GLIFWC Invasive Plant
Risk Assessment/Prioritization Models

for

Ashland, Douglas, Bayfield, and Iron Counties
of Northern Wisconsin.

Project Report 10-01

September 2010

by
Steve Garske, Invasive Plant Specialist
phone: (715) 682-6619 x126
email: steveg@glifwc.org

Contact Person
Miles Falck, Wildlife Biologist
phone: (715) 682-6619 x124
email: miles@glifwc.org

Great Lakes Indian Fish & Wildlife Commission
PO Box 9
72682 Maple Street
Odanah, WI 54861

INTRODUCTION

The Great Lakes Indian Fish and Wildlife Commission (GLIFWC) is an organization exercising delegated authority from 11 federally recognized Ojibwe tribes in Minnesota, Wisconsin, and Michigan. These tribes retain hunting, fishing, and gathering rights in the territories ceded to the United States through various treaties. The degradation of native ecosystems by invasive species poses a serious threat to the continued exercise of these rights, and the traditional lifeways they sustain.

Invasive species are considered by many scientists to be a major cause of biodiversity loss and species extinction worldwide (Wilcove et al. 1998, Enserink 1999, UNEP 2010). A recent United Nations report lists habitat change, overexploitation of natural resources, climate change, pollution, and non-native invasive species as the five major causes of global biodiversity loss (UNEP 2010). Besides physical displacement of native flora and fauna, invasive species can alter fire frequency, hydrologic properties, soil chemistry, and the physical and trophic structure of entire ecosystems (Walker and Smith 1997, Westbrooks 1998).

Invasive plants can even cause significant declines in wildlife populations (Lloyd and Martin 2005, Flanders et al. 2006, Tallamy 2007). Most herbivorous insects are moderately to highly host-specific, and cannot efficiently utilize non-native plants (Tallamy 2004, Conrad et al. 2006). Furthermore, the horticultural industry favors plants that are “pest free” (Tallamy 2004, 2007). Often imported without their suite of natural enemies, introduced plants are able to displace native plants and eventually dominate the landscape (Williamson 1996). Native insects are generally unable to use these plants effectively, sending insect populations into decline (Tallamy 2004, Greenwood et al. 2004, Conrad et al. 2006). Decreases in native insect populations lead to decreases in populations of the birds and small mammals that need insects to survive, leading to adverse impacts that ripple through entire ecosystems (Tallamy 2004, 2007).

The origin and distribution of species (or biogeography) has been of interest to naturalists for literally centuries. Until fairly recently species distribution maps have relied almost entirely on landscape observations, combined with naturalist's knowledge of species’ habitats. Faster, more affordable computers and advances in geographic information system (GIS) technology have led to considerably more sophisticated modelling and prediction of species distributions. Species distribution models (SDMs) correlate spatial information on known occurrences of the species being modelled with information on the environment, to mathematically predict where a species occurs and where it does not or cannot occur (Phillips et al. 2006, Kearney and Porter 2009).

The purpose of this project was to develop an invasive plant risk assessment/prioritization modelling method to support GLIFWC’s ceded territory resource management activities, in conjunction with the Northwoods Cooperative Weed Management Area (NCWMA). We anticipate that this project will lead to more effective allocation of resources and more efficient completion of invasive plant surveys and control activities.

Fundamental vs. realized niche

A species' fundamental niche was first defined as the set of all conditions needed for it to exist indefinitely (Hutchinson 1957). The fundamental niche represents a species' potential distribution, whereas the area that it currently inhabits is its realized niche (Hutchinson 1957). A species typically occupies only a portion of its fundamental niche. Factors preventing a species from fully occupying its entire fundamental niche include competition from other species, physical barriers to dispersal such as waterbodies, deserts, mountain ranges, and cities, and predation and disease (Pulliam 2000). Only the realized niche (or species distribution) can be observed directly (Phillips et al. 2006).

While modelling techniques have advanced rather rapidly in recent years, questions still remain as to what is really being modelled, the degree to which a model captures a species' full niche requirements, and the degree to which the distributions produced by these models represent true fundamental or realized niches (Phillips et al. 2006, Elith and Leathwick 2009). For this reason the term "species distribution model" is preferable (Elith and Leathwick 2009).

Invasive species data

For data sets that include presence and absence data, standard statistical methods can be used. Accurate and reliable absence data are rarely available though (Phillips et al. 2006). Even when reliable absence data are available, their use can be problematic for a number of reasons, most notably because a species may be absent at a site because the habitat is unsuitable, or because it is suitable but unoccupied (Anderson et al. 2003, Elith and Leathwick 2009).

A major problem for most habitat modelling studies is the lack of high-quality data (Guisan et al. 2006, Jones and Reichard 2009). The data used in most habitat modelling studies are often obtained from third-party sources such as museums and data banks, and may contain unknown levels of bias and error. This study is relatively unique because it uses perhaps the most comprehensive set of taxonomically and locationally accurate data for a large area anywhere. This dataset includes 39,746 invasive plant and animal records (generally aquatic invertebrates) from across the western Great Lakes region (see http://maps.glifwc.org/ for more information). These records were obtained through GLIFWC terrestrial and aquatic invasive plant surveys and from regional cooperators. Of these records, 6,009 are of non-cultivated plant populations occurring within the four contiguous Lake Superior counties of Wisconsin (Iron, Ashland, Bayfield and Douglas). Nearly two-thirds of the data for this four-county project area was recorded by the author during invasive plant surveys conducted from 2001 through 2009. The rest were generally obtained by trained botanists and other natural resource workers from other public agencies.

Unless otherwise noted, vascular plant nomenclature follows Gleason and Cronquist (1991).

Species distribution models

The first computer habitat models were based on simple habitat-matching techniques such as BIOCLIM (Busby 1991) and DOMAIN (Carpenter et al. 1993). Within the last 10 years or so these have been joined by “second-generation” models that can fit more complex non-linear relationships to the data. These newer models include Ecological Niche Factor Analysis, or ENFA (Biomapper, Hirzel and Arlettaz 2003), Genetic Algorithm for Rules Set Prediction, or GARP (Stockwell and Peters 1999), Boosted Regression Trees, or BRT (Leathwick et al. 2006), several regression-based methods, and the maximum entropy method (Maxent, Phillips et al. 2006).

In order to sample the available habitats for a species, presence-only methods produce random “background” or pseudo-absence” points to sample the range of habitats occurring within the area being modelled (Elith and Leathwick 2009, Phillips et al. 2009). These methods generally outperform the older habitat matching models (Elith et al. 2006, Phillips et al. 2006).

Maximum entropy

Our investigations have so far relied on the maximum entropy method (see Berger et al. 1996 for an introduction) as implemented by Maxent software (Phillips et al. 2006; Phillips et al. 2009). Maxent was selected for a number of reasons, including: 1) Like most other second-generation modelling programs, Maxent is designed to use presence-only data, 2) Unlike most other modelling programs, Maxent is capable of using categorical as well as continuous data, and can incorporate interactions between different variables, 3) Maxent is capable of using a target-group background layer, to reduce or remove locational bias in the data, 4) Maxent software is reliable and relatively easy-to-use, and 5) Maxent is consistently ranked near or at the top in model performance tests (Elith et al. 2006, Elith and Graham 2009).

Maxent estimates a target probability distribution by finding the distribution of maximum entropy (closest to a uniform distribution), subject to uncertainty constraints enforced by the environmental features (Phillips et al. 2006). Georeferenced occurrence and environmental data are used as input. “Background” points (either randomly-generated or targeted - see below) are used to sample the range of environments available to the species within the project area. Using this information, Maxent finds the suitability of each grid cell as a function of the environmental variables at that grid cell (Phillips 2009). A high value at a particular grid cell indicates that the grid cell is predicted to have suitable conditions for that species. The computed model is a probability distribution over all the grid cells.

The Maxent SDM output includes a number of useful diagnostic tests and statistics (see below), and a GIS ASCII grid coverage. I have chosen a logistic output for the mapped results for this study, resulting in values for each 30 m^2 grid cell that range from 0 to 1. These values represent the probability of an invasive plant species being present at that location. The model output with relevant GIS overlays provides insights into where management activities can be most cost-effective. For example, isolated occurrences in areas of highly favorable habitat can be prioritized for treatment. Conversely, treatment can be deferred for sites within, or adjacent to,

areas of low potential. Additional priorities could be derived by overlaying GIS coverages for other resources potentially threatened by invasive plants, including sensitive habitats or cultural resources.

Environmental variables

The choice of environmental variables is critical in producing environmentally meaningful models (Elith and Leathwick 2009). Major factors limiting the distribution of plant species include minimum winter temperatures (corresponding to a plant's "hardiness"), growing season temperature and length, and moisture availability (Woodward and Williams 1987, Jones and Reichard 2009). Soil characteristics are also important determinants of plant distribution (Fitter 1982, Peet et al. 2003). Only a few modelling studies have made extensive use of soil features.

The project area is large and topographically varied, with topography and Lake Superior influencing climate to a significant degree. The frost-free growing season, for example, ranges from 85-128 days, while the average annual minimum temperature for the period1961 through 1990 ranges from -49 to -30° C. (Unlike most crop plants, most wild plants of temperate climates are tolerant of light to moderate frost, so for these plants the growing season extends somewhat beyond this frost-free period.)

Roads can serve as corridors for the dispersal of invasive species, by providing disturbed habitat for invasives to colonize, and by spreading seeds and other propagules during road maintenance activities and through air turbulence caused by passing vehicles (Christen and Matlack 2006). While some roadside weeds are ruderal species that don't significantly invade natural habitats, others (e.g., purple loosestrife and Eurasian bush honeysuckles) are aggressive invasives that easily disperse along road corridors.

Predictors can represent direct and indirect gradients (Elith and Leathwick 2009). Indirect gradients might include elevation and aspect, for example, while direct gradients might include solar radiation and snow cover. As a general rule, climatic variables such as annual precipitation are appropriate at global and meso-scales, topographic variables at "topographic" scales, and land cover variables such as percent canopy cover at micro-scales (Phillips et al. 2006).

Predictions for new sites are based on a species' location in environmental rather than geographic space (Elith and Leathwick 2009). When species are mapped according to environmental predictors alone, the distribution is mapped in environmental space, and the geographical proximity of species occurrences has no direct effect on the derived model. Environmental layers are typically autocorrelated, however, so clustered species occurrences often result in over-representation of some environmental characteristics and under-representation of others.

PROJECT AREA

The area covered by this project includes the four contiguous Wisconsin counties bordering Lake Superior (Figure 1). This 3.0 million acre region encompasses more than 750 named lakes, 130 rivers, and 51 cities and towns. Land ownership includes significant amounts of federal, state, county, and tribal lands. Our models are aimed at predicting the distributions of selected invasive plants across this area.

Albert (1994) classifies the project area into six ecological regions, based on edaphic, geomorphologic, macroclimatic and vegetational characteristics. These regions can be described as follows:

A band of lacustrine clays and clay till called the Lake Superior Clay Plain extends along the Lake Superior shoreline, across the entire northern portion of the project area. This clay plain is dissected by numerous small rivers. Coastal wetlands are an important feature as well. Before European settlement this region was dominated by mixed hardwood-conifer forests, with pockets of southern boreal forest. Northern hardwood forests composed mostly of sugar maple (*Acer saccharum*) were locally dominant on upland sites. Lowland forests were dominated by white spruce (*Picea glauca*) and balsam fir (*Abies balsamea*). This region has been significantly

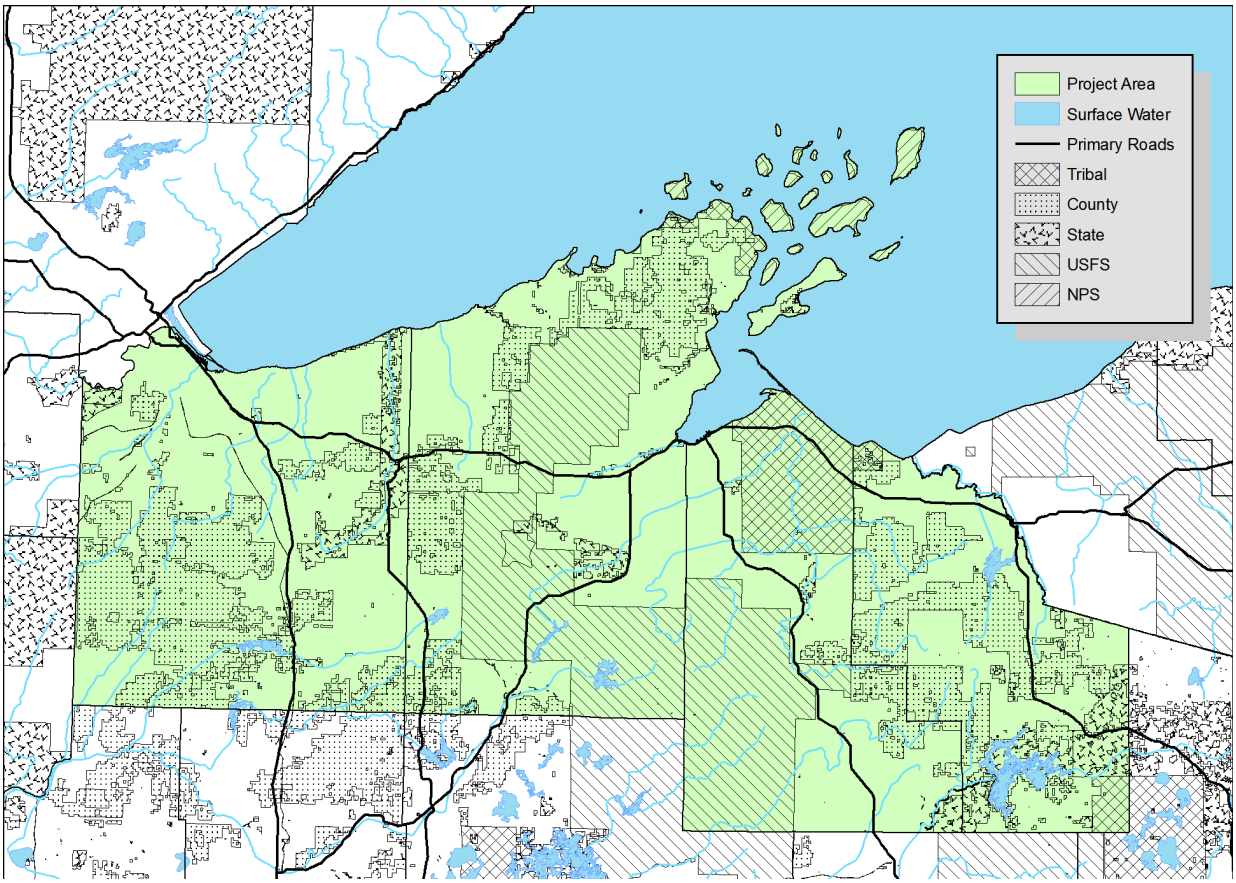

Figure 1. Project area in northern Wisconsin.

altered by farming, urban development and conversion of mixed forests to aspen (mainly *Populus tremuloides*). Nonetheless much of the land remains forested.

The Bayfield Barrens (which include the Moquah Barrens) extend from inland northeast Bayfield County, southwest to southeast Douglas County. This region has deep, droughty, sandy soils low in organic material. It was originally dominated by jack pine (*Pinus banksiana*) barrens, with red pine (*Pinus resinosa*) and white pine (*Pinus strobus*) forests dominating the northern end of this region. Small seepage lakes dot the southern end of the region. This region has been extensively altered by pine plantations, fire suppression and recreational development.

The Gogebic-Penokee Range extends from southeast Bayfield County across central Ashland and northern Iron Counties and into Michigan. Steep ridges of Precambrian basalts and granites are the remnants of an ancient mountain range. Soils are mostly red, sandy loams containing high amounts of iron. Bedrock ridges were and often still are dominated by red pine, white pine, red oak (*Quercus rubra*) and paper birch (*Betula papyrifera*). Northern hardwood forests of sugar maple, hemlock (*Tsuga canadensis*), yellow birch (*Betula alleghaniensis*), basswood (*Tilia americana*), and white ash (*Fraxinus americana*) dominate the uplands, with balsam fir, white spruce, white cedar (*Thuja occidentalis*), and black ash (*Fraxinus nigra*) dominating poorly-drained sites. Northern hardwoods continue to dominate much of the region today.

The Upper Wisconsin/Michigan Moraines is a region of diverse soils and vegetation extending across southern Bayfield, Ashland, and Iron Counties (Albert 1994). Soils range from sandy loams to acid silt loams, and often include significant amounts of rock. Bogs and swamps are common. Uplands were dominated by northern hardwood forests, with hemlock common. Much of the land is still forested.

Southern Iron County falls within the Lac Veaux Desert Outwash Plain. This area is characterized by acidic sands and numerous kettle lakes. White and red pine once dominated the uplands, while extensive hardwood conifer swamps and bogs occupy the lowlands. Logging has converted most of the upland forests to paper birch and aspen.

Finally, much of central and southwestern Douglas County falls within the Mille Lacs uplands. Soils range from loamy in the east to stony sands in the southwest. Vegetation originally ranged from pure white pine stands, to northern hardwoods with white pine as a supercanopy tree, to jack pine forests and upland forests dominated by white oak (*Quercus alba*), red oak, and bur oak (*Quercus macrocarpa*). Large areas of bog and swamp forest still cover the western portion of the region. Much of this region has been converted to pine plantations.

STUDY SPECIES

“Final” models were produced for the following invasive plant taxa. With the exception of alpine oatgrass, soapwort, and yellow iris, these plants have been recorded at more than 60 sites scattered across a substantial portion of the project area. Some of these plants (leafy spurge, common and glossy buckthorn, purple loosestrife, and all three honeysuckle taxa) are listed by the Wisconsin DNR as “restricted” species under NR 40, making it illegal to knowingly transport,

transfer (including through sale), or introduce them anywhere in the state.

Because they are sometimes difficult to identify with certainty and appear to inhabit essentially the same habitats, several closely-related taxa were combined into one for this analysis. These included the Eurasian bush honeysuckles (*Lonicera tatarica*, *L. morrowii*, and their hybrid, *L*. x *bella*) and the Eurasian tree willows (*Salix alba*, *S. fragilis*, and their hybrid, *S*. x *rubens*).

Wetland herbaceous plants

European yellow iris

European yellow iris (*Iris pseudacorus*, Iridaceae) is a showy perennial native to Europe, western Asia and North Africa (Sutherland 1990). Outside cultivation it is generally restricted to wet or waterlogged soils, including salt marshes. It tolerates a wide variety of soil types, from organic peat to gravel. It is also highly tolerant of anaerobic soils and of acid soils, occurring at pH (0-30 cm soil depth) ranging from 3.6 to 7.7. It requires relatively large amounts of nitrogen. Common habitats in its native range (Sutherland 1990) and in North America (pers. obs.) include lakeshores, riverbanks, floodplains, wetlands including bogs, seepy hillsides, and open wet woods. It readily colonizes (and probably contributes to) floating mats of vegetation (pers. obs.). It spreads via its buoyant seeds as well as by thick rhizomes (Sutherland 1990). It is moderately shade tolerant (Sutherland 1990).

Purple loosestrife

Purple loosestrife (*Lythrum salicaria*) is a showy wetland perennial native to Eurasia. It was first reported from North America along the eastern seaboard in 1814, where it probably arrived in shipping ballast (Thompson et al. 1987). Because it did not appear to spread significantly from the coast at first, purple loosestrife was considered to be native (Galatowitsch et al. 1999). But in the late 1800's it began its wholesale spread westward, aided by shipping canals and the shipping and horticultural industries.

Freed from its natural enemies (including two chrysomelid beetles of the genus *Galerucella*, now widely used for biological control), purple loosestrife aggressively invades wetlands, wet meadows, lakeshores, riverbanks, and ditches, forming dense stands and displacing native wetland plant communities. Today purple loosestrife can be found throughout most of North America. Several states now ban the sale and possession of this plant.

Although purple loosestrife prefers rich, organic soils, it will grow in nearly all moist to wet substrates, including clay, sand, and gravel. It can survive in partial shade, but grows best in full sunlight. Populations flourish in wetlands that have been disturbed or degraded through both natural and man-made processes. Banks and mudflats exposed following drawdown or drought are quickly colonized. Mature plants can produce over 2 million seeds annually. These tiny, lightweight seeds are easily dispersed by moving water. They can lie dormant for years until conditions become suitable for germination.

Common valerian

Common valerian (*Valeriana officinalis*, Valerianaceae) is native to temperate Europe and parts of Asia, where it has been grown for centuries for its medicinal properties. It has been cultivated in North America as a medicinal and ornamental plant for more than 150 years (IPANE 2009). It was first collected outside of cultivation in the Lake Superior region in 1938 in Duluth, by Minnesota botanist Olga Lakela (Lakela 1938). Lakela considered it a new record for Minnesota, mentioning that it was already locally common around Duluth. Despite its obviously invasive tendencies, common valerian is still widely sold in garden catalogs.

The modified calyx of common valerian flowers stays attached to the mature seed, acting as a pappus (a sort of a "parachute", similar to dandelion seeds) and allowing the seed to be carried by the wind (Voss 1996). Moving water is probably also a significant vector for seed dispersal in some habitats. Horticultural plantings are undoubtedly a major factor in common valerian's spread to new areas.

Some subspecies of common valerian send out short stolons, allowing them to produce new plants vegetatively (Evstatieva et al. 1993, IPANE 2009). This habit allows common valerian to form dense colonies of rosettes and flowering plants.

In its native range common valerian inhabits both calcareous and siliceous soils (Evstatieva et al. 1993). Typical habitats include “near river and streams, damp meadows, woods, and scrubs, but rarely also on dry rock formations and screes.” (Evstatieva et al. 1993).

Within the project area common valerian inhabits a variety of moist to wet, sunny to partly shaded habitats, including roadsides, ditches, pastures, and moist to wet meadows (pers. obs.). From these disturbed areas it often spreads to more natural habitats including alder-willow swamps, sedge meadows, and moist to wet woods. It often becomes abundant in these habitats, displacing native plant species and the insects, animals, and other organisms that depend on them.

Upland herbaceous plants

Alpine oatgrass

Alpine oatgrass [*Avenula pubescens* (Hudson) Dumort., Poaceae] was first discovered in the midwest in southeastern Ashland county, during the 2008 GLIFWC terrestrial invasives survey (Schimpf et al. 2009). This perennial grass is native to Europe, Siberia and central Asia (Dixon 1991). Across most of Europe alpine oatgrass is a common inhabitant of grasslands and open woods, as well as gravel pits, roadsides, and railway banks (Dixon 1991). Along the European coast it inhabits cliffs, stony sea-slopes and stabilized dunes, and even salt marshes (Dixon 1991). It usually inhabits moist to fairly dry, well-drained soils, but can also survive and reproduce on wet soils (Dixon 1991).

Alpine oatgrass is well-established on the outskirts of Glidden in southeastern Ashland County.

There it is a dominant herb across more than 0.4 ha of a mesic young red pine (*Pinus resinosa*) plantation and adjacent fields (Schimpf et al. 2009). It also occurs in mostly small colonies along Hwy 13 and Hwy 77 (pers. obs.), where it is presumably being spread by the usual road-associated vectors.

Leafy spurge

Leafy spurge (*Euphorbia esula*, Euphorbiaceae) is a strongly rhizomatous perennial native to Europe and Asia (Moore 1958). It is now found nearly throughout temperate North America. It is a major environmental and economic pest on the grasslands of the northern Great Plains. In the upper Great Lakes region leafy spurge inhabits disturbed areas such as roadsides, pastures, and old fields, as well as prairies, dry woodlands, grasslands, riverbanks and other natural habitats. It is found in similar habitats in Europe, but is only a minor agricultural weed there (Selleck et al. 1962). Worldwide, leafy spurge has colonized a wide variety of habitats, ranging from very dry to humid and seasonally wet, and from subtropical to subarctic (Selleck et al. 1962).

Leafy spurge tolerates a wide variety of soil types, but it is most aggressive on coarse, well-drained soils (Selleck et al. 1962). With its deep and extensive rhizome system, it is well-adapted to dry conditions. It can also withstand temporary flooding, as long as the shoots are able to grow above the water surface. It has low shade tolerance (Selleck et al. 1962). In moderate shade leafy spurge can persist vegetatively, but fails to flower. It is unable to survive in shaded, closed-canopy forest.

Woodland sweet pea

Woodland sweet pea (*Lathyrus sylvestris*, Fabaceae) is a sprawling leguminous plant with branches reaching 2.1 m long (USDA-NRCS 2010). It is native to Europe, western Asia, and Morocco (USDA-ARS 2010). It was first recorded in North America in 1827. It is occasionally planted as a forage, cover, or erosion-control crop, and has widely escaped from cultivation. It is now found throughout much of the US and parts of Canada (USDA-NRCS 2010).

According to USDA-NRCS (2010), woodland sweet pea has high drought tolerance, occurs from pH 5.0 to 7.8, is cold-hardy to -39° C, and needs a minimum of 140 frost-free days during the growing season. When our occurrence data for this plant is overlayed on the SSURGO frost-free days layer (USDA-SSDS 2007), though, it becomes evident that woodland sweet pea is well-established in portions of the project area with as few as 105 frost-free days. Although USDA-NRCS (2010) states that this plant is shade-tolerant, it avoids closed woods (pers. obs.).

Woodland sweet pea spreads aggressively by thick rhizomes, forming dense mats along roadsides and power line corridors in some parts of the project area, especially northeastern Bayfield County (pers. obs.). It is also abundant along roadsides in northern Ontonagon County, Michigan, but much less common in southern Ontonagon County (pers. obs.).

Soapwort

Soapwort (*Saponaria officinalis*, Caryophyllaceae) is a hermaphroditic, rhizomatous perennial native to Europe (Davis and Turner-Jones 2008). In North America it inhabits hillsides, riverbanks, roadsides, railroad tracks, meadows and waste areas (Lubke and Cavers 1969, Lokker and Cavers 1995). It tends to grow in moist, well-drained soils, in full sun to partial shade. Reproduction is mostly by seed, but it also spreads clonally by rhizomes, forming large, diffuse to dense patches.

