

## PREDICTION OF LONG-DISTANCE DISPERSAL USING GRAVITY MODELS: ZEBRA MUSSEL INVASION OF INLAND LAKES

JONATHAN M. BOSSENBROEK,<sup>1,3</sup> CLIFFORD E. KRAFT,<sup>2</sup> AND JEFFREY C. NEKOLA<sup>1</sup>

<sup>1</sup>Department of Natural and Applied Sciences, University of Wisconsin, Green Bay, Wisconsin 54311 USA

<sup>2</sup>Department of Natural Resources, Cornell University, Ithaca, New York 14853-3001 USA

**Abstract.** Gravity models are commonly used by geographers to predict migration and interaction between populations and regions. Even though rarely used by ecologists, gravity models allow estimation of long-distance dispersal between discrete points in heterogeneous landscapes. We developed a production-constrained gravity model to forecast zebra mussel (*Dreissena polymorpha*) dispersal into inland lakes of Illinois, Indiana, Michigan, and Wisconsin (USA) based on the site and location of lakes and the number and location of boats within 364 counties. A deterministic form of this model was used to estimate best-fit parameters for distance coefficient, Great Lakes boat-ramp attractiveness, and colonization cutoff threshold. A stochastic model thus developed from these parameters allows for random changes in colonization likelihood. The results of our model are highly correlated with the actual pattern of colonized lakes in southern Michigan and southeastern Wisconsin at the end of 1997. Areas of central Wisconsin and western Michigan, where zebra mussel colonies have not been documented, were also predicted to be colonized, suggesting that future invasions may be imminent in these locations. These analyses suggest that gravity models may be useful in predicting long-distance dispersal when dispersal abilities of species and the attractiveness of potential habitats are known.

**Key words:** colonization; exotic species invasion; gravity model; gravity models vs. diffusion models; invasion of aquatic habitats; lakes, upper Midwest (USA); landscape ecology; long-distance dispersal, modeling; migration; modeling dispersal patterns; spatial interaction; zebra mussels.

### INTRODUCTION

Understanding long-distance dispersal is essential to predicting the spatial and temporal patterns of colonization within heterogeneous landscapes. Patterns of species colonization have primarily been predicted through diffusion models (Shigesada and Kawasaki 1997). While early diffusion models assumed a homogeneous landscape with dispersal resulting from short-range random movements (Skellam 1951), more recently, stratified diffusion models have been implemented to also incorporate long-distance dispersal events (Hengeveld 1989, Shigesada and Kawasaki 1997) that permit organisms to “jump” over or across habitats (Lewis 1997).

Even with these advancements, diffusion models have difficulty in predicting long-range dispersal events and, because of this, organism movement as well. For example, Andow et al. (1990) suggested that macroscale processes, such as air currents or human transport, govern the dispersal of the cereal leaf beetle (*Oulema melanopus*). Surveys of rare, long-distance boater movements provided a better indicator of zebra mussel (*Dreissena polymorpha*) dispersal than diffu-

sion models (Buchan and Padilla 1999). Thus, in spite of their rarity, long-distance dispersal events appear to drive migration patterns for many species (Dean 1998).

Gravity models allow for the prediction of long-distance dispersal events by considering not only the nature of source populations, but also the spatial configuration and nature of potential colonization sites. Because of this, gravity models have the potential to more accurately forecast species movement through heterogeneous landscapes than do diffusion models, which do not explicitly consider the spatial pattern of distant sites. Geographers have used gravity models to predict human dispersal patterns by estimating the flow of people per unit time based on the distance to and attractiveness of destination points (Thomas and Hugget 1980, Sklar and Costanza 1991). Schneider et al. (1998) used a gravity model to assess the relative risk of zebra mussel invasions to the inland lakes and reservoirs of Illinois (USA).

The colonization of inland lakes in the Upper Midwest (USA) by zebra mussels presents an ideal system for testing the use of gravity models for prediction of long-distance dispersal. The spread of zebra mussels across the North American landscape has been closely monitored since their initial North American discovery in 1988 (Hebert et al. 1989). Range expansion quickly occurred throughout commercially navigable waters (Griffiths et al. 1991), but overland dispersal into inland lakes has been slower (Kraft and Johnson 2000). The

Manuscript received 25 February 2000; revised 15 September 2000; accepted 20 October 2000; final version received 13 November 2000.

<sup>3</sup> Present address: Graduate Degree Program in Ecology, Colorado State University, Fort Collins, Colorado 80523 USA. E-mail: bossjm@lamar.colosate.edu

first U.S. inland lake colonization occurred in 1991 in northeast Indiana, and by December 1997 only 56 inland lakes were colonized in Michigan (37 lakes), Indiana (12 lakes), Wisconsin (6 lakes), and Illinois (1 lake) (Kraft and Johnson 2000).

Although many other potential mechanisms exist (Carlton 1993), the overland transport of recreational boats is widely believed to be the primary vector for zebra mussel dispersal into inland lakes (Carlton 1993, Johnson and Carlton 1996, Johnson and Padilla 1996, Schneider et al. 1998, Buchan and Padilla 1999). Recent attempts have been made to compare patterns of boater activities to inland lake invasions of zebra mussels in Wisconsin lakes. Based upon a small number of known lake invasions, the rate of Wisconsin inland-lake zebra mussel invasion appeared to be related to the frequency of recreational boating (Padilla et al. 1996). Additionally, surveys of long-distance boat movements provide better forecasts of zebra mussel dispersal than do diffusion models (Buchan and Padilla 1999). Boater movements have also been incorporated in a risk assessment of the potential zebra mussel invasion of Illinois inland lakes (Schneider et al. 1998). Assessment of these predictions is not currently possible, as only a single Illinois lake has been colonized.

In this study we describe the implementation of a production-constrained gravity model to forecast the overland dispersal of zebra mussels into inland lakes within a four-state region (Illinois, Indiana, Michigan, and Wisconsin). We then describe the use of this model to address: (1) the role of recreational boating in the spread of zebra mussels; (2) regions most prone to future colonization events; and (3) the utility of gravity models for predicting long-distance dispersal.

## METHODS

### *General characteristics of gravity models*

Gravity models, in general, develop a matrix that calculates the flow of individuals that move from a series of origins to a series of destinations based on the distance and attractiveness of the destinations (Thomas and Hugget 1980). For example, variables such as population size, unemployment rate, or crime rate can be used to rate the attractiveness of a given city. Gravity models are functionally different from diffusion models in that diffusion models estimate movement rates by an organism, whereas gravity models estimate the force of attraction between an origin and a destination, with movement rates being a function of this force. Thus, a diffusion model is more appropriate if a movement rate can be estimated, whereas a gravity model is more appropriate when distance to and attractiveness of destinations are known or are of interest.

Different types of gravity models exist based upon prior information; a production-constrained gravity model is used when information about the population

of the site of origin is known, but not the number of people who travel to a particular destination, while a production-attraction-constrained gravity model is used when information is known about both the source and destination populations. Schneider et al. (1998) were able to use a production-attraction-constrained gravity model to estimate zebra mussel dispersal in Illinois because estimates were available for the number of boaters at both origins and destinations for movement.

