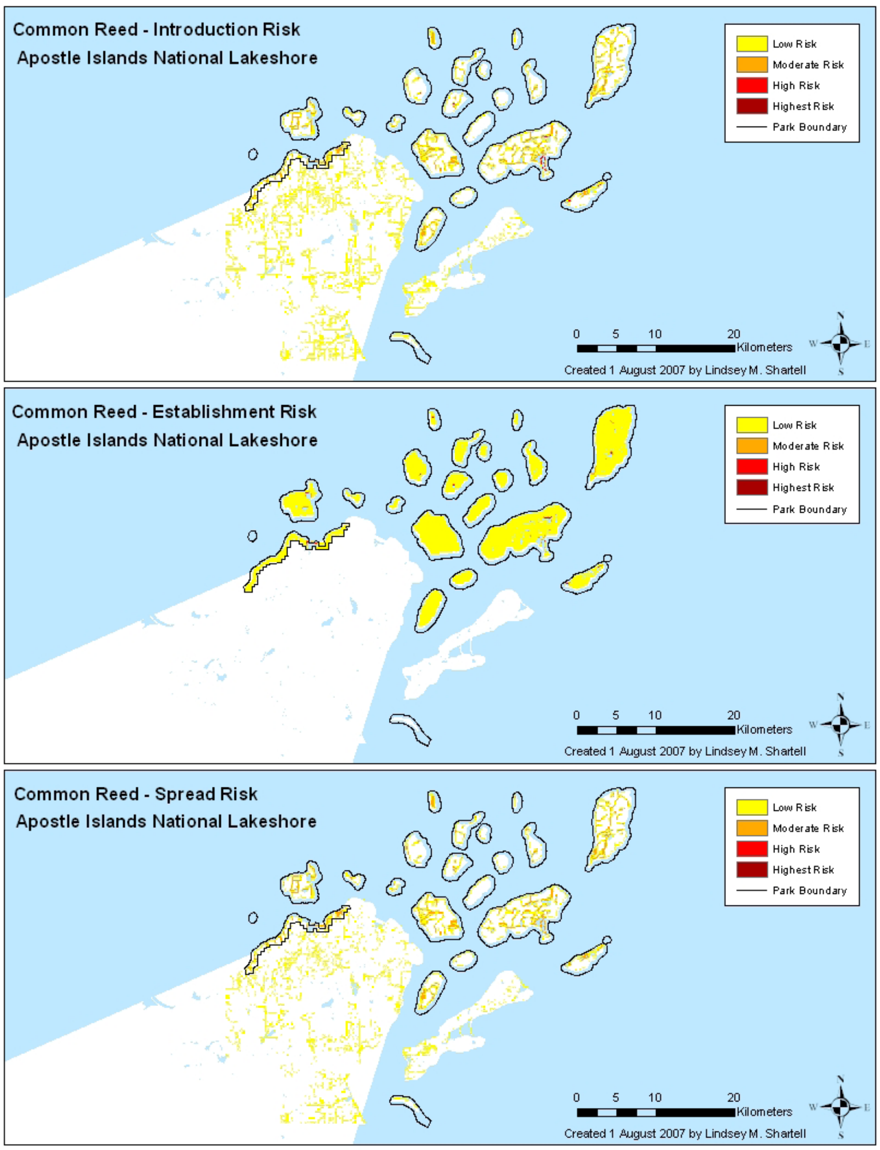
National Park Service
U.S. Department of the Interior

Great Lakes Network Office

Multi-Criteria Risk Models for Invasive Exotic Plant Species within the Great Lakes Network of the U.S. National Park Service

Great Lakes Network Report 2009/06



ON THE COVER
The final model for common reed (*Phalaris arundinacea*) at Apostle Islands National Lakeshore showing the introduction, establishment, and spread risk.

Multi-Criteria Risk Models for Invasive Exotic Plant Species within the Great Lakes Network of the U.S. National Park Service

Great Lakes Network Report GLKN/2009/06

Andrew J. Storer, Linda M. Nagel, Christopher R. Webster, Michael D. Hyslop, Lindsey M. Shartell

School of Forest Resources and Environmental Science
Michigan Technological University
1400 Townsend Dr.
Houghton, MI 49931

April 2009

This report is available from the Great Lakes Inventory & Monitoring Network website (http://science.nature.nps.gov/im/units/glkn/index.cfm).

Please cite this publication as:

Storer, A. J., L. M. Nagel, C. R. Webster, M. D. Hyslop, L. M. Shartell. 2009. Multi-criteria risk models for invasive exotic plant species within the Great Lakes Network of the U.S. National Park Service. Great Lakes Network Report GLKN/2009/06.

April 2009

Multi-Criteria Risk Models for Invasive Exotic Plant Species within the Great Lakes Network of the U.S. National Park Service

Final Report

Principal Investigator: **Andrew J. Storer**, Associate Professor
School of Forest Resources and Environmental Science
Michigan Technological University
1400 Townsend Drive
Houghton, MI 49931
Phone: 906 487-3470
Fax: 906 487-2915
Email: storer@mtu.edu

Co-Investigators: **Linda M. Nagel**, Associate Professor
School of Forest Resources and Environmental Science
Michigan Technological University
Phone: 906 487-2812
Email: lmnagel@mtu.edu

Christopher R. Webster, Assistant Professor
School of Forest Resources and Environmental Science
Michigan Technological University
Phone: 906 487-3618
Email: cwebster@mtu.edu

Michael D. Hyslop, GIS Analyst
School of Forest Resources and Environmental Science
Michigan Technological University
Phone: 906 487-2308
Email: mdhyslop@mtu.edu

Graduate Student: **Lindsey M. Shartell**
School of Forest Resources and Environmental Science
Michigan Technological University
Email: lmsharte@mtu.edu

Introduction

The National Park Service (NPS) uses the annual monitoring of "Vital Signs" to detect ecosystem change within the National Parks. They define a Vital Sign as "a physical, biological, chemical element or process that: indicates the health of a park ecosystem, responds to natural or anthropogenic stresses in a predictable or hypothesized manner, or has high value to the park or the public" (Route and Elias 2007). Vital Signs guide the NPS by indicating ecological changes that create problems and will require management or further research (Route and Elias 2007). Invasive exotic plants are a clear indicator of ecosystem change. They have the ability to impair the long-term health of a natural ecosystem by altering species composition, resource availability, structure, and function, among other things. The NPS ranks plant and animal exotics as a highest priority Vital Sign for the Great Lakes Network (Route and Elias 2007).

Monitoring for invasive plants in the National Parks is a daunting task. The Great Lakes Network is made up of nine Parks totaling 471,264 hectares. The time and resources needed to accomplish in-depth monitoring to detect individual or small invasions or in such large area are often not available. The use of predictive models to would assist with monitoring and management. The models would indicate where monitoring efforts should be focused, and when compared with known invasions, which invaded sites should be given priority for treatment. Similar projects using predictive modeling have been taken on with success. At Yosemite National Park, the potential invasion pattern was predicted for a set of non-native species, and the predictions were used to target monitoring and control efforts to high-priority areas within the Park (Underwood *et al.* 2004).

The overall goal of this research project was to produce multi-criteria risk models that utilize geographic information system (GIS) data to determine the areas at greatest risk for invasive plant species in the Great Lakes Network. This work provides a basis for monitoring invasive plants to detect ecosystem change as part of the Vital Signs monitoring within National Parks. The resulting risk maps can be used as a resource to assist the management of invasive plants through improved monitoring and control efforts. The main objectives were to: 1) obtain and/or create GIS data for nine National Parks and ten target invasive plants, 2) develop multi-criteria risk models that predict invasion at three phases: introduction, establishment, and spread, and 3) test and refine the models using known invasions and field sampling.

Methods

Region of Study

The National Park Service divides its Parks into 32 management Networks. This study was focused on the Great Lakes Network (GLKN). The GLKN manages nine National Parks: Apostle Islands National Lakeshore (APIS), Grand Portage National Monument (GRPO), Indiana Dunes National Lakeshore (INDU), Isle Royale National Park (ISRO), Mississippi National River and Recreation Area (MISS), Pictured Rocks National Lakeshore (PIRO), Saint Croix National Scenic River (SACN), Sleeping Bear Dunes National Lakeshore (SLBE), and Voyageurs National Park (VOYA). These parks encompass a total area of 471,264 ha, and are located in the Great Lakes states of Indiana, Michigan, Minnesota, and Wisconsin (Figure 1). This project focuses on four common habitat types found within the Parks: coastal dune, marsh/bog, woodland, and grassland.

