

A Test of Four Models to Predict the Risk of Naturalization of Non-native Woody Plants in the Chicago Region¹

Mark P. Widrlechner², Janette R. Thompson³, Emily J. Kapler⁴,
Kristen Kordecki⁵, Philip M. Dixon⁶, and Galen Gates⁷
*United States Department of Agriculture — Agricultural Research Service
North Central Regional Plant Introduction Station
Departments of Agronomy and Horticulture
Iowa State University, Ames, Iowa 50011*

Abstract

Accurate methods to predict the naturalization of non-native woody plants are key components of risk-management programs being considered by nursery and landscape professionals. The objective of this study was to evaluate four decision-tree models to predict naturalization (first tested in Iowa) on two new sets of data for non-native woody plants cultivated in the Chicago region. We identified life-history traits and native ranges for 193 species (52 known to naturalize and 141 not known to naturalize) in two study areas within the Chicago region. We used these datasets to test four models (one continental-scale and three regional-scale) as a form of external validation. Application of the continental-scale model resulted in classification rates of 72–76%, horticulturally limiting (false positive) error rates of 20–24%, and biologically significant (false negative) error rates of 5–6%. Two regional modifications to the continental model gave increased classification rates (85–93%) and generally lower horticulturally limiting error rates (16–22%), but similar biologically significant error rates (5–8%). A simpler method, the CART model developed from the Iowa data, resulted in lower classification rates (70–72%) and higher biologically significant error rates (8–10%), but, to its credit, it also had much lower horticulturally limiting error rates (5–10%). A combination of models to capture both high classification rates and low error rates will likely be the most effective until improved protocols based on multiple regional datasets can be developed and validated.

Index words: exotic plant, invasive plants, life history, native range, risk assessment, shrub, tree, validation.

Significance to the Nursery Industry

Nursery and landscape professionals continue to search for new woody landscape plants that will provide consumers with visual interest and diversity in a variety of managed settings. Interest in identifying and propagating additional native species helps to address consumer demand for new plant materials. At the same time, broader use of new, non-native species is likely to continue. Expanding the palette of both native and non-native species used in the landscape is of concern to scientists, managers of natural areas, horticulturists, and some members of the general public due to the potential for widely used introductions to become invasive pests. Although relatively few of these non-native woody species will naturalize, and even fewer are likely to become invasive, invasions can have significant impacts and their control is difficult and costly. One alternative to lengthy test

periods before new woody plants are released for introduction is the use of predictive modeling to assess risk associated with certain plants. Use of models that predict plant invasiveness (which classify species into categories of ‘accept’, ‘reject’, or ‘analyze further’) could allow time-consuming and expensive field screening to be focused only on species of greatest concern or those for which basic information is lacking. This paper describes efforts to find accurate, rapid, and relatively easy-to-use methods for identifying species that could become invasive pests. Our results suggest that the power and accuracy of regional-scale models that include both life-history and geographic-risk components make them very useful as part of an overall risk-management program.

Introduction

Many useful, non-native species of woody plants have been imported and disseminated in the United States for commercial, conservation, and ornamental purposes (15, 22, 23). Although relatively few of these species have become naturalized, and fewer still can be defined as invasive, controlling the spread of those that have disrupted natural plant communities or agricultural systems is very difficult and often costly (28). A high percentage of invaders were originally introduced for horticultural use (1), giving horticulturists a critical role in the prevention of future invasive threats. For all parties involved in the introduction of new landscape plants, this responsibility is particularly salient given the growth of global trade and high demand for new cultivars and the importation of new taxa (4, 15).

Identifying a reliable method for accurately predicting potential invasiveness for otherwise desirable exotic species is a rapidly growing area of research. For example, a recent search of the BIOSIS Previews database identified 64 peer-reviewed articles published in scientific journals between 1990 and 2008 that address the topic of risk assessment for

¹Received for publication May 13, 2009; in revised form August 3, 2009. Journal paper of the Iowa Agriculture and Home Economics Experiment Station, Ames, IA, and supported by Hatch Act, McIntire-Stennis, and State of Iowa funds. The authors thank Candice Gardner, Doria Gordon, Jeffery Iles, Peter Bristol, and two anonymous reviewers for their valuable critiques. We are also grateful for financial support received from USDA-ARS through the Floral and Nursery Crops Research Initiative. Mention of commercial brand names does not constitute an endorsement of any product by the U.S. Department of Agriculture or cooperating agencies.

²USDA-ARS Horticulturist and Assistant Professor (Collaborator), Departments of Agronomy and Horticulture, Iowa State University.

³Associate Professor, Department of Natural Resource Ecology and Management, Iowa State University.

⁴Graduate Research Assistant, Department of Natural Resource Ecology and Management, Iowa State University.

⁵Woody Plants Specialist, Chicago Botanic Garden, Glencoe, IL; current affiliation: Watershed Coordinator, Acadiana Resource Conservation and Development Council, Carencro, LA.

⁶Professor, Department of Statistics, Iowa State University.

⁷Director of Plant Collections, Chicago Botanic Garden.

invasive, terrestrial plants (excluding genetically-modified organisms). Thirty-three of those articles were published between 2006 and 2008. Predictive models support the development of effective risk-management programs for invasive species (16), and are critical for providing information that enables more time-consuming or expensive work (field screening, containment, or eradication) to be focused on the highest-risk species (3, 16).

Among the most frequently cited predictive models for risk assessment are an Australian Weed Risk Assessment (WRA) protocol described by Pheloung et al. (27) and a North American woody plant risk-assessment model developed by Reichard and Hamilton (29). Efforts to apply these models to regions beyond their original scope have rapidly expanded in recent years (5, 12, 16, 17, 21, 44). Other studies have examined the expected economic benefits resulting from the use of risk-assessment protocols (18, 30), and tested the relative ease of model implementation (12, 13, 16), in addition to evaluating the utility and accuracy of the protocols themselves.

Existing predictive models usually assign each species in question to one of three categories: 'accept', where the risk of invasion is low; 'reject', where the risk is high; or 'further analysis', where field monitoring or additional information is needed. These categories then can be tested against *a priori* categories of non-native plants already introduced (12, 13, 29, 44) to estimate both the power and accuracy of the models by determining classification and error rates.

For woody plants, there is considerable evidence that plant adaptation and distribution in the temperate United States strongly reflects regional variation in climate (33), especially as related to moisture balance and extreme temperature events, and to soils and photoperiod regimens (40). Therefore, predictive models that recognize traits important for woody-plant naturalization on a regional basis (11, 26) should be more powerful and accurate than those based solely on national or continental data. For example, limitations in the effectiveness of a continental model were noted when that model was applied to a single state. Widrechner et al. (44) examined the efficacy of Reichard and Hamilton's (29) continental model, by using it to test 100 non-native woody plants widely cultivated in Iowa; the resulting classification rate was only 65% and the error rate was 20% (44).

