Models for Predicting the Risk of Naturalization of Non-Native Woody Plants in Iowa[1]

Mark P. Widrlechner[2], Janette R. Thompson[3], Jeffery K. Iles[4], and Philip M. Dixon[5]
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

Nursery and landscape professionals have introduced many useful non-native woody plants for managed landscapes, but the potential exists for new introductions to escape from cultivation and become pests. The objective of this study was to develop a comprehensive strategy to assess the risk of naturalization of non-native woody plants in Iowa. We examined life-history traits and native distributions of 100 woody plant species, including 28 species known to naturalize in Iowa and 72 other species not known to naturalize in the state. We tested three approaches to predict naturalization of woody plants in Iowa: (1) direct application of a previously developed decision tree designed to predict naturalization on a continental scale within North America; (2) application of the continental model modified to include traits important on a regional scale, and (3) development of a new regional model that included a geographic-risk component along with life-history traits. Our modifications to the continental model and the new regional model all were more powerful than the unmodified, continental model, as measured by their ability to classify species (classification rate) based on risk of naturalization, without reducing accuracy, as measured by the frequency of misclassification (error rate). Classification rates ranged from 65% for the unmodified continental model to 90% for the modified continental model. Error rates varied from 6% for the new regional model to 20% for the unmodified continental model, including horticulturally limiting error rates of 4 to 17%, respectively.

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

Significance to the Nursery Industry

The prospect of introduced plants escaping cultivation to become invasive pests causes concern among proponents and managers of natural ecosystems, nursery and landscape professionals, policy makers and interested citizens. Although only a small proportion of introduced species escape cultivation and even fewer become pests, efforts to control invasions are often difficult and costly. Quarantine programs to identify and exclude potentially invasive species are potentially useful, although lengthy testing periods may prove costly to enterprising nursery operators expecting quick returns from investments in new tree or shrub introductions. Alternatively, predictive models that assess the risk of naturalization provide rapid and affordable information, but an inherent reliance on plant attributes (life-history traits, winter hardiness, etc.) predisposes them to classification errors, specifically, biologically significant, false-negative errors (prediction of no invasive potential for a species that can invade) and horticulturally limiting, false-positive errors (prediction of invasiveness for species that do not escape cultivation). This paper reports on new, predictive models which limit both types of error to the greatest extent possible. Comprehensive regional models that integrate woody plant life-history traits with climatic/geographic risk analysis show promise for increasing the power of prediction while respecting both biological and horticultural concerns. But even the best models should be augmented and revalidated by long-term monitoring of natural areas and other sites that could offer a foothold for newly naturalizing woody plants.

Introduction

Historically, large segments within the agriculture, forestry, horticulture, medicinal plant, seed trade, and companion animal industries, as well as conservation, and fish and game agencies have promoted the importation and dissemination of useful, non-native species (13, 19, 29). And while only a very small number of introduced species become naturalized, and fewer still become invasive, controlling the spread of invasive species can be enormously difficult (24). Given the potential for damage to natural and cultural systems, and the prohibitive cost associated with control efforts, research focus has intensified on developing predictive models to assess risks associated with introducing exotic plant species (11, 21–23, 25, 28, 32). With the rate of both accidental and purposeful introductions accelerating, biotic invasions have joined atmospheric and land-use change impacts as major agents of human-driven global change (19).

Both correlative and experimental approaches have been used to develop risk-assessment models (18, 30). Correlative work has focused on retrospective lists of known invasive and non-invasive species, identifying a number of biological and/or ecological life-history traits, and determining which of those are most strongly associated with invasiveness (11, 21, 28, 32). Traits common to several of these analyses include: reproductive system characteristics, seed crop and seed germination factors, life form (such as tree, shrub, vine, annual herb, etc.), length of time since introduction, invasiveness elsewhere, and location of native range. Critics of this approach warn of

[1]Received for publication September 9, 2003; in revised form December 26, 2003. 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 Peter W. Bristol, Candice A.C. Gardner, Lara Jefferson, J. Michael Kelly, and Marcel Rejmánek for their valuable critiques. Mention of commercial brand names does not constitute an endorsement of any product by the U.S. Department of Agriculture or cooperating agencies.

[2]Horticulturist.

[3]Assistant Professor, Department of Natural Resource Ecology and Management, Iowa State University.

[4]Associate Professor and Chair, Department of Horticulture, Iowa State University.

[5]Professor, Department of Statistics, Iowa State University.

the possibility of coincidental occurrence of traits, and that 'guilt by association' may produce spurious results (19). In addition, lack of knowledge about which species truly are 'non-invasive,' especially for woody plants, for which lag time between introduction and escape can be on the order of 100 or more years, is problematic (16).

Other work has focused on geographic and climatic analyses, examining the relative importance of climatic matches between native ranges and the site of introduction (17, 39, 42) and amplitude of native geographic ranges to potential invasiveness (11). Critics of this approach point out that while habitat compatibility is important, the species in its primary range may have limitations imposed by biotic factors (e.g. herbivory, niche competition, disease) rather than climatic ones (31). It has been demonstrated, however, that climatic factors are important determinants of non-native woody plant survival in the North Central United States (40, 41).

Statistical methods used to develop the risk-assessment tools described above include discriminant analyses (28, 32), cluster analyses (21), multiple logistic regression procedures (11, 25), and classification and regression trees (21, 28). Evaluation of the predictive ability of risk-assessment models has generally included validation via application to a subset of species not included in the original discriminant analysis or regression procedure. Investigators have typically reported overall classification rates, and success rates associated with prediction of invasive species as well as non-invasive species. Unfortunately, accurate predictions are often commingled with spurious results including 'false positives' (prediction of invasiveness for species not known to be invasive, also called 'horticulturally limiting error') and 'false negatives' (prediction of non-invasiveness for species known to be invasive, also called 'biologically significant error').

