

GRAPH THEORY AS A PROXY FOR SPATIALLY EXPLICIT POPULATION MODELS IN CONSERVATION PLANNING

EMILY S. MINOR¹ AND DEAN L. URBAN

Nicholas School of the Environment, Duke University, Durham, North Carolina, USA

Abstract. Spatially explicit population models (SEPMs) are often considered the best way to predict and manage species distributions in spatially heterogeneous landscapes. However, they are computationally intensive and require extensive knowledge of species' biology and behavior, limiting their application in many cases. An alternative to SEPMs is graph theory, which has minimal data requirements and efficient algorithms. Although only recently introduced to landscape ecology, graph theory is well suited to ecological applications concerned with connectivity or movement. This paper compares the performance of graph theory to a SEPM in selecting important habitat patches for Wood Thrush (*Hylocichla mustelina*) conservation. We use both models to identify habitat patches that act as population sources and persistent patches and also use graph theory to identify patches that act as stepping stones for dispersal. Correlations of patch rankings were very high between the two models. In addition, graph theory offers the ability to identify patches that are very important to habitat connectivity and thus long-term population persistence across the landscape. We show that graph theory makes very similar predictions in most cases and in other cases offers insight not available from the SEPM, and we conclude that graph theory is a suitable and possibly preferable alternative to SEPMs for species conservation in heterogeneous landscapes.

Key words: connectivity; conservation; graph theory; habitat patch; *Hylocichla mustelina*; population source; reserve design; SEPMs; site prioritization; spatially explicit population models; stepping stone; Wood Thrush.

INTRODUCTION

In fragmented landscapes, conservation of focal species often involves the protection of land to maintain the long-term health of the population of interest. The protected land may consist of one or multiple habitat patches that can be distributed over a large geographical area and might be chosen for size, quality, and/or connectivity attributes. Often, however, the spatial location of reserve sites, and thus connectivity between sites, is ignored. Here we present a method of site prioritization that incorporates all three criteria and requires only minimal computational capability and knowledge of the focal species.

Reserve site selection has been approached in several different ways. There are multiple site selection strategies in the literature for conservation of biodiversity (Moilanen and Cabeza 2002, Noss et al. 2002, Rothley et al. 2004, Hess et al. 2006, Turner and Wilcove 2006), although most of these are not spatial in nature. A more sophisticated but more complicated alternative, usually done for a single focal species, is a spatially explicit

population model (SEPM; Conroy et al. 1995, Dunning et al. 1995b, Turner et al. 1995, South 1999). SEPMs represent a landscape using patches or cells and explicitly identify the location of every object of interest (e.g., individual, population, cell, or habitat patch). These common models can simulate birth, mortality, and movement of individuals or populations and are often considered the best means of predicting the response of organisms to habitat change or other broadscale landscape processes. However, their utility is constrained by their computational and data requirements. Because they are very computationally intensive, they are limited in both number of habitat patches and individuals they can simulate. They also require knowledge of species' demographic parameters such as mortality and fecundity rates, in addition to knowledge of dispersal and other behaviors. These data are often difficult to obtain and can include much uncertainty, which can render the results unreliable. A third alternative for site prioritization, which incorporates connectivity, yet is valued for its computational ease and minimal data requirements, is graph theory. Graph theory has been used in other fields for years but has just recently been proposed as a tool for predicting species occurrences in fragmented habitats and highlighting important patches for conservation (Keitt et al. 1997, Bunn et al. 2000, Urban and Keitt 2001). However, it remains to be seen whether graphs have the same

Manuscript received 23 June 2006; revised 14 December 2006; accepted 14 February 2007. Corresponding Editor: Y. Luo.

¹ Present address: University of Maryland Center for Environmental Science, Appalachian Laboratory, Frostburg, Maryland 21532 USA. E-mail: emminor@al.umces.edu

predictive capabilities as a SEPM. The objective of this paper is to compare the performance of a graph theoretic model to that of a SEPM, using the Wood Thrush (*Hylocichla mustelina*) in the North Carolina Piedmont (USA) as a case study. We will identify important patches for Wood Thrush conservation using both models and compare the findings to determine whether graph theory can be used as an easy and practical alternative for reserve site selection.

METHODS

Study site and habitat maps

To establish robustness of results across different scales and landscapes, we use two landscapes for this analysis. Landscape 1 is a large-extent, coarse-grained landscape (33 × 50 km) and landscape 2 is a small-extent, fine-grained landscape (13 × 21 km) that is a subsample of the first one (Fig. 1). They are both in Wake County, North Carolina, USA, which is located in the central piedmont region of North Carolina. A classified LANDSAT Thematic Mapper satellite imagery (30-m resolution) from May 2001 was used to identify pixels of hardwood forest, and habitat patches were defined as contiguous hardwood pixels using an eight-neighbor rule. Landscape 1 includes only patches larger than 50 ha and contains 126 patches. Landscape 2 falls within the northeast portion of landscape 1, includes all hardwood patches larger than 5 ha, and contains 172 patches. In both landscapes, we computed the shortest distance to a non-forest edge for each habitat pixel, which was used to calculate habitat quality for the graph model and to calculate nest predation and parasitism rates for the SEPM (*Methods: Focal organism*). (Note that while the final habitat maps are binary [hardwood forest vs. everything else], habitat quality was determined based on a multi-class landcover map. Therefore, distance to non-forest is not necessarily equivalent to distance to the edge of the patch, since hardwood patches might be surrounded by pine or mixed forest rather than a non-forest cover type.) Although both of the landscapes have a large proportion of forest cover, they are highly fragmented, and most patches have a high edge-to-area ratio: the average distance to a non-forest edge for all habitat patches was 58.5 m in landscape 1 and 48.2 m in landscape 2.

Focal organism

The Wood Thrush was selected as the focal species for this study for several reasons. First, it is a well-studied species and there are data in the literature for most of the parameters needed in the SEPM. Second, it is a Neotropical migrant, somewhat of a habitat specialist, and a common host for the brood-parasitic Brown-headed Cowbird (*Molothrus ater*) and so may be especially sensitive to edge effects and forest fragmentation (Roth et al. 1996). Finally, while Wood Thrushes are fairly abundant across most of their range (the eastern United States and southern Canada), they have

been declining in numbers over the past several decades (Sauer et al. 2002) and so are of conservation interest. However, while the Wood Thrush is a compelling subject for this analysis, this is not meant to be an evaluation of Wood Thrush ecology but rather an illustration of a method that could be used with any species.

It is thought that nest parasitism and predation are major factors limiting the breeding success of Wood Thrushes (Friesen et al. 1999, Burke and Nol 2000, Fauth 2000), and these events tend to occur at higher rates near forest edges (Paton 1994, Evans and Gates 1997, Chalfoun et al. 2002). Therefore, any habitat pixel greater than 150 m away from a forest edge was considered interior forest and the highest quality habitat; habitat within 150 m of a forest edge was considered lower quality and was assigned higher rates of nest parasitism and predation in the SEPM. Due to the high edge-to-area ratio of most of the patches, many patches were entirely edge habitat (i.e., low quality) with low reproductive potential.

There is very little information available in the literature about dispersal distance of most songbirds, including Wood Thrushes. However, the little data available indicate that birds may routinely disperse one or more kilometers away from natal territory (Anders et al. 1998, Vega Rivera et al. 1998). Therefore, 1500 m was deemed the maximum distance a dispersing Wood Thrush is likely to travel in any one dispersal bout.

