Biol Invasions (2011) 13:2445–2458
DOI 10.1007/s10530-011-0070-3

ORIGINAL PAPER

Evaluation of stochastic gravity model selection for use in estimating non-indigenous species dispersal and establishment

Jim R. Muirhead · Hugh J. MacIsaac

Received: 23 February 2010 / Accepted: 31 January 2011 / Published online: 17 August 2011

Abstract Predicting dispersal of nonindigenous species (NIS) is an essential component of risk analysis and management, as preventative measures are most readily applied at this stage of the invasion sequence. Gravity models provide one of the most useful techniques available to model dispersal of nonindigenous invasive species (NIS) across heterogeneous landscapes, as these models are able to capture transport patterns of recreational boaters who are dominant vectors responsible for aquatic NIS dispersal. Despite the widespread use of gravity models in forecasting biological invasions, different classes of gravity models have not been evaluated regarding their comparative ability to capture recreational transport patterns and subsequent use in predicting NIS establishment. Here we evaluate model selection between unconstrained, total-flow-constrained, production-constrained and doubly-constrained stochastic gravity models to assess dispersal of the spiny waterflea *Bythotrephes* between Ontario lakes. Differences between the models relate to the amount of data required and constraints under which calculations of source/destination interactions are made. For each class of gravity model, we then estimated the probability of a lake having established *Bythotrephes* populations by modeling the relationship between empirical presence/absence data and inbound recreational traffic (i.e. propagule pressure) via boosted regression. The unconstrained gravity model provided the best fit to observed traffic patterns of recreational boaters. However, when output from the gravity models was used to predict *Bythotrephes* establishment, the doubly-constrained gravity model provided the strongest relationship between inbound recreational traffic and observed *Bythotrephes* presence/absence, followed by the production-constrained model. Our results indicate production-constrained gravity models offer an acceptable balance between modeling recreational boater traffic and their utility to estimate establishment probabilities.

Keywords Nonindigenous species · Invasive species · Biological invasion · Stochastic gravity model

Present Address:
J. R. Muirhead (✉)
Department of Biological Sciences, University of Alberta, Edmonton, AB T6G 3P9, Canada
e-mail: muirheadj@si.edu

H. J. MacIsaac
Great Lakes Institute for Environmental Research, University of Windsor, Windsor, ON N9B 3P4, Canada

Introduction

The global spread of nonindigenous species (NIS) has become a leading environmental issue owing both to the frequency of species introductions and to their profound consequences in some invaded ecosystems (Sala et al. 2000; Davis 2003). Slowing the rate of new invasions requires development of risk assessment and management policies. The first step of risk assessment models is the determination of where NIS are expected to colonize and spread. A major challenge in developing these forecasts lies in tradeoffs between selection of competing risk assessment models, the amount of information required to parameterize these models, and their predictive accuracy. Once accurate forecasts have been developed based on an optimal modeling framework, risk management policies may be developed to reduce rates of spread and establishment. Specifically, forecasts of NIS dispersal that are developed using a process-based or mechanistic approach of NIS transport may be used in risk management to identify putative control measures.

Stage-based invasion models are an important advancement to efforts to predict successful invasions because they logically examine factors affecting success at each step of the invasion process (see Richardson et al. 2000; Kolar and Lodge 2002; Vander Zanden and Olden 2008). Stage-based models begin with quantifying the introduction effort or 'propagule pressure' (i.e. the number of introduction events, and the number and quality of individuals per event) from source populations (Richardson et al. 2000; Kolar and Lodge 2001; Colautti et al. 2006; Lockwood et al. 2009). Upon arrival at a new site, propagules of a NIS must tolerate or exploit ambient environmental conditions (Rouget and Richardson 2003; Hayes and Barry 2008; Herborg et al. 2007; Melbourne et al. 2007). Finally, the surviving propagules must integrate into the community, with possible positive, negative or neutral effects by native residents (Fridley et al. 2007).

There exists considerable evidence that increased propagule pressure is a key determinant of invasion success (Veltman et al. 1996; Forsyth and Duncan 2001; Lockwood et al. 2005; Von Holle and Simberloff 2005; Colautti et al. 2006). Many of these studies derive from biological control, fisheries, bird or mammal introductions (e.g. Jeschke and Strayer 2005). However, the relationship between propagule pressure and the likelihood of a successful invasion is not always straightforward. For example, Lockwood et al. (2005) argued that the location of the introduction event and the composition of the recipient community may interact to alter the relationship between propagule pressure and invasion success.

Gravity models provide an ideal framework to assess human-mediated dispersal of NIS, as they can be used to model movement patterns of vectors and associated propagules of NIS between discrete systems. Initially developed for use in describing immigration patterns (Zipf 1946), flows of economic goods (Linneman 1966), and optimal placement of retail services (Huff 1959), gravity models describe the flow of information between spatially-discrete origins and destinations. The flow from origin to destination is affected by the distance between them, by the amount of outflow and extrinsic 'propulsiveness' from each origin, and by the amount of inflow and 'attractiveness' of different destinations.

Gravity models can be distinguished into five main classes depending on the amount of information available and the interaction flow constraints assumed between origins and destinations. Unconstrained models require the least amount of data to construct, only measures of destination attractiveness and distance between origins and destinations. Total-flow-constrained models require data on only total flow within the system, measures of origin propulsiveness, and destination attractiveness (Haynes and Fotheringham 1984). Production-constrained models require only moderate amounts of effort to collect data, including measures of outflow from origins and measures of destination attractiveness. Production-constrained models have been used to forecast dispersal of NIS into unknown destinations (Leung et al. 2004, 2006; Bossenbroek et al. 2001, 2007). Doubly-constrained gravity models require information about both outflows from origins and inflows to destinations and have also been used to forecast spread of aquatic invaders (Schneider et al. 1998; MacIsaac et al. 2004). Finally, attraction-constrained gravity models provide the same information as doubly-constrained models with respect to inflows to each destination and we do not deal with this class of model further.

Although the models used to forecast NIS dispersal considered the flow between sources and destinations as a deterministic value, stochastic versions of gravity models treat the flow between sources and destinations as a random variable described by a statistical distribution. That is, the data collected on origin–destination flows is treated as a single realization of an underlying stochastic process governing trip distribution. Stochastic gravity models have been used frequently in economic and transportation studies (e.g. Flowerdew and Aitkin 1982, Anas 1983), but their use is not widespread in the ecological literature (although see Potapov et al. 2010).