The most commonly cited predictive model for woody plant evaluation is based primarily on correlative analyses and was developed by Reichard and Hamilton (28). Although validation of the underlying discriminant analysis and classification and regression tree (CART) methods was described, the ultimate 'decision tree' these researchers proposed for use in screening woody plant introductions was not accompanied by a detailed description of its validation. Until now, independent validation (by scientists other than those developing the models) of any of these risk-assessment protocols described herein has not been documented in the scientific literature. Although models tested outside their intended geographic targets cannot be considered as statistically validated, the best example of an independent test of previously published risk-assessment protocols has been conducted for the non-native flora of Hawaii by Daehler and Carino (4). They reported results of the application of three models developed for other parts of the world: North America (28), the South African fynbos (36), and Australia (23).

In addition to variation in the general approaches and statistical methods used to develop predictive models, the scope of the models also has been variable, ranging from continental (23, 28), to national (21, 23), to those with a regional scope within nations (11, 36, 39, 42). Although a continental or national approach is appealing from a regulatory standpoint, regional analyses are likely to be more directly applicable, less exclusionary, and more acceptable to the industries that may be affected (10, 13, 14).

The objective of this study was to develop a comprehensive strategy to assess the risk of naturalization of non-native woody plants in Iowa. We compared three possible approaches to predict naturalization by using an extensive set of non-native woody species that occur in Iowa, including both those that are known to naturalize and those with no record of naturalization in the state. We tested three hypotheses: (1) Predictive models developed with a continental scope and based exclusively on correlative analyses would have less power and accuracy when applied to a regional flora; (2) Incorporating locally important, life-history variables in a continental model would improve the classification rate; and (3) A model that included both geographic and species-based risk components would be more powerful for regional risk assessment.

Materials and Methods

A set of 100 species of woody landscape plants not known to be native to Iowa in the period predating European settlement in the mid-1800s was selected for the purpose of constructing risk-assessment models. This set included 28 species with a history of naturalizing in Iowa (39) (Table 1) and 72 with no record of naturalizing in Iowa (42) (Table 2) and emphasized species with well-characterized natural distributions and relatively long histories of cultivation in the state.

A spreadsheet (summarized in Tables 1 and 2) was compiled containing information on life-history and invasive characteristics required by Reichard and Hamilton's (28) decision tree, as illustrated in Fig. 1. The life-history characteristics are generally connected with ease and speed of vegetative and/or sexual reproduction, and the invasive characteristics focus upon what is known about the spread of the species in other parts of the world and of the invasiveness of its relatives in North America. These data were gleaned from several reports (6, 7, 26, 27, 37, 43) and supplemented with the three senior authors' collective experience observing and cultivating these plants.

Information on two additional life-history characteristics suspected of influencing the ability of woody plants to spread from cultivation, fleshy, bird-dispersed fruits (9) and evergreen foliage (28), also was included.

Another spreadsheet was previously created for all 100 species containing detailed information about their native ranges (see 39, 42) divided into approximately 300 geographic and political subdivisions. For each of these ca. 300 geographical subdivisions, we calculated, P, the ratio of the number of species native to that subdivision known to naturalize in Iowa, N_{nat}, to the total number of species native to that subdivision in the entire 100-species set, N. Thus, $P = N_{nat} / N$. This ratio corresponds to the proportion of naturalizing species as mapped by Widrlechner and Iles (42). A range-wide, geographic-risk value, G, was then calculated for each species (see Tables 1 and 2) by averaging unweighted ratios for all geographical subdivisions included in that species' native range, where

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

and n is the number of geographical subdivisions included in the species' native range.

Once these data were assembled, we subjected all 100 species to a risk assessment with Reichard and Hamilton's (28) decision tree (Fig. 1). The decision tree assigned each spe-

Table 1. Characteristics of 28, non-native woody landscape plants that have naturalized in Iowa used to develop models to assess the risk of naturalization in Iowa.

Species	G^z	Invades outside North America	Requires germination pretreatment	Group invasive in North America	Quick maturity	Sterile hybrid	Quick vegetative spread	North American native	Evergreen foliage	Fleshy, bird dispersed fruits
Acer tataricum L. subsp. *ginnala* (Maxim.) Wesm.	0.433	No	Yes	No	Yes	No	No	No	No	No
Ailanthus altissima (Mill.) Swingle	0.295	Yes	No	No	Yes	No	Yes	No	No	No
Berberis thunbergii DC	0.347	No	Yes	No	No	No	No	No	No	Yes
Campsis radicans (L.) Seem ex Bureau	0.153	Yes	No	No	Yes	No	Yes	Yes	No	No
Catalpa speciosa (Warder ex Barney) Warder ex Engelm.	0.188	No	No	No	Yes	No	No	Yes	No	No
Elaeagnus angustifolia L.	0.542	Yes	Yes	No	Yes	No	No	No	No	Yes
Elaeagnus umbellata Thunb.	0.350	No	Yes	No	Yes	No	No	No	No	Yes
Euonymus alatus (Thunb.) Siebold	0.382	No	Yes	yes/no	Maybe	No	No	No	No	Yes
Lonicera maackii (Rupr.) Maxim.	0.487	No	Yes	Yes	Yes	No	No	No	No	Yes
Lonicera tatarica L.	0.787	Yes	Yes	Yes	Yes	No	No	No	No	Yes
Lycium barbarum L.	0.503	Yes	No	No	Yes	No	Yes	No	No	Yes
Maclura pomifera (Raf.) C.K. Schneid.	0.341	Yes	Maybe	No	No	No	No	Yes	No	No
Malus sylvestris Mill.	0.510	Yes	Yes	Yes	Maybe	No	No	No	No	Yes
Morus alba L.	0.403	Yes	No	No	Yes	No	No	No	No	Yes
Populus alba L.	0.549	Yes	No	Yes	Yes	No	Yes	No	No	No
Prunus tomentosa Thunb.	0.443	No	Yes	Yes	Yes	No	No	No	No	Yes
Rhamnus cathartica L.	0.520	No	Yes	No	Yes	No	No	No	No	Yes
Robinia pseudoacacia L.	0.166	Yes	Yes	Yes	Yes	No	Yes	Yes	No	No
Rosa multiflora Thunb.	0.350	No	Yes	Yes	Yes	No	No	No	No	Yes
Rosa rubinigosa L.	0.470	Yes	Yes	Yes	Yes	No	No	No	No	Yes
Rubus parvifolius L.	0.461	No	Yes	Yes	Yes	No	Yes	No	No	Yes
Salix ×rubens Schrank	0.503	Yes	Sterile	Yes	Yes	Yes	Yes	No	No	No
Salix alba L.	0.561	Yes	No	Yes	Yes	No	Yes	No	No	No
Salix fragilis L.	0.498	Yes	No	Yes	Yes	No	Yes	No	No	No
Sorbus aucuparia L.	0.500	Yes	Yes	Yes	No	No	No	No	No	Yes
Ulmus pumila L.	0.542	No	No	No	Yes	No	No	No	No	No
Viburnum lantana L.	0.451	No	Yes	Yes	No	No	No	No	No	Yes
Viburnum opulus L. var. *opulus*	0.511	No	Yes	Yes	No	No	No	No	No	Yes