Widrechner et al. (44) developed and tested three additional models that combined plant life-history traits (including those from the continental model (29), along with fleshy, bird-dispersed fruits) and geographic/climatic information (42, 43) to predict naturalization of the same 100 woody plants for Iowa. A geographic-risk value (G-value) was calculated for each species based on its native range and the proportion of species from its native range known to naturalize in the test area (43, 44). Models that included both life-history and geographic-risk components were more powerful (classified up to 90% of species) and accurate (error rates of 4 to 20%) for risk assessment in Iowa (44) than was the continental model (29).

Ideally, new models are validated or rejected by testing them against datasets that differ from the dataset used for model development (external validation) or through careful statistical manipulation of the base dataset itself (internal validation) (14). For the types of risk-assessment models discussed here, validation can be accomplished in one of three ways: [1] internal validation during model development

by excluding species in the initial dataset through data-splitting, bootstrapping, and other re-sampling techniques (as described in 14); [2] external validation by applying the model to a test dataset of new or additional species from the same target region; or [3] a more general external validation by applying the model to a dataset from another, similar region (5, 12, 13, 16, 17, 21).

In this paper, we report on an effort to use the third type of validation, by evaluating the performance of four risk-assessment models, first tested in Iowa, on two new species datasets developed for the Chicago region, including 28 counties in northeastern Illinois, northwestern Indiana, southwestern Michigan, and southeastern Wisconsin (Fig. 1).

Materials and Methods

We defined two subsets of the Chicago region (Fig. 1) to test the continental risk-assessment model developed by Reichard and Hamilton (29) and three additional models developed by Widrechner et al. (44) for Iowa. These study areas were of comparable area and internally consistent with respect to climate and soils, with 'Chicago A' comprised of counties in Illinois and Wisconsin with soils that are primarily calcareous clays and loams (37) and 'Chicago B' comprised of counties in Indiana and Michigan with soils that are primarily sandy, organic, or a combination of the two (37). The soils and climate of Chicago A, in terms of moisture balance (41) and mean January temperatures (38), resemble those found in Iowa more closely than do the soils and climate of Chicago B. Thus, we hypothesize that the three

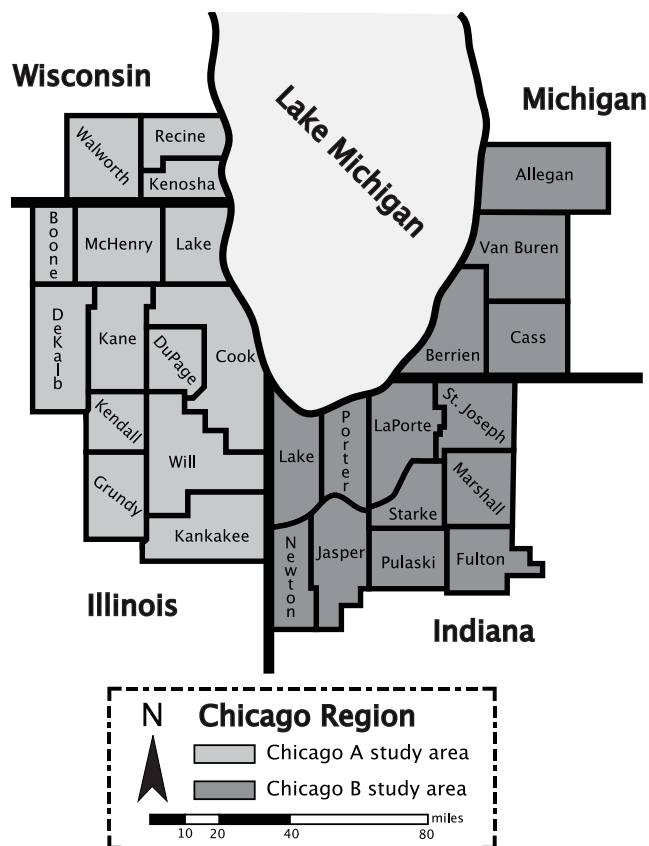


Fig. 1. Map of the Chicago A and Chicago B study areas.

models developed for Iowa will have greater applicability in Chicago A than in Chicago B.

A set of 135 commonly cultivated species of woody landscape plants (Table 1) not known to be native to Chicago A in the period predating European settlement in the early 1800s was identified by staff of the Chicago Botanic Garden for the purpose of testing risk-assessment models. A second set of 142 commonly cultivated woody species (Table 1) not known to be native to Chicago B before the early 1800s was identified through initial input from five nursery and public garden professionals with extensive experience in the region (Steve Bornell, Brian Bunge, Michael Dosmann,

Robert Schutzki, and Tim Woods), and review by the authors. These sets (Table 1) were initially somewhat larger, but were standardized by randomly removing species not known to naturalize until approximately 28% of each set was comprised of species with a record of naturalizing in at least two well-documented sites in their corresponding study areas. This was done so the Chicago datasets could directly use the three models that were developed from Iowa data, where 28% of the analyzed species were known to naturalize (44). Naturalization status was determined by reviewing local floras (6, 32, 39) and consultation with local floristic experts (Richard Rabeler, Gerould Wilhelm, Dennis Woodland and

Table 1. Study area and naturalization status of species in Chicago datasets.

