Predicting the Spread of Purple Loosestrife (*Lythrum salicaria*) in the Prairies

CORY J. LINDGREN[1] and DAVID WALKER[2]

[1]Department of Environment and Geography, University of Manitoba, Winnipeg, Manitoba R3T 2N2 Canada; email: cjllind@mymts.net
[2]Department of Environment and Geography, University of Manitoba, Winnipeg, Manitoba R3T 2N2 Canada

Lindgren, Cory J., and David Walker. 2012. Predicting the spread of Purple Loosestrife (*Lythrum salicaria*) in the Prairies. Canadian Field-Naturalist 126(4): 306–319.

Purple Loosestrife (*Lythrum salicaria*) is an invasive plant introduced into North America in the early 1800s. It has since spread into the prairie provinces of Canada (Manitoba, Saskatchewan, and Alberta). It invades wetland habitats, marshes, riparian areas, and natural areas, and it outcompetes native wetland vegetation. In this study we modelled the potential distribution of Purple Loosestrife in the Prairies, explored which suites of predictive variables produced the best ecological niche models, and explored two different approaches to the partitioning of data in evaluating models. We used a number of performance measures and expert evaluation to select our best models. The best model was developed using a suite of climate variables and growing degree-days as the predictive variables and by partitioning testing and training data using stratified random sampling. The model indicated that Purple Loosestrife has not yet reached its full potential distribution in the Prairies. The modelling techniques presented in this paper may be used to predict the potential distribution of other emerging invasive plants, and the results can be used to optimize early detection and surveillance strategies for Purple Loosestrife in areas of the Prairies.

Key Words: Purple Loosestrife, *Lythrum salicaria*, invasive weed, ecological niche modelling, genetic algorithm for rule-set prediction, GARP, Manitoba, Saskatchewan, Alberta.

Invasive plants threaten the economy, environment, and health, as well as managed and unmanaged systems worldwide (Bradley et al. 2009). Purple Loosestrife (*Lythrum salicaria*; Lythraceae) is an invasive plant that was introduced into North America in the early 1800s and has since spread across Canada (Thompson et al. 1987; Mal et al. 1992; Lindgren 2003; Welk 2004). It invades wetland habitats, marshes, riparian areas, and other natural areas (White et al. 1993*; Mal et al. 1997; Mullin 1998; Anderson et al. 2006), and it outcompetes native wetland vegetation (Gaudet and Keddy 1988; Johansson and Keddy 1991; Gaudet and Keddy 1995; Mal et al. 1997). It is found in all three Canadian prairie provinces (Manitoba, Saskatchewan, and Alberta), and once established it is difficult and costly to eradicate (Ottenbreit 1991; Ottenbriet and Staniforth 1994; Ali and Verbeek 1999*; Lindgren et al. 2001; Lindgren 2003).

Studies have modelled the potential distribution of Purple Loosestrife at continental and regional scales in North America (Welk 2004; Soberon and Peterson 2005; Anderson et al. 2006; Bella 2009*), but there have been no studies that model the potential distribution of Purple Loosestrife specifically in the Prairies. Although Purple Loosestrife has become established in parts of the Prairies (Lindgren 2003), it would be of significant value to know if it has reached the full extent of its geographic distribution within this region, to help determine, for example, which management strategies could be effective. Hence, the first objective of this study was to model the potential distribution of Purple Loosestrife in order to determine whether it has reached its full range potential in the Prairies.

Ecological niche modelling

Increases in international trade have resulted in an increasing number of new invasive weeds being introduced globally. Hence there is a need to explore the use of tools that may support preventive strategies by predicting the potential distribution of a pest in a new area (Pheloung et al. 1999; Zalba et al. 2000; Brasier 2008; Dehnen-Schmutz et al. 2010; Lindgren 2012).

Predicting pest distributions is a topic of interest dating back many years (Cook 1925; Messenger 1959). Modelling geographic distributions has been referred to as ecological niche modelling (Peterson and Cohoon 1999). Ecological niche modelling generally characterizes the abiotic conditions (e.g., climatic conditions) associated with a known location of a species in one area and attempts to predict the potential distribution of that species in a new area based upon those conditions. Ecological niche modelling attempts to approximate Hutchinson's (1957) fundamental niche (Soberon and Peterson 2005).

There are a variety of modelling approaches that have been used to estimate the ecological niche, including BIOCLIM (Busby 1986), Maxent (Phillips et al. 2006), CLIMEX (Sutherst and Maywald 1985; Sutherst et al. 2000), and the Genetic Algorithm for Rule-set Prediction (GARP) (Stockwell and Noble 1992; Stockwell 1997; Stockwell and Peters 1999); see Guisan and Thuiller (2005) for a review of these modelling approaches.

In this study we explored the use of the GARP algorithm, as it has been used successfully across a wide range of disciplines. It has been used to predict the potential distribution of invasive plants (Madsen 1999; Daehler and Carino 2000; Peterson 2001; Welk et al. 2002; Peterson et al. 2003; Sanchez-Flores 2007), mice (Anderson et al. 2002), aphids (Ganeshaiah et al. 2003), owls (Peterson and Robins 2003), butterflies (Oberhauser and Peterson 2003), and diseases (Levine et al. 2004; Adjemian et al. 2006) and to explore how climate change influences potential distributions (Kerns et al. 2009). Specific to invasive plants, it has been successfully used to predict the potential distribution of Garlic Mustard (*Alliaria petiolata*), Russian Olive (*Elaeagnus angustifolia*), Hydrilla (*Hydrilla verticillata*), and Sericea Lespedeza (*Lespedeza cuneata*) (Welk et al. 2002; Peterson et al. 2003).

The GARP algorithm is a machine learning approach that uses algorithms to enable the computer to learn from experience and improve its prediction over time. The GARP algorithm uses decision rules to summarize the ecological niche of a species, as defined by a set of known presence points in one area, and then predicts the potential distribution in a new area based on the summarized ecological niche (Peterson et al. 2003). The rule types used by GARP are atomic, envelope (i.e., based on BIOCLIM rules), and logit. It resamples known occurrence points and pseudo-absence points (e.g., sites at which the species is not known to occur) randomly with replacement to create training and test data sets of up to 1250 points each. It works in an iterative process to develop rules that identify key niche parameters, evaluates their importance and predictivity, and either incorporates them into the model or rejects them (Oberhauser and Peterson 2003). GARP is a superset of individual algorithms that has greater predictive ability than any one of them (Peterson 2001).

When compared to other ecological niche modelling approaches, GARP has several advantages in that (1) it is an algorithm that iteratively evaluates and improves on prediction rules used to generate a predictive risk map (Stockwell and Peters 1999); (2) it is data-driven, producing informative models that allow parameters to be optimized using expert knowledge and errors of omission and commission (Peterson and Cohoon 1999; Stockwell and Peterson 2002); (3) it has been used by a wide variety of practitioners across a number of disciplines (see examples above); (4) it is predictive in that it anticipates a pest's distribution in geographic areas where distribution information is lacking (Peterson 2001); (5) it is a superset of other modelling approaches, providing greater predictive ability than any one individual approach (Peterson 2001); and (6) it is freely available. Free access to a spatial modelling tool allows others to critically scrutinize and replicate risk maps (Kriticos and Randall 2001). Open access is particularly important for countries, for example, that are signatories to the International Plant Pest Convention (see www.ippc.int) and hence have obligations to conduct science-based risk assessments but may not have the resources to purchase expensive modelling applications.

Selecting predictive variables

In developing an ecological niche model, the choice of predictive variables affects the final risk model and how robust it will be. When models perform well, it is generally because predictor variables that are associated with habitat suitability have been selected. When models do not perform well, it suggests that meaningful predictor variables were not selected (Evangelista et al. 2008). It would be of value to know which predictor variables determine a species' distribution; however, such knowledge is generally lacking (Jimenez-Valverde et al. 2011). The second objective of this study was therefore to determine which of two suites of predictive variables and growing degree-days, or combinations of them, produced the most realistic ecological niche models for Purple Loosestrife.