Within the project area it is abundant on the Bad River floodplain north of Mellen in Ashland County (pers. obs.). There it grows on moist, sandy or gravely soils. It is most common along the built-up banks of the river, in partial shade to nearly full sun. Soapwort is only moderately shade-tolerant, and avoids closed-canopy forest (pers. obs.). In its native range it is often abundant in open woods and along rivers and streams (Lubke and Cavers 1969).

According to USDA-NRCS (2010), soapwort has moderate drought-tolerance, no anaerobic tolerance, and occurs at (or at least prefers) a pH of between 5.0 and 7.0. Within the project area, populations occur at sites with mean annual minimum temperatures of as low as -43° C, and as few as 85 frost-free growing days.

Shrubs and trees

Eurasian bush honeysuckles

Eurasian bush honeysuckles include Tartarian honeysuckle (*Lonicera tatarica*, native of Eurasia), Morrow's honeysuckle (*L. morrowii*, native of Japan), and Bell's honeysuckle (*L.* x *bella*), all members of the Caprifoliaceae. These large shrubs have long been cultivated in North America (Rehder 1960). Historically used as hedgerows, ornamentals, and "wildlife" plantings, they have become well-established in a variety of habitats throughout temperate North America. Bell's honeysuckle (the result of hybridization between Tartarian and Morrow's honeysuckles) is the most frequently encountered of the three in the project area (pers. obs.). Within the project area, these large shrubs invade fields, woods edges, open woods and floodplains.

Eurasian bush honeysuckles are moderately shade-tolerant, but avoid deep shade (pers. obs.). According to USDA-NRCS (2010), Tartarian honeysuckle has moderate drought-tolerance, no anaerobic tolerance, and occurs from pH 5.2 to 7.5. It is hardy to -39° C. Within the project area, Eurasian bush honeysuckle populations occur at sites with as little as 85 frost-free days.

Glossy buckthorn

Glossy buckthorn (*Frangula alnus* Miller, Rhamnaceae) is a shrub or small tree with an affinity for moist to wet soils (Godwin 1943b). It avoids anaerobic, waterlogged soils (USDA-NRCS 2010). Godwin (1943b) states that glossy buckthorn is common on “alkaline, neutral and acidic peat” and gives a pH range of 6.5-8.3 for the subsurface peat surrounding roots of plants in one

fen in its native range. Catling and Porebski (1994) consider glossy buckthorn to be aggressive on alkaline as well as acid soils in southern Ontario. Glossy buckthorn needs a minimum of 90 frost-free growing days, according to USDA-NRCS (2010).

While glossy buckthorn does best on moist to wet soils, it tolerates a broad range of environmental conditions, and invades uplands as well (Godwin 1943b, Catling and Porebski 1994, pers. obs.). Like common buckthorn, glossy buckthorn leafs out early and retains its leaves well into the fall, giving it an advantage over most native forest understory shrubs (Harrington et al. 1989).

Glossy buckthorn readily colonizes disturbed habitats such as fencerows and fields, and natural habitats including wet meadows, bogs, lakeshores, riverbanks, woods edges and interior woods (Voss 1985). Glossy buckthorn is shade-tolerant, and is particularly aggressive in moist woods and lowland swamp forest (pers. obs.).

Common buckthorn

Common buckthorn (*Rhamnus cathartica*, Rhamnaceae) is one of the most aggressive invasive plants in North America. It has long been planted as an ornamental and for hedgerows. In Wisconsin it was used as a hedgerow plant as early as 1849 (Hoffman and Kearns 1997). In Michigan, it was first collected outside of cultivation in 1914 (Voss 1985). Two years later it was reported as locally common in southern Michigan (Voss 1985), indicating that it was already naturalized there by the time it was detected. In the early 1900s common buckthorn was promoted as a shelterbelt species on the northern Great Plains, until the realization that it was the alternate host of oat stem rust (Archibold et al. 1997). Today common buckthorn is widely established across most of temperate North America (USDA-NRCS 2010).

Common buckthorn reproduces almost entirely by seed (Godwin 1943a). It is functionally dioecious, with male and female flowers on separate plants (Gleason and Cronquist 1991). (Male flowers may include a vestigial, nonfunctional pistil). The berries are dispersed mainly by birds (Archibold et al. 1997).

Common buckthorn is typically an upland species and strongly associated with calcareous soils in its native Europe (Godwin 1943a, Archibold et al. 1997). In North America it colonizes a variety of soil types and tolerates a broad range of moisture levels. It has moderate shade-tolerance (the seedlings are very shade-tolerant) and is capable of invading mature forest (Godwin 1943a, pers. obs.). It does best at light levels of 12.5% or more of full sunlight (Gourley and Howell 1984). Early leaf-out and late leaf-fall gives this species an advantage over most native shrubs in the forest understory (Harrington et al. 1989).

Habitats colonized by common buckthorn in North America range from old fields, pastures, and roadbanks, to prairies, riverbanks, and interior woods. New populations often get their start where the seeds are deposited by perching birds, including along fencerows and woodland edges, and under isolated trees (Whitford and Whitford 1988). Moist, partly shaded sites appear to be optimal (Archibold et al. 1997). Populations within the project area occur in areas with as

little as 85 frost-free days.

Eurasian tree willows

As defined here the Eurasian tree willows include white willow (*Salix alba*), crack willow (*S. fragilis*), and hybrid crack willow, *S*. x *rubens* (Salicaceae). White willow is native to Europe and the Mediterranean basin, while crack willow is native to Western Asia and naturalized in central and Northern Europe (Skvortsov 1973 in Barcaccia et al. 2003). Both white and crack willow were imported as ornamentals during the British colonial times, and both were recorded as escaping from cultivation in North America by the early to mid 1800s (Mills et al. 1993, Newsholme 1992). Crack and white willow freely hybridize, producing the appropriately-named hybrid crack willow. All three willows are now established across much of temperate North America, including Michigan, Wisconsin and Minnesota (USDA-NRCS 2010).

Eurasian tree willows are typical of riverine habitats in their native range (De Cock et al. 2003). Like most willows, Eurasian tree willows have low shade-tolerance, and grow best in full sun. They prefer moist to wet, rich alluvial soils, but once established do well on average, mesic soils as well. According to USDA-NRCS (2010), white willow has high anaerobic tolerance, occurs from pH 4.5 to 7.8, and is hardy to -36° C. Once established, white willow seedlings and adults are moderately drought-tolerant (Van Splunder et al. 1996). Within the project area, Eurasian tree willow populations occur in areas with as little as 85 frost-free growing days.

Eurasian tree willows are often planted along roadsides, rivers, and lakeshores, as well as in yards and farmlands. They often spread from these plantings to colonize floodplains, stream and riverbanks, lakeshores, and wetland edges. They are well-adapted to seasonally inundated habitats. The twigs of crack willow are especially brittle. Winds and floods can break off twigs and even large branches of all these willows, depositing them downstream where they quickly root and produce new plants.

METHODS

Invasive species data

Invasive plant data were extracted from GLIFWC's online database (http://gisin.glifwc.org/). The 6,237 invasive plant occurrences from within the project area were reprojected from WGS 84 to WTM 83 HARN. Occurrences that listed land use as "residential" and those for which field notes indicated the occurrence was a result of "cultivation" or was "planted" were removed from the analysis. In general these included sites where the species being recorded occurred within a maintained yard and/or was obviously planted and not (yet) spreading significantly to adjacent "wild" areas. Occurrences that were partly planted and partly feral, or recorded as "planted and escaping" were retained. A total of 228 occurrences within the project area were of cultivated populations. These were removed from the data set, leaving 6,009 non-cultivated sites. The number of sites discarded through this process was minimal (generally < 1% for plants of interest). Models calibrated with naturalized populations are generally more accurate than those

based partly on cultivated records (Dullinger et al. 2009).

Of the remaining sites, all those recorded with a precision of at least 10 m (effectively all those recorded with a GPS receiver) were used in the analysis. This eliminated 229 additional sites. The occurrences used for modelling each invasive plant species were then drawn from the remaining 5,710 invasive plant occurrences.

Environmental Layers

Wisconsin Land Cover (WISCLAND) grid

The Wisconsin Statewide Land Cover (WISCLAND) grid (WISCLAND 1998) was downloaded from the Wisconsin DNR website. This layer was derived from LANDSAT Thematic Mapper (TM) satellite imagery acquired from fly-overs during the growing season, between August 1991 and May 1993. It categorically classifies land cover into broad classes (Table 1, Figure 2).

A “shade” layer was derived from the WISCLAND land cover grid, by reclassifying the original WISCLAND categories to approximate shade values, with 1 indicating full sun, 2 light shade, 3 fairly deep shade, and 4 deep shade (Table 1, Figure 3). Because these two layers are obviously highly correlated, they were never used in the same model.

The WISCLAND data has a grid cell size of 30 m. It was clipped with the project boundary layer and used as the template (for cell size and location of the origin) for converting the other environmental layers to grids. As stated by WISCLAND (1998), the data have a minimum mapping unit of 5 acres (2 ha), meaning that most land cover features this size or larger can be resolved in the data. Because our species data are accurate to (ostensibly) within 10 m, I decided to use this relatively fine-scale resolution for all the environmental layers. While it is usually preferable to use data that is consistent in scale, this is not always feasible in practice, and there is still little consensus on how to deal with scale disparities when fitting habitat models (Elith and Leathwick 2009).

NCDC minimum temperature data

Weather station data for average annual minimum temperatures for the period 1961-1990 were downloaded from the National Climatic Data Center website (NOAA-NCDC 2009), and converted to shapefiles. These data were recorded at 394 stations scattered fairly evenly across Minnesota, Wisconsin and Michigan. Eleven of these stations fell within the project area. This layer was converted to an interpolated grid layer using the ArcView Spatial Analyst’s “interpolate grid” function.

The minimum temperature data were interpolated with ArcView, using the inverse distance weighted (IDW) method and 6 nearest neighboring stations. All stations within 96.6 km (60 mile) were included in the interpolation. This layer was then extracted using the project area border polygon. The resulting map is shown in Figure 4.

Table 1. Land cover classes delineated in the WISCLAND data. To produce the "shade" layer, habitats were reclassified as shown. N/A indicates that the corresponding classification value did not occur within the project area.

CLASS VALUE	HABITAT	NUMBER OF MAPPING UNITS	SHADE RECLASS (1 = sun, 2, 3, 4 = deep shade)
100	**URBAN/DEVELOPED**	**N/A**	**N/A**
101	High Intensity	31,999	1
104	Low Intensity	25,917	1
105	Golf Course	434	1
110	**AGRICULTURE**	**858**	**1**
111	Herbaceous/Field Crops	29,703	1
113	Corn	6,250	1
118	Other Row Crops	9,416	1
124	Forage Crops (includes hay and hay/mix)	49,207	1
148	Cranberry Bog	2,695	1
150	**GRASSLAND (includes timothy, rye, pasture, idle, CRP, grass and volunteer)**	**1,129,267**	**1**
160	**FOREST**	**N/A**	**N/A**
162	Jack Pine	365,764	3
163	Red Pine	359,610	3
173	Mixed/Other Coniferous	267,747	4
176	Aspen	2,550,932	3
177	Oak	34,257	3
179	Northern Pin Oak	9,652	3
180	Red Oak	56,441	3
183	Maple	897,015	4
185	Sugar Maple	157,470	4
187	Mixed/Other Broad-leaved Deciduous	2,362,960	4
190	Mixed Deciduous/Coniferous	1,705,266	4
200	**OPEN WATER**	**395,381**	**1**
210	**WETLAND**	**N/A**	**N/A**
211	Emergent/Wet Meadow	165,521	1
217	Lowland Shrub	451,714	3
218	Broad-leaved Deciduous	596,379	3
219	Broad-leaved Evergreen	45,564	2
220	Needle-leaved	3,365	3
223	Broad-leaved Deciduous	322,907	4
229	Coniferous	559,762	4
234	Mixed Deciduous/Coniferous	470,025	4
240	**BARREN**	**67,173**	**2**
250	**SHRUBLAND**	**406,366**	**2**
255	**CLOUD COVER**	**N/A**	**N/A**

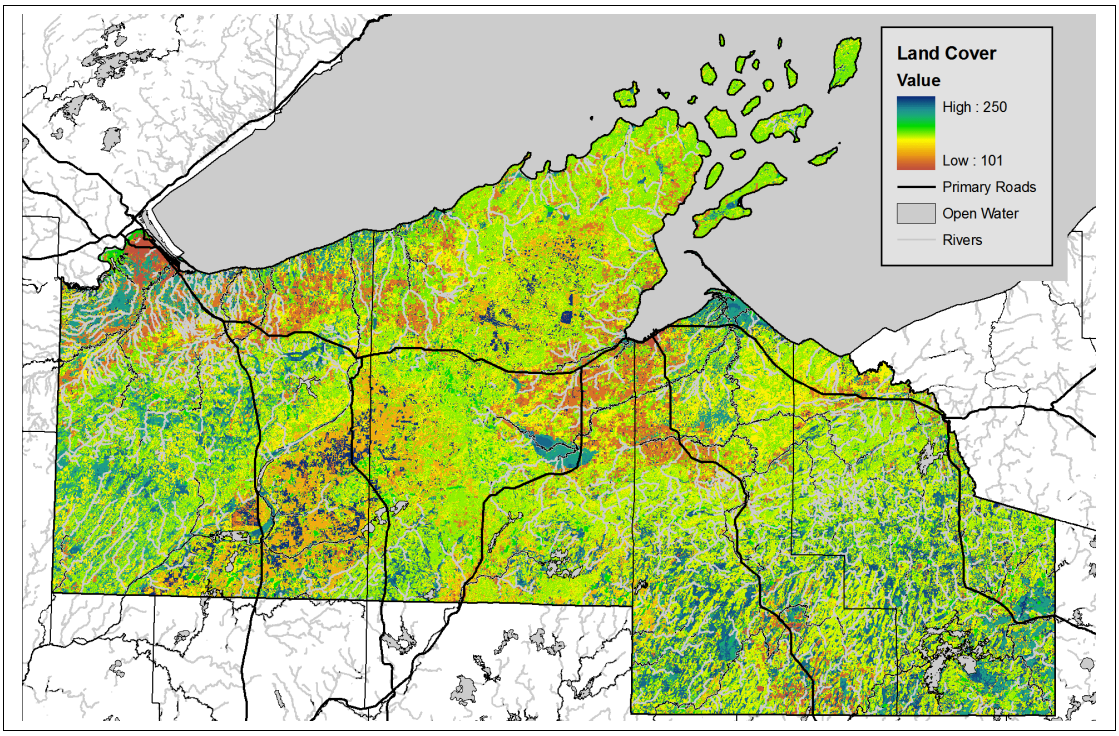

Figure 2. The WISCLAND land cover layer. See Table 1 for a list of land cover classifications.

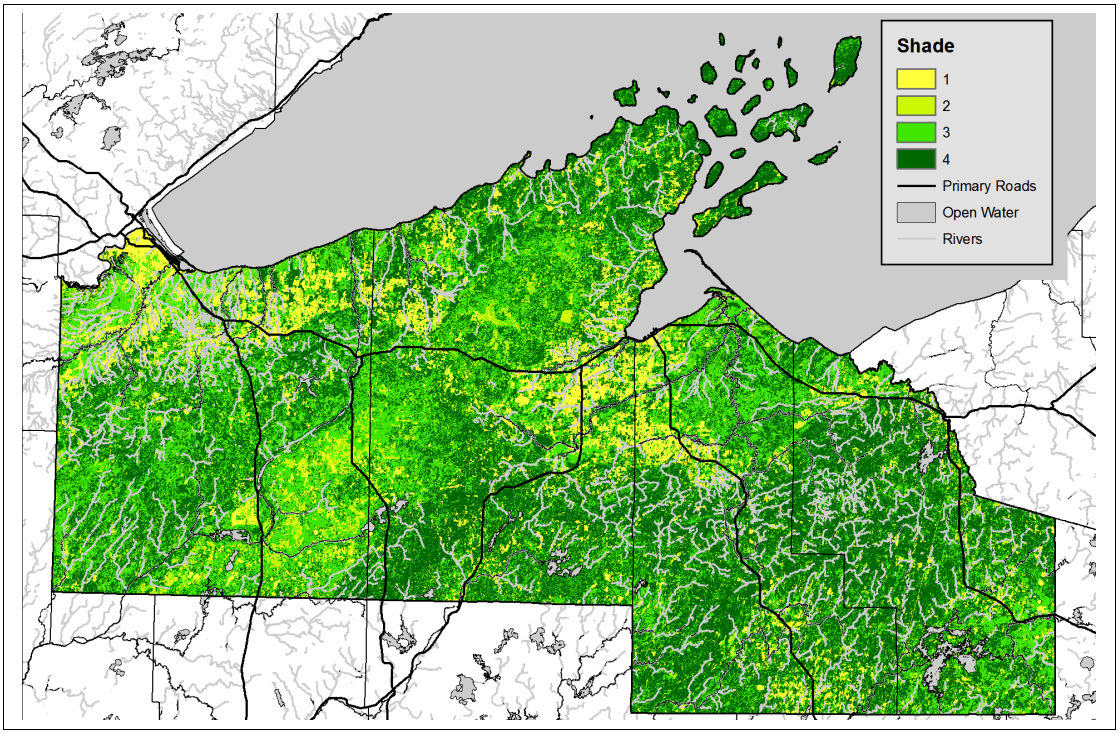

Figure 3. The shade layer. This layer was reclassified from the WISCLAND layer as in Table 1. Shadier, more heavily forested regions include the northern hardwood forests of the Gogebic-Penokee Range, southwest Douglas County, and the Bayfield Peninsula.

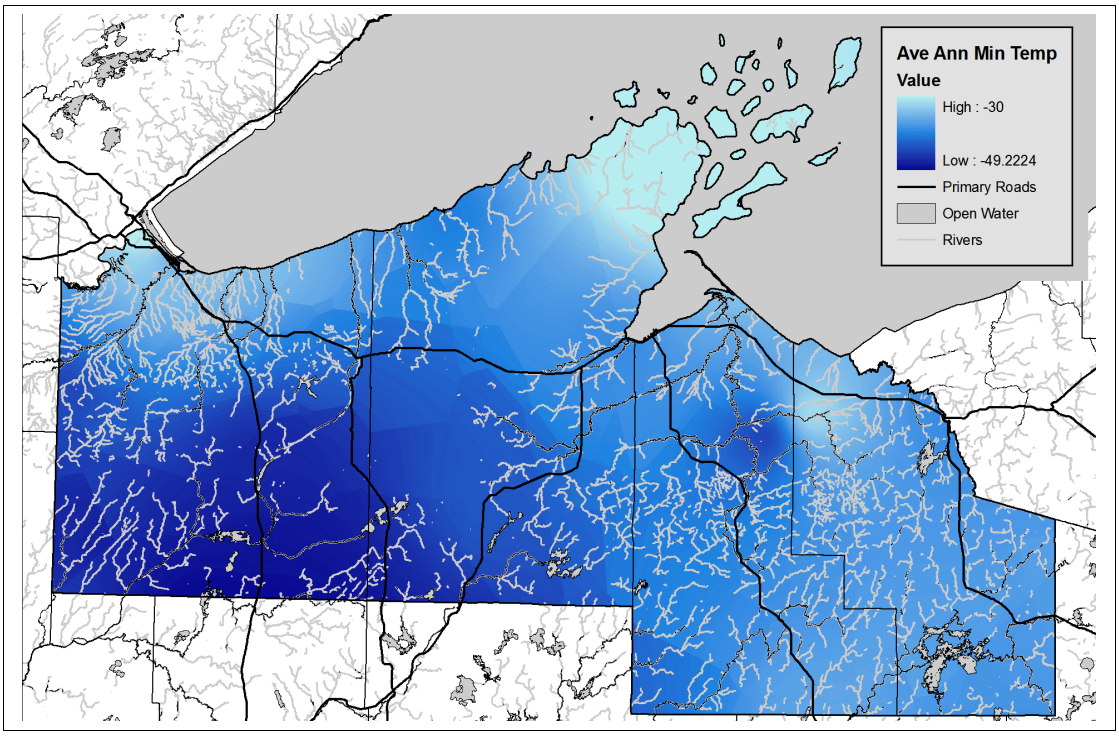

Figure 4. Interpolated map of average annual minimum temperature data, for the period 1961-1990. The western portion of the project area away from Lake Superior experiences the coldest average minimum temperatures, while the Bayfield Peninsula has the warmest.

Distance to nearest open water, river, and Lake Superior

Because models for certain wetland and shoreline species performed poorly, three additional layers were developed to take surface water into consideration. The first was a “distance to nearest open water” layer consisting of distance from lakes and wider, “open water” stretches of rivers (waters represented by polygons in the shapefile) (Figure 5). A “distance to nearest river” layer (represented by polylines in the shapefile) was also constructed (Figure 6). The base shapefiles for both these layers were obtained from WDNR (2010). Finally, a layer for “distance to Lake Superior” was constructed (Figure 7), from a base layer obtained from NOAA-GLGIS (2006).

For each layer, ArcView Spatial Analyst was used to find the distance (in meters) from the nearest respective feature within the project area plus a 20 mile buffer region. A new 30 m grid coverage was then generated, using the WISCLAND grid as a template.

Grid Analyst was used to extract the new distance theme, using the project area boundary as a mask. The 20-mile buffer allowed the effects of surface water within 20 miles of the boundary to be included in the layer, along with those within the boundary.

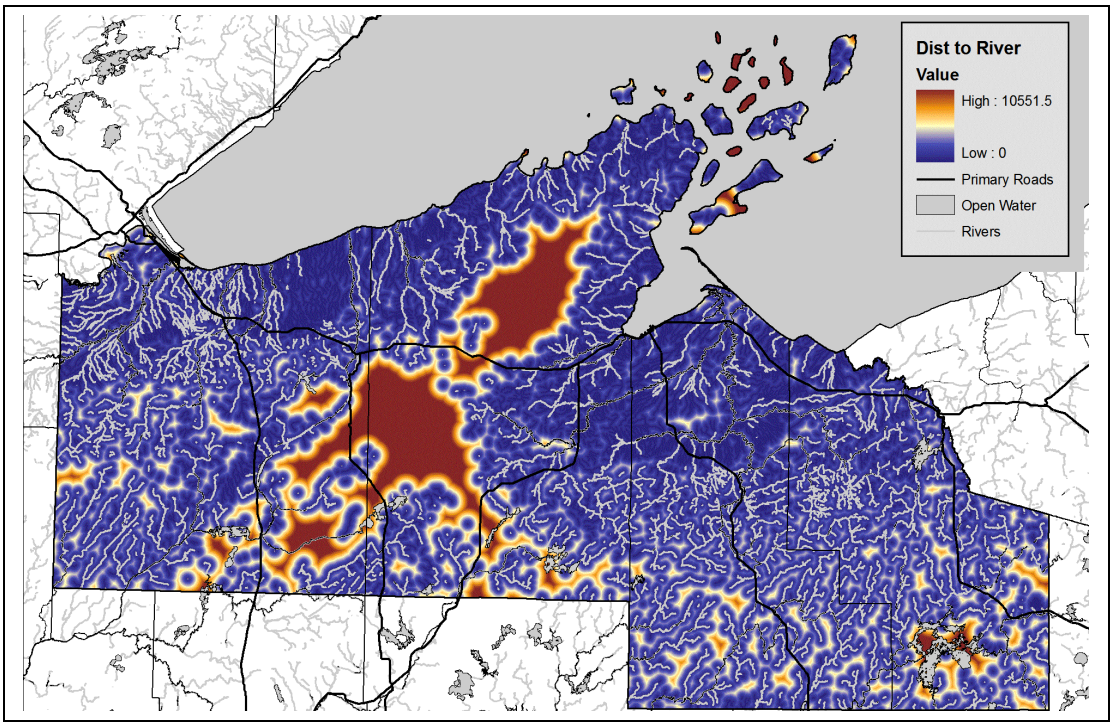

Figure 5. Distance to nearest river (meters), as defined by areas of surface water represented by line features in the original shapefile. Large areas of the Moquah Barrens lack significant river systems.

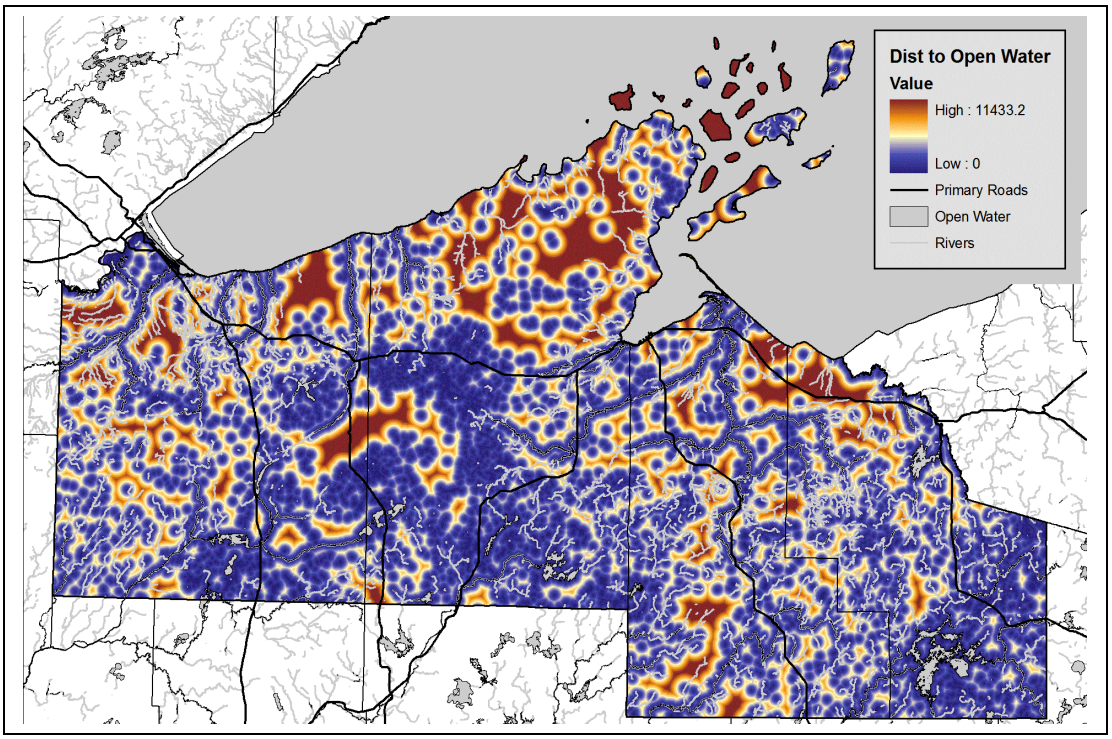

Figure 6. Distance to open water (meters), as defined by areas of surface water represented by polygon features in the original shapefile. Areas farthest from open water include portions of the Moquah Barrens and Lake Superior Clay Plain.