Species ^a	Used in Chicago A	Used in Chicago B	Naturalized in Chicago A	Naturalized in Chicago B
<i>Abies balsamea</i> (L.) Mill.		X		
<i>Abies concolor</i> (Gordon & Glend.) Lindl. ex Hildebr.	X	X		
<i>Abies fraseri</i> (Pursh) Poir.		X		
<i>Abies nordmanniana</i> (Steven) Spach	X			
<i>Acer griseum</i> (Franch.) Pax		X		
<i>Acer palmatum</i> Thunb.		X		
<i>Acer platanoides</i> L.	X	X	X	X
<i>Acer tataricum</i> L. subsp. <i>ginnala</i> (Maxim.) Wesm.	X	X	X	
<i>Actinidia arguta</i> (Siebold & Zucc.) Planch. ex Miq.	X	X		
<i>Aesculus hippocastanum</i> L.	X			
<i>Aesculus parviflora</i> Walter		X		
<i>Aesculus pavia</i> L.		X		
<i>Ailanthus altissima</i> (Mill.) Swingle	X	X	X	X
<i>Alnus glutinosa</i> (L.) Gaertn.	X	X	X	X
<i>Amelanchier alnifolia</i> (Nutt.) Nutt. ex M. Roem.	X	X		
<i>Amelanchier canadensis</i> (L.) Medik.	X	X		
<i>Ampelopsis glandulosa</i> (Wall.) Momiy.	X	X	X	
<i>Aristolochia macrophylla</i> Lam.	X	X		
<i>Aronia arbutifolia</i> (L.) Pers.	X	X		
<i>Berberis thunbergii</i> DC.	X	X	X	X
<i>Betula pendula</i> Roth		X	X ^b	X
<i>Buddleja davidii</i> Franch.		X		
<i>Buxus microphylla</i> Siebold & Zucc.		X		
<i>Buxus sempervirens</i> L.		X		
<i>Callicarpa dichotoma</i> (Lour.) K. Koch		X		
<i>Callicarpa japonica</i> Thunb.		X		
<i>Calycanthus floridus</i> L.	X			
<i>Campsis radicans</i> (L.) Seem. ex Bureau	X	X	X	X
<i>Caragana arborescens</i> Lam.	X	X		
<i>Carpinus betulus</i> L.	X	X		
<i>Castanea mollissima</i> Blume	X			
<i>Catalpa speciosa</i> (Warder ex Barney) Warder ex Engelm.	X	X	X	X
<i>Celastrus orbiculatus</i> Thunb.	X	X	X	X
<i>Cercidiphyllum japonicum</i> Siebold & Zucc.		X		
<i>Chaenomeles japonica</i> (Thunb.) Lindl. ex Spach	X			
<i>Chaenomeles speciosa</i> (Sweet) Nakai		X		
<i>Chamaecyparis obtusa</i> (Siebold & Zucc.) Endl.		X		
<i>Chionanthus virginicus</i> L.	X	X		
<i>Cladrastis kentukea</i> (Dum. Cours.) Rudd	X	X		
<i>Clematis terniflora</i> DC.		X		
<i>Clethra alnifolia</i> L.	X	X		
<i>Cornus kousa</i> Hance	X	X		
<i>Cornus mas</i> L.	X	X		
<i>Corylus avellana</i> L.	X			
<i>Corylus colurna</i> L.		X		
<i>Cotinus coggygria</i> Scop.	X			
<i>Cotoneaster acutifolius</i> Turcz.	X			
<i>Cotoneaster apiculatus</i> Rehder & E.H. Wilson	X	X		
<i>Cotoneaster dammeri</i> C. K. Schneid.		X		
<i>Cotoneaster divaricatus</i> Rehder & E. H. Wilson	X			
<i>Cotoneaster horizontalis</i> Decne.		X		
<i>Cotoneaster multiflorus</i> Bunge	X			
<i>Crataegus laevigata</i> (Poir.) DC.		X		
<i>Crataegus phaenopyrum</i> (L. f.) Medik.		X		

Table 1 continued ...

Table 1. Continued ...

Species ^z	Used in Chicago A	Used in Chicago B	Naturalized in Chicago A	Naturalized in Chicago B
<i>Crataegus viridis</i> L.		X		
<i>Daphne mezereum</i> L.	X			
<i>Deutzia gracilis</i> Siebold & Zucc.	X	X		
<i>Deutzia scabra</i> Thunb.	X			
<i>Diervilla sessilifolia</i> Buckley	X			
<i>Diospyros virginiana</i> L.	X			
<i>Elaeagnus angustifolia</i> L.	X	X	X	
<i>Elaeagnus commutata</i> Bernh. ex Rydb.	X	X		
<i>Elaeagnus umbellata</i> Thunb.	X	X	X	X
<i>Eleutherococcus sieboldianus</i> (Makino) Koidz.	X			
<i>Euonymus alatus</i> (Thunb.) Siebold	X	X	X	X
<i>Euonymus europaeus</i> L.	X		X	
<i>Euonymus fortunei</i> (Turcz.) Hand.-Mazz.	X	X	X	X
<i>Euonymus hamiltonianus</i> Wall.		X		X
<i>Exochorda racemosa</i> (Lindl.) Rehder	X			
<i>Fagus sylvatica</i> L.	X	X		
<i>Forsythia suspensa</i> (Thunb.) Vahl	X			
<i>Fothergilla gardenii</i> L.	X	X		
<i>Frangula alnus</i> Mill.	X	X	X	X
<i>Halesia tetraptera</i> J. Ellis	X			
<i>Hamamelis mollis</i> Oliv.		X		
<i>Hamamelis vernalis</i> Sarg.	X	X		
<i>Hedera helix</i> L.	X	X		X
<i>Heptacodium miconioides</i> Rehder	X			
<i>Hibiscus syriacus</i> L.	X			
<i>Hydrangea anomala</i> D. Don		X		
<i>Hydrangea macrophylla</i> (Thunb.) Ser.		X		
<i>Hydrangea paniculata</i> Siebold		X		
<i>Hydrangea quercifolia</i> W. Bartram	X	X		
<i>Hypericum calycinum</i> L.	X			
<i>Ilex glabra</i> (L.) A. Gray	X	X		
<i>Itea virginica</i> L.	X			
<i>Juglans regia</i> L.	X			
<i>Juniperus chinensis</i> L. var. <i>procumbens</i> Endl.		X		
<i>Juniperus sabina</i> L.		X		
<i>Juniperus scopulorum</i> Sarg.	X	X		
<i>Juniperus squamata</i> Buch.-Ham. ex D. Don	X	X		
<i>Kalmia latifolia</i> L.		X		
<i>Kerria japonica</i> (L.) DC.	X			
<i>Koeleruteria paniculata</i> Laxm.	X	X		
<i>Kolkwitzia amabilis</i> Graebn.	X	X		
<i>Larix decidua</i> Mill.	X	X		
<i>Leucothoe fontanesiana</i> (Steud.) Sleumer		X		
<i>Ligustrum obtusifolium</i> Siebold & Zucc.	X	X	X	X
<i>Ligustrum vulgare</i> L.	X	X	X	X
<i>Liquidambar styraciflua</i> L.	X	X		
<i>Lonicera fragrantissima</i> Lindl. & Paxton	X			
<i>Lonicera japonica</i> Thunb.		X		X
<i>Lonicera korolkowii</i> Stapf	X			
<i>Lonicera maackii</i> (Rupr.) Maxim.		X	X ^y	X
<i>Lonicera morrowii</i> A. Gray	X		X	
<i>Lonicera sempervirens</i> L.	X	X		
<i>Lonicera tatarica</i> L.	X	X	X	X
<i>Lonicera xylosteum</i> L.	X		X	
<i>Lycium barbarum</i> L.		X	X ^y	X
<i>Maclura pomifera</i> (Raf.) C. K. Schneid.	X	X	X	X
<i>Magnolia liliifera</i> (L.) Baill.		X		
<i>Magnolia stellata</i> (Siebold & Zucc.) Maxim.	X	X		
<i>Malus baccata</i> (L.) Borkh.	X		X	
<i>Metasequoia glyptostroboides</i> Hu & W. C. Cheng	X	X		
<i>Microbiota decussata</i> Kom.	X	X		
<i>Morella pensylvanica</i> (Mirb.) Kartesz	X	X		
<i>Morus alba</i> L.	X	X	X	X
<i>Phellodendron amurense</i> Rupr.	X		X	
<i>Phellodendron sachalinense</i> (F. Schmidt) Sarg.	X	X		X
<i>Philadelphus coronarius</i> L.	X			
<i>Picea abies</i> (L.) H. Karst.	X	X		
<i>Picea glauca</i> (Moench) Voss	X	X		
<i>Picea omorika</i> (Pancic) Purk.	X	X		
<i>Picea pungens</i> Engelm.	X	X		

Table 1 continued ...

