

Least-cost transportation networks predict spatial interaction of invasion vectors

D. ANDREW R. DRAKE^{1,3} AND NICHOLAS E. MANDRAK²

¹Department of Ecology and Evolutionary Biology, University of Toronto, 25 Willcocks Street, Toronto, Ontario M5S3B2 Canada

²Great Lakes Laboratory for Fisheries and Aquatic Sciences, Fisheries and Oceans Canada, 867 Lakeshore Road, Burlington, Ontario L7R4A6 Canada

Abstract. Human-mediated dispersal among aquatic ecosystems often results in biotic transfer between drainage basins. Such activities may circumvent biogeographic factors, with considerable ecological, evolutionary, and economic implications. However, the efficacy of predictions concerning community changes following inter-basin movements are limited, often because the dispersal mechanism is poorly understood (e.g., quantified only partially). To date, spatial-interaction models that predict the movement of humans as vectors of biotic transfer have not incorporated patterns of human movement through transportation networks. As a necessary first step to determine the role of anglers as invasion vectors across a land-lake ecosystem, we investigate their movement potential within Ontario, Canada. To determine possible model improvements resulting from inclusion of network travel, spatial-interaction models were constructed using standard Euclidean (e.g., straight-line) distance measures and also with distances derived from least-cost routing of human transportation networks. Model comparisons determined that least-cost routing both provided the most parsimonious model and also excelled at forecasting spatial interactions, with a proportion of 0.477 total movement deviance explained. The distribution of movements was characterized by many relatively short to medium travel distances (median = 292.6 km) with fewer lengthier distances (75th percentile = 484.6 km, 95th percentile = 775.2 km); however, even the shortest movements were sufficient to overcome drainage-basin boundaries. Ranking of variables in order of their contribution within the most parsimonious model determined that distance traveled, origin outflow, lake attractiveness, and sportfish richness significantly influence movement patterns. Model improvements associated with least-cost routing of human transportation networks imply that patterns of human-mediated invasion are fundamentally linked to the spatial configuration and relative impedance of human transportation networks, placing increased importance on understanding their contribution to the invasion process.

Key words: angler movement patterns; angling; biological invasions; gravity model; invasion ecology; least-cost routing; network theory; Ontario (Canada) anglers; secondary spread; spatial-interaction model; transportation network; vector.

INTRODUCTION

Prevention management is a primary goal of invasion research given the paucity of biological invasions whose negative impacts have been reversed (Mack et al. 2000, Simberloff 2003). Because of the abundance of possible invaders, donor and recipient ecosystems, and movement vectors, most preemptive invasion forecasts and risk models aim to evaluate the invaders, regions, or vectors that pose the greatest risk of invasion in order to prioritize prevention-management resources among perceived threats. Estimating the probability of invasion is a multi-stage process that involves quantifying the probability of introduction, establishment, reproduction, spread, and impact to recipient ecosystems (Kolar

and Lodge 2002). Risk models that incorporate only the probability of introduction may be favored in cases where potential invaders and their ecological impacts are well known within a region (e.g., zebra mussel (*Dreissena polymorpha*) and round goby (*Neogobius melanostomus*) impacts within temperate North America) and are expected to occur given the species' recent invasion history. Recent quantitative approaches to forecast invasions include: (1) identifying species or taxonomic groupings that may become future invaders based on their particular ecological or physiological attributes (e.g., Ricciardi and Rasmussen 1998, Kolar and Lodge 2002, Marchetti et al. 2004); (2) determining the similarity of donor and recipient environmental conditions to estimate the potential for establishment of known invaders, given their ecological and physiological tolerances (e.g., Ruesink 2005); (3) identifying relative strength and movement of dispersal pathways or vectors that are capable of translocation to previously uninvaded habitats (e.g., MacIsaac et al. 2004, Leung et al.

Manuscript received 28 October 2009; revised 19 January 2010; accepted 25 January 2010. Corresponding Editor: T. J. Stohlgren.

³ E-mail: andrew.drake@utoronto.ca

2006); (4) estimating the probability of propagule arrival, or relative propagule pressure, among possible donor ecosystems (e.g., Johnson et al. 2001; for reviews see Kolar and Lodge [2002], Lockwood et al. [2005], and Colautti et al. [2006]); or (5) combination of approaches 1–4 to systematically estimate the probability of introduction, establishment, and spread across potential sites (e.g., Bossenbroek et al. 2001, Herborg et al. 2007, Jacobs and MacIsaac 2009; for review see Vander Zanden and Olden [2008]). Fundamental to the applicability of each approach is the ability to successfully predict invader movement from donor to recipient ecosystem, where determining the probability of invasion requires knowledge of vector movements that have the potential to surpass background rates of natural dispersal and promote long-distance invasions (Wilson et al. 2009).

Humans have proved to be effective vectors of aquatic organisms with prevalent global transport of fishes, macrophytes, zooplankton, phytoplankton, mollusks, and pathogens (Johnstone et al. 1985, Dzialowski et al. 2000, Ruiz and Carlton 2003). Vectors may be conspicuous, yet with high dispersal potential. For instance, range expansions of pathogen-bearing mosquito larvae have been reported following the global shipping transport of used tires containing residual freshwater (Eritja et al. 2005, Benedict et al. 2007). Common vectors, such as ballast water, live food, and aquaria transport, have contributed to aquatic invasions within North America and beyond. In particular, the recreational-boating vector and the associated risk of invertebrate transport has been well studied due to the potential for secondary spread of spiny water flea (*Bythotrephes longimanus*) and zebra mussel originating within the Laurentian Great Lakes basin (Bossenbroek et al. 2001, MacIsaac et al. 2004, Muirhead and MacIsaac 2005). However, surprisingly little attention has been directed towards the role of anglers in invasion processes (but see Johnson and Carlton 1996, Ludwig and Leitch 1996, Jacobs and MacIsaac 2007). Although not necessarily mutually exclusive, angler activity patterns differ from those of recreational boaters, which have been the subject of previous spread models. When compared with recreational boaters, anglers have heightened ability to access remote destinations and smaller water bodies, possibly leading to increased movement potential across a landscape. In addition, angling activity may occur throughout the ice-cover season with mobility influenced by motorized backcountry vehicles (e.g., snowmobiles, all-terrain vehicles), further emphasizing possible movement-potential differences through access to a large number of remote locations. Anglers may inadvertently contribute to species mobility and associated range expansions through gear and clothing fouling (Jacobs and MacIsaac 2007), which have potentially disastrous consequences for recipient ecosystems given the recent invasion history of freshwater invertebrates (e.g., zebra

mussel, New Zealand mud snail (*Potamopyrgus anti-podarum*)) associated with rapid ecological change and biodiversity loss (Ricciardi and Rasmussen 1999, Chapin et al. 2000, Kerans et al. 2005). Relatively modest consideration has been given to unintentional species introductions and range expansions associated with live baitfish use, even though this practice constitutes a significant ecological and economic resource across North America (Litvak and Mandrak 1993, 1999, Ludwig and Leitch 1996). The practice of live-bait angling (and subsequent bait-bucket release) has been implicated in the range expansion of fishes, invertebrates, and their pathogens within and beyond the Great Lakes ecosystem (Goodchild and Tilt 1976, Crossman et al. 1992, Litvak and Mandrak 1993, 1999, Baxter and Stone 1995, Ludwig and Leitch 1996, Kerr et al. 2005, Keller et al. 2007). Accordingly, it is necessary to determine the spread potential of anglers as a crucial first step in assessing their possible role in invasion processes.

Vector-based invasion models provide an appropriate spatial measure of spread when human-mediated movements of species have the potential to surpass background rates of natural dispersal across a landscape, or when human movements circumvent biogeographic barriers to dispersal. Recent vector-based approaches (Bossenbroek et al. 2001, 2007, MacIsaac et al. 2004, Leung et al. 2006) used spatial-interaction models of human movements to describe the human-mediated overland dispersal of invertebrates (zebra mussel and spiny water flea) within North America. Spatial-interaction models, also known as “gravity models,” exhibit similarities to ecological distance decay models (e.g., Bell et al. 2006, Green and Bohannon 2006, Morlon et al. 2008, Perez-del-Olmo et al. 2009), but rather than predicting community dissimilarity with increasing geographic distance, they predict decreasing spatial interaction (e.g., aggregate vector movements from origins, i , to destinations, j) with increasing geographic distance, scaled by factors of relative “propulsiveness” and attractiveness across potential origins and destinations, respectively (Thomas and Hugget 1980, Fotheringham and O’Kelly 1989, and first incorporated by Schneider et al. [1998] within invasion ecology). Although spatial-interaction models may predict relatively few of the lengthiest movements across a landscape, their frequent, shorter movements may surpass even the longest species movements via natural dispersal. This distribution of movements emphasizes the ecological appropriateness of spatial-interaction models to explain species range expansions when biogeographic barriers are surpassed by vector activity.