[z]G, the range-wide, geographic-risk value, is a ratio that can vary between 0 and 1 (see Materials and Methods).

Table 2. Characteristics of 72, non-native woody landscape plants that are not known to naturalize in Iowa used to develop models to assess the risk of naturalization in Iowa.

Species	G[z]	Invades outside North America	Requires germination pretreatment	Group invasive in North America	Quick maturity	Sterile hybrid	Quick vegetative spread	North American native	Evergreen foliage	Fleshy, bird dispersed fruits
Abies concolor (Gordon & Glend.) Lindl. ex F. H. Hildebr.	0.000	No	Maybe	No	No	No	No	Yes	Yes	No
Acer campestre L.	0.443	No	Yes	No	No	No	No	No	No	No
Aronia arbutifolia (L.) Pers.	0.082	No	Yes	Yes	Yes	No	Yes	Yes	No	Yes
Berberis koreana Palibin	0.368	No	Yes	No	Yes	No	Yes	No	No	Yes
Betula platyphylla Sukaczev	0.421	No	Yes	No	No	No	No	No	No	No
Betula populifolia Marshall	0.025	No	Yes	No	No	No	No	Yes	No	No
Buxus microphylla Siebold & Zucc.	0.288	No	Yes	No	No	No	No	No	Yes	No
Caragana frutex (L.) K. Koch	0.566	No	Yes	Yes	No	No	Yes	No	No	No
Carpinus betulus L.	0.445	No	Yes	No	No	No	No	No	No	No
Castanea mollissima Blume	0.343	No	Yes	No	No	No	No	No	No	No
Cercidiphyllum japonicum Siebold & Zucc.	0.269	No	Yes	No	No	No	No	No	No	No
Chaenomeles speciosa (Sweet) Nakai	0.240	No	Yes	Yes	No	No	No	No	No	Maybe
Cladrastis lutea (F. Michx.) K. Koch	0.150	No	Yes	Yes	No	No	No	Yes	No	No
Clematis viticella L.	0.391	Yes	Yes	No	Yes	No	Yes	No	No	No
Clethra alnifolia L.	0.067	No	No	No	No	No	No	Yes	No	No
Cornus florida L.	0.079	No	Yes	No	No	No	No	Yes	No	Yes
Cornus mas L.	0.431	No	Yes	No	No	No	No	No	No	Yes
Corylus colurna L.	0.302	No	Yes	No	No	No	No	No	No	No
Cotinus coggygria Scop.	0.357	No	Yes	No	No	No	No	No	No	No
Cotoneaster apiculatus Rehder & E. H. Wilson	0.241	No	Yes	Yes	No	No	No	No	No	Yes
Cotoneaster dammeri C. K. Schneid.	0.221	No	Yes	Yes	No	No	No	No	No	Yes
Cotoneaster lucidus Schltdl.	0.571	No	Yes	Yes	No	No	No	No	No	Yes
Crataegus phaenopyrum (L. f.) Medik.	0.123	No	Yes	Yes	No	No	No	Yes	No	Yes
Deutzia crenata Siebold & Zucc.	0.344	Yes	No	No	No	No	No	No	No	No
Deutzia gracilis Siebold & Zucc.	0.334	No	No	No	No	No	No	No	No	No
Euonymus bungeanus Maxim.	0.353	No	Yes	yes/no	Maybe	No	No	No	No	Yes
Fagus sylvatica L.	0.443	No	Yes	No	No	No	No	No	No	No
Hamamelis vernalis Sarg.	0.204	No	Yes	No	No	No	No	Yes	No	No
Hibiscus syriacus L.	0.290	Yes	No	No	Yes	No	No	No	No	No
Hydrangea arborescens L.	0.096	No	No	No	No	No	Yes	Yes	No	No
Hydrangea paniculata Siebold	0.290	No	No	No	No	No	No	No	No	No
Juniperus chinensis L.	0.316	No	Yes	No	No	No	No	No	Yes	Yes
Juniperus sabina L.	0.431	No	Yes	No	No	No	Yes	No	Yes	Yes
Juniperus scopulorum Sarg.	0.000	No	Yes	No	No	No	No	Yes	Yes	Yes
Koelreuteria paniculata Laxm.	0.323	Yes	Yes	No	No	No	No	No	No	No
Kolkwitzia amabilis Graebn.	0.288	No	No	Yes	No	No	No	No	No	No
Larix decidua Mill.	0.433	No	No	No	No	No	No	No	No	No
Liquidambar styraciflua L.	0.097	No	Maybe	No	No	No	No	Yes	No	No
Magnolia stellata (Siebold & Zucc.) Maxim.	0.353	No	Yes	No	No	No	No	No	No	Maybe
Microbiota decussata Kom.	0.500	No	Yes	No	No	No	No	No	Yes	No
Myrica pensylvanica Mirb.	0.021	Yes	Yes	No	No	No	Yes	Yes	Semi	Yes
Nyssa sylvatica Marshall	0.077	No	Yes	No	No	No	No	Yes	No	Yes
Philadelphus coronarius L.	0.394	No	Yes	No	No	No	No	No	No	No
Picea glauca (Moench) Voss	0.000	No	No	No	No	No	No	Yes	Yes	No
Picea pungens Engelm.	0.000	No	No	No	No	No	No	Yes	Yes	No
Pinus mugo Turra	0.427	No	No	No	No	No	No	No	Yes	No
Pinus nigra J. F. Arnold	0.449	Yes	No	No	No	No	No	No	Yes	No
Pinus ponderosa P. Lawson & C. Lawson	0.000	Yes	No	No	No	No	No	Yes	Yes	No
Prunus avium (L.) L.	0.453	No	Yes	Yes	Yes	No	No	No	No	Yes

