Invited review

Risk assessment and screening for potentially invasive fishes

CINDY KOLAR

   U.S. Geological Survey
   Upper Midwest Environmental Sciences Center
   2630 Fanta Reed Road
   La Crosse, WI 54603
   United States
   email: ckolar@usgs.gov

Abstract Preventing the introduction of potentially invasive species is becoming more important as this worldwide problem continues to grow. The ability to predict the identity or range of potential invaders could influence regulatory decisions and help to optimally allocate resources to deal with ongoing invasions. One screening tool presented here, using species life history and environmental tolerances to identify potential invaders similar to past invaders, can be used to predict potential invading species. Another screening tool, genetic algorithms, can be used to predict the potential range of an invading species. Use and further development of tools such as these, that are quantitative and relatively transparent, would give managers and other decision makers more information for making better-informed decisions.

Keywords GARP; genetic algorithms; habitat niche modelling; invasive species; risk assessment; species characteristics; species screening

INTRODUCTION

Invasive species are a leading threat to global biodiversity (Sala et al. 2000; Novacek & Cleland 2001) and are economically costly (Pimentel et al. 2000). To deal with this growing threat, an increasing number of countries and states require (e.g., New Zealand and Western Australia) or have legislation pending that would require (e.g., United States) the screening of plants and/or animals before importation to identify those that are potentially invasive. Screening tools used for this purpose need to be transparent, relatively easy to use, reliable, and unbiased (i.e., give the same results independent of screener) to be defensible and fair (Kolar & Lodge 2002). Ideally, they should also be quantitative—an attribute that would embody several of the characteristics listed above. Here, I will describe two quantitative and reliable approaches that can be used as screening tools for invading species. The first method, based on the life history characteristics and environmental tolerances of previous invaders in an ecosystem (using historical fish invasions in the Laurentian Great Lakes as an example), can be used to predict whether a species is likely to become established, spread, or to be perceived as a nuisance. The second method, based on ecological niche modelling of potentially invasive species (using General Algorithm for Rule-set Predictions, or GARP, as an example), can be used to predict the potential range and other effects of non-indigenous species.

SCREENING TOOLS BASED ON SPECIES CHARACTERISTICS

Screening tools to predict invading species based on species characteristics assume that invading species somehow differ from species that tend not to invade. If invaders can be accurately distinguished from non-invaders, it follows that the identity of potentially invading species can be predicted. It has been hypothesised that characteristics such as early

maturity, high fecundity, asexual reproduction, having general habitat and diet requirements, and existing well with humans are associated with successful invading species (see more complete listing in Lodge 1993). One synthesis of studies examining characteristics of invading species suggested, however, that although some patterns were present, characteristics associated with invading species were not consistent across locations (Williamson 1996). In a 2001 literature review of quantitative studies that limited potential confounding in the data by examining only one ecosystem, taxon, and stage in the invasion process at a time (e.g., either "establishment" or "invasiveness"), Kolar & Lodge (2001) found that some characteristics of invading species differed from non-invading species across studies. They found, for instance, that invasive plants are more likely to have a history of invasion and to reproduce vegetatively than non-invasive plants in particular ecosystems. In addition, some characteristics, such as having a broad diet and diverse climatic tolerances, previously thought to be associated with invaders were consistently not so associated (Kolar & Lodge 2001). This review located only studies conducted on plants and birds, however, thereby identifying the need for similar studies to be conducted for other taxa. Because invading and non-invading species differed in respect to some species characteristics, results from this study indicate some promise using species characteristics to identify potential invaders.

Until recently, few quantitative species screening models had been developed for aquatic ecosystems. Rather, the method most often used for aquatic species has been deductive risk assessment. This method, based on making inferences about a risk based on what is already known or has occurred, begins with a review of the literature to identify species with a history of invasion and then uses selection criteria to identify those with a higher potential of becoming invasive at a given location (Hayes & Sliwa 2003). Although the method is not based strictly on species characteristics, species attributes can be included in the specific selection criteria. Deductive risk assessment has been used to identify potential invaders to the Laurentian Great Lakes (Ricciardi & Rasmussen 1998), potentially invasive aquatic weeds in New Zealand (Champion & Clayton 2000), and potential marine invaders in Australia (Hayes & Sliwa 2003), among other applications. A weakness in using the deductive risk assessment approach is the assumption that all potential invaders have already invaded elsewhere.