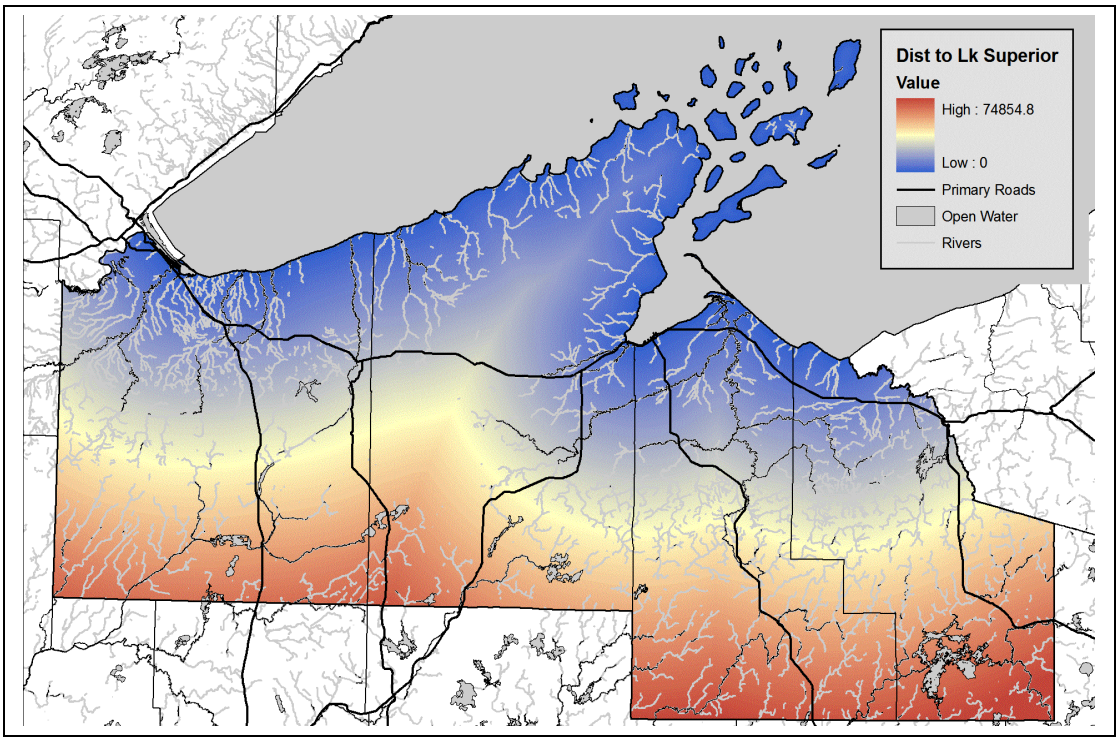

Figure 7. Distance to Lake Superior (meters).

SSURGO soil layers

Soil Survey Geographic (SSURGO) data (version 2.2.3) were downloaded from the USDA-NRCS Soil Data Mart (USDA-SSDS 2007). The scale for these data is 1:12,000. Soil layers used in this analysis are described in Table 2, and maps of these layers are shown in Figures 8 through 14.

County soil data were reprojected to UTM zones 15 and 16N, NAD83 datum. USDA-NRCS Soil Data Viewer 5.2 was used as an extension to ArcGIS 9.3 to select desired soil characteristics, aggregate them and save map units (MUs) and soil characteristics as text files. Layers were aggregated based on dominant condition, higher value for the tie-break rule, and no cutoff percentage. These were converted to shapefiles, reprojected to WTM83 HARN, and exported from ArcGIS. They were then imported into ArcView 3.3, clipped using the 4-county project area boundary, and converted to 30-meter grids.

Shapefiles for each county were aggregated by soil drainage class attribute, using the soil data viewer and ArcMap. Seven classes of natural soil drainage are recognized in the SSURGO data, with lakes, rivers, and other open water areas classified as “missing data”. For each county, drainage categories were reclassified to 1 to 8 as shown in Table 3.

Table 2. SSURGO environmental layers used for analysis. Descriptions condensed from USDA-SSDS 2007.

Attribute	Depth (cm)	Value Range	Description	Figure
Percent clay	0 to 100	0 to 66.2% (by weight)	Mineral soil particles less than 0.002 millimeter in diameter. The estimated clay content of each soil layer is given as a percentage, by weight, of the soil material that is less than 2 millimeters in diameter. The amount and kind of clay affect the fertility and physical condition of the soil and the ability of the soil to adsorb cations and to retain moisture.	8
Drainage class (natural)	N/A	1 to 8 (reclassified as in Table 2)	The frequency and duration of wet periods under conditions similar to those under which the soil formed. Alterations through human activity are not considered unless they have "significantly changed the morphology of the soil."	9
Frost-free days	N/A	85 to 128 days	The expected number of days between the last freezing temperature (0° C) in spring (January-July) and the first freezing temperature in fall (August-December). The number of days is based on the probability that the values for the standard "normal" period of 1961 to 1990 will be exceeded in 5 years out of 10.	10
Organic matter	surface	0.75 to 88.0% (by weight)	Percentage (g) of soil consisting of organic material less than 2 millimeters in diameter.	11
pHw	0 to 100	1.3 to 7.4	Soil acidity or alkalinity. 1:1 water method.	12
Surface texture	surface	0 to 24 (reclassified as in Table 3)	Standard USDA classification, based on percentages of sand, silt, and clay in the fraction of the soil that is less than 2 millimeters in diameter.	13
Water content	0 to 100	1 to 36.7 (% of oven-dry soil weight, 15 bars tension)	The amount of soil water retained at a tension of 15 bars, expressed as a percentage of the oven-dry weight of soil material that is less than 2 millimeters in diameter. Water retained at 15 bars is an estimation of the wilting point. The most important properties affecting water retention are organic matter content, soil texture, bulk density, and soil structure.	14

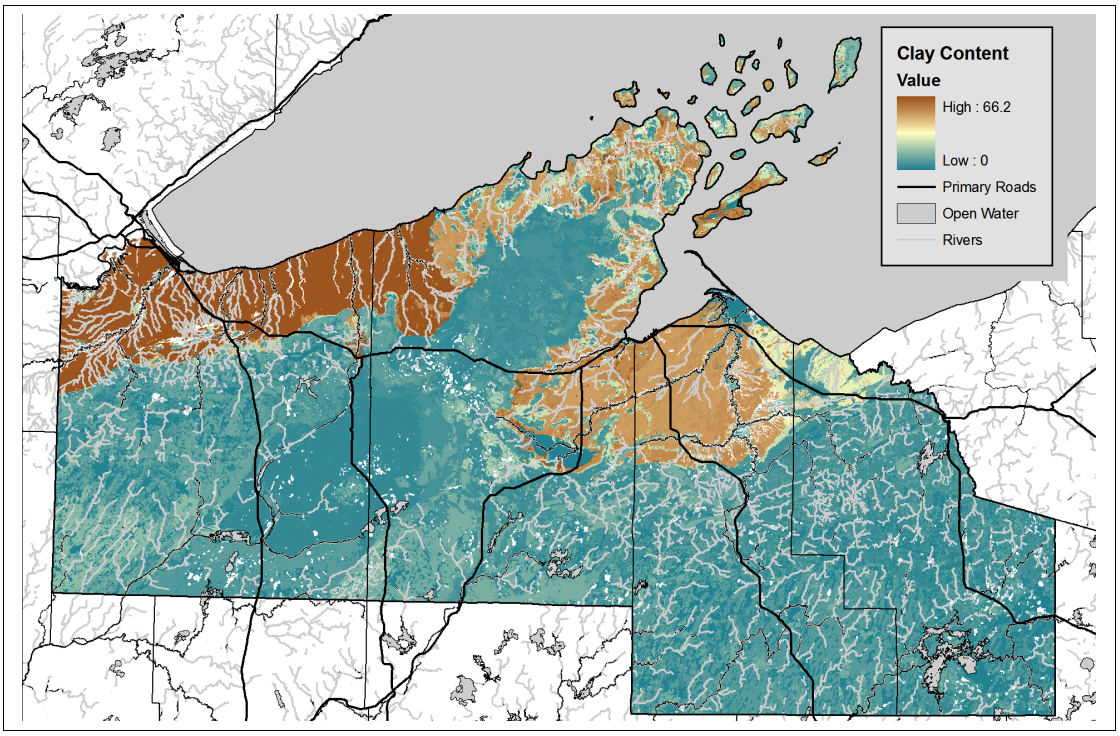

Figure 8. Map of soil clay content. High clay content corresponds primarily to the Lake Superior Clay Plain.

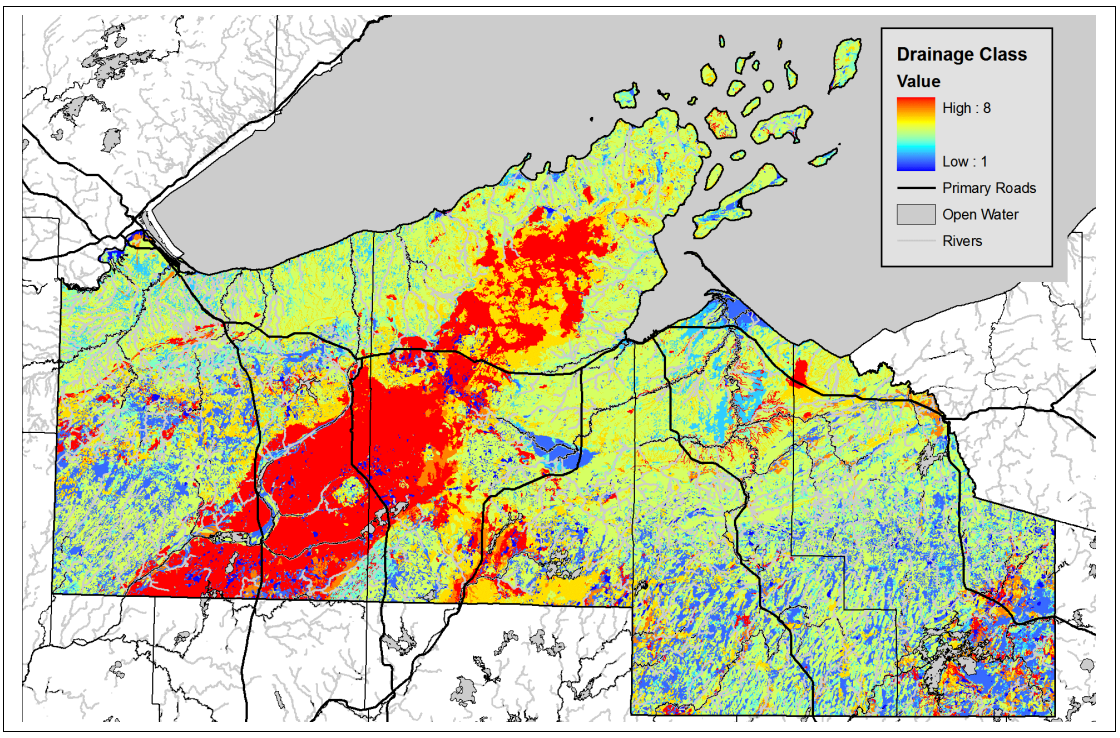

Figure 9. Map of SSURGO soil drainage classes. “Poorly drained” areas (in blue) include the Kakagon Sloughs in northern Ashland County and the Bibbon Swamp in east central Bayfield County. Most of the Bayfield Barrens are classified as “excessively drained”.

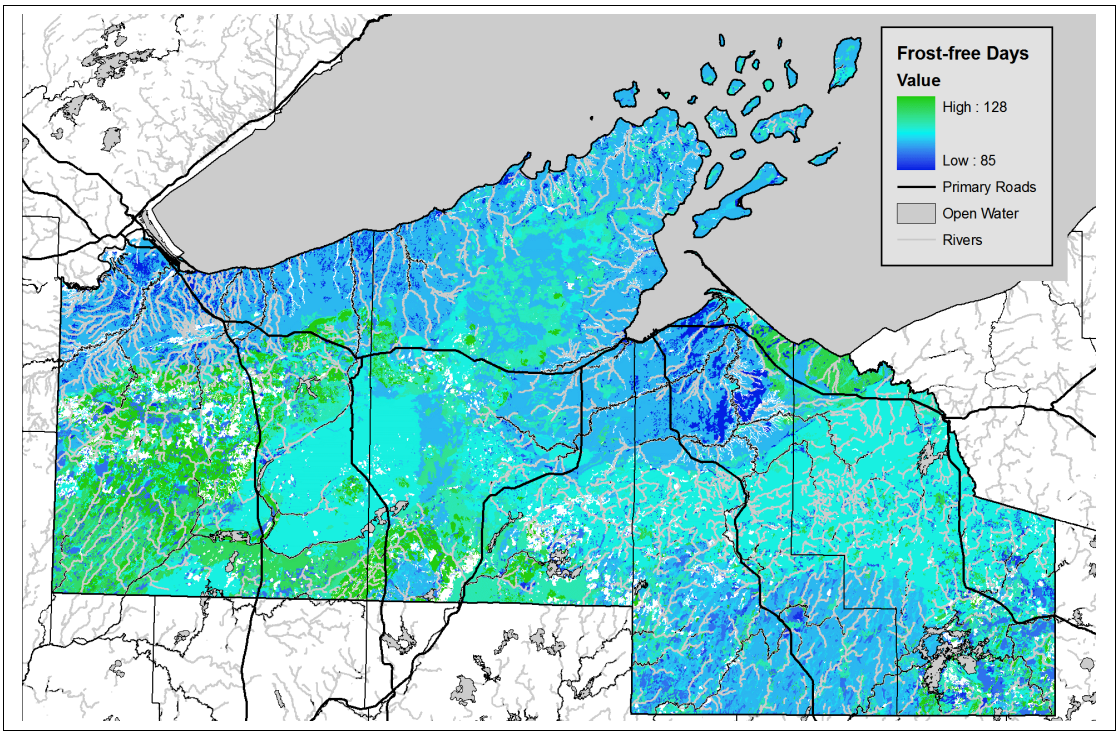

Figure 10. Map of SSURGO average frost-free days. The Lake Superior Clay Plain tends to have the shortest frost-free season, while the longest season areas are concentrated in Douglas, northeast Ashland, and northern Iron Counties.

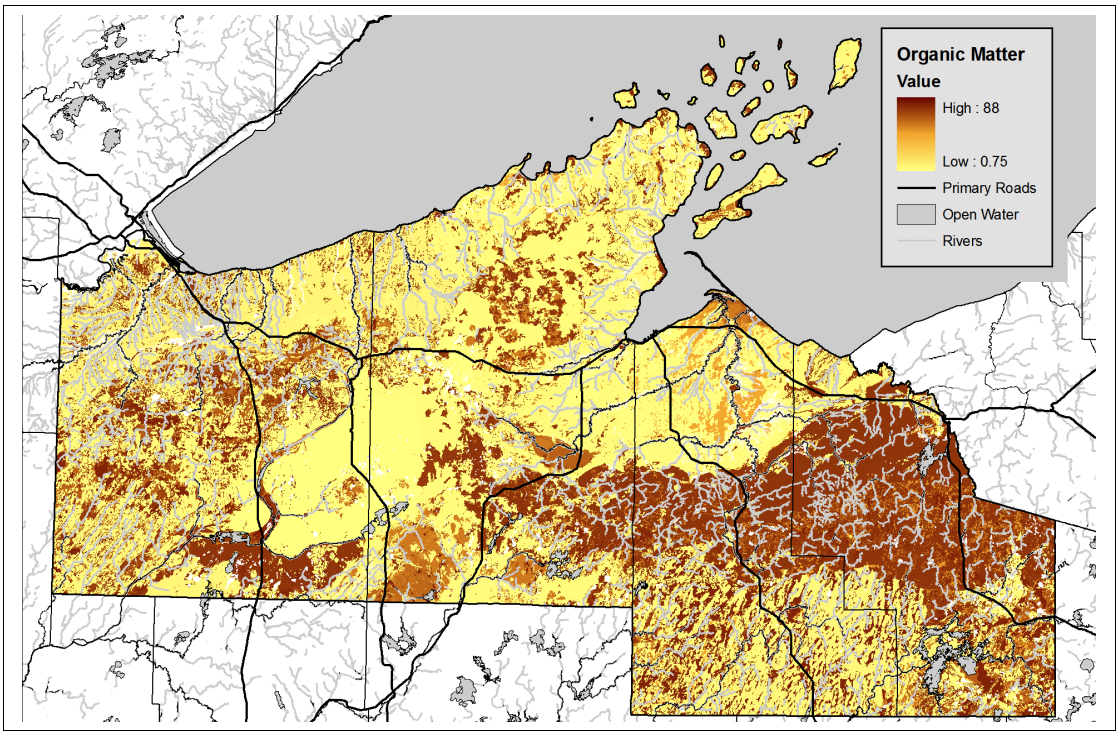

Figure 11. Map of SSURGO soil percent organic matter. Organic matter ranges from low in the Bayfield Barrens and Lake Superior Clay Plain, to high in the Gogebic-Penokee Range.

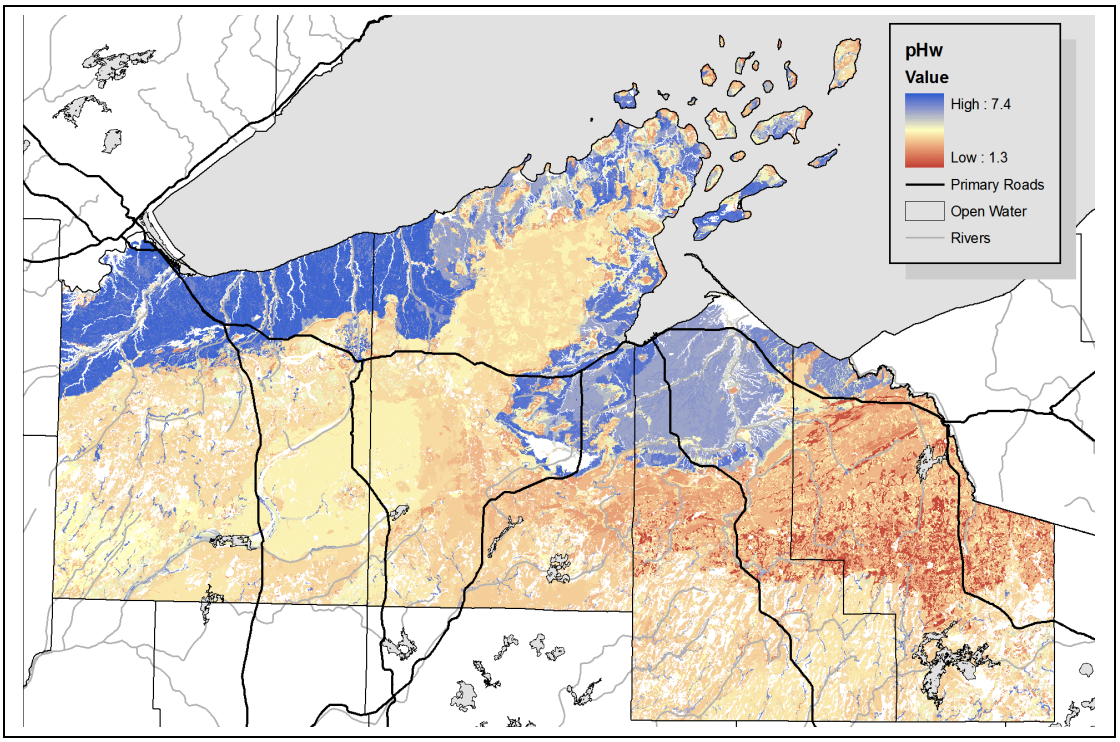

Figure 12. Map of SSURGO soil pH. Soil pH tends to be lowest across the southern portion of the Gogebic-Penokee Range. Soil pH of the clay plain is markedly higher than the rest of the area on average.

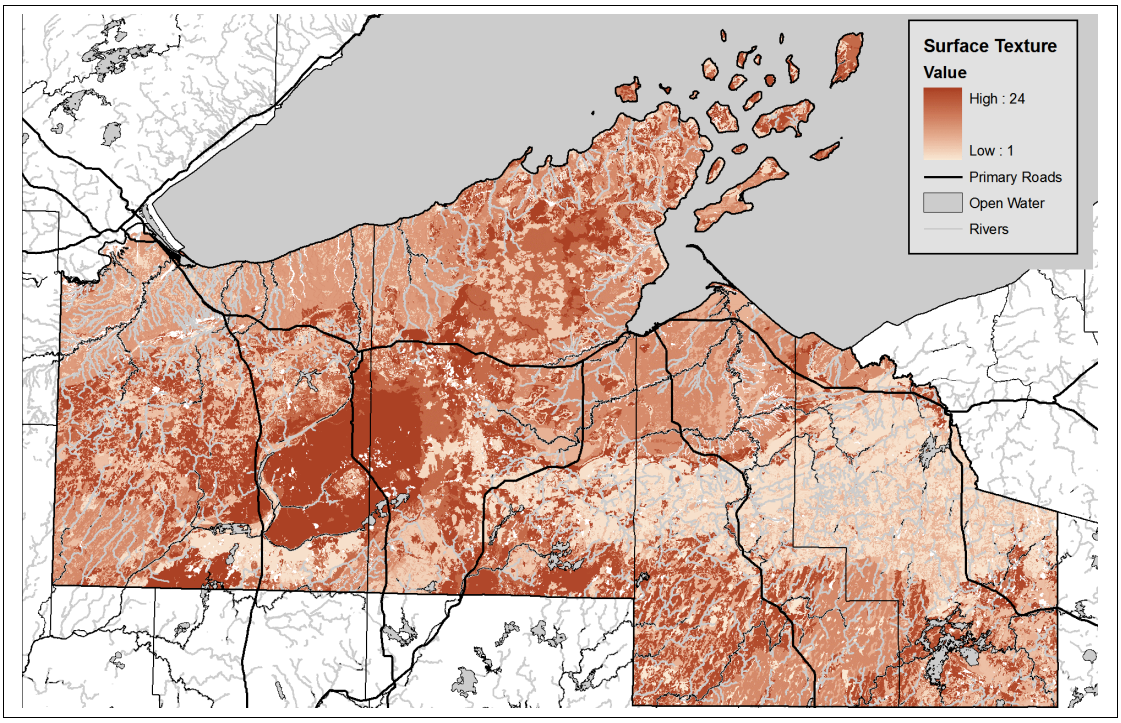

Figure 13. Map of SSURGO soil surface texture. See Table 4 for specific surface texture categories.

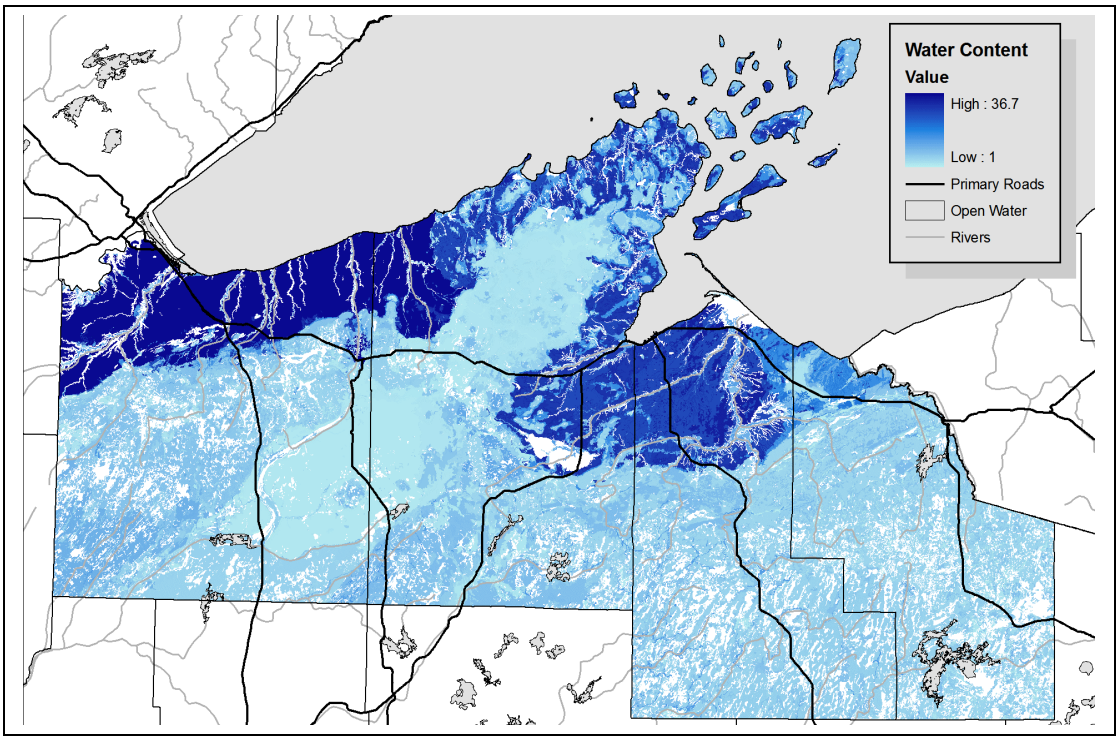

Figure 14. Map of soil water content (15 bars tension). The relatively high water content of the clay plain stands out sharply from the rest of the area, while the largest area of low water content coincides with the Bayfield Barrens.

