Annals of the Entomological Society of America, 113(2), 2020, 100–113
doi: 10.1093/aesa/saz049
Advance Access Publication Date: 11 February 2020
Special Collection

Special Collection: Geospatial Analysis of Invasive Insects

OXFORD

Minimizing Risk and Maximizing Spatial Transferability: Challenges in Constructing a Useful Model of Potential Suitability for an Invasive Insect

Amy C. Morey[1,3] and Robert C. Venette[2]

[1]Department of Entomology, University of Minnesota, 1980 Folwell Avenue, St. Paul, MN 55108-6125, [2]USDA, Forest Service, Northern Research Station, 1561 Lindig Street, St. Paul, MN 55108-6125, and [3]Corresponding author, e-mail: morey041@umn.edu

Subject Editor: Kevin Macaluso

Received 26 April 2019; Editorial decision 19 August 2019

Abstract

Forecasting the spread and potential impacts of invasive, alien species is vital to relevant management and policy decisions. Models that estimate areas of potential suitability are useful to guide early detection and eradication, inform effective budget allocations, and justify quarantine regulations. Machine-learning is a rapidly emerging technology with myriad applications, including the analysis of factors that govern species' distributions. However, forecasts for invasive species often require extrapolation into novel spaces, which may severely erode model reliability. Using the popular machine-learning platform, MaxEnt, we integrate numerous tools and recommendations to demonstrate a method of rigorous model development that emphasizes assessment of model transferability. Our models use *Lymantria dispar dispar* (L.) (Lepidoptera: Erebidae), an insect brought to the United States in the late 1860s from Europe and subsequently well monitored in spread. Recent genetic analyses provide evidence that the eastern North American population originated in Germany, France, and northern Italy. We demonstrate that models built and assessed using typical methodology for invasive species (e.g., using records from the full native geographic range) showed the smallest extent of extrapolation, but the worst transferability when validated with independent data. Conversely, models based on the purported genetic source of the eastern North American populations (i.e., a subset of the native range) showed the greatest transferability, but the largest extent of extrapolation. Overall, the model that yielded high transferability to North America and low extrapolation was built following current recommendations of spatial thinning and parameter optimization with records from both the genetic source in Europe and early North American invasion.

Key words: MaxEnt, species distribution modeling, pest risk mapping, external validation, continuous Boyce Index

The goal of modeling and mapping invasive species distributions often is to visualize areas threatened by a given species. Forecasts of the geographic dynamics of invasive species inform management and policy decision-making, and are a vital component of pest risk assessment (Venette et al. 2010). Maps and models do not necessarily depict comprehensive risk, but may contribute components to its overall assessment (e.g., likelihood of establishment, habitat suitability, pathway analysis, potential impacts) (Venette et al. 2010). A common challenge with modeling invasion risk is the scarcity of biological and ecological data for species new to a region. Little more than observed location data are typically available, which can themselves be sparse. As a consequence, correlative models are commonly used to quantify relationships between an invasive species' known presence (or presence and absence, or abundance) and environmental variables in the those areas (Elith 2017). Then, inference is made about a species occurrence in other regions where similar environmental correlations may occur in space and time. The growing availability of large, open biodiversity and environmental data stores and increased computing power have fueled a rapid rise in the development and use of correlative models (i.e., species distribution models) (Zimmermann et al. 2010, Guisan et al. 2017a).

Currently, MaxEnt is among the most popular methods for species distribution modeling and has been shown to perform well compared to alternatives (Elith et al. 2006, Heikkinen et al. 2012, Venette 2017). Its applications have varied (Elith et al. 2011) and it is increasingly used to forecast future distributions of invasive species (e.g., Sobek-Swant et al. 2012, Zhu et al. 2012, Dos Santos et al. 2017). MaxEnt is a machine-learning algorithm that was specifically developed to approximate an unknown distribution for scenarios where only locations of presence are known (Phillips et al. 2006,

Published by Oxford University Press on behalf of Entomological Society of America 2020.
This work is written by (a) US Government employee(s) and is in the public domain in the US.

Downloaded from https://academic.oup.com/aesa/article/113/2/100/5727917 by U S Dept of Agriculture user on 02 October 2020

Dudik et al. 2007). It uses the information theory of maximum entropy to parse out differences between environmental conditions where the species is present from those at background locations where the species presence is unknown. Further description on the detail and underlying functioning of MaxEnt is well summarized elsewhere (Phillips et al. 2006, Elith et al. 2011, Merow et al. 2013).

With MaxEnt's expanding usage, several a large increase in extensions to its original default user interface have been developed, including support tools and new features (e.g., Warren et al. 2010, Muscarella et al. 2014, Brown et al. 2017, Hijmans et al. 2017, Phillips et al. 2017, Kass et al. 2018). These developments are in part due to the finding that the default settings of the program should not be assumed appropriate (Anderson and Gonzalez 2011, Rodda et al. 2011, Syfert et al. 2013, Yackulic et al. 2013, Radosavljevic and Anderson 2014). Multiple recommendations and critiques have emerged, but how to most appropriately implement MaxEnt for many of its myriad applications remains an active area of discussion and research.

Developing MaxEnt models for novel environments, specifically with respect to invasive species and climate change, remains philosophically and methodologically challenging (Elith et al. 2010, Peterson et al. 2011, Elith 2017). For invasive species, MaxEnt users frequently estimate new occurrences among other known presences (interpolation; e.g., Crall et al. 2013, Steen et al. 2019) or identify new geographic areas that might be environmentally suitable (extrapolation; e.g., Zhu et al. 2012, Kumar et al. 2014). The paramount assumption in both applications is that presence data have come from a random sample of the environment with a consistent probability of detecting the species where it occurs (Merow et al. 2013). Arguably, this assumption is more likely to be satisfied with presence data from the native range where the species might be presumed to be in equilibrium with its environment (Elith et al. 2010). A related assumption is that presence data used to develop the model reflect conditions to which the model will be applied. Using data from the full native range for model development, as is often done for invasive species, assumes not only that the native range niche is conserved with invasion (Broennimann et al. 2007), but that the dynamics of an invading population would be unaffected by any previous genetic divergence (and possible phenotypic variation) within the specific region of the native range where the population originated. To our knowledge, the extent to which the latter is a robust assumption not been investigated in the context of species distribution modeling, particularly for invaded populations that result from only one or a few introductions. In contexts where records from the invaded range exist, whether to include these records remains debatable (Elith 2017), with the common argument that these presences better approximate a species' fundamental niche and future invasion potential while presence records from the native range better reflect the realized niche.

Performance assessment of MaxEnt models for invasive species in novel environments is also challenging. Transferability, the performance of a model when projected into different areas of space and/or time, is critical to the reliability and usefulness of a model (Phillips 2008, Venette et al. 2010). Model transferability can fail for multiple reasons, such as from effects of predictor variable selection (Petitpierre et al. 2017) and violation of the assumption of niche conservatism (Broennimann et al. 2007). Statistical measures such as area under the receiver operating curve (AUC), kappa statistics, or explained deviance remain common to assess transferability though are arguably inappropriate in cases of extrapolation (Elith et al. 2011, Peterson et al. 2011). When the goal is to model the suitability of a new geographic space for an invasive species, measures of performance from typical training and test data may not be informative (Elith 2017). Evaluation using data that are fully independent of model development or verification is the most robust assessment of model results (Fielding and Bell 1997, Araujo et al. 2005). Often, models that perform best in model development regions do not maintain their rank in terms of accuracy in new areas, particularly for models considered to be more complex such as MaxEnt (Guisan et al. 2017b). A recent example shows that modestly informative MaxEnt models (AUC = 0.6–0.7) may perform well (i.e., high True Test Statistic) when applied to new locations (Briscoe Runquist et al. 2019), and another demonstrates how common evaluation metrics favor models overfit to the input data and penalize those that are more transferable to new regions (Fourcade et al. 2018).

Here, we add to the otherwise sparse empirical assessments of spatial transferability for models of invasion risk (Araujo et al. 2005) using an insect, the European gypsy moth *Lymantria dispar dispar* (L.). Native to Europe, this subspecies was intentionally introduced by Leopold Trouvelot to the east coast of the United States in 1868 or 1869 and has since continued to spread (Liebhold et al. 1989, Tobin et al. 2012). Historical records do not indicate where precisely Trovelot obtained *L. dispar dispar* (Liebhold et al. 1989). Genetic analysis of the eastern North American population suggests that *L. dispar dispar* was only introduced once from a region spanning portions of France, Germany, and northern Italy, and that subsequent introductions, if any, had little effect on the genetic composition of current populations (Wu et al. 2015).

With a well-documented extent of occurrence in the United States and Canada, *L. dispar dispar* offers the under-utilized opportunity with invasive species to assess the transferability of species distribution models with spatially independent data from an invaded range (Barbet-Massin et al. 2018). In this study, we rely on temporally shifted assessments of the invasion of *L. dispar dispar* in eastern North America. We investigated which MaxEnt models built with different scenarios of species information (i.e., different portions of the native range and inclusion of the invaded range before 1960) and model construction (i.e., default or optimized settings) performed best at forecasting areas that proved to be climatically suitable for establishment (i.e., newly invaded areas in the North America from 1990 to 2016).

More specifically, we assessed how well the purported genetic source of invasion from the European range informs a portion of the current invaded range in eastern North America, and whether inclusion of a separate portion of the invaded range in model development improves forecasts. Invading populations often result from small founder events that represent a small portion of the genotypic and phenotypic diversity found in the native range (Lee 2002, Dlogosch and Parker 2008). If there is genotypic and phenotypic structuring among *L. dispar dispar* locations within its native range (Wu et al. 2015), then focusing model development on the geographic region that is the most likely source of the North American population should increase transferability to the future invaded range. Including portions of the invaded range in model development was also expected to improve model performance in North America. Previous studies have concluded that species distribution models for invasive species benefit from having representation of the invaded range (Mau-Crimmins et al. 2006, Broennimann and Guisan 2008, Beaumont et al. 2009) because it better reflects the conditions in which the species can occur than the native range alone. However, results to the contrary have also been found (Vaclavık and Meentemeyer 2012, Barbet-Massin et al. 2018) and evaluations using MaxEnt and invasive insects are limited (but see Sobek-Swant et al. 2012).

Downloaded from https://academic.oup.com/aesa/article/113/2/100/5727917 by U S Dept of Agriculture user on 02 October 2020

We also compared how the extent of climatic extrapolation changed with model development dataset. Transferring a model to a different space or time can be unreliable when there is extrapolation; the modeled response of a species to the predictor variables and their interactions can change in new conditions (Elith et al. 2010, Petitpierre et al. 2017). Areas of extrapolation, therefore, also serve to identify regions of uncertainty in model forecasts. Here, we anticipated that the total area of extrapolation would be smaller for models that involved a larger geographic extent of occurrence records. This expectation is based on the assumption that a wider array of environmental conditions may be sampled in these instances.

Materials and Methods

Occurrence Records and Background Selection

Georeferenced occurrence records of *L. dispar dispar* in Europe and North America were obtained from the Global Biodiversity Information Facility (GBIF; www.gbif.org), EDDmapS (www.eddmaps.org), the primary literature, and the Slow the Spread program (STS; Sharov et al. 2002). Records were inspected and removed where needed based on coordinate errors, imprecise coordinates, uncertain sources, or duplicate information. No detections of *L. dispar* from western North American were included because the subspecies involved and establishment status were uncertain. The STS database tracked the invasion front of *L. dispar dispar* in the eastern United States and contained high-density trapping records from 1994 through 2016 (predominantly after 2000, when the program formally began). The STS records used for analysis were limited to those sites where 10 or more moths were captured at a time, which is a trap density considered indicative of an established population (Tobin et al. 2004).