There are many abiotic variables that may determine a species' geographic distribution, but it is likely that climate, topography, and growing degree-days are the primary variables constraining the distribution of Purple Loosestrife in the Prairies. In this study, we explored the influence of climate variables, as these are known to be principal predictive variables in determining species' distributions (Andrewartha and Birch 1954; Peterson and Cohoon 1999; Welk et al. 2002; Pearson and Dawson 2003; Welk 2004; Helaouet and Beaugrand 2009; Kearney and Porter 2009). Specifically, we explored temperature and precipitation as predictive variables, as they have been found to be determining factors in the distribution of invasive plants, including Kudzu (*Pueraria lobata*) (Follak 2011) and Purple Loosestrife (Bella 2009*).

We selected growing degree-days, a thermal measure associated with air temperature, as a predictive variable, as it is considered a spatially dynamic variable (Hassan et al. 2000; Jodoin et al. 2008; Hassan and Bourque 2009) that is known to constrain the distribution of Purple Loosestrife (Lindgren and Walker 2012) and is a driver of species' distributions (Austin et al. 2006). We also explored the influence of topographic variables, as they also have been reported to be determining variables in species' distributions (Kearney and Porter 2009) and are correlated with wetland plants (Welk 2004). Climate, topography, and growing degree-days are also landscape-scale variables which are meaningful in assessing distributions at large spatial scales (Peterson et al. 2011), such as the Prairies.

Geographic partitioning of data

The third objective of this study was to explore two different approaches to data partitioning. Occurrence point data are commonly partitioned into training and testing datasets to evaluate the resulting model. Hence, the way in which the data are partitioned needs to be

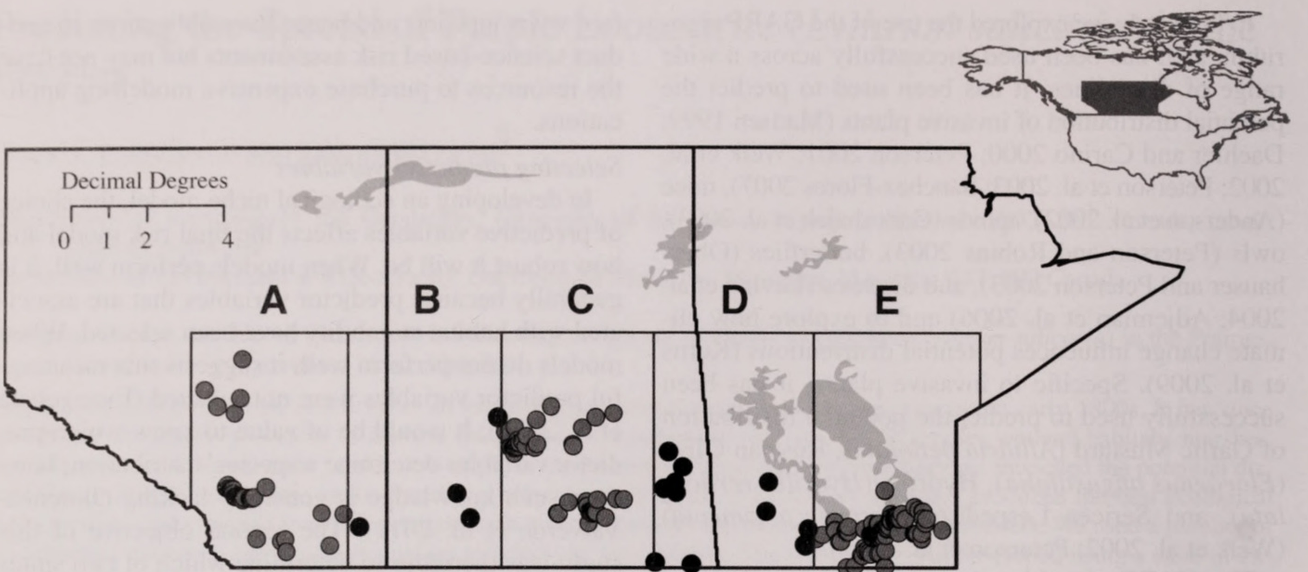

FIGURE 1. Illustration of the quintile data partitioning approach used to evaluate model predictivity for the spread of Purple Loosestrife (*Lythrum salicaria*) in the Prairies. Occurrences in areas B and D were used to train models using occurrence points ($N = 22$) (black dots); areas A, C, and E were used to test the models using occurrence points ($N = 609$) (grey dots). Study area consists of Alberta, Saskatchewan, and Manitoba (from left to right). Major lakes are shown for context.

carefully considered (Fielding and Bell 1997; Peterson and Shaw 2003). Occurrence data are generally partitioned so that one subset is used to train the model and another subset is used to independently test the model (Fielding and Bell 1997; Osborne and Suárez-Seoane 2002; Peterson and Shaw 2003; Heikkinen et al. 2007; Syartinilia and Tsuyuki 2008). To calculate measures of model accuracy, an independent testing dataset is generally withheld from model training (Fielding and Bell 1997; Osborne and Suárez-Seoane 2002; Peterson and Shaw 2003; Peterson et al. 2011). Examples of data partitioning approaches include random sampling (Osborne and Suárez-Seoane 2002), stratified random sampling (Hirzel and Guisan 2002), and partitioning by administrative boundaries (Kapetsky et al. 1988; Anderson et al. 2006).

In this study, we explored the use of two approaches to data partitioning in testing the predictive power of our GARP models: (1) stratified random sampling, in which the data were partitioned by province; and (2) a quintile approach, in which the data were partitioned into five regions of equal size. Partitioning data using these approaches forces the model to predict into broad, unsampled areas from which no input occurrence points are available (Peterson and Shaw 2003; Peterson et al. 2007). Partitioning the testing and training data using a stratified random sampling approach simulated a scenario whereby sampling effort might be planned through provincial survey efforts, with specific administrative boundaries. Both partitioning methods allowed for realistic model evaluation, as the majority of the evaluation area is limited to areas into which Purple Loosestrife might disperse and hence should reduce overall errors of commission in the model (Peterson et al. 2011).

Methods

Study area

The study area covers an area of 360 000 km^2 and includes the provinces of Manitoba, Saskatchewan, and Alberta (Figure 1). The spatial extent of the study area represents accessible area, an important consideration often overlooked in modelling studies (Anderson and Raza 2010; Barve et al. 2011; Peterson 2011). The study area constitutes a geographic space in which Purple Loosestrife has become established and into which it may extend its range, based upon abiotic and dispersal factors. Because of the large scale of the study area, biotic factors should not significantly influence the models.

Occurrence data

One of the challenges of spatial predictive modelling is obtaining accurate occurrence data, as the quality and quantity of these data directly influence modelling results (Welk 2004; Elith et al. 2006; Yemshanov et al. 2010). We obtained 631 geo-referenced occurrence points for the study area. Manitoba occurrence data (i.e., site records) were collected between 1992 and 2004 by CJL (e.g., Lindgren 2003), and the remaining validated data were collected by the Saskatchewan Purple Loosestrife Eradication Project (Sumners 2005) and the Alberta Purple Loosestrife program (Ali and Verbeek 1999*; Cole et al. 2007) (this dataset can be obtained by contacting CJL). Models were developed using presence only data (GARP generates pseudo-absence points in model development), and no occurrence data from garden plantings, herbaria, or museums were included in the dataset.

Predictive variables

We used (1) growing degree-days, (2) a suite of climate variables (mean daily temperature, mean annual diurnal temperature, mean annual precipitation, and mean annual number of wet days), and (3) a suite of topography variables (elevation, slope, and aspect) as the predictive variables. To explore the influence of the predictive variables on predictive accuracy, we ran GARP models with each variable alone (i.e., first order models) as well as with all possible combinations of the variables (i.e., second and third order models).

We were unable to find growing degree-days data that were specific to Purple Loosestrife, so we created a new predictive layer using a T_{base} of 8°C, which is a threshold temperature specific to Purple Loosestrife growth (Shamsi and Whitehead 1974). To calculate growing degree-days (GDD), the following equations were used: $GDD_{daily} = (T_{max} + T_{min}) / 2 - T_{base}$, and cumulative GDD = $\sum GDD_{daily}$, where T_{max} is the maximum daily temperature, T_{min} is the minimum daily temperature, and T_{base} is the base temperature where growth and development are not deemed to occur (McMaster and Wilhelm 1997; Miller et al. 2001; Laswon et al. 2006; Shrestha et al. 2010). Cumulative growing degree-days were calculated by summing those daily growing degree-days that were above 0°C. Temperature data from 48 weather stations that covered the geographic extent of the Prairies were used to interpolate a growing degree-days surface. Data were imported into an Excel spreadsheet to process cumulative growing degree-days.