Most spatial-interaction models that predict biological invasions or vector movements (Bossenbroek et al. 2001, 2007, MacIsaac et al. 2004, Leung et al. 2006) appear to capture the main characteristics of human movement by estimating origin propulsiveness (e.g., the relative frequency of trips leaving origins), destination

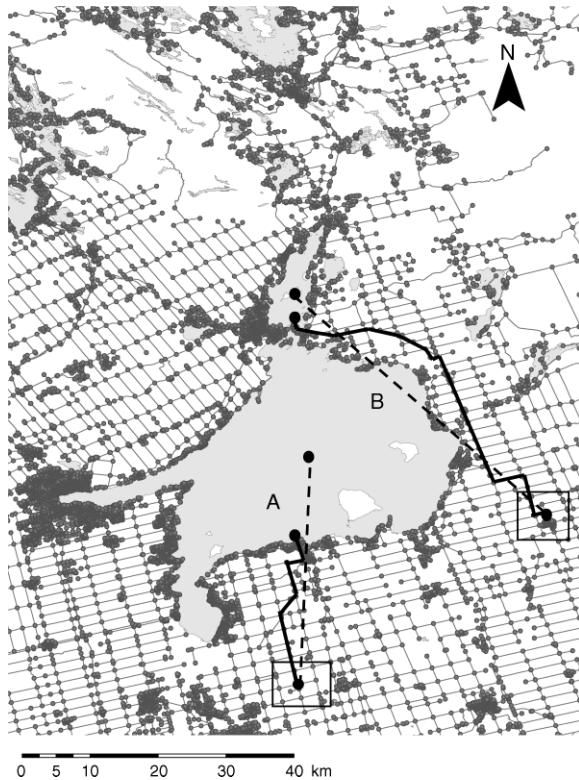


FIG. 1. A comparison of distance measures used to construct spatial interaction models of aggregate angler movements in Ontario, Canada. The comparison depicts examples of Euclidean overestimation and underestimation of distance measures in relation to network routing measures, with examples of geometric network nodes (solid gray circles) and links (gray lines) used in least-cost routing. (A) Overestimation: Euclidean route (dashed line), 30.3 km; least-cost road route (solid line), 26.6 km. (B) Underestimation: Euclidean route (dashed line), 48.5 km; least-cost road route (solid line), 76.1 km.

attractiveness (e.g., lake size) and distance traveled. However, because such models have been developed to predict invasions at coarse spatial scales, distance measures are frequently Euclidean (i.e., straight-line geographic distance), which are easily calculated but potentially misrepresent distances traveled following human movements between origins and destinations. For example, estimating vector movements to inland lakes as the Euclidean distance from the centroid of each source locality to the centroid of each destination lake may potentially overestimate distances traveled to large lakes with direct transportation routes (e.g., Fig. 1A), or underestimated distances traveled to small lakes with indirect transportation routes (e.g., Fig. 1B). Although Euclidean distance measures may provide reasonable approximations of the distance variables that are necessary to predict the distance-decay component of aggregate human movements across a landscape, more accurate estimates may be obtained by incorporating distances related to actual patterns of human movement,

such as distances traveled by road (Leung et al. 2006). The importance of investigating plausible distance measures related to patterns of human movement is emphasized by the sensitivity of previous spatial-interaction model predictions to fluctuations of their distance components (e.g., Bossenbroek et al. 2001).

Modeling approaches involving spatial networks are gaining broad popularity within habitat conservation, metapopulation dynamics, and pollination ecology (Baguette et al. 2000, Bastolla et al. 2009, Kaplan et al. 2009, Planes et al. 2009), but their application to invasion ecology remains sparse. Muirhead and MacIsaac (2005) described patterns of spiny water flea invasion, where lake invasion was a complex process that incorporated temporal and spatial patterns of vector movement throughout a geometric network of lakes. Incorporating network connectivity allowed probability of introduction predictions to be refined based on vector travel history to invaded and uninvaded sites, providing a novel approach to studying landscape-based invasion patterns. Additional network functions may provide similar advances for ecological study. Least-cost routing, which seeks to optimize travel throughout a weighted geometric network, is prevalent within telecommunication, transportation, utility (e.g., pipeline or electrical), and economic applications because it provides robust and efficient solutions to vector routing problems where multiple potential routes exist. The use of least-cost routing is gaining momentum within traditional ecological applications (Baguette and Van Dyck 2007, Gonzalez et al. 2008) despite the conceptual difficulties of assigning relative impedance (i.e., ease of flow) attributes to species movements.

The objectives of this paper are twofold: (1) confirm that movement of anglers using live baitfish can be explained using the distance-decay hypothesis within a spatial-interaction modeling framework, and (2) determine whether uncertainty reductions of spatial-interaction models can be achieved by incorporating least-cost routing of human transportation networks. Such advancements would potentially increase the accuracy and precision of spatial-interaction models within risk forecasts and prevention management. Further, model improvements associated with least-cost routing would imply that patterns of human-mediated invasion are fundamentally linked to the spatial configuration and flow of transportation networks, placing increased importance on their contribution to the invasion process.

METHODS

Developing a spatial-interaction model

To model vector movement in accordance with distance-decay hypotheses, we used an approach similar to those of deterministic spatial-interaction models (see Thomas and Hugget [1980] and Fotheringham and O'Kelly [1989] for reviews). Deterministic spatial-interaction models describing the aggregate movement, or

TABLE 1. Numerical summary of training, validation, and complete data sets used to construct spatial-interaction models of aggregate angler movements, involving live baitfish, in Ontario, Canada.

Model component	<i>N</i>	Range	Median	Mean (SD)
Complete data set				
Spatial interaction, T_{ij}	84 249	0–22	0	0.0228 (0.2987)
Origin propulsiveness, o_i †	207	1–119	4	9.28 (14.28)
Road distance, D_{ij} (km)	n.a.	0.00734–2389	472.2	674.10 (559.91)
Euclidean distance, D_{ij} (km)	n.a.	0.07–2061	356	514.92 (441.18)
Destination attractiveness, $w_{j,1}$ (ha)‡	407	2.2–8 341 300	550	51 678.9 (495 489.4)
Destination attractiveness, $w_{j,2}$ §	407	1–26	6	6.13 (3.51)
Training data set				
Spatial interaction, T_{ij}	58 974	0–22	0	0.0234 (0.3100)
Origin propulsiveness, o_i †	207	1–119	4	9.23 (14.16)
Road distance, D_{ij} (km)	n.a.	0.00734–2389	472	675.00 (561.19)
Euclidean distance, D_{ij} (km)	n.a.	0.07–2061	356	515.70 (442.41)
Destination attractiveness, $w_{j,1}$ (ha)‡	407	2.2–8 241 300	550	52 747.0 (505 149.7)
Destination attractiveness, $w_{j,2}$ §	407	1–26	6	6.13 (3.51)
Validation data set				
Spatial interaction, T_{ij}	25 275	0–9	0	0.0214 (0.2200)
Origin propulsiveness, o_i †	207	1–119	4	9.39 (14.56)
Road distance, D_{ij} (km)	n.a.	0.01–2380	472	671.90 (556.94)
Euclidean distance, D_{ij} (km)	n.a.	0.53–2029	355	513.11 (438.35)
Destination attractiveness, $w_{j,1}$ (ha)‡	407	2.2–8 241 300	550	49 186.9 (472 181.3)
Destination attractiveness, $w_{j,2}$ §	407	1–26	6	6.14 (3.50)

Note: An entry of “n.a.” indicates not applicable.

† Aggregate movements leaving origin.

‡ Lake surface area.

§ Sportfish richness.

flow, of humans between origin and destination as a function of their distance separation, the propulsiveness of an origin, and the attraction of a destination are frequently of the following form:

$$T_{ij} = m o_i w_j D_{ij}^{-\gamma}$$

where T_{ij} is an element of matrix **T** describing aggregate movements between origins, i , and destinations, j , m is a constant, o_i is an element of vector **o** describing the propulsiveness of each origin, w_j is an element of vector **w** describing the attractiveness of each destination, D_{ij} is an element of matrix **D** describing the origin–destination distance and γ is a shape parameter describing the distance relationship. Recent models have been estimated using generalized linear models with Poisson distributions to predict counts of aggregate spatial interactions as the response variable (e.g., Bergkvist and Westin 2001).

Because most origins within our data set (Table 1) interacted with only a small number of destinations, resulting in the absence of T_{ij} across the bulk of origin–destination pairs, the response variable was characterized by more zeros than would be expected by a Poisson distribution. In addition, models incorporating Poisson distributions perform poorly when the response variable is overdispersed such that the variance is greater than the mean (Cameron and Trivedi 1998, 2005, Hilbe 2007), which was a characteristic of our data (Table 1). As an alternative, the negative binomial distribution is suitable for overdispersed data; further, a zero-inflated distribution may account for excess zeros that can occur

within count data (Hilbe 2007). We fit generalized linear models using both Poisson distributions (as a baseline comparison to previous studies) as well as zero-inflated negative binomial distributions to investigate possible model improvements by accounting for overdispersion and zero inflation. We followed the approach of Bergkvist and Westin (2001) and Potapov et al. (*in press*) to develop spatial-interaction models using principles of maximum likelihood to fit model coefficients as

$$T_{ij} = \beta_0 + \beta_1 \log(o_i) + \beta_2 \log(w_j) + \beta_3 \log(D_{ij})$$

where T_{ij} is the response variable, β are estimated model coefficients and o_i , w_j , D_{ij} are explanatory variables describing origin propulsiveness, destination attractiveness, and geographic distance between i and j , respectively (see *Developing a least-cost transportation network and Empirical data used for model parameterization*, below, for a description of empirical model variables). A natural-logarithm link function was used to model counts of spatial interaction as the response variable. Following Burnham and Anderson (2002), Akaike’s information criterion (AIC; Akaike 1974) was used to select the best model among all classes of competing models. To avoid model over-fitting (Olden et al. 2002), coefficients were fitted to a randomly selected training data set (70% of origin–destination pairs) and were subsequently validated within a validation data set (30% of origin–destination pairs). Because the probability of both zero interaction and interaction values greater than zero was of interest, all possible origin–destination