We distinguished four groups of distribution records from the above data (Fig. 1). The first group (n = 820; all blue points in Fig. 1A), designated "Native," were those records considered part of native range in Europe (i.e., west of the Ural Mountains; Hajek and Tobin 2009). Second, a subset (n = 249; light blue points in Fig. 1A) of the native range, designated "Source," was made to include only those records from Germany, France, and Italy that represent the likely geographic source of the eastern North American population (Wu et al. 2015). A third group (n = 2,761; red points in Fig. 1B), designated "Invaded," was occurrence records in North America from 1960 and earlier; this dataset was included when the invaded range was represented in model development (see Model Tuning and Development below). This region was considered generally infested with *L. dispar dispar* (Liebhold et al. 1997), so we assumed the insect would exhibit widespread, dense presence similar to regions formally documented by STS. Using ArcMap (ESRI ArcGIS Desktop 10.6, Redlands, CA), we constructed a 10-km grid within the region considered infested (Liebhold et al. 1997) and selected the centroid of each cell. The final group of points was composed of North American occurrence records that were west of the U.S. invasion front in 1990 (Liebhold et al. 1997). Occurrences in the United States came from the STS program and records for Canada came from the literature and EDDMapS. The combined records were spatially thinned to a minimum distance of 10 km, resulting in a validation dataset of 101,300 records (orange points in Fig. 1B); this dataset was spatially and temporally independent of all other datasets and was used only for model validations.

We limited selection of background sites to a minimum convex polygon (MCP) with a 10 km buffer, the approximate annual distance spread by gypsy moths (Tobin et al. 2015), around the occurrence points for a given region (polygons in Fig. 1) as recommended by Jarnevich and Young (2015). A random sample of 10,000 total locations (i.e., the default for MaxEnt) was taken from these reduced areas to serve as background locations for a given model. For models developed from space in both the native and invaded range, 10,000 background locations were split between each region's MCP, based roughly on the relative size of the MCPs to each other (i.e., 50:50 or 70:30; Table 1).

Environmental Data

We used 19 gridded BIOCLIM variables derived from temperature and precipitation climate summaries (1970–2000) in the WorldClim dataset (v. 2; www.worlclim.org/bioclim) (Hijmans et al. 2005) at 2.5 arc-minute resolution. The climate rasters were clipped to the extent of a given buffered MCP using the R package "raster" (v. 2.6-7) (Hijmans 2017). Then, to reduce the potential confounding effects of collinearity among climate variables (Dormann et al. 2013), we

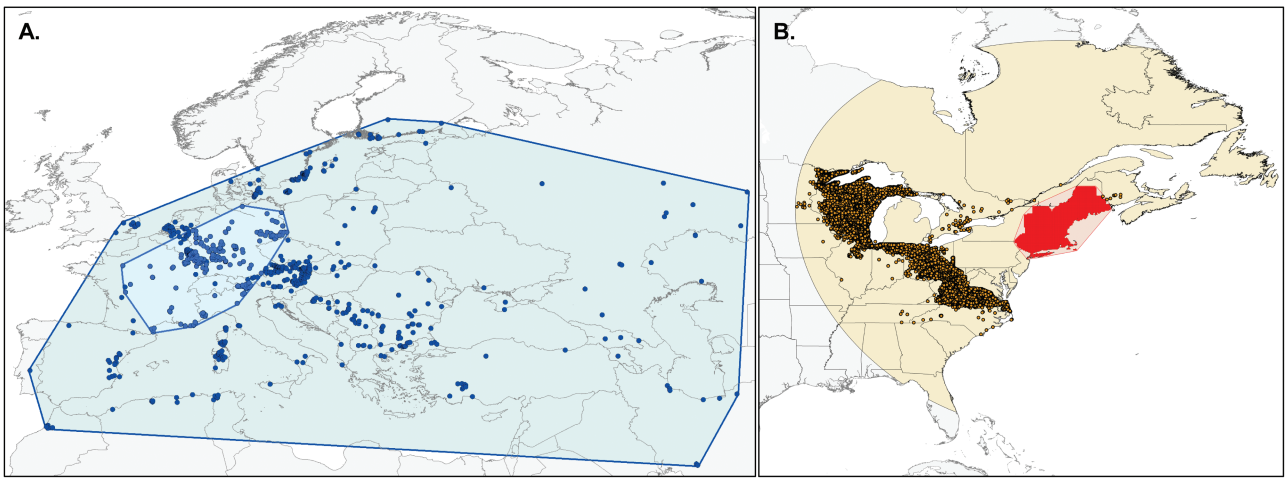

Fig. 1. Records of occurrence for *L. dispar dispar* used for model development (blue and light blue in [A] and red in [B]) and validation (orange in [B]). Colored polygons correspond to the space from where background locations were randomly sampled for a given model: blue was used in the "Native" and "Native + Invaded" models, light blue was used in the "Source" and "Source + Invaded" models, and red was used in the "Native + Invaded" and "Source + Invaded" models. The tan ellipse indicates the space within which the independent validation metric for each model was calculated. Maps are unprojected.

Downloaded from https://academic.oup.com/aesa/article/113/2/100/5727917 by U S Dept of Agriculture user on 02 October 2020

Table 1. Model data and parameters used in MaxEnt models of *L. dispar dispar*

Primary model	Environmental variables[a]	Background records (number; region)	Submodel	*n*	Thinning distance (km)	Regularization multiplier	Feature class(es)
Native	BIO_5, BIO_6, BIO_8, BIO_{12}, BIO_{15}	10,000; Native MCP	A	820	Unthinned	Default	*Default*
			B	820	Unthinned	1.0	Linear
			C	188	69	*Default*	*Default*
			D	188	69	4.5	Hinge
Native + Invaded	BIO_5[b], BIO_6, BIO_7, BIO_8, BIO_{12}, BIO_{15}	7,000; Native MCP	A	3,581	Unthinned	*Default*	*Default*
			B	3,581	Unthinned	2.5	Linear
		3,000; Invaded MCP	C	224	69	*Default*	*Default*
			D	224	69	4.0	Linear, quadratic, hinge, product
Source	BIO_3, BIO_4[b], BIO_5, BIO_8, BIO_9, BIO_{12}, BIO_{15}	10,000; Source MCP	A	249	Unthinned	*Default*	*Default*
			B	249	Unthinned	6.0	Linear
			C	84	32	*Default*	*Default*
			D	84	32	4.5	Linear, quadratic
Source + Invaded	BIO_3[b], BIO_5, BIO_6[c], BIO_8[b,c], BIO_9[c], BIO_{12}[c], BIO_{15}[c]	5,000; Source MCP	A	3,010	Unthinned	*Default*	*Default*
			B	3,010	Unthinned	6.0	Linear
		5,000; Invaded MCP	C	244	32	*Default*	*Default*
			D	244	32	4.5	Hinge

Environmental variables (www.worldclim.org/bioclim) and region(s) within which background locations (10,000 total for each model) were randomly generated were constant within each primary model. See Fig. 1 for depictions of each minimum convex polygon (MCP). Model names refer to the dataset used to develop the model, with *n* occurrences used in each submodel (A–D).

[a]BIO_3 = isothermality; BIO_4 = temperature seasonality; BIO_5 = max. temperature of warmest month; BIO_6 = min. temperature of coldest month; BIO_7 = temperature annual range; BIO_8 = mean temperature of wettest quarter; BIO_9 = mean temperature of driest quarter; BIO_{12} = annual precipitation; BIO_{15} = precipitation seasonality.

[b]Coefficient (i.e., lambda) of zero in the unthinned, optimized tuning model.

[c]Coefficient (i.e., lambda) of zero in the thinned, optimized tuning model.

calculated a Pearson correlation coefficient matrix within each MCP space or pair of MCPs. Only variables with correlations of $|r| < 0.70$ were retained. When selecting variables to remove in instances of correlation, those of combined temperature and precipitation were removed first due to cautions noted by others (e.g., Elith et al. 2013). The final climate variables provided to each model are listed in Table 1. MaxEnt provides coefficients (i.e., lambda values) for the variables used in each model (data not shown). Variables with lambda values of zero (i.e., not contributing to MaxEnt results) were also noted in Table 1.

Model Tuning and Development

We built 16 models for *L. dispar dispar* using MaxEnt version 3.3.3k (Phillips et al. 2006) (http://biodiversityinformatics.amnh.org/open_source/maxent). An overview of procedures is provided in Table 1. Processing and additional analyses were conducted in R version 3.5.2 (R Core Team 2018), RStudio version 1.1.456 (RStudio 2016), and ArcMap.

Four primary models, based on the aforementioned distribution datasets, were produced to compare the effects of modeling using two extents of occurrence within the native European range ("Native" or "Source") and with/without the invaded North American range by 1960 (Table 1). Within each primary model, four submodels were made with different spatial thinning distances (unthinned or thinned) and tuning parameter (default or optimal) combinations, the details of which are described below.

The unthinned and default tuning models represented those practices frequently used in MaxEnt modeling (Morales et al. 2017) despite growing caution. Unthinned datasets were modeled with only one record allowed per grid cell (i.e., the default "remove duplicate presence records" setting in MaxEnt), but were not otherwise modified. In contrast, "thinned" datasets were reduced in size prior to modeling to reduce the confounding effects of sampling bias on model outputs (Aiello-Lammens et al. 2015). Records were thinned by using the R package "spThin," with a thinning distance set to the expected mean distance for a random distribution among the presence points as estimated through Average Nearest Neighbor analysis (ArcMap). The resulting dataset maximizes the number of records retained for a given thinning distance (Aiello-Lammens et al. 2015). Thinning distances were always based on occurrence records from Europe as the European records covered larger areas than those in North America.

The default settings in MaxEnt can potentially result in overfit and poor performing models so species-specific tuning is highly recommended (see Anderson and Gonzalez 2011 and Merow et al. 2013 for further detail and explanation). Briefly, tuning involved selecting values for two of the major settings in MaxEnt, feature classes and the regularization multiplier, which can greatly influence the fitted models. Feature classes are various transformations of the environmental predictor variables that can allow MaxEnt to fit complex, non-linear relationships to the data provided (Merow et al. 2013). The current default is to allow all feature classes except "threshold" features. To reduce overfitting (i.e., when a model fits only the data provided and nothing else), MaxEnt uses regularization to penalize models that include parameters that contribute little information. The intensity of regularization can be increased from the default value of 1.0 by specifying a multiplier that is applied to all regularization values. Optimized settings for the regularization multiplier and feature class combination were selected here by using the "ENMeval" R package (Muscarella et al. 2014) to generate candidate models with regularization values between 1 and 6 (inclusive; 0.5 increments), in combination with the features classes linear, linear + quadratic, hinge, linear + quadratic + hinge, and linear + quadratic

Downloaded from https://academic.oup.com/aesa/article/113/2/100/5727917 by U S Dept of Agriculture user on 02 October 2020

+ hinge + product to represent a range of complexity of *L. dispar dispar* response to spatial variation in the environmental variables. As per Phillips et al. (2017), we did not include "threshold" features in any combination because they are unnecessarily complex and likely unrealistic. "ENMeval" allows multiple methods to partition data into training and testing bins to generate cross-validation statistics. We ran candidate models with "block" spatial partitioning, as recommended when model transferability is of interest (Muscarella et al. 2014). Block partitioning splits occurrence and background localities into four bins based on latitude and longitude, training a model using three of the bins and then testing the results with the fourth. It repeats this four times, using a different bin for testing each time, and provides cross-validation statistics for the average of the four models (Muscarella et al. 2014).