Selecting patches for conservation

In general, it is not feasible to obtain fine-resolution data across an extensive landscape. This holds true for conservation planning; managers rarely know the status of the focal species on every habitat patch in their jurisdiction. However, presence or abundance can be predicted based on size, quality, and connectivity of the patch. Many species have minimum patch size requirements, and larger patches often contain more individuals (Åberg et al. 1995, Hoover et al. 1995, Matthysen 1999, Mörtberg 2001). Quality of a patch affects density and/or reproductive success of a species and might be affected by vegetation type or proximity to human development (Weinberg and Roth 1998, Ortega and Capen 1999, Kristan 2003), among other things. Finally, patches that are connected to other patches by dispersal are more likely to be occupied than isolated patches (Åberg et al. 1995, Dunning et al. 1995a). This is especially true for smaller patches, which may periodically experience local extinction but can be recolonized from neighboring patches if they are well connected. Understanding these three patch attributes (size, quality, and connectivity) is crucial to selecting habitat patches for a protected area network.

Keeping size, quality, and connectivity in mind, there are several kinds of patches that might be good candidates for a reserve: sources, stepping stones, and persistent patches. Sources, which are patches with more

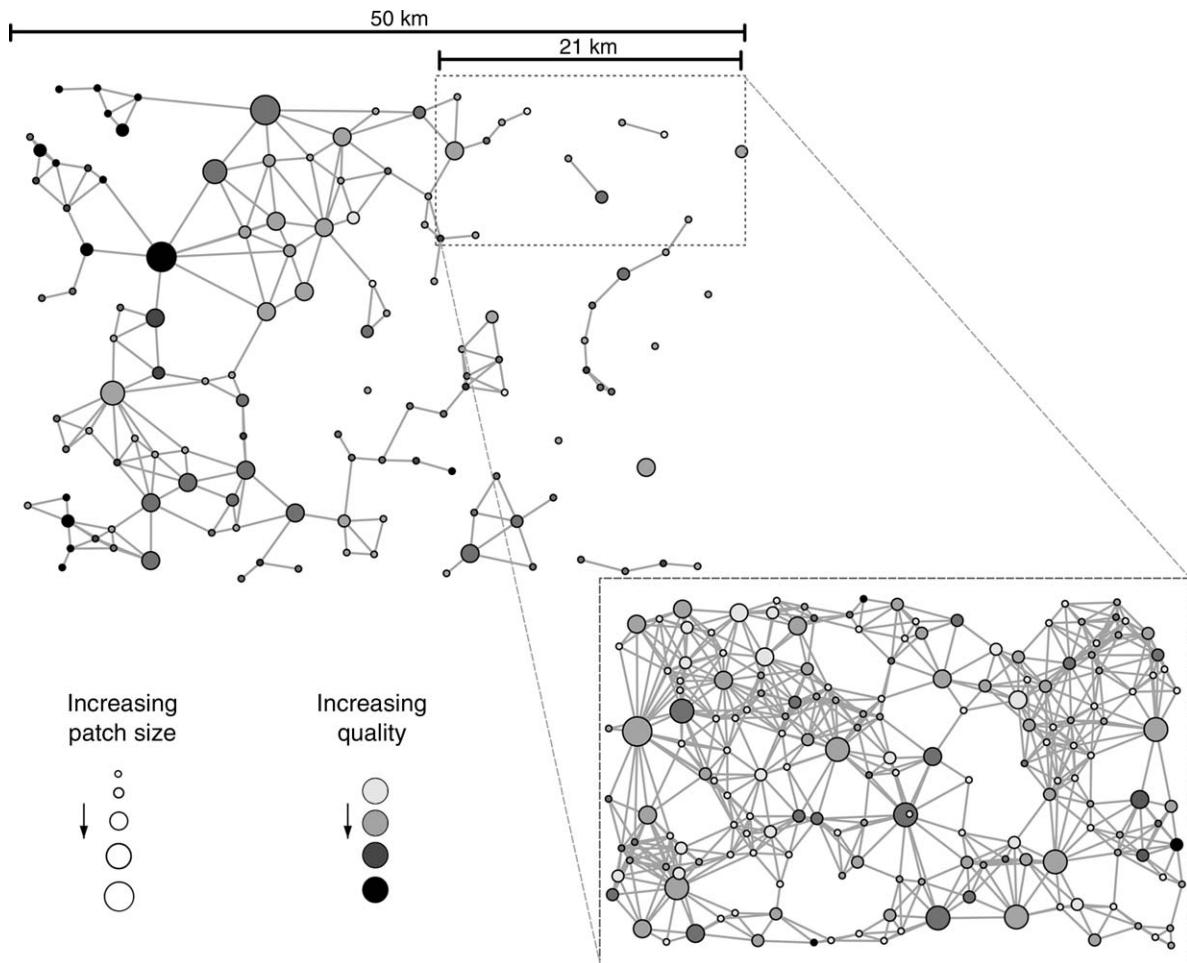


FIG. 1. Landscape 1 (top left) and landscape 2 (inset). Circles of different sizes represent habitat patches of different sizes, and lines between patches (edges) indicate that patches are connected by dispersal. Habitat quality increases from light to dark fill.

births than deaths and more emigrants than immigrants (Pulliam 1988), are usually very desirable additions to a reserve because they are unlikely to go extinct and can sustain the populations of many surrounding patches with their large reproductive output. Stepping stones are patches that are important to connectivity; these are desirable because they allow movement across the landscape and can be critical for connecting source patches to other smaller or lower quality patches. An important stepping stone could be the single link between different parts of the landscape, whose removal would break the landscape into smaller, unconnected sections (this is called a cut-node in graph terminology; *Methods: Graph theory*). Finally, a persistent patch is any patch that maintains a steady population throughout time. A persistent patch could be small and very well connected or simply large, since a very large patch can usually sustain its population without any subsidy from neighboring patches. (Note that these patch types are not mutually exclusive.) A preference for one patch type

over the other in a reserve will depend on the landscape, the focal species, and the conservation goals.

Graph theory

Graph theory (Harary 1969) has traditionally pertained to maximizing flow efficiency in networks or circuits, and consequently has been applied in a variety of disciplines such as information technology and computer science. It also happens to be well suited to ecological applications concerned with connectivity or fluxes and has been used as a framework for food-web theory in ecology (Dunne et al. 2002, Melian and Bascompte 2002). However, it has only recently been used in landscape and metapopulation ecology (Bunn et al. 2000, Urban and Keitt 2001). A graph can represent a landscape of discrete habitat patches as a set of nodes (points) connected to some degree by edges between them (Table 1, Fig. 2). An edge between two nodes implies that there is some flux between those nodes, as in the case of dispersal between two patches. A graph is connected if a path exists between each pair of nodes or

TABLE 1. Definitions of graph terms.