combinations of $T_{ij} \geq 0$ ($n = 84\,249$, given 207 origins and 407 destinations) were computed and randomly allocated into training ($n = 58\,974$) and validation ($n = 25\,275$) data sets for purposes of model validation (Table 1; Fig. 2A–C). During the random allocation process, proportions of $T_{ij} \geq 1$ and $T_{ij} = 0$ within training and validation data sets were forced to be consistent with those proportions of the overall spatial-interaction data set. Comparison of AIC values between Poisson and zero-inflated models derived using both network and Euclidean distance measures was used to select the most parsimonious model, which was subsequently applied to the statistical population of o_i , w_j , and D_{ij} to forecast spatial interaction across Ontario as a function of the relative strength of variables chosen during the model-selection process. Validating model performance within the independent data set allowed model error to be characterized as a rough measure of overall fit. Error rate was described using root mean squared error (RMSE) between actual and predicted T_{ij} , which incorporates a quadratic scoring rule to describe the average magnitude of error in T_{ij} units. Values of RMSE range from zero to infinity, with values closest to zero indicating a relatively low magnitude of average model error. Model residuals were classified based on directional patterns (e.g., over- and under-prediction terms) and plotted geographically to visually identify spatial patterns of prediction error. The validated model's null deviance divided by its residual deviance was subtracted from 1 to determine the total deviance explained, which provided an overall-fit metric similar to Pearson's R^2 . Interaction scores were used to rank destination water bodies based on their projected number of aggregate arrivals following model forecasts across the statistical population. All statistical analyses were performed using the statistical language and software program R, version 2.8.0 (R Development Core Team 2008).

Developing a least-cost transportation network

To model vector movement as a function of distance decay, we required estimates of actual distances that anglers travel to reach their destination. Because most angling events occur in conjunction with vehicular overland travel, we estimated travel distances according to the distance accumulated during travel throughout a provincial road network following the application of a least-cost routing algorithm. A total of 476 869 line segments and 351 930 nodes (e.g., Fig. 1), obtained from Ontario Ministry of Transportation data, were used to construct a spatially explicit geometric road network for Ontario, Canada, within ArcGIS version 9.3 (ESRI 2008). Line segments were originally attributed with roadway classifications (Local/Street, Collector, Arterial, Expressway/Highway, and Freeway) that were subsequently converted into theoretical travel speeds (50, 60, 70, 80, and 100 km/hr, respectively) for each segment, where travel time through each segment was

the product of assigned travel speed divided by segment length. Line segments representing vehicular travel routes onboard provincial ferries used to traverse the Laurentian Great Lakes were excluded from further analysis, due to the uncertainty associated with determining travel speed for these areas. Nodes at segment endpoints provided physical connections between adjoining segments; however, nodes of intersecting line segments with $>10\%$ change in elevation were removed, because these areas likely represented overpasses or tunnels that confine travel. Because of the abundance of possible travel routes throughout the network, we introduced an impedance function to optimize travel routes in accordance with economic travel principles whereby human behaviors minimize travel time and financial expenditure (Wilson 1967, 1998). Therefore, because vectors should select routes that minimize travel time, but not necessarily distance, a least-cost routing algorithm was constructed that selected a single optimal route from all possible routes between i and j to minimize travel time given relative impedance. This time is given by

$$S_{ij} = \min_k \sum_{l=1}^{uijk} uijkl$$

where $uijkl$ is the travel time (cost) of the l th segment along the k th route between origins, i , and destinations, j , and S_{ij} is the optimal route between origins, i , and destinations, j . Geographic travel distance (km) associated with each optimal least-cost route (D_{ij}) was incorporated as the distance variable within the spatial-interaction model. Conventional Euclidean distance estimates (D_{ij}) were also derived between origins and lake centroids, or between origins and lake access points for the Laurentian Great Lakes. Model parsimony and explained deviance were compared for models derived using least-cost routing and Euclidean distance estimates to determine possible model-performance improvements owing to least-cost routing through a geometric network.

Empirical data used for model parameterization

To parameterize our models, we required actual estimates of spatial interaction describing aggregate movements between origins, i , and destinations, j , their geographic separation, and certain characteristics of i and j that may contribute to variation in the relative strength movements. To address these data requirements, a social survey was developed to determine characteristics of angler behavior (e.g., frequency of angling activity, prevalence of bait-bucket release) and movement patterns of anglers using live baitfish in Ontario, Canada (see Appendix for list of survey questions). Hardcopy paper surveys ($n = 5000$) were distributed using the modified Dillman (1978, 2000) method, where surveys were mailed to potential respondents proportionate to the density of licenced anglers (those anglers ≥ 18 and < 65 years of age who

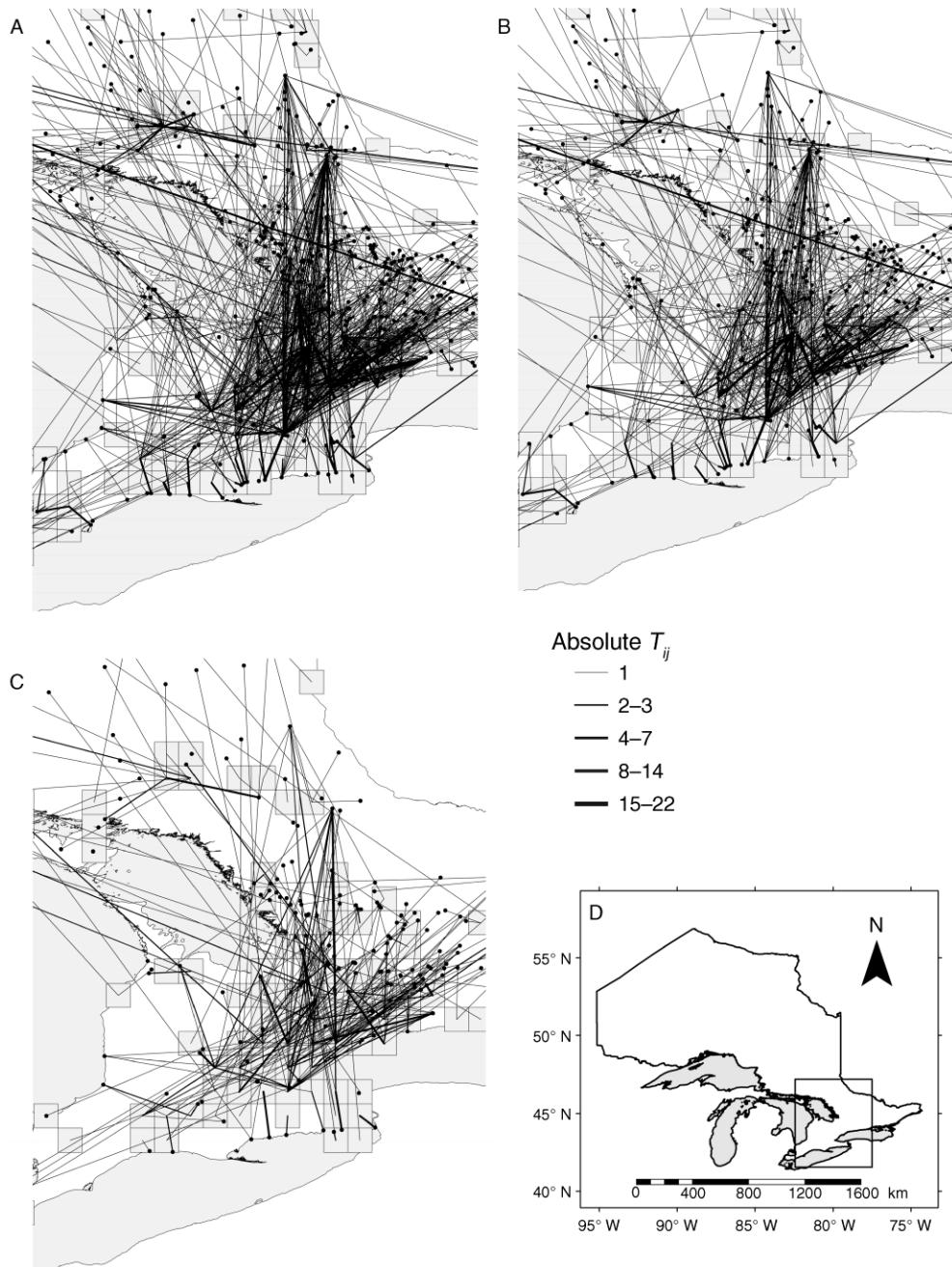


FIG. 2. Spatial-interaction data set (A; $n = 84,249$ origin-destination values [O-D's]) with randomly allocated training (B; $n = 58,974$ O-D's) and validation (C; $n = 25,275$ O-D's) aggregate movement (T_{ij}) observations. Lines represent origin-destination spatial-interaction pairs (T_{ij}) from origin, i (centroid of $0.5^\circ \times 0.5^\circ$ longitude-latitude grid; open squares) to destination water bodies (lake-access points; solid circles), where line thicknesses represent absolute T_{ij} with only $T_{ij} > 0$ displayed. The geographic location of highlighted vector activity is disclosed within an inset map (D) of Ontario, Canada.

held valid fishing licences during the 2007 season) residing in each of five broad postal districts: K (eastern Ontario), L (suburban Toronto), M (metropolitan Toronto), N (southwestern Ontario), and P (northern Ontario). To increase our sample size, an online survey was developed and advertised within angling retailers, sporting magazines, and angling-related web pages to

allow for internet respondents who had not received paper surveys. Additional responses were generated through survey booths at sporting tradeshows in the greater Toronto area between February and May 2006. Only complete surveys were analyzed as an a priori measure of data quality. To account for potential nonresponse bias (Fisher 1996) or geographic bias

across all surveys, a G test (Zar 1999) of maximum likelihood was used to determine that the frequency of respondents from each postal district were consistent with the geographic distribution of licenced anglers in those regions at the $\alpha = 0.05$ level.