Prunus glandulosa Thunb.	0.265	No	Yes	Yes	No	No	No	No	No	Yes
Prunus maackii Rupr.	0.487	No	Yes	Yes	Yes	No	No	No	No	Yes
Pseudotsuga menziesii (Mirb.) Franco	0.000	Yes	No	No	No	No	No	Yes	Yes	No
Pyrus calleryana Decne.	0.294	Yes	Yes	Yes	No	No	No	No	No	Yes
Quercus robur L.	0.466	Yes	No	No	No	No	No	No	No	No
Rhodotypos scandens (Thunb.) Makino	0.318	Yes	Yes	Yes	No	No	No	No	No	No
Ribes alpinum L.	0.498	No	Yes	No	No	No	No	No	No	Yes
Spiraea japonica L. f.	0.292	Yes	No	Yes	Yes	No	No	No	No	No
Styphnolobium japonicum (L.) Schott	0.294	No	No	Yes	No	No	No	No	No	No
Syringa meyeri C. K. Schneid.	0.526	No	Yes	Yes	No	No	No	No	No	No
Syringa reticulata (Blume) H. Hara	0.431	No	Yes	Yes	No	No	No	No	No	No
Syringa villosa Vahl	0.473	No	Yes	Yes	No	No	No	No	No	No
Tamarix ramosissima Ledeb.	0.527	Yes	No	Yes	Yes	No	Yes	No	No	No
Taxus cuspidata Siebold & Zucc.	0.361	No	Yes	No	No	No	No	No	Yes	Yes
Tilia cordata Mill.	0.474	No	Yes	No	No	No	No	No	No	No
Tilia tomentosa Moench	0.407	No	Yes	No	No	No	No	No	No	No
Tsuga canadensis (L.) Carriere	0.045	No	Yes	No	No	No	Yes	Yes	Yes	No
Viburnum carlesii Hemsl.	0.362	No	Yes	Yes	No	No	No	No	No	Yes
Viburnum cassinoides L.	0.038	No	Yes	Yes	No	No	No	Yes	No	Yes
Viburnum rufidulum Raf.	0.179	No	Yes	Yes	No	No	No	Yes	No	Yes
Viburnum sieboldii Miq.	0.347	No	Yes	Yes	No	No	No	No	No	Yes
Vitis labrusca L.	0.060	Maybe	Yes	No	Maybe	No	Yes	Yes	No	Yes
Weigela florida (Bunge) A. DC.	0.403	No	No	Yes	No	No	No	No	No	No

[z]G, the range-wide, geographic-risk value, is a ratio that can vary between 0 and 1 (see Materials and Methods).

cies to one of three categories: reject, accept, or further analysis/monitoring needed. Reject indicated a high risk of invasiveness; accept indicated low-risk species; and the further analysis category was reserved for those species for which the decision tree failed to provide clear advice.

The power and accuracy of the decision tree and all subsequent models were assessed by the following methods. The power of the model to classify species was measured through Reichard and Hamilton's (28) classification rate: the proportion of 'classified' species, in other words, those assigned only to the reject or accept categories, excluding those assigned to the further analysis group. The accuracy of the decision tree was measured by calculating two error rates, one for biologically significant misclassification (false negatives) and the other for horticulturally limiting misclassification (false positives). For purposes of this study, all naturalization events are considered to be biologically significant in that they modify existing plant communities and have the potential to alter them greatly should the naturalized species become invasive. The biologically significant error rate was determined by calculating the ratio of species known to naturalize in Iowa, for which a model indicated acceptance, to all classified species. In that case, a model would indicate that a species should not naturalize but is already known to do so. Application of the model could have biologically damaging consequences if it led to the introduction of a new invasive species. The horticulturally limiting error rate was similarly determined by calculating the ratio of species not known to naturalize in Iowa, for which a model indicates rejection, to all classified species. This type of error indicates where a model would potentially limit horticultural diversity without any *a priori* evidence of naturalization to support it. Comparisons of classification and error rates were statistically tested by using the exact binomial test (33).

After subjecting our dataset to the decision tree (28) and determining its power and accuracy, dissatisfaction with both the power and accuracy of the model, led us to follow three general methods to develop better models.

Modified decision tree. First, we attempted a series of ad-hoc modifications to the structure of the original decision tree, focusing on branches that produced the most unclassified (further analysis) and/or misclassified species. Modifications incorporated information not used in the original decision tree, such as the production of fleshy, bird-dispersed fruits and the range-wide geographic risk value, G, to increase the applicability of the model to our dataset, specific to Iowa conditions. The resulting decision tree is referred to as the 'Modified Decision Tree.'

Decision tree/matrix model. Next, we developed a matrix to help classify species that the original decision tree assigned to the further analysis category. These unclassified species were grouped by their values of G into low, medium and high-risk categories. The low-risk category corresponds to values of G ≤ 0.28, the mean proportion of naturalizing species within the complete set. The medium-risk category corresponds to values of G, $0.28 < G \leq 0.42$, with the upper end set 50% higher than the mean proportion. The high-risk category corresponds to values of $G > 0.42$. For each member of the high-risk group, the importance of including information about fleshy, bird-dispersed fruits also was examined. The resulting decision tree is referred to as the 'Decision Tree/Matrix Model.'