Based on the evaluations metrics from "ENMEval," we used sequential criteria (e.g., Shcheglovitova and Anderson 2013, Galante et al. 2018) to select the optimal settings (i.e., feature classes and regularization multiplier) to be used for each final MaxEnt model. First, we prioritized models that did the best, on average, in predicting test occurrence records as suitable based on the 10% training presence threshold (Pearson et al. 2007) (i.e., prioritized the lowest average test OR_{10}). Then, if a tie, we prioritized the models that balanced goodness-of-fit with the least amount of complexity based on the sample size corrected Akaike information criteria (i.e., the lowest ΔAICc) (Warren and Seifert 2011).

The final default and optimal models were run in the MaxEnt graphical user interface using 5-fold cross-validation with default or optimized feature classes/regularization, respectively. Five-folds allowed for each replicate (fold) across all datasets to have at least 30 occurrence records used in training (Wisz et al. 2008). All other settings were default except the "fadebyclamping" argument was selected as an intermediately conservative extrapolation method (Phillips et al. n.d.), and "addsamplestobackground" was unselected to prevent occurrence records from being included as background samples.

Model Projections

Median logistic MaxEnt outputs (calculated from each 5-fold cross-validation model) were multiplied by 100, rounded to integers, and visualized in ArcMap to show global and region-specific areas of relative climate suitability for *L. dispar dispar*. Because the results from MaxEnt can be unreliable in novel climates (i.e., extrapolated to climate conditions that differ from those areas used to develop the model) (Elith et al. 2010), we identified geographic areas with model extrapolation by examining the Multivariate Environmental Similarity Surface (MESS). The MESS calculates how similar a given point is to a reference set of points (here, the known occurrence records) for a given climate variable, with values less than zero indicating locations where at least one variable was extrapolated (Elith et al. 2011). Not all inputted variables may ultimately contribute to a fitted MaxEnt model, so we produced the MESS maps by using only variables with non-zero lambda values (Phillips et al. n.d.).

For invasive species, forecasts into novel climate conditions may be of interest though such extrapolations are unadvisable from a statistical perspective. To communicate areas of uncertainty for a given model, we specified map regions with MESS values ≥0 as having no extrapolation, 0 > MESS ≥ −10 as having moderate extrapolation, and MESS < −10 as having high extrapolation.

Model Evaluation

The majority of distribution models for invasive species are evaluated based on internal validation, which uses resampling methods to test prediction within the model training region (Guisan et al. 2017a). To mimic this, we compiled conventional metrics, including AUC and an omission rate for the averaged cross-validation results of each model. The AUC is a threshold-independent performance measure that reflects the probability that a randomly chosen presence site will rank above a randomly chosen background site (Phillips et al. 2006). Values near 1.0 indicate high discriminatory ability, whereas values of 0.5 (or less) indicate discrimination no better than random (Elith 2000). Omission rates (ORs) calculate the proportion of test locations with suitability values lower than a specified threshold. The 10% threshold omission rate (OR_{10}) uses the smallest value after excluding the lowest 10% of training suitability values (Anderson and Gonzalez 2011). In ideal models, the expected OR_{10} is 10%, and values higher than expected suggest overfitting. The OR_{10} was chosen instead of the omission rate based on the minimum training threshold because it is less likely to be influenced by outlier occurrence locations (Radosavljevic and Anderson 2014). Additionally, the AICc was included for each model due to its potential as a useful model selection measure, particularly one less impacted by sampling bias (Warren and Seifert 2011, Muscarella et al. 2014, Galante et al. 2018).

To assess transferability, we externally validated performance for all models using the North American occurrences that were not used in any model development (orange points in Fig. 1B). The appropriateness of many commonly used evaluation metrics (including the above) for presence-only models like MaxEnt is questionable (Lobo et al. 2008, Jiménez-Valverde et al. 2011, Sofaer et al. 2019). One measure, the continuous Boyce Index (CBI), has been suggested to evaluate how well a presence-only model predicts the suitability of known occurrences compared to a random expectation within an area (Boyce et al. 2002, Guisan et al. 2017a, Petitpierre et al. 2017). To avoid the shortcomings of arbitrary selection of suitability classification ("bins"), the CBI uses a "moving window" that iteratively evaluates continuous predictions (Hirzel et al. 2006). The CBI is presented as the Spearman rank correlation coefficient between the suitability values and the predicted-to-expected ratio of evaluation points, with values close to 1 indicating good predictions, values near 0 indicating predictions no different from random, and value near −1 indicating poor predictions (i.e., low suitability areas contained the most occurrences) (Hirzel et al. 2006). We used the R package "ecospat" (Di Cola et al. 2017) to calculate the CBI for the dataset of independent occurrences in North America. This index is sensitive to the size of the area used for validation. We delineated the validation area based on an estimate of the space potentially available to the species at present. Given that the majority of the eastern North American population appears to the be result of spread from the initial introduction (Wu et al. 2015), we first estimated the radial distance from Medford, MA (the original site of introduction) to the STS site furthest from Medford with ≥10 captured moths and added an additional 10 km buffer to account for dispersal. As a result, all sites within 1,870 km of Medford were considered accessible to *L. dispar dispar* (tan ellipse in Fig. 1B). We recognize that this estimate of available space comes with an assumption that the processes driving the spread of *L. dispar dispar* are constant in space and time, which may be tenuous. We nonetheless feel it to be a reasonable assumption given lack of defensible alternatives.

Lastly, we compared the relative amount of extrapolated space across models to communicate forecast uncertainty within the area used for model validation (tan ellipse in Fig. 1B). Using ArcMap (ESRI ArcGIS Desktop 10.6), we calculated the change in total area with no extrapolation (i.e., those areas with a MESS ≥0) relative to the model producing the largest area with no extrapolation. We did

Downloaded from https://academic.oup.com/aesa/article/113/2/100/5727917 by U S Dept of Agriculture user on 02 October 2020

the same comparison for areas with high extrapolation (i.e., those areas with a MESS < −10), but calculated the change relative to the model producing the *smallest* area for this category. Therefore, the reference point for both measures (i.e., ΔMESS = 0) was set as the MESS value for the model with the most desirable outcome (i.e., the model with the least amount of uncertain space due to extrapolation or high extrapolation).

Results and Discussion

We generated 16 MaxEnt models that were each applied to North America and Europe, yielding 32 maps of climate suitability for *L. dispar dispar* (Figs. 2–5 and Supp Figs. 1–4 [online only], respectively). The four primary models, designated "Native," "Native + Invaded," "Source," and "Source + Invaded," are distinguished by the source of the presence points that were used to train and test the models. Submodels, designated "A"–" D," differ in methods that were applied during model development: "A," no thinning of presence points and default settings for regularization and feature class; "B," no thinning of presence points and optimized settings for regularization and feature class; "C," thinned presence points and default settings for regularization and feature class; and "D," thinned presence points and optimized settings for regularization and feature class. Our primary intent is to focus on methods that yield transferable models, so we do not discuss projected spatial variation in climate suitability from any one model at length. Rather, we focus on how methodological choices among models affect spatial patterns of suitability, internal validation metrics, external validation, and uncertainty associated with extrapolation. The Native-A model represented methodology often used for invasive species, where all available native range data are used and left unthinned and default MaxEnt settings are selected (Fig. 2A and Supp Fig. 1A [online only]). The relative impact of each of these decisions on model outcomes were then tested against alternatives (i.e., submodels B–D), with particular emphasis on comparing potential effects of local adaptation or population differentiation within the subset of

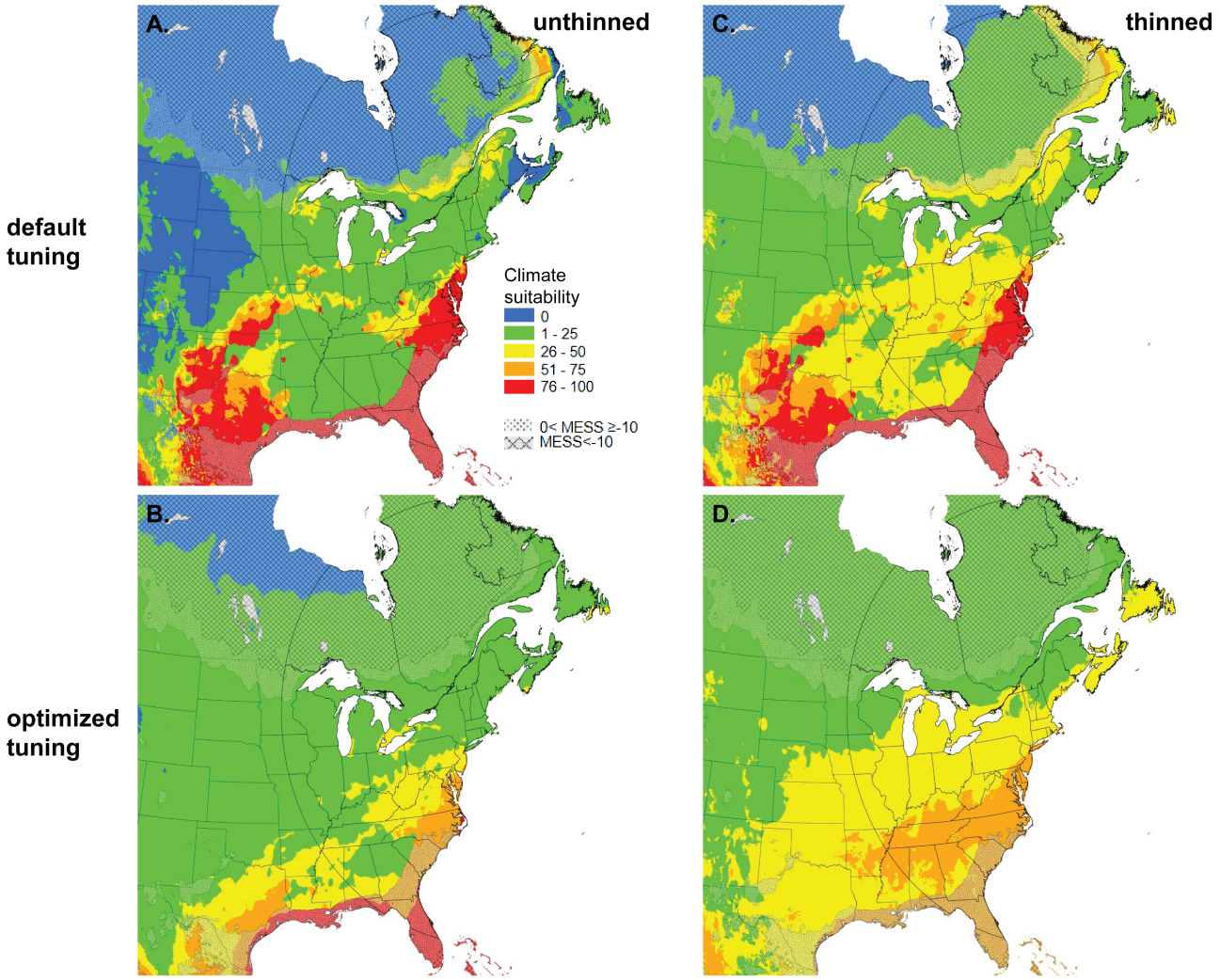

Fig. 2. Forecasted climatic suitability of *L. dispar dispar* in North America using the "Native" occurrence dataset. Four submodels were constructed in MaxEnt using this dataset based on thinning of occurrences and regularization multiplier/feature class combinations: (A) unthinned and default settings, (B) unthinned and 1/linear, (C) 69 km thinning and default settings, and (D) 69 km thinning and 4.5/hinge. The partial ellipse delineates the space within which the independent validation metric for each model was calculated. Colored regions with no patterned overlay had MESS (Multivariate Environmental Similarity Surface) values ≥0 (no extrapolation). Regions with a light gray overlay with dots had MESS values <0 and ≥ −10; light gray with cross-hatching had MESS values < −10. Maps are depicted using the North America Albers Equal Area projection. Refer to Fig. 1B for records of occurrence in North America; they were not included for reference here to avoid obscuring model results.