Term	Definition	Ecological relevance	Example
Betweenness	The number of shortest paths that cross through a node; the frequency with which a node falls between pairs of other nodes in the network.	High betweenness implies linkages or "stepping stones" between subgroups. These patches may control flow across a network and be critical for maintaining connectivity.	Nodes 5, 6, and 7 have very high betweenness (Fig. 2).
Component	Nodes that are connected to each other.	If patches are in the same component, they are all mutually reachable. There is no movement between different components, which prevents a "rescue effect" (emigrants from neighboring populations reducing the probability of local extinction; Brown and Kodric-Brown 1977) and implies eventual genetic divergence. Small components may experience permanent extinction.	Fig. 2 shows two graph components, one consisting of nodes 1–11, and the other of nodes 12 and 13.
Cut-node	A node whose removal breaks a single graph component into several smaller ones. Cut-nodes typically have high betweenness.	See betweenness.	Nodes 5, 6, and 7 are cut-nodes (Fig. 2).
Degree	The number of edges adjoining a node, i.e., the number of neighbors joined to a node.	Small patches with low degree may be vulnerable to extinction if their neighbors are removed. Patches with high degree are highly connected and may be population sources or sinks, depending on the size of patch and direction of movement.	Node 2 has a degree of 4, while node 3 has a degree of 2 (Fig. 2).
Edge	Indicates connectivity between nodes. Edges may have direction associated with them (called arcs) and may have attributes assigned to them (e.g., flux).	Shows dispersal between nodes.	
Flux	An edge attribute that indicates amount of movement between nodes. Flux from patch <i>i</i> to patch <i>j</i> is calculated as the quality-weighted area of patch <i>i</i> multiplied by the probability of dispersal between the two patches. Influx and outflux sum up flux for all incoming and outgoing edges, respectively.	High flux between patches indicates a large number of dispersing individuals and thus a higher degree of genetic similarity and a lower rate of patch extinction. High influx indicates a patch with a large number of immigrants; this may or may not be a "sink" patch. High outflux indicates a patch with a large number of emigrants and is a "source" patch.	
Node	The basic element of a graph; the object of interest.	In landscapes, habitat patches can be represented by nodes.	
Path	A sequence of consecutive edges in a network joining any two nodes.	Represents the possible routes an individual may take while traveling across the landscape.	There are multiple paths between nodes 3 and 1: some alternatives include 3→2→1, 3→4→2→1, 3→4→2→5→1. The shortest path is 3→2→1.
QA (quality-weighted area)	A node characteristic equal to patch size (in hectares) multiplied by patch quality. Quality ranges from 0–1 and is defined by the average distance to a non-forest edge for every pixel in a habitat patch.	It is assumed to be related to reproductive potential, in that patches with higher QA produce more offspring than patches with lower QA.	

if organisms can disperse among all the patches. An unconnected graph may include several connected components (e.g., Fig. 2).

There are minimal data requirements for a graph analysis of any particular species and landscape. These include knowledge of the habitat requirements and maximum dispersal distance of the organism, and *x*, *y* coordinates (and optionally, size and habitat quality) of habitat patches. Unlike most SEPMS, graph theory does not require knowledge of behavior, fecundity, or

mortality parameters. However, these data can be incorporated when available and used to produce a more ecologically rich model.

Computationally, a graph can be defined by two data structures: one that describes its nodes (habitat patches) and one that describes its edges (connections between patches). Nodes are defined by their spatial centroid and might also be described by other attributes such as size or some measure of quality. For this analysis, nodes are described by their quality-weighted area (QA), which is

simply patch size (in hectares) multiplied by patch quality. Quality-weighted area is assumed to be related to reproductive potential, in that patches with higher QA produce more offspring than patches with lower QA. Quality ranges from 0–1 and is defined by the average distance to a non-forest edge for every pixel in a habitat patch. Weighting area by quality ensures that unsuitable patches (quality = 0) have no reproductive potential regardless of their size, while the highest quality patches have the maximum reproductive potential for their size.

Graph edges can be defined in several different ways, the most simple of which is based purely on distance between nodes. Defining connectivity on distance alone creates undirected graphs, which have symmetric relationships between nodes. In other words, if a species can disperse from patch i to patch j , then equal dispersal occurs from patch j to patch i . An alternative method is to define edges based on dispersal flux between nodes, or the relative number of individuals moving from one patch to another, which results in directed graphs with weighted edges. If patch i is larger than patch j , then flux from patch i to patch j will be greater than flux from patch j to patch i . A threshold flux value can then be used to define edges, so that there may be an edge from patch i to patch j but not one from patch j to patch i . To infuse further biology into the graph, edges can also be based on least-cost path distances or some other behaviorally based estimate of connectivity between patches (e.g., Bunn et al. 2000).

In this analysis, edges are defined based on distance between patch boundaries, so that patches within 1500 m of each other (i.e., dispersal distance of the Wood Thrush) are considered connected. Dispersal flux is also calculated between each pair of patches connected by an edge. Flux from a donor patch to a recipient patch is calculated as the probability of dispersal between two patches (P_{ij}) multiplied by the reproductive potential (QA) of the donor patch ($\text{Flux}_{ij} = \text{QA}_i \times P_{ij}$). From Urban and Keitt (2001), the probability of dispersal between two patches i and j can be approximated as negative exponential decay:

$$P_{ij} = \exp(\theta \times d_{ij})$$

where θ is a distance-decay coefficient ($\theta < 0.0$) that determines the steepness of the curve and d_{ij} is the distance between the patches. The dispersal distance corresponding to $P = 0.05$ is equal to $\ln(0.05)/\theta$, which allows us to compute θ given a known tail, or maximum dispersal, distance for a species (in this case, the maximum dispersal distance is assumed to be 1500 m and $\theta = -0.002$).

Node degree is the number of edges connected to each patch (Table 1). In-degree and out-degree can be calculated separately, where in-degree is the number of edges coming into a patch and out-degree is the number of edges going out of a patch. These metrics may differ if edges are defined purely on flux, but in this analysis

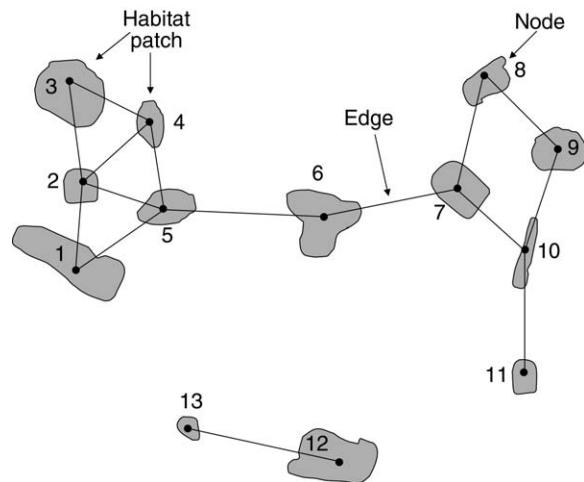


FIG. 2. An illustration of some basic graph terminology. See Table 1 for explanations of numbers (column 4).

edges were defined based on distance and subsequently assigned a flux value. Therefore, if there is an edge from patch i to patch j there is also an edge from patch j to patch i , although those edges are likely to have different flux values. The result is that in-degree and out-degree are equivalent in this analysis (hereafter simply called “degree”).

Influx and outflux are similar to in-degree and out-degree but sum up flux for all incoming and outgoing edges, respectively, rather than just counting the number of edges. Because flux is based on donor patch size (QA) as well as distance between patches, influx is not necessarily equivalent to outflux for any particular patch.