The primary survey responses used to address spatial-interaction patterns were: Question 3, How often do you use live baitfish during the year?; Question 4, Do you catch your own baitfish?; Question 10, List the top-three cities or towns where you buy your bait; and Question 13, List the top-three places (lake or river name, province) in which you use baitfish for angling. Anglers were also asked to provide their six-character postal code, which provided a point-based residence location following latitude-longitude point conversions with Canada Post's postal geography data. To relate survey responses to specific destination characteristics, a georeferenced lake database (the Ontario Freshwater Fish Species Distribution Dataset [Ontario Ministry of Natural Resources, Peterborough, Ontario, Canada]; see Mandrak and Crossman [1992] for description) was used to determine the geographic location, name, surface area (ha), and sportfish richness of 9356 lakes in Ontario. Because the primary location of each lake was represented as a point-based location occupying the position of each lake's centroid, a database of spatial lake extent, obtained from Fisheries and Oceans Canada (Burlington, Ontario, Canada), was used to relate the characteristics of each lake (e.g., sportfish richness, maximum depth) to its specific areal coverage. Lake records whose centroids fell outside of the spatially explicit lake-shape database were omitted from subsequent analyses due to concerns about data quality. To determine specific endpoints in relation to each lake's spatial configuration, a series of lake access points were created by overlaying a spatial buffer of either 100, 300, 1500, 3000, or 5000 m of a lake's spatial coverage onto transportation network nodes. Large lakes close to urban areas (e.g., Lake Ontario, Lake Simcoe) were assigned the smallest (100 m) buffer values due to high abundances of proximal nodes, whereas small, remote lakes were assigned larger values. Overlapped nodes represented specific lake-access points for each corresponding water body. Because a one-to-many spatial join was utilized, a single point could represent access to multiple lakes if several lakes existed in close proximity to a single node. To determine the specific access point utilized for lakes with multiple access possibilities, a simplifying assumption was introduced where each optimized least-cost route was matched with a single access point that minimized travel time between i and j .

For each respondent, up to three possible origin-destination pairs were produced, where origin locations were initially represented by a six-character postal code, and destination locations were represented by up to three water bodies identified in survey Question 13. With the exception of the Ottawa River, which due to its size functions as a lake throughout much of its length, only

lakes were used to construct origin-destination pairs due to the difficulty of positively identifying rivers that were identified by survey respondents. However, because the estimated number of lakes in Ontario exceeds 250 000 (Cox 1978) with many duplicated lake names (e.g., Clear Lake, 26 unique locations; Trout Lake, 24 unique locations; Ontario Freshwater Fish Species Distribution Dataset), we incorporated an optimization process for each possible origin-destination pair to select the most likely location of each destination water body when lakes with duplicate name records were encountered. For each survey respondent, the location of the origin was plotted in addition to up to three baitfish purchase locations (cities, towns). All duplicate lakes were plotted, but a single destination lake (and corresponding access point) that minimized least-cost routing from origin, through baitfish purchase location, to final destination water body was selected as the most likely destination location. The process was repeated for up to two remaining origin-destination pairs if additional duplicate records were encountered. For cases where respondents indicated that they did not purchase baitfish (with purchase locations omitted), the shortest origin-destination route among duplicate lakes was used to select each optimal destination. Respondents indicating that they did not use baitfish were omitted from spatial-interaction analysis.

Variables of origin propulsiveness (o_i) and destination attractiveness (w_j) were selected based on characteristics of angling activity within Ontario, Canada. Due to their fine resolution, relatively few angling trips are generated from geographic areas delineated by single six-character postal codes. To increase the applicability of origin-destination projections during model forecasting at the provincial level, origins were expanded by projecting an $0.5^\circ \times 0.5^\circ$ latitude-longitude grid onto the location of six-character postal codes, where the number of movements leaving each grid cell was equal to the aggregate number of movements leaving all six-character postal codes contained within those areas. Both lake surface area (ha; $w_{j,1}$) and sportfish richness ($w_{j,2}$) were selected as explanatory variables potentially contributing to spatial-interaction patterns because lake size and sportfish richness should enhance the availability, quality and diversity of angling opportunities.

Following model selection and validation, forecasting T_{ij} across the statistical population required additional estimates of o_i , $w_{j,1}$, $w_{j,2}$ and D_{ij} . The geographic distribution of licenced resident anglers during 2007 in Ontario summarized as point-based files using spatial conversion of six-character postal codes, formed the initial estimate of o_i . However, because our model was developed specifically from survey data of angling trips incorporating live baitfish, a linear model of the form $Y = (m)X + b$ was produced to describe the relationship of surveyed anglers within i against the number of origin-destination pairs leaving i to estimate the aggregate number of movements leaving each origin cell at the population level. Survey lakes with previously deter-

mined access points and all additional lakes <5000 m from a road junction and containing at least one sportfish were eligible for model forecasting at the population level. Lakes contained within Algonquin and Quetico Provincial Parks were excluded because fishery management regulations preclude the use of live baitfish within these areas and, accordingly, would influence associated movement patterns.

RESULTS

Survey data

Complete survey responses ($n = 1398$) across each of five postal districts were initially inflated towards the suburban Toronto postal district L (G test of maximum likelihood; G critical 14.86; G statistic following Williams correction Q 16.02; $\alpha = 0.05$; $P > 0.05$). Random selection and subsequent removal of five responses from within the suburban Toronto postal district provided overall response frequencies consistent with the geographic distribution of anglers licenced in Ontario during calendar year 2007 (G test of maximum likelihood; G critical 14.86; G statistic following Williams correction Q 14.78; $\alpha = 0.05$, $P < 0.05$). Because individual survey responses provided up to three destination water body responses, initial origin-destination pairs ($n = 1921$) were greater than the total number of respondents. Spatial overlay and origin summaries using the $0.5^\circ \times 0.5^\circ$ latitude-longitude grid further reduced the number of origin-destination pairs of $T_{ij} > 0$ ($n = 1170$) with 207 cell origins and 407 destination lakes. The spatial buffer of lakes onto road network junctions produced 16 037 unique lake-access locations across destination water bodies identified from survey responses. Pairwise comparison of travel routes obtained from survey data suggested that Euclidean distance measures frequently underestimated network-based measures, especially at the largest geographic distances (Fig. 3). Overestimation by Euclidean distance measures also occurred, albeit at lower frequencies.

Model selection and validation

Four spatial-interaction models were initially selected using Euclidean and network-based distance measures with Poisson and zero-inflated negative binomial distributions. Poisson models selected using maximum likelihood were characterized by relatively high AIC values, with the Euclidean-based model being more parsimonious than the network-based model (network-based model with log link, $T_{ij} = -2.214 + 0.821 \log(o_i) - 0.595 \log(D_{ij}) - 0.330 \log(w_{j,2}) - 0.280 \log(w_{j,1})$, all coefficient $P < 0.05$, AIC = 8558.086; Euclidean-based model with log link, $T_{ij} = -1.727 + 0.921 \log(o_i) - 0.825 \log(D_{ij}) - 0.421 \log(w_{j,2}) + 0.306 \log(w_{j,1})$, all coefficient $P < 0.05$, AIC = 7928.116). Both zero-inflated models were characterized by significantly lower AIC values derived during the fitting process (network-based model with log link, $T_{ij} = -3.957 + 1.563 \log(D_{ij}) - 0.421 \log(o_i) - 0.271 \log(w_{j,1}) - 0.030 \log(w_{j,2})$, all coefficient P

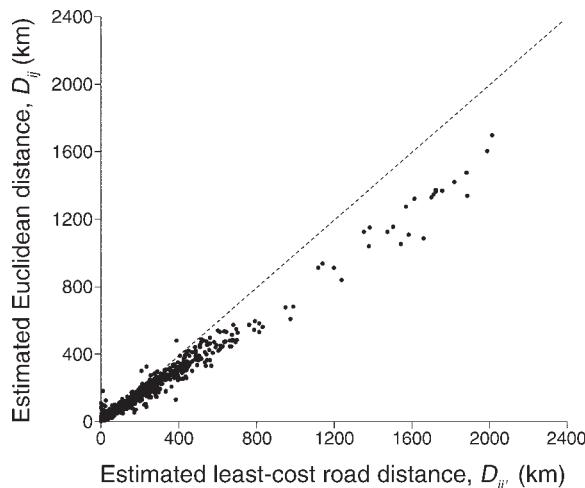


FIG. 3. Comparison of empirical Euclidean (D_{ij}) and least-cost road network (D_{ij}) distance measures across all possible origin-destination pairs where $T_{ij} > 0$ ($n = 1170$ pairs).