Table 1. Continued ...

Species ^z	Used in Chicago A	Used in Chicago B	Naturalized in Chicago A	Naturalized in Chicago B
<i>Pieris japonica</i> (Thunb.) D. Don ex G. Don	X	X		
<i>Pinus flexilis</i> E. James	X			
<i>Pinus mugo</i> Turra		X		
<i>Pinus nigra</i> J.F. Arnold	X	X	X	X
<i>Pinus ponderosa</i> P&C Lawson	X			
<i>Pinus sylvestris</i> L.	X	X	X	X
<i>Populus alba</i> L.	X	X	X	X
<i>Populus simonii</i> Carrière	X			
<i>Populus xcanescens</i> (Aiton) Sm.		X	X ^y	X
<i>Prunus avium</i> (L.) L.		X		X
<i>Prunus cerasifera</i> Ehrh.		X		
<i>Prunus mahaleb</i> L.		X	X ^y	X
<i>Prunus maritima</i> Marshall	X			
<i>Prunus padus</i> L.	X		X	
<i>Prunus serrulata</i> Lindl.	X	X		
<i>Prunus subhirtella</i> Miq.	X	X		
<i>Prunus tomentosa</i> Thunb.	X			
<i>Pseudotsuga menziesii</i> (Mirb.) Franco	X			
<i>Pyracantha coccinea</i> M. Roem.	X	X		
<i>Quercus acutissima</i> Carruth.	X			
<i>Quercus robur</i> L.	X	X		
<i>Quercus shumardii</i> Buckley	X			
<i>Rhamnus cathartica</i> L.		X	X ^y	X
<i>Rhododendron catawbiense</i> Michx.		X		
<i>Rhodotypos scandens</i> (Thunb.) Makino	X		X	
<i>Ribes alpinum</i> L.		X		
<i>Robinia pseudoacacia</i> L.	X	X	X	X
<i>Rosa multiflora</i> Thunb.	X	X	X	X
<i>Rosa rugosa</i> Thunb.	X	X		X
<i>Rosa virginiana</i> Mill.	X			
<i>Salix alba</i> L.	X	X	X	X
<i>Salix caprea</i> L.	X	X		
<i>Salix fragilis</i> L.	X		X	X ^y
<i>Salix integra</i> Thunb.		X		
<i>Salix purpurea</i> L.	X	X	X	X
<i>Sambucus nigra</i> L.		X		
<i>Sorbus alnifolia</i> (Siebold & Zucc.) K. Koch	X			
<i>Sorbus aucuparia</i> L.	X	X	X	X
<i>Spiraea japonica</i> L.f.	X	X		
<i>Spiraea nipponica</i> Maxim.	X	X		
<i>Spiraea prunifolia</i> Siebold & Zucc.	X			
<i>Spiraea thunbergii</i> Siebold ex Blume	X	X		
<i>Stephanandra incisa</i> (Thunb.) Zabel		X		
<i>Symphoricarpos orbiculatus</i> Moench		X	X ^y	X
<i>Syringa meyeri</i> C. K. Schneid.		X		
<i>Syringa pubescens</i> Turcz. subsp. <i>patula</i> (Palib.) M. C. Chang & X. L. Chen		X		
<i>Syringa reticulata</i> (Blume) H. Hara		X		
<i>Syringa vulgaris</i> L.	X	X		
<i>Taxodium distichum</i> (L.) Rich.	X	X		
<i>Taxus baccata</i> L.		X		
<i>Taxus cuspidata</i> Siebold & Zucc.		X		
<i>Thuja plicata</i> Donn ex D. Don		X		
<i>Tilia cordata</i> Mill.	X	X		
<i>Tilia tomentosa</i> Moench	X	X		
<i>Tsuga canadensis</i> (L.) Carrière	X			
<i>Ulmus glabra</i> Huds.	X			
<i>Ulmus parvifolia</i> Jacq.	X	X		
<i>Ulmus pumila</i> L.	X	X	X	X
<i>Viburnum carlesii</i> Hemsl.	X	X		
<i>Viburnum farreri</i> Stearn	X			
<i>Viburnum lantana</i> L.	X	X	X	X
<i>Viburnum opulus</i> L.	X	X	X	X
<i>Viburnum plicatum</i> Thunb.		X		
<i>Viburnum rhytidophyllum</i> Hemsl.		X		
<i>Viburnum sieboldii</i> Miq.		X		
<i>Weigela florida</i> (Bunge) A. DC.		X		
<i>Wisteria floribunda</i> (Willd.) DC.		X		
<i>Wisteria sinensis</i> (Sims) DC.	X			
<i>Zelkova serrata</i> (Thunb.) Makino	X			

^zTaxonomy follows the GRIN database (35).^yNaturalized in study area, but not included in species analyzed.

Kay Yatskievych). This effort results in a total of 193 species for our analyses.

Separate spreadsheets were compiled containing information on nine life-history and invasive characteristics, as described by Widrlechner et al. (44), for each set of species (i.e., for Chicago A and Chicago B):

- Evergreen foliage
- Fleshy, bird-dispersed fruits
- Group invasive in North America (including *Euonymus*)
- Invades outside North America
- Native to North America
- Quick maturity
- Quick vegetative spread
- Germination pretreatment required
- Sterile hybrid

These data were gleaned from several published and online sources, available on request from the authors, and refined through the collective experiences of Chicago Botanic Garden staff members in observing and cultivating these plants in the region by a process of consensus-driven discussions and interviews. 'Group invasive in North America' as defined by Reichard and Hamilton (29) was expanded to include the genus *Euonymus* within this category (Reichard, personal communication, 2002).

Two additional spreadsheets were compiled with detailed information about the native ranges of all 193 species, including 424 geographical and political subdivisions for the ranges of species included in Chicago A dataset and 459 subdivisions for the Chicago B dataset. Native-range data primarily were obtained from the USDA-ARS Germplasm Resources Information Network database (35), supplemented by selective use of the USDA-NRCS PLANTS database (36) and published floras (7, 9, 10, 19, 24, 25, 34).

For every subdivision (and separately for both datasets), we calculated the proportion of native species evaluated that have naturalized in the study area. Following Widrlechner et al. (44), we refer to this proportion as a P ratio. From the P ratios, we calculated range-wide, geographic-risk values, or G values, for each species by using procedures described by Widrlechner et al. (44). For reference, G is the mean of the unweighted P ratios for all n geographic subdivisions included in a species' native range, or

$$G = \left(\sum_{i=1}^n P \right) / n.$$

The nine life-history and invasive characteristics and G-values for the 135 species evaluated in the Chicago A tests and the 142 species in the Chicago B tests (referred to as master datasets) are available at <http://www.ars.usda.gov/Main/docs.htm?docid=17734>.