Soil surface texture was reclassified before being converted to a grid (Table 4). The 24 original categories were reclassified into a continuum of soil textures, running from coarse organic to fine organic to fine inorganic to coarse inorganic, using Chapter 3 of USDA-SSDS (1993) as a guide. Because the reclassification assigned unique values to each data class, this layer could be used as either a categorical or a continuous layer in Maxent.

Unfortunately SSURGO environmental layers include map units with varying amounts of missing data. All the layers lack data for map units without significant surface soil, including open water, rock outcrops and dams. The layers used in this project all included data for essentially all the remaining map units. Layers with substantial missing data for other map units unfortunately could not be used as environmental layers for this analysis.

Table 3. Reclassification of SSURGO soil drainage class values.

Soil Drainage Class	Reclassification Value
None (open water)	1
Very poorly drained	2
Poorly drained	3
Somewhat poorly drained	4
Moderately well-drained	5
Well-drained	6
Somewhat excessively drained	7
Excessively drained	8

Table 4. Reclassification of the SSURGO soil surface texture layer. The 24 original surface texture categories were reclassified into a continuum of soil textures from coarse organic to fine organic to fine inorganic to coarse inorganic as shown.

Original description	Reclass	Original description	Reclass
peat	1	silt loam	13
slightly decomposed plant material	2	channery silt loam	14
moderately decomposed plant material	3	cobbly silt loam	15
moderately/highly decomposed plant material	4	loamy very fine sand	16
highly decomposed plant material	5	loamy fine sand	17
mucky peat	6	loamy sand	18
cobbly mucky peat	7	cobbly loamy sand	19
muck	8	very fine sandy loam	20
stony muck	9	cobbly very fine sandy loam	21
mucky silt loam	10	fine sandy loam	22
silty clay loam	11	sandy loam	23
clay loam	12	sand	24

Distance to nearest city and distance to nearest road

In order to include the effects of roads, cities and towns on the distribution of invasive plants in the project area, grid layers of distance to nearest city and distance to nearest road were produced. The distance to nearest road layer was derived from the 100K TIGER/Line Files (USDC 2000). This road shapefile was imported into ArcView, clipped with a project area layer which included a 5-mile buffer (to include effects of roads just outside the project area), converted to a 30-m grid, and clipped using the project area boundary. The distance to nearest city layer was produced by the same method used for the road layer, except that a 20-mile buffer was used. Maps of distance to nearest road and nearest city appear in Figures 15 and 16.

Background or bias layer

Like most other “presence-only” modelling algorithms, Maxent uses a “background” sample to provide a sample of the range of conditions available in a region (Phillips et al. 2009). Though sometimes referred to as “pseudoabsence” data, these points are not assumed to be true absences. By default, Maxent obtains the background points via a random sample of 10,000 points (here 30 m grid cell values) within the project area.

Like nearly all species occurrence data used for modelling, ours tends to be biased towards roads, cities, towns, and other human-disturbed areas. Approaches to solving this problem (seldom implemented by most habitat modelling studies to date) include attempting to equalize survey intensity by down-weighing or discarding occurrence data in over-sampled regions, or by surveying under-sampled regions. Another approach is to use a “target group” background layer (Dudík et al. 2005, Phillips et al. 2009). Data for broad biological groups (such as all vascular

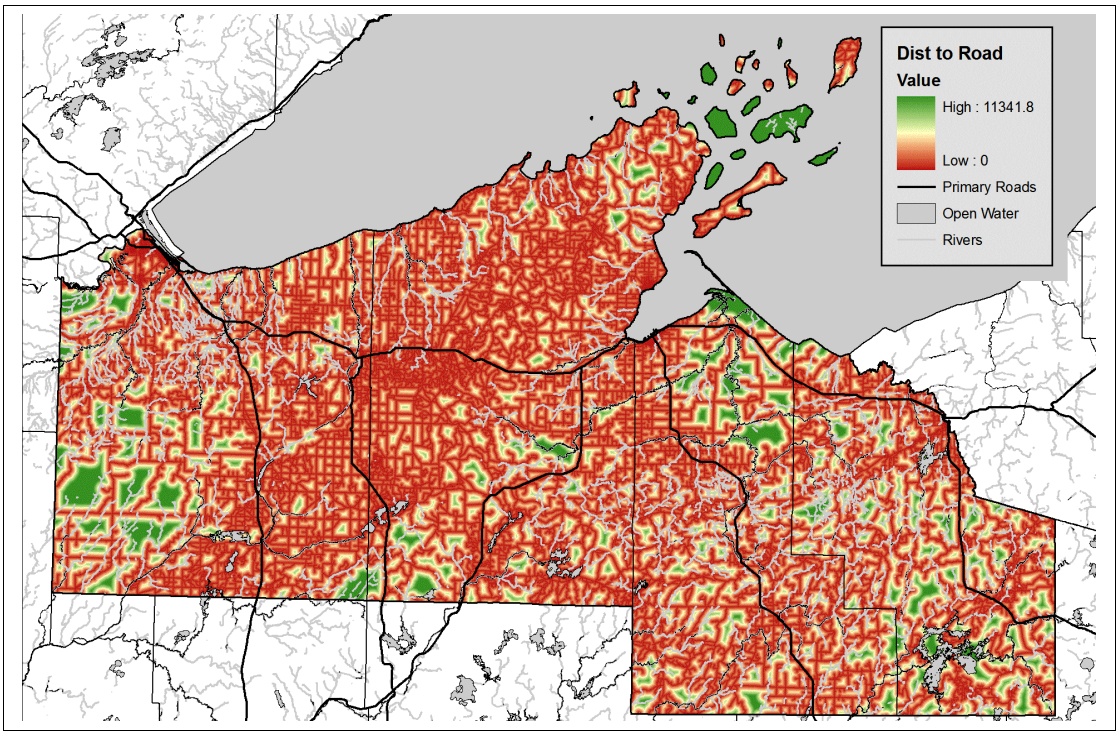

Figure 15. Distance to the nearest road (meters). Areas at the red end of the scale are relatively close to a mapped road; areas at the green end are relatively far away.

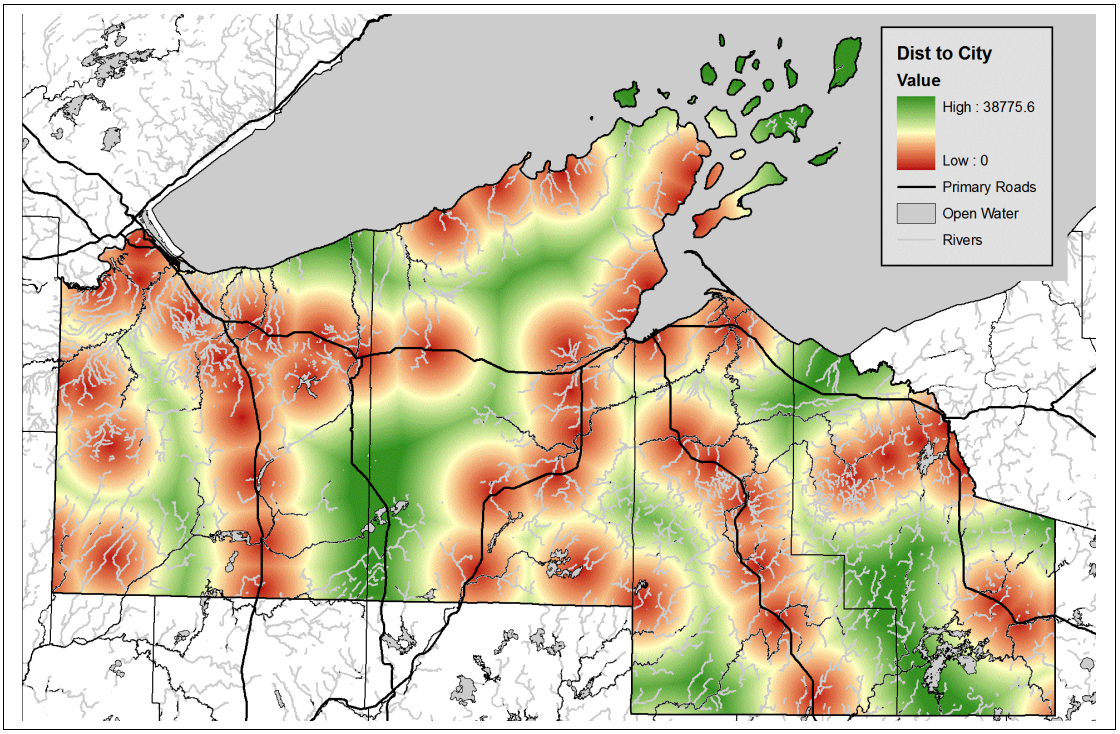

Figure 16. Distance to the nearest city or town (meters). Areas at the red end of the scale are relatively close to a city or town; areas at the green end are relatively far away.

plants sampled during our surveys) that were obtained with the same methods and in the same types of habitats as the species being modelled, can be expected to have the same sampling bias as the study species (Dudík et al. 2005, Phillips et al. 2009). Because our occurrence data is a subset of our target group data, it can treated as a random sample from a distribution with the same locational bias as the occurrence data. Phillips et al. (2009) shows that using such a target group background layer can be very effective in factoring out bias when modelling habitat distributions. The resulting models often have somewhat lower predictive ability, but are usually more realistic and do a substantially better job of modelling regions well away from the known occurrences.

For this project we were fortunate to have a relatively large target-group data set, appropriate for use as a background grid. These data were collected by the same methods and obtained from the same sources as the data for the species being modeled, so they should have the same bias as the modelled data (Phillips and Dudík 2008; Phillips et al. 2009). The set of 5,710 non-cultivated invasive plant occurrences within the project area was converted to a grid layer. Because two or more invasives sites sometimes fell within the same 30-meter grid cell, the resulting layer contained only 4,904 “presence” cells. This layer was used as background data for factoring out locational bias in the data.

Environmental and background layers were exported from ArcView as 30 m ASCII grid files. For the background layer, background points are denoted by 1, with missing data values (-9999) for the rest of the points.

Model analysis

The possible outcomes for each prediction of presence or absence can be summarized in a confusion matrix (Figure 17). Given the fact that a species will be either present or absent in each discrete location of an area (here defined by 30-meter grid cells), a prediction of presence or absence can lead to four possible outcomes.

The omission rate is the proportion of absences predicted as present (type I error rate), while the commission rate is defined as the proportion of presences predicted as absent (type II error rate). Prevalence [a/(a+b)] was found to be the best predictor of model effectiveness by Liu et al. (2005).

The accuracy and reliability of the models was assessed using several accepted statistical methods. The “omissions curves” produced by Maxent plot the cumulative threshold for background data, training and (if applicable) test curves on the x-axis, versus the fractional predicted area (cumulative proportion of total project area where the species is predicted as present) on the y-axis. The predicted omissions rate is a straight line from the origin to point (1.0, 1.0), as defined by the logistic output format. The omissions rate (proportion of false negatives, or type II error rate) should be close to the predicted omissions rate. If the test curve is well below the predicted omissions rate, the graph indicates potential problems with the model. Such a test curve indicates lack of independence between the training and test data, often due to spatial autocorrelation of the presence data (Phillips 2009).

		OBSERVED	
		Present	**Absent**
PREDICTED	**Present**	a: true positive	b: false positive (type I error)
	Absent	c: false negative (type II error)	d: true negative

Figure 17. A confusion matrix.

The receiver operating characteristic (ROC) plot and associated area-under-the-curve (AUC) statistic is probably the single most relied-upon measure of model efficacy in the area of SDM today (Peterson et al. 2008). The standard ROC technique plots the false positive fraction values (1-specificity, or type II error rate) on the x-axis, versus the true positive fraction (sensitivity) values on the y-axis (Fielding and Bell 1997). For the presence-only data typical of most habitat modelling studies (including this one), no species absence data exists. The ROC curve is therefore modified to plot the true positive fraction against the fractional predicted area (Phillips et al. 2006).

The area under the ROC curve can be interpreted as the probability that a randomly chosen presence site is ranked above a randomly chosen background site (Phillips et al. 2006). Models are usually constructed using a random subset (typically 60-75%) of the data ("training data"), and tested using the remainder of the data ("test data"). ROC curves are then calculated for both sets of the data, with the area under the curve (AUC) statistic giving a measure of the predictive ability of the model. For standard ROC curves this area can range from 0.5 to 1.0, but because fractional predicted area is substituted for true commission the maximum is somewhat less than one (Phillips et al. 2006). As a general rule, Elith et al. (2006) consider models with ROC curve values above 0.75 to be useful, while Phillips and Dudík (2008) consider models with a minimum value of 0.70 to be useful.

Maxent also produces several other useful analyses. A set of response curves for each environmental variable (layer) show how each variable affects the prediction as all the other variables are kept at their average (background) sample value. Another set of response curves shows how each variable would affect the model if it were the only one used. A table gives the heuristic estimate of the percent contribution of each variable to the model. Plots show jackknife training gain, test gain and AUC values for each variable when the variable is the only one used to build the model, and when each variable is omitted from the model in turn. Finally a map of the model is produced, showing the probabilities of species presence across the project area.

Current versus potential distribution

For each taxa, two conceptually different models were produced. First, a model was produced using only environmental layers representing the natural environment. This model shows areas the plant in question may be capable of colonizing, and areas that are likely to be unsuitable for it. This type of model is intended to show a plant's *potential* distribution. (The concept of "potential distribution" is closely related to what ecologists call a species' "fundamental niche.")

After this “potential distribution model” was produced, a model reflecting the species’ “current distribution ” (analogous to its “realized niche”) was constructed. Depending on factors such as amount of time since introduction, predisposition to dispersal on vehicles and equipment, and affinity to disturbed habitats, some species are much more likely to be found in the general area of human activity as opposed to more remote natural areas. And all other things being equal, an invasive species is more likely to spread to nearby areas than to distant sites. Therefore these “current distribution” models generally included the environmental layers used for the corresponding potential distribution models, plus one or both human-imposed habitat layers of distance to roads and distance to cities.

The distinction between these two types of models has been discussed by Jones and Reichard (2009), Soberon and Nakamura 2009, and others. Soberon and Nakamura (2009) note that species distribution methods relate ‘‘niches’’ to ‘‘areas of distribution.’’ They also point out that rigorous definitions of these concepts have not as yet been formulated.

Modelling methodology

Maxent version 3.3.2 was used to construct the final models. It was downloaded from www.cs.princeton.edu/~schapire/maxent/.

Construction of potential distribution models began with all the environmental layers included, except that the original land cover layer was used instead of the reclassified “shade” layer mentioned above. Typically a second run would be completed, with the shade layer substituted for the land cover layer. The two models were evaluated, and subsequent runs were done with the best-performing of the two layers (usually the land cover layer). The omissions curves, ROC curves and AUC statistic, percent contribution, variable response curves, and jackknife gain plots (especially the test gain plots) were used to evaluate each model. One-by-one, variables with low percent contribution to the model (typically under 2%) and with level or nearly-level response curves were eliminated and a new run was completed. As the environmental layers were narrowed down, layers that were previously eliminated were sometimes reintroduced or substituted for a related layer (e.g., water content for drainage class) to ascertain their effects. (Usually but not always a variable that performed poorly early on performed similarly poorly when reintroduced.) Particular attention was paid to variables suspected or known to be important to each species (e.g., water content and drainage class for wetland species). This process was repeated until a model was produced where all variables positively contributed to the model.

Typically 20-30 models were attempted before arriving at a “good” model for each plant. All these models were initially run using the target-group background layer, 25% of data used for testing, a maximum of 1,000 iterations for algorithm convergence (default is 500), and one replication for each run.

Once a “good” model (or several comparably “good” models) was obtained, these models were rerun using 10 replications for each model. A final “potential distribution” model was chosen. The distance to road and distance to city layers were then added to this “potential distribution”

model, and the new model was run with 10 replications and evaluated. Finally, a model with one or both of these layers included was chosen as the final “current distribution” model.

RESULTS

Modelling strategy

Habitat modeling is an interplay between the relevance of the environmental layers, the test statistics for the model, and the “reasonableness” of the model. Ecological relevance of environmental layers relative to the spatial distribution of each species is an important consideration. For example, when attempting to model a still highly-localized species such as alpine oatgrass, climate layers such as minimum temperature are obviously irrelevant to the model (though they may still be relevant to the species' potential habitat). Also, environmental layers may interact with each other and thus indirectly affect the model, even though they may have little influence on their own.

The target-group layer had a substantial positive effect on most of the models. This was obvious both visually (from the maps) and from the diagnostic data. The most obvious effect was often a large reduction in influence of the “distance to nearest road” layer. This road layer typically dominated the current distribution models that were constructed using the default 10,000 randomly-located background points, but usually had a subordinate or even minor influence when the target-group background layer was used.

When used as a continuous predictor, the soil surface texture layer was usually ineffective. Therefore this layer was only used as a categorical layer for these models.

Variables such as soil drainage class and soil water content are presumably correlated to some degree. Yet these two layers often seemed to act independently in the models.

Invasive plant models

Wetland herbaceous plants

European yellow iris: The test data AUC values for both the current and potential distribution models for this species were much higher than the 50% (0.500) value expected of the corresponding random models (Figures 18a and b). Inclusion of the distance to open water layer improved the potential distribution model somewhat, increasing the AUC for training and test data from the original 0.927 and 0.874±0.062 to the values in Figure 18b.

Addition of the distance to river layer lowered the AUC of both models slightly. The response curves for the river layer (not shown) clearly indicated increasing influence with distance to roughly 600 m from the nearest river, and decreasing influence beyond 600 m. Because the river layer seemed to have a consistent effect on both models, it was retained in the final models.

Unfortunately nearly half of the 40 known sites for this plant were dropped from the analysis due to missing environmental data in the three SSURGO soil layers. Missing environmental data was an inordinately large problem for this species because it frequently occurs along the edges of waterbodies. The SSURGO layers classify open water as “missing data”, and slight locational inaccuracies in the occurrence data and/or environmental layers resulted in some occurrences falling within open-water areas.

Distance to road was a minor factor in the predicted current distribution of yellow iris (Figure 18a). Distance to city had almost no effect on the model so this layer was excluded.

It should be noted that while the maps in Figures 18a and b appear to show nearly uniform low probabilities of occurrence for yellow iris across the project area, narrow bands of high probability (barely visible at this scale) border most of the waterbodies across the region.

Purple loosestrife: Despite a large amount of data for this species across the project area, a completely satisfactory model for purple loosestrife proved to be elusive. On the positive side, the region of well-drained, coarse soils extending from southern Douglas County northeast into the Bayfield peninsula (generally poor loosestrife habitat) clearly shows up as low-probability on both the current and potential distribution models (Figures 19a and b). Also as expected, the high-probability (pink to red) areas in northern Ashland County and the wetter areas of Bayfield County near Lake Superior correspond to areas of high water content.

On the negative side, much of southern and western Douglas County consists of large open wetlands, floodplains and swamp forest (pers. obs.), but the western portion of this county in particular was predicted to be relatively poor habitat for purple loosestrife. The same is true for the clay plain of northern Douglas and western Bayfield County, where (especially on the western end of this feature, in northwest Douglas County) much suitable habitat for this plant would appear to exist.

Water content was a major factor for both models, increasing in influence until about 28% water, then dropping off. Probability of occurrence also increased with percent organic matter, which is presumably positively correlated with high soil water content. The significance of the high influence of soil pH on the models is unclear: the diagnostic curves show the influence of pH to be moderate overall, and increasing only slightly as pH goes from highly acid to just above neutral.

The relatively high probabilities of occurrence assigned to northern Ashland County may to some degree be an artifact of greater sampling intensity in that area. During the early days of the GLIFWC invasive species program, the Bad River Reservation and surrounding lands were the focus of loosestrife mapping and control activities, so this area was surveyed for loosestrife earlier and with greater intensity than areas farther away.

PARAMETERS	
Presence points:	21-22
Background points:	4123
Training AUC:	0.987
Test AUC±SD:	0.952±0.020
LAYERS	**PCT**
Distance to open water	67.1
Land cover	15.4
Water content (15 bars)	5.8
Distance to road	5.6
Distance to river	2.6
pH (1:1 water)	1.8
Clay (%)	1.7

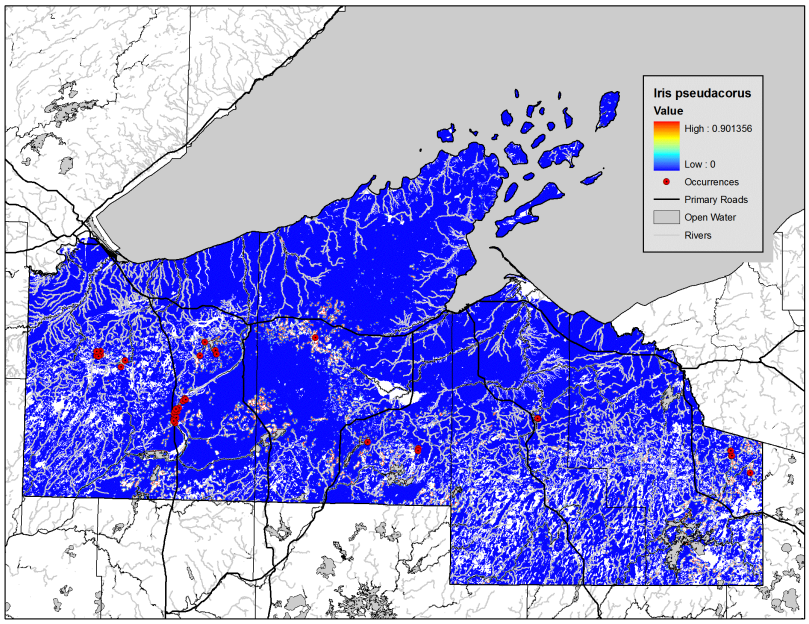

Figure 18a. Current distribution for yellow iris.

PARAMETERS	
Presence points:	21-22
Background points:	4123
Training AUC:	0.985
Test AUC±SD:	0.955±0.020
LAYERS	**PCT**
Distance to open water	70.0
Land cover	16.7
Water content (15 bars)	5.4
Distance to river	3.2
pH (1:1 water)	2.7
Clay (%)	2.1

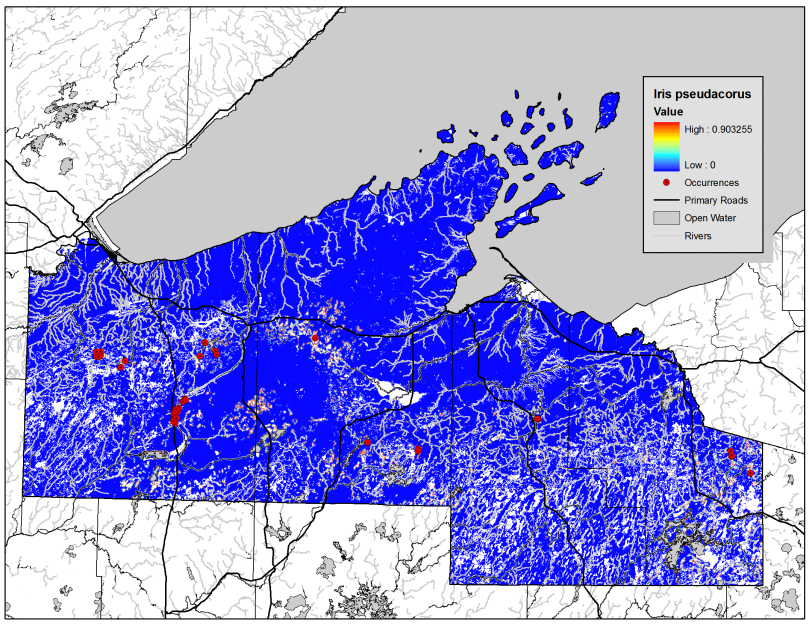

Figure 18b. Potential distribution for yellow iris.

PARAMETERS	
Presence points:	566-567
Background points:	4011
Training AUC:	0.831
Test AUC±SD:	0.795±0.018
LAYERS	**PCT**
pH (1:1 water)	19.6
Water content (15 bars)	15.6
Organic matter (%)	15.0
Clay (%)	13.2
Distance to city	11.6
Land cover	8.0
Drainage class	4.9
Distance to road	4.4
Ave frost-free days	2.8
Distance to open water	2.7
Distance to river	2.2

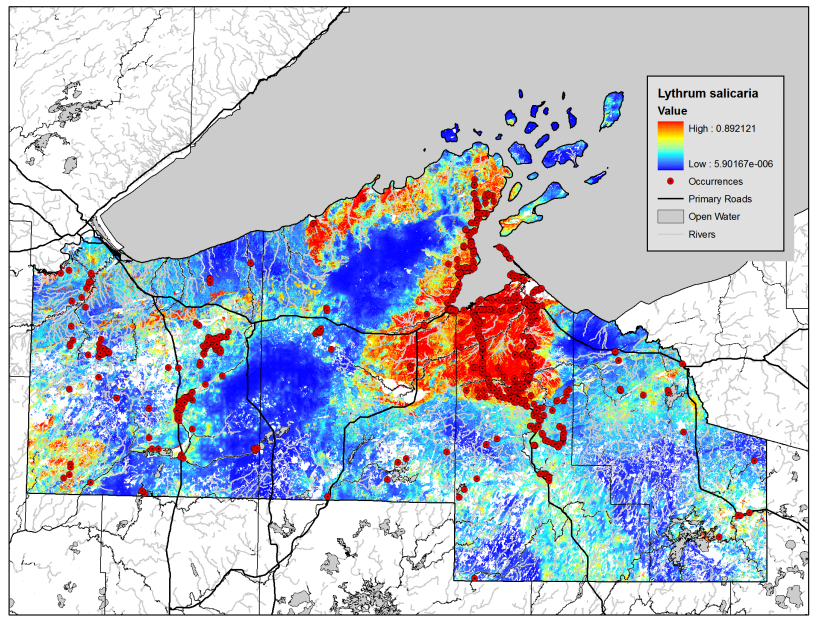

Figure 19a. Current distribution for purple loosestrife.