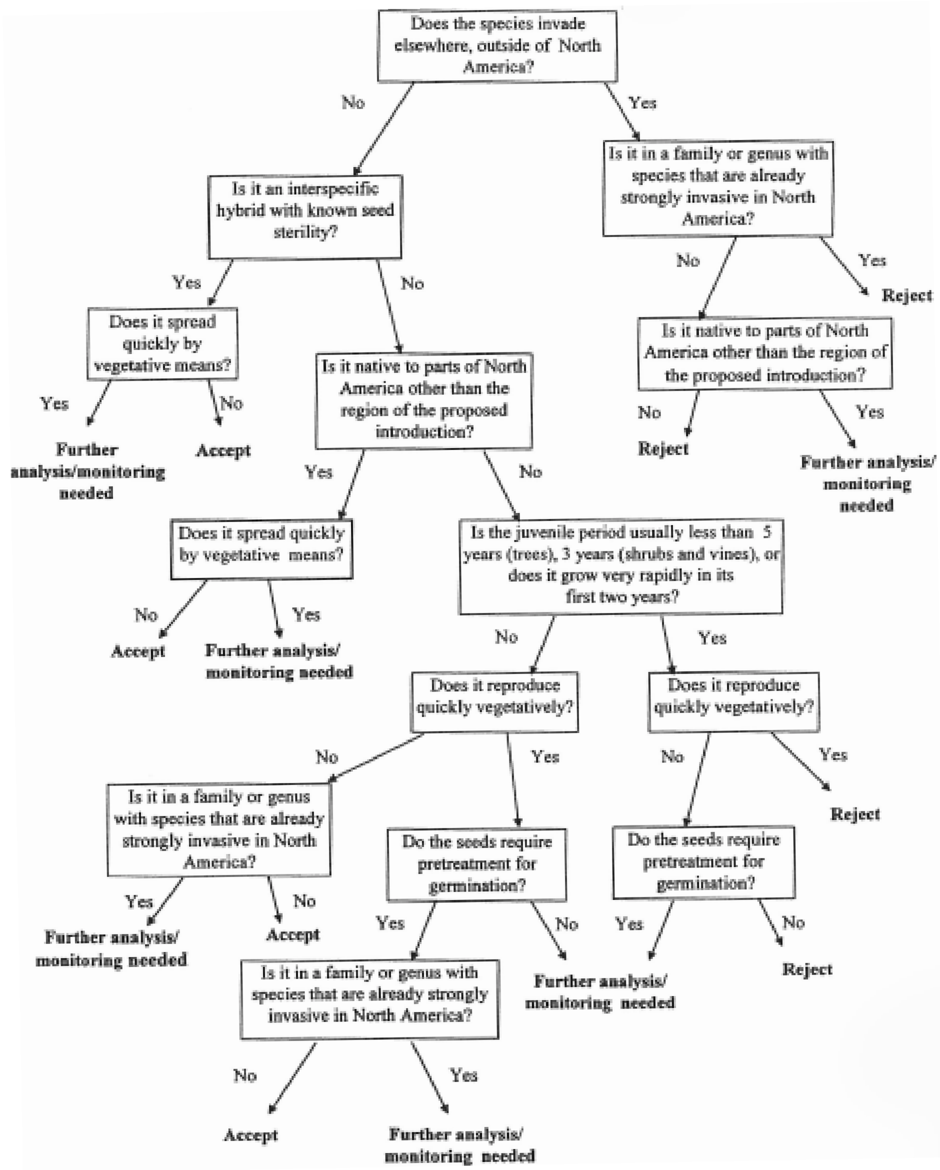

Fig. 1. Reichard and Hamilton's (28) decision tree.

New CART model. Finally, we constructed a new classification tree based on our entire dataset of variables required for the original decision tree, as well as the two additional life-history characteristics and values of G. A classification and regression tree (CART) results from a recursive partitioning of the data (3). The first step is based on the entire dataset. Among those 100 species, 28 have naturalized in Iowa. The CART algorithm searches all possible binary partitions of the data set based on one variable. For example, one possible partition is based on length of juvenile period. All species with short juvenile periods (noted as 'Quick Maturity' in Tables 1 and 2) are partitioned into one group. The remaining species form the second group. When the partition is useful, the two groups have different proportions of naturalizing species. The quality of the partition can be quantified using the change in deviance, a quantity related to the Chi-square statistic for a 2 × 2 contingency table. The optimal partition is the one with the largest change in deviance,

Table 3. Summary of classification and error rates for four risk-assessment models, as tested upon 100 non-native woody landscape plants in Iowa.

Model	Classification rate (%)	Biologically significant error rate (%)	Horticulturally limiting error rate (%)
Original Decision Tree	65	3.1	16.9
Modified Decision Tree	90	3.3	13.3
Decision Tree/Matrix Model	85	3.5	16.4
New CART Model	81	2.5	3.7

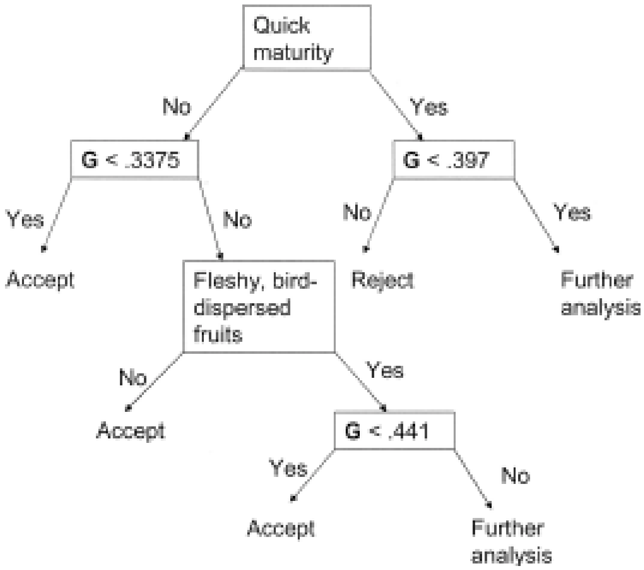

Fig. 2. **New CART Model derived by the tree function in S-Plus (20) from species characteristics presented in Tables 1 and 2.**

i.e. creating two groups with very different risk ratings. Each group from the first partition is partitioned repeatedly, until either the group is homogeneous (all species are naturalizing or not) or the group contains 5 or fewer species.