Target Invasive Plants

Ten exotic species were identified as target invasive plants: baby's breath (*Gypsophila paniculata* L.), common buckthorn (*Rhamnus cathartica* L.), common reed (*Phragmites* spp.), garlic mustard (*Alliaria petiolata* [Bieb] Cavara and Grande), glossy buckthorn (*Frangula alnus* P. Mill.), honeysuckle (*Lonicera* spp.), leafy spurge (*Euphorbia esula* L.), multiflora rose (*Rosa multiflora* Thunb. ex Murr.), purple loosestrife (*Lythrum salicaria* L.), and spotted knapweed (*Centaurea biebersteinii* DC.) (Table 1, Appendix 1). These ten plants are thought to present the highest risk to natural ecosystems and pose the greatest challenge to management. The species identified are diverse as each invades different habitat types and has unique biological and environmental requirements. However, all of the ten species can be found in one or more of the common habitat types associated with the GLKN.

Model Development

To develop the multi-criteria risk models, the invasion process was broken into three phases: introduction, establishment, and spread. Definitions of each phase were developed to determine which factors would be used to predict invasion at each phase. The definitions were adapted from several previous definitions (see Williamson 1996, Richardson *et al*. 2000, Kolar and Lodge 2001, Sakai *et al*. 2001). The invasion process was considered to follow an S-shaped population growth curve (Figure 2). Introduction was defined as the arrival of a species in an area where it was not currently present. Introduction involves the dispersal of seeds to an area as well as successful germination of the seeds and survival of seedlings. Establishment follows introduction and was defined as the development of a free-living, reproducing population of a species. Establishment requires the survival of plants past the seedling stage, combined with successful reproduction. Most factors affecting establishment are environmental characteristics of the landscape, but other factors such as fire regime, human land use, and disturbance play roles as well. Spread was defined as the increase in population size of an established population. Spread involves short- and long-distance dispersal of seeds and survival and growth of the established population. A population can grow by increasing the number of individuals or by increasing the size of the area invaded. The potential for population spread can be greater in large areas of connected suitable habitat or on disturbed land.

The risk of invasion for a given area varied by phase since each phase had different requirements. An initial list of risk factors affecting invasion was created for each phase based on a literature review for each species (for bibliography see Appendix 2). The preferences and requirements of each invasive plant were determined for each risk criteria. Important criteria were incorporated into the models as parameters. Criteria were unimportant if there was little variation within Parks, or if the invasive plant had no requirements or preferences discernable from the literature for that criterion. Furthermore, to be used in the model, the criteria needed to have obtainable or creatable spatial data across the nine National Parks. The NPS provided all existing GIS data layers for the National Parks. This generally consisted of spatial data layers for Park boundary, hydrology, hypsography, roads, railroads, trails, vegetation, recreation features, land use, and disturbance features. Additionally, Soil Survey Geographic (SSURGO) data was obtained from the Natural Resources Conservation Service. When not available, coarser scale State Soil Geographic (STATSGO) data was used. From the soil data, drainage and average pH were extracted for the soil types present. All spatial data were converted to or created in coverage or shapefile format, and included relative attribute data indicating source, date, and species if applicable. A common datum and coordinate system was used within each Park.

An additional parameter used in the model was the presence of known invasions. Most of the National Parks monitor invasive plants and develop GIS data each year based on known locations of target species. The NPS provided any GIS data available for the target invasive plants. These data occasionally included information on areas where invasions occurred and had been treated. Such data were included as known locations since these sites were known to be suitable for invasion, and because it was not possible to determine whether treatment had been successful. Every species had invasion data for at least one Park and many species were recorded at multiple parks (Table 2). Only one park, GRPO, did not have any GIS data on invasive plant presence. Despite this, it is likely that some of these plants are present at this site. The known invasion data were used after model development to test the initial accuracy of the models and did not influence criteria selection or the assignment of risks to parameters.

Using the lists of significant criteria, a unique risk model was developed for each invasive plant using ModelBuilder within ArcGIS 9.2 (ESRI 2006, Figure 3). The information from the literature search guided the assignment of risks to parameter levels. Risk values ranged from zero to ten, with zero being equivalent to no risk and ten equivalent to high risk. When possible, a typical curve shape, such as a normal or sigmoid curve, was applied using information on when risk begins, peaks, and ends for a particular parameter. For categorical data, such as vegetation, risk was determined by considering the suitability of each category to both invasion and survival. Assigned risks varied by the phase of invasion being considered. For categorical data, reclassification tables were created for every species and Park combination, which were entered into the model as a parameter. For example, vegetation data was grouped by cover type and had to be reclassified into the correct risk value for each category. The appropriate risk values for categorical data were determined using the information gained from the literature search, and were altered if needed during model adjustment. Each parameter in the model was assigned a rank and a confidence level. These values were weighted, with the rank having three times more importance than the confidence level, to calculate an influence for each parameter. The influence was the percent weight calculated for the parameter, with the sum of all influences equal to 100%. Within the models the influence was entered into the weighted overlay table for each phase of invasion. During model development, the models were run using the data collected and compiled for each Park to find the optimal model structure.

Model Analysis and Adjustment

Once completed, the models were run for each of the 90 species and Park combinations. The output for each combination consisted of three 10 m x 10 m raster grids, one for each phase of invasion. The raster contained risk values associated with each pixel of the Park and the surrounding area. The risks range from 0 to 10, with 10 representing highest risk. The invasive plant data provided by the Parks were used to assess the initial accuracy of the models. For each species at each Park where invasion data were available (Table 2), the risk ratings for areas with a known invasion were extracted from the model output. The percent of pixels with a known invasion and a high risk rating (risk ≥ 7) was calculated. This was considered to be a measure of the accuracy of the model to correctly predict a high risk for areas both suitable to invasion and with a high probability of invasion. A model was considered to be sufficiently accurate if it correctly assigned a high risk rating to at least 70% of the known invaded pixels. This value was selected as a general rule to guide model adjustment and did not carry any statistical significance. An overall accuracy was calculated for all phases of invasion, all Parks, and all species by weighting the individual Park and species results by the number of invaded pixels and averaging this over the three phases of invasion. The results were further compared at various scales based on overall accuracy, accuracy by species, accuracy by Park, and accuracy by invasion phase. Models with insufficient accuracy were first examined to look for unnoticed errors in model structure and function. Then the presence data were compared to each of the model parameters to determine if incorrect or absent parameter data could be the cause of decreased model accuracy. When possible, missing or inaccurate data were replaced. Finally, the weights and risks assigned to the parameters were adjusted to obtain a consistent accuracy across all nine Parks and invasion phases for each species model.

Field Sampling

During the summer 2006 field season initial ground-truthing took place within two National Parks, INDU and SLBE. This consisted of sampling randomly-generated points within these areas for the presence and abundance of the ten target invasive plants. The points were generated within the boundaries of each Park, excluding open water, using the Random Point Generator (Sawada 2002). Based on an estimate of required time and effort, an initial set of 75 points was generated for each Park. The points were spatially stratified, with the sample size based on the area of contiguous Park sections. The points were also assessed to ensure that each of the four common habitat types found in the Parks of the GLKN were represented. The GPS coordinates of each point were downloaded onto a Garmin GPS Map 76 unit. Using maps of the random points a field crew of three navigated to within 20 m of each point. In some rare cases points were inaccessible to within 20 m and sampling was done as close to the random point as possible. This was most often due to areas being restricted by fencing or open water. At each point, the exact coordinates were recorded as well as the accuracy of the GPS unit. A decrease in accuracy was noticeable under full cloud cover or dense canopy cover. The locations of the random points were later adjusted to the actual coordinates recorded in the field.

At each point a 40 m x 40 m plot was assessed for the presence of the ten invasive plants. The random point served as the center of the plot, and a compass was used to align the plot with the cardinal directions. The plot was divided into sixteen 10 m x 10 m blocks. Within these blocks, the presence as well as percent cover was recorded for each of the ten invasive plants. Percent cover was based on a visual assessment and was ranked on the following scale: rare (R, <1% cover), occasional (O, 1-10% cover), common (C, 10-25% cover), abundant (A, 25-50% cover), dominant (D, 50-100% cover). Additional notes were taken describing the site characteristics including presence of roads, railroads, and water, general vegetation type, presence of target invasive plants just outside of a plot or while traversing to a plot, and evidence of invasive plant treatment or removal. This information was used to evaluate incorrect predictions at single points to determine the cause of error.