< 0.001, AIC = 6480.752; Euclidean-based model with log link, $T_{ij} = -2.597 + 1.385 \log(D_{ij}) - 0.312 \log(o_i) - 0.300 \log(w_{j,1}) - 0.2685 \log(w_{j,2})$, all coefficient $P < 0.001$, AIC = 6564.538). Poisson model validation revealed higher average model error (network-based Poisson model RMSE [root mean square error] $T_{ij} = 5.232$, Euclidean-based Poisson model RMSE $T_{ij} = 5.510$) when compared with the average error of zero-inflated models (network-based zero-inflated negative binomial model RMSE $T_{ij} = 0.270$, Euclidean-based zero-inflated negative binomial model RMSE $T_{ij} = 0.374$). The network-based zero-inflated model was selected as the most parsimonious model due to its low AIC value, and subsequently performed best at forecasting survey data as indicated by a relatively low average error rate and the highest proportion of deviance explained (0.477; Fig. 4A) when compared to the Euclidean model (0.422; Fig. 4B).

As a common characteristic of Poisson and zero-inflated negative binomial residuals, both network- and Euclidean-based residuals were dominated by negative values, indicating over-prediction of T_{ij} . However, the many, negative residuals were characterized by low absolute values (mean network-based $T_{ij} = 0.019675$ and mean Euclidean-based $T_{ij} = 0.019671$) when actual T_{ij} was zero. When actual $T_{ij} \geq 1$, models were dominated by positive residuals for both network-based ($n = 344$ positive residuals; mean absolute $T_{ij} = 1.086$; Fig. 5A) and Euclidean ($n = 345$ positive residuals; mean absolute $T_{ij} = 1.119$; Fig. 5B) forms, indicating under-prediction of spatial interaction in those specific geographic areas. Over-prediction for network-based ($n = 26$ negative residuals; mean absolute $T_{ij} = 4.099$; Fig. 5C) and Euclidean ($n = 25$ negative residuals; mean absolute $T_{ij} = 5.341$; Fig. 5D) forms occurred relatively infrequently when actual $T_{ij} \geq 1$, but with considerably higher magnitude.

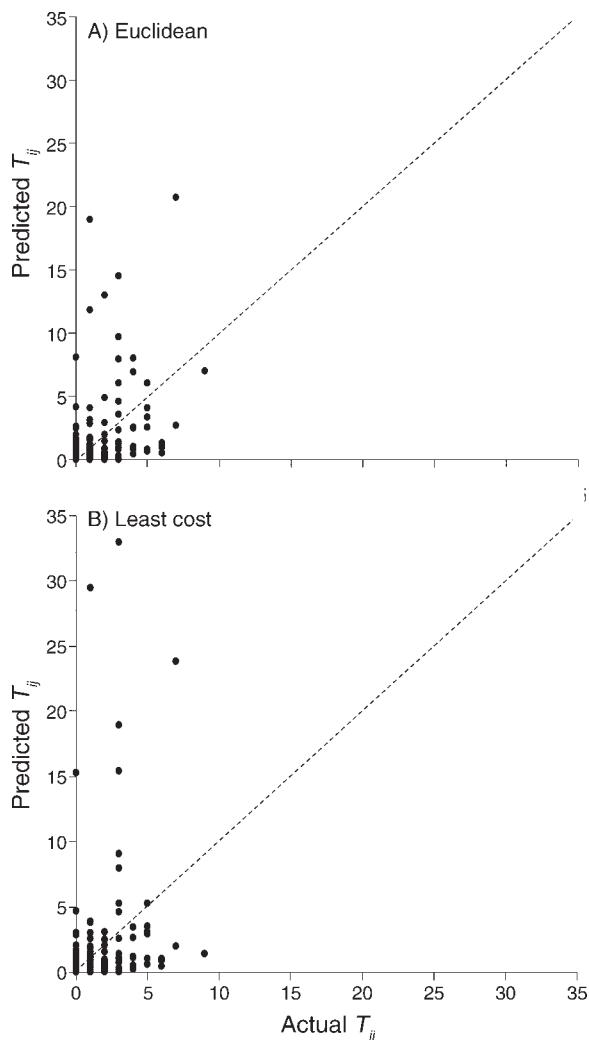


FIG. 4. Comparison of model predictions vs. observed spatial interaction (T_{ij}) using (A) Euclidean (D_{ij}) and (B) least-cost road network (D_{ij}) distance measures. The dashed line represents the 1:1 equivalence line.

Model forecasting across the statistical population

The most parsimonious model (network-based with zero-inflation) was selected to forecast spatial-interaction patterns at the provincial level. The linear model used to determine the relationship between aggregate movements leaving each 0.5×0.5 grid cell and the number of registered anglers within each grid cell was population $a_i = 0.229 + 1.34$ (count per grid cell); F statistic = 4876 with $df = 245$; $P < 0.001$; adjusted $R^2 = 0.952$. Absolute population-level spatial-interaction values ($n = 1\,246\,840$) from 427 origins to 2920 destination water bodies ranged from 0 to 567 213, with mean T_{ij} of 2.48. A distance–frequency histogram (Fig. 6) of origin–destination movements where $T_{ij} \geq 1$ suggested that most vectors undergo relatively short to modest travel distances (median = 292.60 km) with

certain vectors traveling significantly further (75th percentile = 484.58 km, 95th percentile = 775.20 km). The summation of interaction values summarized across each destination water body suggested that Ontario's five largest water bodies (lakes Ontario, Superior, Huron, and Erie and Georgian Bay) and the Ottawa River receive the bulk of spatial interaction across the province (Table 2). Only two of the ranked top-20 lakes were <6000 ha in size, further emphasizing the importance of lake size and its contribution to movement patterns.

DISCUSSION

Angler movement was explained across the provincial level using the distance-decay hypothesis within a spatial-interaction modeling framework. The most suitable model, based on least-cost routing of human transportation networks and a zero-inflated negative binomial distribution, ranked distance traveled, origin outflow, lake size, and sportfish richness as influential variables (listed in decreasing order of importance). Influential variables confirmed the association between angler movement and expected patterns of spatial interaction. The importance of lake size suggested that large lakes influence angler movement by enhancing sportfishing opportunities due to increased size of fish populations, as well as through access to angling facilities (e.g., prevalence of retailers, lodging, docking facilities). However, analysis of the 20 water bodies receiving greatest aggregate movement (Table 2) suggested that large outflows from origins will result in substantial movement to small, nearby lakes. Associations between invasion vectors and distance-influenced patterns of spatial interaction have been found for the recreational boating pathway (Bossenbroek et al. 2001, 2007, MacIsaac et al. 2004, Leung et al. 2006), suggesting that although the specific contribution of distance, lake attractiveness, and origin outflows will change among vector types based on geographic variation and user mobility, these variables collectively influence vector movements across large spatial scales.

Model results suggest that roughly half of aggregate angler movement cannot be explained using patterns of distance decay. Most model errors occurred with the under-prediction of lengthy movements to relatively small water bodies (Fig. 5A, B), suggesting that the best-fit model describes a conservative scenario of angler movement at the provincial level. The prevalence of under-prediction error was likely the result of the model's inability to capture unique trip scenarios that do not fit expected patterns of movement (e.g., anglers that travel long distances to angle in small water bodies, such as those owning recreational property, or anglers that preferentially choose to angle within the wilderness regions of northern Ontario, Canada). However, capturing these movements within the model framework would have been difficult without the collection of more detailed survey data concerning the perceived quality of angling in relation

