

Global hot spots of biological invasions: evaluating options for ballast-water management

John M. Drake* and David M. Lodge

Department of Biological Sciences, University of Notre Dame, Notre Dame, IN 46556, USA

Biological invasions from ballast water are a severe environmental threat and exceedingly costly to society. We identify global hot spots of invasion based on worldwide patterns of ship traffic. We then estimate the rate of port-to-port invasion using gravity models for spatial interactions, and we identify bottlenecks to the regional exchange of species using the Ford–Fulkerson algorithm for network flows. Finally, using stochastic simulations of different strategies for controlling ballast-water introductions, we find that reducing the per-ship-visit chance of causing invasion is more effective in reducing the rate of biotic homogenization than eliminating key ports that are the epicentres for global spread.

Keywords: ballast water; biological invasion; biotic homogenization; Ford–Fulkerson algorithm

1. INTRODUCTION

Biological invasions are a leading cause of species extinctions and biotic homogenization worldwide (Sala *et al.* 2000; Lockwood & McKinney 2001). Therefore, describing existing patterns of invasion and forecasting future invasions are essential for the more effective management of global biodiversity resources. Yet, invasions are difficult to quantify because organisms at low density are often overlooked. Furthermore, because invasions are commonly initiated during trade, quantification requires the integrated analysis of ecological factors and commercial pathways. Accounting for movements of 7000–10 000 species simultaneously (Carlton 1999), ballast water is the principal source of invasive species in coastal freshwater and marine ecosystems. Here, we estimate current invasion rates and show that the relative effectiveness of eliminating key port-to-port shipping linkages is negligible compared with moderate reductions in the per-ship-visit chance of introducing an invasive species. We used data on the number of ships visiting 243 ports worldwide and port-to-port travel probabilities to identify invasion hot spots and key port-to-port linkages. We then identified the subset of ports that contribute most to inter-region exchange of non-indigenous species, using network flows analysis. Finally, we used a stochastic simulation to estimate times to invasion from secondary (stepping-stone) invasions under two alternative management approaches: elimination of organisms in ballast water at critical ports (such as might be achieved with onshore treatment facilities); or reduction in the number of organisms in every ship globally (such as might be achieved with on-board ballast-water treatment).

2. INVASION HOT SPOTS

Since the rate of biological invasion in coastal aquatic ecosystems depends on the volume of ship traffic to a region (Ricciardi 2001), ports that enjoy a disproportionate share of global commerce are likely to be the

most common sources and recipients of introduced species. Therefore, using data on the number of ships calling at 243 frequently visited ports (the subset of the ports listed in Zachcial & Heideloff (2001) for which there were sufficient data; see electronic Appendix A available on The Royal Society's Publications Web site, for a list of ports used in this study), we determined the annual number of ship visits within a 500 km radius (figure 1a) and the change between 1996 and 2000 (figure 1b). A 500 km radius was chosen so that the resulting hot spots are visually distinguishable on maps. The results do not depend greatly on this choice. We converted the numbers of ship visits to probabilities of invasion with the formula

$$\pi = 1 - (1 - p)^n, \quad (2.1)$$

the complement of the probability of all ship visits failing to cause an invasion, where $p = 0.000\ 44$, the per-ship-call probability of initiating invasion (see § 5a) and n is the number of ships. Hot spots on the map of ship density result from the interaction between the number of ships arriving at a port and the density of ports in an area. The southeast Asia hot spot, for example, is the result of the presence of a few frequently visited ports. By contrast, the European hot spots result from larger clusters of ports with intermediate levels of trade.