PARAMETERS	
Presence points:	566-567
Background points:	4011
Training AUC:	0.814
Test AUC±SD:	0.781±0.019
LAYERS	**PCT**
pH (1:1 water)	23.4
Water content (15 bars)	18.9
Organic matter (%)	16.7
Clay (%)	15.1
Land cover	9.2
Drainage class	7.0
Distance to open water	4.6
Ave frost-free days	3.2
Distance to river	1.9

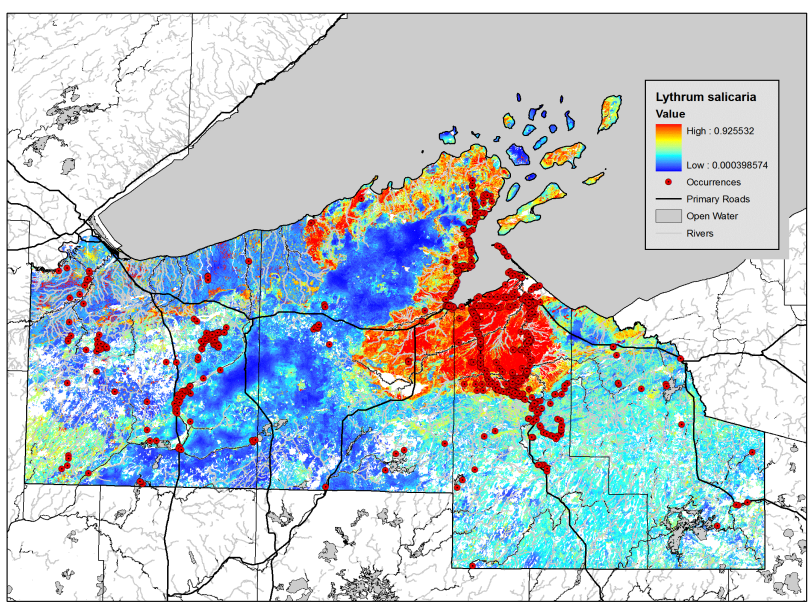

Figure 19b. Potential distribution for purple loosestrife.

Upland herbaceous plants

Alpine oatgrass: The current distribution model for alpine oatgrass (Figure 20a) reflects the fact that most of the known colonies of this plant (currently the only ones recorded in Wisconsin) occur along Highways 13 and 77 in southeastern Ashland County. While alpine oatgrass is clearly capable of invading open to moderately shaded habitats away from roads (Dixon 1991, pers. obs.), its current expansion seems to be primarily along these road corridors at present (pers. obs.).

The potential distribution model for alpine oatgrass shows high scores for both the training and test curves (Figure 20b). The 16 points available to this model are all in relatively close proximity to one another, though, with most found along a 20 km stretch of Hwy 13 and 77, north and west of Glidden. The habitat currently occupied by this population appears to be fairly homogenous (pers. obs.), making this relatively small part of the project area relatively "unique" compared to the much more diverse project-wide background layer.

For both models the land cover layer gives higher probability of occurrence with grassland, maple and broad-leaved deciduous woods. The high probability of these last two categories likely reflects the habitat on either side of some of the road populations. The same is true for the high probability of occurrence associated with "peat" in the surface texture layer - while none of the populations grow on peat (pers. obs.), some do occur along roads that run through peaty wetland areas. More appropriately, high probability is also associated with several categories of coarser soils.

The relatively limited area in which alpine oatgrass is currently found tends to have fairly coarse soils. This is consistent with the strongly decreasing probability of occurrence with increasing percent clay seen in both models. The current distribution model shows probability dropping off sharply from about 0.8 to nearly zero with distance from the nearest road.

Because so little data are available for alpine oatgrass in the project area, models for this species should be considered "exploratory".

PARAMETERS	
Presence points:	15-16
Background points:	4329
Training AUC:	0.927
Test AUC±SD:	0.874±0.062
LAYERS	**PCT**
Land cover	41.5
Surface texture	35.0
Clay (%)	14.8
Distance to road	8.7

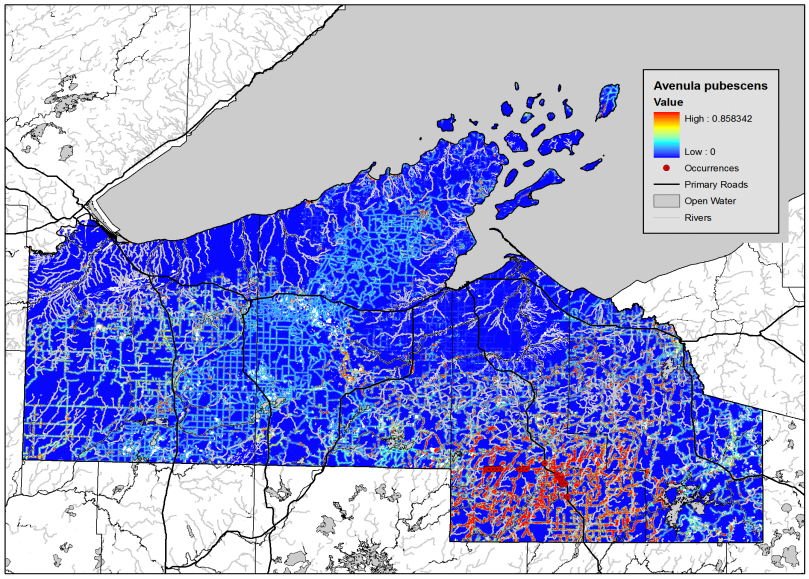

Figure 20a. Current distribution for alpine oatgrass.

PARAMETERS	
Presence points:	15-16
Background points:	4329
Training AUC:	0.972
Test AUC±SD:	0.918±0.024
LAYERS	**PCT**
Land cover	46.0
Surface texture	37.7
Clay (%)	16.4

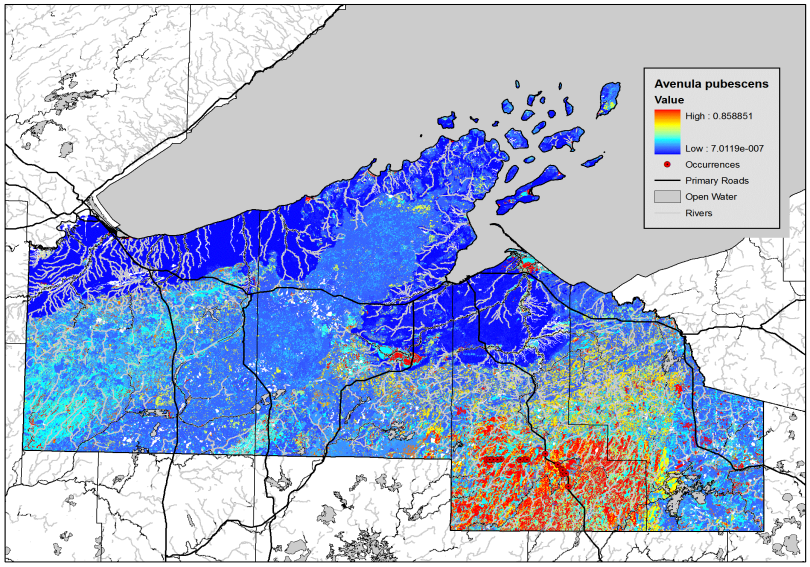

Figure 20b. Potential distribution for alpine oatgrass.

Leafy spurge: For both the current and potential distribution models, surface texture was a major environmental predictor for leafy spurge, followed by land cover, soil organic matter, and soil water content (Figures 21a and b). The highest probability of occurrence was on loamy fine sand, with (oddly) highly and moderately decomposed plant material second and third. “Loamy sand” was fourth. The land cover layer class associated with the highest probability of occurrence was “herbaceous/field crops”, followed by “mixed deciduous/coniferous”, with “barrens” and “broad-leaved deciduous” tying for third. Probability of occurrence increased with increasing water content in both models.

When added to the model, distance to road became the third most important factor (Figure 21b). Judging from our surveys, existing populations do appear to be strongly associated with roads (pers. obs.). Distance to city (model not shown) was an ineffective predictor for leafy spurge, with a nearly level response curve and a 3.0% contribution to the model.

Woodland sweet pea: Because of the concentration of points near Lake Superior, the soil layers seemed to have relatively little effect on the distribution of woodland sweet pea. I finally settled on models dominated by average annual minimum temperature and distance to Lake Superior (Figures 22a and b). For both models, probability of occurrence increased from near zero to around 0.8 with increasing average annual low temperatures. As shown in the models, probability was highest in the Lake Superior region.

The high probability of occurrence near Lake Superior does reflect the on-the-ground distribution of woodland sweet pea. Within the project area sweet pea is most abundant on the northern end of the Bayfield Peninsula (pers. obs.), where the average annual minimum temperatures are the highest in the region. It is also common east of the project area, in northern Ontonagon County, Michigan (pers. obs.). It seems possible that this plant’s present distribution and abundance may be due in part to the microclimatic influence of Lake Superior.

Even though any bias towards roads in the data was presumably neutralized by the target-group background layer, the current distribution model still shows higher probabilities for woodland sweet pea along road corridors. This is consistent with our field observations.

PARAMETERS	
Presence points:	80-81
Background points:	4129
Training AUC:	0.824
Test AUC±SD:	0.734±0.046
LAYERS	**PCT**
Surface texture	36.3
Land cover	23.9
Distance to road	21.2
Organic matter (%)	13.9
Water content (15 bars)	4.7

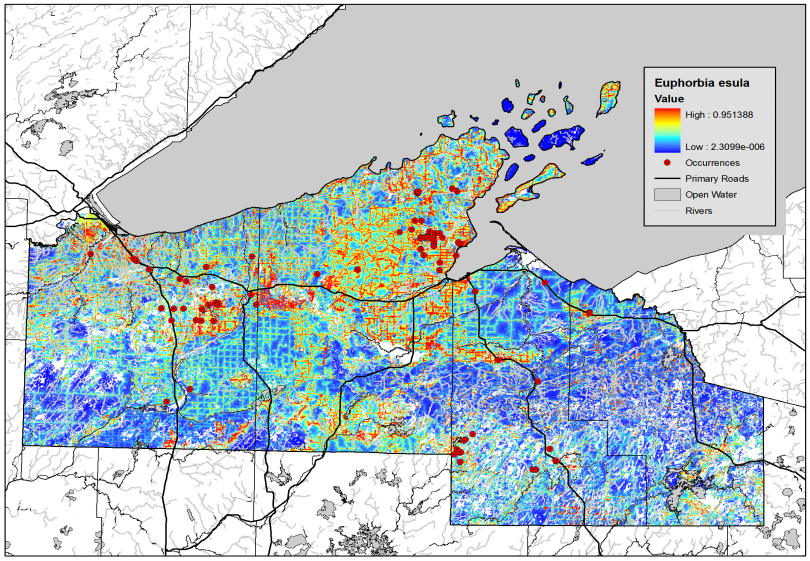

Figure 21a. Current distribution for leafy spurge.

PARAMETERS	
Presence points:	80-81
Background points:	4129
Training AUC:	0.810
Test AUC±SD:	0.707 ± 0.050
LAYERS	**PCT**
Surface texture	45.4
Land cover	30.5
Organic matter (%)	17.5
Water content (15 bars)	6.6

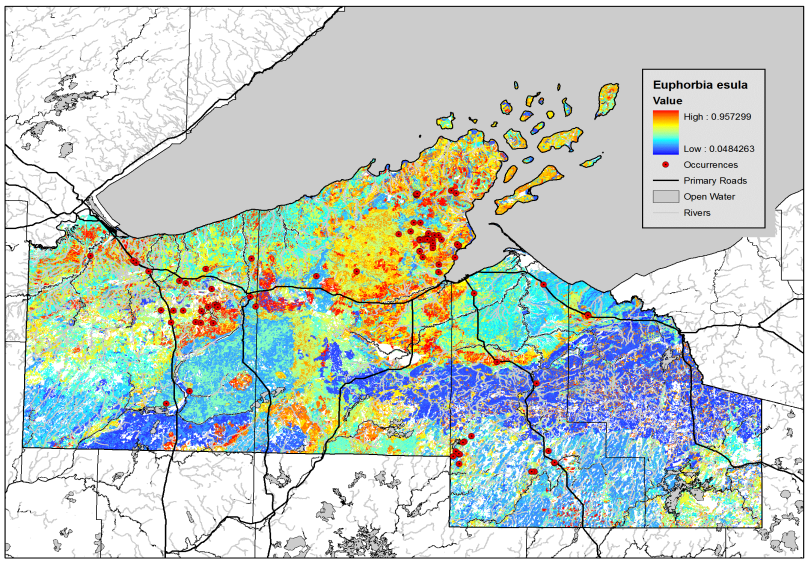

Figure 21b. Potential distribution for leafy spurge.

PARAMETERS	
Presence points:	60-61
Background points:	4329
Training AUC:	0.928
Test AUC±SD:	0.892±0.037
LAYERS	**PCT**
Ave annual min temp	41.5
Distance to Lk Superior	29.8
Land cover	11.6
Distance to road	7.4
Distance to city	4.9
Surface texture	4.9

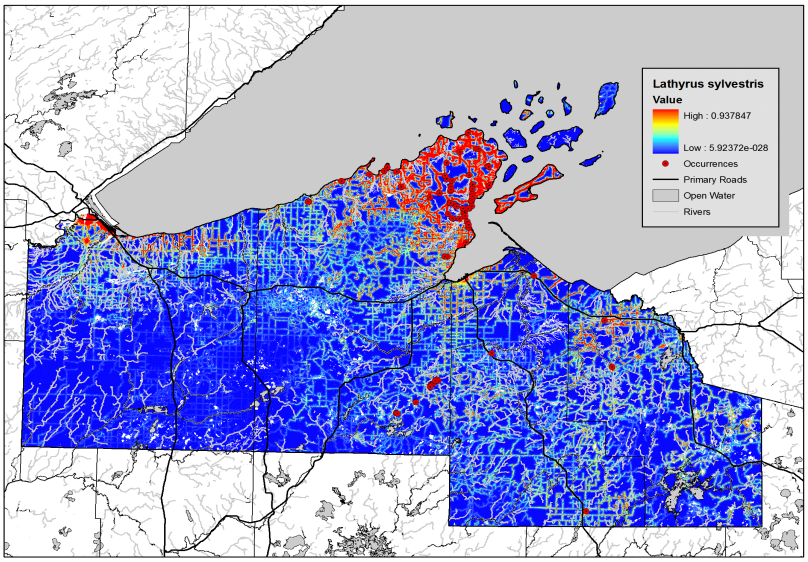

Figure 22a. Current distribution for woodland sweet pea.

PARAMETERS	
Presence points:	60-61
Background points:	4329
Training AUC:	0.900
Test AUC±SD:	0.876±0.044
LAYERS	**PCT**
Ave annual min temp	47.9
Distance to Lk Superior	33.5
Land cover	12.9
Surface texture	5.7

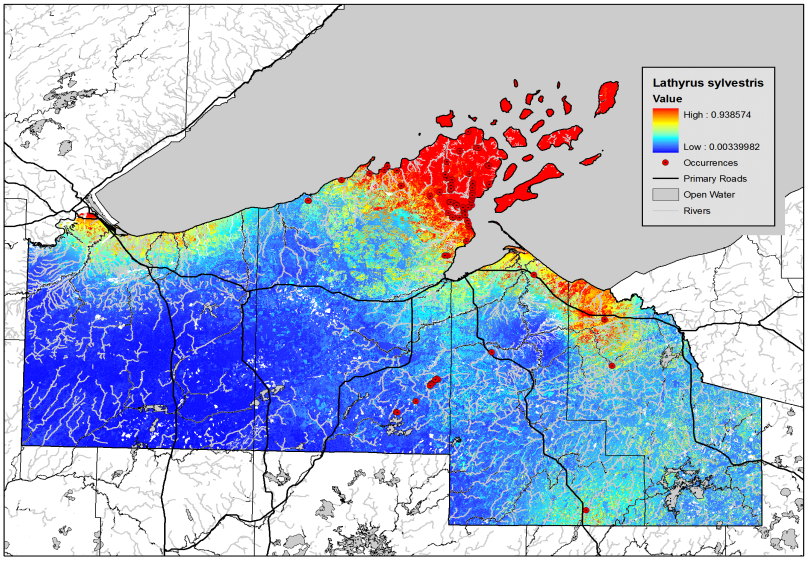

Figure 22b. Potential distribution for woodland sweet pea.

Soapwort: Soapwort proved difficult to model, with most of the layers for most of the attempted models having negative gain. This included all models with pHw (the most influential layer by far) omitted. These poor results were presumably an artifact of the small data set and the fact that most of the known occurrences are clustered along a relatively short stretch of the Bad River floodplain in east central Ashland County.

The best model in terms of AUC and other diagnostics was one dominated by soil pH (Figures 23a and b). The pHw response curve for this model indicated that the probability of occurrence was highest (around 0.7) at the low end of the scale, decreasing slightly as pH increased, then dropping off rapidly in reverse-sigmoid fashion at about pHw 3.4, until hitting a probability of nearly zero at around pH 7. This result conflicts with the pH range of 5.0-7.0 given for this plant by USDA-NRCS (2010). Nonetheless the (rather limited) set of occurrences in our database nearly all occupy areas of low pH, so it is not surprising that Maxent produced a model reflecting this.

Despite the concentration of data along the Bad River floodplain, the soapwort model seems reasonable in some ways. It predicts the probability of soapwort in the northern hardwood region in the east central portion of the project area to be fairly low, which is what would be expected of this moderately (at best) shade-tolerant species. Probability across the sand country (southwest-trending from the Bayfield peninsula) is fairly high, which is also consistent with the habitat requirements of soapwort.

Common valerian: Valerian was so abundant in northwest Douglas County (both along and away from road corridors) that during the 2008 survey of that county I was reduced to marking it at the midpoints between crossroads. Even then many midpoint occurrences went unrecorded, especially those close to Superior, Wisconsin. These data collection shortcomings are the cause of the artificially “patterned” appearance of data for this species in northern Douglas County evident from the maps (Figures 24a and b).

Soil water content had the strongest influence on both the current and potential habitat models. The response curves for both models (all data) shows probability increasing sharply from 1.2% to around 2.0% water content, then roughly leveling off.

The second most influential factor in both models was pHw. For both models, probability of occurrence remained near zero until around pH 5, when it shot up to around 0.2 and 0.3 for the current and potential (habitat) models, respectively. At around pH 6.8, the probability jumped again, to about 0.3 and 0.5, respectively.

Distance to nearest road and nearest city had only a moderate influence on the model (Figure 24a). This canceling out of the “road effect” is no doubt due to the effectiveness of the target-group background layer.

PARAMETERS	
Presence points:	28-29
Background points:	4152
Training AUC:	0.880
Test AUC±SD:	0.770±0.076
LAYERS	**PCT**
pH (1:1 water)	67.6
Shade	20.5
Drainage class	6.2
Distance to road	2.9
Clay (%)	1.5
Distance to city	1.4

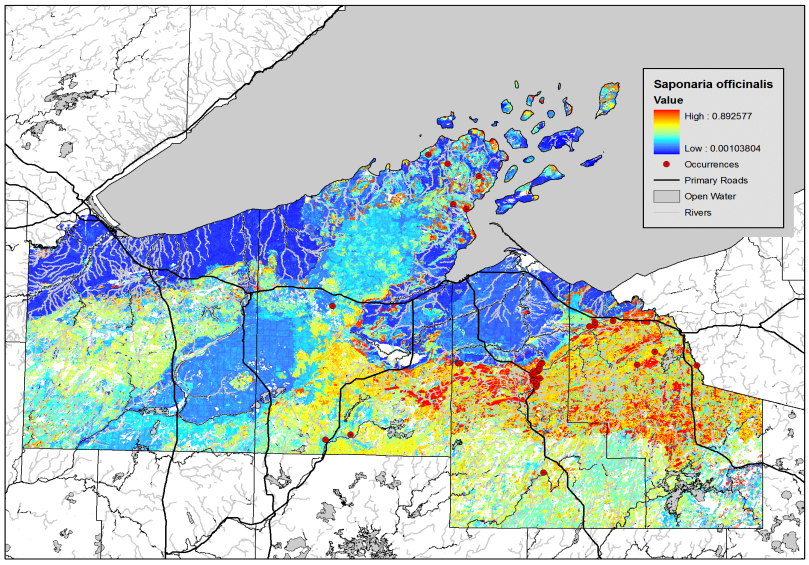

Figure 23a. Current distribution for soapwort.

PARAMETERS	
Presence points:	28-29
Background points:	4152
Training AUC:	0.847
Test AUC±SD:	0.808±0.066
LAYERS	**PCT**
pH (1:1 water)	81.7
Drainage class	7.5
Shade	5.8
Distance to road	3.5
Clay (%)	1.4

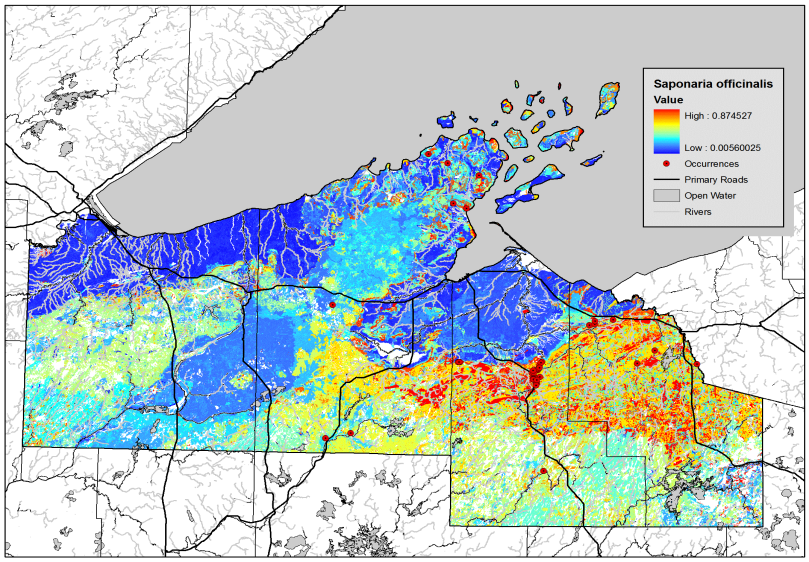

Figure 23b. Potential distribution for soapwort.

PARAMETERS	
Presence points:	335-336
Background points:	4122
Training AUC:	0.823
Test AUC±SD:	0.781±0.024
LAYERS	**PCT**
Water content (15 bars)	35.7
pH (1:1 water)	13.7
Distance to city	13.6
Land cover	12.6
Clay (%)	9.9
Distance to road	6.9
Organic matter (%)	4.4
Drainage class	3.2

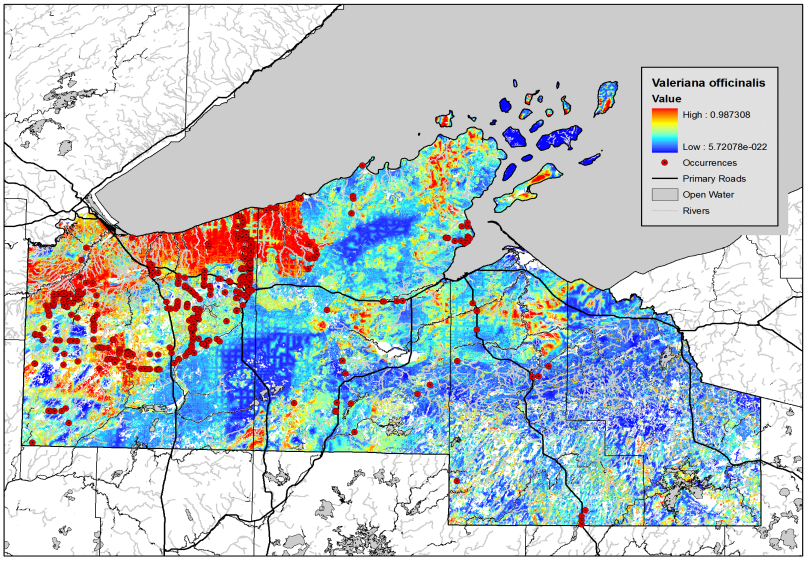

Figure 24a. Current distribution for common valerian.

PARAMETERS	
Presence points:	335-336
Background points:	4122
Training AUC:	0.796
Test AUC±SD:	0.767±0.026
LAYERS	**PCT**
Water content (15 bars)	45.0
pH (1:1 water)	17.5
Land cover	16.3
Clay (%)	12.4
Organic matter (%)	4.8
Drainage class	4.0

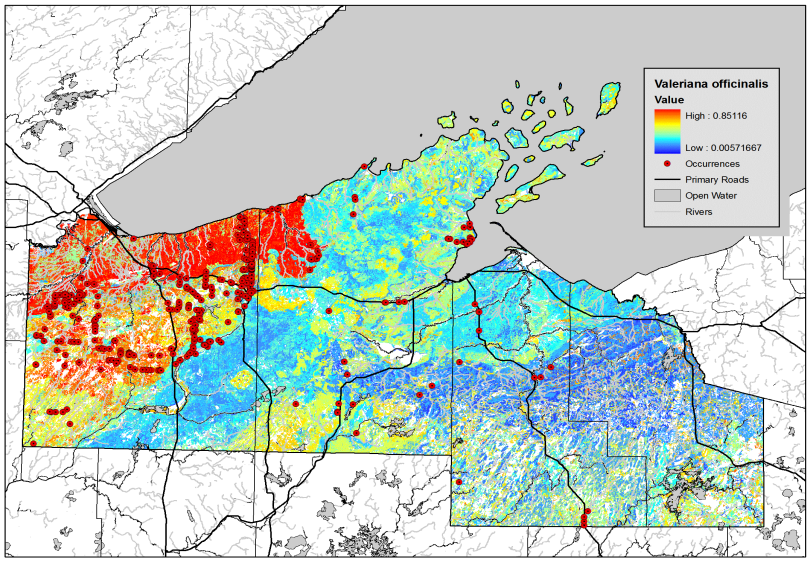

Figure 24b. Potential distribution for common valerian.