Once the master datasets were assembled and verified, we subjected the species therein to four risk-assessment models. These models included [1] the 'Continental Decision Tree' defined by Reichard and Hamilton (29), and three additional models developed by Widrlechner et al. (44) based on an evaluation of data collected in Iowa: [2] the 'Modified Decision Tree', which adds ten steps to the Continental Decision Tree, and [3] the 'Decision Tree/Matrix Model', a less complex refinement focused only on evaluating species

that required further analysis as a result of the Continental Decision Tree, and [4] a new model, generated independently from the others, the 'CART Model', based on an application of Classification and Regression Trees (CART) to the Iowa dataset (44).

Each of these four approaches assigned species to one of three categories: reject, accept, or further analysis. 'Reject' indicated a high risk of invasiveness; 'accept' indicated low-risk species; and the 'further analysis' category was reserved for species for which the model failed to provide clear guidance.

The power and accuracy of the models were assessed separately for the Chicago A and Chicago B master datasets by methods described in Widrlechner et al. (44). Briefly, the power of each model was measured by calculating Reichard and Hamilton's (29) classification rate, the proportion of evaluated species assigned to the 'reject' or 'accept' categories. The accuracy of the models was measured by calculating two error rates, which document biologically significant and horticulturally limiting misclassifications, i.e., the proportions of false negatives and false positives, respectively. In other words, the assignment of a naturalizing species to the 'accept' category creates a biologically significant error by allowing the introduction of a new, potentially invasive species, and the assignment of a species with no record of naturalization to the 'reject' category creates a horticulturally limiting error by excluding a new species with little or no risk to natural habitats. Both error rates are expressed as proportions of the total number of classified species.

Differences in classification and error rates between the Iowa, Chicago A or Chicago B data sets and all external data sets were evaluated by using a Chi-square test (8); however, this test is not appropriate for comparing classification or error rates among the Iowa, Chicago A and Chicago B data sets, because the Chi-square test assumes independence among observations. Such an assumption is invalid when the same species are components of two data sets and/or are tested in two models. To compare classification rates between two models when applied to the same data set, we used McNemar's test for paired responses (8). This test allows analysis of paired data and the difference in resulting classification and error rates.

For all other rate comparisons, such as those of horticulturally limiting error rates in the Chicago A and Chicago B data sets, the data are a mix of paired responses (for species evaluated in both data sets) and unpaired responses (for those evaluated in only one data set). There is no named statistical test for this situation. A likelihood-ratio test was developed to test the null hypothesis of no difference in error rates between two models or two data sets (available from the authors on request). Because the sample sizes were often small, p-values were computed by randomization (31).

In addition, for those species included in at least two of three comparable datasets, i.e., in Iowa, Chicago A, and Chicago B, we calculated correlation coefficients for G-values in pairs of data sets to evaluate the degree of similarity between these three regions in risk directly related to the species' native ranges. A correlation of 1 indicates a perfect correspondence between G-values. This would be true if regional differences in the underlying P-values (risk values for specific geographic subdivisions) that serve as the basis for the calculation of G-values are small. The hypothesis that the correlation between Chicago A and Iowa should be

Table 2. Summary of classification and error rates for four risk-assessment models by data set.

Model	Classification rate (%)	Biologically significant error rate (%)	Horticulturally limiting error rate (%)
Continental Decision Tree			
Chicago A	71.9	5.2	23.7
Chicago B	76.1	5.6	20.4
Iowa Data	65.0	3.1	16.9
Modified Decision Tree			
Chicago A	92.6	8.0	21.6
Chicago B	88.7	7.1	15.9
Iowa Data	90.0	3.3	13.3
Decision Tree/Matrix Model			
Chicago A	87.4	5.9	21.2
Chicago B	85.2	5.0	18.2
Iowa Data	85.0	3.5	16.4
CART Model			
Chicago A	70.4	8.4	9.5
Chicago B	72.5	9.7	4.5
Iowa Data	81.0	2.5	3.7

higher than that for Chicago B and Iowa was tested by using a Fisher Z transformation (31).

Results and Discussion

Continental decision tree. Both sets of non-native, woody species (representing Chicago A and Chicago B) (Table 1) were evaluated with the Continental Decision Tree developed by Reichard and Hamilton (29). The decision tree was able to classify 71.9% of species in Chicago A and 76.1% of species in Chicago B for a mean classification rate of 74% (Table 2). This rate approximates the 76% classification rate reported in a recent study of non-native woody plants in the Czech Republic by Křivánek and Pyšek (21). Chicago classification rates approach those reported by Reichard and Hamilton (29) (80% classification) more closely than did rates reported in the Iowa study by an average of 9%, but the differences are not statistically significant ($p > 0.15$).

While increased classification rates are generally desirable, in the case of the Chicago datasets, we did observe small increases in error rates compared to the Iowa study. Only 69 of the 97 (71.1%) classifications in Chicago A and 80 of the 108 (74.1%) classifications in Chicago B were correct, producing a mean error rate of 27.4%. Mean error rates determined by Widrlechner et al. (44) in Iowa and Křivánek and Pyšek (21) in the Czech Republic were 20%, not statistically significant from the findings in our tests of Chicago A ($p > 0.09$) or Chicago B ($p > 0.2$). Most misclassifications from the Chicago A and Chicago B datasets were rejections of species not known to naturalize, or horticulturally limiting errors (Table 2). This is analogous to findings by Křivánek and Pyšek (21) who also reported a higher proportion of such rejections. Of the species misclassified in the Chicago datasets, about 60% resulted from the short, right side of the decision tree (29). This suggests that additional steps incorporated into this branch of the tree during the Iowa study (44) (producing the Modified Decision Tree) would also be warranted for the Chicago datasets.

The remaining misclassifications (a mean value of 5.4% for Chicago A and B, representing 5 and 6 species, respectively) resulted in the acceptance of species already known to naturalize, or biologically significant errors. Those species that resulted in biologically significant errors when evaluated are listed by model in Table 3. Although the estimated biologically significant error rates are slightly larger for the

two Chicago datasets than for the Iowa dataset, the differences are not statistically significant ($p > 0.50$).

Overall, our assessment of the Chicago datasets reinforces the need for revisions to the Continental Decision Tree in order to enhance both classification rates and accuracy.

Modified decision tree. Widrlechner et al. (44) developed the Modified Decision Tree to increase the power of the Continental Decision Tree by reexamining the more problematic unclassified and/or misclassified species. This involved two primary modifications: [1] reevaluating the ‘further analysis’ species based upon G-values (risk related to the species’ native range) and the presence or absence of fleshy, bird-dispersed fruits with the goal of increasing classification rates, and [2] reevaluating ‘reject’ species from the right-hand side of the decision tree to reduce horticulturally limiting errors. Both of these modifications produced results for the Chicago datasets that resembled results for the Iowa data: increased classification rates and decreased horticulturally limiting errors. Chicago A’s classification rate increased to 92.6% and Chicago B’s classification rate increased to 88.7% (Table 2), representing a mean increase of 16.7% over the Continental Decision Tree, both differences highly significant ($p < 0.0001$ for Chicago A and $p = 0.0005$ for Chicago B). This is a smaller increase than observed in the Iowa data (25.0%) (44), but these increases in classification rates are not significantly different from each other ($p = 0.38$).