Betweenness is a measure that describes the frequency with which a patch falls between other pairs of patches in the network. It is calculated by finding the shortest paths between every pair of patches in the landscape, then counting the number of times those paths cross each node. Nodes with high betweenness are often linkages or stepping stones between different parts of the landscape. They may control flow through the network and are very important for connectivity.

Graph theory analysis

We used the metrics described above (QA, degree, influx, outflux, and betweenness) to identify patches that are sources, stepping stones, or persistent patches for Wood Thrushes. All graph metrics other than QA were computed using Pajek 1.12 (Batagelj and Mrvar 1996), a free program available for analyzing large networks.

Outflux approximates the source strength of a patch by measuring emigration potential, while influx measures the immigration potential to a patch and should predict a patch’s persistence. Two simple graph metrics, QA and degree, may also be useful indicators of both source strength and persistence. Source patches must

TABLE 2. Parameters used in SEPMs (spatially explicit population models).

Parameter	Value	References
Maximum nest parasitism rate	75%	Donovan et al. (1995), Fauth (2001)
Maximum nest predation rate	65%	Hoover et al. (1995), Brawn and Robinson (1996)
Annual adult survival	0.65	Powell et al. (2000), Roth et al. (1996)
Annual juvenile/floater survival	0.3	Anders et al. (1997)
Clutch size	2.5 eggs	Trine (1998)
No. broods attempted annually	2	Roth et al. (1996)
Territory size	1 ha	Twomey (1945), Weaver (1949)
Dispersal distance	1.5 km	Anders et al. (1998)
Juvenile/floater site fidelity	0.10	Estimated data
Breeder site fidelity	0.90	Estimated data

have high QA as well as be connected to a large number of patches, and persistent patches must either be self-sustaining (i.e., have high QA) or be constantly maintained by incoming migrants (i.e., be connected to a large number of patches). Finally, a patch's value as a stepping stone can be indexed by betweenness.

The spatially explicit population model (SEPM)

The spatially explicit population model used in this analysis (updated from Urban and Shugart 1986) was mechanistic, stochastic, and individual based. It required the same landscape data as the graph model (x , y coordinates of patch centroid, size [or carrying capacity] of each patch, and distance between each patch), but also required an additional nine parameters defining life history traits of the Wood Thrush (Table 2). For this analysis, we used two different initialization schemes. The first one, random initialization, initiated a model run by filling the landscape randomly with half of its carrying capacity (determined by QA of the habitat patches). For example, if the carrying capacity of the landscape is 1000 birds, random initialization will place 500 birds randomly on the landscape (while not giving any one patch more birds than its carrying capacity allows). The second method, single patch initialization, began by filling only one patch to carrying capacity and leaving the others unoccupied. This was done individually for each patch on the landscape, with separate model runs for each patch. After model initialization, both versions proceeded the same way. Based on the

parameters in Table 2, the model simulated natality for each bird, and offspring were created; then, mortality occurred, affecting the adults differently than the juveniles. Average distance to non-forest edge ("edginess") is used in this model as a way to modify natality rates. Edginess is inversely and nonlinearly related to the probability of nest success by a modified Weibull function,

$$P(\text{success}) = b_1 + \exp[-(d/b_2)^5]$$

where d is the average distance to an edge, b_1 is the nest success rate at the edge, b_2 is the distance into the forest that the edge extends (150 m in this case), and the exponent 5 dictates the steepness of the curve. As a result, patches vary in their predation and parasitism rates based on their proximity to non-forested edges.

The next stage was dispersal: if a patch has more birds than its carrying capacity allows, birds must disperse to a nearby patch. Dispersal is a stochastic event based on the size of the patches and the distance between them. Juveniles have lower site fidelity than adults, so they have a higher probability of dispersal. After dispersal, all juveniles became adults and the model began the next cycle; this can be repeated for as many years as desired. The model usually reaches equilibrium (i.e., output stops changing) in ~ 50 years for the random initialization version of the model, but it takes much longer following single-patch initialization. Fifty replicates were run for each version of the model.

TABLE 3. Output metrics from the SEPM.

SEPM output metrics	Definition
BL (number of breeders on whole landscape)	From the single-patch initialization version of the model, the number of birds on the landscape after 100 years following initialization from the focal patch.
BP (number of birds on each patch after 100 years)	From the random initialization version of the model, the number of birds on the focal patch after the model reaches equilibrium.
CV (coefficient of variation of patch population over last 10 years of model run)	From the random initialization version of the model, the coefficient of variation of the patch population in the last 10 years of model runs.
YO (how many of the last 10 years the patch was occupied)	From the random initialization version of the model, the number of the last 10 years that the focal patch was occupied.
Source strength	(rank BL + rank BP)/2
Persistence	(rank CV + rank YO)/2
Overall goodness	(source strength + persistence)/2

TABLE 4. Spearman correlation coefficients between graph metrics and SEPM metrics for landscape 1.

Metric	Graph				SEPM		Graph goodness
	Outflux	Degree	QA	Betweenness	Source strength	Persistence	
Graph							
Influx	0.89	0.77	0.43	0.60	0.81	0.87	
Outflux		0.83	0.66	0.69	0.86	0.79	
Degree			0.53	0.78	0.75	0.68	
QA				0.54	0.72	0.27	
Betweenness					0.69	0.45	
SEPM							
Source strength						0.67	
SEPM goodness							0.93

Note: All coefficients shown are significant at $P = 0.05$.

SEPM analysis

We used a variety of metrics from each initialization scheme to identify patches that were sources, stepping stones, or persistent patches (Table 3). From the random initialization version of the model, we computed the number of birds on each patch after 100 years (BP), how many of the last 10 years the patch was occupied (YO), and coefficient of variation (CV) of the patch population over the last 10 years of the model run. We also used an output metric from the single-patch initialization method to assess patch connectivity and ability to contribute immigrants to the landscape: after 100 model runs initiated from each patch, we totaled the number of breeders on the whole landscape (BL).

A patch’s source strength was determined by combining two metrics: BP, which indicates the size and quality of a patch, and BL, which is an index of how well the patch is connected to the rest of the landscape. We ranked each patch according to its value of BP and BL and then averaged the two ranks to determine a single value of source strength. To measure a patch’s persistence, we combined the metrics YO and CV by ranking each patch according to YO and CV and averaging the two ranks. The spatially explicit population model does not provide an easy way to measure a patch’s value as a stepping stone, so there is no metric representing this aspect of site selection.

Model comparison

For each model (graph and SEPM), patches are ranked according to their potential as a source, stepping stone, and persistent patch. A correlation is then computed for the rankings from each model. Often, source strength and persistence are of equal value when designing a reserve. Therefore, we also computed an overall “goodness” metric for each model by averaging influx and outflux ranks in the graph model and source strength and persistence ranks from the SEPM. A rank correlation between the measures of overall goodness from each model was then computed.

RESULTS

The graph metric outflux was well correlated with the SEPM metric for source strength (landscape 1, Spearman’s $\rho = 0.86$; landscape 2, $\rho = 0.87$; Tables 4 and 5). In landscape 1, 7 of the 10 highest ranked patches for each metric were the same; this increased to 8 out of 10 for landscape 2 (Fig. 3). The variability in the relationship between outflux and SEPM source strength is mostly due to the relationship between outflux and the number of breeders on the whole landscape (BL; Fig. 4). While patches with large outflux consistently contributed a large number of birds to the larger landscape (i.e., high BL), patches with lower outflux were extremely variable; some patches with fairly low outflux still managed to

TABLE 5. Spearman correlation coefficients between graph metrics and SEPM metrics for landscape 2.