3. NETWORK FLOWS MODEL

As the ports identified in our analysis of invasion hot spots develop communities of non-indigenous species, they will probably serve as stepping-stones in a process of global spread causing widespread biotic homogenization (e.g. Apte *et al.* 2000). Using the framework of gravity models for movement behaviour (Fotheringham & O'Kelly 1989; Bossenbroek *et al.* 2001) and data on 28 748 ship arrivals in the US (data from the National Ballast Water Information Clearinghouse; <http://invasions.si.edu/ballast.htm>), we constructed a transportation network model to determine the most likely pathways of introduction. This model was fitted to data on ship traffic at 243 ports worldwide for the year 2000 (data obtained from Zachcial & Heideloff (2001)).

* Author for correspondence (drake.4@nd.edu).

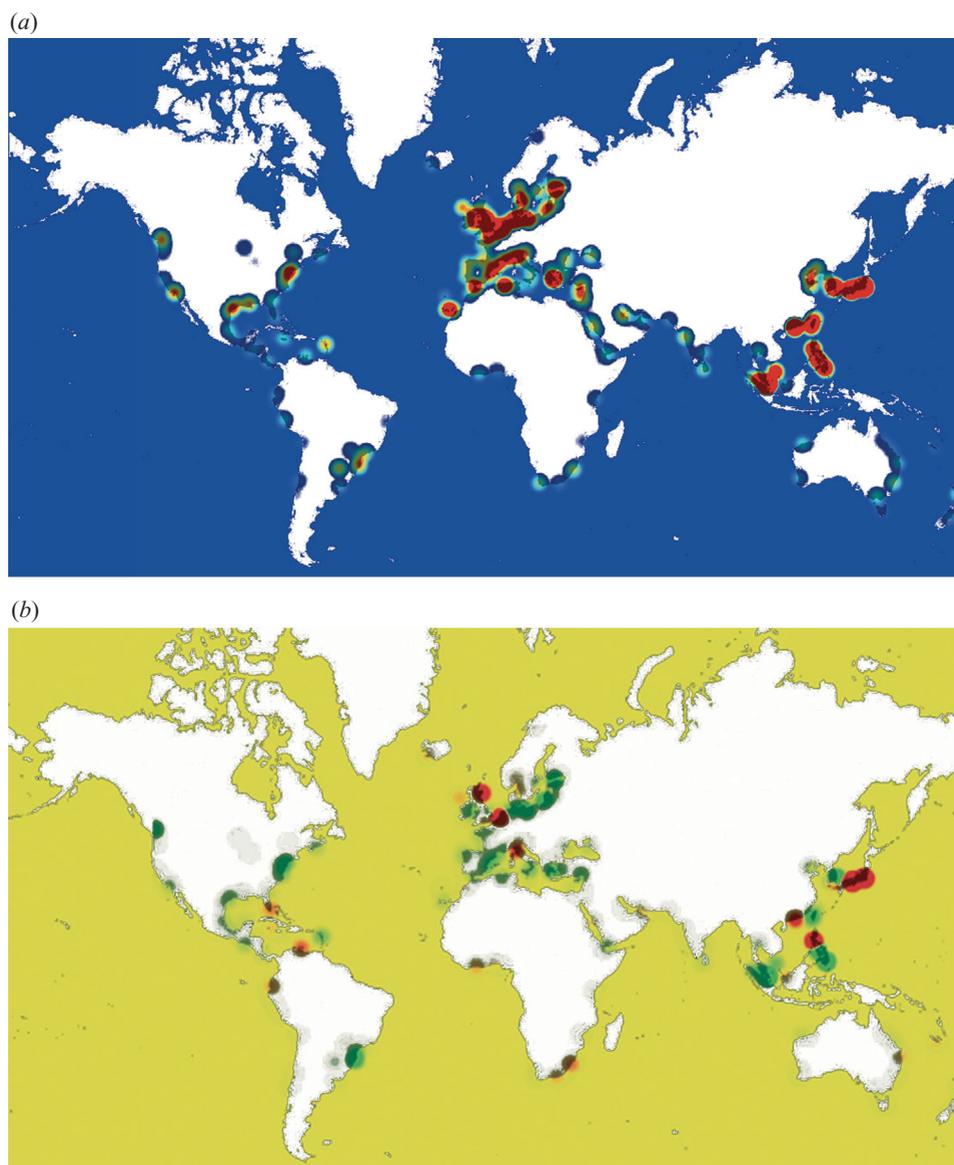


Figure 1. (a) Global hot spots for biological invasion from ballast water. The density of ship visits was estimated by a quartic approximation to a Gaussian kernel (McCoy & Johnston 2001). Expected invasion rates range from 0 (blue) to 2.94×10^{-4} species $\text{km}^{-2} \text{yr}^{-1}$ (red). (b) Changes in the expected rate of invasion from 1996 to 2000 range from -4.58×10^{-8} species $\text{km}^{-2} \text{yr}^{-1}$ (green) to 2.11×10^{-8} species $\text{km}^{-2} \text{yr}^{-1}$ (red). Yellow indicates background rates of no change.

In all but the simplest networks, it is not feasible to measure flow between every pair of nodes empirically since the number of connections increases quadratically with the number of nodes. Therefore, gravity models are often used in the analysis of spatial interactions to estimate expected flows from known inflows and outflows, based on a deterrence function that relates the location-to-location travel probability to the distance between locations (Fotheringham & O’Kelly 1989). If the deterrence function is exponential, the mathematical expression is identical to Newton’s model of gravitational attraction (hence ‘gravity model’). However, since global trade is highly heterogeneous owing to social, political and geographical obstacles, traffic to a distant trading partner might be more common than traffic to certain nearby ones. Therefore, using moment recursion relations (Cobb *et al.* 1983), we fitted the multimodal gamma probability distribution