Additional validation is found in decreased horticulturally limiting error rates for both Chicago datasets (Table 2). On average, the Chicago horticulturally limiting error rate decreased by 3.3% compared to the Continental Decision Tree, similar to Iowa’s 3.6% decrease. Where this model fails to perform adequately is in the biologically significant error rate, with rates of 7.1 to 8.0%. These values exceed the rate reported from the Iowa dataset, but the differences are not statistically significant ($p = 0.51$ for Chicago A and Iowa, $p = 0.52$ for Chicago B and Iowa). The list of species misclassified in this way for the Chicago datasets resembled that from the Continental Decision Tree with some additions (Table 3). Taking this shortcoming into account, the overall accuracy of the Modified Decision Tree for the Chicago datasets was little improved over the Continental Decision Tree. Of the 125 classified species for Chicago A, only 88 were correct (70.4%); of the 126 classified species for Chicago B, 97 were

Table 3. Species producing biologically significant errors for Chicago A and Chicago B.

Continental Decision Tree	Modified Decision Tree
<i>Berberis thunbergii</i> DC. <i>Catalpa speciosa</i> (Warder ex Barney) Warder ex Engelm. <i>Elaeagnus umbellata</i> Thunb. <i>Frangula alnus</i> Mill. <i>Phellodendron amurense</i> Rupr. <i>Philadelphus coronarius</i> L. <i>Symphoricarpos orbiculatus</i> Moench	<i>Acer platanoides</i> L. <i>Berberis thunbergii</i> DC. <i>Campsis radicans</i> (L.) Seem ex Bureau <i>Catalpa speciosa</i> (Warder ex Barney) Warder ex Engelm. <i>Elaeagnus umbellata</i> Thunb. <i>Frangula alnus</i> Mill. <i>Maclura pomifera</i> (Raf.) C. K. Schneid. <i>Phellodendron amurense</i> Rupr. <i>Philadelphus coronarius</i> L. <i>Pinus nigra</i> J.F. Arnold <i>Pinus sylvestris</i> L. <i>Symphoricarpos orbiculatus</i> Moench
Decision Tree/Matrix Model	CART Model
<i>Berberis thunbergii</i> DC. <i>Campsis radicans</i> (L.) Seem. ex Bureau <i>Catalpa speciosa</i> (Warder ex Barney) Warder ex Engelm. <i>Elaeagnus umbellata</i> Thunb. <i>Frangula alnus</i> Mill. <i>Maclura pomifera</i> (Raf.) C. K. Schneid. <i>Phellodendron amurense</i> Rupr. <i>Philadelphus coronarius</i> L. <i>Symphoricarpos orbiculatus</i> Moench	<i>Acer platanoides</i> L. <i>Berberis thunbergii</i> DC. <i>Elaeagnus umbellata</i> Thunb. <i>Euonymus alatus</i> (Thunb.) Siebold <i>Euonymus fortunei</i> (Turcz.) Hand.-Mazz. <i>Euonymus hamiltonianus</i> Wall. <i>Maclura pomifera</i> (Raf.) C. K. Schneid. <i>Philadelphus coronarius</i> L. <i>Pinus nigra</i> J.F. Arnold <i>Pinus sylvestris</i> L. <i>Rhodotypos scandens</i> (Thunb.) Makino

correct (77%). This produced a mean error rate of 26.3% — not significantly different from the overall error rates for the original decision tree ($p = 0.91$ for Chicago A and $p = 0.75$ for Chicago B). In summary, the Modified Decision Tree yielded the highest classification rate of the tested models (Table 2), but did not improve overall accuracy in these tests.

Decision tree/Matrix model. A second refinement of the Continental Decision Tree based on Iowa data produced a model focused on reevaluating ‘further analysis’ species by using G-values and fleshy, bird-dispersed fruits (44). Model classification rates for both Chicago datasets resembled those for the Iowa data (Table 2). The Decision Tree/Matrix Model was able to classify a mean of 86.2% species for Chicago A and B compared to Iowa’s classification rate of 85%, generally supporting the model’s ability to increase classification rates. However, the improvement in classification rates over those produced by the Continental Decision Tree for the Chicago A dataset (15.5%) resembled that reported in the Iowa study (20.0%) ($p = 0.92$), while the improvement for the Chicago B dataset (9.1%) was significantly less ($p = 0.002$). In part, this can be explained by the higher classification rate of the Continental Decision Tree for the Chicago B dataset.

Application of the Decision Tree/Matrix Model to the Iowa data yielded relatively little change in accuracy. This also was true for the Chicago datasets with respect to biologically significant errors. Of the three new models tested herein, the Decision Tree/Matrix Model produced the lowest biologically significant error rate; these rates were not reduced, however, from those produced by the Continental Decision Tree. Horticulturally limiting errors decreased slightly more for both Chicago datasets than they did for the Iowa data. On average, errors for Chicago A and Chicago B decreased by 2.3%, similar to the 0.5% decrease for Iowa (44). The species misclassified in both the Continental Decision Tree and the Decision Tree/Matrix Model are almost identical (Table 3), suggesting that the increased classification rates

were primarily responsible for the decreased error rates in the Chicago datasets. Overall, our analysis confirms that the Decision Tree/Matrix Model from the Iowa study improves classification rates over the Continental Decision Tree but offers little in the way of improved accuracy.

CART model. This five-step decision tree was developed independently of the Continental Decision Tree and relies on only three characteristics: quick vegetative spread, fleshy, bird-dispersed fruits, and G-values. This model was very effective in the Iowa study (44), but was not subject to external validation prior to our study. Although this model is much simpler than the others, its classification rates for the Chicago A and Chicago B datasets were only slightly lower than those obtained by the Continental Decision Tree (Table 2), and these differences were not statistically significant ($p = 0.89$ for Chicago A, $p = 0.54$ for Chicago B). As noted in the Iowa study (44), encouragingly, it also produced significantly lower rates of horticulturally limiting errors than did any other model. Chicago A had 9 species (9.5%) and Chicago B had only 5 species (4.9%) that resulted in horticulturally limiting errors. On average, this represents a 15.1% decrease in the error rates generated by the Continental Decision Tree ($p = 0.027$ for Chicago A, $p = 0.037$ for Chicago B); greater than the change reported for the Iowa data (13.2%, $p = 0.023$).