Downloaded from https://academic.oup.com/aesa/article/113/2/100/5727917 by U S Dept of Agriculture user on 02 October 2020

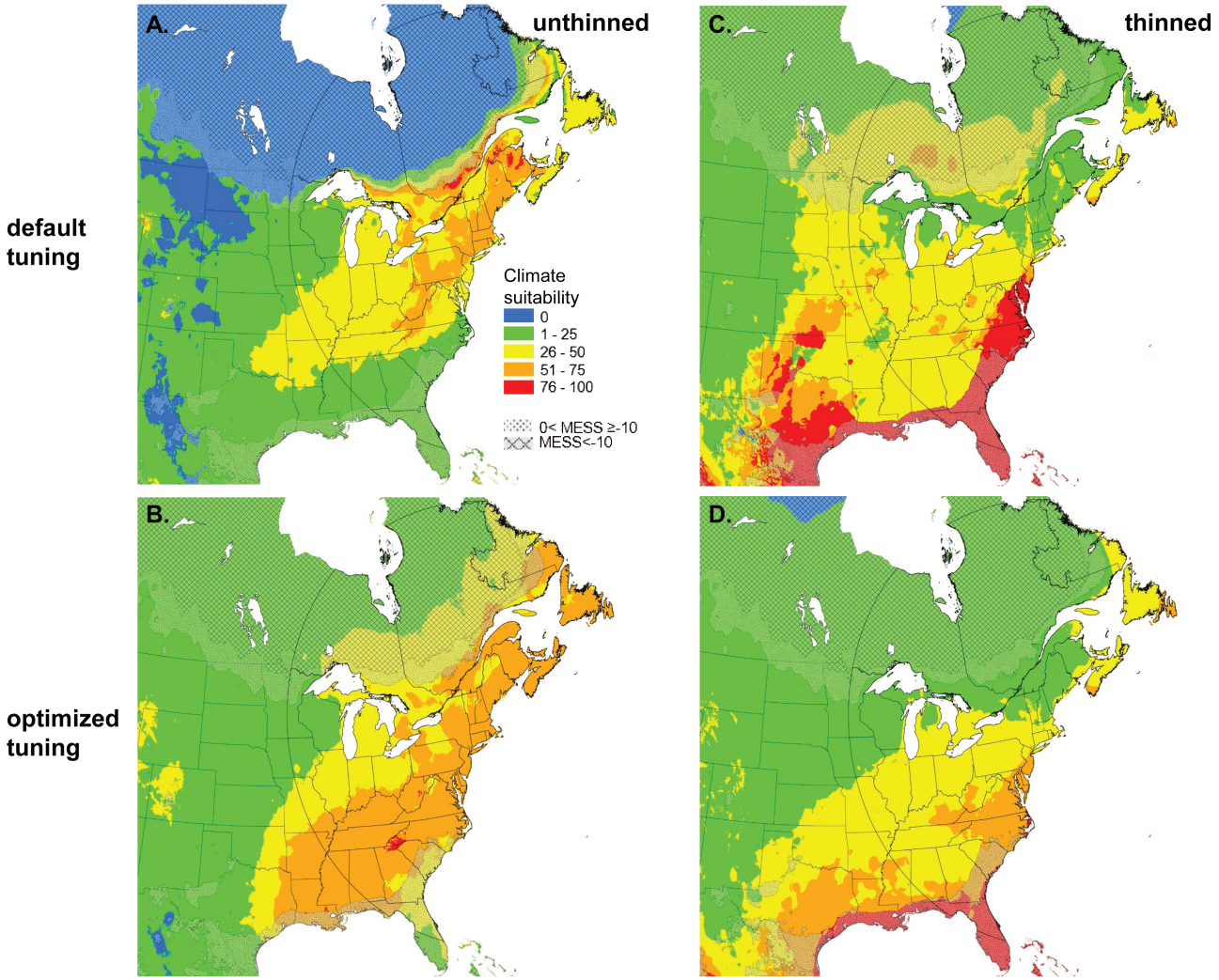

Fig. 3. Forecasted climatic suitability of *L. dispar dispar* in North America using the "Native + Invaded" occurrence dataset. Four submodels were constructed in MaxEnt using this dataset based on thinning of occurrences and regularization multiplier/feature class combinations: (A) unthinned and default settings, (B) unthinned and 2.5/linear, (C) 69 km thinning and default settings, and (D) 69 km thinning and 4/linear + quadratic + hinge + product. The partial ellipse delineates the space within which the independent validation metric for each model was calculated. Colored regions with no patterned overlay had MESS (Multivariate Environmental Similarity Surface) values ≥0 (no extrapolation). Regions with a light gray overlay with dots had MESS values <0 and ≥ −10; light gray with cross-hatching had MESS values < −10. Maps are depicted using the North America Albers Equal Area Conic projection. Refer to Fig. 1B for records of occurrence in North America.

the native range where the invaded population is estimated to be from versus the whole native range (i.e., the primary models using "Native" and "Source" data). Overall, based on current recommendations for model construction and theoretical justification, our expectation was that the model showing the greatest spatial transferability to North America would be the Source + Invaded-D model (Fig. 5D and Supp Fig. 4D [online only]).

All 16 models performed well in their ability to internally discriminate occurrence points from background points. The average AUC for all models ranged from 0.629 to 0.869 (Table 2). Common interpretation of AUC values follows that: AUC >0.9 as "excellent"; 0.9 > AUC > 0.8 as "good"; 0.8 > AUC > 0.7 as "fair"; 0.7 > AUC > 0.6 as "poor"; 0.6 > AUC > 0.5 as "fail"; and AUC < 0.5 as counter-predictions (Araujo et al. 2005, Guisan et al. 2017c). Here, most AUC values were 0.7 or more ("fair" to "good"). The highest average AUC_{test} within each primary model was >0.75, always by using the unthinned data with default MaxEnt settings (i.e., each submodel A in Table 2). This result is consistent with the reputation of MaxEnt for generating models with high discrimination (Elith et al. 2006). The OR_{10} estimates were typically near 0.10 and differences in AUC between training and test datasets were <0.04, both providing evidence that models were generally not overfitted to model training data (Jarnevich and Young 2015).

Because the primary models were built with different datasets, and therefore different extents, direct comparisons of internal validation statistics across primary models were not appropriate (Lobo et al. 2008, Radosavljevic and Anderson 2014). However, when using a common, independent dataset for external validation (orange points in Fig. 1B) comparisons can be made. The CBI, one of the few metrics appropriate to presence-only (or presence-background) data, revealed patterned differences in the ability of different primary and submodels to forecast climate suitability for *L dispar dispar* into novel space in North America. Below, we compare the effects of data filtering (thinned vs unthinned), model parameters (optimized vs default), and the extent of occurrence data (invaded range vs no invaded range; full native range vs source of invasion) on model transferability and extrapolation.

Downloaded from https://academic.oup.com/aesa/article/113/2/100/5727917 by U S Dept of Agriculture user on 02 October 2020

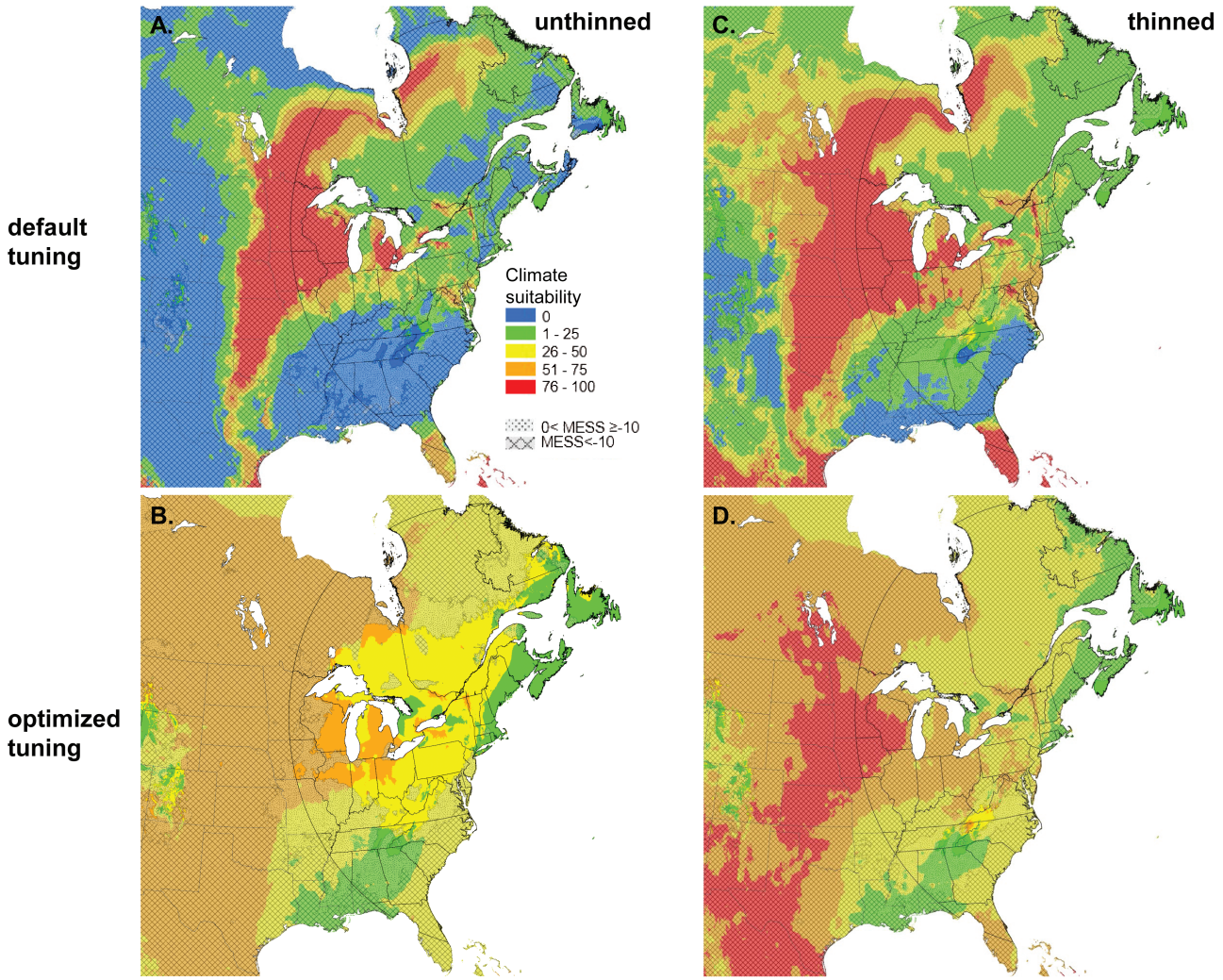

Fig. 4. Forecasted climatic suitability of *L. dispar dispar* in North America using the "Source" occurrence dataset. Four submodels were constructed in MaxEnt using this dataset based on thinning of occurrences and regularization multiplier/feature class combinations: (A) unthinned and default settings, (B) unthinned and 6/linear, (C) 32 km thinning and default settings, and (D) 32 km thinning and 4.5/linear +quadratic. The partial ellipse delineates the space within which the independent validation metric for each model was calculated. Colored regions with no patterned overlay had MESS (Multivariate Environmental Similarity Surface) values ≥0 (no extrapolation). Regions with a light gray overlay with dots had MESS values <0 and ≥ −10; light gray with cross-hatching had MESS values < −10. Maps were depicted using the North America Albers Equal Area Conic projection. Refer to Fig. 1B for records of occurrence in North America.