$$G_R(x) = \xi x^{\alpha-1} \exp(\theta_1 x + \theta_2 x^2 + \dots + \theta_k x^k) \quad (3.1)$$

to data on 28 748 ships arriving in the continental US. The distribution of distances travelled has two modes, at *ca.* 1100 and 8275 km, evidently corresponding to intra- and intercontinental traffic, respectively (figure 2).

Global biotic homogenization is primarily a function of the rate at which species are translocated outside their native ranges, particularly to distant regions. Using equation (3.1) for the deterrence function, we fitted a doubly constrained gravity model to these data to estimate the total number of ships travelling between each pair of ports for the year 2000. Viewing the gravity model for ship traffic as a system of network flows, we used the Ford–Fulkerson algorithm for network flows with multiple sources and multiple sinks (Ford & Fulkerson 1962) to identify the set of bottlenecks to the invasion process in five global regions: Asia, North America, South America,

Table 1. Dominant pathways for biological invasions from ballast water.

(The most important pathways for region-to-region global exchange of non-indigenous species in ballast water, the number of ships expected to make the voyage each year (based on the year 2000) and the estimated port-to-port invasion rate, determined from a global gravity model for ballast-water invasions.)

regional exchange	pathway	expected number of ships per year	expected number of invasions per year
Asia–North America	Chiba–Long Beach	402	0.18
Asia–South America	Las Palmas de Gran Cana–Tubarao	176	0.08
Asia–Africa	Durban–Singapore	839	0.37
Asia–Europe	Singapore–Piraeus	3646	1.60
Africa–North America	Durban–Singapore	839	0.37
Africa–South America	Las Palmas de Gran Cana–Tubarao	176	0.08
Africa–Europe	Durban–Singapore	839	0.37
Europe–North America	Chiba–Long Beach	402	0.18
Europe–South America	Las Palmas de Gran Cana–Tubarao	176	0.08
North America–South America	Las Palmas de Gran Cana–Tubarao	176	0.08