However, biologically significant error rates for the CART Model were among the highest of all models evaluated in our tests (Table 2), which was unexpected since it had the lowest such error rate when evaluated on the Iowa dataset that was used to develop it (44). The actual number of species misclassified, a total of 11 (Table 3), was actually one less than the total number of species that produced biologically significant errors from the Modified Decision Tree, but lower classification rates in the CART Model resulted in a proportionately higher error rate. Therefore, we cannot recommend the CART Model *per se* for use in the Chicago region given its rather low classification rates and higher

rates of biologically significant errors. However, it has fewer steps and is easier to apply than the Modified Decision Tree or the Decision Tree/Matrix Model, suggesting it might be worthy of further refinement or use in conjunction with another model that ultimately would result in fewer biologically significant errors.

Geographic-risk value comparison. Differences in geographic-risk values between the Iowa data and the Chicago datasets could be one source of decreased power and/or accuracy in our evaluation. However, there are strong, highly significant correlations in G-values between each dataset (Table 4), suggesting that these G-values are consistent enough among the datasets that the models that use them should not be weakened on this basis.

Although the correlations are strong (Table 4), they are not perfect. Interesting, but subtle differences in correlations between the three pairwise comparisons were discovered. As predicted based on relative differences in climatic and edaphic conditions, Chicago A was most similar to Iowa with a correlation of 0.97 (Table 4). Chicago B and Iowa had a lower correspondence with a correlation of 0.88. These two correlations are significantly different ($p < 0.0001$), implying there are greater regional differences between Iowa and Chicago B than between Iowa and Chicago A.

Factors that contribute to deviations from a perfect correlation between sets of G-values can be studied by examining the ‘outlier’ species with G-values that fall farthest from the linear relationship. In general, these outliers were more likely to have a small ‘n’ in terms of the number of geographic subdivisions where each species was native. Of the seven most extreme outliers, six species had ranges of 15 or fewer geographic subdivisions, while the average species in this study had a native range with a mean of 35 geographic subdivisions in the Chicago A dataset and a mean of 38 subdivisions in the Chicago B dataset. These species with limited geographic ranges displayed more highly divergent G-values than those with broader ranges, which would tend to diminish the effects of any local differences in the underlying P-values among datasets.

Comparative model performance. Since Chicago A and Iowa are more analogous to each other in terms of environmental conditions and in G-values, one might also expect to see this reflected in results from the four risk-assessment models. However, there were no consistent differences in the power and accuracy of models developed for Iowa when applied to the Chicago A dataset than when applied to the Chicago B dataset (Table 2). In five of six cases, error rates for the Iowa models applied to the Chicago B dataset were actually lower than for Chicago A. These inconsistencies lead us to reject our geographic hypothesis when applied to model performance, where we expected our tests of the

Chicago A dataset to more closely resemble past performance in Iowa (44). We recognize that many factors in addition to G-values can affect model outcomes, which complicates the determination of true confounding factors.

Of the models we used to evaluate non-native woody species in the Chicago region, the Decision Tree/Matrix Model gave the most favorable results, with a high classification rate (a result analogous to the Iowa data) and a slight decrease in the horticulturally limiting error rate from that of the Continental Decision Tree without increasing the biologically significant error rate. Although its results were not as consistently positive, the CART model presented advantages in its simplicity and low rates of horticulturally limiting errors, making it a promising model for further refinement, with a focus on ways to reduce its relatively high rate of biologically significant errors.

Another variation on models developed by Widrechner et al. (44) from Iowa data was evaluated by Jefferson et al. (16) at the Chicago Botanic Garden alongside the Australian Weed Risk-Assessment (WRA) protocol. They reported their variation on the Iowa model was 80–100% accurate when species placed in the ‘further analysis’ category were assessed correctly (16). In some circumstances, lower classification rates may not be a significant impediment, especially when those using a model have the capacity to conduct the further analyses needed. However, such testing is typically time-consuming, expensive, and as yet, poorly defined.

To provide a larger context, the classification and error rates presented in our study can also be compared to the classification and error rates produced by the Australian WRA in a larger meta-analysis presented by Gordon et al. (12). In seven geographically diverse studies, 1183 species tests were conducted with the WRA resulting in a mean classification rate of 83.1% (range 71 to 92%). Of the 983 species that were accepted or rejected by the WRA, the mean false negative rate (analogous to our biologically significant error rate) was 7.0% (range 0 to 12.7%), and the mean false positive rate (analogous to our horticulturally limiting error rate) was 4.2% (range 1.9 to 10.5%). Our results generally fall within the ranges of WRA classification rates and false negative rates, but only the CART model has comparable false positive rates.

With these considerations in mind, we intend to assemble and evaluate additional datasets from the Midwestern United States to refine and validate existing risk-assessment models for non-native woody plants and to define new ones, with the assistance of a new statistical approach based on Random Forest modeling (2). Our ultimate goal is to produce regional models that are more accurate, powerful, and easy to use than those currently available. However, it is important to remember that not all species that will naturalize in a region over time have already done so. On average, woody species in Brandenberg, Germany, had been cultivated for 147 years before becoming naturalized (20). Thus, some of the reported horticulturally limiting errors that models generate may actually portend future naturalization events. In contrast, some species that were included in our study as having naturalized may only do so under limited conditions that do not lead to invasions or other serious threats. Future analyses should incorporate the degree of invasiveness, as discussed by Gordon et al. (12), to determine how the most invasive species differ from those with well-documented records of limited naturalization without invasion.

Table 4. Correlations comparing geographic risk values of common species in the Chicago and Iowa datasets.

Comparison	Species in common	Correlation	Fisher Z_r ± s.e.
Chicago A to Chicago B	84	0.92	1.58 ± 0.11
Chicago A to Iowa	53	0.97	2.17 ± 0.14
Chicago B to Iowa	63	0.88	1.36 ± 0.13