Although this is often true, high impact invading species in an ecosystem sometimes have no history of previous invasion (e.g., sea lamprey *Petromyzon marinus* in the Laurentian Great Lakes).

Screening approaches based on species characteristics to date have been dominated by efforts for terrestrial plants. Two of the more widely used weed risk assessments, the Weed Risk Assessment of Australia (Groves et al. 2001) and the Weed-Initiated Pest Risk Assessment guidelines developed by the Animal and Plant Health Inspection Service of the U.S. Department of Agriculture (2002), are based on qualitative or semi-quantitative categorisations of diagnostic characteristics. Two quantitative approaches of screening using species characteristics for plants include Reichard & Hamilton (1997) to predict invasive woody plants in North America, and Daehler et al. (2004) to predict invasive plants in Hawaii. Below is an example of a quantitative risk assessment constructed using species life history characteristics and environmental tolerances in one ecosystem (the Laurentian Great Lakes), for one taxon (fishes), and for several stages of the invasion sequence (establishment, spread, and impact), examined independently.

**Historical fish invasions
in the Laurentian Great Lakes as an example**

Kolar & Lodge (2002) used the historical establishment of non-indigenous fishes in the Laurentian Great Lakes to determine if successfully introduced fishes differed predictably from fishes that were introduced but did not become established. They developed statistical models that differentiated fishes introduced successfully ($n = 24$ species) from those that failed to become established in the Great Lakes ($n = 21$ species). Data on 26 variables (including 14 life history characteristics, five environmental tolerances, six aspects of invasion history, and degree of human use) were obtained for each species. Statistical models developed from those data identified four variables that accurately discriminated successful invaders from unsuccessful invaders with 87–96% accuracy using discriminant analysis and categorical and regression tree analysis (see Kolar & Lodge 2002). Discriminant functions were also developed to distinguish quickly-spreading from slowly-spreading non-indigenous fishes, and those fishes perceived as a nuisance from those perceived as a non-nuisance (see Kolar & Lodge 2002).

Kolar & Lodge (2002) used models based on historical fish invasions in the Great Lakes and life history information from Eurasian fishes to identify

18 fishes from the Ponto-Caspian Basin, a donor region for introduced species in the Great Lakes, with a high risk of becoming established in the Great Lakes if introduced. Further identified were 5 of the 18 that were also predicted to spread quickly and to be perceived as a nuisance in the Great Lakes if they were introduced (see Kolar & Lodge 2002). Although the species characteristics found to be associated with invading fishes and the predicted high-risk species are specific to the Great Lakes, a major advantage to using this modelling approach as a screening tool is its applicability to a variety of ecosystems, taxa, and stages of the invasion process. Other advantages are that the models are not difficult to construct using presently available software, and, although substantially more species characteristic data are needed to construct the models, only limited data are required to use them as screening tools. A disadvantage of this modelling approach is the limited availability of data on failed introductions, although these data are probably more available than may be perceived (e.g., from agency stocking records, grey literature). Other disadvantages include data-intensive model development and geographic limitations to model predictions to the ecosystem for which the model was developed. Also, the reliability of these models (and all other models) depends not only on the accuracy of the prediction, but also the frequency with which the event occurs at all (i.e., the "base rate"). Models predicting rare events can be susceptible to substantial base rate error (Smith et al. 1999). The effect of base rate on the models in this example are minimal (see Kolar & Lodge 2002).

SCREENING TOOLS BASED ON ECOLOGICAL NICHE REQUIREMENTS

Another quantitative method to screen for potentially invasive species is based on the concept of the ecological niche, which refers to the "set of tolerances and limits in multidimensional space that define where a species is potentially able to maintain populations" (based on Grinell 1904, 1917; quoted from Peterson & Vieglais 2001) and includes ecological tolerances and other factors limiting distribution. Ecological niche models are developed to explain geographic distributions and typically include parameters such as temperature, precipitation, aspect, elevation, and vegetation (Peterson & Vieglais 2001). Thus the goal of ecological niche modelling is to describe the fundamental niche of a species as completely as possible using geographical and ecological data. Once an ecological niche model is developed for a species, the model can then be used as a screening tool to predict invasive potential by projecting the niche parameters outside of the native range where these requirements can be met. Areas providing necessary niche requirements would be expected to support the species if it were introduced. Thus, ecological niche models and resulting potential distribution maps can be used to evaluate the invasive potential of species not yet found in a given ecosystem.