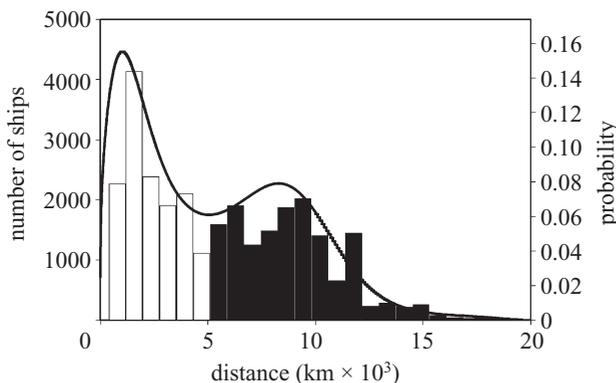


Figure 2. The probability density of port-to-port distances travelled is a bimodal distribution, corresponding to intracontinental (open bars) and intercontinental (filled bars) traffic. This distribution is fitted by a $G_6(x)$ multimodal gamma distribution with parameters $\alpha = 1.2505$, $\theta_1 = 0.5520$, $\theta_2 = -0.5795$, $\theta_3 = 0.1428$, $\theta_4 = -0.0150$, $\theta_5 = 7.0261 \times 10^{-4}$ and $\theta_6 = -1.2202 \times 10^{-3}$; ξ is a normalization parameter. This distribution is used for the deterrence function in the doubly constrained gravity model used to estimate the pairwise flows between ports.

Africa and Europe. Only seven cities are involved in the most important pathways for the exchange of species among these world regions (table 1): Chiba (Japan), Durban (South Africa), Las Palmas de Gran Cana (Spain), Long Beach (USA), Piraeus (Greece), Singapore (Singapore) and Tubarao (Brazil). Because the model is based solely on distance, inflows and outflows, these pathways are also the most important for flow in the opposite direction (i.e. Chiba–Long Beach is the most important pathway from Asia to North America and vice versa).

4. STOCHASTIC SIMULATION

Finally, a stochastic simulation shows the relative effectiveness of two possible strategies for managing ballast water: (i) eliminating key port-to-port linkages; or (ii) reducing the per-ship probability of introducing an invasive species. A stochastic model is required because invasion is often a stepping-stone process: once one

location has been invaded, it serves as a source for secondary introductions. A simulation based on the stochastic model requires an estimate of the per-species rate of establishment, which is the product of the chances of entrainment in ballast water, survival during transport and establishment after release (MacIsaac *et al.* 2002). Except for a few well-studied cases, far less is known about the invasion rates of individual species than about overall rates. Moreover, because of heterogeneity among species, establishment rates estimated from the few well-studied organisms, which are typically highly invasive, would be biased. Therefore, as a first approximation, we make the order of magnitude assumption that the total number of potentially invasive species inhabiting each port is between 100 and 1000. For comparison, one study found 367 taxa in 159 ships originating in Japan and arriving in Coos Bay, Oregon, USA (Carlton & Geller 1993). Dividing the overall invasion rate obtained above (see §§ 2 and 5a) by the total number of species suggests that a reasonable per-species invasion rate obtained above is somewhere between 4.4×10^{-6} and 4.4×10^{-7} per ship.

Biological invasions on the network were simulated 1000 times for each scenario, with the following 10 ports serving as hypothetical sources for invasions of the other ports in the global network: Yokohama (Japan), Nhava Sheva (India), Antwerp (Belgium), Narvik (Norway), Ceuta (Spain), Cotonou (Benin), Houston (USA), Jacksonville (USA), Itaquí (Brazil) and Valparaiso (Chile). These ports were selected to be representative both of prominent contributors of non-indigenous propagules and of less dominant sources from all five global regions, and were exclusive of the seven most important ports (in table 1), which are candidates for elimination as a hypothetical strategy for management. To compare different strategies for managing ballast water, we compared three scenarios: (i) no action (invasion rate of 4.4×10^{-6} species per ship); (ii) elimination of key ports from the network, i.e. removal of ports in table 1 (invasion rate of 4.4×10^{-6} species per ship); and (iii) a reduction in the per-ship propagule pressure by a factor of 10 (invasion rate of 4.4×10^{-7} species per ship). Eliminating key ports (blue lines in figure 3) has a relatively small effect on overall invasion rates (cf. no action rates; black lines in figure 3). By contrast, reducing