Since it was reported in the Laurentian Great Lakes in the early 1980's, *Bythotrephes longimanus* has spread to more than 150 inland lakes throughout the province of Ontario (Johannsson et al. 1991; MacIsaac et al. 2004; N. Yan, pers. comm.). In addition to advective dispersal through connected waterways, *Bythotrephes* dispersal between lakes is facilitated by human-mediated transport associated with recreational boating and fishing (Boudreau and Yan 2004; MacIsaac et al. 2004; Branstrator et al. 2006). One of the key life-history traits that may facilitate *Bythotrephes'* rapid range expansion is its production of diapausing or resting eggs. These sexually-produced, diploid eggs remain viable after passage through fish gastrointestinal tracts (Jarnagin et al. 2000), and may survive overland transport (Ketelaars and Gille 1994). If the fishing gear is subsequently used on another lake without cleaning, dried masses of females and resting eggs may fall off the line; viable resting eggs could then hatch and cause a new invasion. *Bythotrephes* may also be introduced to new lakes in transported water, recreational or scientific gear, or in transplanted fish.

In this study, we evaluate model selection between unconstrained, total-flow, production- and doubly-constrained stochastic gravity models to describe the pattern of recreational movement among Ontario lakes as a proxy for *Bythotrephes* dispersal. We then relate inbound recreational traffic as a measure of propagule pressure to observed *Bythotrephes* occurrences via boosted regression establishment models to illustrate tradeoffs in gravity model choice affecting the ability to fit recreational data versus predicting invasion outcomes. Finally, we make recommendations about gravity model choice based on management goals and the effort required to collect the necessary data.

Materials and methods

To implement a two-stage conceptual model of the invasion sequence, we have modeled *Bythotrephes* dispersal as a submodel nested within a model estimating the probability of establishment for a given lake. First, we modeled recreational traffic of trailered boats travelling among lakes with different classes of gravity models as a proxy for *Bythotrephes* dispersal. We then estimated the mean number of trips predicted to arrive at a series of lakes within a single boating season as a measure of propagule pressure, and modeled the relationship between propagule pressure and *Bythotrephes* presence/absence with boosted regression trees to determine probability of establishment.

Data collection

We mailed 10,000 surveys in July 2004 to owners of fishing licenses registered with the Ontario Ministry of Natural Resources to assess movement patterns of recreationalists within the province. Approximately 218 surveys were sent to households in each of 46 zones based on the first two digits of their postal code, thereby minimizing bias due to differences in population density across the province. Overall response rate for returned surveys was 7.8%. From the surveys, we were able to collect data on 1,576 pairwise recreational trips made between 15 May 2004 and 7 September 2004 between 49 invaded origin lakes i and 191 invaded and non-invaded destination lakes j. Since we were interested in transient recreationalist movement, we excluded data where recreationalists only visited a single lake and recorded trips only where the origin and destination lakes were different ($i \neq j$).

Dispersal models

Stochastic gravity models

To model recreational boat movement between lake pairs i and j, the expected (i.e. mean) number of trips can be built using different classes of gravity model depending on model constraints and data requirements for model parameterization and extrapolation of pairwise trips from invaded origins to additional destinations. Although it is not a condition of gravity model use for modeling aquatic NIS dispersal, we specifically used lakes pairs to examine the effect of increasing model constraints on estimated boater movement. Unlike Random Utility Models that address individual trip behaviour to competing destinations (e.g. Bockstael et al. 1989), gravity models are zonal aggregate models and are suitable for comparing relative numbers of trips between origins and destinations over a single boating season. Hence, we are not considering individual multi-trip movement such as trips from their house to each of the lakes, although additional models describing each leg of a multi-trip movement may be used (Bossenbroek et al. 2001; Leung et al. 2004.)

We modeled the number of recreational trips between origin and destination pairs as a random (stochastic) variable with error, in contrast to deterministic gravity models used to model biological invasions that have a fixed number of trips calculated as a function of model inputs (e.g. Bossenbroek et al. 2001; Leung et al. 2006). For the gravity models, we assume that the mean number of *Bythotrephes* individuals transported per trip is the same for each trip, and that propagule pressure of *Bythotrephes* scales with the mean number of trips arriving at the lake. In addition, we are making the simplifying assumption that the mean number of *Bythotrephes* individuals per trip is the same regardless of its population size in the origin lake.

Gravity model classes

A generalized gravity model describing the number of trips between lakes i and j can be modeled as a function of observed trips leaving lake i (O_i) or arriving at lake j (D_j), measures of destination "attractiveness" (w_j), and road distance between lakes i and j ($d_{ij}^{-\beta_1}$):

$$\hat{\mu}_{ij} = cO_iD_jw_jd_{ij}^{-\beta_1}. \tag{1}$$

Here, c can be one or more fitted constants or "balancing factors" to ensure that the total number of expected trips, leaving lake i or arriving at lake j are constrained to equal to observed numbers of trips for different classes of gravity model (discussed below). Destination attractiveness is described by $w_j = e^{\beta_2}\log_{10}(a_j + 1)$, in which a_j is the destination lake area (hectares). Parameters β_1 and β_2 are fitted during maximum likelihood estimation, and e^{β_2} ensures that lake attractivity is constrained from 0 to ∞.

Gravity model classes

Depending on the type of information available on outbound or inbound traffic, and various constraints on under which to estimate traffic between origins and destinations, the general form of the gravity model may be modified into different classes.

The unconstrained gravity model requires the least amount of information to estimate pairwise traffic—only an external measure of destination lake attractiveness, w_j, and the expected number of pairwise trips is not under constraint. Here, the generalized model is reduced to:

$$\hat{\mu}_{ij} = w_jd_{ij}^{-\beta_1} \tag{2}$$

For this class of gravity model, extrapolation to other destinations requires data on only lake area and the road distance between origins and new destinations.

The total-flow-constrained model requires information on the total number of pairwise trips in the system and a measure of destination attractiveness. Expected flow between origins and destinations for the total-flow-constrained model is calculated as:

$$\hat{\mu}_{ij} = cw_jd_{ij}^{-\beta_1} \tag{3}$$

where c is calculated as $c = T\Big/\sum_i\sum_j w_jd_{ij}^{-\beta_1}$ to ensure that the total of predicted flows, $\sum_i\sum_j\hat{\mu}_{ij}$, is equal to observed total flow, T (Haynes and Fotheringham 1984). Destination attractiveness and road distance is calculated as before in Eq. 1. In this formulation of total-flow-constrained gravity model, the term for origin "propulsiveness" was not included (e.g. Haynes and Fotheringham 1984). Similar to the unconstrained model, extrapolated number of trips to lakes not in the data set based on the total-flow constrained gravity model requires measures of attractiveness and distance from origins.