During the summer 2007 field season, additional random points were sampled at SLBE to more intensively assess the accuracy of the final versions of the models. In order to do so, an appropriate sample size was determined using a standard sample size formula (Levy and Lemeshow 1999). Considering the available time and effort, a confidence level of 0.90 and a relative error of 0.20 were selected. This suggested a sample size of approximately 168. The points were generated using the same methods as used previously, and stratified spatially with equal amounts given to the three sections of the Park. The points were assessed to ensure that each of the four common habitat types and varying levels of risk for each species were represented. The field methods were similar to the previous sampling effort, although only one measure of presence and percent cover was taken for the entire 40 m x 40 m plot. The occurrence data were used to assess the accuracy of the models following adjustment. The cover data were used to assess the differences predicted for the three phases of invasion.

The data collected during the summer 2006 field season were assessed using the output from the adjusted models. From this the models were adjusted, if needed, and the final models were reassessed and validated using the summer 2007 field data. For each sample point, the predicted risk value was extracted from the output for each species. The value was interpolated so that it also represented the eight pixels surrounding that of the sampling point. The percentage of points with a high risk and occurrence of a known invasion was calculated for each species. The species accuracies were combined to create an average accuracy for each phase by weighting the species accuracies by the number of observations. These values were then averaged over the three phases to calculate an overall accuracy. Models that correctly predicted high risks for known invasions were considered to show a high degree of accuracy. To evaluate the differences in risk among phases, risk values were extracted using interpolation for each phase. The change in risk across phases for each species was assessed qualitatively. Points with a rare or occasional cover rating were considered to be in the introduction phase, points with a common cover rating were considered to be in the establishment phase, and points with an abundant or dominant rating were considered to be in the spread phase.

Risk Map Creation

The final outputs produced by each model were used to create risk maps by species and Park highlighting the areas at greatest risk for the three phases of invasion. The risk maps indicated the areas with risk ratings from seven to ten using a yellow, orange, red, and burgundy color scheme. An overall risk map was also created for each Park in order to determine if any areas of the Park were particularly more at risk for every species. These maps could also guide multi-species monitoring if it were not possible to monitor each species individually. The overall risk maps were created by averaging the establishment risk for all ten species. The areas with the greatest combined risk value were highlighted in the same color scheme as the species-specific risk maps.

Results

Multi-Criteria Risk Models

All ten multi-criteria risk models utilized the parameters of Park boundary, transportation (distance to roads, railroads, and trails), hydrology, vegetation type, disturbance features, soil drainage, and connectivity of suitable habitat. Invasive plant presence was also used in each of the models. Distance to the nearest known invasion was a parameter for the spread phase. The models for common reed, garlic mustard, glossy buckthorn, honeysuckle, and leafy spurge also utilized soil pH as a model parameter. Transportation and hydrology were usually combined to form one parameter, which represented potential pathways for dispersal. Each phase of invasion and species model made use of a different set of the parameters, and each parameter had a unique influence weight (Table 3). Introduction was based on Park boundary, dispersal, vegetation type, disturbance, soil drainage, and soil pH (when applicable). Establishment took into account the assigned introduction risk, as well as the parameters vegetation type, disturbance, soil drainage, and when important to a given species, soil pH, hydrology, and/or transportation. Spread included the assigned establishment risk, dispersal, connectivity of suitable habitat, and the locations of known invasions.

Park Data Analysis

There were 32 combinations of species and Parks with known invasion data. Preliminary assessment of model accuracy using the data provided by the Parks showed that 75.8% of pixels with the presence of a given invasive plant were correctly assigned a high risk rating (risk ≥ 7) by the models. This was an average of the three phases of invasion, for which the individual results were 60.1% for introduction, 80.3% for establishment, and 87.1% for spread. Breaking the data down by species and Park, there was a range of results from 0% to 100% (Table 4). Following final adjustments, the overall accuracy increased to 86.4%. This was based on a total of 130,718 invaded pixels across all ten species and eight Parks with known invasions. Broken down by phase, a high risk was correctly assigned for 77.5% of invaded pixels for introduction, 90.4% for establishment, and 91.2% for spread. Two individual accuracies were lower than the target accuracy of 70% (spotted knapweed introduction and establishment at VOYA) but this was based on a sample size of only two, so was not considered an error of the model. All other individual accuracies ranged between 70.2% and 100% (Table 5).

Field Sampling

During field work in summer 2006, 75 points were sampled at INDU, and 76 points were sampled at SLBE (Figures 4 and 5). Eight of the ten target invasive plants were identified at

INDU and seven were identified at SLBE (Table 6). A range of abundance levels was encountered across both Parks. At INDU multiflora rose and garlic mustard were identified most frequently, found at 25 and 21 points respectively, and with the greatest abundances. At SLBE spotted knapweed was most frequent, being found at 14 points. During additional field work at SLBE in summer 2007, 162 points were sampled (Figure 6). Six of the ten invasive plants were identified during this sampling (Table 6). Spotted knapweed was again most frequent, found at 42 points. In 2007, honeysuckle and leafy spurge were found more often and with higher abundance than in 2006. Common reed and purple loosestrife were encountered in 2006 but were not found in 2007. Garlic mustard was absent from the random sample points in 2006, but occurred at three of the 2007 random sample points.

The initial overall model accuracy by Park for the 2006 field data was 64.1% for INDU and 78.4% for SLBE. Following model correction and adjustments, the model accuracy increased for both Parks. INDU had an overall model accuracy of 86.8%, while SLBE had an overall model accuracy of 87.0%. By phase the results were 89.7% for introduction, 82.8% for establishment, and 87.9% for spread at INDU, and 91.7% for introduction, 91.7% for establishment, and 77.8% for spread at SLBE. This was based on 58 observations of target invasive plants at INDU and 36 at SLBE. For the 2007 SLBE data the overall model accuracy was 83.3%. Based on phase, the results were 88.8% for introduction, 85.0% for establishment, and 76.3% for spread. This was based on 80 observations of target invasive plants.

Model Predictions and Risk Maps

The risk maps created from the model output (Appendix 3) offered the best glimpse at the level of invasion risk predicted across Parks. The risk maps for APIS showed garlic mustard and multiflora rose as having the greatest risk across the Park. At GRPO, common buckthorn, garlic mustard, honeysuckle, multiflora rose, and spotted knapweed risk maps all showed large portions of the Park at risk. However, GRPO is the smallest Park, covering only 287 hectares. At INDU the species showing the most extensive areas at risk were garlic mustard, honeysuckle, multiflora rose, and spotted knapweed. Common reed and honeysuckle had the greatest areas of risk at ISRO. At MISS, garlic mustard, honeysuckle, and spotted knapweed risk maps all showed the majority of the Park at risk. Multiflora rose and honeysuckle risk maps had the greatest areas at risk at PIRO. At SACN common buckthorn, garlic mustard, honeysuckle, and spotted knapweed showed the greatest areas at risk. The variety of habitat types at SLBE created a range of risks across the park for each species. The SLBE risk maps for baby's breath, glossy buckthorn, garlic mustard, honeysuckle, and spotted knapweed all showed large areas at risk. At VOYA the species showing the greatest areas of risk were common buckthorn, honeysuckle, multiflora rose, and spotted knapweed. In addition, purple loosestrife showed high risk around all bodies of water at VOYA.

The overall risk maps, which reflected all ten species, showed small areas at high risk and the majority of most Parks at low or moderate risk (Appendix 4). Overall risk maps for GRPO, INDU, MISS, and PIRO showed distinct areas of high risk, despite these making up only a small portion of the Parks. Overall maps for SACN and SLBE showed a few small areas of high risk, and overall maps for APIS, ISRO, and VOYA showed no noticeable areas of high risk.

Discussion

The multi-criteria risk models correctly identified invaded areas as high risk for introduction, establishment, and spread for each of the ten target invasive plants. The models also predicted high risks for areas identified as invaded through field sampling. The success of the models throughout initial testing, adjustments, and final assessment indicates that they will be of use to assist with monitoring invasive plant presence and managing known populations of these ten invasive plants. The accuracy results were similar to those obtained by Underwood *et al*. (2004), who correctly predicted 76% of invasive plant species presence at Yosemite National Park. Underwood *et al*. (2004) also noted that areas with a high level of human activity and disturbance, such as campgrounds, roads, and trails, were predicted to have higher probability of invasive plant occurrence, despite not being included as a factor in their model. The success of the models created in this project is due in part to the addition of these anthropogenic parameters.

The parameters selected for the multi-criteria risk models differed notably from other similar predictive models (Peterson *et al.* 2003, Dark 2004, Underwood *et al.* 2004). Human activity was stressed through the addition of transportation and disturbance features. Elevation was not included, but has been a common factor in other predictive models (Dark 2004, Underwood *et al*. 2004). It was not included here since it offered no limitation to plant invasion by the target invasive plants. This was also encountered by Peterson *et al*. (2003), who eliminated elevation from a predictive model for garlic mustard in California. Additionally, elevation may not have had an effect due to the lack of variation across the Great Lakes Region and within Parks. Other common factors were also excluded, such as climate, which had little variation across the region, and slope and aspect, which were eliminated due to insufficient evidence of association with the target invasive plants.