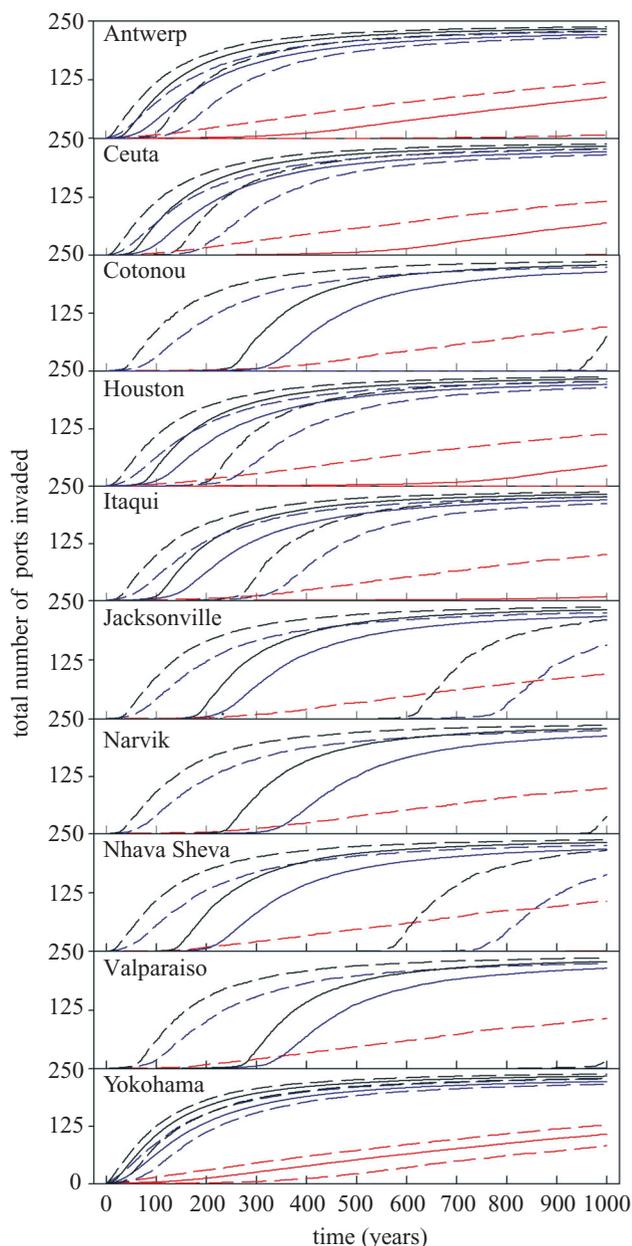


Figure 3. Results from stochastic simulations of the cumulative numbers of ports that have been invaded over time with a species originating in each of the 10 ports shown (labels on subfigures). Shown are both the median number of invasions (solid lines) and the 2.5 and 97.5 percentiles (dashed lines). Reducing the per-ship-visit probability of initiating invasion by a factor of 10 (red lines) is more effective than eliminating the seven most important ports from the network (blue lines). Compare both with the background rate of invasion resulting from no management action (black lines).

per-ship probabilities of initiating invasion by a factor of 10 was much more effective (red lines). This pattern holds both for dominant ports (e.g. Yokohama, Antwerp) and for lesser ones (e.g. Jacksonville). Thus, these results indicate that even key ports are individually relatively insignificant as stepping-stones for invasion compared with the generally high rate of invasion expected among ports overall.

