussion

This PATH application reveals that high-quality interstitial habitat is more likely to be associated with corridors than less-preferred habitat, and that short connectors develop more often than longer ones. The model assumes that movement through non-habitat areas is semi-random, with the walker's current heading having some

persistence. Actual animals may functionally select resources at a finer scale than is used in this model, and thus may not follow the same route as one optimized by PATH at a coarser resolution. However animals also may respond to coarser-grained landscape cues than are represented by this model's habitat map, especially when migrating or dispersing across long distances. Animal movement choices and behaviors may vary with age, gender, pregnancy and nutrition status, and the nature and composition of traveling groups. The importance of those factors may differ between species.

The implementation of PATH described here attempts to illuminate patterns of successful inter-habitat migration with a computationally low-overhead modeling approach. It encapsulates essential migration activities and costs into bare fundamentals—a binary habitat indicator, a movement parameter, a randomness parameter, an energy-accounting function, and a mortality probability. The model also includes output basic functions to track and display successful crossings. The map interface shows only the metapopulations and the strength of population connections within them. This information can be concisely output to reveal whether given populations are likely sources and sinks (Fig. 12.3).

Although this PATH-based model is very simple, it is important to note that it is technically sufficient to capture the essential elements of real-world migration path formation for a variety of species. While simpler analytical approaches are possible, such as a least-cost method for predicting path formation, they do not account for multiple terrain types that differ in terms of crossing-energy costs or lethality probability, and are too simplistic to capture all essential considerations. More importantly, because least-cost path modeling methods show the solution to be a single pathway, they illustrate only the current optimal route but do not show secondary routes that could be improved over the current best route through small modifications to the habitat map. The PATH tool shows all feasible connectivity routes, not just the single current best one. Paths that have good potential, but are currently used as secondary routes by the metapopulation, are exactly the ones that resource managers need to know about in order to consider where strategic management alterations can dramatically improve habitat connectivity. Such routes are evident in a PATH output corridor map as strong potential corridors, which may be impeded only in one or two locations by passing through low-quality patches.

The walkers in a simulation, considered individually, have almost no sophistication at finding migration paths as compared with individual real-world animals. However, considered collectively, the behavior of all successful walkers represents a spatial optimization process that can be used to reproduce the optimum pathways we would expect well-adapted individuals of the subject species to use most often. For this reason, it is not a problem that the walkers, endowed only with the ability to "see" land types immediately adjacent to their current location, represent animals that may have much greater sensory scope and range. The simplified walker mode of functioning in the model does not constrain the optimization of the potential corridors found by PATH. The optimization in this model results from the collective action of the large number of successful dispersers and the weighting of the most efficient potential dispersal paths the most heavily. Even if individual walkers in the

model had a greater look-ahead capability, the same optimal potential corridors would be predicted by the PATH tool. Consider, for example, a situation in which the local environment is inhospitable, but just beyond this there is a high-quality habitat pathway to another patch. Although short-sighted, a few walkers will make rare, immediately suboptimal choices and cut through the inhospitable bottleneck to discover the optimized pathway beyond. Becoming successful dispersers, their tracks will become part of the final map, showing the connection just as surely as if they had been able to look beyond the local problem with terrain. A converse situation is also true: pockets of higher-quality habitat that are surrounded by barriers of poor-quality habitat will attract many individual walkers, but it is not likely that those walkers will successfully reach a second patch. Since only the travel paths of successful dispersers are used, no potential corridors will pass through this attractive dead-end area, just as if walkers had been able to see the barriers to migration that lie beyond.

For purposes of real-world application to habitat-management decisions, corridors identified using randomized walkers in this PATH tool must be verified against actual movement corridors that have been observed in the field. Radio tracking and telemetry studies would be appropriate for empirically testing walker-generated corridors against migration paths established by the subject species in the natural environment.

## 12.6 Conclusions

The original PATH model, designed to handle very large numbers of walkers on very large landscapes, was implemented on a supercomputer. Few working resource managers actually have access to a supercomputer. This implementation of the PATH algorithm as a simplified NetLogo model makes PATH available to a much wider user community because it works well on a standard desktop computer. This PATH-driven model identifies and highlights areas in a landscape that contribute to the natural connections among populations, identifies the metapopulation structure, and indicates the relative strength of connections holding a metapopulation together. This information is essential to making effective habitat-management decisions that support robust populations of species at risk.

## References

- Applied Biomathematics (2003) Ramas software. Copyright 1982–2007. Setauket. <http://www.ramas.com/> Accessed 2/26/2010
- Beier P, Noss RF (1998) Do habitat corridors provide connectivity? *Conserv Biol* 12(6):1241–1252
- Bennett AF (1999) Linkages in the Landscape: The Role of Corridors and Connectivity in Wildlife Conservation. International Union for Conservation of Nature and Natural Resources (IUCN) Forest Conservation Programme. *Conserv Forest Ecosystems*, ISBN 2831702216, 9782831702216, 24(1):244

- Hanski IA, Gilpin ME (eds) (1997) *Metapopulation biology*. Academic, San Diego
- Hargrove WW, Hoffman FM, and Efronson RA (2005) A practical map-analysis tool for detecting potential dispersal corridors. *Landsc Ecol* 20(4):361–373
- Noss RF (1987) Corridors in real landscapes: a reply to Simberloff and Cox. *Conserv Biol* 1(2): 159–164
- Wilensky U (1999) *NetLogo*. Computer software. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston. <http://ccl.northwestern.edu/netlogo/>. Accessed 2/26/2010