The accuracy for introduction was lower than that of establishment and spread during analysis of the known invasion data. This could be due to the stochastic nature of invasive plant introductions, making them more difficult to model. Propagule pressure is particularly difficult to measure and represent spatially as GIS data (Lockwood *et al*. 2005). In the analysis of the field data, the spread accuracy was generally lower than the spread accuracy for the Park data. This was expected, since distance to the nearest invasion was included as a parameter for the spread phase, and was based on the same dataset being used to analyze the model. This parameter is important, however, since it will create higher spread risks surrounding known invasions, where propagule pressure is expected to be highest, and where management is most needed.

Parks with a high density of roads generally showed a greater extent of area at risk than Parks with few roads. This was expected since roads provide pathways of dispersal, areas of disturbance, and in general, a habitat with ample light and water (Forman and Alexander 1998). Dark (2004) also found that roads played an important role in the locations of plant invasions. ISRO, which has only a few trails and no roads, had few areas with a high risk of invasion. INDU, MISS, and SACN are located in more urban locations and had grids of roads visible in the risk maps since the roads ranked higher in risk relative to the surrounding areas. It is likely correct to assume an overall lower risk for ISRO since there are fewer opportunities for dispersal and less disturbance, while it is accurate for INDU, MISS, and SACN to have higher risks due to high levels of disturbance and many opportunities for dispersal by humans.

The methods used for this project helped to avoid one common obstacle to predictive modeling: the amount of data points required to both create a model and validate its accuracy. By using a literature review to determine the environmental preferences of each invasive species, fewer data points were needed overall. This permitted an initial accuracy assessment of the models, model adjustment, and a final assessment of accuracy. During the literature review, information was collected from many sources, across diverse sites. As a result, the models were useful within different areas and not limited to the extent of known invasions. This allowed for risk assessment in National Parks where the invasive plant is not yet present.

The setup of the multi-criteria risk models allows for the parameter influences to be easily adjusted within the ArcGIS ModelBuilder program (ESRI 2006). It is also possible to manipulate the lower and upper limits of a parameter, as well as each level within this range. This will be useful for tailoring the model to the specific site where it is being applied. For example, at INDU, multiflora rose was planted along railroads and roads and is continuing to spread further along these routes. In the model, the influence of dispersal (roads, railroads, and trails) could be increased since it is strongly related to the presence of multiflora rose. The ease of tweaking the models also makes them easy to update as new research and management information becomes available.

The risk maps created from the model output will be useful for invasive plant management. However, predictive models should not be used as the only resource when making management decisions. The results are only predictions. The invasion of exotic species is difficult to predict accurately because of the complex relationship between predictors and the chance occurrences of dispersal and disturbance that facilitate introduction and establishment. More detailed information on ecosystem properties, such as species present, population sizes, and resource availability, are necessary to determine the probability of invasion in a specific area (Stohlgren and Schnase 2006). This information is not readily available as GIS data, and would be difficult to obtain over a large area. Interspecies interactions, such as competition, herbivory, and predation, are also important factors that are difficult to quantify and, therefore, are not included (Stohlgren and Schnase 2006). A clearer understanding of dispersal is also needed. Lockwood *et al*. (2005) stressed the need for propagule pressure to be incorporated into the predictive modeling of invasive species. Unfortunately, propagule pressure is difficult to quantify, as pathways and rates of introduction are complex and poorly understood (Lockwood *et al*. 2005). What is more, the degree of propagule pressure required to establish populations may have an association with disturbance and ecosystem properties, further confounding the ability to measure and utilize these aspects as predictors (Lockwood *et al*. 2005).

The process of predictive modeling inevitably has flaws and limitations. Error can occur during parameter risk estimation, or can arise from limitations in the input data. The predictions must be balanced between over- and under-fitting, which, when modeling invasive plants, is difficult to assess since their full invasion potential is unknown. The standard strategy for assessing actual accuracy is to test for statistical significance of randomly selected subsets of the data (Stockwell and Peters 1999). With invasive plants this is not possible since they have not reached their full potential distribution, and therefore sample points with the absence of an invasive plant may in reality be an optimal invasion site.

Predictive models are also limited by the quality of the data (e.g. GIS layers) being used. Data may not be complete due to missing information or from estimates and generalizations made during surveying, such as listing the vegetation type as the dominant species when a variety of other species exist at the site (Stohlgren and Schnase 2006). In this project, missing or incomplete data affected the ability of the models to correctly predict risk. For example, 4.68%, nearly 40 hectares, of SLBE had undefined soil pH. The precision and scale of the data should also be considered. In the case of some Parks, coarse scale STATSGO data had to be used rather than more detailed SSURGO data, since these data were not available for every area. The Parks affected by this were ISRO, GRPO, PIRO and the southern portion of SLBE. In the case of ISRO, the soil data only provided one soil type and therefore only one soil pH and soil drainage for the entire Park. This may have affected the ability of the model to predict areas of risk precisely, resulting in similar risk predictions across the Park. Once available, running the models with completed SSURGO data may provide more accurate and precise results.

Although the models can be adjusted to a specific Park, a multi-criteria risk model created for one specific Park may be more efficient and result in higher accuracy. Risk assessment focused on a particular species in a specific area has had more success than larger-scale, multiple species predictions (Stohlgren and Schnase 2006). Working with smaller areas allows for a more in-depth assessment of ecosystem properties and processes that affect the probability of invasion. The overall risk maps indicated that there is likely an association between the occurrences of these invasive plants. Areas along roads were clearly at high risk for all species at Parks such as INDU and MISS, and more than likely at risk at other Parks as well. However, this may not be the case for wetland species, which would rely more on the locations of wet habitat along lakes and streams. Grouping the target invasive plants by common habitat type and creating combined risk maps should provide improved results for monitoring and managing multiple species.

At many of the National Parks, the target invasive plants have already invaded, established populations, and begun to spread. Once established, control can be used to remove the plants or to stop the spread of the population either in size or to new locations. Time, effort, and funding are not always available to implement control methods for every population. In this case the best management option is to treat populations with the highest potential for spread and the greatest threat to ecosystems. By consulting the models and risk maps, known populations can be prioritized for treatment based on their risk of establishment and spread. When multiple species of invasive plants are present, combined species risk maps should be utilized.

Conclusions

Despite some limitations, the multi-criteria risk models developed in this project will be useful tools. Assessment of the models should be continued as they are utilized and adjusted. Three tasks were suggested by Underwood *et al*. (2004) to continue improving their predictive model for plant invasions at Yosemite National Park: rapid ground validation, known occurrence mapping, and experimental studies to understand the relationship between species occurrence and predictors. The models created for the Great Lakes Network were field tested at only two of the nine National Parks, Indiana Dunes National Lakeshore and Sleeping Bear Dunes National Lakeshore. Further ground-truthing would be beneficial toward understanding the function and accuracy of the predictive models. Mapping new locations of the target invasive plants will also be important for assessing the accuracy of the model predictions over time. This includes identifying locations where invasive plants have been removed or treated, as this will affect the accuracy when ground-truthing the models (Higgins *et al.* 1999).

Monitoring National Parks for ecosystem change is a difficult but important task, since these areas protect threatened and endangered plants, animals, and ecosystems, show historical conditions, and provide recreation and enjoyment to visitors. The models will be able to assist the NPS with the management of these ten invasive species and can be easily adapted to fit other areas. Used as one of many resources, the risk maps created from the model output can assist with management of invasive plants through improved monitoring and control efforts by focusing on areas at highest risk for individual invasive plants or for groups of similar species.

Literature Cited

Dark, S. J. 2004. The biogeography of invasive alien plants in California: an application of GIS and spatial regression analysis. Diversity Distrib 10:1-9.

Devine, B. 1999. Clear-cut mission. Communities unite to free native landscapes from the grip of invasive species. Nature Conservancy 49:12-17.

ESRI. 2006. ArcGIS 9.2. Badlands, CA.

Forman, R. T. T. and L. E. Alexander. 1998. Roads and their major ecological effects. Ann Rev Ecol Syst 29:207-231.

Higgins, S. I., D. M. Richardson, R. M. Cowling, and T. H. Trinder-Smith. 1999. Predicting the landscape-scale distribution of alien plants and their threat to plant diversity. Conserv Biol 13:303-313.