Effect of Occurrence Extent on Suitability Forecasts and Extrapolation

Suitability

Five of the eight models that used occurrence data from the purported geographic source of the North American infestation, alone or with some occurrence records from North America, gave near perfect predictions of the validation dataset in North America (CBI close to 1), implying that occurrences fell in areas of higher suitability more often than expected by chance alone (Table 2; Supp Fig. 5 [online only]). The highest performing model by this metric was produced with occurrence records from the purported geographic source of the North American infestation alone and by spatial thinning and optimizing tuning parameters (Source-D; CBI = 0.995). The remaining three of eight models (Source-A, Source + Invaded-A and -B) had slightly lower CBI values, but still showed better than random prediction (CBI > 0). Models using the invasion source data, whether alone or in combination with a portion of the invaded range, tended to forecast high relative suitability (>50) throughout the North Central region of the United States and into Canada (Figs. 4 and 5), which is where most of our validation points were located. These models also generally forecasted lower suitability (<26) in the south and southeastern United States.

In contrast, half of the eight models that used occurrence records throughout the native range gave negative CBI scores, which indicates the counter-prediction; a greater proportion of occurrences fell in low suitability values than expected by random chance (Table 2; Supp Fig. 6 [online only]). One model indicated predictions near random (CBI ~0; Native-D), and the remaining three models (Native-A and -C, Native + Invaded-C) showed better than random predictions in similar proportion to the lowest performing Source/Source + Invaded models. These models using the broader native range generally showed high suitability limited to the east and along the southeast coast of the United States (Figs. 2 and 3).

Contrary to previous work (e.g., Beaumont et al. 2009), no clear patterns were seen regarding model transferability after including a portion of the invaded range in model development. In many instances, the addition of invaded space clearly worsened model

Downloaded from https://academic.oup.com/aesa/article/113/2/100/5727917 by U S Dept of Agriculture user on 02 October 2020

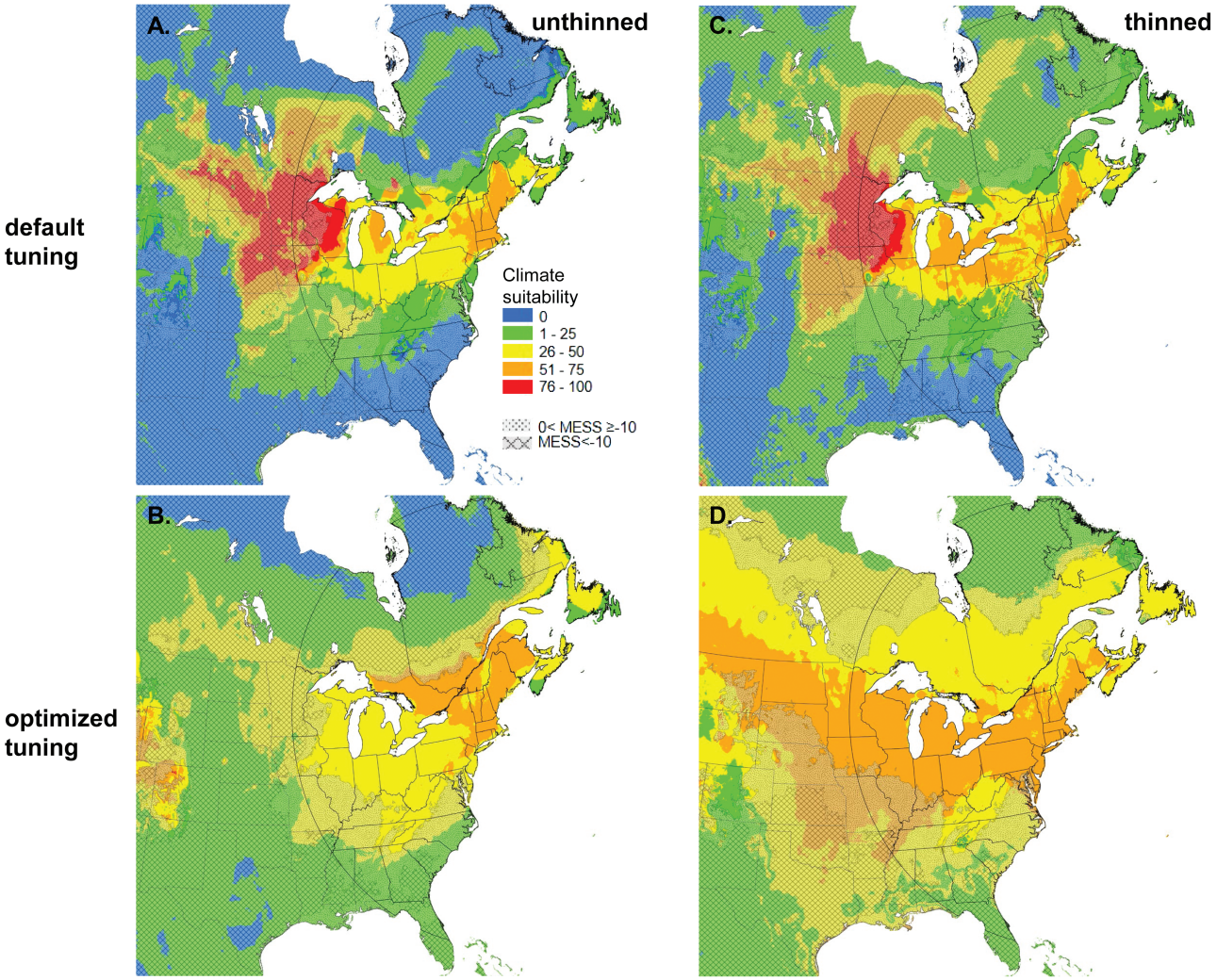

Fig. 5. Forecasted climatic suitability of *L. dispar dispar* in North America using the "Source +Invaded" occurrence dataset. Four submodels were constructed in MaxEnt using this dataset based on thinning of occurrences and regularization multiplier/feature class combinations: (A) unthinned and default settings, (B) unthinned and 6/linear, (C) 32 km thinning and default settings, and (D) 32 km thinning and 4.5/hinge. The partial ellipse delineates the space within which the independent validation metric for each model was calculated. Colored regions with no patterned overlay had MESS (Multivariate Environmental Similarity Surface) values ≥0 (no extrapolation). Regions with a light gray overlay with dots had MESS values <0 and ≥ −10; light gray with cross-hatching had MESS values < −10. Maps were depicted using the North America Albers Equal Area Conic projection. Refer to Fig. 1B for records of occurrence in North America.

performance based on CBI (e.g., Native-A vs Native + Invaded-A; Source-B vs Source + Invaded-B in Table 2). This result may be because the North American population is not at equilibrium, so by including an early portion of the invasion space, represented by a high density of occurrence, we were potentially biasing our model to these non-equilibrium conditions. Lack of equilibrium violates an underlying assumption of species distribution modeling (Elith and Leathwick 2009) but is unavoidable when forecasting invasive species. How to contend with this violation is an active area of research. Briscoe Runquist et al. (2019) suggest that only the invasion stage within the last 10 years may be useful for forecasting present-day distributions for an invasive plant. However, we felt this short timeframe would confound the independence of our validation dataset with the development dataset in North America, both spatially and temporally. We selected ≥1990 and ≤1960 to define the temporal boundaries of the datasets in North Americas because they allowed for temporal separation of occurrence based on an approximation for the longevity of a pest risk map as 30 years (Venette 2013).

Extrapolation

The occurrence dataset extent also impacted the amount of extrapolation space as measured by the MESS (Table 2). To reduce uncertainty, the ideal forecast would be into space that does not require extrapolation (i.e., regions of MESS ≥0) or to space where the degree of extrapolation is relatively minimal. The models with the greatest amount of unextrapolated space within the area used for external validation (i.e., smallest ΔMESS ≥0) were developed using the Native dataset, with the Native + Invaded models having the next lowest. Interestingly, increasing the geographic coverage of the model by including occurrences from the invaded range reduced the amount of unextrapolated space. The opposite would be expected if inclusion of the invaded range better reflected the broader fundamental niche of the species, as was seen with the Source and Source + Invaded models, albeit to a modest degree. The reason for the drop in unextrapolated space is likely because of the slight difference in environmental variables selected between the two primary models; the inclusion of the invaded range data to the Native dataset changed the correlation matrix such that a sixth variable (BIO_7) was added.

Downloaded from https://academic.oup.com/aesa/article/113/2/100/5727917 by U S Dept of Agriculture user on 02 October 2020

Table 2. Model evaluation metrics for MaxEnt models of *L. dispar dispar*

Model		Thinning distance (km)	Regularization/ feature class	Internal validation				External validation	ΔMESS area within validation space (km²)	
				OR_{10}[a]	AUC_{train}[b]	AUC_{test}[c]	AICc[d]	CBI[e]	≥0	< −10
Native	A	Unthinned	*Default*	0.108	0.869	0.863	20,350.24	0.231	0.00	898,714.80
	B	Unthinned	1/L	0.099	0.818	0.816	20,844.16	−0.522	0.00	898,714.80
	C	69	*Default*	0.128	0.783	0.745	4,943.23	0.633	34,973.55	898,714.84
	D	69	4.5/H	0.107	0.761	0.746	4,924.55	0.028	34,973.55	898,714.84
Native + Invaded	A	Unthinned	*Default*	0.101	0.811	0.808	94,175.14	−0.732	67,183.16	885,037.47
	B	Unthinned	2.5/L	0.101	0.777	0.776	98,426.26	−0.400	67,143.85	885,037.47
	C	69	*Default*	0.152	0.782	0.752	6,739.94	0.566	67,183.16	885,037.47
	D	69	4/LQHP	0.098	0.756	0.745	6,380.99	−0.404	67,183.16	885,037.47
Source	A	Unthinned	*Default*	0.127	0.834	0.808	5,146.43	0.592	2,443,736.47	3,622,201.12
	B	Unthinned	6/L	0.114	0.692	0.680	5,367.13	0.949	1,715,512.08	592,102.47
	C	32	*Default*	0.132	0.773	0.734	1,827.74	0.963	2,451,217.63	3,594,126.78
	D	32	4.5/LQ	0.121	0.699	0.690	1,810.73	0.995	2,451,217.63	3,594,126.78
Source + Invaded	A	Unthinned	*Default*	0.102	0.763	0.756	86,201.78	0.665	1,172,727.10	1,234,470.08
	B	Unthinned	6/L	0.101	0.734	0.733	84,080.17	0.156	954,443.74	1,172,424.46
	C	32	*Default*	0.151	0.689	0.632	7,289.13	0.973	1,172,727.10	1,234,470.08
	D	32	4.5/H	0.107	0.629	0.625	7,015.80	0.872	597,592.06	0.00

Internal validation metrics were averaged across five replicates for each model. External validation was conducted on the same independent dataset in North America for all models (see Fig. 1B). The change in Multivariate Environmental Similarity Surface space (ΔMESS) shows how much a given model increased the area of extrapolation within the model validation ellipse, relative to the model with the least amount of extrapolated space. ΔMESS values of zero indicate the model with the greatest amount of space with MESS values ≥0 (no extrapolation), or the least amount of space with MESS values < −10.