We used ArcGIS 9.3's (Environmental Systems Research Institute, Redlands, California) Geostatistical Analyst tool to krige (i.e., spatial correlation modelling) (Cousens et al. 2002; Kleijnen 2009) a surface using latitude, longitude, and the growing degree-days value, and we converted this into a digital raster layer. Digital raster GIS datasets for climate and topography predictor variables were obtained from the Intergovernmental Panel on Climate Change (Intergovernmental Panel on Climate Change 2001). Data were sampled at a pixel resolution of $0.2^0 \times 0.2^0$ or about 20×20 km grid. We created a study area mask of the Prairies to prevent GARP from selecting pseudo-absence data from areas where Purple Loosestrife was absent for possible abiotic or dispersal reasons, and the mask forced models to be trained in the geographic area of interest (Barve et al. 2011; Peterson et al. 2011).

Model building

We used GARP to model the potential distribution of Purple Loosestrife across the Prairies (Stockwell and Peters 1999). In producing each GARP model, we set optimization parameters to 100 runs (i.e., each run produced a unique model) or a convergence limit of 0.01 with the number of maximum iterations set to 1000. Occurrence points were divided randomly by GARP: 50% into training data and 50% into testing datasets.

For each task, atomic, range, negated range, and logistic regression rule sets were selected. GARP selects a rule set method and applies it to the training data, and a rule is developed. Rules then evolve through an iterative process to maximize predictivity, where the change in predictive accuracy from one iteration to the next is used to determine whether a specific rule should be incorporated and each rule set represents a different method of characterizing the ecological niche (Peterson and Shaw 2003; Peterson et al. 2007). Predictive accuracy is evaluated based upon 1250 points resampled from the test data and 1250 pseudo-absence points from the Prairies (Peterson and Shaw 2003). For all GARP models, a subset of data was withheld to allow for independent testing of the model.

As each of the 100 models produced was unique and varied in quality, we selected a best subset of 10 models and summed these together to create one composite model (see Anderson et al. 2003 for discussion on best subset approach). The best subset approach minimizes overfitting by prioritizing errors of omission over errors of commission (Peterson et al. 2008). In optimizing our modelling parameters, we selected an extrinsic omission threshold of 10% so that models with greater that 10% of testing points omitted would be excluded from the final composite model; we set the commission threshold at 50%. Using extrinsic training data, we calculated the median commission index across models with the lowest number of errors of omission, and the models with indices closest to the commission median were chosen as the best subset.

Evaluating model performance

Model evaluation is used to identify models that predict into either excessively small or excessively large areas (Raxworthy et al. 2007). As there is no accepted single way to measure model performance (Fielding and Bell 1997; Peterson et al. 2011; Tarkesh and Jetschke 2012), we employed a number of measures, as well as expert evaluation (Table 1). The GARP algorithm produces binary models that allow performance measures to be calculated from elements of a 2×2 confusion matrix (Fielding and Bell 1997; Welk 2004). In the confusion matrix, element *a* represents pixels where the species is known to occur and the model correctly identifies as present, element *b* represents pixels where the species is not known to occur but are incorrectly identified as present (i.e., errors of commission) (false positives), element *c* represents pixels of known distribution incorrectly identified as absent by the model (i.e., errors of omission) (false negatives), and element *d* represents pixels where the species has not been found and the model correctly identifies as absent.

TABLE 1. Measures used to evaluate performance and accuracy of best subset composite models used to predict the spread of Purple Loosestrife (*Lythrum salicaria*) in the Prairies. Letters *a* to *d* represent elements of the 2 × 2 confusion matrix (see Methods).

Measure	Calculation
Intrinsic errors of omission (performance)	$c / (a + c)$
Intrinsic commission index (performance)	$b / (b + d)$
Sensitivity (performance)	$a / (a + c)$
Specificity (performance)	$d / (d + b)$
Kappa (accuracy)	$\dfrac{(a + d) - ((a + c)(a + b) + (b + d)(c + d)) / N}{N - ((a + c)(a + b) + (b + d)(c + d)) / N}$
Testing accuracy	$(a + b) / (a + b + c + d)$
Extrinsic accuracy	outside model$_{\text{test points}}$ / $N_{\text{test points}}$

We used extrinsic accuracy, testing accuracy, sensitivity, specificity, omission error rate, and commission index to measure the accuracy of our models. Sensitivity is the proportion of observed true positives, correctly indicating how good the model is at detecting a pest or an occurrence point (Fielding and Bell 1997). Specificity is the proportion of observed true negatives or absences that are predicted as absent, indicating how good the model is at detecting absences or predicting no presence (Fielding and Bell 1997). The intrinsic omission error is the proportion of known localities that fall outside the predicted area (i.e., the false negative rate), and the intrinsic commission index (i.e., the false positive rate) is the proportion of pixels predicted as present by the model (Anderson et al. 2003). In general, models with zero or low errors of omission that are sensitive are desired (Peterson et al. 2011).

We also used the kappa statistic (κ), as it corrects the overall accuracy of model predictions by the accuracy expected to occur by chance, and it also accounts for both errors of commission and omission in one parameter (Landis and Koch 1977; Fielding and Bell 1997; Liu et al. 2005; Allouche et al. 2006; Zhu et al. 2007; Tarkesh and Jetschke 2012). However, the kappa statistic should be used with caution, as it weights errors of omission and commission equally and hence may not be a good measure of performance for invasive species models, where errors of omission are considered more serious than errors of commission (Guisan and Thuiller 2005; Peterson et al. 2011).

We also measured model accuracy by using independently withheld testing data to calculate the percentage of the number of known occurrence points predicted correctly.

While it is important that practitioners quantitatively evaluate models, the importance of expert evaluation in evaluating models cannot be overlooked (Anderson et al. 2003), and expert evaluation should be incorporated into final model selection methodology (Thuiller 2003). Expert evaluation is required to determine whether the ecological niche model is geospatially realistic and make senses both intuitively and biologically. Expert evaluation has been found to be very informative, for example, in risk analysis frameworks (Pheloung et al. 1999; Therriault and Herborg 2008).

In this study, we defined an expert as someone who has both extensive biological knowledge of Purple Loosestrife (i.e., the species) and knowledge of the biogeography of the Prairies (i.e., the study area). Our definition of an expert is similar to that of Anderson et al. (2003). We, the authors of this study, served as the expert evaluators. We interpreted the composite models and evaluated them as either good or poor. We defined a good model (in this paper we refer to a good model also as a realistic model) as one which excluded unsuitable areas where Purple Loosestrife could not exist (i.e., high elevations of the Rocky Mountains in Alberta or areas of the Boreal Plains ecozone (Ecological Stratification Working Group 1996) where Purple Loosestrife cannot become established) or disperse (i.e., areas where there would be no known pathways for possible introduction). A poor model was one that included large unsuitable areas (i.e., areas where Purple Loosestrife could not disperse or become established due to biotic or abiotic events). For example, a good model accurately delimited the current distribution as well as predicting potential distribution into novel areas where expert opinion determined there were suitable abiotic conditions as well as a potential to disperse into these areas.

We employed an error cost criterion, where errors of omission (i.e., false negatives) were considered to be more costly that errors of commission (i.e., false positives); an error of omission is more serious than an error of commission, as it indicates a model has failed to predict known occurrence points (Raxworthy et al. 2003; Chen et al. 2007; Peterson et al. 2008).

Geographic partitioning of data

To test the predictive power of the ecological niche models, we partitioned the occurrence data geographically using a quintile approach and a stratified random sampling approach (i.e., by province). Both approaches tested the models' ability to predict into unknown geographic space. In the quintile approach, data were partitioned for model testing and training into five regions (i.e., quintiles) of about four degrees longitude each (Figure 1). To test the accuracy of the model in predicting across unsampled areas of the study area, we used quintiles A, C, and E to independently test models (number of occurrence points) (N = 609) and we used quintiles B and D to train (N = 22) GARP models. In the stratified random sampling approach, occurrence data were first partitioned among the three provincial boundaries, and then about 30% of the data from each province were randomly selected for model testing (N = 441 for model training and N = 190 for model testing).