Kolar, C. S. and D. M. Lodge. 2001. Progress in invasion biology: predicting invaders. Trends Ecol Evol 16:199-204.

Levy, P. S. and S. Lemeshow. 1999. *Sampling of Populations*. Wiley-Interscience; New York.

Lockwood, J. L., P. Cassey, and T. Blackburn. 2005. The role of propagule pressure in explaining species invasions. Trends Ecol Evol 20:223-228.

Peterson, A. T., M. Papes, and D. A. Kluza. 2003. Predicting the potential invasive distributions of four alien plant species in North America. Weed Sci 51:863-868.

Petrella, S., N. Shutt, and R. McNeill. 1999. A Survey of Invasive Exotic Plants in Seney National Wildlife Refuge. U. S. Fish and Wildlife Service Puplication.

Richardson, D. M., P. Pysek, M. Rejmanek, M. G. Barbour, F. D. Panetta, and C. J. West. 2000. Naturalization and invasion of alien plants: concepts and definitions. Diver Distrib 6:93-107.

Route, B. and J. Elias (eds). 2007. Long-term Ecological Monitoring Plan: Great Lakes Inventory and Monitoring Network. National Park Service, Fort Collins, Colorado.

Sakai, A. K., F. W. Allendorf, J. S. Holt, D. M. Lodge, J. Molofsky, K. A. With, S. Baughman, R. J. Cabin, J. E. Cohen, N. C. Ellstrand, D. E. McCauley, P. O'Neil, I. M. Parker, J. N. Thompson, and S. G. Weller. 2001. The population biology of invasive species. Ann Rev Ecol Syst 32:305-332.

Sawada, M. 2004. Random Point Generator. Available from ESRI ArcScripts (http://arcscripts.esri.com/details.asp?dbid=12098).

Stockwell, D., and D. Peters. 1999. The GARP modeling system: problems and solutions to automated spatial prediction. Int J Geogr Inf Sci 13:143–158.

Stohlgren, T. J. and J. L. Schnase. 2006. Risk analysis for biological hazards: what we need to know about invasive species. Risk Anal 26:163-173.

Underwood, E. C., R. Klinger, and P. E. Moore. 2004. Predicting patterns of non-native plant invasions in Yosemite National Park, California, USA. Diver Distrib 10:447-459.

United States Forest Service, Hiawatha National Forest. 1998. Michigan's Upper Peninsula Weeds. Collaborative brochure.

Van Driesche, R. 2002. Introduction. *In* Van Driesche, R. Biological Control of Invasive Plants in the Eastern United States. USDA Forest Service Publication FHTET-2002-04.

Voss, E.G. 1985. Michigan Flora. A Guide to the Identification and Occurrence of the Native and Naturalized Seed Plants of the State. Part II. Dicots. Kingsport Press; Bloomfield Hills, Michigan.

Voss, E.G. 1996. Michigan Flora. A Guide to the Identification and Occurrence of the Native and Naturalized Seed Plants of the State. Part III. Dicots Concluded. Kingsport Press; Bloomfield Hills, Michigan.

Williamson, M. 1996. *Biological Invasions*. Chapman & Hall, London.

Table 1. The ten target invasive plants within nine National Parks in the Great Lakes Network and the code used to identify each.

Common Name	Scientific Name	Code
Baby's Breath	*Gypsophila paniculata*	BB
Common Buckthorn	*Rhamnus cathartica*	CB
Common Reed	*Phragmites* spp.	CR
Garlic Mustard	*Alliaria petiolata*	GN
Glossy Buckthorn	*Frangula alnus*	GB
Honeysuckle	*Lonicera* spp.	HS
Leafy Spurge	*Euphorbia esula*	LS
Multiflora Rose	*Rosa multiflora*	MR
Purple Loosestrife	*Lythrum salicaria*	PL
Spotted Knapweed	*Centaurea biebersteinii*	SK

Table 2. Known occurrence of the ten target invasive plants within the National Parks of the Great Lakes Network (X = present, - = not yet detected). See Tables 1 and 2 for Park and species codes.

National Park	BB	CB	CR	GM	GB	HS	LS	MR	PL	SK
APIS	-	-	-	-	-	-	-	-	-	X
GRPO	-	-	-	-	-	-	-	-	-	-
INDU	-	-	X	X	X	X	X	X	X	X
ISRO	-	-	-	X	-	-	-	-	-	X
MISS	-	X	-	X	-	X	X	-	-	X
PIRO	-	-	-	-	-	-	-	-	-	X
SACN	-	-	-	-	-	-	X	-	X	X
SLBE	X	-	X	X	-	X	X	X	X	X
VOYA	-	-	-	-	-	-	-	-	X	-

Table 3. Parameters and influence weights used in the multi-criteria models. (PB = Park Boundary, DP = Dispersal, SD = Soil Drainage, SP = Soil pH, DT = Disturbance, HD = Hydrology, TR = Transportation, VT = Vegetation Type, I = Introduction, E = Establishment, CN = Connectivity of Suitable Habitat, IS = Invasive Species Presence)

Species	Weighted Overlay Formula
Introduction	
Baby's Breath	0.28*DP + 0.28*SD + 0.27*VT + 0.09*DT + 0.08*PB
Common Buckthorn	0.26*DP + 0.25*SD + 0.25*VT + 0.14*DT + 0.10*PB
Common Reed	0.26*DP + 0.26*SD + 0.24*VT + 0.09*SP + 0.08*DT + 0.07*PB
Garlic Mustard	0.27*DP + 0.26*VT + 0.16*SD + 0.16*SP + 0.09*DT + 0.06*PB
Glossy Buckthorn	0.24*SD + 0.23*VT + 0.14*TR + 0.14*HD + 0.10*VT + 0.10*SP + 0.07*PB
Honeysuckle	0.26*DP + 0.26*VT + 0.16*SD + 0.16*SP + 0.08*DT + 0.08*PB
Leafy Spurge	0.25*DP + 0.24*VT + 0.24*SD + 0.11*SP + 0.09*DT + 0.07*PB
Multiflora Rose	0.33*DP + 0.31*VT + 0.19*SD + 0.09*PB + 0.08*DT
Purple Loosestrife	0.33*DP + 0.31*VT + 0.19*SD + 0.09*PB + 0.08*DT
Spotted Knapweed	0.32*DP + 0.31*VT + 0.19*SD + 0.09*DT + 0.09*PB
Establishment	
Baby's Breath	0.34*SD + 0.32*VT + 0.19*DT + 0.15*I
Common Buckthorn	0.34*VT + 0.34*SD + 0.16*DT + 0.16*I
Common Reed	0.33*SD + 0.32*VT + 0.14*SP + 0.11*DT + 0.10*I
Garlic Mustard	0.27*SD + 0.26*SP + 0.25*VT + 0.12*DT + 0.10*I
Glossy Buckthorn	0.26*SD + 0.26*VT + 0.26*SP + 0.13*HD + 0.09*I
Honeysuckle	0.25*SD + 0.25*SP + 0.24*VT + 0.14*DT + 0.12*I
Leafy Spurge	0.31*VT + 0.31*SD + 0.18*SP + 0.10*DT + 0.10*I
Multiflora Rose	0.34*VT + 0.32*SD + 0.19*DT + 0.15*I
Purple Loosestrife	0.37*VT + 0.22*SD + 0.17*HD + 0.13*DT + 0.11*I
Spotted Knapweed	0.35*SD + 0.34*VT + 0.16*I + 0.15*DT
Spread	
Baby's Breath	0.37*E + 0.22*DP + 0.19*CN + 0.12*DT + 0.10*IS
Common Buckthorn	0.36*E + 0.22*DP + 0.20*CN + 0.12*DT + 0.10*IS
Common Reed	0.37*E + 0.22*DP + 0.20*CN + 0.12*DT + 0.09*IS
Garlic Mustard	0.38*E + 0.21*CN + 0.17*DP + 0.15*DT + 0.09*IS
Glossy Buckthorn	0.36*E + 0.22*DP + 0.20*CN + 0.12*DT + 0.10*IS
Honeysuckle	0.37*E + 0.22*DP + 0.19*CN + 0.12*DT + 0.10*IS
Leafy Spurge	0.37*E + 0.22*DP + 0.19*CN + 0.12*DT + 0.10*IS
Multiflora Rose	0.37*E + 0.22*DP + 0.21*CN + 0.11*DT + 0.09*IS
Purple Loosestrife	0.37*E + 0.22*DP + 0.19*CN + 0.12*DT + 0.10*IS
Spotted Knapweed	0.36*E + 0.22*DP + 0.20*CN + 0.12*DT + 0.10*IS

Table 4. The percent of invaded pixels correctly assigned a high risk for each Park, species, and invasion phase for the initial version of the models.