Production-constrained gravity models require additional information about the outflows from each origin. Flows between origins and destinations are calculated as:

$$\hat{\mu}_{ij} = c_iO_iw_jd_{ij}^{-\beta_1} \tag{4}$$

where O_i is the observed number of trips leaving each origin i and c_i is a balancing factor for each origin,

$c_i = \left[\sum_j w_j d_{ij}^{-\beta_1}\right]^{-1}$, ensuring that total predicted outflow from each origin is equal to observed outflow $\left(\sum_j \hat{\mu}_{ij} = O_i\right)$ (Haynes and Fotheringham 1984). Propagule pressure forecast to additional lakes with production-constrained models again requires measures of destination attractiveness and road distance between origin and destinations.

In the doubly-constrained gravity model, predicted total outflows for each origin and total inflows for each destination are constrained to match observed values. Here pairwise flow is formulated as:

$$\hat{\mu}_{ij} = c_i c_j O_i D_j d_{ij}^{-\beta_1} \tag{5}$$

where D_j is the inflow to destination j, and other terms are as previously defined. The parameters c_i and c_j are balancing factors, $c_i = \left[\sum_j c_j D_j d_{ij}^{-\beta_1}\right]^{-1}$ and $c_j = \left[\sum_i c_i D_i d_{ij}^{-\beta_1}\right]^{-1}$, added to the model to ensure that the sum of the interaction flow, $\hat{\mu}_{ij}$, for each origin is equal to the total outflow, O_i, and likewise, the sum of the interaction flow for each destination is equal to the total inflow for each destination, D_j (Haynes and Fotheringham 1984).

$$\Pr(Y_{ij}|\hat{\mu}_{ij}, \omega_{ij}, k) = \begin{cases} \omega_{ij} + (1-\omega_{ij})\left(1+\hat{\mu}_{ij}/k\right)^{-k}, & Y_{ij} = 0 \\ (1-\omega_{ij})\frac{\Gamma(Y_{ij}+1/k)}{\Gamma(1+Y_{ij})\Gamma(1/k)}\left[k\hat{\mu}_{ij}\right]^{Y_{ij}}\left[1+k\hat{\mu}_{ij}\right]^{-(Y_{ij}+1/k)}, & Y_{ij} = 1, 2, 3, \ldots, n \end{cases} \tag{6}$$

Unlike the three previous classes of gravity models, the doubly-constrained model does not require measures of destination attractiveness, w_j. It is, however, the most data intensive, as observed outflows from each origin and inflows to each destination are required for model parameterization. Extrapolation to other destinations is not possible because although information may be collected on inflows to new destinations, D_j and road distances d_{ij}, pairwise flows requires re-calculation of fitted parameters c_i and c_j.

For the purpose of model comparison in this analysis, because the doubly-constrained gravity model is limited to predicting inflow to destinations for which there are observations, the total-flow and production-constrained models were constrained to the same set of destinations. Here, we make a simplifying assumption that we had complete data on the number of trips in the system in order to calculate recreational traffic under gravity model constraints.

Zero-inflated negative binomial distribution

For each class of gravity model, the expected number of pairwise trips between lakes, $\hat{\mu}_{ij}$ is modeled as a random variable described by a zero-inflated negative binomial distribution (ZINB). A ZINB distribution is a frequently used to model situations where count data is overdispersed and may contain an excess of zeros that are either structural in nature or arise due to sparse sampling effort. The observed number of pairwise trips (Y_{ij}) is expected to follow a ZINB distribution with a probability mass function:

where $\hat{\mu}_{ij}$ is the expected traffic between lakes for each of the gravity models (Eqs. 2–5) and is a function of fitted parameters, β and γ (Eqs. 1, 6a). In this form, $\hat{\mu}_{ij}$ and k are the mean and dispersion parameter of the negative binomial (NB) distribution, and ω_{ij} is a parameter that describes the probability that only $Y_{ij} = 0$ can occur (i.e. zero-inflation), and $(1-\omega_{ij})$ the probability that $Y_{ij} \sim NB(\hat{\mu}_{ij}, k)$ is occurring. The probability of zero-inflation was calculated from the logistic equation:

$$\log\left(\frac{\omega_{ij}}{1-\omega_{ij}}\right) = \gamma_0 + \gamma_1 \log_{10}(a_{.j}+1) \tag{6a}$$

where $a_{.j}$ is destination lake area (in ha) and γ_0 and γ_1 are fitted parameters (Jansakul 2005).

Estimates for the fitted parameters β, γ and k were solved using maximum likelihood with an Expectation-Maximization algorithm (Dempster et al. 1977).

Model fitting comparison and confidence limits

Given the same baseline data set on the number of pairwise trips, the four classes of gravity models were evaluated using corrected Akaike Information Criteria (AICc). Because maximum likelihood estimation assuming ZINB requires estimating the components for zero-inflation (ω), the gravity model (β) and dispersion (k) simultaneously, the AICc was calculated as $-2(ll_m + ll_z) + 2p + (2p(p+1)/(n-p-1))$ where ll_m and ll_z are the log-likelihoods for the weighted negative binomial (Eq. 6 for $y > 0$) and zero-inflation (Eq. 6a), respectively; p is the number of fitted parameters, and n is the sample size. Confidence limits for the fitted parameters were estimated by bias-corrected bootstrapping data that were resampled 1,000 times (Efron and Tibshirani 1986).

Establishment model

Prior to establishment model development and validation, destination lakes in the data set were randomly divided 50:50 into model training and testing subsets. Based on the training data only, the probability of specific lakes having established populations is determined by:

$$P(X = 1) = f\left(\hat{\mu}_{ij}\right), \tag{7}$$

where the functional relationship between observed *Bythotrephes* presence/absence (1,0) and the expected number of inbound trips, $\hat{\mu}_{ij} = \sum_i (1 - \omega_{ij})\hat{\mu}_{ij}$, is determined by boosted regression trees. Boosted regression trees use an iterative approach and do not make any parametric assumptions about the relationships between predictor variables and species occurrence. They have been shown to be strong performers in predicting species occurrence and abundance relative to other methods (Elith et al. 2006). To avoid over-fitting the models, the optimum number of iterations for each model was based on a stopping-rule at which there was zero improvement (Ridgeway, G. 2007. gbm: Generalized Boosted Regression Models. R package).

Finally, as boosted regression is a non-parametric procedure, and maximum likelihood estimation is not used to optimize the overall decision tree structure as it would be for used in logistic regression, we calculated 95% prediction intervals for probability of establishment by incorporating variability in the fitted parameters from the gravity model. Here, prediction intervals were estimated by fitting two additional boosted regression trees to estimates of the minimum and maximum number of inbound trips determined by 95% confidence limits of the fitted parameters of the gravity models and raw data. These prediction intervals allowed us to assess how variability in the fit of gravity models to observed recreational traffic influences the predictive ability of establishment models based on the different gravity model classes.