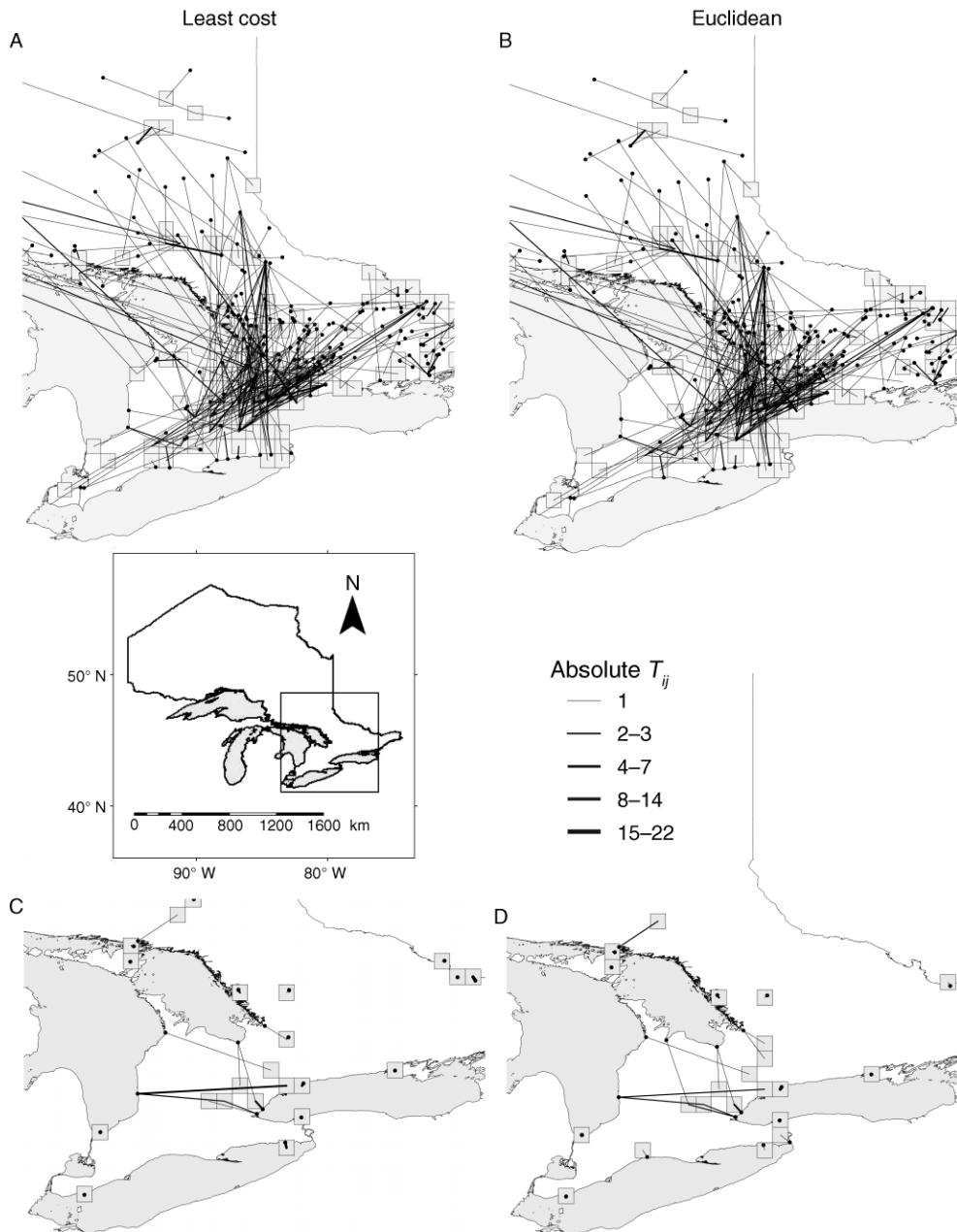


FIG. 5. Spatial-interaction model residuals of underpredictions (A, least-cost road network; B, Euclidean) and overpredictions (C, least-cost road network; D, Euclidean) when actual $T_{ij} \geq 1$. Residuals are shown as absolute spatial interactions (T_{ij}) units. T_{ij} movements define origins (i) and destinations (j).

to specific geographic areas. We chose not to collect these data due to their subjective nature, and because we wanted relatively simple model metrics (e.g., lake size) that were available across the provincial level and, therefore, could easily be forecast at large spatial scales.

Although most spatial interactions were relatively short compared with the lengthiest movements (Fig. 6), median distance values were 292.6 km, emphasizing the movement potential of anglers as invasion vectors across a landscape in which drainage basins have a mean primary

axis of 65.5 km. Although model results show that anglers in Ontario display movement patterns with the potential to circumvent natural dispersal barriers, long-distance vector movements are only one of several prerequisites necessary for biotic transfer. Successful biological invasions associated with angler movement require the progression of specific risk activities (e.g., fouling of invasive biota or purchase of invasive baitfishes; movement to uninvaded system; release; survival; reproduction; impact). The importance of the progression of

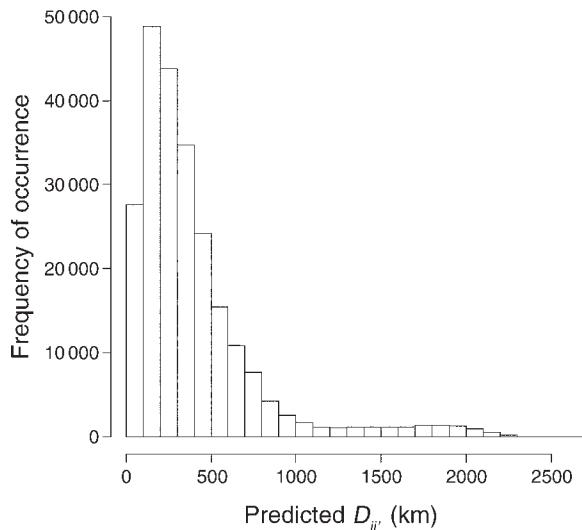


FIG. 6. Distance–frequency histogram of population-level spatial-interaction model predictions (T_{ij} , where $T_{ij} \geq 1$) derived using least-cost road network distance measures. D_{ij} is the geographic distance associated with each optimal origin-to-destination least-cost route.

pathway steps for successful invasion requires that future research focus on quantifying risk activities (e.g., the probability of bait-bucket release of non-indigenous species) as they relate to movement patterns to determine the distribution of risk activities across large spatial scales. Such studies are necessary to prioritize recipient ecosystems in anticipation of their invasion (Lockwood et al. 2005, Colautti et al. 2006, Herborg et al. 2007). However, as a necessary first step to determine the spread potential of anglers given their history as invasion vectors (Crossman et al. 1992, Litvak and Mandrak 1993, 1999,

Baxter and Stone 1995, Jacobs and MacIsaac 2007, Keller et al. 2007), we highlight the mobility of anglers to circumvent biogeographic barriers to dispersal across a landscape where angling activity is prevalent and has a rich history. Although specific distance-decay parameters of anglers may differ globally across geographic jurisdictions, these results highlight the general mobility of anglers and emphasize the movement potential of likely future invaders in other ecosystems (e.g., rock snout *Didymosphenia geminata*; Bothwell et al. 2009, Kilroy et al. 2009; New Zealand mudsnail *Potamopyrgus antipodarium*; Kerans et al. 2005) whose mobility into novel habitats are enhanced dramatically by overland movement vectors.

Routing optimization approaches to travel have been used extensively within human and transportation geography. Least-cost routing approaches to road-network travel with automobiles were initially developed for the “traveling salesman problem” (e.g., Lawler et al. 1985), where routing algorithms attempted to find optimal solutions to minimize total travel time among several cities. Similar applications are used to solve vehicle routing problems (e.g., Braysy et al. 2009), where a fleet of vehicles are used as efficiently as possible given travel times between service locations. In both examples, least-cost routing of transportation networks mimics actual patterns of human movements, which frequently minimize energy expenditure, thereby resulting in reduced financial output or time required for travel (Wilson 1967, 1998, Kolbl and Helbing 2003). Such theoretical underpinnings have formed the basis for archaeological route tracing, where spatial pathways derived from human metabolic optimizations of energetically variable landscapes predict historic travel routes (e.g., Wood and Wood 2006). Although human travel may, in certain

TABLE 2. Priority-ranked (top 20) destination water bodies receiving greatest spatial interaction at the provincial level. Predictions were summarized following application of the network-based spatial-interaction model that incorporated least-cost routing.

Destination water body	Surface area, $w_{j,1}$ (ha)	Sportfish richness, $w_{j,2}$	Interaction score, ΣT_j (absolute)
Lake Ontario	1 952 900	25	1 047 718
Ottawa River	127 100	15	195 209
Lake Erie	2 574 500	26	133 094
Georgian Bay	1 500 000	26	97 104
Lake Huron	4 459 600	26	54 619
Lake Superior	8 241 300	14	30 331
Lake Nipissing	87 330	14	15 490
Lake Simcoe	74 400	14	15 150
Lake Muskoka	12 206	13	10 822
Lake Manitou	10 400	6	10 795
Lake of the Woods	435 000	13	8 524
Otter Lake	602	10	8 447
Lake of Bays	6 904	9	5 756
Lake St. Clair	110 000	20	5 713
Lake Scugog	8 256	8	5 333
Rainy Lake	89 400	12	5 207
Cranberry Lake	227	3	5 010
Lake Wanapitei	62 200	10	4 940
Rice Lake	10 017	9	4 611
Panache Lake	8 796	5	4 563

cases, follow alternative routing patterns (e.g., scenic routing or freeway avoidance), optimized routing by travel time represents the most likely distance estimate based on observed patterns of human movements.

Utilizing least-cost routing of human transportation networks provided measurable model improvements over standard Euclidean measures. Models derived from Euclidean distance measures were less parsimonious and explained less deviance than those incorporating transportation networks and least-cost routing, suggesting that Euclidean-based models generally suffer from methodological deficiencies when used to explain human movement patterns. Although Euclidean-based measures may be reasonable when forecasting movement at large spatial scales, they appear to digress when forecasting high levels of interaction between short distances to large lakes, and low levels of interaction between long distances to small lakes. In both of these scenarios, Euclidean measures performed more poorly than when incorporating network approaches because of an apparent nonuniform distance-smoothing effect that influences origin–destination pairs. Network approaches were superior in both the explanatory and forecasting stage of model development, suggesting that they provide a robust improvement over previous approaches. Greater predictive capability is likely the result of incorporating concepts such as travel impedance that more accurately reflect human movement processes. Uncertainty reductions of movement models for alternate vectors may be possible when utilizing an approach based upon human transportation networks. Model comparison results imply that patterns of human-mediated invasion are fundamentally linked to the spatial configuration and relative impedance of transportation networks, placing increased importance on understanding their contribution to the invasion process.

Other potential applications of least-cost routing within invasion ecology may involve predicting the secondary spread of invaders via natural dispersal through disjunctly connected aquatic habitats, following their establishment within large water bodies (e.g., the Laurentian Great Lakes). Here, least-cost routing would require classifying impedance of connecting channels by species that are expected to increase their invaded range through natural dispersal. The least-cost invasion route would represent the initial foray into uninvaded waters, allowing dispersal routes to be preventively managed to discourage movement into novel habitats.