Species	Introduction	Establishment	Spread	Pixels
Apostle Islands National Lakeshore (APIS)				
Spotted Knapweed	4.56%	27.72%	4.91%	285
Indiana Dunes National Lakeshore (INDU)				
Common Reed	87.03%	84.50%	88.65%	987
Garlic Mustard	76.46%	87.24%	99.29%	19414
Honeysuckle	89.02%	89.95%	90.05%	965
Multiflora Rose	19.72%	46.48%	46.48%	71
Purple Loosestrife	64.67%	48.15%	69.31%	8145
Spotted Knapweed	100.00%	100.00%	100.00%	51
Isle Royale National Park (ISRO)				
Common Reed	84.19%	83.33%	83.33%	54
Spotted Knapweed	20.47%	20.47%	51.97%	197
Mississippi National River and Recreation Area (MISS)				
Common Buckthorn	23.60%	73.09%	76.79%	30101
Garlic Mustard	25.70%	59.66%	90.65%	1284
Honeysuckle	17.65%	41.18%	52.94%	17
Leafy Spurge	49.25%	62.75%	65.50%	400
Spotted Knapweed	39.14%	78.62%	92.26%	1188
Pictured Rocks National Lakeshore (PIRO)				
Spotted Knapweed	77.37%	99.42%	99.88%	3424
Saint Croix National Scenic River (SACN)				
Common Buckthorn	0.00%	100.00%	100.00%	5
Garlic Mustard	35.76%	37.75%	88.74%	151
Glossy Buckthorn	45.00%	60.00%	95.00%	20
Honeysuckle	27.27%	27.27%	27.27%	11
Leafy Spurge	100.00%	100.00%	100.00%	1
Purple Loosestrife	5.45%	10.91%	10.91%	55
Spotted Knapweed	0.00%	33.33%	66.67%	3
Sleeping Bear Dunes National Lakeshore (SLBE)				
Baby's Breath	81.27%	81.40%	81.15%	785
Common Reed	71.65%	71.65%	70.87%	127
Garlic Mustard	37.04%	15.62%	80.84%	621
Honeysuckle	94.69%	94.91%	96.15%	9258
Leafy Spurge	83.14%	92.49%	92.82%	14358
Multiflora Rose	72.31%	90.88%	94.65%	6688
Purple Loosestrife	100.00%	90.91%	100.00%	11
Spotted Knapweed	71.60%	71.60%	71.60%	81
Voyageurs National Park (VOYA)				
Purple Loosestrife	33.33%	33.33%	33.33%	3
Spotted Knapweed	50.00%	100.00%	100.00%	2
Average	**60.12%**	**80.29%**	**87.08%**	

Table 5. The percent of invaded pixels correctly assigned a high risk for each Park, species, and invasion phase for the final version of the models.

Species	Introduction	Establishment	Spread	Pixels
Apostle Islands National Lakeshore (APIS)				
Spotted Knapweed	73.33%	72.28%	81.05%	285
Indiana Dunes National Lakeshore (INDU)				
Common Reed	83.01%	70.17%	92.49%	865
Garlic Mustard	73.37%	74.99%	81.00%	28803
Honeysuckle	99.42%	99.33%	100.00%	1041
Multiflora Rose	92.49%	90.06%	93.71%	493
Purple Loosestrife	96.42%	92.15%	96.37%	8205
Spotted Knapweed	100.00%	100.00%	100.00%	51
Isle Royale National Park (ISRO)				
Common Reed	84.19%	83.33%	83.33%	54
Spotted Knapweed	83.61%	83.61%	95.38%	238
Mississippi National River and Recreation Area (MISS)				
Common Buckthorn	72.04%	96.36%	89.98%	47311
Garlic Mustard	90.63%	96.68%	99.88%	3223
Honeysuckle	100.00%	100.00%	100.00%	134
Leafy Spurge	70.81%	95.32%	95.00%	620
Spotted Knapweed	72.08%	83.18%	97.33%	4001
Pictured Rocks National Lakeshore (PIRO)				
Spotted Knapweed	78.18%	99.42%	99.56%	3424
Saint Croix National Scenic River (SACN)				
Common Buckthorn	100.00%	100.00%	100.00%	5
Garlic Mustard	88.74%	86.09%	99.34%	151
Glossy Buckthorn	90.00%	100.00%	95.00%	20
Honeysuckle	100.00%	88.89%	100.00%	9
Leafy Spurge	100.00%	100.00%	100.00%	1
Purple Loosestrife	98.18%	98.18%	100.00%	55
Spotted Knapweed	87.10%	83.87%	100.00%	31
Sleeping Bear Dunes National Lakeshore (SLBE)				
Baby’s Breath	97.20%	97.20%	98.22%	785
Common Reed	99.21%	100.00%	100.00%	127
Garlic Mustard	93.60%	98.85%	100.00%	609
Honeysuckle	97.39%	94.97%	99.39%	9288
Leafy Spurge	77.21%	96.98%	97.56%	14061
Multiflora Rose	70.50%	88.18%	96.98%	6698
Purple Loosestrife	100.00%	100.00%	100.00%	11
Spotted Knapweed	72.62%	85.71%	88.10%	84
Voyageurs National Park (VOYA)				
Purple Loosestrife	100.00%	96.97%	100.00%	33
Spotted Knapweed	50.00%	50.00%	100.00%	2
Average	**77.50%**	**90.29%**	**91.21%**	

Table 6. Number of points with presence of the target invasive plant at random points sampled within Indiana Dunes (n=75) and Sleeping Bear Dunes (n=76) during summer 2006 and within Sleeping Bear Dunes (n=162) during summer 2007.

Species	INDU	SLBE (2006)	SLBE (2007)
Baby's Breath	1	7	7
Common Buckthorn	-	-	-
Common Reed	4	1	-
Garlic Mustard	21	-	3
Glossy Buckthorn	2	-	-
Honeysuckle	2	6	15
Leafy Spurge	-	2	11
Multiflora Rose	25	4	2
Purple Loosestrife	2	3	-
Spotted Knapweed	3	14	42

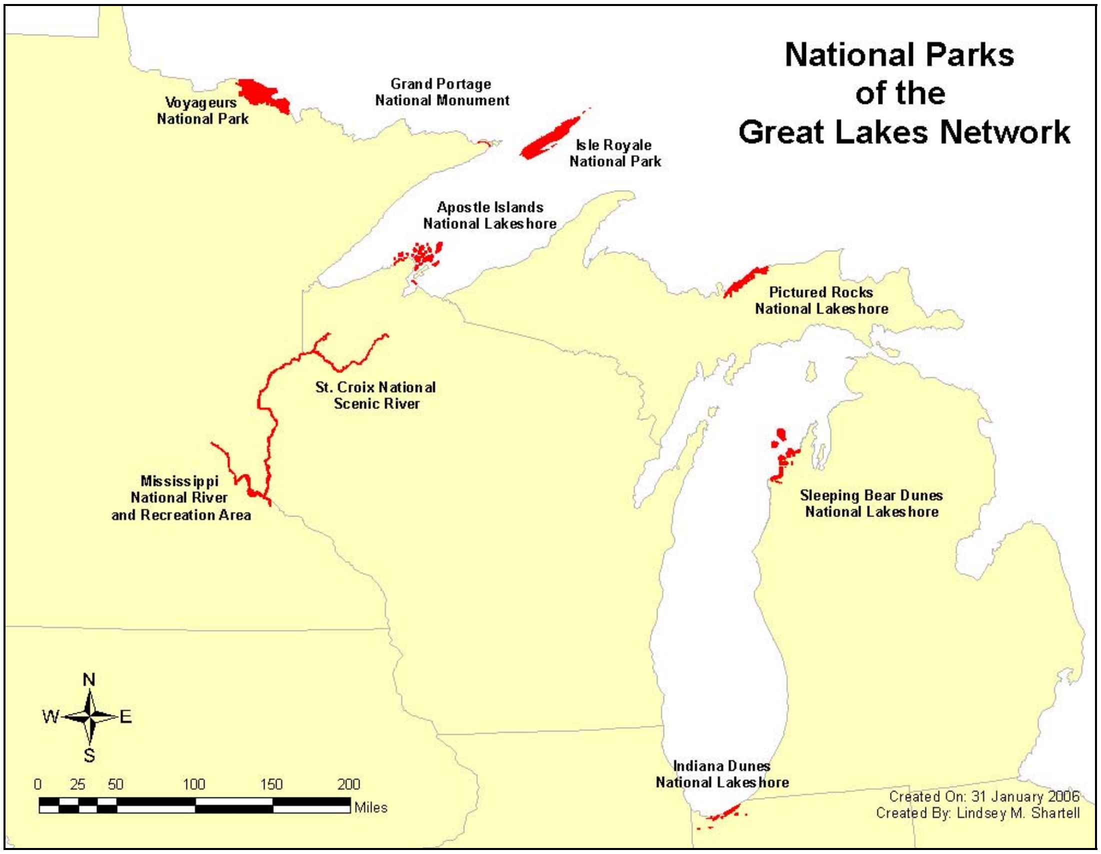


Figure 1. The locations of the nine National Parks of the Great Lakes Network.