Geographic information processing

ArcGIS 9.3 was used to process and project the GARP models. The 10 best models were imported into ArcGIS, converted from ASCII files to raster grid files, and projected onto a map of the three prairie provinces. For each model, GARP predicts Purple Loosestrife as either present or absent within a pixel. The ArcGIS Spatial Analyst tool (i.e., local cell statistics) was used to sum all 10 best subset models together to create a final composite model. Using ArcGIS 9.3, we reclassified the modelling results into one of six categories representing probable risk: 0 (i.e., no models predicted presence), <25%, 26–50%, 51–75%, 76–99%, and 100% (i.e., all 10 models agreed). Projection of the composite model onto a map of the Prairies provided a final invasive risk map.

Results and Discussion

Geographic partitioning of data

We found that the way in which we partitioned the data (i.e., into training and testing subsets) influenced the modelling results. Evaluation data for the composite models developed using the quintile data partitioning approach are found in Table 2. Using the performance measures and expert evaluation to assess the quintile models, we determined that the single variable climate model (Figure 2A) was the best model.

The next best model was the three variable model (Figure 2G), which had good performance and accuracy measures but suffered from errors of omission (i.e., the model failed to predict as suitable areas of central Alberta where there are known established populations). Using our adopted error cost criterion, where an error of omission is considered to be the most serious error (Wiley et al. 2003), we evaluated the three variable model as poor.

The remaining models also had errors of omission (Figures 2C, 2D, and 2E), errors of commission (Figures 2B and 2F), and in some cases both (Figures 2B, 2C, 2D, 2F), and were hence evaluated as poor. While errors of commission, or overprediction, may be desirable in invasive species models (Stockman et al. 2006), the topography model and the topography and climate model (Figures 2B and 2F) predicted potential distributions into areas of northern Alberta that were beyond what we expected.

TABLE 2. Performance data for ecological niche models for Purple Loosestrife (*Lythrum salicaria*) in the Prairies developed using the quintile data partitioning approach. Measures represent mean values for the best subset composite model. Predictive variables used to develop GARP models were climate, topography (Topo), and growing degree-days (GDD). Models were evaluated using specificity, sensitivity, testing accuracy, the kappa statistic, and expert evaluation, number of occurrence points found outside the model, independent test accuracy using points found outside the model, and errors of commission and omission.

	Model evaluation							Error rate	
Predictor dataset	Specificity (%)	Sensitivity (%)	Testing accuracy (%)	Kappa	Expert evaluation	Occurrence points outside the model	Independent test accuracy (%)	Commission (%)	Omission (%)
Three variable model									
Climate + Topo + GDD	69	97	59	0.98	poor	114	82	35	0
Two variable models									
Climate + GDD	25	90	73	0.98	poor	185	70	25	0
Topo + GDD	34	78	71	0.99	poor	183	70	34	0
Topo + Climate	65	95	61	0.98	poor	0	100	54	0
Single variable models									
Climate	33	75	70	0.85	good	0	100	34	0
Topography	35	70	67	0.99	poor	188	70	34	2
GDD	44	100	77	0.98	poor	189	69	43	0

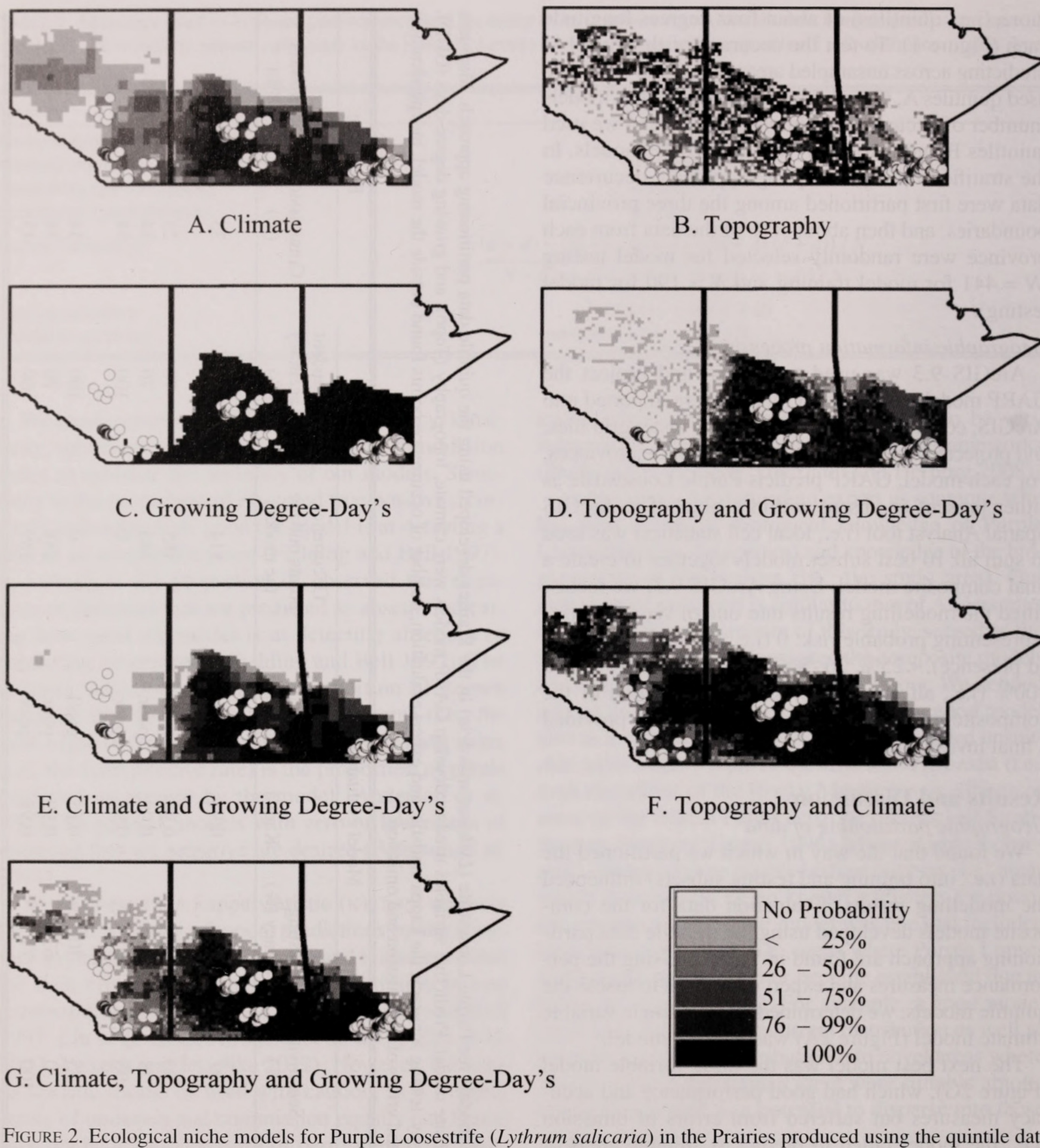

FIGURE 2. Ecological niche models for Purple Loosestrife (*Lythrum salicaria*) in the Prairies produced using the quintile data partitioning approach. White circles represent occurrence points use to test the model.

Evaluation data for composite models developed using the stratified random sampling partitioning approach are found in Table 3. Using the performance measures and expert evaluation, we selected the climate and growing degree-days model (the two variable model) as the best model developed when data were partitioned using a stratified random sampling approach (Figure 3E). The model had a low intrinsic commission index as well as high values for testing accuracy and sensitivity. Expert evaluation determined it appeared to be the most realistic in predicting the potential distribution of Purple Loosestrife in the study area.

The next best model was determined to be the single variable climate model (Figure 3A), which had a low commission index and high values for sensitivity and testing accuracy. Expert evaluation determined it to be a good model, but, when it was compared with the climate and growing degree-days model, it over-predicted into areas of northwestern Alberta where climate conditions as well land use (i.e., boreal forest) would most likely prevent Purple Loosestrife from becoming established.