Metric	Graph				SEPM		Graph goodness
	Outflux	Degree	QA	Betweenness	Source strength	Persistence	
Graph							
Influx	0.44	0.24	0.19	ns	0.52	0.73	
Outflux		0.54	0.86	0.35	0.87	0.82	
Degree			0.31	0.51	0.42	0.56	
QA				0.37	0.84	0.54	
Betweenness					0.31	0.23	
SEPM							
Source strength						0.76	
SEPM goodness							0.93

Note: All coefficients shown are significant at $P = 0.05$; “ns” indicates nonsignificant.

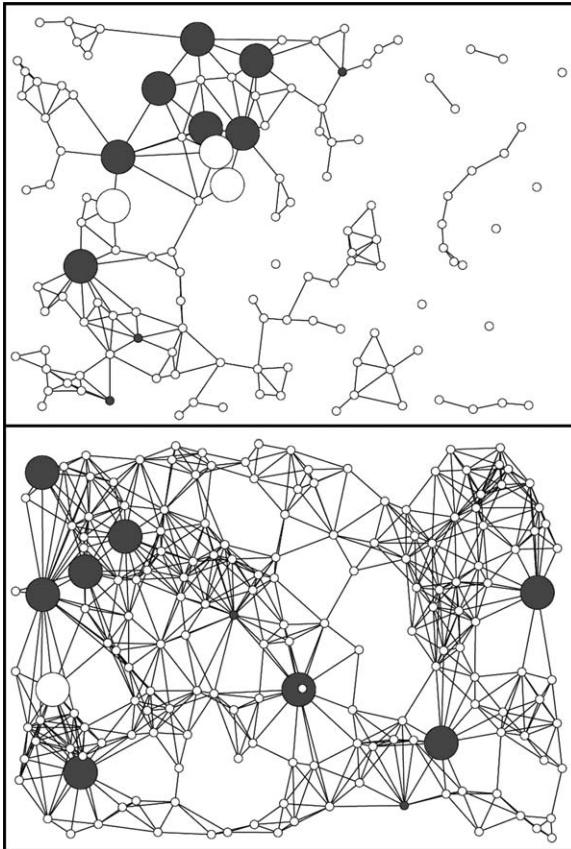


FIG. 3. A comparison of graph outflux and the SEPM metric "source strength." The large landscape is shown in the upper panel, and the small one in the lower panel. The top 10 patches for graph outflux are shown as dark circles, while the top 10 patches for source strength are indicated by the large circles. Patches that are both dark and large are in the top 10 for both models.

contribute large numbers of birds to the landscape while others contributed very few.

The graph metric influx was well correlated with the SEPM metric for persistence (landscape 1, Spearman's $\rho = 0.87$; landscape 2, $\rho = 0.73$; Tables 4 and 5). However, there was little overlap in the top 10 patches for either landscape. In landscape 1, two of the 10 highest ranked patches for each metric were the same; this increased to 4 out of 10 for landscape 2 (Fig. 5). The variability in the relationship between influx and SEPM persistence was mostly due to the relationship between influx and CV (Fig. 6). While patches with small influx consistently had a large amount of interannual variability in population size, patches with large influx were not as consistent. Some patches had large influx but high variability while others had low variability.

In general, QA and degree were fairly good proxies for source strength but not for persistence (Tables 4 and 5). The correlation between QA and the SEPM metric for source strength was 0.72 for landscape 1 and 0.84 for landscape 2. The correlation between degree and source

strength was 0.75 for landscape 1 and 0.42 landscape 2 (Spearman's ρ used for all correlation data). SEPM persistence had a weaker relationship with QA and degree. The correlation between QA and persistence was 0.54 for landscape 1 and 0.27 for landscape 2. The correlation between degree and persistence was 0.68 for landscape 1 and 0.56 for landscape 2.

Finally, the measures of overall goodness from each model were very highly correlated (landscape 1, Spearman's $\rho = 0.93$; landscape 2, $\rho = 0.93$; Tables 4 and 5). In landscape 1, 5 of the top 10 overall patches chosen for the SEPM were also selected by the graph model (Fig. 7); in landscape 2, this increased to 8 out of 10.

In landscape 1, all patches with high betweenness are in the large graph component (Fig. 8). In addition, many of the patches scoring highest for betweenness are cut-nodes (their removal would break the graph into smaller components). In landscape 2, which was more connected than the large landscape, the patches scoring highest for betweenness would not be as easy to identify visually. However, they tended to be patches that were centrally located in the landscape.

DISCUSSION

While it is easy to assume that the SEPM provides more biologically realistic output than the graph model because of its level of complexity, it should be recognized that with complexity comes potential for huge error in model output. Therefore, it is important to remember that without field data to test the models, there is no way of knowing which one is actually more accurate. However, the utility of SEPMs (despite the

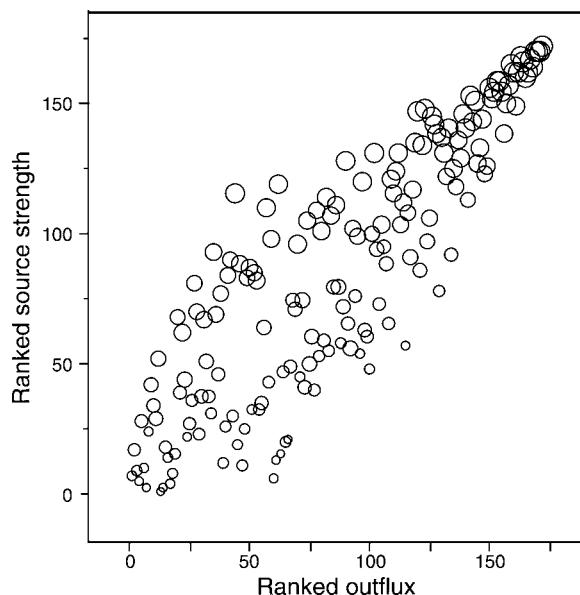


FIG. 4. Source strength (SEPM metric) vs. outflux (graph metric) for the small landscape. Note that different sizes of circles represent BL (the number of breeders on whole landscape), which is an index of how well the patch is connected to the rest of the landscape.