Evaluating establishment predictions

The ability of the establishment models based on each gravity model class to predict invasions were validated by comparing each model's probability of occurrence derived from boosted regression trees to observed *Bythotrephes* presence/absence data using the testing data subset. Establishment models were evaluated using Area under the Receiver Operating Characteristic Curve (AUC) as well as measures of concordance between model predictions and observations in 2 $\times$ 2 confusion tables. The AUC provides a useful metric to validate whether the model is able to detect a true signal (hit rate) from noise (false alarm) across a range of probability thresholds. The AUC metric has been widely used in evaluating species distribution models (e.g. McPherson et al. 2004; Liu et al. 2005; Allouche et al. 2006; Elith et al. 2006), but should be used with caution (Lobo et al. 2008). In addition, a probability threshold calculated from the proportion of lakes invaded by *Bythotrephes* in the testing data set was used to classify lakes to be predicted as invaded if the estimated probability of establishment is greater or equal to this threshold ($\tau = 0.252$). Hit rates, false alarm rates, percent correct classifications and Cohen's kappa (κ) were used to assess concordance between model predictions and observed presence/absence based on the threshold probability for establishment.

Results

Comparison of gravity model fits to recreational traffic

Stochastic gravity models were greatly influenced by the constraints under which the pairwise number of

Table 1 Summary of goodness of gravity model fit (corrected Akaike Information Criteria, AICc) and parameter estimates with bootstrapped confidence intervals

Gravity model class	AICc	k	β_1 (distance)	β_2 (lake area, attractivity)	γ_0 (zero-inflation)	γ_1 (zero-inflation)
Unconstrained	1,2870	1.22	−0.05	1.51	2.71	−0.08
		[1.200, 1.505]	[−0.061, 0.123]	[1.222, 2.168]	[2.647, 3.066]	[−0.083, −0.035]
Total-flow constrained	12,976	1.13	−0.06	−1.61	2.71	−0.08
		[1.076, 1.428]	[−0.060, 0.115]	[−1.597, -0.776]	[2.646, 2.955]	[−0.089, −0.041]
Production-constrained	15,141	0.23	−0.01	−0.11	2.71	−0.08
		[0.186, 0.267]	[−0.080, 0.126]	[−0.211, 0.236]	[2.643, 3.132]	[−0.090, −0.042]
Doubly-constrained	17,557	0.06	−0.10	NA	2.70	−0.08
		[0.047, 0.066]	[−0.234, −0.055]		[2.638, 3.010]	[−0.093, −0.029]

Lake area was not considered in the doubly-constrained gravity model. Parameter k is the measure of dispersion for the ZINB (Eq. 5), β_1 corresponds to the distance decay component of gravity models (Eqs. 1–4), β_2 corresponds to lake area in destination attractivity (Eq. 1) and γ_0 and γ_1 are solved during estimation of zero-inflation (Eq. 6a)

trips was estimated. Of the four classes of models, the unconstrained model provided the best fit to observed trip data as indicated by lowest AICc, followed by the total-flow constrained, production-constrained and finally doubly-constrained model (Table 1). Here, large differences in AICc between the gravity model types (>100) imply significant differences between fits for each of the gravity model classes because there is one fewer parameter needed to be estimated in the doubly-constrained relative to the other types of gravity models.

The use of the same baseline data set across gravity model classes allows us to compare values of the fitted parameters across models. Differences in the fitted parameters associated with lake attractiveness, the effect of distance on the expected number of trips, and the dispersion parameter followed trends related to the level of gravity model constraint. The relationship between lake area and destination attractiveness decreased in strength as model constraints increased. For the unconstrained gravity model, lake area was more than five times as influential in determining lake attractiveness (w_j) than the production-constrained model as measured by the magnitude of the fitted parameter β_2 (Table 1). In contrast, the strength of the relationship between gravity flow and distance between origins and destinations (parameter β_1) was greatest for the doubly-constrained model and negligible for the other gravity model classes (Table 1). Next, as the level of constraint increased, there was a marked decrease in the level of dispersion used to describe the ZINB. For all four gravity model classes, the dispersion parameter k was significantly different from 0 as indicated by the 95% confidence limits (Table 1). The use of a ZINB over other commonly used trip distributions like Poisson is therefore justified for this dataset. Finally, in terms of the parameters fitted in the estimation of zero-inflation, γ_0 and γ_1, differences in their values between models were not significant.

Establishment model

The estimated inbound number of trips for both invaded and noninvaded destinations decreased with increasing constraints across gravity models (Fig. 1). More importantly, differences in estimated inbound number of trips arriving at invaded versus noninvaded lakes varied between the four classes of gravity

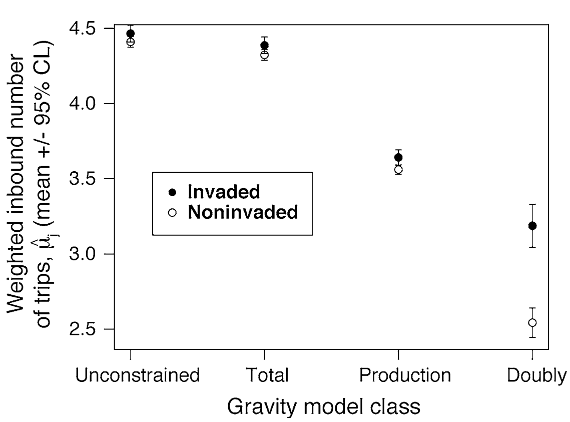

Fig. 1 Estimated weighted inbound number of trips $\hat{\mu}_j = \sum_{i=1}^{n} (1 - \omega_{ij})\hat{\mu}_{ij}$ to invaded and noninvaded lakes for four classes of gravity models

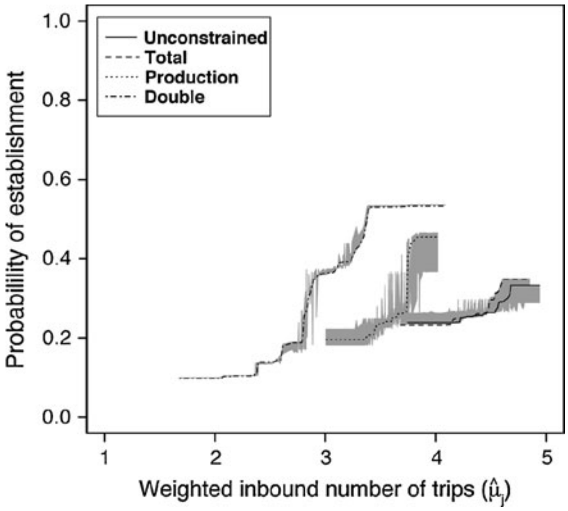

Fig. 2 Estimated probability of establishment as a function of weighted inbound number of trips. *Black lines* indicated boosted regressions fitted to the mean number of trips, and *grey areas* indicate the 95% prediction interval in the probability of establishment based on variability propagated from fitting gravity models. The relationship between probability of establishment and propagule pressure was strongest for the doubly-constrained model and weakest for the unconstrained model

models. Differences were negligible for the unconstrained and total-flow constrained gravity models, but increased significantly for the production- and doubly-constrained models (Fig. 1). As a result, the strength in the relationship between establishment probability based on presence/absence data also varied widely between the four classes of gravity models.