[a]OR_{10}: omission rate for the proportion of test locations with suitability values lower than the smallest value after excluding the lowest 10% of training suitability values (10% training omission rate); overfitting is indicated by deviation from the expectation 0.10.

[b]AUC_{train} measures model ability to distinguish presence points used in model training from background points, with 1.0 being perfect discrimination.

[c]AUC_{test} measures model ability to distinguish presence points used in model testing from background points, with 1.0 being perfect discrimination.

[d]AICc: sample corrected Akaike information criteria. Smaller values are considered better at balancing model complexity and goodness-of-fit.

[e]CBI: continuous Boyce Index, which calculates the frequency of occurrences predicted to those expected across a continuous (moving window) prediction gradient. The value provided is the Spearman rank correlation coefficient between the mean suitability and the predicted:expected ratio of occurrences for a given suitability bin. Values close to 1 indicate good predictions (i.e., high suitability areas contained the most occurrences), values near 0 indicate predictions no different from random, and value near −1 indicate counter-predictions (i.e., low suitability areas contained the most occurrences).

Conversely, the models with the least amount of unextrapolated space were developed using the Source dataset. This dataset also produced the model with greatest amount of "highly extrapolated" space (i.e., biggest ΔMESS < −10; Source-A). Appearing somewhat anomalously, Source + Invaded-D was the model with the least amount of highly extrapolated space (i.e., ΔMESS < −10 = 0.0). Though extrapolation was still common, this model had a relatively high amount of unextrapolated space compared to the other models using data from the invasion source. A possible reason for this model's low extrapolation is that MaxEnt found few of the environmental variables to be influential (i.e., lambda >0: Table 1), which is what the MESS is based upon. With fewer variables, the likelihood of extrapolation would be expected to decrease.

In sum, focusing the representation of the native range to the area of purported invasion source appeared to increase greatly the transferability of models to North America. However, including an early stage of the invaded range in model development did not. From an uncertainty standpoint, models using occurrences from the full native range yielded the largest area for which extrapolation will not threatened model reliability. In general, models using occurrences from the invasion source showed the greatest extent of uncertainty due to extrapolation. However, two models, Source-B and Source + Invaded-D showed the two smallest amounts of highly extrapolated space across all models, despite having comparatively low amounts of unextrapolated space.

Effect of Spatial Thinning and Model Tuning on Suitability Forecasts and Extrapolation

Suitability

Generally, model optimization (comparison of panels A to B and C to D in Figs. 2–5 and Supp Figs. 1–4 [online only]) and spatial thinning (comparison of panels A to C and B to D in Figs. 2–5 and Supp Figs. 1–4 [online only]) broadened areas forecasted to be climatically suitability (index >0) compared to default and unthinned models, respectively. These procedures also tended to reduce the amount of space forecast within the highest suitability index (76–100; red), though in some instances, thinning had the opposite effect (e.g., Figs. 2A–C and 3B–D).

Thinning tended to decrease performance compared to unthinned data within a primary model based on the cross-validation results (i.e., increased OR_{10} and decreased AUC values). However, it resulted in consistently less complex models based on relative AICc values within a primary model. Thinning is intended to reduce the spatial autocorrelation often associated with occurrence record data. An apparent loss in model accuracy is expected when comparing thinned to unthinned data due to the tendency of spatial autocorrelation to generate overly optimistic model assessments (Veloz 2009).

Few strong patterns of the effects of optimal tuning parameters on evaluation metrics were seen when compared to unthinned and/or default parameters (Table 2). Within each primary model, optimizing parameters lowered the OR_{10} slightly when compared to default models, indicating a decrease in the already low overfitting to the training data. Our first criterion for selecting the optimal parameters

Downloaded from https://academic.oup.com/aesa/article/113/2/100/5727917 by U S Dept of Agriculture user on 02 October 2020

was the lowest OR_{10}, so this change is expected. Optimization also consistently lowered AUC values when compared to default models, which indicates a decrease in discriminatory ability. This corroborates the qualitative patterns of broadened suitability described above, and is the ultimate goal of optimization; optimization tends to result in "smoother" and simpler functions, thus making the model less specific (overfit) to the model development data (Radosavljevic and Anderson 2014).

We did not see any general patterns of the effect of thinning and/or optimal tuning on external validation across models. However, among the highest performing models (based on CBI), thinned models consistently gave the highest results (Source-C and -D, Source + Invaded-C and -D). There, the models using thinned data (Figs. 4C and D, and 5C and D) generally showed a greater extent of moderate and higher suitability (i.e., >26) across the eastern and Great Lakes region than their unthinned counterparts, with portions of particularly high relative suitability (76–100; red) in the North Central United States. These regions corresponded to the location of much of our validation dataset, so higher suitability scores focused here resulted in better performance. Of note, these thinned models would be considered among the lowest performers based on the standard evaluation metrics of AUC and OR.

Extrapolation

Thinning had little to no effect on the amount of unextrapolated or extrapolated space. This indicates that our method of spatial thinning did not appreciably change the boundaries of climatic space within a given model. In cases where optimal parameter tuning resulted in a different amount of unextrapolated or extrapolated space compared to default settings, it was always for the better; a gain in unextrapolated space (i.e., smaller ΔMESS ≥0) or loss in highly extrapolated spaced (i.e., smaller ΔMESS < −10) (e.g., Source-A to Source-B; Table 2).

In sum, parameter optimization and spatial thinning had variable effects on model performance individually (Table 2). Taken together (i.e., the "D" submodels), these procedures reduced model complexity (AICc) and increased predictive ability (OR_{10}), but the difference was generally small compared to other submodels. Importantly, there was no clear impact on transferability (CBI) other than that ignoring thinning (i.e., the A and B submodels) resulted in the worst or second-worst transferable models. Sometimes optimization of unthinned models mitigated this effect, as in Source-A to -B, but other times not (e.g., Native-A to -B, Source + Invaded-A to -B).

Conclusions

In this study, we offer an example of modeling an invasive species using a popular machine-learning approach (MaxEnt) when different amounts of information are available. The spatial extent of the occurrence dataset used in model development seemed to matter most in determining the spatial transferability of our models. Models constructed with occurrences limited to the purported European source of invasion for *L. dispar dispar* in North America, either alone (Source) or in conjunction with occurrences from a portion of the invaded range (Source + Invaded) performed very well on independent data (CBI values near 1.0). Within that, the recommended practice of spatial thinning helped to consistently predict independent occurrences with high suitability (Table 2; CBI). However, the most transferable models had the greatest extents of uncertainty as measured by MESS (i.e., smaller values in ΔMESS ≥0 and/or larger values in ΔMESS < −10). Optimization of parameters (i.e., regularization multiplier and features classes) did not consistently impact how well a model would perform in new regions, nor did inclusion of a portion of the invaded range. Our results and subsequent conclusions here are limited to one species and whether patterns hold for others remains to be seen. Additional studies that evaluate multiple circumstances of information typically available to a risk practitioner, such as those in our study, for many different invasive species may help form broader generalizations to guide model selections.

We note that our study did not explicitly address environmental variable selection, other than reducing the number used based on correlation. We recognize that the choice of variables has been shown to have important impact on the model results (e.g., Rödder et al. 2009, Braunisch et al. 2013). Though much is known about the biology of *L. dispar dispar* at this point in its invasion, we opted to mimic the circumstance of most invasive species distribution models and begin with few a priori assumptions about species-specific climatic drivers upon which to limit variable selection. The general assumption that variation in temperature and moisture shapes geographic distributions for polyphagous insects is well accepted. For similar reasons, we chose the widely used WorldClim database for our climate data. Though there have been some useful expansions of this original dataset (e.g., Kriticos et al. 2012, Abatzoglou et al. 2018, Title and Bemmels 2018), the availability of high-resolution, global datasets that include a diversity of recently measured biotic and abiotic variables is lacking.

The process of constructing a MaxEnt model requires a number of subjective decisions to be made, affecting the uncertainty in final forecasts. Each decision could be considered an axis in multidimensional modeling space. The number of potential models that can be created within MaxEnt then depends on the number of axes and the number of options along each axis. In this case, we focused on four axes each with two options. As a result, it is inappropriate to think of just one MaxEnt model for a species, and a general description of good or bad model performance becomes difficult, if not impossible. Given that applications of any modeling technique to invasive species are typically rife with idiosyncratic limitations and caveats, it is difficult to imagine a robust set of recommendations that will ensure a rigorous and reliable model for all applications. In spite of this, we find it impractical and premature to suggest abandoning correlative models such as MaxEnt for forecasting invasive species distributions. The intent of our current study is not to suggest a "best" approach, but rather further illustrate particular sources of uncertainty that may be present when using this tool. When circumstances of data limitation do not allow for conformity to current recommendations in methodology (e.g., lack of independent data for validation, too few occurrences to spatially filter, etc.), our study adds to the evidence a pest risk practitioner can evaluate in order to make more informed decisions regarding the appropriateness of MaxEnt for their situation, and the uncertainties and possible consequences that need to be communicated if a forecast is produced.

The possibility to improve the transferability of MaxEnt models with information from the genetic origin of an invading population is an intriguing concept that requires further testing with more species. The source population in our study was defined by genetic divergence observed in a survey of *L. dispar dispar* populations, but whether this divergence corresponds to phenotypic differences in environmental tolerances has not been shown. The results seen here suggest there may be biological relevance to refining species distribution models with population genetic information, but experimental evidence is still needed to establish a robust link.

Perhaps one of the greatest challenges in developing models to forecast suitable habitats for invasive species is model validation, the formal assessment of model transferability. Rigorous validation

Downloaded from https://academic.oup.com/aesa/article/113/2/100/5727917 by U S Dept of Agriculture user on 02 October 2020

is prerequisite to judge model reliability and usefulness (e.g., Bleeker et al. 2003). A fundamental assumption of validation of any model is independence of the data used for validation from the data used in development. Though often acknowledged as the true test of validation, independent data are often unavailable, so are rarely used to validate species distribution models (Araujo et al. 2005). Forms of internal validation (e.g., resubstitution) are used, or a subset of occurrence records are withheld from model development (e.g., bootstrapping or data-splitting), but the data used in these approaches typically are not spatially independent from the dataset used in development. In our study, models that performed the best internally failed external validation tests, whereas models that might otherwise be considered poor by internal performance statistics did well. Because the primary models in our study were built with different datasets, direct comparisons of internal validation statistics across primary models were not appropriate (Lobo et al. 2008). This realization makes it difficult to assign an a priori model selection metric to identify the "best" model with which to support or refute with external validation. Using external validation (here, the CBI) to decide model selection confounds the role of independent data; the data then become involved in model development thus negating their independence. We suggest that in instances where data limitations prevent proper validation of a habitat suitability model in novel space for an invasive species, this limitation should be clearly stated and future monitoring efforts for the species should be structured to provide independent records for evaluating the usefulness of the model.

Supplementary Data

Supplementary data are available at *Annals of the Entomological Society of America* online.