We have developed a production-constrained gravity model to simulate zebra mussel dispersal over a larger region for which data regarding the number of registered boats per county were available, but data were not generally available regarding the number of boats traveling to given lakes.

### *Deterministic model*

The colonization of an inland lake by zebra mussels is the result of a three-step process. First, boats travel to a colonized lake or boat ramp and pick up juvenile or adult zebra mussels. Second, these infested boats travel to an uncolonized lake on a subsequent outing, inadvertently releasing zebra mussels into this water body. Third, these transported individuals recruit a new colony based upon the physical nature of the lake (water chemistry, depth) and stochastic demographic events (Johnson and Padilla 1996). As a result, our model will estimate the potential for colonization based upon three factors: (1) the probability of a boat traveling to a zebra mussel source, (2) the probability of the same boat making a subsequent outing to an uncolonized lake, and (3) the probability of zebra mussels becoming established once released in that uncolonized lake.

The first step of the model calculates the number of boats from each county that travel to a zebra mussel source and thus have the potential to transport zebra mussels to an uncolonized lake on a subsequent outing. The number of boats,  $T$ , that travel from county  $i$  to a lake or boat ramp,  $j$ , is estimated as

$$T_{ij} = A_i O_i W_j c_{ij}^{-\alpha} \quad (1)$$

where,  $A_i$  is a scalar,  $O_i$  is the number of boats in county  $i$ ,  $W_j$  is the attractiveness of location  $j$ ,  $c_{ij}$  is the distance from county  $i$  to location  $j$ , and  $\alpha$  is the distance coefficient.  $A_i$  ensures that all the boats from county  $i$  reach some lake. Without  $A_i$  a production-constrained gravity model calculates the proportion of boats that move from county  $i$  to each lake. Such scalars are referred to as "balancing factors" in the spatial interaction literature (Fotheringham and O'Kelly 1989).  $A_i$  is estimated via

$$A_i = 1 / \sum_{j=1}^N W_j c_{ij}^{-\alpha} \quad (2)$$

where  $N$  represents the total number of lakes and boat ramps and  $j$  represents each lake in the study region.

The number of potentially infested boats for each county is expressed as  $P_i$  and calculated by

$$P_i = \sum_{s=1}^n T_{is} \quad (3)$$

where  $T_{is}$  is the subset of  $T_{ij}$  which consists of those boaters who travel from county  $i$  to a source of zebra mussels,  $s$ .  $T_{is}$  is summed for each county over the total number of zebra mussel sources,  $n$ .

In the second step of our model, the "infested" boats,  $P_i$ , make a second excursion during which they transport zebra mussels to other lakes.  $T_{iu}$  represents the number of these boats that travel from county  $i$  to an uncolonized lake  $u$ :

$$T_{iu} = A_i P_i W_u c_{iu}^{-\alpha} \quad (4)$$

The total number of infested boats that arrive at a given uncolonized lake,  $Q_u$ , is calculated by summing over all counties,  $M$ :

$$Q_u = \sum_{i=1}^M T_{iu} \quad (5)$$

Thus,  $Q_u$  is the relative number of infested boats that visit lake  $u$  in one iteration (year) of the model.

The third step of the model determines whether overland-dispersal events lead to the establishment of zebra mussel colonies. Establishment of new colonies is based on two factors: (1) physical characteristics of the lake and (2) stochastic demographic events.

The single transfer of a few zebra mussel adults or juveniles to an environmentally appropriate lake will not guarantee development of a new colony. Because of environmental and demographic factors, multiple zebra mussel deliveries by infested boats are likely necessary before successful colonization occurs. This statement is supported by the observation that the delivery estimates of adult zebra mussels to inland waters exceed the number of invasions (Johnson et al. 2001). The relative number of infested boats required to guarantee establishment of a new colony,  $f$ , was determined through a best-fit parameterization of the data (see *Best-fit parameterization*, below). In subsequent trials, lakes with values of  $Q_u$  (number of infested boats visiting a lake per year) greater than this colonization threshold,  $f$ , were designated as "colonized" and became new zebra mussel sources for subsequent model iterations, while lakes with values of  $Q_u$  below this limit remained uncolonized. To generate a deterministic distribution of zebra mussel-colonized lakes, the deterministic model was run for seven iterations (years) using the best-fit parameters (see *Best-fit parameterization*, below).

#### Stochastic model

Given the stochastic nature of the transport and deposition of zebra mussels, the establishment of new

colonies will not be static. Analysis of the deterministic model showed that for many lakes, the value of  $Q_u$  was just slightly above or below the colonization threshold ( $f$ ). For the deterministic model, lakes with values of  $Q_u$  slightly below the colonization threshold ( $f$ ) were never designated as colonized; consequently a stochastic process was incorporated into the model. Even though multiple deliveries of zebra mussels are most likely necessary for colonization to occur, theoretically it is possible for a single boat to cause a lake to become colonized. Therefore, we estimated the relative probability that a single boat would cause a lake to become colonized,  $p_f$ . For this stochastic model, the probability of an uncolonized lake becoming colonized,  $x_u$ , was estimated as

$$x_u = Q_u p_f \quad (6)$$

For each model iteration the probability of colonization of a given lake is based on the number of infested boats that arrive at each lake. At the end of each model iteration, each lake is assigned a probability of colonization,  $x_u$ , and then subjected to a Bernoulli trial, by which each lake is either designated as colonized (a score of 1 from the Bernoulli trial) or remains uncolonized, based on the probability  $x_u$ . Newly colonized lakes then become sources for the subsequent iterations (years).

To estimate  $p_f$ , probabilities ranging from 0.0000118 to 0.000235 (equivalent to a 1 to 20% chance of colonization when 850 boaters arrive at a lake) were incorporated into the stochastic model. Each probability was used in 100 trials of the model to determine which probability resulted in an average of 47 colonized lakes after seven iterations of the model. The selected probability,  $p_f$  was then used in the 2000 trials of the stochastic model.

To generate a probabilistic distribution of zebra mussel-colonized lakes, 2000 trials of the stochastic model were conducted over a simulated period of seven years (seven iterations). From these, the probability of colonization for each lake was determined by dividing the number of times each lake was predicted to become colonized by the number of trials (2000). The number of colonized lakes for each county was determined by summing the individual colonization probabilities of each lake within that county.

#### Data sets

The region over which the model was tested includes all of Michigan and Wisconsin and those Illinois, Indiana, and Ohio counties within 300 km of the western Great Lakes shoreline (Fig. 1). Empirical data incorporated in the model consisted of the number of registered recreational boats per county,  $O_i$ ; the location and area of lakes,  $W_j$ ; the location of public-access boat ramps along the Great Lakes, Mississippi River, and Illinois River, the observed pattern of zebra mussel-colonized lakes; and limnological data for inland lakes.