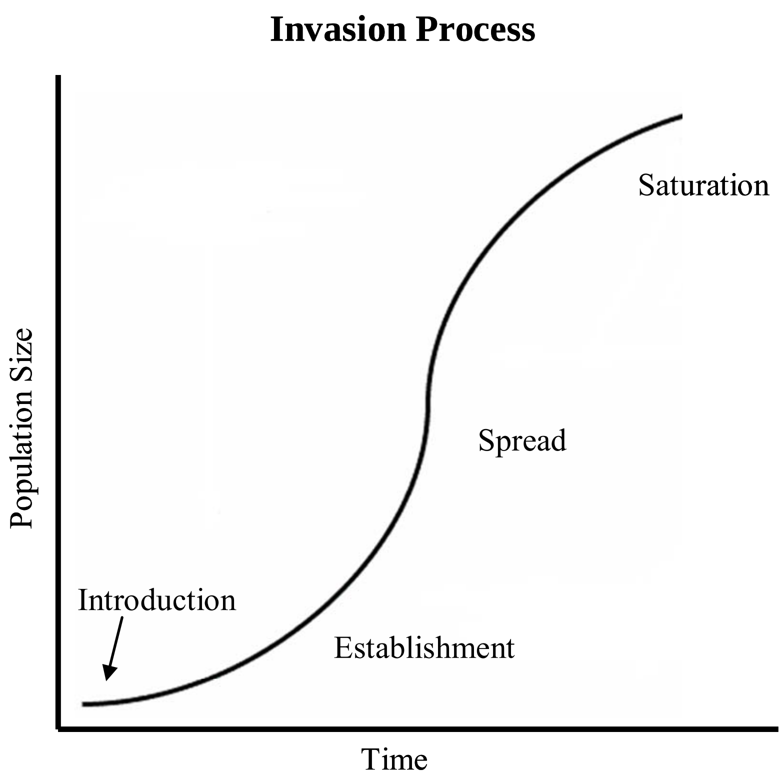


Figure 2. The phases of invasion follow an S-shaped population growth curve.

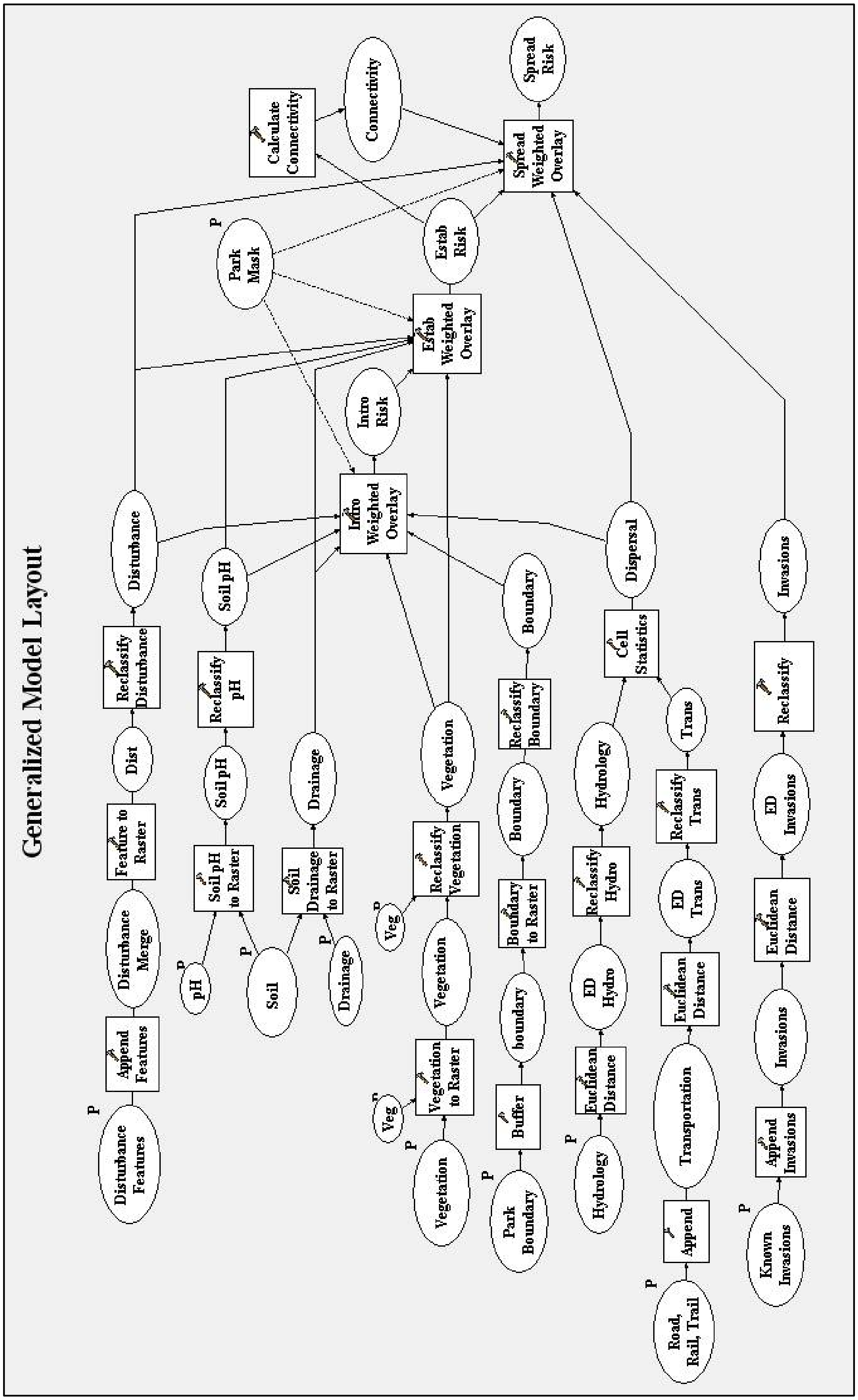


Figure 3. Generalized model layout for the multi-criteria risk models.

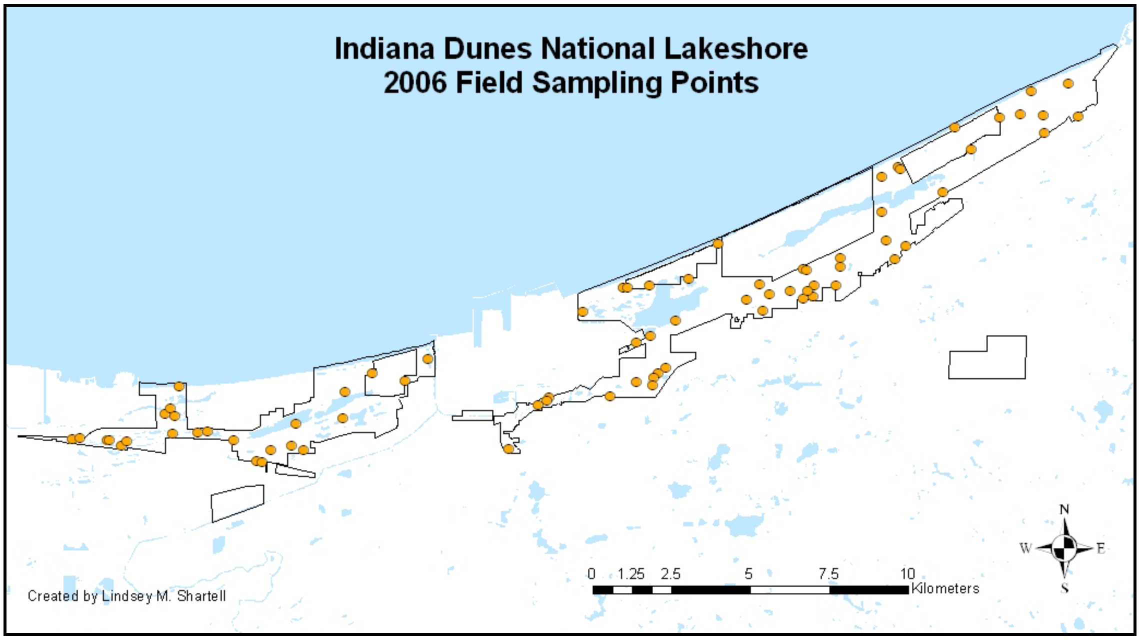


Figure 4. The location of field sampling points at Indiana Dunes National Lakeshore during summer 2006.