Supplementary Fig. 1. Forecasted climatic suitability of *L. dispar dispar* in Europe using the "Native" occurrence dataset. Four submodels were constructed in MaxEnt using this dataset based on thinning of occurrences and regularization multiplier/feature class combinations: (A) unthinned and default settings, (B) unthinned and 1/linear, (C) 69 km thinning and default settings, and (D) 69 km thinning and 4.5/hinge. The partial ellipse delineates the space within which the independent validation metric for each model was calculated. Colored regions with no patterned overlay had MESS (Multivariate Environmental Similarity Surface) values ≥0 (no extrapolation). Regions with a light gray overlay with dots had MESS values <0 and ≥ −10; light gray with cross-hatching had MESS values < −10. Maps are depicted using the Europe Albers Equal Area Conic projection.

Supplementary Fig. 2. Forecasted climatic suitability of *L. dispar dispar* in Europe using the "Native + Invaded" occurrence dataset. Four submodels were constructed in MaxEnt using this dataset based on thinning of occurrences and regularization multiplier/feature class combinations: (A) unthinned and default settings, (B) unthinned and 2.5/linear, (C) 69 km thinning and default settings, and (D) 69 km thinning and 4/linear + quadratic + hinge + product. The partial ellipse delineates the space within which the independent validation metric for each model was calculated. Colored regions with no patterned overlay had MESS (Multivariate Environmental Similarity Surface) values ≥0 (no extrapolation). Regions with a light gray overlay with dots had MESS values <0 and ≥ −10; light gray with cross-hatching had MESS values < −10. Maps are depicted using the Europe Albers Equal Area Conic projection.

Supplementary Fig. 3. Forecasted climatic suitability of *L. dispar dispar* in Europe using the "Source" occurrence dataset. Four submodels were constructed in MaxEnt using this dataset based on thinning of occurrences and regularization multiplier/feature class combinations: (A) unthinned and default settings, (B) unthinned and 6/linear, (C) 32 km thinning and default settings, and (D) 32 km thinning and 4.5/linear + quadratic. The partial ellipse delineates the space within which the independent validation metric for each model was calculated. Colored regions with no patterned overlay had MESS (Multivariate Environmental Similarity Surface) values ≥0 (no extrapolation). Regions with a light gray overlay with dots had MESS values <0 and ≥ −10; light gray with cross-hatching had MESS values < −10. Maps are depicted using the Europe Albers Equal Area Conic projection.

Supplementary Fig. 4. Forecasted climatic suitability of *L. dispar dispar* in Europe using the "Source + Invaded" occurrence dataset. Four submodels were constructed in MaxEnt using this dataset based on thinning of occurrences and regularization multiplier/feature class combinations: (A) unthinned and default settings, (B) unthinned and 6/linear, (C) 32 km thinning and default settings, and (D) 32 km thinning and 4.5/hinge. The partial ellipse delineates the space within which the independent validation metric for each model was calculated. Colored regions with no patterned overlay had MESS (Multivariate Environmental Similarity Surface) values ≥0 (no extrapolation). Regions with a light gray overlay with dots had MESS values <0 and ≥ −10; light gray with cross-hatching had MESS values < −10. Maps are depicted using the Europe Albers Equal Area Conic projection.

Supplementary Fig. 5. Plots of the predicted-to-expected ratio of occurrences used to calculate the continuous Boyce Index (CBI; "ecospat" package) for the "Source" and "Source + Invaded" models. Calculations were made within the validation space in North America (tan space; main text Fig. 1) and using the validation dataset (orange points; main text Fig. 1). See Table 2 in main text for CBI values.

Supplementary Fig. 6. Plots of the predicted-to-expected ratio of occurrences used to calculate the continuous Boyce Index (CBI; "ecospat" package) for the "Native" and "Native + Invaded" models. Calculations were made within the validation space in North America (tan space; main text Fig. 1) and using the validation dataset (orange points; main text Fig. 1). See Table 2 in main text for CBI values.

Acknowledgments

This research was supported by the Minnesota Invasive Terrestrial Plants and Pests Center through the Minnesota Environment and Natural Resources Trust Fund. We are grateful to Denise Dodd (Virginia Tech) and Andrew Liebhold (USDA-FS) for graciously providing historical Slow the Spread data and additional records in west-central Asia, respectively.

References Cited

Abatzoglou, J. T., S. Z. Dobrowski, S. A. Parks, and K. C. Hegewisch. 2018. TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015. Sci. Data 5: 1–12.

Aiello-Lammens, M. E., R. A. Boria, A. Radosavljevic, B. Vilela, and R. P. Anderson. 2015. spThin: an R package for spatial thinning of species occurrence records for use in ecological niche models. Ecography (Cop.) 38: 541–545.

Anderson, R. P., and I. Gonzalez. 2011. Species-specific tuning increases robustness to sampling bias in models of species distributions: an implementation with Maxent. Ecol. Modell. 222: 2796–2811.

Araujo, M. B., R. G. Pearson, W. Thuiller, and M. Erhard. 2005. Validation of species-climate impact models under climate change. Glob. Chang. Biol. 11: 1504–1513.

Barbet-Massin, M., Q. Rome, C. Villemant, and F. Courchamp. 2018. Can species distribution models really predict the expansion of invasive species? PLoS ONE 13: e0193085.

Beaumont, L. J., R. V. Gallagher, W. Thuiller, P. O. Downey, M. R. Leishman, and L. Hughes. 2009. Different climatic envelopes among invasive populations may lead to underestimations of current and future biological invasions. Divers. Distrib. 15: 409–420.

Bleeker, S. E., H. A. Moll, E. W. Steyerberg, A. R. Donders, G. Derksen-Lubsen, D. E. Grobbee, and K. G. Moons. 2003. External validation is necessary in prediction research: a clinical example. J. Clin. Epidemiol. 56: 826–832.

Boyce, M. S., P. R. Vernier, S. E. Nielsen, and F. K. A. Schmiegelow. 2002. Evaluating resource selection functions. Ecol. Modell. 157: 281–300.

Braunisch, V., J. Coppes, R. Arlettaz, R. Suchant, H. Schmid, and K. Bollmann. 2013. Selecting from correlated climate variables: a major source of

Downloaded from https://academic.oup.com/aesa/article/113/2/100/5727917 by U S Dept of Agriculture user on 02 October 2020

uncertainty for predicting species distributions under climate change. Ecography (Cop.) 36: 971–983.

Briscoe Runquist, R. D., T. Lake, P. Tiffin, and D. A. Moeller. 2019. Species distribution models throughout the invasion history of Palmer amaranth predict regions at risk of future invasion and reveal challenges with modeling rapidly shifting geographic ranges. Sci. Rep. 9: 1–12.

Broennimann, O., and A. Guisan. 2008. Predicting current and future biological invasions: both native and invaded ranges matter. Biol. Lett. 4: 585–589.

Broennimann, O., U. A. Treier, H. Müller-Schärer, W. Thuiller, A. T. Peterson, and A. Guisan. 2007. Evidence of climatic niche shift during biological invasion. Ecol. Lett. 10: 701–709.

Brown, J. L., J. R. Bennett, and C. M. French. 2017. SDMtoolbox 2.0: the next generation Python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. PeerJ 5: e4095.

Crall, A. W., C. S. Jarnevich, B. Panke, N. Young, M. Renz, and J. Morisette. 2013. Using habitat suitability models to target invasive plant species surveys. Ecol. Appl. 23: 60–72.

Di Cola, V., O. Broennimann, B. Petitpierre, F. T. Breiner, M. D'Amen, C. Randin, R. Engler, J. Pottier, D. Pio, A. Dubuis, et al. 2017. ecospat: an R package to support spatial analyses and modeling of species niches and distributions. Ecography (Cop.) 40: 774–787.

Dlogosch, K. M., and I. M. Parker. 2008. Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. Mol. Ecol. 17: 431–449.

Dormann, C. F., J. Elith, S. Bacher, C. Buchmann, G. Carl, G. Carré, J. R. G. Marquéz, B. Gruber, B. Lafourcade, P. J. Leitão, et al. 2013. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography (Cop.) 36: 27–46.

Dos Santos, L. A., M. F. Mendes, A. P. Krüger, M. L. Blauth, M. S. Gottschalk, and F. R. Garcia. 2017. Global potential distribution of *Drosophila suzukii* (Diptera, Drosophilidae). PLoS ONE 12: e0174318.

Dudik, M., S. J. Phillips, and R. E. Schapire. 2007. Maximum entropy density estimation with generalized regularization and an application to species distribution modeling. J. Mach. Learn. Res. 8: 1217–1260.

Elith, J. 2000. Quantitative methods for modeling species habitat: comparative performance and an application to Australian plants, pp. 39–58. *In* S. Ferson and M. Burgman, (eds.), Quantitative methods for conservation biology. Springer-Verlag, New York.

Elith, J. 2017. Predicting distributions of invasive species, pp. 93–129. *In* A. P. Robinson, T. Walshe, M. A. Burgman, and M. Nunn, (eds.), Invasive species: risk assessment and management. Cambridge University Press, Cambridge, UK.

Elith, J., and J. R. Leathwick. 2009. Species distribution models: ecological explanation and prediction across space and time. Annu. Rev. Ecol. Evol. Syst. 40: 677–697.

Elith, J., C. H. Graham, R. P. Anderson, M. Dudík, S. Ferrier, A. Guisan, R. J. Hijmans, F. Huettmann, J. R. Leathwick, A. Lehmann, et al. 2006. Novel methods improve prediction of species' distributions from occurrence data. Ecography (Cop.) 29: 129–151.

Elith, J., M. Kearney, and S. Phillips. 2010. The art of modelling range-shifting species. Methods Ecol. Evol. 1: 330–342.

Elith, J., S. J. Phillips, T. Hastie, M. Dudík, Y. E. Chee, and C. J. Yates. 2011. A statistical explanation of MaxEnt for ecologists. Divers. Distrib. 17: 43–57.

Elith, J., J. Simpson, M. Hirsch, and M. A. Burgman. 2013. Taxonomic uncertainty and decision making for biosecurity: spatial models for myrtle/ guava rust. Australas. Plant Pathol. 42: 43–51.

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

Fourcade, Y., A. G. Besnard, and J. Secondi. 2018. Paintings predict the distribution of species, or the challenge of selecting environmental predictors and evaluation statistics. Glob. Ecol. Biogeogr. 27: 245–256.

Galante, P. J., B. Alade, R. Muscarella, S. A. Jansa, S. M. Goodman, and R. P. Anderson. 2018. The challenge of modeling niches and distributions for data-poor species: a comprehensive approach to model complexity. Ecography (Cop.) 41: 726–736.

Guisan, A., W. Thuiller, and N. E. Zimmermann. 2017a. Assessing model performance: which data to use?, pp. 270–296. *In* A. Guisan, W. Thuiller, and N. E. Zimmermann (eds.), Habitat suitability and distribution models with applications in R. Cambridge University Press, Cambridge, UK.

Guisan, A., W. Thuiller, and N. E. Zimmermann. 2017b. Habitat suitability and distribution models with applications in R. Cambridge University Press, Cambridge, UK.

Guisan, A., W. Thuiller, and N. E. Zimmermann. 2017c. Measuring model accuracy: which metrics to use?, pp. 241–269. *In* A. Guisan, W. Thuiller, and N. E. Zimmermann, (eds.), Habitat suitability and distribution models with applications in R. Cambridge University Press, Cambridge, UK.