5. METHODS

(a) Invasion hot spots

To convert ship number to number of introduced species, we re-analysed data (Ricciardi 2001) on net tonnage of shipping traffic and the numbers of non-indigenous species discovered in the Great Lakes in each decade since 1900. If bias arises from treating this relationship in the Great Lakes as a global average, it is unknown and therefore unaccounted for in our model. A generalized linear regression with Poisson errors showed that shipping traffic was significantly related to the number of non-indigenous species discovered ($y = 8.47 \times 10^{-8}x$ where y is the number of non-indigenous species discovered and x is the shipping volume, in metric tonnes; $p < 0.0001$). To obtain a per-ship probability of invasion, we rescaled this relationship by the average tonnage of ships in this dataset (5197 tons). The estimated per-ship probability of causing an invasion was therefore 0.000 44 (95% CI = $\pm 0.000 08$) or approximately one in 2275 ships. The number of ship visits for each port was multiplied by this value to estimate port-specific annual invasion rates (i.e. the expected total number of new non-indigenous species from all sources).

(b) Network flows model

The doubly constrained gravity model is defined by the equation $F_{ij} = A_i B_j O_i D_j f(d_{ij})$, where F_{ij} is the flow between ports i and j , O_i is the total flow from the origin i , D_j is the total flow to destination j , $f(d_{ij})$ is the deterrence function based on the distance from i to j , and A_i and B_j are balancing factors that define the constraints of the system (Fotheringham & O'Kelly 1989). A_i and B_j are defined implicitly by the following relations:

$$A_i = \left(\sum_j B_j D_j f(d_{ij}) \right)^{-1} \quad \text{and} \quad B_j = \left(\sum_i A_i O_i f(d_{ij}) \right)^{-1}.$$

The model was fitted by iteratively adjusting all A_i and B_j until the model converged. The matrix of estimated pairwise flows is obtainable upon request from the authors.

The gravity model obtained above can be viewed as a system of network flows by employing the standard graph-theoretic apparatus (van Lint & Wilson 2001). Accordingly, each port in the dataset is represented by a vertex ($v \in V$) on the graph while each pairwise linkage is represented by an edge ($e \in E$). The resulting graph $G[V, E]$ is complete in the sense that it is fully connected. From the perspective of potentially transported propagules, the capacity of each edge $c(x, y)$ to facilitate invasion is the total flow of ship traffic between two vertices x and y , i.e. the flow of ships estimated with the gravity model above. To reduce complexity, we used only the 100 most frequently visited ports out of the total dataset of 243, constituting ca. 84.8% of total ship traffic. Based on the classifications of the original dataset, each of these ports was assigned to one out of five global regions: Asia, Africa, North America, South America or Europe. (No ports from Oceania, the sixth global region in the dataset, were among the 100 most frequently visited ports.) For the purposes of this analysis, we define global homogenization as the movement of species between regions, but not invasion within regions. Therefore, each region alternately represents a set of possible sources and a set of possible sinks for species invasions. Let S be the set of sources, T the set of sinks and R the set of intermediate vertices (i.e. $R = V - (S \cup T)$). We wish to quantify the total flow of species from S to T for each S and T , i.e. with each region alternately as each source and sink. Ford & Fulkerson (1962) demonstrated that the problem of the

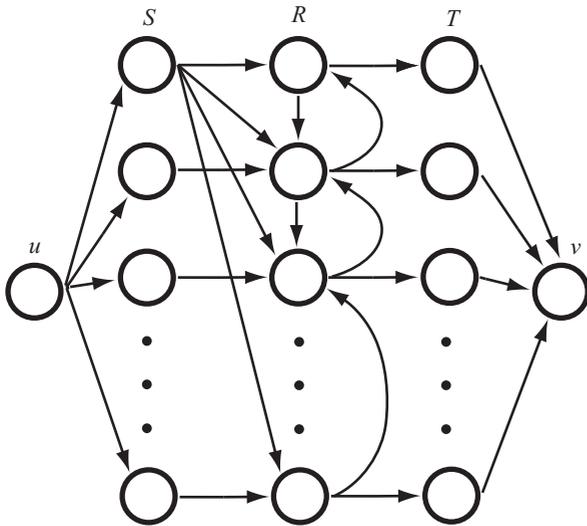


Figure 4. The network composed of multiple sources (S), multiple sinks (T) and intermediate vertices (R) can be extended to the familiar single-source, single-sink network by the addition of two new vertices, u and v . The extended network can be analysed in the usual way to obtain the minimum cut separating S and T .