General Algorithm for Rule-set Predictions (GARP) as an example

General Algorithm for Rule-set Predictions (GARP) is a general machine-learning algorithm that has been used to predict species distributions from geographical and ecological data (Peterson & Vieglais 2001; Drake & Bossenbroek in press) (available for download at http://www.lifemapper.org/desktopgarp). Using species presence or absence data and spatially explicit environmental niche parameters (such as precipitation, elevation, aspect, and temperature, available on a global scale at a fairly fine resolution), GARP searches iteratively for non-random correlations between species presence/absence and the niche parameter values using several different types of rules (atomic, logistic regression, bioclimatic envelopes, and negated bioclimatic envelopes) (Stockwell & Noble 1992; Stockwell 1999; Stockwell & Peters 1999). The resulting ecological niche model can then be tested for accuracy using point location data from the native range. Predictions of the potential range of a species can then be made by projecting the ecological niche model to identify areas where the species may become established if introduced (see manual by Payne & Stockwell 2002, for instructions on use of GARP).

GARP has been used for a variety of applications, including modelling spatial patterns of rare species for conservation purposes (Ortega-Huerta & Peterson 2004), examining realised and potential effects of global climate change on species' distributions (Peterson 2003a,b; Thomas et al. 2004), investigating evolution and speciation (Peterson et al. 1999; Barber et al. 2004; Peterson 2004), predicting the potential ranges of invading species (garlic mustard and Russian olive, Peterson & Robins 2003; Asian longhorned beetles, Peterson et al. 2004; and mosquitoes, Levine et al. 2004), and examining interaction between range of invasive species and conservation (barred and spotted owls,

Peterson et al. 2003a) and bird migration (Peterson et al. 2003b). GARP is a powerful analytical tool that can be used for any taxon and allows predictions to be made up to a global scale at a fairly fine resolution. Drake & Bossenbroek (in press) state that GARP is relatively reliable and well validated compared with similar algorithms for predicting species occurrences. They cite the benefits of the algorithm as including the need for as few as five factors (Peterson & Cohoon 1999), being insensitive to the lack of absence data (Stockwell & Peterson 2002a), converging using only c. 50 observations (Stockwell & Peterson 2002b), and use of a best subset technique to improve model selection and forecasting (Anderson et al. 2003).

There are several disadvantages to using this approach, however. Ecological niche modelling predicts potential range of the non-indigenous species, but cannot provide information about its impact. Also, ecological niche modelling assumes that niche parameters for which geospatial coverages available are those that limit the distribution of species, and that the realised niche of a species in its native range is similar to the fundamental niche of the species. Ecological niche modelling also requires extensive knowledge of the native range of the species and the availability of geospatial and ecological data sets for the desired variables. Although the availability of geospatial data sets continues to improve, there is a need to develop and make available more coverages to continue to improve the utility and power of algorithms such as GARP for predicting potential ranges and other aspects of the ecology of invasive species.

Drake & Bossenbroek (in press) overcame the lack of geological geospatial coverages for the native range of zebra mussels in Europe and Asia and were able to use GARP to predict the potential range of zebra mussels in North America based on their current distribution in North America ($n = 2416$ observations). They were able to do so by assuming that zebra mussels have had access to all regions of the United States, based on their slow rate of expansion since 1993. Although this method requires more data on the invaded range of the species, Drake & Bossenbroek (in press) were able to show that virtually the same range prediction would be made with half of the distributional observations. To date, applications of GARP have been limited primarily to terrestrial ecosystems. Besides Drake & Bossenbroek (in press), the only other published account of using GARP for invasive aquatic species that I located was the modelling of the potential distribution of *Hydrilla* in Peterson et al. (2003a). There is ongoing work in the United States, however, using habitat niche modelling to predict the potential ranges of over 50 freshwater fishes (J. Williams, U.S. Geological Survey pers. comm.). The limited use of GARP and similar algorithms to understand and predict aspects of the invasion of aquatic species is probably caused by the lack of geospatial data for these systems. Currently, climatic and geological data must often be used as surrogates for variables in the aquatic environment.