The remaining models were evaluated as poor, as they suffered from errors of omission (3B, 3C, 3D),

TABLE 3. Performance data for ecological niche models for Purple Loosestrife (*Lythrum salicaria*) in the Prairies using the stratified random sampling approach to partition the data. Measures represent mean values for the best subset composite model. Predictive variables used to develop GARP models were climate, topography (Topo), and growing degree-days (GDD). Models were evaluated using specificity, sensitivity, testing accuracy, the kappa statistic, and expert evaluation, number of occurrence points found outside the model, independent test accuracy using points found outside the model, and errors of commission and omission.

	Model evaluation							Error rate	
Predictor dataset	Specificity (%)	Sensitivity (%)	Testing accuracy (%)	Kappa	Expert evaluation	Number of occurrence points outside the model	Independent test accuracy (%)	Commission (%)	Omission (%)
Three variable model									
Climate + Topo + GDD	56	96	69	0.77	poor	0	100	50	2
Two variable models									
Climate + GDD	31	94	82	0.83	good	0	100	32	2
Topo + GDD	42	92	74	0.82	poor	1	99	42	9
Topo + Climate	32	92	73	0.83	poor	0	100	32	3
Single variable models									
Climate	25	89	79	0.76	good	0	100	25	7
Topography	59	84	63	0.79	poor	7	97	58	10
GDD	58	99	70	0.77	poor	1	99	59	1

errors of commission (Figure 3B, 3C, 3D, 3F, 3G), or both (Figure 3B, 3C, and 3D).

Overall, we concluded that the most realistic models were produced when data were partitioned using a stratified random sampling approach (see expert evaluation discussion below).

Selecting predictive variables

As we expected, the selection of predictive variables used to build our models significantly influenced the final results. The single variable models produced using topography and growing degree-days as a single predictive variable were evaluated to be of poor quality. Using either of the data partitioning approaches, we found that the topographic models (Figures 2B and 3B) overpredicted suitable area across the majority of the study area except areas of high elevation in the Rocky Mountains of western Alberta and areas of low elevation in northern Manitoba and Saskatchewan. The growing degree-days models (Figures 2C and 3C) developed using either of the data partitioning methods also suffered from errors of omission, as they failed to predict as suitable areas in central Alberta where known occurrence point data existed. In addition, each pixel in the potential distribution was determined to be 100% at risk. In other words, all ten models forming the composite model agreed that the pixel was suitable, illustrating the limitations of using only one predictive variable in developing a model.

Based on performance measures and expert evaluation, the GARP models using the suite of climate variables (Figure 2A and 3A) were found to be of good quality, as they correctly captured the current distribution and predicted a realistic potential distribution, where the northern extent of the potential distribution of Purple Loosestrife would be constrained by the physiographic features of the Boreal Plains ecozone (Ecological Stratification Working Group 1996).

The two variable topography and growing degree-days models (Figures 2D and 3D) both suffered from errors of omission, as they did not predict potential areas in central Alberta as suitable habitat when in fact there were known occurrence points. We judged these models to be poor. These models also suffered from errors of commission, as they predicted potential distribution into northern areas of Saskatchewan and Alberta where expert opinion determined Purple Loosestrife could not become established or disperse.

The climate and growing degree-days models (Figure 3E) were judged to be good when stratified random sampling was used to partition the data, but there were errors of omission when the quintile approach was used (i.e., did not predict suitable areas through Alberta). The ecological niche model produced with topography and climate variables using stratified random sampling to partition the data (Figure 3F) was of good quality; the model produced using the quintile partitioning (Figure 2F) approach unrealistically overpredicted suitable area into northwestern Alberta. This

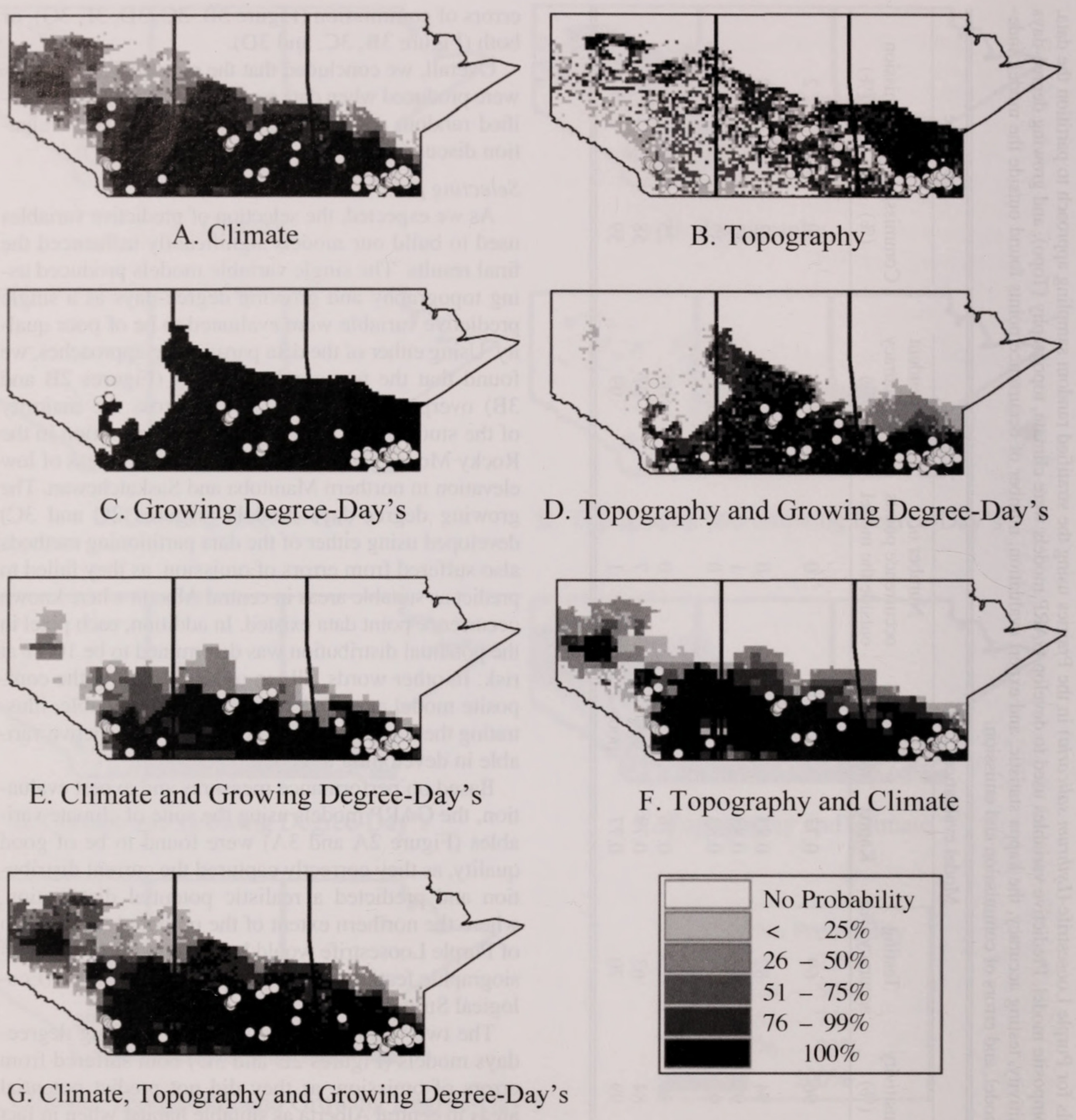

FIGURE 3. Ecological niche models for Purple Loosestrife (*Lythrum salicaria*) in the Prairies produced using stratified random sampling to partition the data. White circles represent occurrence points use to test the model.

is a good example of how different data partitioning methods produce different potential distributions.

In this study, more variables did not produce better models, as the models using all three predictive variables overpredicted into areas of northwestern Alberta and were evaluated as poor. When stratified random sampling was used to partition the data (Figure 3G), the amount of area predicted as suitable by all 10 models in the composite model seemed unrealistic. When the quintile data partitioning approach was used, the three variable model suffered from errors of omission (it failed to predict areas of central Alberta as suitable where there are known occurrence points) and was evaluated as poor (Figure 2G). It also overpredicted into northern areas of Alberta where expert opinion determined Purple Loosestrife could not become established or disperse.