We used the tree function in S-Plus (20) to construct decision trees. Some partitions did not provide an unambiguous classification, e.g. a node with three naturalizing species and two that were not. We classified the species resulting from these partitions as requiring 'further analysis,' in a sense analogous to that used by Reichard and Hamilton (28). The resulting decision tree (Fig. 2) is referred to as the 'New CART Model.'

Results and Discussion

The set of 100 woody-plant species (Tables 1 and 2) was subjected to Reichard and Hamilton's (28) decision tree. The decision tree indicated that 39 of these species should be accepted for introduction, 26 should be rejected, and the remaining 35 species subjected to further analysis. The power of the model to classify species was indicated by a 65.0% classification rate (Table 2). This rate is significantly lower (exact binomial test, $p = 0.0007$) than the 80.0% classification rate calculated from data reported by Reichard and Hamilton (28), where 233 of 291 species were subjected to their decision tree as part of an internal validation procedure or the 86.0% classification rate calculated from 57 species evaluated by Daehler and Carino (4) for Hawaii.

Error rates for our 100-species set were 3.1% for biologically significant errors and 16.9% for horticulturally limiting errors (Table 3). Error rates were then calculated from Reichard and Hamilton's (28) validation data. Their results indicated rates of only 1.3% for biologically significant errors and 6.9% for horticulturally limiting errors. Similar results were obtained for this model by Daehler and Carino (4) for Hawaii, who reported no biologically significant errors and a horticulturally limiting error rate of only 6.1% (3/49). The biologically significant error rate for our data is based on a relatively small number of species, and there is no evidence that our rate differs from Reichard and Hamilton's (28). However, our rate of horticulturally limiting errors is significantly greater than that of Reichard and Hamilton (28) (exact binomial test, $p = 0.0094$).

Our application of the continental-scale decision tree to Iowa data resulted in much lower power and accuracy than that reported as part of that tree's internal validation, levels low enough for us to question the utility of the decision tree for Iowa conditions. These differences can be attributed to at least two distinct factors. First, the application of a decision tree designed to predict invasiveness on a continental scale to one small part of that continent would logically contribute to our higher horticulturally limiting error rate, since rejected species that have not naturalized in Iowa may be adapted to, and invasive in, other parts of North America. This effect is likely to be more pronounced in those parts of North America with particularly harsh climates, limiting the range of adapted woody plants. Second, differences between validation based on data used to develop that same model and validation based on novel data sets, such as ours, can contribute to observed reductions in both power and accuracy (12). Finally, we tested a smaller sample than that used by Reichard and Hamilton (28), which should not bias estimates of power and accuracy, but would make those estimates less precise. However, the exact binomial test (33) accounts for such differences in sample size.

Modified decision tree. Dissatisfaction with the power and accuracy of Reichard and Hamilton's (28) decision tree led to the development of a system to classify the 'further analysis' species and to an analysis of the sources of misclassification. We began by examining values of G for the 35 unclassified species to determine whether species that had naturalized in Iowa were from higher risk regions than those that had not. Twelve of the 35 unclassified species had values of $G < 0.28$, the mean proportion of naturalizing species within the complete set. Of those 12, 11 had not naturalized. A step was then added to the decision tree to test all 'further analysis' species for G, and 'accept' those with values < 0.28.

The remaining 23 unclassified species were then evaluated for fleshy, bird-dispersed fruits, based on the implication of that characteristic as a factor contributing to the invasiveness of woody plants in Iowa (9). Those species with fleshy, bird-dispersed fruits were placed in the 'reject' category, with the others remaining in the 'further analysis' group.

The first two refinements greatly increased the classification rate from 65 to 91%, but the horticulturally significant error rate was not decreased. The decision tree was then examined to determine whether there was a pattern in the sources of rejected species that had not naturalized in Iowa. Eleven of the 17 misclassified rejects came from the short, right branch of the decision tree (Fig. 1). Characteristics used in the right side of the tree focus on native range and invasion history, not specific biological attributes, and are oriented toward continental risk, not necessarily applicable to environmental conditions in Iowa. The tree was then modified to test the 24 rejected species from the right side of the tree for fleshy, bird-dispersed fruits. As done earlier, the species (9 of 24) with fleshy, bird-dispersed fruits were placed in the 'reject' category, with the other 15 classified as requiring further analysis. Finally, we tested three key biological attributes of the 15 species requiring further analysis by linking them into the left side of the decision tree beginning with evaluation of the length of the juvenile period.

The resulting Modified Decision Tree included 24 steps, with a much more complex pattern of questions than the original, 14-step tree, suggesting that its practical application may be limited. However, on the positive side, these changes did increase the classification rate from 65 to 90% (Table 3) and lower the horticulturally limiting error rate from 16.9 to 13.3% with little change in the biologically significant error rate (3.1 vs. 3.3%).

Decision tree/matrix model. Our second attempt to refine Reichard and Hamilton's (28) decision tree focused solely on classification of the 35 species that had been assigned to the 'further analysis' category. It applied a matrix approach, first categorizing the 35 species into three geographic-risk groups based on G values (low: $G \leq 0.28$, medium: $0.28 < G \leq 0.42$, high: $G > 0.42$). Because 11 of 12 members of the low-risk group have not naturalized in Iowa (as noted above), those 12 were assigned to the 'accept' category. The 10 members of the medium-risk group were retained in the 'further analysis' category, and the 13 members of the high-risk group were subjected to an additional test, for fleshy, bird-dispersed fruits. There were five species that did not have fleshy, bird-dispersed fruits, and four of them have not naturalized. The group of five was retained in the 'further analysis' category. Five of the remaining eight species with fleshy, bird-dispersed fruits have naturalized in Iowa. The group of eight was then assigned to the 'reject' category.