From the boosted regression, the doubly-constrained model exhibited the widest range of establishment probability at lower levels of propagule pressure (Fig. 2). As propagule inflow was estimated from gravity models having fewer constraints, the relationship between establishment probability and propagule pressure became weaker. The unconstrained and total-flow constrained gravity models showed similar a similar, but weak, relationship between establishment probability and the inbound number of trips.

The size of prediction intervals for establishment probability also followed a similar trend and varied according to gravity model type (Fig. 2). Here, variability around establishment probability was lowest for the doubly-constrained gravity model due to smaller confidence limits of fitted parameters propagated from the dispersal models. In particular, variability in establishment probability was apparent only for lakes with moderate levels of propagule pressure ($\hat{\mu}_{ij} = 2.5$–3.5) as estimated from the doubly-constrained model. Prediction intervals around establishment probability estimated from the other three gravity model classes were similar in magnitude. In general, variability around estimated probabilities of establishment was high throughout the ranges of propagule pressure and tended to increase with increasing levels of propagule pressure (Fig. 2).

When predictions from the establishment models were validated with the testing portion of the data, the predictive ability between probability of establishment and observed presence/absence data differed widely between the four classes of gravity model. As the level of constraint increased between the gravity model classes, the area under the receiver operating curves likewise increased, indicating a stronger ability of the model to correctly predict an invasion if it had occurred (hit rate) relative to incorrectly predicting an invasion had occurred when it in fact had not (false alarms) (Table 2). Here, only the production-constrained and doubly-constrained gravity models had significant AUC scores (P(AUC) = 0.011 and $\ll$ 0.001, respectively).

Table 2 Summary statistics for validation of the establishment model based on the testing data subset ($n = 95$)

Model	AUC	P(AUC)	Hit rate	False alarm rate	Percent correct	Cohen's κ
Unconstrained	0.579	0.115	0.577	0.465	0.546	0.088
Total-flow constrained	0.593	0.080	0.577	0.437	0.567	0.113
Production-constrained	0.647	0.011	0.654	0.408	0.608	0.197
Doubly-constrained	0.810	$\ll$0.001	0.731	0.211	0.773	0.473

The baseline probability of invasion based on *Bythotrephes* prevalence in the testing data subset required to calculate hit rates, false alarm rates, percent correct and Cohen's κ was 0.252

Similar trends in predictive power and level of gravity model constraint were observed when a baseline level for the probability of establishment was applied to establishment probability for the testing data subset. Here, lakes were predicted as invaded if the estimated probability of establishment was greater or equal to this baseline probability based on the proportion of lakes invaded by *Bythotrephes* in the testing data set. As the level of constraint increased in the gravity models, hit rates, overall percentage correctly classified and Cohen's κ increased with a corresponding decrease in false alarm rates (Table 2).

The corresponding spatial pattern of establishment probability also differs between the four classes of gravity models (Fig. 3). The wide range of establishment probabilities corresponding to levels of incoming propagule pressure for the doubly-constrained gravity model (Fig. 2) allow for a stronger identification of invasion hot spots (i.e. areas of high relative probability of invasion) when these probabilities are projected spatially (Fig. 3). That is, high levels for establishment probability are concentrated in smaller areas, unlike the lower levels of establishment probability for the unconstrained and total-flow gravity models that were spread out over a greater area. For the doubly-constrained model, pockets of high (> 0.5) establishment probability are focused in areas north of Lakes Erie and Ontario, to the east of Georgian Bay (Lake Huron) as well as two large inland lakes in northern Ontario that were reported invaded. For the production-constrained gravity model, areas of higher establishment probabilities were focused north of Lake Erie, the southern half of Ontario, and north of Lake Superior. Spatial patterns of establishment probability were similar between the unconstrained and total-flow models. Here, estimates of moderate establishment probability (between 0.26 and 0.30) were uniformly distributed throughout most of the province with the exception of pockets surrounding non-invaded lakes in northern Ontario and north of Lake Erie (Fig. 3).

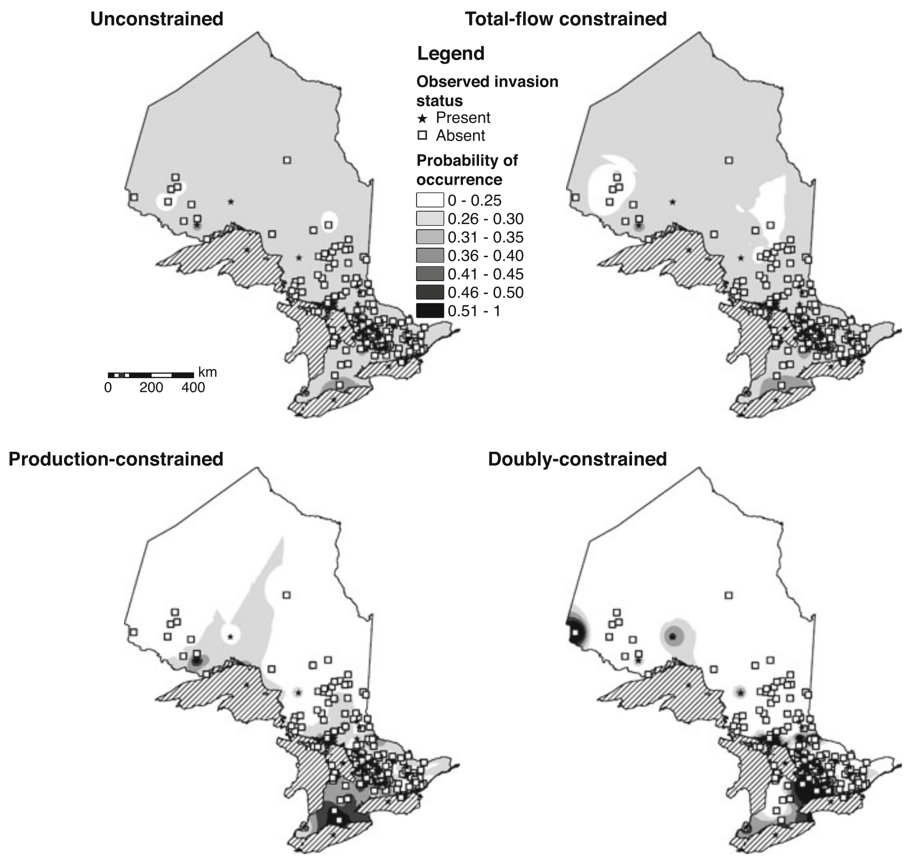

Fig. 3 Spatial pattern of estimated probability of establishment with invaded and non-invaded lakes in Ontario, Canada

Discussion

Estimating the dispersal of a NIS is the significant first step in risk management because preventative measures are most readily applied at the introduction stage of the invasion sequence (Lodge et al. 2006). Heretofore, ecologists have used a variety of gravity models without consideration of the importance of method used, with selection usually based upon the nature of available data. In this study, we examined the tradeoff between the relative ability of four classes of gravity models to fit recreational traffic data and the ability of establishment models based on these gravity models to estimate the probability of establishment by the spiny waterflea, *Bythotrephes*. Simulation results indicate that ecologists should be cognizant that the gravity method used may have strong consequences on forecasted patterns of spread.