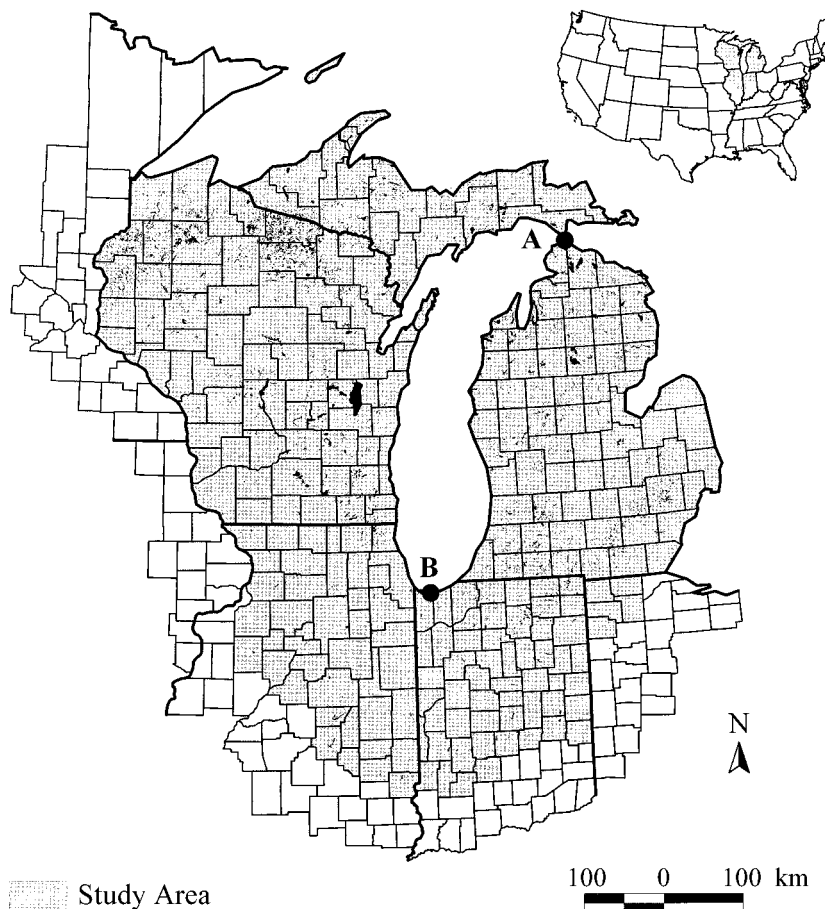


FIG. 1. Region of study (Michigan and Wisconsin, and portions of Illinois, Indiana, and Ohio [USA]). The shaded counties contain lakes included in the gravity model. All the registered boaters in the shaded counties were considered in the model, while only 10% of the registered boaters in the non-shaded counties were incorporated. The black polygons are the lakes >25 ha in surface area within the study region. "A" represents the point through which all travels between Area 3 and Area 2 were routed; "B" represents the point through which all travels between Area 3 and Area 1 were routed (see Table 1 caption).

Recreational-boat registration records were obtained for 368 counties: all counties within the study region, as well as Minnesota, Iowa, and Ohio counties within 50 km of the region's boundary (Bossenbroek 1999). Only 10% of the boats beyond the primary study region were included, as 90% of boater movements are less than 50 km (Buchan and Padilla 1999). As such, we assumed that 90% of boater movements for these counties either took place in that county, or outside of our study region.

The surface area and location of inland lakes were based on data sets acquired from the Wisconsin Department of Natural Resources (Wisconsin DNR, Bureau of Fisheries Management and Habitat Protection, unpublished data), Michigan State University (MSU, Center for Remote Sensing and Geographic Information Science, unpublished data), and EPA River Reach files (U.S. Environmental Protection Agency, Office of Water 1999), while the locations of water bodies too large to be considered single entities (the Great Lakes,

Mississippi River, and Illinois River) were designated by the location of 399 public access boat ramps. In order to limit the number of potential-destination lakes to a tractable amount, we only included those lakes that have a surface area >25 ha (3600 lakes). The largest lake in the region was Lake Winnebago in Wisconsin (53 504 ha).

The attractiveness of a given lake,  $W_j$ , is assumed to be positively correlated with lake area, with large lakes being more attractive than small ones (Reed-Anderson et al. 2000). As such, the value of  $W_j$  for a given inland lake is equal to that lake's surface area in hectares. For rivers, the surface area was estimated by multiplying the length of a river within the study region by the average width and then dividing by the number of boat ramps on each river. (The estimated  $W_j$  values for the Illinois and Mississippi River ramps are 269 ha and 6835 ha, respectively). The value of  $W_j$  for Great Lakes boat ramps was determined through best-fit parameterization (see Best-fit parameterization, below).

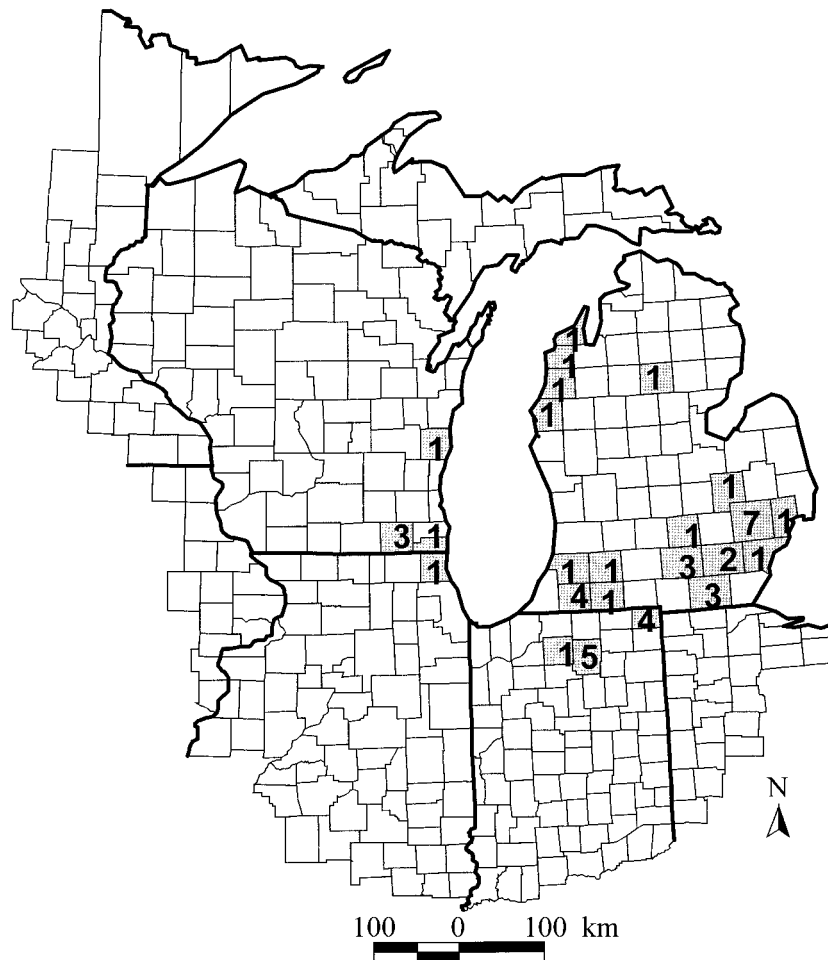


FIG. 2. The observed distribution of zebra-mussel-colonized lakes in the Western Great Lakes States as of December 1997. This distribution only includes lakes thought to have been colonized as a result of overland transport. The numbers indicate the number of colonized lakes within that county. Data are from Kraft and Johnson (2000).