Expert evaluation

We found that expert evaluation was a useful discriminatory measure in selecting the best or most realistic models. Overall, expert evaluation determined that 78% of the models were poor and 22% were good. Selecting the best models using performance measures alone would have resulted in different results. It was evident that a model could have high values for performance measures but could conversely be evaluated by an expert as poor or unrealistic, for example, the models in Figures 2F and 3A. Based on the results of this study, a triage approach is recommended to select the best ecological niche model using (1) measures of predictive accuracy, (2) performance measures, and (3) expert evaluation as the final discriminatory measure.

Expert evaluation was also found to be important in evaluating data partitioning approachs. Expert evaluation determined that both approaches had errors of commission in that they overpredicted into areas of northwestern Alberta or northern Saskatchewan where Purple Loosestrife would not be expected to become established or disperse. The GARP algorithm has been reported to overpredict (Peterson et al. 2007); however, the choice of predictive variables and data partitioning methods influences the degree of overprediction.

In this study, expert evaluation concluded that using stratified random sampling to partition the testing and training data produced more realistic models than the quintile data partitioning approach. This conclusion is also supported by performance measures using the independently withheld dataset (Tables 2 and 3), where the overall mean independent tests of accuracy for the quintile and stratified random sampling approach were 80% and 99%, respectively.

Selecting the best overall model

Using expert evaluation as the final discriminatory measure, we determined that the best overall model used the suite of climate variables and growing degree-days as the predictive variables and stratified random sampling to partition the data (Figure 3E). The model had good performance measures, including low errors of omission and commission (Table 3). The potential distribution of Purple Loosestrife generally follows the extent of the Prairies ecozone (Ecological Stratification Working Group 1996) in Canada. For example, the potential distribution of Purple Loosestrife is constrained by the Rocky Mountains in western Alberta and by the Boreal Plains ecozone in all three provinces. Areas predicted as 100% probable for invasion (e.g., all 10 models agree) follow the 49th parallel across the prairie provinces and a semicircular pattern north to Strathmore (Alberta), Prince Albert (Saskatchewan), and Yorkton (Saskatchewan), and then east to the southeastern part of Manitoba. The model indicated that there is suitable habitat in the Prairies for Purple Loosestrife to continue to expand its distribution northward.

Practical applications of the model

Spatial and temporal characterizations of risk or models of potential distribution are required in order to develop strategies to respond to an invasive plant (Venette et al. 2010). Models of potential distribution are very useful in preparing response strategies for invasive plants, as limited resources can then be prioritized for prevention, eradication, or control strategies (Waage and Mumford 2008). Prevention is the preferred strategy, and predictive models provide the spatial information required to develop response strategies. Therefore, based on the predictive map developed in this study, we make the following recommendations: (1) to prevent Purple Loosestrife from becoming established in areas of the Prairies predicted by the model, authorities should develop regulations to prohibit horticultural sales of Purple Loosestrife (to prevent human-mediated dispersal); (2) provinces should develop regional programs that target either the eradication or the containment of localized populations; and (3) provinces should focus early detection programs on areas predicted as suitable by the model into which Purple Loosestrife has not yet dispersed or where Purple Loosestrife has not yet become established.

The spatial predictive model can be used to optimize early detection programs by identifying high-risk areas for surveillance, leading to efficient allocation of survey resources. For example, the model suggests that early detection efforts should be directed to areas near Grande Prairie, Alberta (i.e., an area predicted by the model as having suitable habitat for Purple Loosestrife but where Purple Loosestrife has not yet become established). A city about 460 km northwest of Edmonton, Grande Prairie has a population of over 50 000, and garden centres there may retail ornamental plants that could provide dispersal pathways. The area also has wetlands, reservoirs, and rivers that would provide suitable aquatic habitat if an ornamental planting of Purple Loosestrife escaped. Early detection strategies should also consider using field naturalists in their survey efforts.

Conclusions

Based on the results of this study, we conclude that GARP is a useful tool that can be successfully used to model the potential distribution of invasive plants, in this case, Purple Loosestrife. Our model indicates that although Purple Loosestrife has been established in the Prairies for some time, there is considerable potential for further invasion. It will be interesting to observe how the distribution of Purple Loosestrife changes over the coming years. Changes in the distribution may support the GARP models that we determined overpredicted the potential distribution or they may support the models that we selected as good models. It could be argued, for example, that the overpredicted areas represent areas where Purple Loosestrife will become established and spread. In predicting the potential distribution of an invasive plant, it may be wise to err on the side of caution and accept a reasonable amount of error of commission, which may represent geographic space into which a species simply has not yet dispersed (Jimenez-Valverde et al. 2011).

Documents Cited (marked * in text)

Ali, S., and **C. Verbeek.** 1999. The Alberta Purple Loosestrife Eradication Program 1999 Status Report. Alberta Agriculture, Food and Rural Development, Edmonton, Alberta.

Bella, E. M. 2009. Invasive plant species response to climate change in Alaska: bioclimatic models of current and predicted future ranges. Report prepared for the U.S. Fish and Wildlife Service, Anchorage, Alaska. 33 pages.

White, D. J., E. Haber, and **C. Keddy.** 1993. Invasive Plants of Natural Habitats in Canada: An Integrated Review of Wetland and Upland Species and Legislation Governing Their Control. Report prepared for the Canadian Wildlife Service, Environment Canada, Ottawa, Ontario.

Literature Cited

Adjemian, J. C. Z., E. H. Girvetz, L. Beckett, and **J. E. Foley.** 2006. Analysis of genetic algorithm for rule-set production (GARP) modeling approach for predicting distributions of fleas implicated as vectors of plague, *Yersinia pestis*, in California. Journal of Medical Entomology 43: 93–103.

Allouche, O., A. Tsor, and **R. Kadmon.** 2006. Assessing the accuracy of species distribution models: prevalence, kappa and the true sill statistic (TSS). Journal of Applied Ecology 43: 1223–1232.

Anderson, R. P., and **A. Raza.** 2010. The effect of the extent of the study region on GIS models of species geographic distributions and estimates of niche evolution: preliminary tests with montane rodents (genus *Nephelomys*) in Venezuela. Journal of Biogeography 37: 1378–1393.

Anderson, R. P., M. Gomez-Laverde, and **A. T. Peterson.** 2002. Geographical distributions of spiny pocket mice in South America: insights from predictive models. Global Ecology and Biogeography 11: 131–141.

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

Anderson, R. P., A. T. Peterson, and **S. L. Egbert.** 2006. Vegetation-index models predict areas vulnerable to purple loosestrife (*Lythrum salicaria*) invasion in Kansas. Southwestern Naturalist 51: 471–480.

Andrewartha, H. G., and **L. E. Birch.** 1954. The Distribution and Abundance of Animals. University of Chicago Press, Chicago, Ill. 782 pages.

Austin, M. P., L. Belbin, J. A. Meyers, M. D. Doherty, and **M. Luoto.** 2006. Evaluation of statistical models for predicting plant species distributions: role of artificial data and theory. Ecological Modelling 199: 197–216.

Barve, N., V. Barve, A. Jimenez-Valverde, A. Lira-Noriega, S. P. Maher, A. T. Peterson, J. Soberon, and **F. Villalobos.** 2011. The crucial role of the accessible area in ecological niche modeling and species distribution modeling. Ecological Modelling 222: 1810–1819.

Bradley, B. A., M. Oppenheimer, and **S. A. Wilcove.** 2009. Climate change and plant invasions: restoration opportunities ahead? Global Change Biology 15: 1511–1521.

Brasier, C. M. 2008. The biosecurity threat to the UK and global environment from international trade in plants. Plant Pathology 57: 792–808.

Busby, J. R. 1986. A biogeographical analysis of *Nothofagus cunninghamii* (Hook.) Oerst. in southeastern Australia. Australian Journal of Ecology 11: 1–7.

Chen, P., E. O. Wiley, and **K. M. McNyset.** 2007. Ecological niche modeling as a predictive tool: silver and bighead carps in North America. Biological Invasions 9: 43–51.