USE OF SPECIES CHARACTERISTICS AND ALGORITHMS TO PREDICT POTENTIAL INVADING SPECIES

Applications of both of the species screening methods described herein are dominated to date by species invasions in terrestrial ecosystems. One explanation for this discrepancy may be based on the differential histories of management and control of invasive species in aquatic and terrestrial habitats. The desire to protect and enhance crop and pasture productivity has led to a long history of weed and crop pest control. In the United States, for example, the Western Society of Weed Science was formed in 1938, followed by the Southern Weed Science Society in 1947. To date, there is no professional society in the United States focused on aquatic invasive species (although the International Conference on Aquatic Invasive Species which held its 13th annual conference in 2004 is a venue dedicated to this topic). The direct economic benefit of protection of crop and livestock forage, estimated to be US$33 billion annually (Pimentel et al. 2000), has driven the development of prevention and control methodologies. Sociological and political will exists in the United States for continued research and development for the control of terrestrial weeds. Aquatic habitats are less visible and economic benefits derived from these ecosystems are often more indirect and for non-market ecosystem services. In addition, sociological and political will has diminished in the United States for the widespread use of chemical or biological control in aquatic ecosystems. Thus, pressure for making progress in prevention, control, and management of invasive species has been greater for terrestrial than for aquatic habitats. Another explanation for greater use of species screening systems in terrestrial versus aquatic ecosystems may be the availability of relevant data: whether regarding species

characteristics or geospatial in nature, data for terrestrial ecosystems are often more available than for corresponding data in aquatic ecosystems.

There are two patterns emerging in the literature that should be mentioned in regard to their relevancy to species screening methodologies: the correlation between propagule pressure and establishment and biotic resistance as a determinant of invasion success. Several studies have shown that the likelihood of successful introduction increases with propagule pressure (or the number of individuals introduced and the number of introduction events) (Kolar & Lodge 2001; Duncan et al. 2003; Rouget & Richardson 2003). Differences in propagule pressure would not directly affect the predictions made from species screening models as presented here. The higher the propagule pressure, the more likely it would be that an introduced species would be introduced into a habitat amenable to survival given the potential range identified through modelling (or the more likely the species would have suitable characteristics).

VALUE OF PREDICTION

Assuming that potentially invading species can be identified, either by the methods described herein or by others, what is the benefit of doing so? First, identifying potentially invasive species before their importation and introduction provides the opportunity to block entry for those high-risk species by way of legislation and regulation. Second, for those non-indigenous species already established, particularly those in the early stages of invasion, identifying those that pose the greatest risk allows for the timely allocation of monetary and personnel resources, and the development of species-specific management strategies for eradication, containment, or control.

Neither of the approaches described here to predict invading species are perfect; errors would inevitably be made if management decisions were based solely upon any single type of model predictions. Given this reality, what is the value of using modelling techniques to predict invading species? Any technique used to screen species for importation or any method used to ascertain high priority species for focused management will be imperfect. Mistakes will be made whether decisions are based wholly on qualitative or quantitative assessments. Some of the benefit of using quantitative modelling approaches to characterise potential invaders are realised in increased transparency, reliability, and tractability, and decreased opportunity for subjective influence. The greatest benefit, however, of using quantitative approaches for identifying potentially invading species is knowing the probability of error associated with model predictions. Understanding that a model has a 5% or 10% chance of inaccurate predictions provides managers and policy makers with more information on which to base decisions rather than a qualitative assessment or indicator without knowledge of probable error. Given that after establishment, few invading species are successfully eradicated, and that eradication, containment, and control are economically costly, prudent management and policy decisions should be based on the predictions and recommendations of all screening resources—including expert opinion, qualitative and quantitative predictions, and new screening tools as they become available.

ACKNOWLEDGMENTS

REFERENCES

Anderson, R. P.; Lew, D.; Peterson, A. T. 2003: Evaluating predictive models of species' distributions: criteria for selecting optimal models. *Ecological Modelling 162*: 211–232.