The distribution of zebra mussel-colonized lakes within the study region was based on known colonization patterns as of December 1997 (Fig. 2; Kraft and Johnson 2000). Only the 47 lakes considered to be colonized via overland dispersal were used, eliminating the 9 lakes that are thought to have been secondarily colonized by such means as stream flow or water pumped from a colonized lake into an uncolonized lake (L. E. Johnson, *personal communication*). The Great Lakes, Mississippi River, and Illinois River boat ramps were considered the initial zebra mussel sources. Lakes colonized during an iteration of the model were considered sources during subsequent iterations.

The appropriateness of a given lake's physical characteristics for establishment of zebra mussel colonies was based on Ramcharan et al. (1992), in which lakes with low pH and calcium levels were not colonized. They estimated the suitability of a given lake through the equation:

$$B = 1.246 \text{ pH} + 0.045[\text{Ca}] - 11.696. \quad (7)$$

Lakes with  $B$  values exceeding  $-0.638$  were considered suitable for zebra mussel colonization. Based on this model, those lakes in our study with  $B$  values below this threshold were designated as unsuitable for colonization. Additionally, lakes with a maximum depth  $< 4$  m were considered uninhabitable (Strayer 1991), and were excluded from analysis.

To estimate  $B$ , pH and Ca levels were obtained for 50% and 30%, respectively, of the lakes within the study region via the EPA STORET data bank (U.S. Environmental Protection Agency, Office of Water 1999). For the study region, lake pH and calcium levels are closely tied to bedrock formation, and general regional patterns among lakes are apparent (Omernik and Powers 1983, Omernik et al. 1988). Because of this spatial autocorrelation, pH and calcium values for lakes not included in this database were estimated through punctual kriging (Burgess and Webster 1980) of the nearest 20 recorded lakes, using the GS+ software package (Gamma Design Software 1992). Based on this

TABLE 1. Measures for calculating distance between each county and each lake within the study region.

Distance to be calculated		Distance calculation
From area:	To area:	
1	2,4	County <i>i</i> to Lake <i>j</i>
1	3	County <i>i</i> to Pt. B + Pt. B to Lake <i>j</i>
2	1,4	County <i>i</i> to Lake <i>j</i>
2	3	County <i>i</i> to Pt. A + Pt. A to Lake <i>j</i>
3	1	County <i>i</i> to Pt. B + Pt. B to Lake <i>j</i>
3	2	County <i>i</i> to Pt. A + Pt. A to Lake <i>j</i>
3	4	County <i>i</i> to Lake <i>j</i>
4	1,2,3	County <i>i</i> to Lake <i>j</i>

Notes: The study region was divided into four areas. Area 1 = the portion of Wisconsin south of the southern tip of Green Bay and northwest Illinois; Area 2 = northern Wisconsin and the upper peninsula of Michigan; Area 3 = the lower peninsula of Michigan, northeast Indiana, and northwest Ohio; and Area 4 = the portion of Illinois not in Area 2 and the portion of Indiana not in Area 3. Point A is located at the Straits of Mackinac (northern end of Lake Michigan), and point B is located at the south end of Lake Michigan.

analysis, 1843 lakes in the region were considered to be environmentally inappropriate for zebra mussels. Most of these lakes were concentrated in northern Wisconsin, where igneous bedrock and alluvium predominates. Although these lakes were excluded for establishment of zebra mussel populations, they were still considered as destinations for boater movements.

Distances from each lake and boat ramp to each county were calculated from lake centroids to the center of lake polygons within each county. The center of lake polygons within a county was used instead of the county centroid because it was assumed that recreational boat owners are more likely to live close to lakes. For those counties with no lakes, the county centroid was used.

Although simple Euclidean distance estimates were used in most cases, these estimates were not appropriate when the lake/boat ramp and county in question were located on opposite sides of Lake Michigan, which is a barrier to straight-line travel. To address this problem, the study region was divided into four areas. For boats traveling from a county to a lake within the same area, or between areas on the same side of Lake Michigan, Euclidean distance estimates were used. For boats traveling between areas on opposite sides of Lake Michigan, distances were calculated by routing trips around Lake Michigan. The paths used to calculate these trips are shown in Table 1.

#### Best-fit parameterization

Three model parameters were estimated using least-sum-of-squares (LSS) parameterization: distance coefficient ( $\alpha$ ); colonization threshold ( $f$ ); and attractiveness of Great Lakes boat ramps ( $W$ ). Best-fit values were determined through LSS comparison of model predictions with the observed distribution of zebra

mussels at the county level. Comparisons were made at this scale because the county is the minimum resolution of boat registration data. In the parameterization routine,  $\alpha$  was varied from 1.0 to 7.0 in increments of 0.1,  $f$  ranged from 400 to 1500 in increments of 50, and  $W$  ranged from 5000 to 70 000 ha in increments of 5000 ha. Parameter estimation was conducted deterministically without inclusion of stochastic variables.

#### Sensitivity analysis and model robustness

To analyze the sensitivity of parameter estimation with respect to the number of lakes predicted to be colonized, the deterministic model was run while varying the values of the three estimated parameters from  $-20\%$  to  $+20\%$  of their best-fit estimate across seven model iterations. Since 47 lakes were included in the fitted data set, and the best-fit parameterization predicted 44 colonized lakes, we defined the least sensitive region to parameter adjustments as that range of predicted colonized lakes from 44 to 47 lakes.

Two additional tests were used to evaluate model robustness. First, the distance of the predicted colonized lakes to Great Lakes boat ramps and to county centroids was compared with observed distributions to assess the ability of the model to simulate regional and within-county distribution patterns. Second, correlation coefficients were calculated for the predicted and observed number of colonized lakes per county and compared to a random selection of 47 suitable lakes over 1000 trials

## RESULTS

The best-fit results of the deterministic model designated 44 lakes as colonized with a least sum of squares (LSS) of 85. The best-fit parameterizations for this model were: distance coefficient of 1.9, colonization threshold of 850 boats, and attractiveness value of 55 000 ha for Great Lakes boat ramps. For the stochastic model, it was determined that an infested boat has a probability of 0.0000411 to establish a zebra mussel colony; this translates into a 3.5% chance of a zebra mussel colony becoming established when visited by 850 infested boats.