Cole, D. E., J. R. King, D. A. Oyarzun, T. H. Dietzler, and **A. McClay.** 2007. Experiences with invasive plant management and ecology in Alberta. Canadian Journal of Plant Science 87: 1013–1022.

Cook, W. C. 1925. The distribution of the alfalfa weevil (*Phytonomus pesticus* Gyll.): a study in physical ecology. Journal of Agricultural Research 30: 479–491.

Cousens, R. D., R. W. Brown, A. B. McBratney, B. Whelan, and **M. Moerkerk.** 2002. Sampling strategy is important for producing weeds maps: a case study using kriging. Weed Science 50: 542–546.

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.

Dehnen-Schmutz, K., O. Holdenrieder, M. J. Jegar, and **M. Pautasso.** 2010. Structural change in the international horticultural industry: some implications for plant health. Scientia Horticulturae 125: 1–15.

Ecological Stratification Working Group. 1996. A national ecological framework for Canada. Agriculture and Agri-Food Canada, Research Branch, Centre for Land and Biological Resources Research, and Environment Canada, State of the Environment Directorate, Ecozone Analysis Branch, Ottawa.

Elith, J., C. Graham, R. Anderson, M. Dudik, S. Ferrier, A. Guisan, R. Hijmans, F. Huettmann, R. Leathwick, A. Lehmann, J. Lucia, G. Lohman, B. Loiselle, G. Manion, C. Moritz, M. Nakamura, Y. Nakazawa, J. Overton, A. Townsend Peterson, S. Phillips, K. Richardson, R. Scachetti-Pereira, E. Schapire, J. Soberon, S. Williams, M. Wisz, and **N. E. Zimmerman.** 2006. Novel methods improve prediction of species distributions from occurrence data. Ecography 29: 129–151.

Evangelista, P. H., S. Kumar, T. J. Stohlgren, C. S. Jarvevich, A. W. Crall, J. B. Norman III, and **D. T. Barnett.** 2008. Modelling invasion for a habitat generalist and a specialist plant species. Diversity and Distributions 14: 808–817.

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

Follak, S. 2011. Potential distribution and environmental threat of *Pueraria lobata*. Central European Journal of Biology 6: 457–469.

Ganeshaiah, K. N., N. Barve, N. Nath, K. Chandrashekara, M. Swamy, and **R. Shaanker.** 2003. Predicting the potential geographic distribution of the sugarcane woolly aphid using GARP and *DIVA*-GIS. Current Science 85: 1526–1528.

Gaudet, C. L., and **P. A. Keddy.** 1988. A comparative approach to predicting competitive ability from plant traits. Nature 334: 242–243.

Gaudet, C. L., and **P. A. Keddy.** 1995. Competitive performance and species distribution in shoreline plant communities: a comparative approach. Ecology 76: 280–291.

Guisan, A., and **W. Thuiller.** 2005. Predicting species distributions: offering more than simple habitat models. Ecological Letters 8: 993–1009.

Hassan, Q., and **C. Bourque.** 2009. Potential species distribution of Balsam Fir based on the integration of biophysical variables derived with remote sensing and process-based methods. Remote Sensing 1: 393–407.

Hassan, Q., C. Bourque, F. Meng, and **W. Richards.** 2007. Spatial mapping of growing degree days: an application of MODIS-based surface temperatures and enhanced vegetation index. Journal of Applied Remote Sensing 1: 1–12.

Heikkinen, R. K., M. Luoto, R. Virkkala, R. G. Pearson, and **J. H. Korber.** 2007. Biotic interactions improve prediction of boreal bird distributions at macro-scales. Global Ecology and Biogeography 16: 754–763.

Helaouet, P., and **G. Beaugrand.** 2009. Physiology, ecological niches and species distributions. Ecosystems 8: 1235–1245.

Hernandez, P. A., C. H. Graham, L. L. Master, and **D. L. Albert.** 2006. The effect of sample size and species characteristics on performance of different species distribution modeling methods. Ecography 29: 773–785.

Hirzel, A., and **A. Guisan.** 2002. Which is the optimal sampling strategy for habitat suitability modelling? Ecological Modelling 157: 331–341.

Hutchinson, G. E. 1957. Concluding remarks. Population Studies: Animal Ecology and Demography, Cold Spring Harbour Symposium on Quantitative Biology 22: 415–457.

Intergovernmental Panel on Climate Change. 2001. Climate data archive. Geneva, Switzerland. http://www.ipcc.ch. (Accessed May 2010).

Jimenez-Valverde, A. T. Peterson, J. Soberon, J. M. Overton, P. Aragon, and **J. M. Lobo.** 2011. Use of niche models in invasive species risk assessments. Biological Invasions 13: 2785–2797.

Jodoin, Y., C. Lavoie, P. Villeneuve, M. Theriault, J. Beaulieu, and **F. Belzile.** 2008. Highways as corridors and habitat for the invasive common reed *Phragmites australis* in Quebec, Canada. Journal of Applied Ecology 45: 459–466.

Johansson, M. E., and **P. A. Keddy.** 1991. Intensity and asymmetry of competition between plant pairs of different degrees of similarity: an experimental study on two guilds of wetland plants. Oikos 60: 27–34.

Kapetsky, J. M., J. M. Hill, and **L. D. Worthy.** 1988. A geographic information system for catfish farming development. Aquaculture 68: 311–320.

Kearney, M., and **W. Porter.** 2009. Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. Ecology Letters 12: 334–350.

Kerns, B. K., B. J. Naylor, M. Buonopane, C. G. Parks, and **B. Rogers.** 2009. Modeling Tamarix (*Tamarix spp.*) habitat and climate change effects in the northwestern United States. Invasive Plant Science and Management 2: 200–215.

Kleijnen, J. P. 2009. Kriging metamodeling in simulation: a review. European Journal of Operational Research 192: 707–716.

Kritocos, D. J., and **R. P. Randal.** 2001. A comparison of systems to analyze potential weed distributions. Pages 61–79 *in* Weed Risk Assessment. *Edited by* R. H. Groves, F. D. Panetta, and J. G. Virtue. CSIRO Publishing, Collingwood, Victoria, Australia.

Landis, J. R., and **G. C. Koch.** 1977. The measure of observer agreement for categorical data. Biometrics 33: 159–174.

Lawson, A. N., R. C. Van Acker, and **L. F. Friesen.** 2006. Emergence timing of volunteer canola in spring wheat fields in Manitoba. Weed Science 54: 873–882.

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

Lindgren, C. J. 2003. A brief history of Purple Loosestrife, *Lythrum salicaria*, in Manitoba and its status in 2001. Canadian Field-Naturalist 117: 100–109.

Lindgren, C. J. 2012. Biosecurity policy and the use of geospatial predictive tools to address invasive plants: updating the risk analysis toolbox. Risk Analysis 32: 9–15.

Lindgren, C. J., and **D. Walker.** 2012. Growth rate, seed production, and assessing the spatial risk of *Lythrum salicaria* using growing degree-days. Wetlands 32: 885–893.

Lindgren, C. J., J. Corrigan, and **R. A. DeClerk-Floate.** 2001. *Lythrum salicaria* L., Purple Loosestrife (Lythraceae). Pages 383–390 *in* Biological Control Programmes in Canada, 1981–2000. *Edited by* G. Mason and J. T. Huber. CABI Publishing, Wallingford, U.K.

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

Madsen, J. D. 1999. Predicting the invasion of Eurasian watermilfoil into northern lakes. U.S. Army Corps of Engineers Waterways Experiment Station. Technical Report A-99-2, February 1999.

Mal, T. K., J. Lovett-Doust, L. Lovett-Doust, and **G. A. Mulligan.** 1992. The biology of Canadian weeds. 100. *Lythrum salicaria*. Canadian Journal of Plant Science 72: 1305–1330.

Mal, T. K., J. Lovett-Doust, and **L. Lovett-Doust.** 1997. Time-dependent competitive displacement of *Typha angustifolia* by *Lythrum salicaria*. Oikos 79: 26–33.

McMaster, G. S., and **W. W. Wilhelm.** 1997. Growing degree-days: one equation, two interpretations. Agricultural and Forest Meteorology 8: 291–300.