The Decision Tree/Matrix Model uses all of the same characters as the Modified Decision Tree but is much simpler in its design and use, adding only two steps to the original tree. In comparison to results from the original tree, the classification rate increased from 65 to 85% (Table 3) with only minor changes in accuracy. The horticulturally limiting error rate decreased from 16.9 to 16.4%, and the biologically significant error rate increased from 3.1 to 3.5% (Table 3).

New CART model. The application of CART to our entire data set resulted in a very simple five-step tree (Fig. 2). The branching points of the New CART Model rely upon only three characteristics: quick vegetative spread, fleshy, bird-dispersed fruits, and G values. Again, in relation to results from the original Reichard and Hamilton (28) tree, the classification rate increased from 65 to 81% (Table 3), somewhat less than improvements made with the first two modifications, but both error rates decreased. The horticulturally limiting error rate decreased from 16.9 to 3.7% (much more than for the other models), and the biologically significant error rate decreased from 3.1 to 2.5% (Table 3).

Assessment of misclassification. Each of the four risk-assessment models resulted in misclassifications (Table 3). Biologically significant misclassifications were extremely rare. Of our 100-species set, only four naturalizing species were placed in the 'accept' category by one or more of the models. Three of those species, *Campsis radicans* (Common Trumpetcreeper), *Catalpa speciosa* (Northern Catalpa), and *Maclura pomifera* (Osageorange), were accepted by one to three models. They are native to nearby regions and may, with the assistance of cultivation, disturbance in recipient habitats, and/or changes in climate, be expanding their post-glacial ranges into Iowa. Of these three, only *Maclura pomifera* has been identified as a particularly invasive plant in Iowa (9).

Disturbingly, the fourth species, *Berberis thunbergii* (Japanese Barberry), was accepted by all four models. This native of northeastern Asia is known to be invasive in the northeastern United States (8, 15), naturalized widely in the Chicago region between 1940 and 1974 (5, 35), and has also been identified as a problematic, invasive plant in Iowa (9). Either its pattern of naturalization differs from other invasive woody plants or perhaps the information we used in our data set is incorrect or incomplete for some important aspect of its life history; for example, other models have indicated the importance of length of time since introduction (21) and persistence of fruit (23) as important determinants for escape from cultivation.

Horticulturally limiting misclassifications were relatively common, with 19 non-naturalizing species from our 100-species set rejected by one or more models. Given that the models used in this study were designed to identify patterns of characteristics for naturalizing species, we examined these misclassified species further to see if they might be the source of future problems. All models rejected *Tamarix ramossissima* (Fivestamen Tamarisk). This species is known to disrupt riparian plant communities in the western United States (2), but may be at a competitive disadvantage in regions with a positive moisture balance, such as Iowa (38).

Five other species were rejected by at least three models**.** Of these, three are only marginally winter-hardy in Iowa (*Hibiscus syriacus* (Shrubalthea), *Pyrus calleryana* (Callery Pear), and *Quercus robur* (English Oak)). Our models predict that populations of these species hardier than those currently under cultivation could naturalize in Iowa. Another misclassified species, *Spiraea japonica* (Japanese Spirea), is widely cultivated throughout Iowa. It is known to be invasive in the southeastern United States (34), but our experience with *Spiraea* seed propagation suggests that seedling establishment is optimal on peat-based media, quite unlike typical Iowa soils, which may significantly limit its ability to naturalize.

Discriminant analyses and classification and regression tree approaches have generally resulted in classification rates between 60–90% (4, 21, 23, 25, 28). In most of these models, horticulturally limiting, false positives have accounted for a relatively high proportion of errors (19). Although these errors would limit potential new introductions, there is widespread support among scientists and conservation agencies to error on the side of prevention (1, 10, 29), especially given the uncertainty associated with lag times between introduction and confirmed naturalization for woody plants (16, 29).

The Modified decision tree, Decision tree/matrix model, and New CART model developed in this study have been validated only by evaluating the *a posteriori* classification of the plant species upon which they were based. A more robust evaluation of these models through external validation (12) seems warranted. One way to accomplish this would be to apply these models to data sets assembled for plant naturalization observed in nearby areas outside Iowa. This would serve multiple purposes. It would show which of these models is most powerful and accurate with independent data sets and indicate whether our observed improvements over the original continental decision tree are statistically significant. Second, the results could be used to develop regional risk-assessment models applicable to Iowa and surrounding states.

Literature Cited

1. Bell, C., C. Wilen, and A. Stanton. 2003. Invasive plants of horticultural origin. HortScience 30:14–16.

2. Brock, J.H. 1994. *Tamarix* spp. (salt cedar), an invasive exotic woody plant in arid and semi-arid riparian habitats of western USA. pp 27–44. *In*: L.C. de Waal, L.E. Child, P.M. Wade, and J.H. Brock (eds.). Ecology and Management of Invasive Riverside Plants. John Wiley, Chichester, England.

3. Clark, L.A. and D. Pregibon. 1992. Tree-based Models. Chapter 9. *In*: J.M. Chambers and T.J. Hastie (eds.). Statistical Models in S. Wadsworth and Brooks, Pacific Grove, CA.

4. Daehler, C.C. and D.A. Carino. 2000. Predicting invasive plants: prospects for a general screening system based on current regional models. Biological Invasions 2:93–102.

5. Deam, C.C. 1940. Flora of Indiana. Indiana Department of Conservation, Indianapolis, IN.

6. Dirr, M.A. 1998. Manual of Woody Landscape Plants, 5th ed. Stipes, Champaign, IL.

7. Dirr, M.A. and C.W. Heuser, Jr. 1987. The Reference Manual of Woody Plant Propagation: From Seed to Tissue Culture. Varsity Press, Athens, GA.