Evaluation of the unconstrained, total-flow-, production- and doubly-constrained gravity models revealed differences in each model's ability to fit recreational data as goodness of model fit decreased as more constraints were applied to the gravity models. The unconstrained model was best able to describe the pattern of pairwise recreational traffic between source and destination lakes, with the doubly-constrained model performing the worse. Here, goodness of model fit is a direct result of the optimization process during ML estimation. As more constraints are added to the gravity models, convergence on an optimal solution during ML estimation becomes more difficult (Anderson 1979). These results contrast qualitatively with those from Black and Salter (1975), in which they compare gravity model performance for different model constraints and distance deterrence functions for car and bus journeys to work. In their study, there was a marginal improvement in the correlation between the model trip matrix and survey trip data for doubly-constrained model over the production-constrained and unconstrained models. These qualitative differences between the studies may be the result of differences in methodology of model fitting as well as characteristics of the data set. For example, even though the distance decay parameter became larger (i.e. more negative) with increasing model constraint in both studies (parameter β_1, Table 1), model fit was highly affected by the ZINB dispersion parameter k, whereas they used strictly deterministic forms of gravity models.

Data requirements for the unconstrained gravity model are minimal, however, and this class of stochastic gravity model is recommended for describing pairwise trip distribution between lakes. Since data is required on only extrinsic measures of attraction, such a data set would be relatively easy to acquire from government agencies. Although not commonly used to model dispersal of NIS, unconstrained gravity models have been used extensively in the economic, geographic and social literature since their inception (e.g. Zipf 1946; Anderson 1979; Fotheringham 1981; Johnston 1983). In a terrestrial environment, unconstrained and total-flow constrained models may be used to model trade flows between cities where population sizes may be used as measures of propulsiveness and attractiveness. In a scenario where data on only the total amount of trade volume is available, the total-flow constrained model estimates how the total trade is partitioned among sources and destinations.

Establishment models developed with inbound recreational traffic derived from the gravity models displayed the opposite trend—prediction of establishment success increased as more constraints were imposed on the nested gravity models. Although the unconstrained model was the optimal form when fitted to recreational trip data, this class of gravity model was unable to detect hits any better than at random when invasions occurred. The amount of propagule pressure arriving at non-invaded and invaded lakes account for the lack of fit. Relative to the more constrained forms of gravity models, there were smaller differences between the levels of propagule pressure arriving at non-invaded versus invaded lakes. That is, propagule pressure was not as strong as a discriminator for invasion status in this class of gravity model as per the doubly-constrained model. As a consequence, estimated probabilities of establishment in each lake are more reflective of the proportion of invaded lakes (i.e. baseline probability) than levels of inbound propagule pressure.

Depending on the NIS, the "cost" of not detecting an invasion (miss) is likely greater than the cost of a false alarm, thus the doubly-constrained model would offer the most protective management option. It is also noteworthy that some false alarms may, over the long term, correctly predict invasions even if they are incorrect in the short-term. For example, lakes identified as vulnerable may already be invaded (though at non-perceptible population levels), or they may be the most susceptible to future invasion. Thus, if longer time scales are considered, many of the lakes incorrectly predicted to be invaded by current biological surveys may eventually become invaded and properly classified.

However, the doubly-constrained gravity model is also the most data intensive, since observations are required for both recreational inflows and outflows to each lake in the system. Here, such data is usually not readily available from previous studies or government agencies and additional costs and efforts are required for collection. One consequence of choosing the doubly-constrained gravity model, therefore, is that a decision must be made between minimizing costs either to collect the larger amounts of data relative to other gravity model types or the costs of missing an invasion. Such a decision could be addressed by optimal control models described in Leung et al. (2002).

The third alternative lies in a compromise between the use of gravity models to model recreational traffic and, as a sub-model, in estimating the probability of establishment. Here, we recommend the use of production-constrained gravity models if we consider that predicting invasions (i.e. estimating the probability of establishment) is more useful in NIS management than strictly modeling recreational traffic. In addition, production-constrained gravity models (e.g. Bossenbroek et al. 2001, 2007; Leung et al. 2006) provide a balance between predictive ability as measured by the AUC or Cohen's κ and the amount of information required to populate the model, thus offering important advantages over data-intensive, doubly-constrained models (e.g. Schneider et al. 1998; MacIsaac et al. 2004). The data required to produce a production-constrained model is often readily available from government agencies or non-governmental organizations, thereby reducing both cost and time in development of new models. For example, the US Travel Data Center coordinates surveys on US tourist traffic including point-of-origin, length of duration and average amount of money spent on the trip. Preliminary risk assessment models may be developed by combining specifics on the distribution of NIS that may be transported by human-mediated mechanisms, and patterns of human recreational movement.

In this study, we based our comparison of gravity model classes assuming we had complete data on the number of trips in the system and excluded unsampled lakes from the analysis. Unsampled lakes may serve as propagule sources if invaded or additional destinations for recreational traffic, and as such, the ability to predict invasion status may be influenced by the proportion of lakes sampled. Theoretical simulations by Leung and Delaney (2006) in which they compared various approaches to dealing with unsampled data suggest that sampling a small proportion of sites tends to create a bias in underestimating the probability of establishment for given levels of propagule pressure when assuming complete data on recreational trips. They recommended an approach (MCMSAM) where they combine the observed invasion status for sampled sites and Monte Carlo simulations for unknown sites to determine the invasion probability in the next time interval. In this study, underestimating invasion probability is unlikely an issue when fitting the gravity models to invasion data due to differences in methodology from Leung and Delaney (2006). One possible advantage of using non-parametric boosted regression trees in lieu of specified functional relationships (e.g. asymptotic equations) between propagule pressure and establishment probability lies in the functional relationship of the regression, in which the shape is determined by data (e.g. Fig. 1). However, the regression trees also did not extend to establishment probabilities of 0 and 1 as with other asymptotic equations, and thus should not be used to extrapolate results beyond the range of data.

