Quantifying the expansion of an invasive plant species, Dog-strangling Vine (*Vincetoxicum rossicum*), in environmental and geographic space over the past 130 years

Sharla Foster

Thesis submitted to the University of Ottawa
In partial fulfillment of the requirements for the Master of Sciences (M.Sc.) degree in the
Department of Biology, University of Ottawa, Ottawa, Ontario, Canada, K1N 6N5

Thèse soumise à l'Université d'Ottawa
En vue de l'obtention du diplôme de maîtrise (M.Sc.) au
Département de Biologie, Université d'Ottawa, Ottawa, Ontario, Canada, K1N 6N5

Acknowledgements

Firstly, I am deeply grateful for my supervisors, whose guidance, support, and expertise were invaluable during this process. To Dr Heather Kharouba, thank you for always working to provide opportunities for your students to grow and for fostering a constructive and friendly lab environment. To Dr. Tyler Smith, thank you for making your excitement for coding contagious; you and your sense of humour made the learning process more enjoyable and for that I'll always be grateful. I would like to extend my appreciation to my supervisory committee, Dr. Naomi Cappuccino and Dr. David Currie, for your insight and constructive feedback. Special thanks to current and former Kharouba lab members for engaging discussions and diversions, especially Phil Tremblay, Stephanie Rivest, Sydney Gilmour, Sam Lalonde, and Sarah Dolson for the games, guidance, lunches and laughs. Thanks as well to the Forrest and Sargent labs for your feedback during joint lab meetings. I gratefully acknowledge the assistance of Dr. Subbaiah Mechanda & Jean-François Beaulnes for verifying over 1000 iNaturalist records, Stephen Darbyshire and Dr. Paul Catling who provided herbarium record data, and Dr. Richard Dickinson, Dr. Mhairi McFarlane, Dr. Jenny McCune, Michael Oldham and Paul O'Hara who provided survey data. Thanks to Grace Simonds and everyone else who georeferenced occurrence data used in this project. I'd like to thank the University of Ottawa and NSERC for financial support, and my parents, family, and friends for their emotional support and understanding whenever I had to miss events. Finally, thanks to my partner, Kevin, for your patience, unwavering belief in me and for reading over all my drafts.

Abstract

Invasive plant species are an increasing global threat to native biodiversity. Effective management depends on accurate predictions of their spread. However, modelling the geographic distribution of invasive species, particularly with methods like correlative species distribution models (SDMs), is challenging. SDMs operate under the assumption that species are in equilibrium with their environment (i.e., they occur in all suitable environments); this assumption is more likely to be violated for a species that is still in the process of colonizing suitable environments. SDMs also assume that environmental constraints are the most important factors determining a species' distribution. However, these assumptions are not commonly assessed, and when violated can have consequences for model reliability. I investigated SDM performance and equilibrium in the invasive *Vincetoxicum rossicum* vine in northeastern North America. *Vincetoxicum rossicum* has a long, detailed history of occurrence records in its invaded range, which enabled me to observe trends in equilibrium and model performance over a relatively long time scale. I tested the hypotheses that: 1) invasive species approach equilibrium in environmental and geographic space over time; 2) SDM performance will increase as *V. rossicum* approaches environmental equilibrium; and 3) range expansion in the early stages of an invasion is primarily a function of dispersal rather than environmental constraints, while the reverse is true in later stages. I found that *V. rossicum* has reached equilibrium in environmental space, but is still expanding its geographic range. SDM performance was poor in the first 30 years following introduction, but then improved as *V. rossicum* approached environmental equilibrium. SDMs were outperformed by spatial dispersal models in the earliest time period, however, the reverse was true for all subsequent time periods. Overall, these results suggest that *V. rossicum*'s distribution is becoming more stable and more predictable over time and that models built using the most recent data for this species, will be the most transferable across time and space. Additionally, my findings highlight the need for researchers modelling invasive species' distributions to consider the inherent assumptions, biases, and unique features related to SDMs and SDMs of invasive species.