Sensitivity analysis (Fig. 3) showed that the model produced 44 to 47 colonized lakes over a wide range of parameter values. Changes in the distance coefficient produced the largest changes in estimated number of infected lakes, while ramp attractiveness and colonization threshold produced the least.

The results of the two tests for model robustness showed that the model was successful in duplicating the actual patterns of zebra mussel colonization. First, on a regional scale the number of colonized lakes at various distances from Great Lakes boat ramps was calculated and compared between each model (deterministic and stochastic) and the observed data (Table 2), and no significant differences were found ( $P = 0.8014$  for the deterministic and  $P = 0.9974$  for the

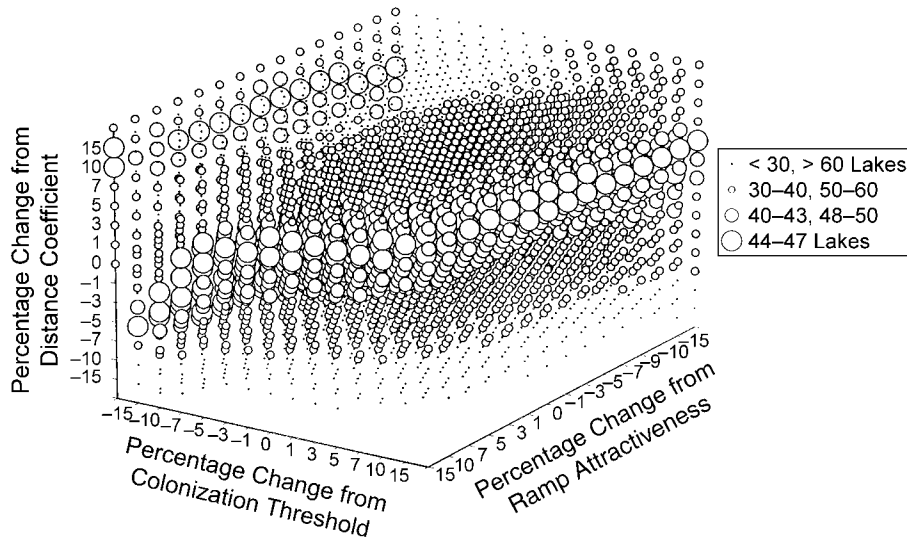


FIG. 3. Results of the sensitivity analysis. Each circle represents the number of lakes designated as colonized for combinations of values for the three test parameters: distance coefficient, colonization threshold, and attractiveness of a Great Lakes boat ramp. The size of the circles represents the number of infected lakes produced by each test parameter. The axis numbers are the percentage change of the best-fit values of the three parameters.

stochastic model). At a local scale, the number of colonized lakes within given distances of county centroids was also compared (Table 3). These results showed that patterns predicted by the models do not differ significantly from the observed data at either regional or local scales. Second, the number of colonized lakes predicted per county is highly correlated with the actual number of colonized lakes ( $r^2 = 0.639$  and  $0.681$  for the deterministic and stochastic models, respectively). These correlation coefficients are outside the 95% confidence intervals generated from correlation of observed vs. random numbers of colonized lakes per county based on 1000 simulations (Fig. 4).

*Deterministic model*

The deterministic model was successful in matching the pattern of colonization within three areas of the study region (Fig. 5A): southeast Wisconsin, southwest Michigan, and the western suburbs of Detroit. In southeast Wisconsin, the observed distribution of zebra mussels consisted of five colonized lakes in three counties of Wisconsin and northern Illinois. The model pre-

dicted three colonized lakes in the same counties, including Lake Geneva, which is colonized, with a fourth lake only 5 km outside one of the counties. Four counties in southwest Michigan have colonized lakes, while the model predicted colonized lakes in three of these counties. In the Detroit region, the most colonized county in 1997 was Oakland County, which had seven colonized lakes, and five more colonized lakes were found in neighboring counties. The model predicted that seven colonized lakes would occur in Oakland County, with three additional lakes found in surrounding counties. (Of the lakes known to be colonized in Oakland and surrounding counties, the model predicted Elizabeth Lake, Lake Maceday, Union Lake, Belleville Lake, and Stony Creek Lake to be colonized.)

The model was less successful in predicting the observed distribution of zebra mussels in central Wisconsin, southeast Michigan, and northern Indiana. The model predicted six lakes to become colonized from Dane County to Shawano County in eastern Wisconsin, where no colonized lakes had been observed. In south-

TABLE 2. A comparison of the number of observed and predicted zebra-mussel-colonized lakes within different distances to a Great Lakes boat ramp.

Distance to ramps	No. of lakes colonized		
	Observed	Deterministic model†	Stochastic model‡
<50 km	21	23	19
<100 km	40	35	38
<150 km	47	41	44

†  $\chi^2 = 0.4429$ ,  $df = 2$ ,  $P = 0.8014$ .  
‡  $\chi^2 = 0.0053$ ,  $df = 2$ ,  $P = 0.9974$ .

TABLE 3. A comparison of the number of observed and predicted zebra-mussel-colonized lakes within the specified distance of county centroids.

Distance to centroid	No. of lakes colonized		
	Observed	Deterministic model†	Stochastic model‡
<20 km	42	42	43
<10 km	14	27	21
<5 km	6	19	11
<2 km	2	3	2

†  $\chi^2 = 6.5783$ ,  $df = 3$ ,  $P = 0.0866$ .  
‡  $\chi^2 = 1.6751$ ,  $df = 3$ ,  $P = 0.6425$ .

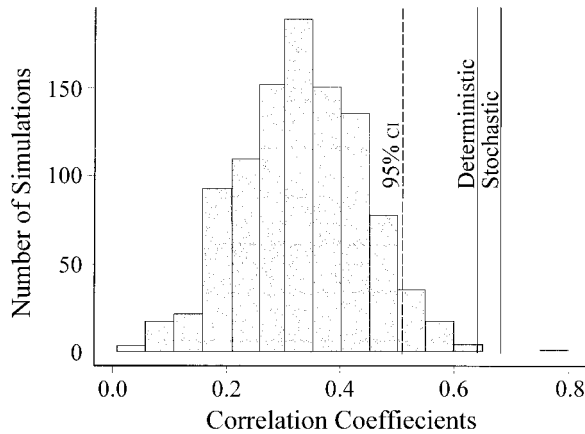


FIG. 4. Histogram of the correlation coefficients comparing the observed number of colonized lakes per county with 1000 simulations of randomly sampled lakes. The vertical lines represent the value of the correlation coefficients for the deterministic and stochastic model results compared with the observed distribution of zebra-mussel-colonized lakes. The correlation coefficient was calculated based on the number of colonized lakes per county. Both the deterministic and stochastic model correlations exceed the 95th percentile for the random distribution.

east Michigan, nine colonized lakes were found in four counties, whereas the model only predicted two colonized lakes in that region. The model also predicted four colonized lakes in northern Indiana, where 10 were observed.