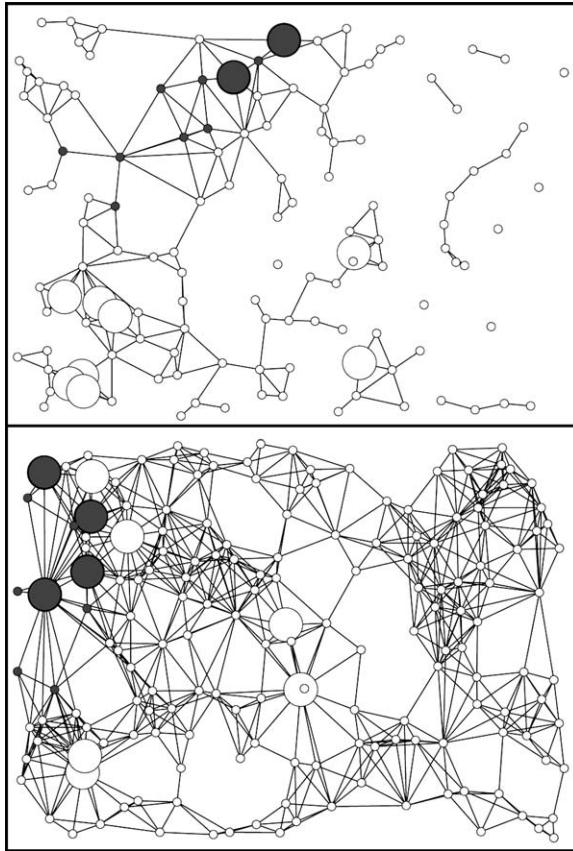


FIG. 5. A comparison of graph influx and the SEPM metric “persistence.” The large landscape is shown in the upper panel, and the small one in the lower panel. The top 10 patches for graph influx are shown as dark circles, while the top 10 patches for persistence are indicated by the large circles. Patches that are both dark and large are in the top 10 for both models.

potential for error) has generally been accepted and thus the goal of this analysis was not to test whether the graph model performs *as well* as the SEPM but whether it performs *similarly*.

In general, the graph model does perform similarly to the SEPM. Most of the graph metrics correspond well with the SEPM metrics, and there is good overlap between the important patches selected by each model. Further, graph theory provides a way of identifying stepping stone patches, which can be important additions to reserves but are not easily identified with a SEPM.

The relationship between SEPM source strength and graph outflux is reasonably strong, but outflux, by definition, is limited in its ability to capture emigration beyond immediate neighbors. In other words, patches with small outflux can still contribute large numbers of immigrants to the surrounding landscape in the SEPM (i.e., have a high number of breeders on the whole landscape [BL]). In particular, this might happen if a patch were mid-sized and had only one neighbor (small outflux), but that neighbor was a high-quality patch that

was connected to many other patches. In the single patch initialization version of the SEPM, only one immigrant from the original patch would need to reach its high-quality neighbor and the number of birds on the landscape could quickly increase. Whether or not that mid-sized patch with only one neighbor would be a good target for conservation is questionable, however. Instead, the high-quality neighbor might be a preferable target and is also more likely to be identified as an important patch by the graph model. It is possible that running the single-patch initialization version of the SEPM for 100 years overemphasizes the importance of neighboring patches’ characteristics and that reducing the length of the model run would provide a better measure of source strength as well as correspond better to graph outflux.

The relationship between SEPM persistence and graph influx is also fairly strong for landscape 1 but less so for landscape 2. Similar to outflux, influx is limited in its ability to measure movement beyond that from immediate neighbors. The SEPM measure of persistence, however, is calculated from two metrics that reflect more complicated dynamics than the number of birds in a patch after 100 years (BP) and the number of breeders on the whole landscape (BL): e.g., number of the last 10 years occupied (YO) and interannual variation in population size (CV). A plot of influx vs. SEPM persistence (Fig. 6) shows that influx is not very strongly associated with CV, while it is more strongly associated with YO (data not shown). Patches with large influx can either have large CV or small CV, implying that CV may incorporate dynamics at a larger distance

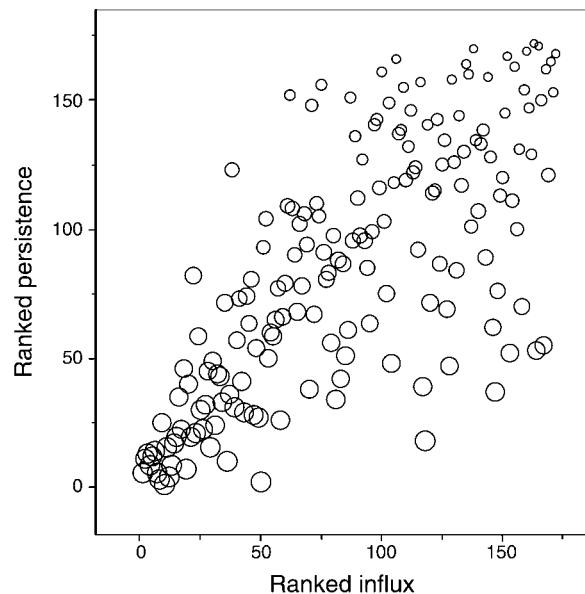


FIG. 6. Persistence (SEPM metric) vs. influx (graph metric) for the small landscape. Note that different sizes of circles represent CV, the coefficient of variation of the patch population over the last 10 years of the model run.

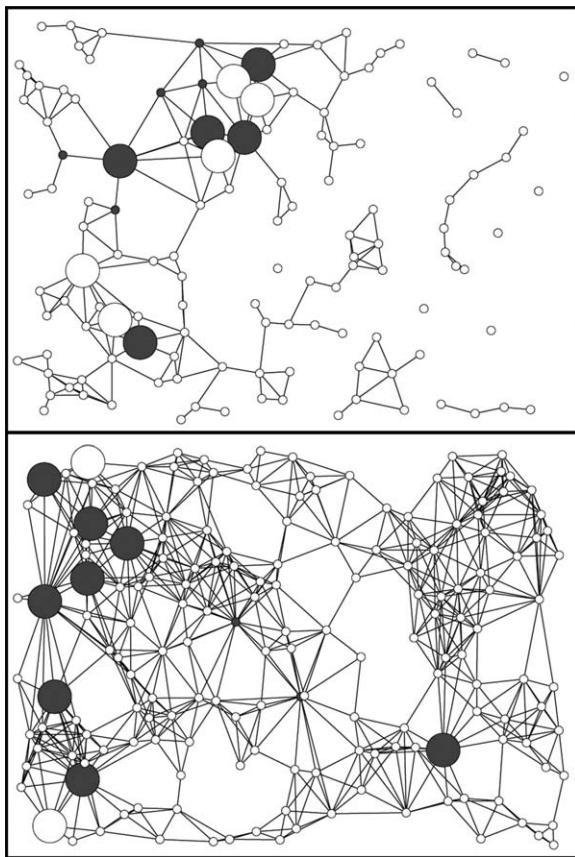


FIG. 7. A comparison of overall graph "goodness" and SEPM "goodness." The large landscape is shown in the upper panel, and the small one in the lower panel. The overall top 10 patches for the graph model are shown as dark circles, while the overall top 10 patches for the SEPM are indicated by the large circles. Patches that are both dark and large are in the top 10 for both models.

away from the focal patch. The map of persistence and influx on landscape 1 (Fig. 5, top panel) shows an interesting geographic trend: the top 10 patches for influx are all in the upper left corner of the map, while most of the best patches for persistence are in the lower left corner. Another potentially useful graph metric is the clustering coefficient (Newman 2001, Dunne et al. 2002), which indicates the degree to which a node's (or patch's) neighbors tend to be near each other. A clustering analysis in Pajek 1.12 (Batagelj and Mrvar 1996) revealed that patches in the lower left corner of the map are highly clustered, meaning that they share many neighbors with each other. In addition, landscape 2 is more clustered than landscape 1 and influx is less strongly related to persistence in landscape 2. Further analysis is required before reaching any conclusions, but high clustering may provide a patch with greater population stability (and therefore be a better proxy for persistence) than simply high influx.