Animal and Plant Health Inspection Service, U.S. Department of Agriculture 2002: Weed-initiated pest risk assessment guidelines for qualitative assessments, version 5.2. Plant Protection and Quarantine, Riverdale, Maryland.

Barber, B. R.; Martinez-Gomez, J. E.; Peterson, A. T. 2004: Systematic position of the Socorro mockingbird *Mimodes graysoniii*. *Journal of Avian Biology 35*: 195–198.

Champion, P. D.; Clayton, J. S. 2000: Border control for potential aquatic weeds—Stage 1 weed risk model. *Science for Conservation 141*. Wellington, New Zealand, Department of Conservation.

Daehler, C. C.; Denslow, J. S.; Ansari, S.; Kuo, H. C. 2004: A risk-assessment system for screening out invasive pests plants from Hawaii and other Pacific Islands. *Conservation Biology 18*: 360–368.

Drake, J. M.; Bossenbroek, J. M. 2004: The potential distribution of zebra mussels (*Dreissena polymorpha*) in the U.S.A. *BioScience*, in press.

Duncan, R.P.; Blackburn, T.M.; Sol, D. 2003: The ecology of bird introductions. *Annual Review of Ecology, Evolution, and Systematics 34*: 71–98.

Grinnell, J. 1904: The origin and distribution of the chestnut-backed chickadee. *Auk 21*: 364–382.

Grinnell, J. 1917: Field tests of theories concerning distributional control. *American Naturalist 51*: 115–128.

Groves, R. H.; Panetta, F. D.; Virtue, J. G. 2001: Weed risk assessment. Collingwood, Victoria, Australia, CSIRO Publishing.

Hayes, K. R.; Sliwa, C. 2003: Identifying potential marine pests—a deductive approach applied to Australia. *Marine Pollution Bulletin 46*: 91–98.

Kolar, C. S.; Lodge, D. M. 2001: Progress in invasion biology: predicting invaders. *Trends in Ecology and Evolution 19*: 199–204.

Kolar, C. S.; Lodge, D. M. 2002: Ecological predictions and risk assessment for alien species. *Science 298*: 1233–1236.

Levine, R. S.; Peterson, A. T.; Benedict, M. Q. 2004: Geographic and ecologic distributions of the *Anopheles gambiae* complex predicted using a genetic algorithm. *American Journal of Tropical Medicine and Hygiene 70*: 105–109.

Lodge, D. M. 1993: Biological invasions: lessons for ecology. *Trends in Ecology and Evolution 8*: 133–137.

Novacek, M. J.; Cleland, E. E. 2001: The current biodiversity extinction event: scenarios for mitigation and recovery. *Proceedings of the National Academy of Sciences 98*: 5466–5470.

Ortega-Huerta, M. A.; Peterson, A. T. 2004: Modelling spatial patterns of biodiversity for conservation prioritization in north-eastern Mexico. *Diversity and Distributions 10*: 39–54.

Payne, K.; Stockwell, D. R. B. 2002: GARP modelling system user's guide and technical reference. San Diego, United States, San Diego Supercomputing Center. (Online at: http://biodi.sdsc.edu/Doc/GARP/Manual/manual.html)

Peterson, A. T. 2003a: Subtle recent distributional shifts in great plains bird species. *Southwest Naturalist 48*: 289–292.

Peterson, A. T. 2003b: Projected climate change effects on Rocky Mountain and Great Plains birds: generalities of biodiversity consequences. *Global Change Biology 9*: 647–655.

Peterson, A. T. 2004: The speciation and biogeography of birds. *Science 304*: 1249–1249.

Peterson, A. T.; Cohoon, K. P. 1999: Sensitivity of distributional prediction algorithms to geographic data completeness. *Ecological Modelling 117*: 159–164.

Peterson, A. T.; Robins, C. R. 2003: Using ecological-niche modelling to predict barred owl invasions with implications for spotted owl conservation. *Conservation Biology 17*: 1161–1165.

Peterson, A. T.; Vieglais, D. A. 2001: Predicting species invasions using ecological niche modelling: new approaches from bioinformatics attack a pressing problem. *BioScience 51*: 363–371.

Peterson, A. T.; Soberon, J.; Sanchez-Cordero, V. 1999: Conservatism of ecological niches in evolutionary time. *Science 285*: 1265–1267.