Distance–decay models have gained wide popularity within ecological study because they represent patterns of many intraspecific and interspecific ecological interactions that occur throughout time and space. We extend the application of distance–decay models to movement patterns of invasion vectors, and, in addition, suggest that broad concepts of distance decay have considerable future application to ecology. As the development and interest within metapopulation and metacommunity dynamics evolves (e.g., Facon and David 2006,

Labonne et al. 2008, Morrissey and de Kerckhove 2009, Rodriguez-Iturbe et al. 2009), the ability to describe the effect of geographic distance between interacting species becomes fundamental to understanding how and why species interact, with what frequency, and the role of spatiotemporal variability across patches of suitable habitat (Hanski and Simberloff 1997, Ovaskainen 2002, Labonne et al. 2008, Planes et al. 2009). Because species interactions may occur throughout multiple physical habitat connections, incorporating geometric networks of species dispersal in conjunction with least-cost routing to determine optimal pathways may provide fundamental advances over previous approaches (e.g., mean pathway distance) with the understanding that additional pathway choices exist, but likely are selected with reduced frequency (Proulx et al. 2005). Alternate uses, such as those of pollination ecology and the development of conservation reserves (e.g., Sala et al. 2002, Lopezarazola-Mikel et al. 2007), rely on multiple spatial linkages between patch habitats and their species, where least-cost routing may provide methodological improvements when determining distance and its effect upon ecological interactions among patches. We present here one of many possible uses of geometric networks and least-cost routing within invasion ecology, and emphasize that the additional ecological applications discussed herein are numerous and may benefit through the application of routing optimization.

ACKNOWLEDGMENTS

We thank H. H. Harvey and two anonymous reviewers whose suggestions greatly improved earlier versions of the manuscript. We also thank S. Walker and D. Jackson, who provided modeling advice. L. Bouvier, A. Boyko, C. Boyko, B. Cudmore, M. Finch, and D. Marson administered social surveys. Funding was provided to D. A. R. Drake through a University of Toronto Fellowship Grant, an Ontario Graduate Scholarship and a Natural Sciences and Engineering Research Council (NSERC) Postgraduate Scholarship. Funding to N. E. Mandrak was provided by Fisheries and Oceans Canada's Center of Expertise for Aquatic Risk Assessment and through an NSERC Canadian Aquatic Invasive Species Network Grant. In-kind and financial support was provided by the Ontario Ministry of Natural Resources and the Ontario Federation of Anglers and Hunters.

LITERATURE CITED

- Akaike, H. 1974. New look at statistical-model identification. *IEEE Transactions on Automatic Control* 19:716–723.
- Baguette, M., S. Petit, and F. Queva. 2000. Population spatial structure and migration of three butterfly species within the same habitat network: consequences for conservation. *Journal of Applied Ecology* 37:100–108.
- Baguette, M., and H. Van Dyck. 2007. Landscape connectivity and animal behaviour: functional grain as a key determinant for dispersal. *Landscape Ecology* 22:1117–1129.
- Bastolla, U., M. A. Fortuna, A. Pascual-Garcia, A. Ferrera, B. Luque, and J. Bascompte. 2009. The architecture of mutualistic networks minimizes competition and increases biodiversity. *Nature* 458:1018–1021.
- Baxter, G. T., and M. D. Stone. 1995. *Fishes of Wyoming*. Wyoming Game and Fish Department, Cheyenne, Wyoming, USA.
- Bell, G., M. J. Lechoqicz, and M. J. Waterway. 2006. The comparative evidence relating to functional and neutral

- interpretations of biological communities. *Ecology* 87:1378–1386.
- Benedict, M. Q., R. S. Levine, W. A. Hawley, and L. P. Lounibos. 2007. Spread of the tiger: global risk of invasion by the mosquito *Aedes albopictus*. *Vector-borne and Zoonotic Diseases* 7:76–85.
- Bergkvist, E., and L. Westin. 2001. Estimation of gravity models by OLS estimation, NLS estimation, Poisson and neural network specifications. Pages 135–155 in M. Gastaldi and A. Reggiani, editors. *New analytical advances in transportation and spatial dynamics*. Ashgate, Burlington, Vermont, USA.
- Bossenbroek, J. M., L. E. Johnson, B. Peters, and D. M. Lodge. 2007. Forecasting the expansion of zebra mussels in the United States. *Conservation Biology* 21:800–810.
- Bossenbroek, J. M., C. E. Kraft, and J. C. Nekola. 2001. Prediction of long-distance dispersal using gravity models: zebra mussel invasion of inland lakes. *Ecological Applications* 11:1778–1788.
- Bothwell, M. L., D. R. Lynch, J. Deniseger, and H. Wright. 2009. Blooms of *Didymosphenia geminata* in rivers on Vancouver Island 1990 to present: a new invasive species. *Journal of Phycology* 45(s1):33.
- Brasly, O., W. Dullaert, and P. Nakari. 2009. The potential of optimization in communal routing problems: case studies from Finland. *Journal of Transportation Geography* 17:484–490.
- Burnham, K. P., and D. R. Anderson. 2002. *Model selection and multimodal inference: a practical information-theoretic approach*. Second edition. Springer Publishers, New York, New York, USA.
- Cameron, A. C., and P. K. Trivedi. 1998. *Regression analysis of count data*. Cambridge University Press, Cambridge, UK.
- Cameron, A. C., and P. K. Trivedi. 2005. *Microeconometrics: methods and applications*. Cambridge University Press, Cambridge, UK.
- Chapin, F. S., III, E. S. Zavaleta, V. T. Eviner, R. L. Naylor, P. M. Vitousek, H. L. Reynolds, D. U. Hooper, S. Lavore, O. E. Sala, S. E. Hobbie, M. C. Mack, and S. Diaz. 2000. Consequences of changing biodiversity. *Nature* 405:234–242.
- Colautti, R. I., I. A. Grigorovich, and J. J. MacIsaac. 2006. Propagule pressure: a null model for biological invasion. *Biological Invasions* 8:1023–1037.
- Cox, E. T. 1978. *Counts and measurements of Ontario Lakes*. Ontario Ministry of Natural Resources, Toronto, Ontario, Canada.
- Crossman, E. J., E. Holm, R. Cholmondeley, and K. Tuininga. 1992. First record for Canada of the rudd, *Scardinius erythrophthalmus*, and notes on the introduced round goby, *Neogobius melanostomus*. *Canadian Field Naturalist* 106: 206–209.
- Dillman, D. A. 1978. *Mail and telephone surveys: the total design method*. John Wiley and Sons, New York, New York, USA.
- Dillman, D. A. 2000. *Mail and internet surveys: the tailored design method*. Second edition. John Wiley and Sons, New York, New York, USA.
- Dzialowski, A. R., W. J. O'Brien, and W. M. Swaffer. 2000. Range expansion and potential dispersal mechanisms of the exotic cladoceran *Daphnia lumholzi*. *Journal of Plankton Research* 22:2205–2223.
- Eritja, R., R. Escosa, J. Lucientes, E. Marques, D. Roiz, and S. Ruiz. 2005. Worldwide invasion of vector mosquitoes: present European distribution and challenges for Spain. *Biological Invasions* 7:1573–1464.
- ESRI [Environmental Systems Research Institute]. 2008. ArcGIS, version 9.3. Environmental Systems Research Institute, Redlands, California, USA.
- Facon, B., and P. David. 2006. Metapopulation dynamics and biological invasions: a spatially explicit model applied to a freshwater snail. *American Naturalist* 168:769–783.
- Fisher, M. R. 1996. Estimating the effect of nonresponse bias on angler surveys. *Transactions of the American Fisheries Society* 125:118–126.
- Fotheringham, A. S., and M. E. O'Kelly. 1989. *Spatial interaction models: formulations and applications*. Kluwer Academic Publishers, Dordrecht, The Netherlands.
- Gonzalez, J. R., G. del Barrio, and B. Duguy. 2008. Assessing functional landscape connectivity for disturbance propagation on regional scales—A cost-surface model approach applied to surface fire spread. *Ecological Modelling* 211:121–141.
- Goodchild, G. A., and J. C. Tilt. 1976. A range extension of *Nocomis micropogon*, the river chub, into eastern Ontario. *Canadian Field-Naturalist* 90:491–492.
- Green, J., and B. J. M. Bohannan. 2006. Spatial scaling of microbial biodiversity. *Trends in Ecology and Evolution* 21: 501–507.
- Hanski, I. A., and D. Simberloff. 1997. The metapopulation approach, its history, conceptual domain, and application to conservation. Pages 5–26 in I. A. Hanski and M. E. Gilpin, editors. *Metapopulation biology: ecology, genetics, and evolution*. Academic Press, San Diego, California, USA.
- Herborg, L.-M., C. L. Jerde, D. M. Lodge, G. M. Ruiz, and H. J. MacIsaac. 2007. Predicting invasion risk using measures of introduction effort and environmental niche models. *Ecological Applications* 17:663–674.
- Hilbe, J. M. 2007. *Negative binomial regression*. Cambridge University Press, Cambridge, UK.
- Jacobs, M. J., and H. J. MacIsaac. 2007. Fouling of fishing line by the waterflea *Cercopagis pengoi*: a mechanism of human-mediated dispersal of zooplankton? *Hydrobiologia* 583:119–126.
- Jacobs, M. J., and H. J. MacIsaac. 2009. Modelling spread of the invasive macrophyte *Cabomba caroliniana*. *Freshwater Biology* 54:296–305.
- 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., A. Ricciardi, and J. T. Carlton. 2001. Overland dispersal of aquatic invasive species: a risk assessment of transient recreational boating. *Ecological Applications* 11: 1789–1799.
- Johnstone, I. M., B. T. Coffey, and C. Howard-Williams. 1985. The role of recreational boat traffic in interlake dispersal of macrophytes: a New Zealand case study. *Journal of Environmental Management* 20:263–279.
- Kaplan, D. M., L. W. Botsford, M. R. O'Farrell, S. D. Gaines, and S. Jorgensen. 2009. Model-based assessment of persistence in proposed marine protected area designs. *Ecological Applications* 19:433–448.
- Keller, R. P., A. N. Cox, C. Van Loon, D. M. Lodge, L.-M. Herborg, and J. Rothlisberger. 2007. From bait shops to the forest floor: earthworm use and disposal by anglers. *American Midland Naturalist* 158:321–328.
- Kerans, B. L., M. E. Dybdhal, M. M. Gangloff, and J. E. Jannot. 2005. *Potamopyrgus antipodarum*: distribution, density, and effects on native macroinvertebrate assemblages in the Greater Yellowstone Ecosystem. *Journal of the North American Benthological Society* 24:123–138.
- Kerr, S. J., C. S. Brousseau, and M. Muschett. 2005. Invasive aquatic species in Ontario: a review and analysis of potential pathways for introduction. *Fisheries* 30:21–30.
- Kilroy, C., S. T. Larned, and B. J. F. Biggs. 2009. The non-indigenous diatom *Didymosphenia geminata* alters benthic communities in New Zealand rivers. *Freshwater Biology* 54: 1990–2002.
- Kolar, C. S., and D. M. Lodge. 2002. Ecological predictions and risk assessment for alien fishes in North America. *Science* 298:1233–1236.
- Kolbl, R., and D. Helbing. 2003. Energy laws in human travel behaviour. *New Journal of Physics* 5:1–12.
- Labonne, J., V. Ravigne, B. Parisi, and C. Gaucherel. 2008. Linking dendritic network structures to population demogenetics: the downside of connectivity. *Oikos* 117:1479–1490.
- Lawler, E. L., J. K. Lenstra, A. H. G. Rinnooy Kan, and D. B. Shmoys. 1985. The traveling salesman problem: a guided tour