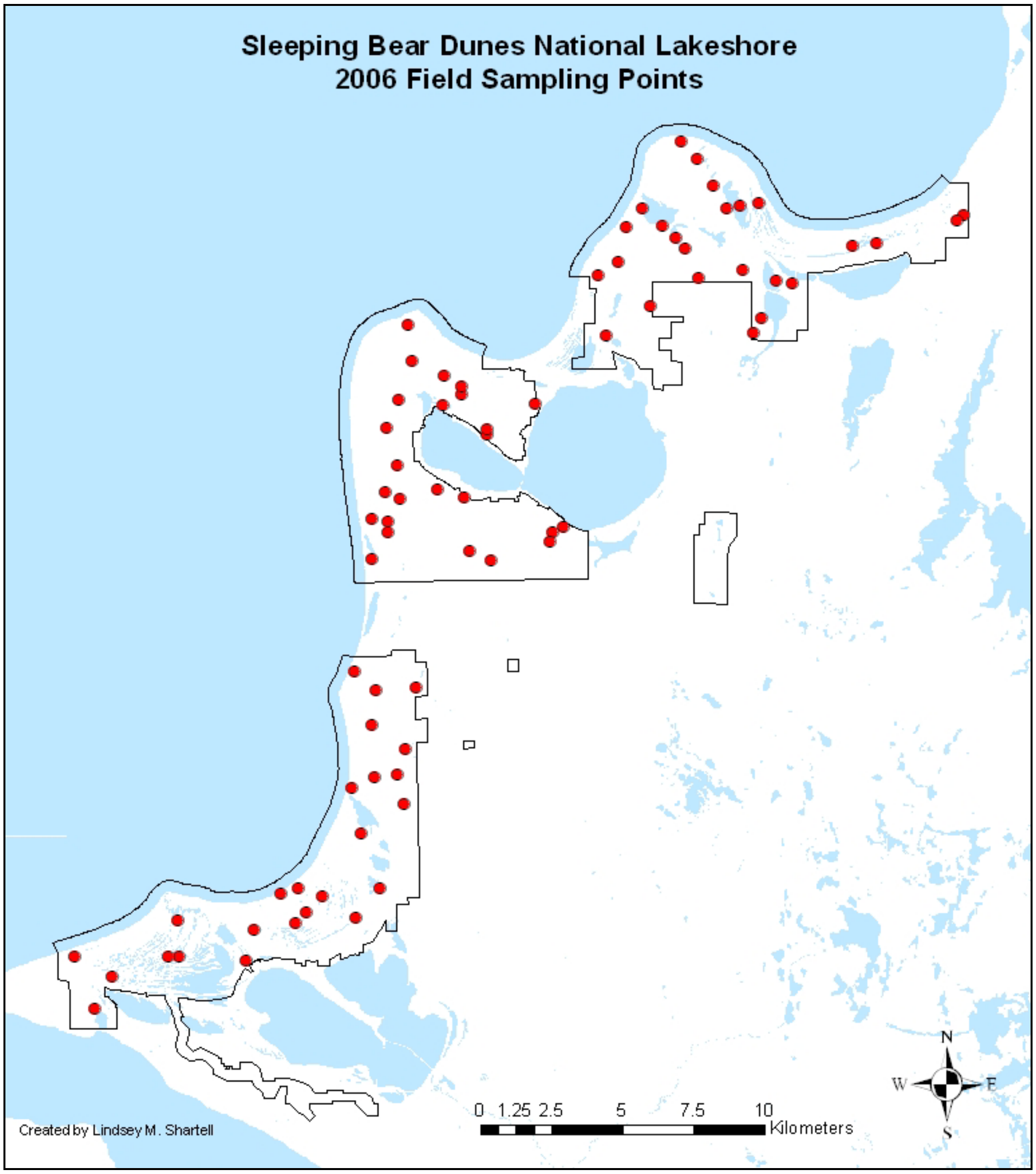


Figure 5. The location of field sampling points at Sleeping Bear Dunes National Lakeshore during summer 2006.

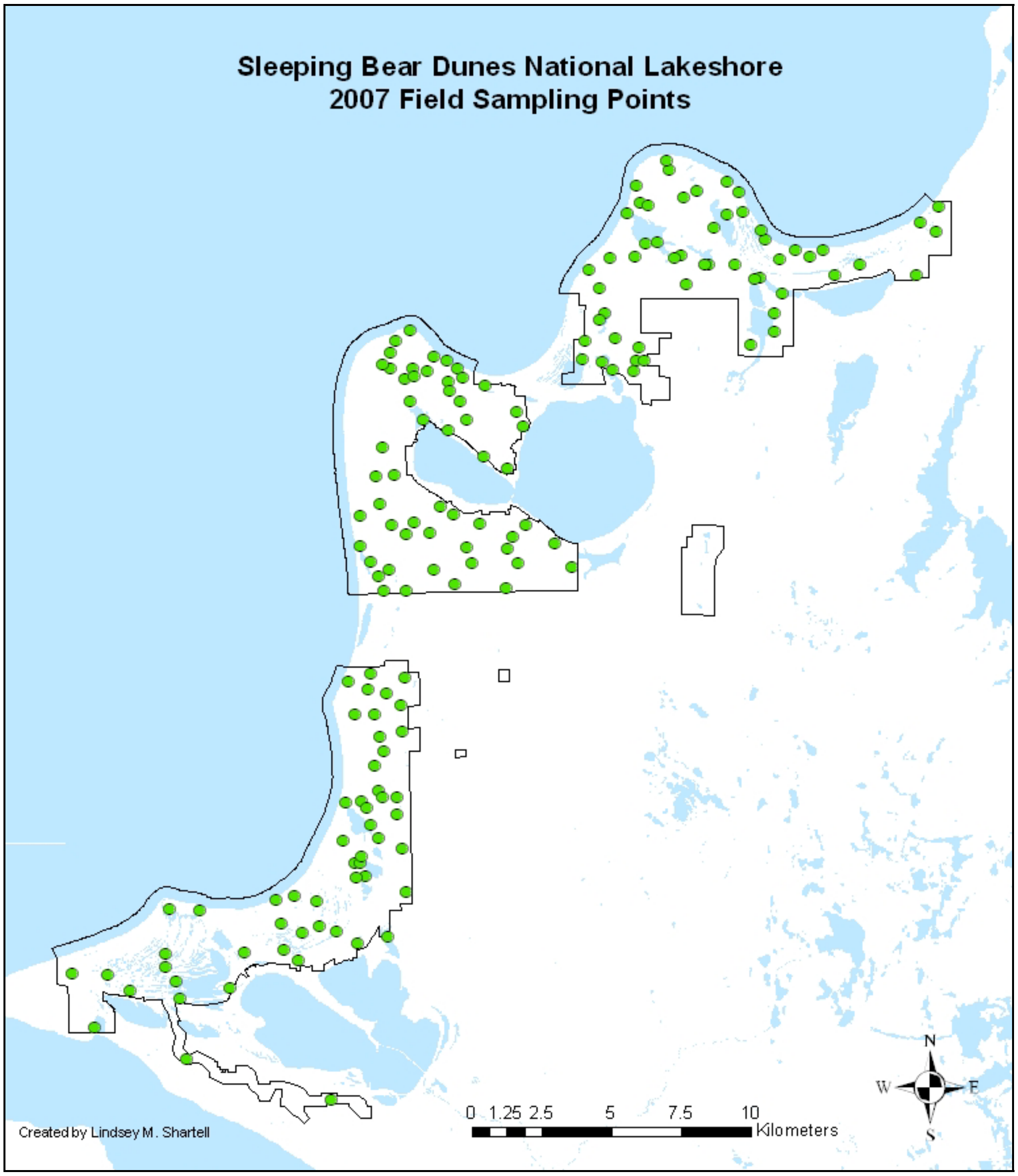


Figure 6. The location of field sampling points at Sleeping Bear Dunes National Lakeshore during summer 2007.

Due to the large number of images within the appendix, the file size of this complete report is 52 MB. Thus, we have deleted the appendix from this version posted on the web. If you are interested in obtaining a copy of the complete report, please contact the Great Lakes Network office at 715-682-0631.

National Park Service
U.S. Department of the Interior

Great Lakes Network Office
2800 Lake Shore Dr. East
Ashland, WI 54806

www.nps.gov