*Stochastic model*

Similarities between observed and stochastically modeled distributions were evident in southern Michigan and southeast Wisconsin (Fig. 5B). In these regions, for all counties that had >1 colonized lake, the model predicted at least 0.61 colonized lakes, and the model showed similar clusters of colonized counties as were evident in the observed distribution.

Differences between the predicted and observed pattern of colonized-lake distribution were again evident in central Wisconsin, where counties that had no record of colonization were predicted to have colonized lakes. As with the deterministic model, the stochastic model predicted that colonized lakes would occur in four counties in west-central Michigan, but no lakes were known to be colonized in that region as of December 1997.

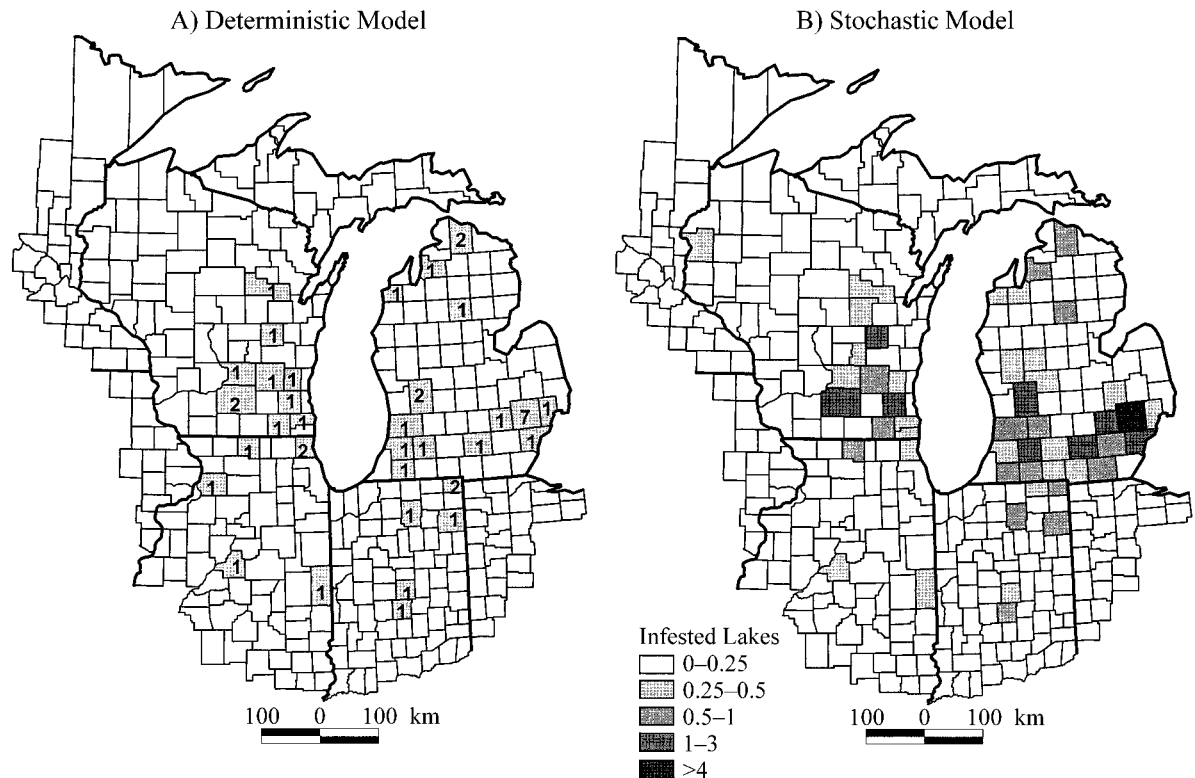


FIG. 5. The predicted distribution of zebra-mussel-colonized lakes for the (A) deterministic and (B) stochastic models. The deterministic model results were based on the best-fit parameters, and the numbers indicate the number of colonized lakes within that county. The stochastic model used the same parameters, but colonization of a particular lake was based on a colonization probability. The stochastic model was run for 2000 iterations. The number of colonized lakes for each county was determined by summing the number of lakes predicted to become colonized through all iterations and dividing by the number of trials (2000).



## DISCUSSION

*Advances over previous models*

Initial predictions of zebra mussel colonization patterns in North America were limited to forecasts of the potential geographic range at continental (Strayer 1991) or provincial (Neary and Leach 1992) scales, based upon general trends in climate, bedrock, water chemistry, and proximity to roads. Subsequent efforts have attempted to predict invasion susceptibility of lakes within more limited geographic regions (Padilla et al. 1996, Schneider et al. 1998, Buchan and Padilla 1999). The production-constrained gravity model presented in this paper builds upon these previous efforts by predicting colonization probabilities based only upon the number of registered recreational boats per county and the size, location, and water chemistry of lakes and location of boat ramps as input data. These input data are much more readily available than are data for models based upon boater surveys (Padilla et al. 1996, Schneider et al. 1998, Buchan and Padilla 1999). Boater surveys are not available for most of the study region, are expensive to generate, and are rarely repeated. For example, the boater data used by Padilla et al. (1996) and Buchan and Padilla (1999) resulted from an intense, year-long, social survey, which involved contacting 58 000 licensed Wisconsin boaters (Penaloza 1991). The boater data used by Schneider et al. (1998) was gathered from a landscape with a limited number of inland lakes (55) and zebra mussel sources (~60 sources, which included boat ramps on Lake Michigan and the Illinois, Ohio, and Mississippi rivers). Similar data are unavailable from most neighboring states with numerous lakes, even for a single year.

*Avenues for future investigation*

During model development, we were restricted by available data to fitting a model to 47 colonized lakes out of 1700+ suitable lakes. As such, not enough data were available to parameterize the model using a subset of these 47 colonized lakes, and then test the validity of the model on the remaining subset. Even though we were not able to validate our model using traditional methods, achieving a 0.67 correlation between the model forecast and fitted data suggests that the model is useful. It is important to point out that we did not simply choose one particular model while ignoring alternative ones. During parameter estimation, we essentially rejected numerous "poorer fitting" models. Our method of model selection follows the philosophy described by Hilborn and Mangel (1997), in that we selected a best-fit model from numerous alternative models with alternative parameter values. Future efforts to validate this model will only become possible when more lakes become colonized within the study region. It is interesting to note that all of the reported additional colonized lakes that have occurred since 1997 occurred in counties our model identifies as hav-

ing high colonization potentials, and that the uncolonized lake deemed most likely to become colonized (Lake Winnebago) became recognized as colonized in 1999.

We also recognize that other factors not incorporated in our model likely influence zebra mussel transport, including local boater behavior, boat type, and lake access (Penaloza 1991, Reed-Anderson et al. 2000). Some of the discrepancies between observed county-wide colonization patterns and model results suggest that such additional factors might enhance or detract from the likelihood of colonization at this spatial scale. For example, the over-prediction of colonized lakes in west-central Michigan and central Wisconsin could result from reduced attractiveness of Great Lakes boat ramps to boaters living in this region. By contrast, the under-prediction of colonized lakes in northeastern Indiana could result from increased lake attractiveness in this area.