- of combinatorial optimization. John Wiley and Sons, Chichester, UK.
- Leung, B., J. M. Bossenbroek, and D. M. Lodge. 2006. Boats, pathways, and aquatic biological invasions: estimating dispersal potential with gravity models. *Biological Invasions* 8:241–254.
- Litvak, M. K., and N. E. Mandrak. 1993. Ecology of freshwater baitfish use in Canada and the United States. *Fisheries* 18:6–13.
- Litvak, M. K., and N. E. Mandrak. 1999. Baitfish trade as a vector of aquatic introductions. Pages 163–180 in R. Claudi and J. J. Leach, editors. *Nonindigenous freshwater organisms: vectors, biology, and impacts*. Lewis Publishers, Boca Raton, Florida, USA.
- Lockwood, J. L., P. Cassey, and T. Blackburn. 2005. The role of propagule pressure in explaining species invasions. *Trends in Ecology and Evolution* 20:223–228.
- Lopezariza-Mikel, M. E., R. B. Hayes, M. R. Whalley, and J. Memmott. 2007. The impact of an alien plant on a native plant–pollinator network: an experimental approach. *Ecology Letters* 10:539–550.
- Ludwig, H. R., Jr., and J. A. Leitch. 1996. Interbasin transfer of aquatic biota via anglers bait buckets. *Fisheries* 21:14–18.
- MacIsaac, H. J., J. V. M. Borbely, J. R. Muirhead, and P. A. Graniero. 2004. Backcasting and forecasting biological invasions of inland lakes. *Ecological Applications* 14:773–783.
- Mack, R. N., D. Simberloff, W. M. Lonsdale, H. Evans, M. Clout, and F. A. Bazzaz. 2000. Biotic invasions: causes, epidemiology, global consequences, and control. *Ecological Applications* 10:689–710.
- Mandrak, N. E., and E. J. Crossman. 1992. A checklist of Ontario freshwater fishes annotated with distribution maps. Royal Ontario Life Science Museum, Toronto, Ontario, Canada.
- Marchetti, M. P., P. B. Moyle, and R. Levine. 2004. Invasive species profiling? Exploring the characteristics of non-native fishes across invasion stages in California. *Freshwater Biology* 49:646–661.
- Morlon, H., G. Chuyong, R. Condit, S. Hubbell, D. Kenfack, D. Thomas, R. Valencia, and J. L. Green. 2008. A general framework for the distance-decay of similarity in ecological communities. *Ecology Letters* 11:904–917.
- Morrissey, M. B., and D. T. de Kerckhove. 2009. The maintenance of genetic variation due to asymmetric gene flow in dendritic metapopulations. *American Naturalist* 174:875–899.
- Muirhead, J. R., and H. J. MacIsaac. 2005. Development of inland lakes as hubs in an invasion network. *Journal of Applied Ecology* 42:80–90.
- Olden, J. D., D. A. Jackson, and P. R. Peres-Neto. 2002. Predictive models of fish species distributions: a note on proper validation and chance predictions. *Transactions of the American Fisheries Society* 131:329–336.
- Ovaskainen, O. 2002. The effective size of a metapopulation living in a heterogeneous patch network. *American Naturalist* 160:612–628.
- Perez-del-Olmo, A., M. Fernandez, J. A. Raga, A. Kostadinova, and S. Morand. 2009. Not everything is everywhere: the distance decay of similarity in a marine host–parasite system. *Journal of Biogeography* 36:200–209.
- Planes, S., G. P. Jones, and S. R. Thorrold. 2009. Larval dispersal connects fish populations in a network of marine protected areas. *Proceedings of the National Academy of Sciences USA* 106:5693–5697.
- Potapov, A., J. R. Muirhead, S. R. Lele, and M. A. Lewis. *In press*. Stochastic gravity models for modeling lake invasions. *Ecological Modelling*.
- Proulx, S. R., D. E. L. Promislow, and P. C. Phillips. 2005. Network thinking in ecology and evolution. *Trends in Ecology and Evolution* 20:345–353.
- R Development Core Team. 2008. R: a language and environment for statistical computing. Version 2.8.0. R Foundation for Statistical Computing, Vienna, Austria. (<http://www.R-project.org>)
- Ricciardi, A., and J. B. Rasmussen. 1998. Predicting the identity and impact of future biological invaders: a priority for aquatic resource management. *Canadian Journal of Fisheries and Aquatic Sciences* 55:1759–1765.
- Ricciardi, A., and J. B. Rasmussen. 1999. Extinction rates of North American freshwater fauna. *Conservation Biology* 13:1220–1222.
- Rodriguez-Iturbe, I., R. Muneeppeerakul, E. Bertuzzo, S. A. Levin, and A. Rinaldo. 2009. River networks as ecological corridors: a complex systems perspective for integrating hydrologic, geomorphologic, and ecological dynamics. *Water Resources Research* 45:W01413.
- Ruesink, J. 2005. Global analysis of factors affecting the outcome of freshwater fish introductions. *Conservation Biology* 19:1883–1893.
- Ruiz, G. M., and J. T. Carlton. 2003. *Invasive species: vectors and management strategies*. Island Press, Washington, D.C., USA.
- Sala, E., O. Aburto-Oropeza, G. Paredes, I. Parra, and J. C. Barrera. 2002. A general model for designing networks of marine reserves. *Science* 298:1991–1993.
- 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.
- Simberloff, D. 2003. How much information on population biology is needed to manage introduced species? *Conservation Biology* 17:83–92.
- Thomas, R. W., and R. J. Hugget. 1980. *Modeling in geography*. Barnes and Noble Books, Totowa, New Jersey, USA.
- Vander Zanden, M. J., and J. D. Olden. 2008. A management framework for preventing the secondary spread of aquatic invasive species. *Canadian Journal of Fisheries and Aquatic Sciences* 65:1512–1522.
- Wilson, A. G. 1967. A statistical theory of spatial distribution models. *Transportation Research* 1:253–269.
- Wilson, A. G. 1998. Land-use/transport interaction models. Past and future. *Journal of Transport Economics and Policy* 32:3–26.
- Wilson, J. R. U., E. E. Dormontt, P. J. Prentis, A. J. Lowe, and D. M. Richardson. 2009. Something in the way you move: dispersal pathways affect invasion success. *Trends in Ecology and Evolution* 24:136–144.
- Wood, B. M., and Z. J. Wood. 2006. Energetically optimal travel across terrain: visualizations and a new metric of geographic distance with anthropological applications. Pages 6000–6007 in R. F. Erbacher, J. C. Roberts, M. T. Grohn, and K. Borner, editors. *Visualisation and data analysis. Proceedings of the Society of Photo-Optical Instrumentation Engineers (SPIE)*, San Jose, California, USA. SPIE Publications, Bellingham, Washington, USA.
- Zar, J. H. 1999. *Biostatistical analysis*. Prentice-Hall Publishers, New Jersey, USA.

APPENDIX

List of social survey questions used to address movement potential of anglers in Ontario, Canada (*Ecological Archives* A020-087-A1).