Peterson, A. T.; Sánchez-Cordero, V.; Soberón, J.; Bartley, J.; Buddemeier, R. H.; Navarro-Siguenza, A. G. 2001: Effects of global climate change on geographic distributions of Mexical Cracidae. *Ecological Modelling 144*: 21–30.

Peterson, A. T.; Papes, M.; Kluza, D. A. 2003a: Predicting the potential invasive distributions of four alien plant species in North America. *Weed Science 51*: 863–868.

Peterson, A. T.; Vieglais, D. A.; Andresean, J. K. 2003b: Migratory birds modeled as critical transport agents for West Nile Virus in North America. *Vector-Borne and Zoonotic Diseases 3*: 27–37.

Peterson, A. T.; Scachetti-Pereira, R.; Hargrove, W. W. 2004: Potential geographic distribution of *Anoplophora glabripennis* (Coleoptera: Cerambycidae) in North America. *American Midland Naturalist 151*: 170–178.

Pimentel, D.; Lach, L.; Zuniga, R.; Morrison, D. 2000: Environmental and economic costs of non-indigenous species in the United States. *BioScience 50*: 53–65.

Reichard, S. H.; Hamilton, C. W. 1997: Predicting invasions of woody plants introduced into North America. *Conservation Biology 11*: 193–203.

Ricciardi, A.; Rasmussen, J. B. 1998: Predicting the identity and impact of future biological invaders: a priority for aquatic resources management. *Canadian Journal of Fisheries and Aquatic Sciences 55*: 1759–1965.

Rouget, M.; Richardson, D. M. 2003: Inferring process from pattern in plant invasions: a semi-mechanistic model incorporating propagule pressure and environmental factors. *American Naturalist 162*: 713–724.

Sala, O. E.; Chapin, F. S. III; Armesto, J. J.; Berlow, R.; Bloomfield, J.; Dirzo, R.; Huber-Sanwald, E.; Huenneke, L. F.; Jackson, R. B.; Kinzig, A.; Leemans, R.; Lodge, D.; Mooney, H. A.; Oesterheld, M.; Poff, N. L.; Sykes, M. T.; Walker, B. H.; Walker, M.; Wall, D. H. 2000: Global biodiversity scenarios for the year 2100. *Science 287*: 1770–1774.

Smith, C. S.; Lonsdale, W. M.; Fortune, J. 1999: When to ignore advice: invasion predictions and decision theory. *Biological Invasions 1*: 89–96.

Stockwell, D. R. B. 1999: Genetic algorithms II. *In*: Fielding, A. H. *ed*. Machine learning methods for ecological applications. Boston, Kluwer Academic Publishers. Pp. 123–144.

Stockwell, D. R. B.; Noble, I. R. 1992: Induction of sets of rules from animal distribution data: a robust and informative method of data analysis. *Mathematical Computer Simulations 32*: 249–254.

Stockwell, D. R. B.; Peters, D. 1999: The GARP modelling system: problems and solutions to automated spatial prediction. *International Journal of Geographical Information Science 13*: 143–158.

Stockwell, D. R. B.; Peterson, A. T. 2002a: Controlling bias in biodiversity data. *In*: Scott, J. M.; Heglund, P. J.; Morrison, M. L.; Haufler, J. B.; Raphael, M. G.; Wall, W. A.; Samson, F. B. *ed*. Predicting species occurrences: issues of accuracy and scale. Washington, Island Press. Pp. 537–546.

Stockwell, D. R. B.; Peterson, A. T. 2002b: Effects of sample size on accuracy of species distribution models. *Ecological Modelling 148*:1–13.

Thomas, C. D.; Cameron, A.; Green, R. E.; Bakkenes, M.; Beaumont, L. J.; Collingham, Y. C.; Erasmus, B. F. N.; de Siqueira, M. F.; Grainger, A.; Hannah, L.; Hughes, L.; Huntley, B.; van Jaarsveld, A. S.; Midgley, G. F.; Miles, L.; Ortega-Huerta, M. A.; Peterson, A. T.; Philips, O. L.; Williams, S. E. 2004: Extinction risk from climate change. *Nature 427*: 145–148.

Williamson, M. 1996: Biological Invasions. *Population and Community Biology Series 15*. London, Chapman & Hall.