Literature Cited

1. Bell, C.E., C.A. Wilen, and A.E. Stanton. 2003. Invasive plants of horticultural origin. *HortScience* 38:14–16.
2. Breiman, L. 2001. Random forests. *Machine Learning* 45:5–32.
3. Byers B.E., S. Reichard, J.M. Randall, I.M. Parker, C.S. Smith, W.M. Lonsdale, I.A.E. Atkinson, T.R. Seastedt, M. Williamson, E. Chornesky, and D. Hayes. 2002. Directing research to reduce the impacts of nonindigenous species. *Conservation Biology* 16:630–640.
4. Caley, P., W. Lonsdale, and P. Pheloung. 2006. Quantifying uncertainty in predictions of invasiveness. *Biological Invasions* 8:277–286.
5. Daehler, C.C., J.S. Denslow, S. Ansari, and H. Kuo. 2004. A risk assessment system for screening out invasive pest plants from Hawaii and other Pacific islands. *Conservation Biology* 18:360–368.
6. Deam, C.C. 1940. *Flora of Indiana*. Indiana Department of Conservation, Indianapolis.
7. eFloras 2008. *Flora of China* online volumes. Accessed August 2008. http://www.efloras.org/flora_page.aspx?flora_id=2
8. Fleiss, J.L. 1981. *Statistical Methods for Rates and Proportions*, 2nd ed. Wiley, New York.
9. *Flora of North America* Editorial Committee. 1993+. *Flora of North America North of Mexico*. Oxford University Press, New York.
10. *Flora of Taiwan* Editorial Committee. 1975–1979. *Flora of Taiwan*, 6 vols. Epoch Publ., Taipei.
11. Frappier, B. and R.T. Eckert. 2003. Utilizing the USDA PLANTS database to predict exotic woody plant invasiveness in New Hampshire. *Forest Ecology & Management* 185:207–215.
12. Gordon, D.R., D.A. Onderdonk, A.M. Fox, and R.K. Stocker. 2008a. Consistent accuracy of the Australian weed risk assessment system across varied geographies. *Diversity and Distributions* 14:234–242.
13. Gordon, D.R., D.A. Onderdonk, A.M. Fox, and R.K. Stocker. 2008b. Predicting invasive plants in Florida using the Australian weed risk assessment. *Invasive Plant Science and Management* 1:178–195.
14. Harrell, F.E., Jr. 2001. *Regression Modeling Strategies with Applications to Linear Models, Logistic Regression, and Survival Analysis*. Springer-Verlag, New York.
15. Harrington, R., R. Kujawski, and H. Ryan. 2003. Invasive plants and the green industry. *J. Arboriculture* 29:42–48.
16. Jefferson, L., K. Havens, and J. Ault. 2004. Implementing invasive screening procedures: The Chicago Botanic Garden model. *Weed Technology* 18:1434–1440.
17. Kato, H., K. Hata, H. Yamamoto, and T. Yoshioka. 2006. Effectiveness of the weed risk assessment system for the Bonin Islands. p. 65–72. *In: Assessment and Control of Biological Invasion Risk*, F. Koike, M. Clout, M. Kawamichi, M. De Poorter, and K. Iwatsuki, eds., Kyoto, Japan: Shoukadoh Book Sellers.
18. Keller R.P., D.M. Lodge, and D.C. Finnoff. 2007. Risk assessment for invasive species produces net bioeconomic benefits. *Proc. National Academy of Sciences (USA)* 104:203–207.
19. Komarov, V.L. et al. (eds.) 1934–1964. *Flora of the USSR*, 30 vols. Academy of Sciences of the USSR, Moscow and Leningrad.
20. Kowarik, I. 1995. Time lags in biological invasions with regard to the success and failure of alien species. p. 15–38. *In: Plant Invasions: General Aspects and Special Problems*, P. Pyšek, K. Prach, M. Rejmánek, and M. Wade, eds., Amsterdam: SPB Academic Publishing.
21. Křivánek, M. and P. Pyšek. 2006. Predicting invasions by woody species in a temperate zone: A test of three risk assessment schemes in the Czech Republic (Central Europe). *Diversity and Distributions* 12:319–327.
22. Mack, R.N. and M. Erneberg. 2002. The United States naturalized flora: Largely the product of deliberate introductions. *Annals of the Missouri Botanical Garden* 89:176–189.
23. Mack, R.N., D. Simberloff, W. Lonsdale, H. Evans, M. Clout, and F. Bazzaz. 2000. Biotic invasions: Causes, epidemiology, global consequences, and control. *Ecological Applications* 10:689–710.
24. Mosyakin, S.L. and M.M. Fedoronchuk. 1999. *Vascular Plants of Ukraine: A Nomenclatural Checklist*. M.G. Kholodny Institute of Botany, Kiev.
25. Ohwi, J. 1965. *Flora of Japan* (in English). Smithsonian Institution, Washington, DC.
26. Peterson, A.T. 2003. Predicting the geography of species' invasions via ecological niche modeling. *Quarterly Review of Biology* 78:419–433.
27. Pheloung, P.C., P.A. Williams, and S.R. Halloy. 1999. A weed risk assessment model for use as a biosecurity tool for evaluating plant introductions. *J. Environ. Management* 57:239–251.
28. Pimentel, D., R. Zuniga, and D. Morrison. 2005. Update on the environmental and economic costs associated with alien-invasive species in the United States. *Ecological Econ.* 52:273–288.
29. Reichard, S.H. and C.W. Hamilton. 1997. Predicting invasions of woody plants introduced into North America. *Conservation Biology* 11:193–203.
30. Smith, C., W. Lonsdale, and J. Fortune. 1999. When to ignore advice: Invasion predictions and decision theory. *Biological Invasions* 1:89–96.
31. Snedecor, G.W. and W.G. Cochran. 1989. *Statistical Methods*, 8th ed. Iowa State University Press, Ames, IA.
32. Swink, F. and G. Wilhelm. 1994. *Plants of the Chicago Region*, 4th ed. Indiana Academy of Science, Indianapolis.
33. Thompson, R.S., K.H. Anderson, and P.J. Bartlein. 2000. *Atlas of relations between climatic parameters and distributions of important trees and shrubs in North America*. 2 vols. USGS Professional Paper 1650.
34. Tutin, T.G. et al. (eds.) 1964–1994. *Flora Europaea*, 5 vols. Cambridge University Press, Cambridge.
35. U.S. Department of Agriculture, Agricultural Research Service. 2008. *Germplasm Resources Information Network* database. Accessed August 2008. <http://www.ars-grin.gov/npgs/>
36. U.S. Department of Agriculture, Natural Resources Conservation Service. 2008. *PLANTS* database. Accessed August 2008. <http://plants.usda.gov/>
37. U.S. Department of Agriculture, Natural Resources Conservation Service. 2008. *Tables of Acreage and Proportionate Extent of the Soils from the Soil Survey Geographic (SSURGO) database*. USDA-NRCS, Fort Worth, Texas. Accessed March 2009. <http://SoilDataMart.nrcs.usda.gov/>
38. U.S. Environmental Data Service. 1968. *Climatic Atlas of the United States*. US Department of Commerce, Washington, DC. 80 pp.
39. Voss, E.G. 1972–1996. *Michigan Flora: A Guide to the Identification and Occurrence of the Native and Naturalized Seed-plants of the State*, 3 vols. Cranbrook Institute of Science and University of Michigan Herbarium, Bloomfield Hills, Michigan.
40. Widrlechner, M.P. 1994. Environmental analogs in the search for stress-tolerant landscape plants. *J. Arboriculture* 20:114–119.
41. Widrlechner, M.P. 1999. A zone map for mean annual moisture balance in the North Central United States. *Landscape Plant News* 10(2):10–14.
42. Widrlechner, M.P. 2001. The role of environmental analogs in identifying potentially invasive woody plants in Iowa. *J. Iowa Acad. Sci.* 108:158–165.
43. Widrlechner, M.P. and J.K. Iles. 2002. A geographic assessment of the risk of naturalization of non-native woody plants in Iowa. *J. Environ. Hort.* 20:47–56.
44. Widrlechner, M.P., J.R. Thompson, J.K. Iles, and P.M. Dixon. 2004. Models for predicting the risk of naturalization of non-native woody plants in Iowa. *J. Environ. Hort.* 22:23–31.