Efforts to forecast the likelihood of colonization for specific lakes will need to take these additional factors into account. We believe our model is resilient to these factors because colonization rates are predicted only at a countywide scale. However, because colonization predictions for specific lakes will be important to resource managers, incorporating such additional information regarding boater behavior into a gravity-modeling approach will be an important future advancement.

*Implications*

Our ability to model the invasion of zebra mussels into the inland lakes of the upper Midwest provides insight into a number of important ecological issues regarding zebra mussel ecology, exotic species invasion, and the modeling of dispersal.

First, this analysis helps to resolve the debate over which vectors are likely responsible for the dispersal of zebra mussels to North American inland lakes. The mechanisms invoked in our model to reproduce observed patterns of inland zebra mussel invasions support assertions by Johnson and Carlton (1996) that recreational boat use is a more important vector of spread than are other vectors, such as waterfowl. Our analysis also helps to clarify the scale at which long-distance dispersal operates for this organism. Although 106-km distances were previously identified to represent long-distance movements for zebra mussels (Buchan and Padilla 1999), 43 of the 47 observed colonized lakes are located within this distance of a zebra mussel source. Using the definition of "long-distance" dispersal from Shigesada and Kawasaki (1997), we conclude that movement over any amount of land (no matter how short) will constitute a "long-distance" dispersal event for zebra mussels. Our model demonstrates that, at least over countywide scales (<45 km), the regional effects of such relatively short "long-distance" dispersal events are predictable.

Second, our modeling effort helps identify why the North American zebra mussel invasion has not and will not occur as a moving-wave front. Within our study region, regional aggregations of colonized-lakes are evident in observed colonized-lake distributions (Kraft and Johnson 2000) as well as in our model forecasts. It appears that the expansion of inland-lake zebra mussel range has occurred through the development of isolated centers of distribution, which have led to further aggregations of colonized lakes. This pattern is similar to that reported for many other invasive species, in which new colonies are formed beyond the boundaries of previous colonies due to long-distance dispersal across barriers (Shigesada and Kawasaki 1997). Plant invasions often demonstrate this type of spatial colonization pattern, in which satellite populations occur away from a center of introduction (Baker 1986). Such colonization patterns have not been previously described for invasive aquatic species colonizing a heterogeneous landscape of hydrologically isolated lakes, although they have been observed for wetland butterflies (Nève et al. 1996).

Last, this model demonstrates that a gravity model better predicts the overland spread of zebra mussels as compared to the diffusion model evaluated and rejected by Buchan and Padilla (1999). In general, the characteristics of gravity models provide potential advantages over diffusion models in forecasting organism movement under certain conditions. For instance, gravity models operate when potential destinations, such as inland lakes or islands, represent isolated habitats in heterogeneous landscapes where the structure and spatial pattern of potential habitat is known. Since diffusion models assume that all of a landscape is accessible, they may not be effective in situations where parts of the landscape are uninhabitable. Like stratified diffusion models (Hengeveld 1989), gravity models enable dispersing organisms to "leap-frog" suitable habitat. However, unlike stratified diffusion models, gravity models also permit organisms to jump over unsuitable habitats and dispersal barriers. For this reason, stratified diffusion models are not likely to be as useful in forecasting the dispersal of aquatic organisms to lakes within a terrestrial landscape.

Since humans are the principal overland vector for zebra mussel dispersal, the use of a model developed to forecast human movement appears clearly justified for this invasive organism. A larger question is whether gravity models have utility beyond human-dispersed taxa. We believe that gravity models are appropriate in any situation where movement between areas is negatively correlated with intersite distance and positively correlated with site "attractiveness." We suggest that such conditions apply to other taxa and landscapes. Across a wide array of systems, the biological interaction between areas can be shown to co-vary negatively with increasing intersite distance (Okubo and Levin 1989, Nekola and White 1999). Many organisms

have been shown to exhibit nonrandom movement patterns, particularly when moving a long distance (Zollner and Lima 1999). If these organisms are responding to environmental cues from potential habitat, such as the preferential uphill movement of checkerspot butterflies (*Euphydryas editha*; Turchin 1998) or female chrysomelid beetles (*Trirhabda virgata*) responding to lush host patches (Herzig and Root 1996), gravity models can be used to incorporate an attractiveness value to destination sites. The concept of attraction may also be expanded to include organisms that do not exhibit behaviors associated with neural processes, such as plants. For instance, it could be argued that passively dispersing propagules are more "attracted" to larger habitats simply because they represent larger catchment zones. It is thus possible that the two underlying conditions of gravity models exist for other examples of natural and anthropogenically mediated dispersal.

Dispersal processes occur at different scales (Levin 1992), and diffusion is obviously at work at local scales (e.g., the spread of zebra mussels throughout a lake). However, at larger landscape or regional scales, long-distance dispersal results from different mechanisms. We have demonstrated that the pattern and structure of suitable habitat is essential for modeling long-distance dispersal by zebra mussels. If an organism's dispersal is affected by distance and the attractiveness of specific destinations, gravity models may represent another important class of dispersal models that, along with diffusion models, can be used to predict the spread of native and non-native species through heterogeneous landscapes.

#### ACKNOWLEDGMENTS

We thank Evan Cooch, Greg Davis, Linda Puth, Peter Weisberg, John Wiens, two anonymous reviewers, and especially Ladd Johnson for comments on earlier drafts, which helped improve the manuscript significantly. We thank Pat Sullivan for assistance with parameter estimation and model implementation, and Christopher DuPont for helping determine centroid locations for lake polygons within counties. We also thank the Departments of Natural Resources of Illinois, Indiana, Iowa, Ohio, Minnesota, and Wisconsin, and the Michigan Department of State for furnishing us with boater registration records. Funding for this research was provided by The Institute for Land and Water Studies of the University of Wisconsin-Green Bay.

#### LITERATURE CITED

- Andow, D. A., P. M. Kareiva, S. A. Levin, and A. Okubo. 1990. Spread of invading organisms. *Landscape Ecology* 4:177-188.
- Baker, H. G. 1986. Patterns of plant invasion in North America. Pages 44-57 in H. A. Mooney and J. A. Drake, editors. *Ecology of biological invasions of North America and Hawaii*. Springer-Verlag, Berlin, Germany.
- Bossenbroek, J. M. 1999. Modeling the dispersal of *Dreissena polymorpha* (zebra mussel) into the inland lakes of the western Great Lakes region using a gravity model. Thesis. University of Wisconsin, Green Bay, Wisconsin, USA.
- Buchan, L. A. J., and D. K. Padilla. 1999. Estimating the probability of long-distance overland dispersal of invading aquatic species. *Ecological Applications* 9:254-265.
- Burgess, T. M., and R. Webster. 1980. Optimal interpolation