Shrubs and trees

Eurasian bush honeysuckles: The potential distribution model for the rather generalist Eurasian bush honeysuckles showed moderate to fairly high probability of occurrence across most of the project area, except near Lake Superior where probability was somewhat lower. This reflects the fact that known occurrences are much less frequent close to the lake than away from it (Figure 25b). This pattern of occurrence seems to coincide with both the lower (on average) percent clay and the greater number of consecutive frost-free days away from the lake. When frost-free days was included, however, the test AUC was nearly unaffected, going to 0.617±0.030 and 0.612±0.030 for the current and potential distribution models, respectively. In both cases frost-free days comprised less than 5% of the model.

Whether the apparent pattern of avoidance of Lake Superior is due to stochastic factors related to the introduction and expansion of honeysuckles in the region, incomplete sampling, or environmental conditions is not clear, but deserves further investigation.

Distance to cities and towns dominated the current distribution model (Figure 25a), with honeysuckles much more likely to occur close to cities and towns than farther away. This would seem consistent with the fact that these honeysuckles have historically been commonly planted in urban areas, eventually escaping (often with the help of fruit-eating birds) to the surrounding landscape.

Glossy buckthorn: The predicted current and potential distribution models for glossy buckthorn turned out to be very similar, with distance to roads and cities having little effect on the model (Figures 26a and b). For both models, soil surface texture was the most influential factor on the distribution of glossy buckthorn, followed by distance to open water, soil drainage class, and soil pH.

Even though inclusion of the water content layer gave a slightly higher test AUC value than soil drainage class, soil drainage class was used because more occurrences were retained in the model, and the response curve showed drainage class increasing with increasing soil moisture.

Adding the distance to open water and distance to river layers improved the models substantially. This would seem to be consistent with our perception that this species frequently occurs along woods edges bordering lakeshores and in wetlands (which often border lakes and rivers).

Distance to cities was more influential than distance to road for glossy buckthorn. The response curve for distance to city climbed steeply from zero to a peak at around 1,380 m, dropped off to about 5,000 m, and climbed slowly and continuously after that. This peak may be indicative of the fact that we (and other contributors) often did not survey in or on the outskirts of cities and towns. That land cover was not a good predictor may in part reflect the fact that glossy buckthorn readily invades open, disturbed sites as well as nearly pristine natural areas, ranging mesic to wet, and from full sun to fairly deep shade.

PARAMETERS	
Presence points:	386-387
Background points:	4123
Training AUC:	0.722
Test AUC±SD:	0.622±0.026
LAYERS	**PCT**
Distance to city	44.3
Clay (%)	15.0
pH (1:1 water)	13.8
Land cover	11.9
Water content (15 bar)	8.0
Distance to road	5.0
Drainage class	2.0

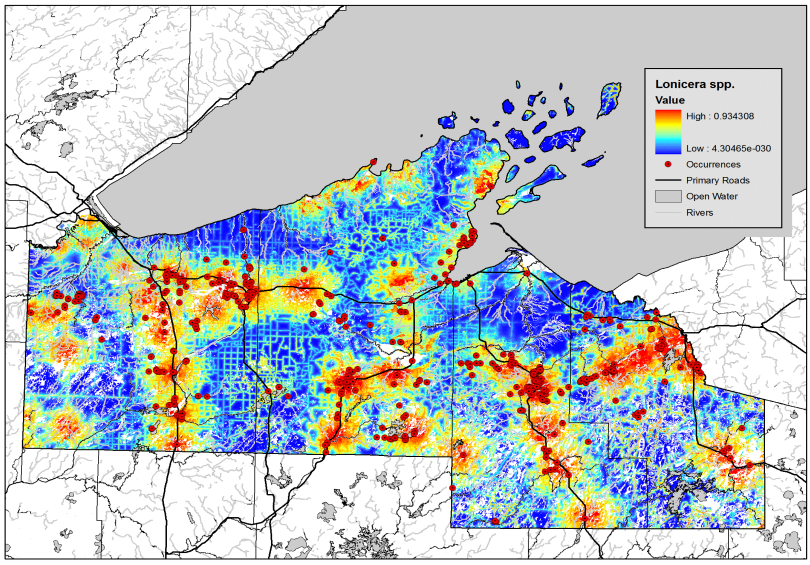

Figure 25a. Current distribution for Eurasian bush honeysuckles.

PARAMETERS	
Presence points:	386-387
Background points:	4123
Training AUC:	0.663
Test AUC±SD:	0.617±0.029
LAYERS	**PCT**
pH (1:1 water)	33.6
Land cover	30.1
Clay (%)	20.7
Water content (15 bar)	10.9
Drainage class	4.7

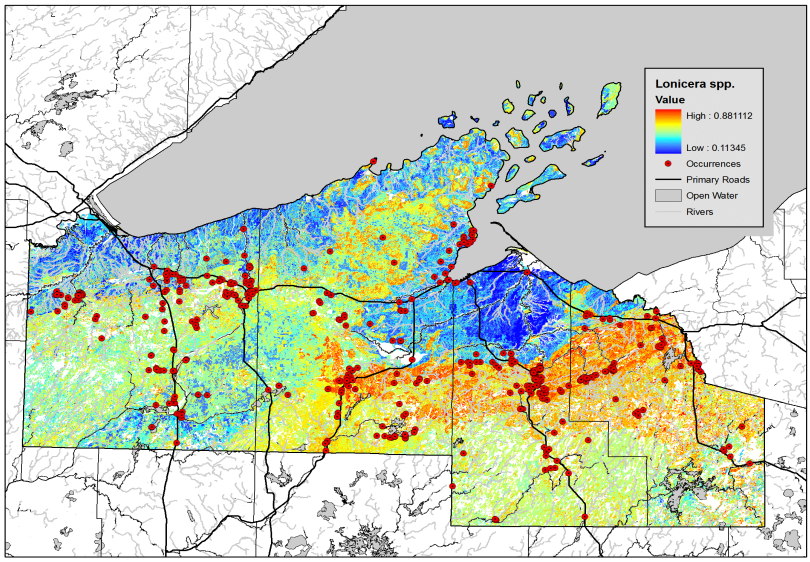

Figure 25b. Potential distribution for Eurasian bush honeysuckles.

PARAMETERS	
Presence points:	85-86
Background points:	4152
Training AUC:	0.813
Test AUC±SD:	0.732±0.061
	PCT
LAYERS	
Surface texture	35.3
Distance to open water	23.0
Drainage class	13.7
pH (1:1 water)	10.9
Distance to river	8.1
Distance to city	7.5
Distance to road	1.4

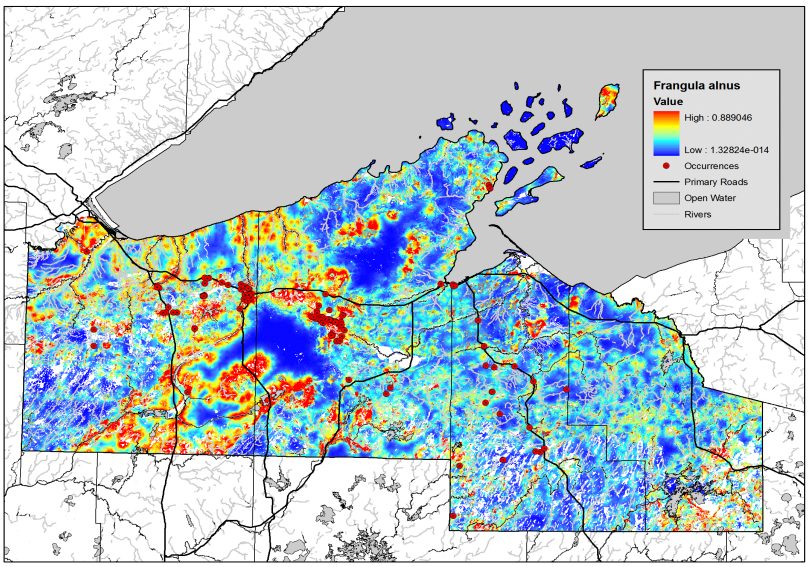

Figure 26a. Current distribution for glossy buckthorn.

PARAMETERS	
Presence points:	85-86
Background points:	4152
Training AUC:	0.798
Test AUC±SD:	0.726±0.062
LAYERS	**PCT**
Surface texture	38.0
Distance to open water	25.3
Drainage class	14.4
pH (1:1 water)	12.3
Distance to river	10.0

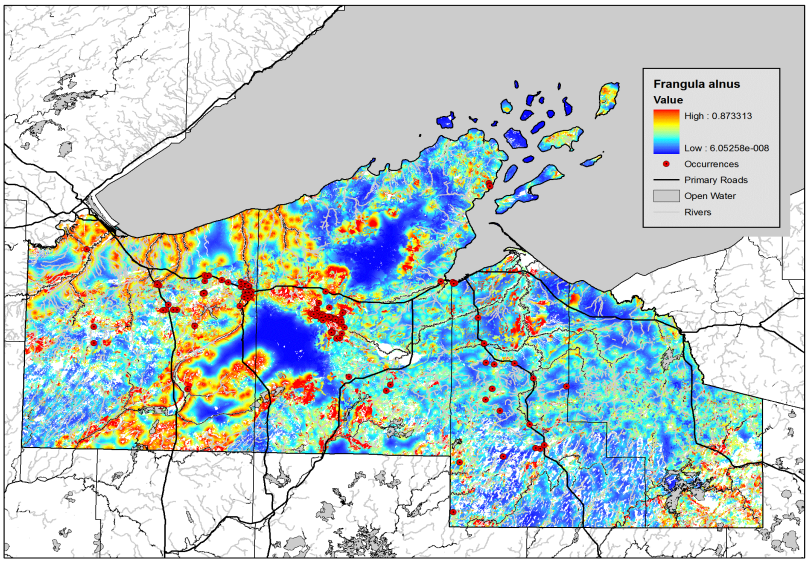

Figure 26b. Potential distribution for glossy buckthorn.

Common buckthorn: Common buckthorn proved difficult to model (Figures 27a and b). The original, unreplicated "best" potential distribution model included drainage class, frost-free days, organic matter, pHw, surface texture, and vegetation cover, resulting in acceptable diagnostic curves and AUC values for training and test data of 0.718 and 0.729, respectively. However, the replicated run resulted in an unsatisfactory model with training and test AUC values of 0.742 and 0.634±0.051, respectively.

The final (replicated) potential distribution model (Figure 27b) is dominated by surface texture. The full-model response curve for this variable (categorical, but with categories essentially ordered) shows high-effect categories interspersed with lower-effect categories, but with an apparent rough trend of mid-range categories (loams) higher than low-range (highly organic soils) and high range (sandy soils) categories.

The best "current distribution" model for common buckthorn (Figure 27a) included both roads and cities. This model was the only one to include higher-probability "rings" around cities and towns. I suspect that these are a result of residual bias in the data, resulting from two factors: 1) I usually did not sample in and on the outskirts of cities and towns, and data from other agencies is generally also from rural areas, and 2) common buckthorn is often very common in and on the outskirts of at least the larger population centers in the project area (pers. obs.). While my surveys did not include the often high-abundance areas for this species within and just outside these population centers, they apparently did include areas of moderate abundance just outside these high-abundance areas, resulting in the higher-probability rings evident in Figure 27a.

Eurasian tree willows: Distance to nearest river was the most influential in predicting both the current and potential distribution of Eurasian tree willows (Figure 28a and b). This result is consistent with our observations as well as with habitats they invade in other parts of the world (Greenwood et al. 2004).

Distance to open water was rather ineffective in predicting the distribution of these species. This was a bit surprising as these trees are often common along developed lakeshores, where they presumably escape from plantings (pers. obs.) The low importance of lakeshores in the models is almost certainly due to undersampling in these areas.

As would be expected for a taxa with an affinity towards floodplains, wetlands and lakeshores, the current and potential distribution models both predict low probability of occurrence for Eurasian tree willows in the sandy barrens areas extending into the Bayfield Peninsula. Most of the project area would appear to provide fair to excellent habitat for these willows, though.

PARAMETERS	
Presence points:	126
Background points:	4152
Training AUC:	0.815
Test AUC±SD:	0.713±0.045
LAYERS	**PCT**
Distance to city	49.5
Surface texture	27.2
Distance to road	10.8
Drainage class	5.2
Organic matter	4.2
pH (1:1 water)	3.1

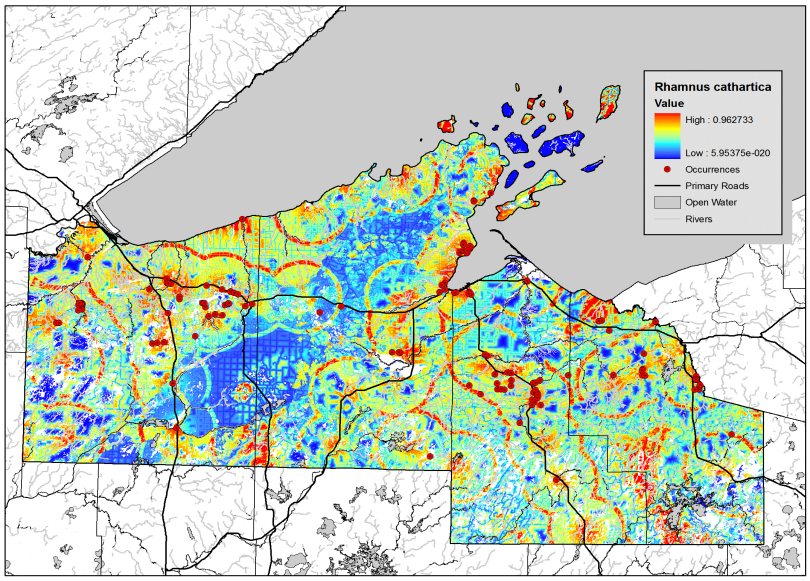

Figure 27a. Current distribution for common buckthorn.

PARAMETERS	
Presence points:	126
Background points:	4152
Training AUC:	0.711
Test AUC±SD:	0.634±0.047
LAYERS	**PCT**
Surface texture	68.1
Drainage class	25.3
Organic matter	7.4
pH (1:1 water)	6.5

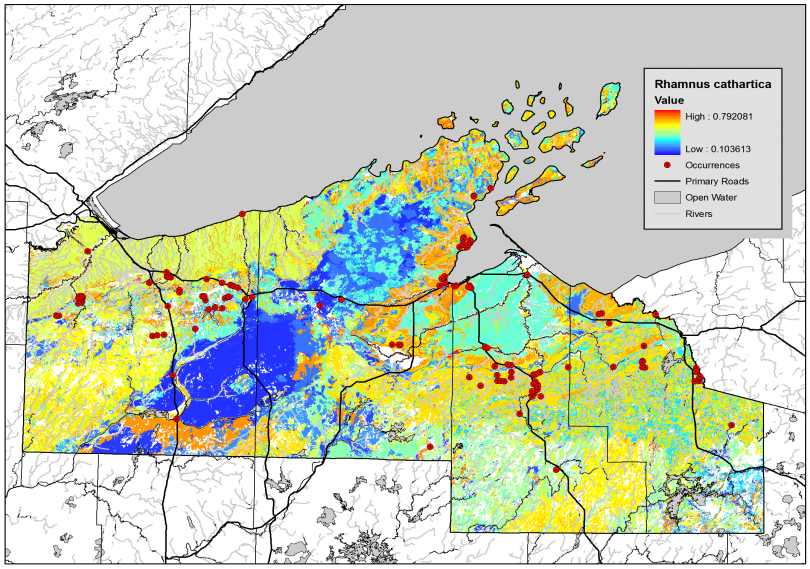

Figure 27b. Potential distribution for common buckthorn.

PARAMETERS	
Presence points:	214-215
Background points:	4151
Training AUC:	0.825
Test AUC±SD:	0.731±0.032
LAYERS	**PCT**
Distance to river	32.7
Surface texture	17.5
Drainage class	13.0
Organic matter (%)	10.5
Distance to open water	7.4
Distance to city	6.8
Distance to road	6.3
pH (1:1 water)	5.7

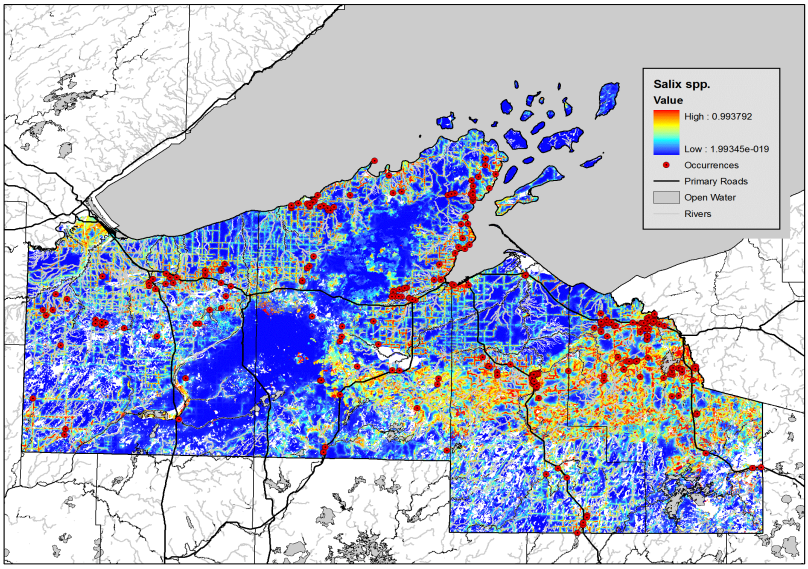

Figure 28a. Current distribution for Eurasian tree willows.

PARAMETERS	
Presence points:	214-215
Background points:	4151
Training AUC:	0.791
Test AUC±SD:	0.715±0.035
LAYERS	**PCT**
Distance to river	39.4
Surface texture	20.9
Drainage class	16.1
Organic matter (%)	11.7
Distance to open water	6.5
pH (1:1 water)	5.3

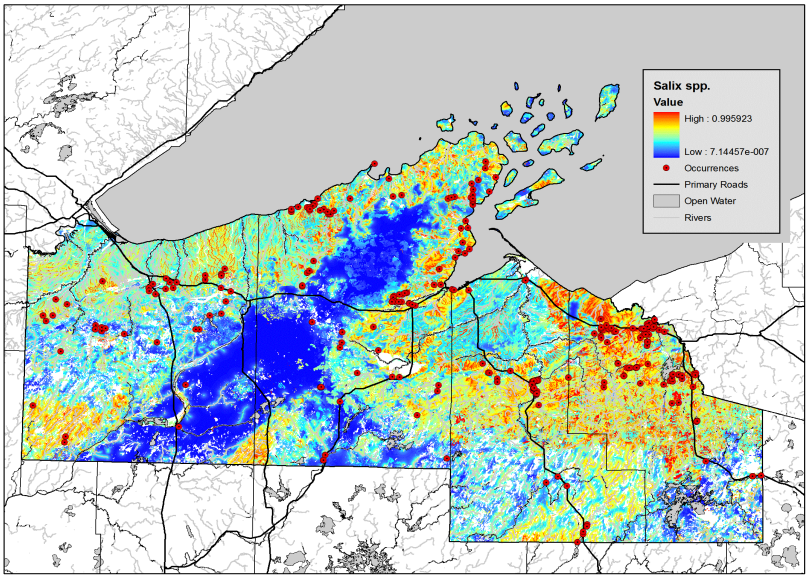

Figure 28b. Potential distribution for Eurasian tree willows.

DISCUSSION

Environmental layers

Unfortunately the SSURGO soil layers typically lack data for a small percentage of their area. While we were able to fill in the gaps in the drainage class data in a reasonable way, data missing from the other SSURGO layers is simply not available. The main disadvantage of using these layers appears to be that when a species occurrence lacked corresponding data for one or more layers, Maxent simply drops that occurrence from the entire analysis. The same was true for any target-group background points that fell in locations with missing data.

Because we have relatively large data sets for most of the species we modelled, the loss of a small minority of occurrences presumably has little effect on the models for these species. For species with few occurrences, though, loss of even one or two points can have a serious detrimental effect on the accuracy and reliability of the model. Maxent does have an option to include points in locations with missing data in one or more layers, but using certain combinations of layers with missing data causes Maxent to crash, so models with those combinations cannot be run. Unfortunately the best option at this writing is to simply uncheck the “allow missing data” option and deal with loss of a (hopefully) small percentage of data points (S. Phillips, pers. comm.).

For future models it may be possible to fill in missing data in the SSURGO layers using other sources. The coarser State Soil Geographic Database (STATSGO) data (USDA-SSDS 2007) might even include this information. This possibility requires further investigation.

Judging from their performance in these models, the shade and drainage class layers (used as continuous layers) generally were not as useful or effective as the original categorical layers from which they were derived. This is probably due to the fact that relevant information is distorted or lost by reclassification. For the soil surface texture layer, the original categories were reclassified to integers in an ordered one-to-one correspondence (Table 4). When this layer was used as a categorical instead of a continuous layer, its effect on most of the models increased substantially. In contrast, the original WISCLAND land cover layer was “simplified”, by “pigeonholing” the 32 original categories into four shade categories (Table 1). The resulting “shade” layer generally performed poorly compared to the original layer.

The target-group background layer seemed to improve the models substantially, by reducing or eliminating the locational “road effect” and thereby reducing overfitting of the models. However, populations along roads still appear to introduce “noise” into results for another reason - the SSURGO soil layers do not include road effects on the land. Roadside populations generally grow in highly compacted, gravelly soil, with moisture levels potentially ranging from dry along the roadbanks to wet in the ditches. Whatever the wetness, soil texture, pHw, or other characteristics of the surrounding soil, they often do not correspond well to conditions experienced by roadside populations.

A similar situation exists with the WISCLAND land cover layer, though perhaps to a lesser degree. The WISCLAND layer is detailed enough to show most major highways with wide

corridors as “open”, even though they may be surrounded by forest. It is not detailed enough to show secondary corridors such as county roads, though. Thus, an occurrence along a county road may be tagged as “shaded”, when in fact it is growing in nearly full sun. This shortcoming could be dealt with by making a “road habitat” layer. This layer would include shoulders for class 1 and 2 road coverages. Alternatively (and perhaps concurrently) the areas (cells) in other layers corresponding to road corridors could be replaced with the appropriate data type. For example, roadbeds could be reclassified to “very well-drained” in the drainage class layer, and road corridors to “open” in the land cover layer.

Additional environmental layers should be sought to complement the ones we have used so far. Miller at al. (2007) point out that bioclimatic layers such as potential solar radiation, mean relative humidity and potential evapotranspiration may be more directly related to plant distribution than more commonly used layers such as average precipitation and temperature. Growing degree days also have a significant influence on plant growth and distribution (Barbour et al. 1987), so a growing-degree layer might be useful as well.

Phillips et al. (2006) suggest using recent land-cover data to exclude highly altered habits from predicted distribution models, when those alterations have made the environment unsuitable for the species being modelled.

Species sampling and data

Species data

An almost universal problem that arises when trying to model the distribution of an invasive species is that the species in question often has not yet spread throughout its potential habitat. SDMs assume that the species being modelled is at equilibrium with its environment (Elith and Leathwick 2009). For introduced species this is often not the case. When species have not yet occupied all of their fundamental niche, habitat models generally underestimate the potential habitat (Phillips et al. 2006). Poor dispersers may on average have colonized less of their potential habitat and therefore be harder to model than good dispersers (Peterson et al. 2007). Models can still be useful in identifying suitable habitats that have not yet been invaded, though.

Because invasive species are generally still colonizing suitable habitats in their new environment, they are frequently more constrained by propagule availability than by habitat requirements (Rouget and Richardson 2003). Several studies (e.g., Dormann et al. 2007, Rouget and Richardson 2003) have developed techniques to include geographic and biological influences on species dispersal into models, and our models would likely benefit from these types of techniques as well.

Species data are often autocorrelated. This is not a problem for the models directly, as each occurrence point is simply matched to the value of each environmental layer at that point. But because environmental layers (reflecting the real world) are generally autocorrelated, and because invasive species have typically not reached equilibrium with their environment over large areas, clusters of occurrence points have a strong tendency to artificially emphasize some

habitats over others in a way that may have little to do with the species' habitat requirements (Elith and Leathwick 2009).

During the early stages of this project, layers for distance to nearest population of the study species were constructed for several of the more common invasives in the region. These layers can be used to predict present distribution based on known occurrences, and represent a simple way to include dispersal in the models. Because species occurrences generally show a strong tendency towards clumped distributions, these layers usually had large effects on their respective models. Because our goal was to define the habitats that these species are likely to occur in, and because I wanted to (eventually) use a more integrated method for including spatial autocorrelation in our models, I did not use these “distance to taxa” layers in the final models.

Unfortunately there is simply not enough occurrence data to produce reliable models for a substantial number of species in the project area directly. These include species which are rare or newly-established in the region but which possess many of the classic biological and life-history characteristics typical of invasives. Species whose native range has a similar climate to the project area, and which have become invasive in other regions of the world with similar climates, are also more likely to become invasive here (e.g., Williamson 1996, Reichard and Hamilton 1997).

De Siqueira et al. (2009) describe a method whereby new occurrences of a rare South American plant were found using environmental distance calculations, based on the habitat of the single known occurrence of the species in the region. First, maps were produced based on environmental distance to the one known population. Floristic surveys were then conducted in 9 of the predicted high-probability areas and (as a control) 8 of the low-probability areas. These surveys uncovered 5 new populations in the high-probability areas and 1 in the low-probability areas. With the (now) 7 known sites, De Siqueira et al. (2009) were able to produce “standard” species distribution models using GARP.