Hajek, A. E., and P. C. Tobin. 2009. North American eradications of Asian and European gypsy moth, chapter: 5, pp. 71–89. *In* A. E. Hajek, T. R. Glare and M. O'Callaghan, (eds.), Use of microbes for control and eradication of invasive arthropods. Springer Netherlands, Dordrecht.

Heikkinen, R. K., M. Marmion, and M. Luoto. 2012. Does the interpolation accuracy of species distribution models come at the expense of transferability? Ecography (Cop.) 35: 276–288.

Hijmans, R. J. 2017. 'raster': Geographic data analysis and modeling. R package version 2.9-5. https://CRAN.R-project.org/package=raster (accessed on 16 September 2019).

Hijmans, R. J., S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis. 2005. Very high resolution interpolated climate surfaces for global land areas. Int. J. Climatol. 25: 1965–1978.

Hijmans, R. J., S. J. Phillips, J. R. Leathwick, and J. Elith. 2017. 'raster': Geographic data analysis and modeling. R package. https://CRAN.R-project.org/package=dismo (accessed on 16 September 2019).

Hirzel, A. H., G. Le Lay, V. Helfer, C. Randin, and A. Guisan. 2006. Evaluating the ability of habitat suitability models to predict species presences. Ecol. Modell. 199: 142–152.

Jarnevich, C. S., and N. Young. 2015. Using the MAXENT program for species distribution modelling to assess invasion risk, pp. 65–81. *In* R. C. Venette, (ed.), Pest risk modelling and mapping for invasive alien species. CAB International, Wallingford, UK.

Jiménez-Valverde, A., A. T. Peterson, J. Soberón, J. M. Overton, P. Aragón, and J. M. Lobo. 2011. Use of niche models in invasive species risk assessments. Biol. Invasions 13: 2785–2797.

Kass, J. M., B. Vilela, M. E. Aiello-Lammens, R. Muscarella, C. Merow, and R. P. Anderson. 2018. Wallace: a flexible platform for reproducible modeling of species niches and distributions built for community expansion. Methods Ecol. Evol. 9: 1151–1156.

Kriticos, D. J., B. L. Webber, A. Leriche, N. Ota, I. Macadam, J. Bathols, and J. K. Scott. 2012. CliMond: global high-resolution historical and future scenario climate surfaces for bioclimatic modelling. Methods Ecol. Evol. 3: 53–64.

Kumar, S., L. G. Neven, and W. L. Yee. 2014. Assessing the potential for establishment of western cherry fruit fly using ecological niche modeling. J. Econ. Entomol. 107: 1032–1044.

Lee, C. E. 2002. Evolutionary genetics of invasive species. Trends Ecol. Evol. 17: 386–391.

Liebhold, A. M., K. W. Gottschalk, E. R. Luzader, D. A. Mason, R. Bush, and D. B. Twardus. 1997. Gypsy moth in the United States: an atlas. General Technical Report NE-233. U.S. Department of Agriculture, Forest Service, Northeastern Forest Experiment Station, Radnor, PA. 36 p.

Liebhold, A., V. Mastro, and P. W. Schaefer. 1989. Learning from the legacy of Leopold Trouvelot. Bull. Entomol. Soc. Am. 35: 20–22.

Lobo, J. M., A. Jiménez-valverde, and R. Real. 2008. AUC: a misleading measure of the performance of predictive distribution models. Glob. Ecol. Biogeogr. 17: 145–151.

Mau-Crimmins, T. M., H. R. Schussman, and E. L. Geiger. 2006. Can the invaded range of a species be predicted sufficiently using only native-range data? Lehmann lovegrass (*Eragrostis lehmanniana*) in the southwestern United States. Ecol. Modell. 193: 736–746.

Merow, C., M. J. Smith, and J. A. Silander. 2013. A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. Ecography (Cop.) 36: 1058–1069.

Downloaded from https://academic.oup.com/aesa/article/113/2/100/5727917 by U S Dept of Agriculture user on 02 October 2020

Morales, N. S., I. C. Fernández, and V. Baca-González. 2017. MaxEnt's parameter configuration and small samples: are we paying attention to recommendations? A systematic review. PeerJ 5: e3093.

Muscarella, R., P. J. Galante, M. Soley-Guardia, R. A. Boria, J. M. Kass, M. Uriarte, and R. P. Anderson. 2014. ENMeval: an R package for conducting spatially independent evaluations and estimating optimal model complexity for Maxent ecological niche models. Methods Ecol. Evol. 5: 1198–1205.

Pearson, R. G., C. J. Raxworthy, M. Nakamura, and A. T. Peterson. 2007. Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. J. Biogeogr. 34: 102–117.

Peterson, A. T., J. Svoberón, R. G. Pearson, R. P. Anderson, E. Martinez-Meyer, M. Nakamura, and M. B. Araújo. 2011. Ecological niches and geographic distributions, pp. 49–6266. *In* S. A. Levin and H. S. Horn (eds.), Monographs in population biology. Princeton University Press, Princeton, NJ.

Petitpierre, B., O. Broennimann, C. Kueffer, C. Daehler, and A. Guisan. 2017. Selecting predictors to maximize the transferability of species distribution models: lessons from cross-continental plant invasions. Glob. Ecol. Biogeogr. 26: 275–287.

Phillips, S. J. 2008. Transferability, sample selection bias and background data in presence-only modelling: a response to Peterson *et al*. (2007). Ecography (Cop.) **31**: 272–278.

Phillips, S. J., M. Dudík, and R. E. Schapire. n.d. Maxent software for modeling species niches and distributions. Version 3.4.1. Help documentation. http://biodiversityinformatics.amnh.org/open_source/maxent/ (accessed on 16 September 2019).

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

Phillips, S. J., R. P. Anderson, M. Dudík, R. E. Schapire, and M. E. Blair. 2017. Opening the black box: an open-source release of Maxent. Ecography (Cop.) 40: 887–893.

R Core Team. 2018. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Version 3.5.2. http://www.R-project.org/.

Radosavljevic, A., and R. P. Anderson. 2014. Making better Maxent models of species distributions: complexity, overfitting and evaluation. J. Biogeogr. 41: 629–643.

Rodda, G. H., C. S. Jarnevich, and R. N. Reed. 2001. Challenges in identifying sites climatically matched to the native ranges of animal invaders. PLoS ONE 6: e14570.

Rödder, D., S. Schmidtlein, M. Veith, and S. Lötters. 2009. Alien invasive slider turtle in unpredicted habitat: a matter of niche shift or of predictors studied? PLoS ONE 4: e7843.

RStudio. 2016. RStudio: integrated development for R. Version 1.1.456. RStudio Inc., Boston, MA. http://www.rstudio.com/

Shcheglovitova, M., and R. P. Anderson. 2013. Estimating optimal complexity for ecological niche models: a jackknife approach for species with small sample sizes. Ecol. Modell. 269: 9–17.

Sobek-Swant, S., D. A. Kluza, K. Cuddington, and D. B. Lyons. 2012. Potential distribution of emerald ash borer: what can we learn from ecological niche models using Maxent and GARP? For. Ecol. Manage. 281: 23–31.

Sofaer, H. R., J. A. Hoeting, and C. S. Jarnevich. 2019. The area under the precision-recall curve as a performance metric for rare binary events. Methods Ecol. Evol. 10: 565–577.

Sharov, A. A., D. Leonard, A. M. Liebhold, E. A. Roberts, and W. Dickerson. 2002. Slow the spread: a national program to contain the gypsy moth. J. For. Res. 100: 30–36.

Steen, B., A. C. Cardoso, K. Tsiamis, K. Nieto, J. Engel, and E. Gervasini. 2019. Modelling hot spot areas for the invasive alien plant *Elodea nuttallii* in the EU. Manag. Biol. Invasions. 10: 151–170.

Syfert, M. M., M. J. Smith, and D. A. Coomes. 2013. The effects of sampling bias and model complexity on the predictive performance of MaxEnt species distribution models. PLoS ONE 8: e55158.

Title, P. O., and J. B. Bemmels. 2018. ENVIREM: an expanded set of bioclimatic and topographic variables increases flexibility and improves performance of ecological niche modeling. Ecography (Cop.) 41: 291–307.

Tobin, P. C., A. A. Sharov, A. A. Liebhold, D. S. Leonard, E. A. Roberts, and M. R. Learn. 2004. Management of the gypsy moth through a decision algorithm under the STS project. Am. Entomol. **50**: 200–209.

Tobin, P. C., B. B. Bai, D. A. Eggen, and D. S. Leonard. 2012. The ecology, geopolitics, and economics of managing *Lymantria dispar* in the United States. Int. J. Pest Manag. 58: 195–210.

Tobin, P. C., A. M. Liebhold, E. A. Robert, and L. M. Blackburn. 2015. Estimating spread rates of non-native species: the gypsy moth as a case study, pp. 131–144. *In* R. C. Venette, (ed.), Pest risk modelling and mapping for invasive alien species. CAB International, Wallingford, UK.

Vaclavık, T., and R. Meentemeyer. 2012. Equilibrium or not? Modelling potential distribution of invasive species in different stages of invasion. Divers. Distrib. 18: 73–83.

Veloz, S. D. 2009. Spatially autocorrelated sampling falsely inflates measures of accuracy for presence-only niche models. J. Biogeogr. 36: 2290–2299.

Venette, R. 2013. Incorporating climate change into pest risk models for forest pathogens: a role for cold stress in an era of global warming? NeoBiota 18: 131–150.

Venette, R. C. 2017. Climate analyses to assess risks from invasive forest insects: simple matching to advanced models. Curr. For. Reports 3: 255–268.

Venette, R. C., D. J. Kriticos, R. D. Magarey, F. H. Koch, R. H. A. Baker, S. P. Worner, N. N. Gómez Raboteaux, D. W. McKenney, E. J. Dobesberger, D. Yemshanov, et al. 2010. Pest risk maps for invasive alien species: a roadmap for improvement. Bioscience **60**: 349–362.

Warren, D. L., and S. N. Seifert. 2011. Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. Ecol. Appl. 21: 335–342.

Warren, D. L., R. E. Glor, and M. Turelli. 2010. ENMTools: a toolbox for comparative studies of environmental niche models. Ecography (Cop.) 33: 607–611.

Wisz, M. S., R. J. Hijmans, J. Li, A. T. Peterson, C. H. Graham, and A. Guisan. 2008. Effects of sample size on the performance of species distribution models. Divers. Distrib. 14: 763–773.

Wu, Y., J. J. Molongoski, D. F. Winograd, S. M. Bogdanowicz, A. S. Louyakis, D. R. Lance, V. C. Mastro, and R. G. Harrison. 2015. Genetic structure, admixture and invasion success in a Holarctic defoliator, the gypsy moth (*Lymantria dispar*, Lepidoptera: Erebidae). Mol. Ecol. 24: 1275–1291.

Yackulic, C. B., R. Chandler, E. F. Zipkin, J. A. Royle, J. D. Nichols, E. H. Campbell Grant, and S. Veran. 2013. Presence-only modelling using MAXENT: when can we trust the inferences? Methods Ecol. Evol. 4: 236–243.

Zhu, G., W. Bu, Y. Gao, and G. Liu. 2012. Potential geographic distribution of brown marmorated stink bug invasion (*Halyomorpha halys*). PLoS ONE. 7: e31246.

Zimmermann, N. E., T. C. Edwards, C. H. Graham, P. B. Pearman, and J. C. Svenning. 2010. New trends in species distribution modelling. Ecography (Cop.) 33: 985–989.

Downloaded from https://academic.oup.com/aesa/article/113/2/100/5727917 by U S Dept of Agriculture user on 02 October 2020