Further, results from Leung and Delaney (2006) suggest that the invasion pattern across the entire system influenced the probability of establishment for a given destination, and in particular, that the number of invaded sites sampled was the most important. If we consider the situation where an unsampled lake is invaded and serves as a source for propagules, we expect the doubly-constrained model to be the best performer in predicting invasion status as there is already a significant difference in the levels of estimated propagule flow to invaded versus noninvaded lakes relative to the other methods (e.g. Fig. 1).

Gravity models provide a flexible modeling framework and can be used in conjunction with other models in multistage invasion models. Although we consider only propagule flow from gravity models as a sole predictor of establishment in this study, gravity models can be combined with data on habitat suitability and biotic integration into the recipient assemblage. Measures of lake physicochemistry such as lake area and depth, calcium, pH, water clarity and dissolved oxygen often define the quality of available habitat for *Bythotrephes* (MacIsaac et al. 2000; Branstrator et al. 2006; Weisz and Yan 2010) and may be used as additional predictors in boosted regression trees to model *Bythotrephes* establishment. *Bythotrephes* integration into the native community may be modeled by incorporating presence/absence or abundance data of native species as predictive variables in *Bythotrephes* establishment models.

Similar approaches to combining dispersal and establishment models have been commonly been used for other NIS. For example, Herborg et al. (2007) used the volume of ballast water transported in transoceanic vessels as a proxy of propagule pressure, in combination with an environmental niche model (Genetic Algorithm for Rule-set Prediction) to assess habitat suitability for the Chinese mitten crab *Eriocheir sinensis* in North America. Rouget and Richardson (2003) provided semi-mechanistic models for the percentage cover and species occurrences for three nonindigenous plant species in South Africa. Their multistage model indicated that propagule pressure—as measured by the distance from invasion foci—and environmental variables were able to account for 70% of successful invasions.

A key consideration of gravity models is that the basic formulation does not specifically incorporate the individual-decision making process for recreationalists when deciding between competing destinations. Thus, measures of destination attractiveness should not be ascribed as influencing individual recreational choice, but rather the overall pattern of pairwise flows. To address this issue, a hierarchical approach may be used where individual choice can be modeled by a sub-model such as a Random Utility Model. The probability of choosing one destination over another is weighted by the cost to travel to those lakes, which is then nested as input into the gravity model (Siderelis and Moore 1998). Additional avenues of research in the development of gravity models for use in risk assessment require the inclusion of NIS population dynamics such as growth rate and mortality. This would enable gravity models to serve as a bridge from describing dispersal of individual propagules to that describing population spread of NIS at a landscape level. For example, models describing population growth can be coupled with gravity models describing immigration and emigration among systems, in a manner analogous to metapopulation models. Currently, gravity models are used primarily to describe relative vector traffic to different lakes, usually in terms of boater numbers (Schneider et al. 1998; Bossenbroek et al. 2001, 2007; Leung et al. 2006).

In conclusion, our empirical data set provided an opportunity to evaluate model selection between competing classes of gravity models to model recreational boat traffic, as well as to explore the relationship between inbound propagule pressure and invasion status. Our study provides guidance for choosing the best model to describe propagule pressure and the amount of effort required to collect data in the planning stage of an experiment. Our analyses indicate that the production-constrained gravity model offers the best compromise between describing recreational traffic and estimating establishment probability. Because data required to formulate production-constrained gravity models is often readily accessible, we recommend this procedure over other possibilities for future studies.

Acknowledgments We are grateful for discussions with M. A. Lewis and A. Potapov, and for financial support from the Canadian Aquatic Invasive Species Network, an OGS scholarship and NSERC postdoctoral fellowship to JRM, and by an NSERC Discovery Grant and a DFO Invasive Species Research Chair to HJM.

References

Allouche O, Tsoar A, Kadmon R (2006) Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). J Appl Ecol 43:1223–1232

Anas A (1983) Discrete choice theory, information-theory and the multinomial logit and gravity models. Trans Res B Methodol 17:13–23

Anderson JE (1979) Theoretical foundation for the gravity equation. Am Econ Rev 69:106–116

Black JA, Salter RJ (1975) A statistical evaluation of the accuracy of a family of gravity models. Proc Instit Civil Eng Part 2 Res Theor 59:1–20

Bockstael NE, McConnell KE, Strand IE (1989) A random utility model for sportfishing: some preliminary results for Florida. Mar Res Econ 6:245–260

Bossenbroek JM, Kraft CE, Nekola JC (2001) Prediction of long-distance dispersal using gravity models: zebra mussel invasion of inland lakes. Ecol Appl 11:1778–1788

Bossenbroek JM, Johnson LE, Peters B, Lodge DM (2007) Forecasting the expansion of zebra mussels in the United States. Con Biol 21:800–810

Boudreau SA, Yan ND (2004) Auditing the accuracy of a volunteer-based surveillance program for an aquatic invader Bythotrephes. Environ Monitor Assess 91:17–26

Branstrator DK, Brown ME, Shannon LJ, Thabes M, Heimgartner K (2006) Range expansion of *Bythotrephes longimanus* in North America: evaluating habitat characteristics in the spread of an exotic invader. Biol Invas 8:1367–1379

Colautti RI, Grigorovich IA, MacIsaac HJ (2006) Propagule pressure: a null model for biological invasions. Biol Invas 8:1023–1037

Davis MA (2003) Biotic globalization: does competition from introduced species threaten biodiversity? Bioscience 53:481–489

Dempster AP, Laird NM, Rubin DB (1977) Maximum likelihood from incomplete data via the EM algorithm. J R Stat Soc Ser B (Methodological) 39:1–38

Efron B, Tibshirani R (1986) Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. Stat Sci 1:54–77

Elith J, Graham CH, Anderson RP, Dudik M, Ferrier S, Guisan A, Hijmans RJ, Huettmann F, Leathwick JR, Lehmann A, Li J, Lohmann LG, Loiselle BA, Manion G, Moritz C, Nakamura M, Nakazawa Y, Overton JM, Peterson AT, Phillips SJ, Richardson K, Scachetti-Pereira R, Schapire RE, Soberón J, Williams S, Wisz MS, Zimmerman NE (2006) Novel methods improve prediction of species' distributions from occurrence data. Ecography 29:129–151

Flowerdew R, Aitkin M (1982) A method of fitting the gravity model based on the Poisson distribution. J Reg Sci 22:191–202

Forsyth DM, Duncan RP (2001) Propagule size and the relative success of exotic ungulate and bird introductions to New Zealand. Am Nat 157:583–595

Fotheringham AS (1981) Spatial structure and distance-decay parameters. Ann Assoc Am Geogr 71:425–436