8. Ehrenfeld, J.G. 1997. Invasion of deciduous forest preserves in the New York metropolitan region by Japanese barberry (*Berberis thunbergii*). Journal of the Torrey Botanical Society 124:210–215.

9. Farrar, D.F. 2001. Exotic and invasive woody plant species in Iowa. J. Iowa Acad. Sci. 108:154–157.

10. Fox, A.M., D.R. Gordon, and R.K. Stocker. 2003. Challenges of reaching consensus on assessing which non-native plants are invasive in natural areas. HortScience 38:11–16.

11. Goodwin, B.J., A.J. McAllister, and L. Fahrig. 1999. Predicting invasiveness of plant species based on biological information. Conservation Biology 13:422–426.

12. Harrell, Jr., F.E. 2001. Regression Modeling Strategies. pp 90–91. Springer Verlag, New York.

13. Harrington, R.A., R. Kujawski, and H.D.P. Ryan. 2003. Invasive plants and the green industry. J. Arboric. 29:42–48.

14. Heffernan, B. 2002. Managing a tough issue. American Nurseryman 196(10):34–39.

15. Kourtev, P.S., J.G. Ehrenfeld, and W.Z. Huang. 1998. Effects of exotic plant species on soil properties in hardwood forests of New Jersey. Water Air and Soil Pollution 105:493–501.

16. Kowarik, I. 1995. Time lags in biological invasions with regard to the success and failure of alien species. pp 15–38. *In*: P. Pyšek, K. Prach, M. Rejmánek , and M. Wade (eds.). Plant Invasion: General Aspects and Special Problems. Academic Publishing, Amsterdam, The Netherlands.

17. Kriticos, D.J. and R.P. Randall. 2001. A comparison of systems to analyze potential weed distributions. pp 61–79. *In*: R.H. Groves, F.D. Panetta and J.G. Virtue (eds.). Weed Risk Assessment. CSIRO Publishing, Collingwood, Australia.

18. Mack, R.N. 1996. Predicting the identity and fate of plant invaders: Emergent and emerging approaches. Biological Conservation 78:107–121.

19. 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.

20. MathSoft. 1998. S-Plus 5 for UNIX. MathSoft, Seattle, WA.

21. Maillet, J. and C. Lopez-Garcia. 2000. What criteria are relevant for predicting the invasive capacity of a new agricultural weed? The case of invasive American species in France. Weed Research 40:11–26.

22. Panetta, F.D. 1993. A system for assessing proposed plant introductions for weed potential. Plant Protection Quarterly 8:10–14.

23. 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. Journal of Environmental Management 57:239–251.

24. Pimentel, D., L. Lach, R. Zuniga, and D. Morrison. 2000. Environmental and economic costs of nonindigenous species in the United States. BioScience 50:53–65.

25. Prinzig, A., W. Durka, S. Klotz, and R. Branell. 2002. Which species become aliens? Evolutionary Ecology Research 4:385–405

26. Randall, R. 2002. Weed science. The big weed list. Retrieved August–September 2002 from: <http://www.agric.wa.gov.au/progserv/plants/weeds/weeds/weedlist.htm>.

27. Reichard, S.E. 1994. Assessing the potential of invasiveness in woody plants introduced in North America. Ph.D. dissertation, University of Washington, Seattle.

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

29. Reichard, S.H., and P. White. 2001. Horticulture as a pathway of invasive plant introductions in the United States. BioScience 51:103–113.

30. Rejmánek, M. 1999. Invasive plant species and invasible systems. pp 79–102 *In*: O. Sandlund, P. Schei, and A Viken (eds.). Invasive Species and Biodiversity Management. Kluwer Academic Publishers, Boston, MA.

31. Rejmánek, M. 2000. Invasive plants: Approaches and predictions. Austral Ecology 25:497–506.

32. Rejmánek, M., and D.M. Richardson. 1996. What attributes make some plant species more invasive? Ecology 77:1655–1661.

33. Siegel, S. 1956. Nonparametric Statistics. McGraw Hill, New York. 312 pp.

34. Southeast Exotic Pest Plant Council. 2001. Invasive Exotic Pest Plants in Tennessee. Accessed on 28 May 2003 on the Internet at <http://www.se-eppc.org/states/TN/TNIList.html>.

35. Swink, F. 1974. Plants of the Chicago Region, 2nd ed. The Morton Arboretum, Lisle, IL.

36. Tucker, K.C. and D.M. Richardson. 1995. An expert system for screening potentially invasive alien plants in South African fynbos. Journal of Environmental Management 44:309–338.

37. Tutin, T.G., V.H. Heywood, N.A. Burges, D.M. Moore, D.H. Valentine, S.M. Walters, and D.A. Webb. (eds.) 1964–1980. Flora Europaea. Cambridge University Press, Cambridge. 5 volumes.

38. 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.

39. Widrlechner, M.P. 2001. The role of environmental analogs in identifying potentially invasive woody plants in Iowa. J. Iowa Acad. Sci. 108:158–165.

40. Widrlechner, M.P., E.R. Hasselkus, D.E. Herman, J.K. Iles, J.C. Pair, E.T. Paparozzi, R.E. Schutzki, and D.K. Wildung. 1992. Performance of landscape plants from Yugoslavia in the north central United States. J. Environ. Hort. 10:192–198.

41. Widrlechner, M.P., J.B. Hebel, D.E. Herman, J.K. Iles, G.J. Kling, A.P. Ovrom, J.C. Pair, E.T. Paparozzi, S.R. Poppe, N. Rose, R.E. Schutzki, C. Tubesing, and D.K. Wildung. 1998. Performance of landscape plants from northern Japan in the north central United States. J. Environ. Hort. 16:27–32.

42. 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.

43. Young, J.A. and C.G. Young. 1992. Seeds of Woody Plants in North America, Rev. ed. Dioscorides Press, Portland, OR.