Unsampled areas

Despite our intention of “finishing” sampling the project area in 2009, some portions remain inadequately sampled or unsampled altogether (Figure 29). These include parts of southern Douglas County, southwestern Bayfield County, southern Iron County, and scattered smaller portions of the project area. When occurrence records are drawn from too small a geographic area, these records are likely to encompass only a small fraction of the habitats the species can potentially occupy, resulting in a significant underestimation of potential distribution (Phillips et al. 2006). Additional sampling of underrepresented areas has the potential to improve results substantially, especially for more widespread species (Jones and Reichard 2009).

Cities and towns and the often highly-disturbed natural areas surrounding them frequently support large populations of common and glossy buckthorn, Eurasian bush honeysuckles, and other invasives (pers. obs.). Because our focus and that of others who have contributed species occurrence data has primarily been on populations of invasives occurring in natural areas

(especially on and near public lands), these data seldom include occurrences within and on the outskirts of cities and towns. That distance to nearest city was often not a significant factor in the models for these species may very well be due to this lack of data.

Perhaps the best solution to this problem would be to survey the population centers within the project area, and record at least one occurrence point within each center for each species found. Alternatively a “pseudo-occurrence” point could be assigned to each city. This second approach would almost certainly result in some occurrences being erroneously assigned, though, particularly for smaller towns and for species not commonly cultivated.

Model evaluation

The omissions curves, ROC curve and associated AUC statistic are useful in assessing the overall effectiveness of the models in predicting the distribution of the species being modelled. They are of little use in diagnosing the underlying causes of common model deficiencies, however. As with any statistical procedure, exploratory data analysis (EDA) should be an important part of data analysis and model evaluation.

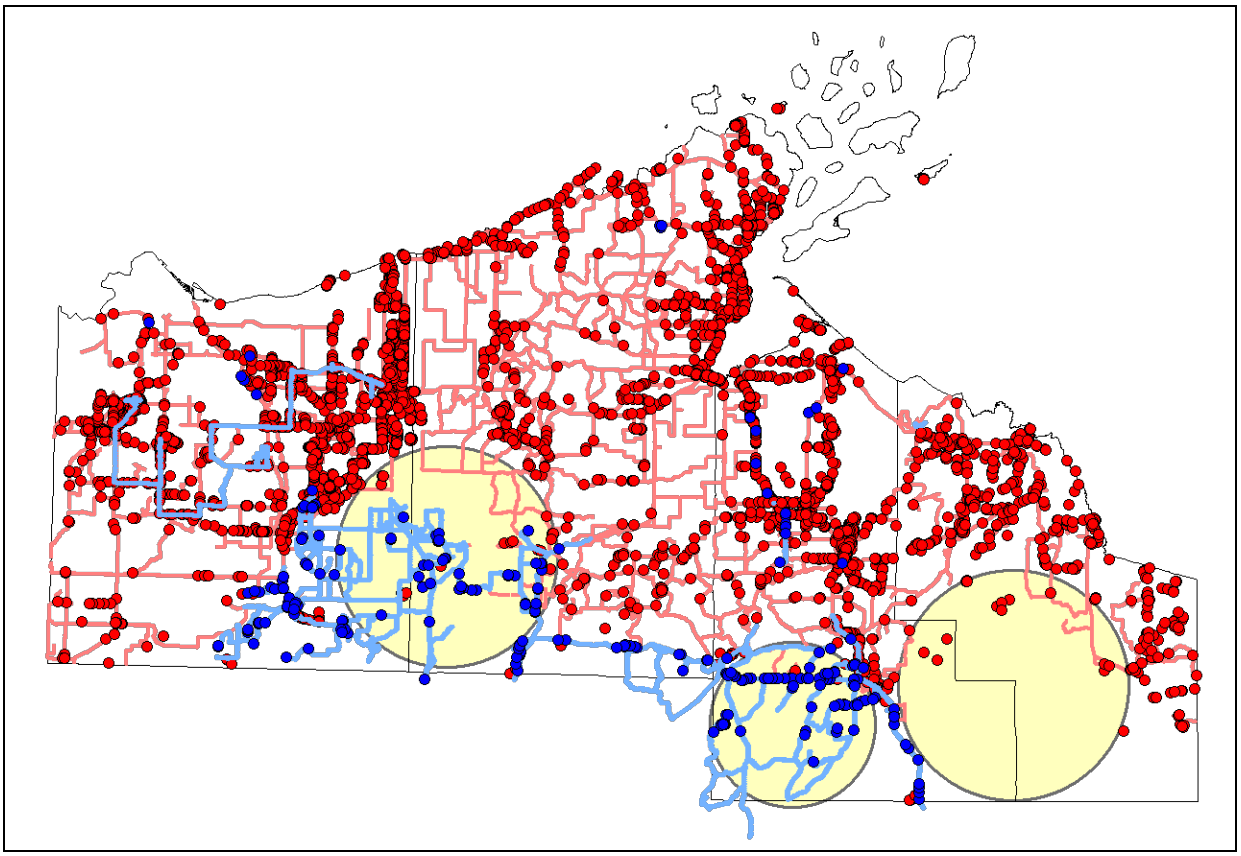

Figure 29. Invasive plant distribution records and survey effort within the project area. Red lines and circles show search routes completed and occurrences recorded from 2000 through 2008. Yellow circles highlight portions of the project area targeted for sampling in 2009. Blue lines and circles show routes taken and occurrences recorded in 2009.

The residuals of the predicted probabilities resulting from the model should be analyzed for patterning and autocorrelation (Elith and Leathwick 2009). A prediction generally results in patterned residuals when the environmental variables are inadequate (Leathwick and Whitehead 2001), when the model is mis-specified, or when geographic factors are influential (Miller at al. 2007, Dormann 2007). These residuals can be tested using Moran's I statistic or other diagnostic tests. These tests might be performed with Spatial Analysis in Macroecology (SAM), a free software package developed by Rangel et al. (2006, 2010). Strong patterns in the residuals generally indicate that key environmental predictors are missing from the model (Leathwick and Whitehead 2001) or that geographic factors are influencing the model (Dormann et al. 2007).

APPLICATION

These species distribution models provide a useful tool for planning and prioritizing invasive plant management at nominal scales of 1:40,000 and smaller. The raster format of the model output facilitates the use of standard GIS software to visualize, summarize and analyze these models in relation to associated spatial data. We believe that these models will be useful in gaining insights into the habitat requirements of these species, predicting which areas are most suitable for them, and for highlighting portions of the project area at greatest risk of invasion.

The results for leafy spurge provide a good example of how these models might be used. The potential distribution model (Figure 21b) indicated that soil surface texture was an important predictor of leafy spurge habitat. The Maxent response curves for surface texture (not shown) indicated that loamy fine sands provided optimum conditions for spurge, followed by moderately decomposed plant material, highly decomposed plant material, and loamy sands. Assuming that the association with soil high in decomposed plant material is an artifact of spurge populations occurring along roads that pass through wetlands, loamy fine sands provide the best surface soil type for leafy spurge. Higher probabilities of spurge occurrence were also associated with the land cover classification “herbaceous/field crops”, followed by “mixed deciduous/coniferous”, “barrens” and “broad-leaved deciduous”.

Despite these somewhat conflicting results, the potential distribution map for leafy spurge appears to coincide quite closely with “conventional wisdom” regarding the habitats most at risk from this plant (sunny, mesic to dry sites with coarse soils). Areas of dry, coarse soils appear likely to support spurge, while wetter, heavily wooded areas show low probability of occurrence. Road corridors appear to be a major factor in facilitating the spread of spurge (Figure 21a).

The largest high-probability area for spurge appears to be the Moquah Barrens of northern Bayfield County (the northern end of the Bayfield Barrens, as defined by Albert 1994). The Moquah Barrens are an important hunting and berry-gathering area for local residents (both tribal and non-tribal) and are increasingly being managed for their ecological value. Existing spurge sites in and on the fringes of the Moquah Barrens should be a high priority for treatment. Except for road corridors and agricultural areas, the more heavily wooded areas in the southern and eastern portions of the project area are predicted to be at much less risk from invasion by leafy spurge.

When survey routes are overlayed on the potential distribution map for leafy spurge, areas that remain poorly surveyed but that have high potential for supporting spurge become evident (Figure 30). These areas include large blocks of land in the Bayfield Barrens of southern Bayfield County and southeast Douglas County, as well as areas in northwest Douglas County southern Iron County. While leafy spurge may still be absent from these areas, significant suitable habitat exists. These areas would be good targets for future surveys. Because leafy spurge appears to be spreading mainly along roads at the present time (pers. obs., also see Figure 21a), any unsurveyed roads or ATV trails running through these high-probability areas should be high priority for future surveys.

Annual surveys might be warranted in areas of suitable habitat that are adjacent to known spurge populations. For example, the high probability area in southwest Bayfield County has a known population of spurge on its southeastern border, and may be at imminent risk from invasion. The same is true of the high-probability area in northwest Douglas County. Annual surveys could be useful in detecting nascent spurge invasions in these areas, before they spread out of control. Conversely, the high-probability area in southeast Douglas County may be "buffered" to some degree by the apparently unsuitable habitat between this area and the nearest known spurge populations to the north.

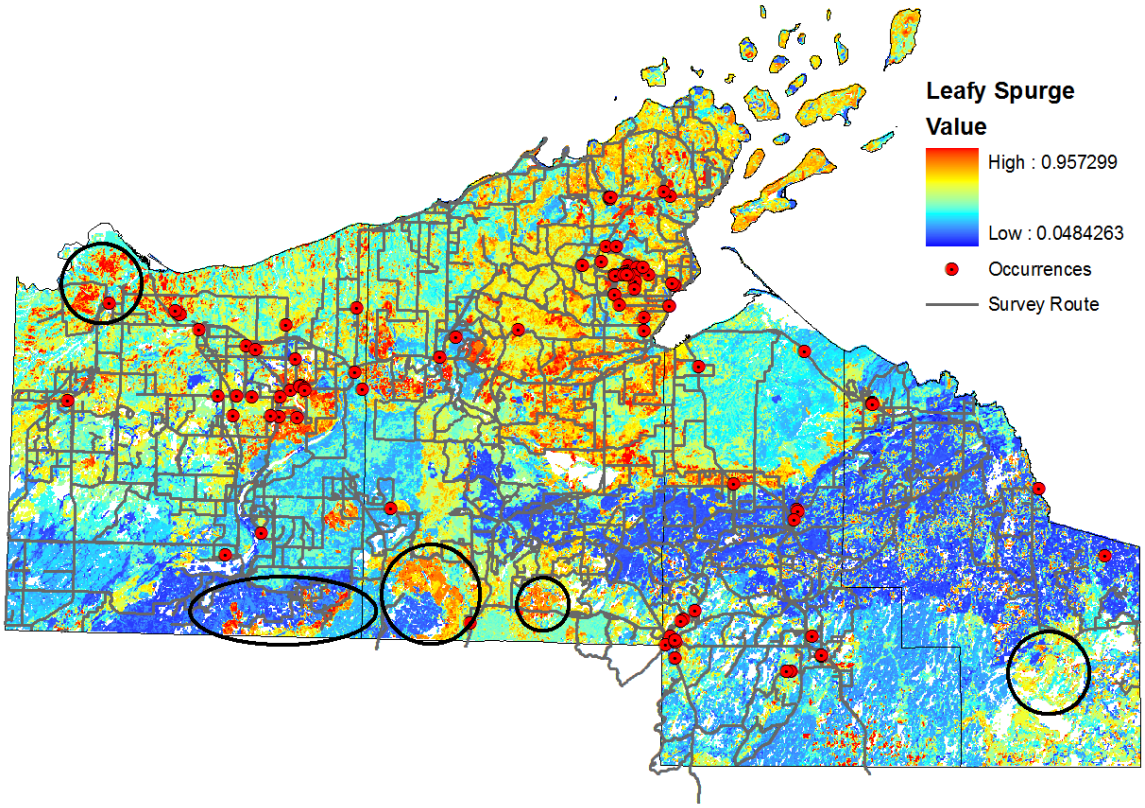

Figure 30. Potential distribution of leafy spurge from Figure 21b, with survey routes included. Relatively undersurveyed areas of high predicted probability are circled.

In addition to prioritizing areas for surveys and treatment, these models can be useful in comparing the susceptibility of various land management areas to invasion and domination by species of concern. Overlaying land ownerships on the potential distribution map for leafy spurge (Figure 31a) highlights susceptibility of these ownerships to spurge invasion.

The distribution of susceptibility leafy spurge across the different land ownerships (measured in terms of number of 30 m^2 cells in each category) can also be illustrated using a histogram (Figure 31b). Private lands comprise the largest percentage of the project area, and have the largest total land area falling into the moderately to highly susceptible range in term of invasion by spurge. County and USFS lands also include large areas that appear to be moderately to highly susceptible to invasion. By comparison a relatively small amount of susceptible lands fall under tribal, state, and National Park Service ownership.

By normalizing the land area of each landowner, the relative probability of occurrence of leafy spurge for each ownership category can be highlighted (Figure 31c). National Park Service lands have the greatest percentage of area in the moderate to high-risk category (here interpreted as cells having over 50% probability of occurrence), while most of the land within the other ownerships falls below the 50% level of probability. Tribal lands appear to be the least susceptible of all the ownerships to invasion by spurge. For various reasons only a small proportion of tribal and National Park lands were surveyed for spurge, though, which may have affected the results for these ownerships in ways that are difficult to predict.

A similar scenario is evident for purple loosestrife (Figures 32a, b and c). Most of the highly susceptible lands appear to be in private ownership, with substantial susceptible areas in county and National Forest ownership (Figure 32b). The normalized histogram shows tribal lands at greatest risk to invasion by loosestrife, with the other ownerships at relatively lower risk (Figure 32c). In this case tribal lands (particularly the Bad River Reservation) have been relatively well-surveyed for loosestrife, so the results for these lands may be more reliable.

Finally, models may be useful in comparing the relative risk these landowners face from various invasive plants. With a higher percentage of 30-meter grid cells falling into the moderate to high probability range for purple loosestrife (Figure 32c) than for leafy spurge (Figure 31c), tribal lands would appear to be at greater risk of colonization by loosestrife than by spurge. The risk of colonization by loosestrife versus spurge appears to be roughly similar for the other ownerships.

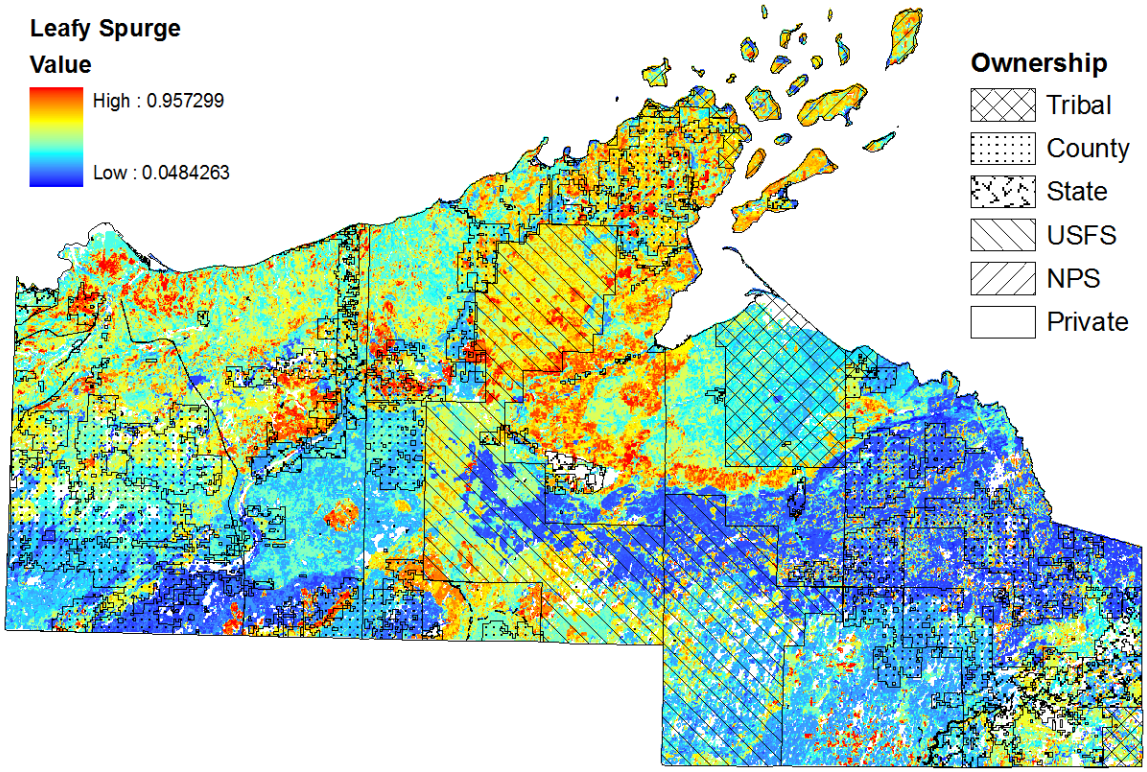

Figure 31a. Potential distribution of leafy spurge, with land ownerships superimposed.

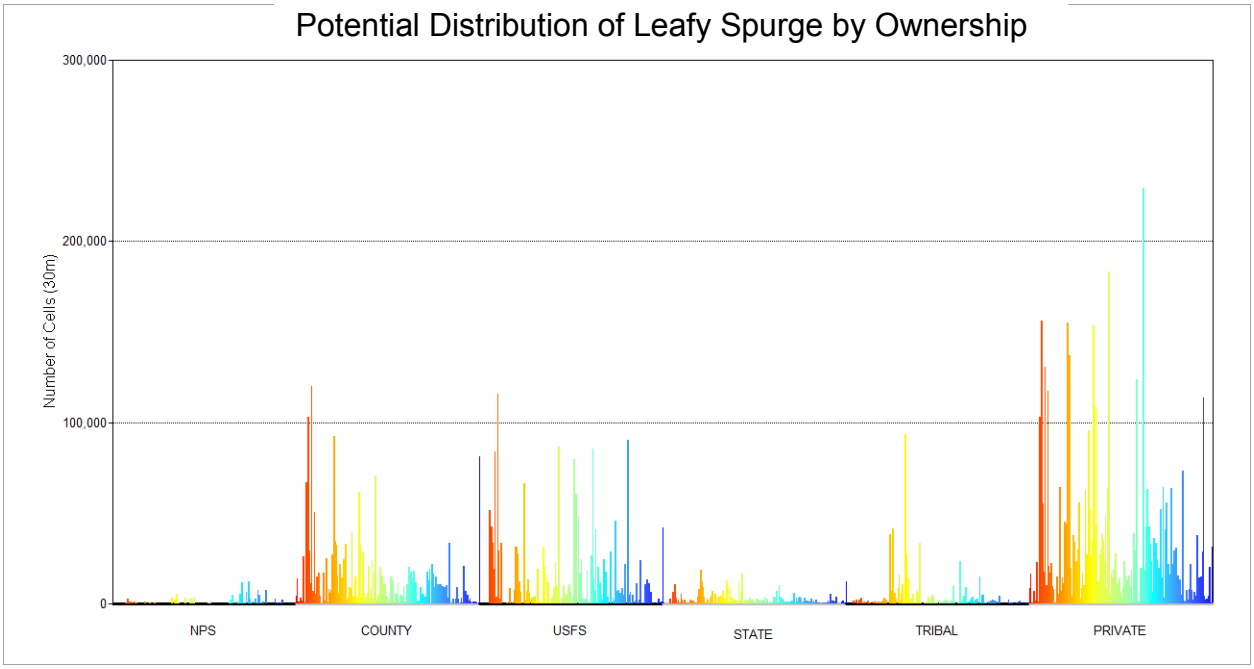

Figure 31b. Probability of occurrence of leafy spurge by land ownership, as calculated with ArcView Spatial Analyst. For each ownership category, bars indicate the number of 30-meter grid cells falling into each probability category. Histogram color scheme corresponds to Figure 31a.

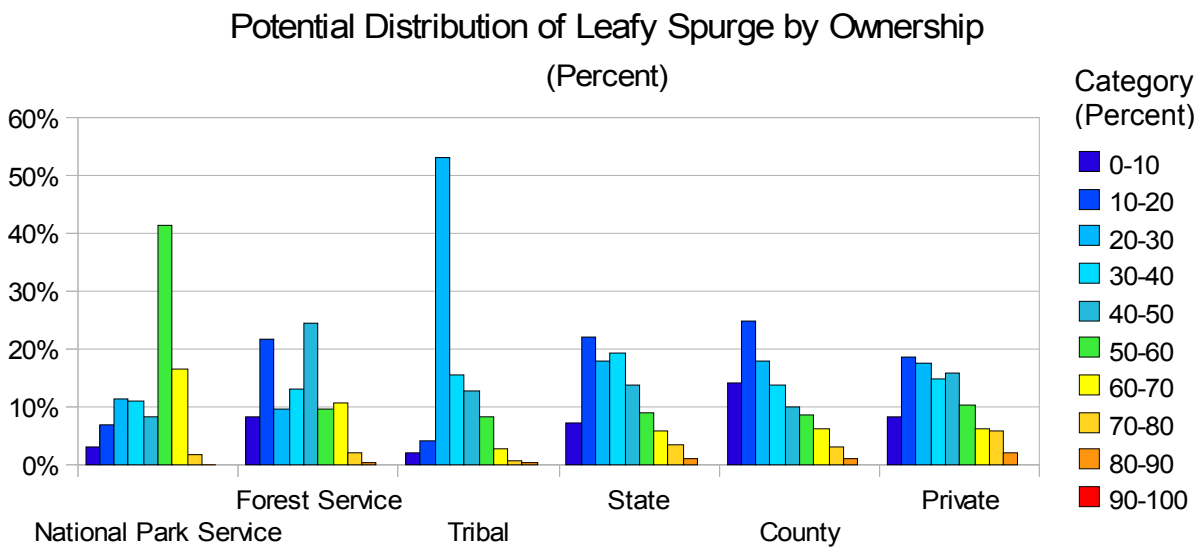

Figure 31c. Probability of occurrence of leafy spurge by land ownership category. For each ownership category, bars indicate the number of 30-meter grid cells falling into each 10% probability category, starting with interval 0-10%.

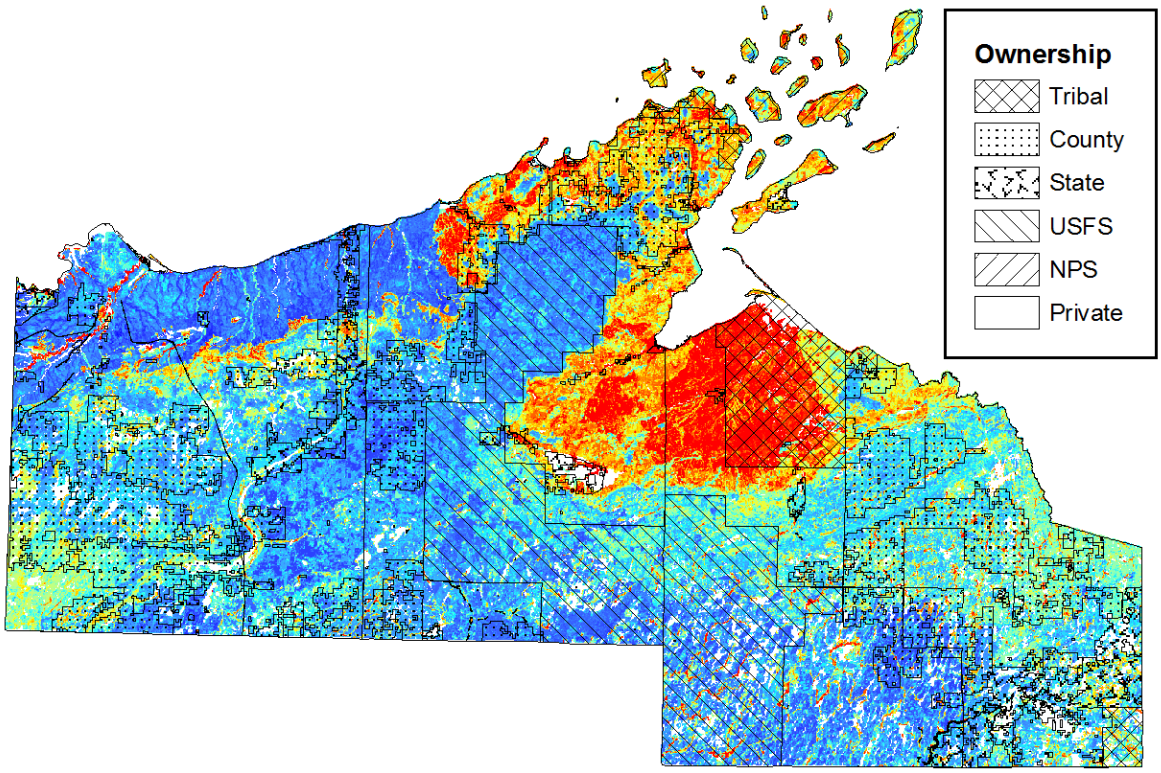

Figure 32a. Potential distribution of purple loosestrife, with land ownerships superimposed.