Fridley JD, Stachowicz JJ, Naeem S, Sax DF, Seabloom EW, Smith MD, Stohlgren TJ, Tilman D, Von Holle B (2007) The invasion paradox: reconciling pattern and process in species invasions. Ecology 88:3–17

Hayes KR, Barry SC (2008) Are there any consistent predictors of invasion success? Biol Invas 10:483–506

Haynes KE, Fotheringham AS (1984) Gravity and spatial interaction models. Sage, Beverly Hills

Herborg L-M, Jerde CL, Lodge DM, Ruiz GM, MacIsaac HJ (2007) Predicting invasion risk using measures of introduction effort and environmental niche models. Ecol Appl 17:663–674

Huff DL (1959) Geographical aspects of consumer behavior. U Wash Bus Rev 18:27–37

Jansakul N (2005) Fitting a zero-inflated negative binomial model via R. In: Proceedings of the 20th international Wksh Stat Model, pp 277–284

Jarnagin ST, Swan BK, Kerfoot WC (2000) Fish as vectors in the dispersal of *Bythotrephes cederstroemi*: diapausing eggs survive passage through the gut. Freshw Biol 43:579–589

Jeschke JM, Strayer DL (2005) Invasion success of vertebrates in Europe and North America. Proc Natl Acad Sci USA 102:7198–7202

Johannsson OE, Mills EL, Ogorman R (1991) Changes in the nearshore and offshore zooplankton communities in Lake Ontario—1981–88. Can J Fish Aquat Sci 48:1546–1557

Ketelaars HAM, Gille L (1994) Range extension of the predatory cladoceran *Bythotrephes longimanus* Leydig 1860 (Crustacea, Onychopoda) in Western Europe. Neth J Aquat Ecol 28:175–180

Kolar CS, Lodge DM (2001) Progress in invasion biology: predicting invaders. Trends Ecol Evol 16:199–204

Kolar CS, Lodge DM (2002) Ecological predictions and risk assessment for alien fishes in North America. Science 298:1233–1236

Leung B, Delaney DG (2006) Managing sparse data in biological invasions: a simulation study. Ecol Model 198:229–239

Leung B, Lodge DM, Finnoff D, Shogren JF, Lewis MA, Lamberti GA (2002) An ounce of prevention or a pound of cure: bioeconomic risk analysis of invasive species. Proc R Soc Lond Ser B 269:2407–2413

Leung B, Drake JM, Lodge DM (2004) Predicting invasions: propagule pressure and the gravity of Allee effects. Ecology 85:1651–1660

Leung B, Bossenbroek JM, Lodge DM (2006) Boats, pathways, and aquatic biological invasions, estimating dispersal potential with gravity models. Biol Invas 8:241–254

Linneman HV (1966) An econometric study of international trade flows. North-Holland, Amsterdam

Liu CR, Berry PM, Dawson TP, Pearson RG (2005) Selecting thresholds of occurrence in the prediction of species distributions. Ecography 28:385–393

Lobo JM, Jiménez-Valverde A, Real R (2008) AUC: a misleading measure of the performance of predictive distribution models. Global Ecol Biogeogr 17:145–151

Lockwood JL, Cassey P, Blackburn T (2005) The role of propagule pressure in explaining species invasions. Trends Ecol Evol 20:223–228

Lockwood JL, Cassey P, Blackburn T (2009) The more you introduce the more you get: the role of colonization pressure and propagule pressure in invasion ecology. Divers Distrib 15:904–910

Lodge DM, Williams S, MacIsaac HJ, Hayes KR, Leung B, Reichard S, Mack RN, Moyle PB, Smith M, Andow DA, Carlton JT, McMichael A (2006) Biological invasions: recommendations for US policy and management. Ecol Appl 16:2035–2054

MacIsaac HJ, Ketelaars HAM, Grigorovich IA, Ramcharan CW, Yan ND (2000) Modeling *Bythotrephes longimanus* invasions in the Great Lakes basin based on its European distribution. Arch Hydrobiol 149:1–21

MacIsaac HJ, Borbely J, Muirhead JR, Graniero P (2004) Backcasting and forecasting biological invasions of inland lakes. Ecol Appl 14:773–783

McPherson JM, Jetz W, Rogers DJ (2004) The effects of species' range sizes on the accuracy of distribution models: ecological phenomenon or statistical artefact? J Appl Ecol 41:811–823

Melbourne BA, Cornell HV, Davies KF, Dugaw CJ, Elmendorf S, Freestone AL, Hall RJ, Harrison S, Hastings A, Holland M, Holyoak M, Lambrinos J, Moore K, Yokomizo H (2007) Invasion in a heterogeneous world: resistance, coexistence or hostile takeover? Ecol Lett 10:77–94

Potapov A, Muirhead JR, Lele SR, Lewis MA (2010) Stochastic gravity models for modeling lake invasions. Ecol Model. doi:10.1016/j.ecolmodel.2010.07.024

Richardson DM, Pyšek P, Rejmánek M, Barbour M, Panetta FD, West CJ (2000) Naturalization and invasion of alien plants: concepts and definitions. Divers Distrib 6:93–107

Rouget M, Richardson DM (2003) Inferring process from pattern in plant invasions: a semimechanistic model incorporating propagule pressure and environmental factors. Am Nat 162:713–724

Sala OE, Chapin FS III, Armesto JJ, Berlow E, Bloomfield J, Dirzo R, Huber-Sanwald E, Huenneke LF, Jackson RB, Kinzig A, Leemans R, Lodge DM, Mooney HA, Oesterheld M, Poff NL, Sykes MT, Walker BH, Walker M, Wall DH (2000) Global biodiversity scenarios for the year 2100. Science 287:1770–1774

Schneider DW, Ellis CD, Cummings KS (1998) A transportation model assessment of the risk to native mussel communities from zebra mussel spread. Conserv Biol 12:788–800

Siderelis C, Moore RL (1998) Recreation demand and the influence of site preference variables. J Leisure Res 30:301–318

Vander Zanden MJ, Olden JD (2008) A management framework for preventing the secondary spread of aquatic invasive species. Can J Fish Aquat Sci 65:1512–1522

Veltman CJ, Nee S, Crawley MJ (1996) Correlates of introduction success in exotic New Zealand birds. Am Nat 147:542–557

Von Holle B, Simberloff D (2005) Ecological resistance to biological invasion overwhelmed by propagule pressure. Ecology 86:3212–3218

Weisz EJ, Yan ND (2010) Relative value of limnological, geographic, and human use variables as predictors of the presence of *Bythotrephes longimanus* in Canadian Shield lakes. Can J Fish Aquat Sci 67:462–472

Zipf GK (1946) The P1P2/D hypothesis: on the intercity movement of persons. Am Sociol Rev 11:677–686