Messenger, P. S. 1959. Bioclimatic studies with insects. Annual Review of Entomology 4: 183–206.

Miller, P., W. Lanier, and **S. Brandt.** 2001. Using growing degree days to predict plant stages. Montana State University Extension Service, Bozeman, Montana.

Mullin, B. H. 1998. The biology and management of Purple Loosestrife (*Lythrum salicaria*). Weed Technology 12: 397–401.

Oberhauser, K., and **A. T. Peterson.** 2003. Modeling current and future potential wintering distributions of eastern North American monarch butterflies. Proceedings of the National Academy of Sciences of the United States of America 100: 14063–14068.

Osborne, P. E., and **S. Suárez-Seoane.** 2002. Should data be partitioned spatially before building large scale distribution models? Ecological Modelling 157: 249–259.

Ottenbreit, K. 1991. The distribution, reproductive biology, and morphology of *Lythrum* species, hybrids and cultivars in Manitoba. M.S. thesis, University of Manitoba, Winnipeg, Manitoba.

Ottenbreit, K., and **R. J. Staniforth.** 1994. Crossability of naturalized and cultivated *Lythrum* taxa. Canadian Journal of Botany 72: 337–341.

Pearson, R. G., and **T. P. Dawson.** 2003. Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? Global Ecology and Biogeography 12: 361–371.

Peterson, A. T. 2001. Predicting species' geographic distributions based on ecological niche modeling. Condor 103: 599–605.

Peterson, A. T. 2011. Ecological niche conservatism: a time-structured review of evidence. Journal of Biogeography 38: 817–827.

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

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

Peterson, A. T., and **J. Shaw.** 2003. *Lutzomyia* vectors for *Cutaeous leishmaniasis* in southern Brazil: ecological niche models, predicted geographic distributions, and climate change effects. International Journal for Parasitology 33: 919–931.

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

Peterson, A. T., R. Williams, and **G. Chen.** 2007. Modeled global invasive potential of Asian gypsy moths, *Lymantria dispar*. Entomologia Experimentalis et Applicata 125: 39–44.

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

Peterson, A. T., J. Soberon, R. G. Pearson, P. Anderson, E. Martinez-Meyer, M. Nakamura, and **M. Bastos Araujo.** 2011. Ecological Niches and Geographic Distributions. Princeton University Press, Princeton, New Jersey. 316 pages.

Pheloung, P. C., P. A. Williams, and **S. R. Halloy.** 1999. A weed risk assessment model for use as a biosecurity tool evaluating plant introductions. Journal of Environmental Management 57: 239–251.

Phillips, S. J., R. P. Anderson, and **R. E. Schapire.** 2006. Maximum entropy modeling of species geographic distributions. Ecological Modelling 190: 231–259.

Raxworthy, C. J. C. M. Ingram, N. Rabibisoa, and **R. G. Pearson.** 2007. Applications of ecological niche modeling for species delimitation: a review and empirical evaluation using day geckos (*Phelsuma*) from Madagascar. Systematic Biology 56: 907–923.

Sanchez-Flores, E. 2007. GARP modeling of natural and human factors affecting the potential distribution of the invasives *Schismus arabicus* and *Brassica tournefortii* in "El Pinacate y Gran Desierta de Alter" Biosphere Reserve. Ecological Modelling 204: 457–474.

Shamsi, S. R. A., and **F. H. Whitehead.** 1974. Comparative eco-physiology of *Epilobium hirsutum* L. and *Lythrum salicaria* L. I. General biology, distribution and germination. Journal of Ecology 62: 279–290.

Shrestha, A., B. D. Hanson, W. Fidelibus, and **M. Alcorta.** 2010. Growth, phenology, and intraspecific competition between glyphosate-resistant and glyphosate-susceptible horseweeds (*Conyzes canadensis*) in the San Joaquin Valley of California. Weed Science 58: 147–153.

Soberon, J., and **A. T. Peterson.** 2005. Interpretation of models of fundamental ecological niches and species' distributional areas. Biodiversity Informatics 2: 1–10.

Stockwell, D. R. B. 1997. Generic predictive systems: an empirical evaluation using the learning base system (LBS). Expert Systems with Applications 12: 301–310.

Stockwell, D. R. B. 2000. The GARP modelling system: problems, solutions and automated spatial prediction. International Journal of Geographic Information Science 13: 143–158.

Stockwell, D. R. B., and **I. R. Noble.** 1992. Induction of sets of rules from animal distribution data: a robust and informative method of data analysis. Mathematics and Computers in Simulation 33: 385–390.

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

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

Sumners, W. H. 2005. Exotic plant species in the mixed-wood section of the southern boreal forest of Saskatchewan. M.Sc. thesis, University of Saskatchewan, Saskatoon, Saskatchewan. 138 pages.

Sutherst, R, G. F. Maywarld, and **B. L. Russell.** 2000. Estimating vulnerability under global change: modular modelling of pests. Agriculture, Ecosystems and Environment 82: 303–319.

Sutherst, R. W. 2003. Prediction of species geographical ranges. Journal of Biogeography 30: 805–816.

Sutherst, R. W., and **G. F. Maywald.** 1985. A computerised system for matching climates in ecology. Agriculture, Ecosystems and Environment 13: 281–299.

Syartinilia, S., and **S. Tsuyuki.** 2008. GIS-based modeling of the Javan Hawk-Eagle distribution using logistic and autologistic regression models. Biological Conservation 141: 756–769.

Tarkesh, M., and **G. Jetschke.** 2012. Comparison of six correlative models in predictive vegetation mapping on a local scale. Environmental and Ecological Statistics 19: 437–457.

Therriault, T. W., and **L. Herborg.** 2008. A qualitative biological risk assessment for vase tunicate *Ciona intestinalis* in Canadian waters: using expert knowledge. ICES Journal of Marine Science 65: 781–787.

Thompson, D. Q., R. L. Stuckey, and **E. Thompson.** 1987. Spread, Impact and Control of Purple Loosestrife (*Lythrum salicaria*) in North American Wetlands. U.S. Fish and Wildlife Service, Fish and Wildlife Research 2. 55 pages.

Thuiller, W. 2003. BIOMOD—optimizing predictions of species distributions and projecting potential future shifts under global change. Global Change Biology 9: 1353–1362.

Venette, R., D. Kriticos, D. Magarey, F. Koch, R. Baker, S. Worner, N. Raboteau, D. McKenny, E. Dobesberger, D. Yemshanov, P. De Barro, W. Hutchison, G. Fowler, T. Kalaris, and **J. Pedlar.** 2010. Pest risk maps for invasive alien species: a roadmap for improvement. BioScience 60: 349–362.

Waage, J. K., and **J. D. Mumford.** 2008. Agricultural biosecurity. Philosophical Transactions of the Royal Society B: Biological Sciences 363: 863–876.

Welk, E. 2004. Constraints in range predictions of invasive plant species due to non-equilibrium distribution patterns:

purple loosestrife (*Lythrum salicaria*) in North America. Ecological Modelling 19: 551–567.

Welk, E., K. Schubert, and **M. H. Hoffmann.** 2002. Present and potential distribution of invasive garlic mustard (*Alliaria petiolata*) in North America. Diversity and Distributions 8: 219–233.

Wiley, E. O., K. M. McNyset, A. T. Peterson, C. R. Robins, and **A. M. Stewart.** 2003. Niche modeling and geographic range predictions in the marine environment using a machine-learning algorithm. Oceanography 16: 120–127.

Yemshanov, D., F. H. Koch, Y. Ben-Haim, and **W. D. Smith.** 2010. Robustness of risk maps and survey networks to knowledge gaps about a new invasive pest. Risk Analysis 30: 261–276.

Zalba, S. M., M. I. Sonaglioni, C. A. Compagnoni, and **C. J. Belenguer.** 2000. Using a habitat model to assess the risk of invasion by an exotic plant. Biological Conservation 93: 203–208.

Zhu, L., O. Sun, W. Sang, L. Zhenyu, and **K. Ma.** 2007. Predicting the spatial distribution of an invasive plant species (*Eupatorium adenophorum*) in China. Landscape Ecology 22: 1143–1154.

Received 27 June 2012
Accepted 4 December 2012