maximum flow for multiple sources and multiple sinks could be reduced to the single-source single-sink case by the addition of two new vertices u and v (figure 4). The extended network $G^*[V^*, E^*]$ is assigned the following capacities:

$$c^*(u, x) = \infty \quad x \in S,$$

$$c^*(x, v) = \infty \quad x \in T,$$

$$c^*(x, y) = c(x, y) \quad x \in R \text{ or } y \in E.$$

Notice that the extended network $G^*[V^*, E^*]$ is not fully connected. The original implementation of the Ford–Fulkerson algorithm was employed to identify the minimum cut, i.e. the separation of all vertices into two subsets $X(S \subseteq X)$ and $Y(T \subseteq Y)$ such that the sum of all flows $f(x, y)$ ($x \in X, y \in Y$) is at a minimum. The maxflow–mincut theorem (Ford & Fulkerson 1962) states that the maximum flow from all sources to all sinks in the network is equal to the minimum flow across all ‘cuts’, where a cut is a pair of subsets partitioning all vertices into one or the other subset. The total flow from sources to sinks is the sum of all flows across the minimum cut. It follows that a reduction in flows across the minimum cut results in an equivalent reduction in the maximum flow from sources to sinks. Therefore, the greatest single action that can be taken to reduce the maximum flow of the network is to eliminate the greatest flow across the minimum cut.

(c) Stochastic simulation

The expected number of ships, n , travelling from i to j in each time-step (year) is given by the gravity model. For the port-removal experiment, $n = 0$ for any i or $j \in$ (Chiba, Durban, Las Palmas de Gran Cana, Long Beach, Piraeus, Singapore, Tubarao). The per-species probability of invasion from i to j at each time-step is given by $1 - (1 - p)^n$, where p is either 4.4×10^{-6} or 4.4×10^{-7} , depending on the hypothetical reduction in propagule pressure. Thus, the probability that an uninvaded port j is invaded at t over all sources i is $1 - \prod_i (1 - p)^{n_i}$. Simulations were run 1000 times and the results

tabulated to provide the 2.5, 50 and 97.5 percentiles for the number of ports invaded by t .

6. CONCLUSIONS

Biological invasions from ballast water are a leading environmental concern in navigable freshwater and marine ecosystems (Carlton 1999; Sala *et al.* 2000). In this paper, we have identified global hot spots for invasion including large regions of southeast Asia, northern Europe and the Mediterranean Sea, and some coastal areas in North and South America (figure 1a). Our analysis suggests that some areas known to have a large number of non-indigenous species, including the North American Great Lakes (Ricciardi 2001) and San Francisco Bay and Delta (Cohen & Carlton 1998) ecosystems, are of relatively minor importance. We speculate that this reflects differential investment in research and that non-indigenous species in areas less thoroughly studied may persist undetected. Alternatively, ecosystems may exhibit differential susceptibility to invasion by species indigenous to the ecosystems of primary trading partners. Additionally, although the expected rate of invasion is increasing in some places, it is declining in other places (figure 1b). This results primarily from changing patterns of ship traffic. Most areas of no expected change (yellow areas in figure 1b) are areas exhibiting low projected invasion rates (blue areas in figure 1a).