- and isarithmic mapping of soil properties. I. The semivariogram and punctual kriging. *Journal of Soil Science* **31**: 315–331.
- Carlton, J. T. 1993. Dispersal mechanisms of the zebra mussel. Pages 677–697 in D. Schloesser and T. Nalepa, editors. *Zebra mussels: biology, impacts and control*. Lewis, Ann Arbor, Michigan, USA.
- Dean, W. R. 1998. Space invaders: modeling the distribution, impacts and control of alien organisms. *Trends in Ecology & Evolution* **13**:256–258.
- Fotheringham, A. S., and M. E. O'Kelly. 1989. *Spatial interaction models: formulations and applications*. Kluwer Academic Publishers, Dordrecht, The Netherlands.
- Gamma Design Software. 1992. *GS+ statistical package*. Gamma Design Software, Plainwell, Michigan, USA.
- Griffiths, R. W., D. W. Schloesser, J. H. Leach, and W. P. Kovalak. 1991. Distribution and dispersal of the zebra mussel (*Dreissena polymorpha*) in the Great Lakes region. *Canadian Journal of Fisheries and Aquatic Sciences* **48**: 1381–1388.
- Hebert, P. D. N., B. W. Muncaster, and G. L. Mackie. 1989. Ecological and genetic studies on *Dreissena polymorpha* (Pallas): a new mollusc in the Great Lakes. *Canadian Journal of Fisheries and Aquatic Science* **48**:1381–1388.
- Hengeveld, R. 1989. *Dynamics of biological invasions*. Chapman & Hall, London, UK.
- Herzig, A. L., and R. B. Root. 1996. Colonization of host patches following long-distance dispersal by a goldenrod beetle, *Triphaba virgata*. *Ecological Entomology* **21**:344–351.
- Hilborn, R., and M. Mangel. 1997. *The ecological detective: confronting models with data*. Princeton University Press, Princeton, New Jersey, USA.
- Johnson, L. E., and J. T. Carlton. 1996. Post-establishment spread in large-scale invasions: dispersal mechanisms of the zebra mussel (*Dreissena polymorpha*). *Ecology* **77**: 1686–1690.
- Johnson, L. E., and D. K. Padilla. 1996. Geographic spread of exotic species: ecological lessons and opportunities from the invasion of the zebra mussel *Dreissena polymorpha*. *Biological Conservation* **78**:22–33.
- Johnson, L. E., A. Ricciardi, and J. T. Carlton. 2001. Overland dispersal of aquatic invasive species: a risk assessment of mechanisms associated with transient recreational boating. *Ecological Applications* **11**, in press.
- Kraft, C. E., and L. E. Johnson. 2000. Regional differences in rates and patterns of North American inland lake invasions by zebra mussels (*Dreissena polymorpha*). *Canadian Journal of Fisheries and Aquatic Sciences* **57**:993–1001.
- Levin, S. A. 1992. The problem of pattern and scale in ecology. *Ecology* **73**:1943–1967.
- Lewis, M. A. 1997. Variability, patchiness, and jump dispersal in the spread of an invading population. Pages 46–69 in D. Tilman and P. Kareiva, editors. *Spatial ecology: the role of space in population dynamics and interspecific interactions*. Princeton University Press, Princeton, New Jersey, USA.
- Neary, B. P., and J. H. Leach. 1992. Mapping the potential spread of the zebra mussel (*Dreissena polymorpha*) in Ontario. *Canadian Journal of Fisheries and Aquatic Sciences* **49**:406–415.
- Nekola, J. C., and P. S. White. 1999. The distance decay of similarity in biogeography and ecology. *Journal of Biogeography* **26**:867–878.
- Nève, G., B. Barascud, R. Hughes, J. Aubert, N. Descimon, P. Lebrun, and M. Baguette. 1996. Dispersal, colonization power and metapopulation structure in the vulnerable butterfly *Proclissiana eunomia* (Lepidoptera: Nymphalidae). *Journal of Applied Ecology* **33**:14–22.
- Okubo, A., and S. A. Levin. 1989. A theoretical framework for data analysis of wind dispersal of seeds and pollen. *Ecology* **70**:329–338.
- Omernik, J. M., D. P. Larsen, C. M. Rohm, and S. E. Clarke. 1988. Summer total phosphorus in lakes: a map of Minnesota, Wisconsin, and Michigan, USA. *Environmental Management* **12**:815–825.
- Omernik, J. M., and C. F. Powers. 1983. Total alkalinity of surface waters—a national map. *Annals of the Association of American Geographers* **73**:133–136.
- Padilla, D. K., M. A. Chotkowski, and L. A. J. Buchan. 1996. Predicting the spread of zebra mussels (*Dreissena polymorpha*) to inland waters using boater movement patterns. *Global Ecology and Biogeography Letters* **5**:353–359.
- Penaloza, L. J. 1991. Boating pressure on Wisconsin's lakes and rivers. Results of the 1989–1990 Wisconsin recreational boater survey, phase I. Technical Bulletin 174. Wisconsin Department of Natural Resources, Madison, Wisconsin, USA.
- Ramcharan, C. W., D. K. Padilla, and S. I. Dodson. 1992. Models to predict the potential occurrence and density of the zebra mussel, *Dreissena polymorpha*. *Canadian Journal of Fisheries and Aquatic Sciences* **49**:2611–2629.
- Reed-Anderson, T., E. M. Bennett, B. S. Jorgensen, G. Lauster, D. B. Lewis, D. Nowacek, J. L. Riera, B. L. Sanderson, and R. Stedman. In press. The distribution of recreational boating across lakes: Do landscape variables affect recreational use? *Freshwater Biology*. **43**:439–448.
- Schneider, D. W., C. D. Ellis, and K. S. Cummings. 1998. A transportation model assessment of the risk to native mussel communities from zebra mussel spread. *Conservation Biology* **12**:788–800.
- Shigesada, N., and K. Kawasaki. 1997. *Biological invasions: theory and practice*. Oxford University Press, Oxford, UK.
- Skellam, J. G. 1951. Random dispersal in theoretical populations. *Biometrika* **38**:196–218.
- Sklar, F. H., and R. Costanza. 1991. The development of dynamic spatial models for landscape ecology: a review and prognosis. Pages 239–288 in M. G. Turner and R. H. Gardner, editors. *Quantitative methods in landscape ecology: the analysis and interpretation of landscape heterogeneity*. Springer-Verlag, New York, New York, USA.
- Strayer, D. L. 1991. Projected distribution of the zebra mussel, *Dreissena polymorpha*, in North America. *Canadian Journal of Fisheries and Aquatic Sciences* **48**:1389–1395.
- Thomas, R. W., and R. J. Hugget. 1980. *Modeling in geography*. Barnes & Noble Books, Totowa, New Jersey, USA.
- Turchin, P. 1998. *Quantitative analysis of movement*. Sinauer Associates, Sunderland, Massachusetts, USA.
- U.S. Environmental Protection Agency, Office of Water. 1999. *Environmental and program information systems compendium*. U.S. Environmental Protection Agency, Washington, D.C., USA.
- Zollner, P. A., and S. L. Lima. 1999. Search strategies for landscape-level interpatch movements. *Ecology* **80**:1019–1030.