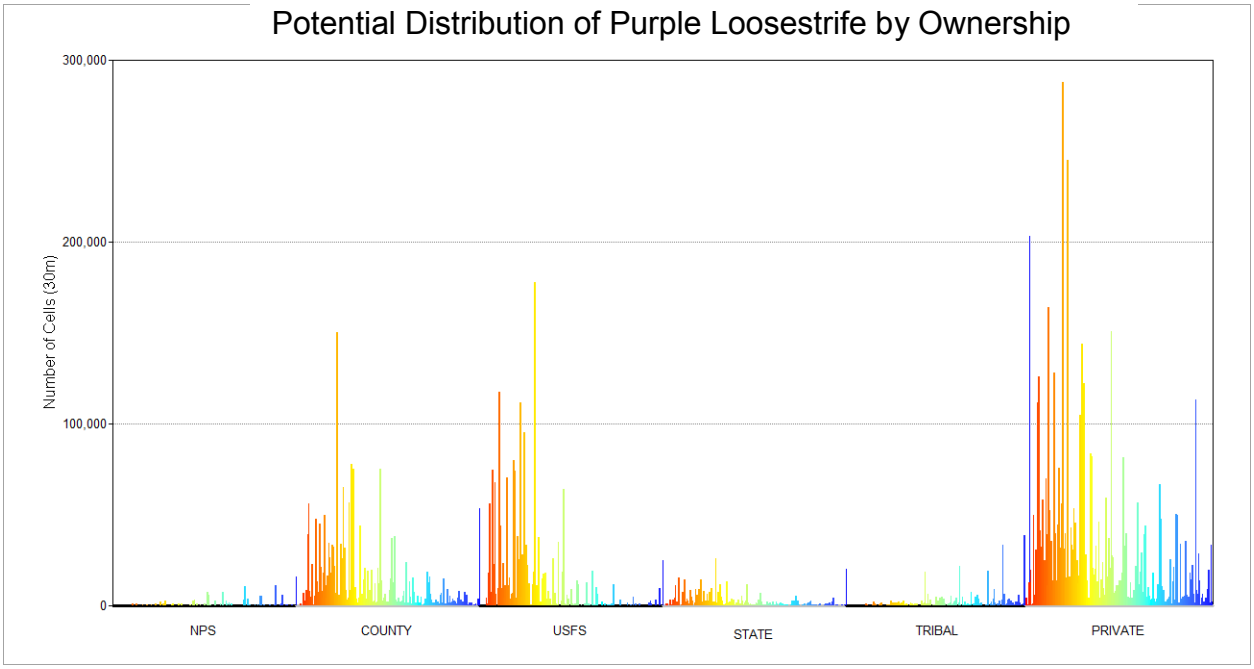

Figure 32b. Probability of occurrence of purple loosestrife by land ownership category, as calculated with ArcView Spatial Analyst. For each ownership category, bars indicate the number of 30-meter grid cells falling into each probability category. Histogram color scheme corresponds to Figure 32a.

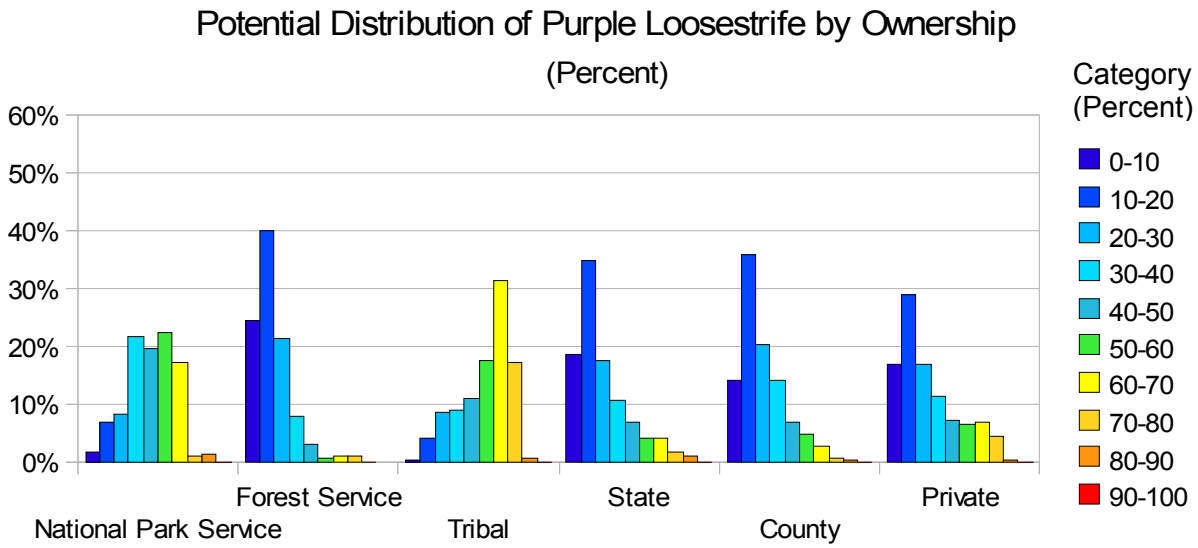

Figure 32c. Probability of occurrence of purple loosestrife by land ownership category. For each ownership category, bars indicate the number of 30-meter grid cells falling into each 10% probability category, starting with interval 0-10%. (For this model, no cells had a predicted probability of occurrence higher than 90%.)

CONCLUSIONS

Species distribution modelling is a rapidly evolving field, with an ever-expanding array of modelling and evaluation methods being developed to address an ever-widening array of research problems (Elith and Graham 2009). We plan on keeping up-to-date with advances in species distribution modelling and model evaluation. We also plan on further researching the habitat requirements and tolerances of the species we are attempting to model, in order to better evaluate the appropriateness of the environmental features used in our models.

We expect that the "final" invasive species models presented in this report will continue to be refined and improved. Additional environmental layers will be developed to refine the modelling effort. Ecological requirements of the species themselves might also be integrated into the models. Relatively simple functions to take into account key factors such as autocorrelation and dispersal have the potential to substantially improve these models and should be included as well.

We also intend to become proficient with additional methods to evaluate our models, including use of Moran's I statistic to evaluate prediction residuals. Systematic field verification of predictions should be carried out, to evaluate the real-world accuracy and effectiveness of our "final" models. Finally, areas that have been substantially undersampled should be surveyed, to assess the accuracy of existing models and improve future models.

As we continue to gain proficiency at various aspects of species distribution modelling, we hope that some of the less common species in the project area can be reliably modelled as well. The results for some of the more abundant invasive plant species in the project area provide promise that these modelling efforts will be useful in guiding and prioritizing invasive species management and control in the ceded territories.

ACKNOWLEDGEMENTS

Thanks to Dara Olson, whose help with ArcView and ArcGIS was invaluable and whose patience with my many questions was seemingly limitless. Thanks also to Miles Falck for help with GIS software and for providing helpful suggestions on the manuscript. This project was funded by the Wisconsin Tribal Conservation Advisory Council and the US Department of Interior – Bureau of Indian Affairs.

REFERENCES

Albert, D. A. 1994. Regional landscape ecosystems of Michigan, Wisconsin and Minnesota: A working map and classification. General technical report NC-178. North Central Forest Experiment Station, US Forest Service, St. Paul, MN, USA.

Anderson, R. P., D. Lew, and A. T. Peterson. 2003. Evaluating predictive models of species' distributions: Criteria for selecting optimal models. Ecological Modelling 162 (3): 211-232.

Archibold, O. W., D. Brooks, and L. Delanoy. 1997. An investigation of the invasive shrub European buckthorn, *Rhamnus cathartica* L., near Saskatoon, Saskatchewan. Canadian Field Naturalist 111 (4): 617-621.

Barbour, M. G., J. H. Burke, and W. D. Pitts. 1987. Terrestrial Plant Ecology, 2nd edition. The Benjamin/Cummings Publishing Company, Inc., Menlo Park, CA, USA.

Barcaccia, G., S. Meneghetti, E. Albertini, L. Triest, and M. Lucchin. 2003. Linkage mapping in tetraploid willows: Segregation of molecular markers and estimation of linkage phases support an allotetraploid structure for *Salix alba* x *Salix fragilis* interspecific hybrids. Heredity 90 (2): 169-180.

Berger, A.L., S. A. Della Pietra, and V. J. Della Pietra. 1996. A maximum entropy approach to natural language processing. Computational Linguistics 22 (1): 39–71.

Busby, J. R. 1991. BIOCLIM - a bioclimate analysis and prediction system. Pages 64-68 *in*: Margules, C. R. and M. P. Austin, eds. Nature conservation: Cost effective biological surveys and data analysis. CSIRO, Canberra, Australia.

Carpenter, G., A. N. Gillison, and J. Winter. 1993. DOMAIN: A flexible modeling procedure for mapping potential distributions of animals and plants. Biodiversity and Conservation 2: 667-680.

Catling, P. M. and Z. S. Porebski. 1994. The history of invasion and current status of glossy buckthorn, *Rhamnus frangula*, in Southern Ontario. Canadian Field Naturalist 108 (3): 305-310.

Christen, D. and G. Matlack. 2006. The role of roadsides in plant invasions: A demographic approach. Conservation Biology 20 (2): 385-391.

Conrad, K. F., M. S. Warren, R. Fox, M. S. Parsons, and I. P. Wiowod. 2006. Rapid declines of common, widespread British moths provide evidence of a British biodiversity crisis. Biological Conservation 132: 279-291.

Davis, S. L. and L. Turner-Jones. 2008. Potential for mixed mating in the protandrous perennial *Saponaria officinalis* (Caryophyllaceae). Plant Species Biology 23 (3): 183-191.

De Cock, K., B. Lybeer, K. Vander Mijnsbrugge, A. Zwaenepoel, P. Van Peteghem, P. Quataert, P. Breyne, P. Goetghebeur, and J. Van Slycken. 2003. Diversity of the willow complex *Salix alba* - *S.* x *rubens* - *S. fragilis*. Silvae Genetica 52 (3-4): 148-153.

De Siqueira, M. F., G. Durigan, P. de Marco Jr., and A.T. Peterson. 2009. Something from nothing: Using landscape similarity and ecological niche modeling to find rare plant species. Journal for Nature Conservation 17 (1): 25-32.

Dixon, J. M. 1991. Biological flora of the British Isles: *Avenula* (Dumort.) Dumort. Journal of Ecology 79: 829–865.

Dormann, C. F. 2007. Promising the future? Global change projections of species distributions. 8 (5): 387-397.

Dormann, C. F., J. M. McPherson, M. B. Araújo, R. Bivand, J. Bolliger, G. Carl, R. G. Davies, A. Hirzel, W. Jetz, W. D. Kissling, I. Kühn, R. Ohlemüller, P. R. Peres-Neto, B. Reineking, B. Schröder, F. M. Schurr, and R. Wilson. 2007. Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. Ecography 30 (5): 609-628.

Dudík, M., S. J. Phillips, and R. E. Schapire. 2005. Correcting sample selection bias in maximum entropy density estimation. Pages 323–330 *in*: Advances in neural information processing systems 18. MIT Press, Cambridge, Massachusetts, USA.

Dullinger, S., I. Kleinbauer, J. Peterseil, M. Smolik, and F. Essl. 2009. Niche based distribution modelling of an invasive alien plant: effects of population status, propagule pressure and invasion history. Biological Invasions 11 (10): 2401-2414.

Elith, J., C. H. Graham, R. P. Anderson, M. Dudík, S. Ferrier, A. Guisan, R. J. Hijmans, F. Huettmann, J. R. Leathwick, A. Lehmann, J. Li, L. G. Lohmann, B. A. Loiselle, G. Manion, C. Moritz, M. Nakamura, Y. Nakazawa, J. McC. M. Overton, A. T. Peterson, S. J. Phillips, K. Richardson, R. Scachetti-Pereira, R. E. Schapire, J. Soberón, S. Williams, M. S. Wisz, and N. E. Zimmermann. 2006. Novel methods improve prediction of species' distributions from occurrence data. Ecography 29 (2): 129-151.

Elith, J. and C. Graham. 2009. Do they? How do they? WHY do they differ? On finding reasons for differing performances of species distribution models. Ecography 32: 66–77.

Elith, J. and J. R. Leathwick. 2009. Species distribution models: Ecological explanation and prediction across space and time. Annual Review of Ecology Evolution and Systematics 40: 677-697.

Enserink, M. 1999. Biological invaders sweep in. Science 285: 1834-1836.

Evstatieva, L. N., N. V. Handjieva, S. S. Popov, and P. I. Pashankov. 1993. A biosystematic study of *Valeriana officinalis* (Valerianaceae) distributed in Bulgaria. Plant Systematics and Evolution 185 (3-4): 167-179.

Fielding, A. H. and J. F. Bell. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environmental Conservation 24 (1): 38-49.

Fitter, A. H. 1982. Influence of soil heterogeneity on the coexistence of grassland species. Journal of Ecology 70 (1): 139-148.

Flanders, A. A., W. P. Kuvlesky Jr., D. C. Ruthven III, R. E. Zaiglin, R. L. Bingham, T. E. Fulbright, F. Hernandez, and L. A. Brennan. 2006. Effects of invasive exotic grasses on south Texas rangeland breeding birds. The Auk 123: 171-182.

Galatowitsch, S. M., N. O. Anderson, and P. D. Ascher. 1999. Invasiveness in wetland plants in temperate North America. Wetlands 19 (4): 733-755.

Gleason, H. A. and Cronquist, A. C. (1991). Manual of vascular plants of Northeastern United States and adjacent Canada, second edition (corrected 2004). New York Botanical Garden, New York, NY, USA. xlvi + 993 pp.

Godwin, H. 1943a. Biological flora of the British Isles. Rhamnaceae L. Journal of Ecology 31 (1): 66-68.

Godwin, H. 1943b. Biological flora of the British Isles. *Frangula alnus* Miller. Journal of Ecology 31(1): 77-92.

Gourley, L.C. and E. Howell. 1984. Factors in buckthorn invasion documented; control measure checked (Wisconsin). Restoration and Management Notes 2 (2): 87.

Greenwood, H. D., J. O'Dowd, and P. S. Lake. 2004. Willow (*Salix* x *rubens*) invasion of the riparian zone in south-eastern Australia: reduced abundance and altered composition of terrestrial arthropods. Diversity and Distributions 10 (5-6): 485-492.

Guisan, A., A. Lehmann, S. Ferrier, M. Austin, J. Mc. C. Overton, R. Aspinall, and T. Hastie. 2006. Making better biogeographical predictions of species' distributions. Journal of Applied Ecology 43 (3): 386-392.

Harrington, R. A., B. J. Brown, and P. B. Reich. 1989. Ecophysiology of exotic and native shrubs in southern Wisconsin (USA): I. Relationship of leaf characteristics, resource availability, and phenology to seasonal patterns of carbon gain. Oecologia 80 (3): 356-367.

Hirzel, A. H. and R. Arlettaz. 2003. Modeling habitat suitability for complex species distributions by environmental-distance geometric mean. Environmental Management 32 (5): 614-623.

Hoffman, R. and K. Kearns, eds. 1997. Wisconsin manual of control recommendations for ecologically invasive plants. Wisconsin Department of Natural Resources. Madison, WI, USA. 102 pp.

Hutchinson, G. E. 1957. Concluding Remarks. Pages 415-427 *in*: Cold Springs Harbor Symposia on Quantitative Biology 22: 415-427.

[IPANE] Invasive Plant Atlas of New England. 2009. *Valeriana officinalis* (garden heliotrope, garden valerian). University of Connecticut and partners, Storrs, Connecticut, USA. http://nbii-nin.ciesin.columbia.edu/ipane/icat/browse.do?specieId=96 (June 2010).

Jones, C. C. and S. Reichard. 2009. Current and potential distributions of three non-native invasive plants in the contiguous USA. Natural Areas Journal 29 (4): 332-343.

Kearney, M. and W. Porter. 2009. Mechanistic niche modelling: Combining physiological and spatial data to predict species' ranges. Ecology Letters 12 (4): 334-350.

Lakela, O. 1938. Specimen 2560, DUL, MINN.

Leathwick, J. R. and Whitehead 2001. Soil and atmospheric water deficits and the distribution of New Zealand's indigenous tree species. Functional Ecology 15 (2): 233-242.

Leathwick, J. R., J. Elith, M. Francis, T. Hastie, and P. Taylor. 2006. Variation in demersal fish species richness in the oceans surrounding New Zealand: an analysis using boosted regression trees. Marine Ecology - Progress Series 321: 267-281.

Liu, C. R., P. M. Berry, T. P. Dawson, and R. G. Pearson. 2005. Selecting thresholds of occurrence in the prediction of species distributions. Ecography 28 (3): 385-393.

Lloyd, J. D. and T. E. Martin. 2005. Reproductive success of chestnut-collared longspurs in native and exotic grassland. Condor 107: 363-374.

Lokker, C. and P. B. Cavers. 1995. The effects of physical damage on seed production in flowering plants of *Saponaria officinalis*. Canadian Journal of Botany 73 (2): 235-243.

Lubke, M. A. and P. B. Cavers. 1969. Germination ecology of *Saponaria officinalis* from riverside gravel banks. Canadian Journal of Botany 47 (4): 529-535.

Mateo, R. G., T. B. Croat, Á. M. Felicísimo, and J. Muñoz. 2009. Profile or group discriminative techniques? Generating reliable species distribution models using pseudo-absences and target-group absences from natural history collections. Diversity and Distributions, 16, 84–94.

Miller, J., J. Franklin, and R. Aspinall. 2007. Incorporating spatial dependence in predictive vegetation models. Ecological Modelling 202 (3-4): 225-242.

Mills, E. L., J. H. Leach, J. T. Carlton, and C. L. Secor. 1993. Exotic species in the Great Lakes: A history of biotic crises and anthropogenic introductions. Journal of Great Lakes Research 19 (1): 1-54.

Moore, R. J. 1958. Cytotaxonomy of *Euphorbia esula* in Canada and its hybrid with *Euphorbia cyparissias*. Canadian Journal of Botany 36: 547-559.

Newsholme, C. 1992. Willows: The genus *Salix*. B. T. Batsford Ltd, London, UK, and Timber Press, Portland, OR, USA. 224 pp.

[NOAA-GLGIS] Great Lakes Geographic Information System, National Oceanographic and Atmospheric Administration, U. S. Department of Commerce. 2006. glgis Great Lakes shore noaa 70k. Downloaded from http://gis.glin.net/ogc/services.php?by=topic#inland (February 2010).

[NOAA-NCDC] National Climatic Data Center, National Oceanographic and Atmospheric Administration, U. S. Department of Commerce. 2009. No title. http://www.ncdc.noaa.gov/oa/ncdc.html (Decmber 2009).

Peet, R. K., J. D. Fridley, and J. M. Gramling. 2003. Variation in species richness and species pool size across a pH gradient in forests of the southern Blue Ridge Mountains. Folia Geobotanica 38 (8): 391–401.

Peterson, A. T., M. Papes, and M. Eaton. 2007. Transferability and model evaluation in ecological niche modeling: A comparison of GARP and Maxent. Ecography 30 (4): 550-560.

Peterson, A. T., M. Papes, and J. Soberon. 2008. Rethinking receiver operating characteristic analysis applications in ecological niche modeling. Ecological Modelling 213 (1): 63-72.

Phillips, S. J. 2009. A brief tutorial on Maxent. http://www.cs.princeton.edu/~schapire/maxent/tutorial/tutorial.doc (June 2010).

Phillips, S. J., R. P. Anderson, and R. E. Schapire. 2006. Maximum entropy modeling of species geographic distributions. Ecological Modelling 190 (3-4): 231-259.

Phillips, S. J. and M. Dudík. 2008. Modeling of species distributions with Maxent: New extensions and a comprehensive evaluation. Ecography 31: 161-175.

Phillips, S. J., M. Dudík, J. Elith, C. H. Graham, A. Lehmann, J. Leathwick, and S. Ferrier. 2009. Sample selection bias and presence-only distribution models: Implications for background and pseudo-absence data. Ecological Applications 19 (1): 181-197.

Pulliam, H. R. 2000. On the relationship between niche and distribution. Ecology Letters 3: 349-361.

Rangel, T. F. L. V. B., Diniz-Filho, J. A. F., and L. M. Bini. 2006. Towards an integrated computational tool for spatial analysis in macroecology and biogeography. Global Ecology and Biogeography 15: 321-327.

Rangel, T. F. L. V. B., Diniz-Filho, J. A. F., and L. M. Bini. 2010. SAM: A comprehensive application for Spatial Analysis in Macroecology. Ecography 33: 46-50.

Rehder, A. 1960, c1940. Manual of cultivated trees and shrubs hardy in North America, exclusive of the subtropical and warmer temperate regions. 2d edition, rev. and enl. The Macmillan Company, New York.

Reichard, S. H. and C. W. Hamilton. 1997. Predicting invasions of woody plants introduced into North America. Conservation Biology 11 (1): 193-203.

Rouget, M. and D. M. Richardson. 2003. Inferring process from pattern in plant invasions: a semimechanistic model incorporating propagule pressure and environmental factors. The

American Naturalist 162 (6): 713-724.

Schimpf, D. J., D. L. Pomroy, S. C. Garske, and J. C. Green. 2009. Noteworthy collections. The Michigan Botanist 48: 49-60.

Selleck, G. W., R. T. Coupland, and C. Frankton. 1962. Leafy spurge in Saskatchewan. Ecological Monographs 32 (1): 1-29.

Skvortsov A. K. 1973. Problems of Biogeocoenology, Geobotany and Plant Geography, Russian Academy of Science, Hayka Publishers: Leningrad, USSR (in Russian).

Soberon, J. and M. Nakamura. November 17, 2009. Niches and distributional areas: concepts, methods, and assumptions. Proceedings of the National Academy of Sciences 106: suppl. 2., 19644-19650. http://www.pnas.org/content/106/suppl.2/19644 (September 2010).

Stockwell, D.R.B. and D.P. Peters. 1999. The GARP modeling system: Problems and solutions to automated spatial prediction. International Journal of Geographic Information Systems 13: 143-158.

Sutherland, W. J. 1990. Biological flora of the British Isles (UK): *Iris pseudacorus* L. Journal of Ecology 78 (3): 833-848.

Tallamy, D. W. 2004. Do alien plants reduce insect biomass? Conservation Biology 18 (6): 1689-1692.

Tallamy, D. W. 2007. Bringing nature home: How you can sustain wildlife with native plants. Timber Press, Portland, OR, USA. 358 pp.

Thompson, D. Q., R. L. Stuckey, and E. B. Thompson. 1987. Spread, impact, and control of purple loosestrife (*Lythrum salicaria*) in North American wetlands. Publications Unit, U.S. Fish and Wildlife Service, U.S. Department of Interior, Washington, DC.

[UNEP] Secretariat of the Convention on Biological Diversity, United Nations Environmental Program. 2010. Global Biodiversity Outlook 3. Montréal, Canada. 94 pages. http://gbo3.cbd.int/ (May 2010).

[USDA-ARS] U.S. Department of Agriculture, Agricultural Research Service, Germplasm Resources Information Network. 2010. National Germplasm Resources Laboratory, Beltsville, Maryland, USA. *Lathyrus sylvestris* L. http://www.ars-grin.gov/cgi-bin/npgs/html/taxon.pl?21619 (June 2010).

[USDA-NRCS] U.S. Department of Agriculture, Natural Resources Conservation Service. 2010. The PLANTS database. National Plant Data Center, Baton Rouge, LA, USA. http://plants.usda.gov (June 2010).

[USDA-SSDS] Soil Survey Division Staff, Soil Conservation Service, U. S. Department of

Agriculture. 1993. Soil survey manual. Soil Conservation Service, U.S. Department of Agriculture Handbook 18. http://soils.usda.gov/technical/manual/ (June 2010).

[USDA-SSDS] Soil Survey Division Staff, Natural Resources Conservation Service, U. S. Department of Agriculture. 2007. Soil Data Mart. http://soildatamart.nrcs.usda.gov/ (March 2010).

[USDC] Geography Division, Bureau of the Census, U. S. Department of Commerce. 2000. TIGER/Line files, Redistricting Census 2000. Downloaded from http://www.esri.com/data/download/census2000-tigerline/index.html (February 2010).

[UWSP] Freckmann Herbarium, University of Wisconsin - Stevens Point. 2010. Vascular plants. http://wisplants.uwsp.edu/VascularPlants.html (May 2010).

Van Splunder, I., L. A. C. J. Voesenek, H. Coops, X. J. A. De Vries, and C. W. P. M. Blom. 1996. Morphological responses of seedlings of four species of Salicaceae to drought. Canadian Journal of Botany 74 (12): 1988-1995.

Voss, E. G. 1985. Michigan Flora, Part II. Dicots (Saururaceae-Cornaceae). Cranbrook Institute of Science and University of Michigan Herbarium, Ann Arbor, MI, USA.

Voss, E. G. 1996. Michigan Flora, Part III. Dicots (Pyrolaceae-Compositae). Cranbrook Institute of Science and University of Michigan Herbarium, Ann Arbor, MI, USA.

Walker, L. R. and S. D. Smith. 1997. Impacts of invasive plants on community and ecosystem properties. Pages 69-86 *in*: J. O. Lukin and J. W. Thieret, editors. Assessment and management of plant invasions. Springer-Verlag, New York, NY, USA.

Westbrooks, R. G. 1998. Invasive plants: Changing the landscape of America: fact book. Federal Interagency Committee for the Management of Noxious and Exotic Weeds, Washington, DC, USA.

Whitford, P.C. and P. B. Whitford. 1988. A note on nurse trees and browsing. Michigan Botanist 27 (4): 107-110.

Williamson, M. 1996. Biological Invasions. Population and community biology series: 15. Chapman and Hall, London, UK. 244 pp.

Wilcove, D. S., D. Rothstein, J. Dubow, A. Phillips, and E. Losos. 1998. Quantifying threats to imperiled species in the United States. BioScience 48: 607-615.

[WDNR] Wisconsin Department of Natural Resources. 2010. Surface water (hydrography) data. Wisconsin rivers and shorelines from 1:24,000-scale sources. Wisconsin Department of Natural Resources, Madison, Wisconsin, USA. http://dnr.wi.gov/maps/gis/datahydro.html (July 2010).

[WISCLAND] Wisconsin Initiative for Statewide Cooperation on Landscape Analysis and Data.

1998. WISCLAND Land Cover (WLCGW930). Wisconsin Department of Natural Resources (WDNR), Madison, Wisconsin, USA. http://dnr.wi.gov/maps/gis/datalandcover.html (March 2010).

Woodward, F. I. and B. G. Williams. 1987. Climate and plant distribution at global and local scales. Vegetatio 69 (1-3): 189-197.