Potential macro-scale alternatives to managing global biotic homogenization through ballast-water discharge are available. Key ports that function as epicentres for inter-region exchange of species are Chiba (Japan), Durban (South Africa), Las Palmas de Gran Cana (Spain), Long Beach (USA), Piraeus (Greece), Singapore (Singapore) and Tubarao (Brazil). However, stochastic simulations show that reducing the per-ship-visit probability of initiating invasion is paramount in reducing the rate of global biotic homogenization through ballast-water transport and that the elimination of invasion epicentres is likely to be less effective. The methods used here for ballast-water assessment could also be used for similar analyses of container shipping and other commercial pathways that introduce non-indigenous species to terrestrial ecosystems.

The authors thank A. Ricciardi for data about invasion rates in the Great Lakes, A. Reid for consultation about the implementation of the Ford–Fulkerson algorithm, and R. Keller, M. Xenopoulos, T. Kreps, S. Rosenthal and two reviewers for comments about the manuscript. C. Arango assisted with figure 1. This work was funded by grants from US EPA STAR (to D.M.L.), US Great Lakes Fishery Commission (to D.M.L.) and US NSF Integrated Research Challenges in Environmental Biology (to D.M.L.), and by an EPA STAR Graduate Research Fellowship (to J.M.D.) and a University of Notre Dame Schmitt Graduate Research Fellowship (to J.M.D.).

REFERENCES

- Apte, S., Holland, B. S., Godwin, L. S. & Garder, J. P. A. 2000 Jumping ship: a stepping stone event mediating transfer of a non-indigenous species via a potentially unsuitable environment. *Biol. Invasions* 2, 75–79.
- Bossenbroek, J. M., Kraft, C. E. & Nekola, J. C. 2001 Prediction of long-distance dispersal using gravity models: zebra mussel invasion of inland lakes. *Ecol. Appl.* 11, 1778–1788.

- Carlton, J. T. 1999 The scale and ecological consequences of biological invasions in the world's oceans. In *Invasive species and biodiversity management* (ed. O. Sandlund, P. Schei & Å. Viken), pp. 195–212. Dordrecht, The Netherlands: Kluwer.
- Carlton, J. T. & Geller, J. B. 1993 Ecological roulette: the global transport and invasion of nonindigenous marine organisms. *Science* **261**, 78–82.
- Cobb, L., Koppstein, P. & Chen, N. H. 1983 Estimation and moment recursion relations for multimodal distributions of the exponential family. *J. Am. Stat. Assoc.* **78**, 124–130.
- Cohen, A. N. & Carlton, J. T. 1998 Accelerating invasion rate in a highly invaded estuary. *Science* **279**, 555–558.
- Ford, L. R. & Fulkerson, D. R. 1962 *Flows in networks*. Princeton University Press.
- Fotheringham, A. S. & O'Kelly, M. E. 1989 *Spatial interaction models: formulations and applications*. Dordrecht, The Netherlands: Kluwer.
- Lockwood, J. L. & McKinney, M. L. 2001 *Biotic homogenization*. Dordrecht, The Netherlands: Kluwer.
- McCoy, J. & Johnston, K. 2001 *Using ArcGIS spatial analyst*. Redlands, CA, USA: Environmental Systems Research Institute, Inc.
- MacIsaac, H. J., Robbins, T. C. & Lewis, M. A. 2002 Modeling ships' ballast water as invasion threats to the Great Lakes. *Can. J. Fish. Aquat. Sci.* **59**, 1245–1256.
- Ricciardi, A. 2001 Facilitative interactions among aquatic invaders: is an 'invasional meltdown' occurring in the Great Lakes? *Can. J. Fish. Aquat. Sci.* **58**, 2513–2525.
- Sala, O. E. (and 18 others) 2000 Global biodiversity scenarios for the year 2100. *Science* **287**, 1770–1774.
- van Lint, J. H. & Wilson, R. M. 2001 *A course in combinatorics*, 2nd edn. Cambridge University Press.
- Zachcial, M. & Heideloff, C. 2001 *ISL shipping statistics yearbook 2001*. Bremen, Germany: Institute of Shipping Economics and Logistics.

Visit <http://www.journals.royalsoc.ac.uk> to see an electronic appendix to this paper.