Degree and quality-weighted area (QA) do not contain enough information to be good proxies for

SEPM persistence and source strength: degree reveals nothing about the size of the patch and QA reveals nothing about its neighbors. However, it is interesting that SEPM source strength is more strongly correlated with both QA and degree than SEPM persistence. This may reflect the complexity of the persistence metric and the landscape dynamics that are related to CV and YO.

Overall graph and SEPM "goodness" are very highly correlated. However, there is surprisingly little overlap among the top 10 patches selected by each model for landscape 1 (Fig. 7). Similar to influx, most of the important patches for overall graph goodness are in the top left portion of the map. These patches are among the largest and highest quality patches in landscape 1, and the graph metrics seem to be cuing in on this fact, while the SEPM output may be reflecting larger scale dynamics.

The betweenness metric identifies patches that are cut-nodes in landscape 1, but it is not obvious why it identifies some of the patches in landscape 2. Landscape 2 is much more highly connected than landscape 1 (mean degree for landscape 1 = 6.3, mean degree for landscape 2 = 17.6) and does not contain any cut-nodes. When there

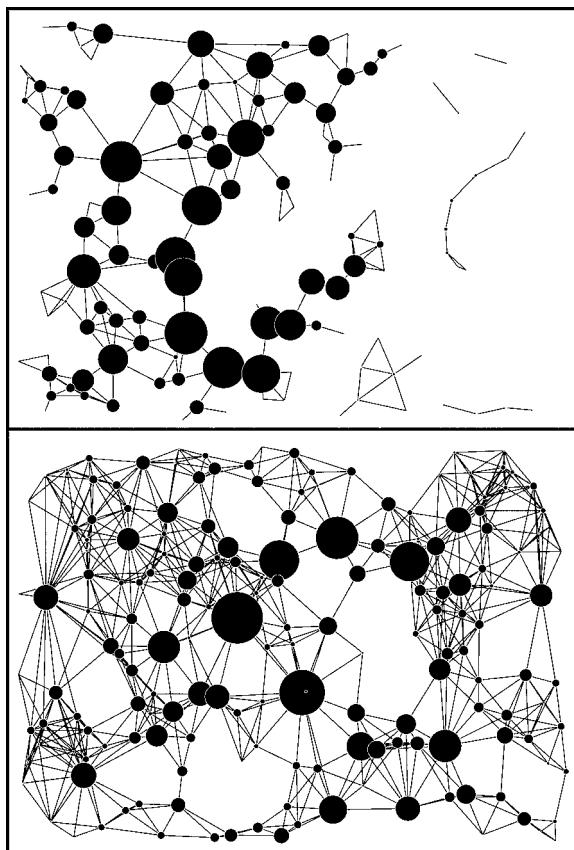


FIG. 8. Betweenness for the large landscape (top) and small landscape (bottom). The size of the circles indicates value of betweenness; large circles are important for network connectivity.

are multiple alternate paths across the landscape, as in landscape 2, connectivity is not usually a big concern and betweenness may not be a very useful metric.

The correlation between the graph model and the SEPM remains strong across a range of parameter values. While a thorough sensitivity analysis is beyond the scope of this paper, preliminary analyses (data not shown) indicate that the graph model is able to replicate the SEPM results for dispersal distances ranging from 500 m to 3000 m, as well as for birds with increased clutch size. However, when clutch size is reduced so that population levels continuously decline in the SEPM, the graph model output no longer corresponds to the SEPM output. This suggests that graph theory may be most suitable when a species' distribution is limited by the landscape rather than biological parameters.

Summary and future applications

This analysis has shown that graph models can closely reproduce many of the results from a SEPM. In particular, the graph metric outflux does a good job of ranking and identifying patches according to their source strength, and overall graph goodness is highly correlated with overall SEPM goodness. The ability of the graph metric influx to identify patches with high persistence seems to vary according to landscape characteristics such as connectivity. Preliminary analyses suggest that patch clustering may increase patch persistence without affecting influx, thus limiting the correlation between influx and persistence. In general, it seems that the SEPM is capable of capturing more complicated and larger scale dynamics than the graph metrics used here. However, it is important to remember that the SEPM is not necessarily the gold standard by which any other model should be measured, as the large number of parameters has the potential to cause large errors in model output. Discrepancies between the graph model and the SEPM output should not be automatically considered flaws in the graph model but rather events worthy of further examination. Additionally, there is a suite of additional graph metrics (not discussed here) that may be equally or more useful for selecting habitat patches for conservation. Therefore, we believe that the benefits and ease of a graph theory analysis for conservation planning outweigh any potential shortcomings.

Graph theory is an emerging conservation tool that can be applied to a variety of taxa. For example, Bunn et al. (2000) compared landscape graphs for two species with similar habitat requirements but different dispersal abilities. They showed that the Coastal Plain of North Carolina is functionally connected for mink but disconnected for the Prothonotary Warbler. More recently, graph theory has been applied to marine organisms as well. Treml et al. (*in press*) created marine graphs for corals in the Tropical Pacific by estimating connectivity between reefs based on ocean currents. Patterns in these graphs revealed connected populations

and critical stepping stones, and they suggested areas that might be prioritized for marine conservation efforts. The potential for graph theory to be applied to other organisms and landscapes is unlimited.

Graph theory in general, and the program Pajek in particular, can also be used to analyze landscapes for multiple species simultaneously, even for species with different dispersal abilities. While this is beyond the scope of this paper, it is possible to create separate networks for each species and to combine them in a variety of ways, depending on the question of interest. Another use of graph theory in conservation planning is the evaluation of networks as a whole, rather than the focus on individual patches seen in this paper. Integrative network properties, such as rates of movement and vulnerability to disturbance, can be measured and compared using techniques traditionally applied to social sciences and other disciplines.

ACKNOWLEDGMENTS

We thank the members of the Duke University Nicholas School graph group for discussion that led to improvements to this manuscript. We also thank the anonymous reviewers for their comments on this manuscript.

LITERATURE CITED

- Åberg, J., G. Jansson, J. E. Swenson, and P. Angelstam. 1995. The effect of matrix on the occurrence of Hazel Grouse (*Bonasa bonasia*) in isolated habitat fragments. *Oecologia* 103:265–269.
- Anders, A. D., D. C. Dearborn, J. Faaborg, and F. R. Thompson. 1997. Juvenile survival in a population of neotropical migrant birds. *Conservation Biology* 11:698–707.
- Anders, A. D., J. Faaborg, and F. R. Thompson. 1998. Postfledging dispersal, habitat use, and home-range size of juvenile Wood Thrushes. *Auk* 115:349–358.
- Batagelj, V., and A. Mrvar. 1996. Pajek 1.12. (<http://vlado.fmf.uni-lj.si/pub/networks/pajek/>)
- Brawn, J. D., and S. K. Robinson. 1996. Source–sink population dynamics may complicate the interpretation of long-term census data. *Ecology* 77:3–12.
- Brown, J. H., and A. Kodric-Brown. 1977. Turnover rates in insular biogeography: effect of immigration on extinction. *Ecology* 58:445–449.
- Bunn, A. G., D. L. Urban, and T. H. Keitt. 2000. Landscape connectivity: a conservation application of graph theory. *Journal of Environmental Management* 59:265–278.
- Burke, D. M., and E. Nol. 2000. Landscape and fragment size effects on reproductive success of forest-breeding birds in Ontario. *Ecological Applications* 10:1749–1761.
- Chalfoun, A. D., M. J. Ratnaswamy, and F. R. Thompson. 2002. Songbird nest predators in forest–pasture edge and forest interior in a fragmented landscape. *Ecological Applications* 12:858–867.
- Conroy, M. J., Y. Cohen, F. C. James, Y. G. Matsinos, and B. A. Maurer. 1995. Parameter-estimation, reliability, and model improvement for spatially explicit models of animal populations. *Ecological Applications* 5:17–19.
- Donovan, T. M., F. R. Thompson, J. Faaborg, and J. R. Probst. 1995. Reproductive success of migratory birds in habitat sources and sinks. *Conservation Biology* 9:1380–1395.
- Dunne, J. A., R. J. Williams, and N. D. Martinez. 2002. Food-web structure and network theory: the role of connectance and size. *Proceedings of the National Academy of Sciences (USA)* 99:12917–12922.

- Dunning, J. B., R. Borgella, K. Clements, and G. K. Meffe. 1995a. Patch isolation, corridor effects, and colonization by a resident sparrow in a managed pine woodland. *Conservation Biology* 9:542–550.
- Dunning, J. B., D. J. Stewart, B. J. Danielson, B. R. Noon, T. L. Root, R. H. Lamberson, and E. E. Stevens. 1995b. Spatially explicit population-models—current forms and future uses. *Ecological Applications* 5:3–11.
- Evans, D. R., and J. E. Gates. 1997. Cowbird selection of breeding areas: the role of habitat and bird species abundance. *Wilson Bulletin* 109:470–480.
- Fauth, P. T. 2000. Reproductive success of Wood Thrushes in forest fragments in northern Indiana. *Auk* 117:194–204.
- Fauth, P. T. 2001. Wood Thrush populations are not all sinks in the agricultural midwestern United States. *Conservation Biology* 15:523–527.
- Friesen, L., M. D. Cadman, and R. J. MacKay. 1999. Nesting success of neotropical migrant songbirds in a highly fragmented landscape. *Conservation Biology* 13:338–346.
- Harary, F. 1969. Graph theory. Addison-Wesley, Boston, Massachusetts, USA.
- Hess, G. R., F. H. Koch, M. J. Rubino, K. A. Eschelbach, C. A. Drew, and J. M. Favreau. 2006. Comparing the potential effectiveness of conservation planning approaches in central North Carolina, USA. *Biological Conservation* 128:358–368.
- Hoover, J. P., M. C. Brittingham, and L. J. Goodrich. 1995. Effects of forest patch size on nesting success of Wood Thrushes. *Auk* 112:146–155.
- Keitt, T. H., D. L. Urban, and B. T. Milne. 1997. Detecting critical scales in fragmented landscapes. *Conservation Ecology* [online] 1:4.
- Kristan, W. B. 2003. The role of habitat selection behavior in population dynamics: source–sink systems and ecological traps. *Oikos* 103:457–468.
- Matthysen, E. 1999. Nuthatches (*Sitta europaea*: Aves) in forest fragments: demography of a patchy population. *Oecologia* 119:501–509.
- Melian, C. J., and J. Bascompte. 2002. Complex networks: two ways to be robust? *Ecology Letters* 5:705–708.
- Moilanen, A., and M. Cabeza. 2002. Single-species dynamic site selection. *Ecological Applications* 12:913–926.
- Mörtberg, U. M. 2001. Resident bird species in urban forest remnants; landscape and habitat perspectives. *Landscape Ecology* 16:193–203.
- Newman, M. E. J. 2001. Scientific collaboration networks. I. Network construction and fundamental results. *Physical Review E* 6401.
- Noss, R. F., C. Carroll, K. Vance-Borland, and G. Wuerthner. 2002. A multicriteria assessment of the irreplaceability and vulnerability of sites in the Greater Yellowstone Ecosystem. *Conservation Biology* 16:895–908.
- Ortega, Y. K., and D. E. Capen. 1999. Effects of forest roads on habitat quality for ovenbirds in a forested landscape. *Auk* 116:937–946.
- Paton, P. W. C. 1994. The effect of edge on avian nest success—how strong is the evidence? *Conservation Biology* 8:17–26.
- Powell, L. A., J. D. Lang, M. J. Conroy, and D. G. Krentz. 2000. Effects of forest management on density, survival, and population growth of wood thrushes. *Journal of Wildlife Management* 64:11–23.
- Pulliam, H. R. 1988. Sources, sinks, and population regulation. *American Naturalist* 132:652–661.
- Roth, R. R., M. S. Johnson, and T. J. Underwood. 1996. Wood Thrush. In A. Poole and F. Gill, editors. *The birds of North America*, No. 246. The Academy of Natural Sciences, Philadelphia, Pennsylvania, USA.
- Rothley, K. D., C. N. Berger, C. Gonzalez, E. M. Webster, and D. I. Rubenstein. 2004. Combining strategies to select reserves in fragmented landscapes. *Conservation Biology* 18:1121–1131.
- Sauer, J. R., J. E. Hines, and J. Fallon. 2002. The North American breeding bird survey, results and analysis 1966–2001. Version 2002.1. USGS Patuxent Wildlife Research Center, Laurel, Maryland, USA.
- South, A. 1999. Dispersal in spatially explicit population models. *Conservation Biology* 13:1039–1046.
- Treml, E. A., P. N. Halpin, D. L. Urban, and L. F. Pratson. *In press*. Modeling population connectivity by ocean currents a graph theoretic approach for marine conservation. *Landscape Ecology*.
- Trine, C. L. 1998. Wood Thrush population sinks and implications for the scale of regional conservation strategies. *Conservation Biology* 12:576–585.
- Turner, M. G., G. J. Arthaud, R. T. Engstrom, S. J. Hejl, J. G. Liu, S. Loeb, and K. McKelvey. 1995. Usefulness of spatially explicit population models in land management. *Ecological Applications* 5:12–16.
- Turner, W. R., and D. S. Wilcove. 2006. Adaptive decision rules for the acquisition of nature reserves. *Conservation Biology* 20:527–537.
- Twomey, A. C. 1945. The bird populations of an elm–maple forest with special reference to aspektion, territorialism, and coactions. *Ecological Monographs* 15:173–205.
- Urban, D., and T. Keitt. 2001. Landscape connectivity: a graph-theoretic perspective. *Ecology* 82:1205–1218.
- Urban, D. L., and H. H. Shugart. 1986. Avian demography in mosaic landscapes: modeling paradigm and preliminary results. Pages 273–279 in J. Verner, M. L. Morrison, and C. J. Ralph, editors. *Wildlife 2000: modeling habitat relationships of terrestrial vertebrates*. University of Wisconsin Press, Madison, Wisconsin, USA.
- Vega Rivera, J. H., J. H. Rappole, W. J. McShea, and C. A. Haas. 1998. Wood Thrush postfledging movements and habitat use in northern Virginia. *Condor* 100:69–78.
- Weaver, F. G. 1949. Wood Thrush. Pages 101–123 in A. Bent, editor. *Life histories of North American thrushes, kinglets, and their allies*. U.S. National Museum Bulletin 196.
- Weinberg, H. J., and R. R. Roth. 1998. Forest area and habitat quality for nesting Wood Thrushes. *Auk* 115:879–889.