Résumé

Les plantes envahissantes constituent une menace croissante pour la biodiversité indigène à l'échelle mondiale. Une gestion efficace dépend de prévisions précises sur leur propagation. Cependant, il est difficile de modéliser la distribution géographique des espèces envahissantes, en particulier avec des méthodes telles que les modèles de distribution corrélative des espèces (SDM). Les SDM fonctionnent sous l'hypothèse que les espèces sont en équilibre avec leur environnement (c'est-à-dire qu'elles occupent tous les environnements appropriés); cette hypothèse est particulièrement susceptible d'être violée pour les espèces qui font activement la colonisation d'environnements appropriés. Les SDM supposent également que les contraintes environnementales sont les facteurs les plus importants déterminant la distribution d'une espèce. Cependant, ces hypothèses ne sont pas généralement évaluées et, lorsqu'elles ne sont pas respectées, peuvent avoir des conséquences sur la fiabilité du modèle. J'ai étudié les performances et l'équilibre de SDM chez la vigne invasive *Vincetoxicum rossicum* dans le nord-est de l'Amérique du Nord. La *Vincetoxicum rossicum* a un vaste historique de données d'occurrences dans son aire d'envahissement, ce qui m'a permis d'observer les tendances de l'équilibre et de la performance du modèle sur une échelle de temps relativement longue. J'ai testé les hypothèses que: 1) les espèces envahissantes s'approchent de l'équilibre dans leur espaces environnementales et géographiques au fil du temps; 2) les performances du SDM s'amélioreront à mesure que la distance par rapport à l'équilibre diminuera; et 3) l'expansion de l'aire de répartition aux premiers stades d'une invasion est principalement relié à la dispersion plutôt qu'aux contraintes environnementales, alors que l'inverse est vrai aux stades ultérieurs. J'ai trouvé que *V. rossicum* a atteint l'équilibre dans l'espace environnemental, mais qu'elle étend encore son aire géographique. Les performances des SDMs étaient médiocres au cours des 30 premières années suivant son introduction, mais elles se sont ensuite améliorées à mesure que *V. rossicum* s'approchait de l'équilibre environnemental. Les SDMs ont été surclassés par des modèles de dispersion spatiale au cours de la période la plus ancienne, mais l'inverse était vrai pour toutes les périodes subséquentes. Dans l'ensemble, ces résultats suggèrent que la distribution de *V. rossicum* devient plus stable et plus prévisible au fil du temps et que les modèles construits à l'aide des données les plus récentes pour cette espèce seront les plus transférables dans le temps et dans l'espace. De plus, mes résultats soulignent la nécessité pour les chercheurs qui modélisent la distribution des espèces envahissantes de prendre en compte les hypothèses, les biais et les caractéristiques uniques inhérents aux SDMs et les SDMs des espèces envahissantes.

Contents

List of Tables

List of Figures

as yellow open circles superimposed on the spatial dispersal models (panels E–H). The extent of all models is shown in white.

Introduction

Invasive plants threaten ecosystem function and native biodiversity globally (Vilà *et al.*, 2011). Their impact is likely to be exacerbated by increasing trade and travel, which facilitates their introduction and spread (Jenkins, 2002; Meyerson & Mooney, 2007; Seebens *et al.*, 2015). Invasive plants negatively impact native plant and animal biodiversity by competing for resources with native species (Vilà *et al.*, 2011), and altering or degrading physical characteristics (e.g., soil pH, nutrient concentrations; Gibbons *et al*., 2017) and species interactions (Dangremond, Pardini & Knight, 2010). Together, these changes displace native plant and animal species from sites that were historically suitable (Powell, Chase & Knight, 2011).

Fortunately, invasive plants can be managed—to varying degrees of effectiveness—using practices such as chemical (e.g., herbicides; Carlson and Gorchov 2004), mechanical (e.g., mowing; Wilson and Clark 2001) and biological (e.g., herbivorous natural enemy; Clewley et al., 2012) controls. A useful tool for invasive species management is to describe and track the geographic distributions of invasive plant species to identify and target areas for the application of these management practices. Therefore, one of the primary aims of invasion biology is to predict the expansion of invasive plants in their introduced ranges.

Correlative species distribution modelling is a commonly used approach for describing the geographic distribution of invasive species. Species distribution models (SDMs) correlate occurrence data (presence-only or presence/absence) with environmental characteristics to model environmental suitability across a defined geographic extent. SDMs have been used to identify the most important abiotic and/or biotic factors that determine the geographic distributions of invasive species (Lewis *et al.*, 2017). This is particularly important for invasive species, because these factors can define potential constraints on their spread (Kearney & Porter, 2004; Dainese, Kühn & Bragazza, 2014). For example, SDMs revealed that elevation, temperature, and bedrock geology could be used to identify unsuitable areas for zebra mussel (*Dreissena polymorpha*) range expansion in the US (Drake & Bossenbroek, 2004). SDMs have also been used to predict the success of potential biological controls (e.g., herbivorous insects) by determining the amount of spatial overlap between their geographic distribution and that of their target invasive plant species (Trethowan, Robertson & McConnachie, 2011). Finally, SDMs have been used to predict how the distribution of invasive species may change under future climate change scenarios (Taylor & Kumar, 2013).

Regardless of the objective for using SDMs, building accurate, reliable models can be difficult for invasive species. The consequences of using inaccurate models to inform invasive species management decisions about where range expansions are likely to be costly. Management of invasive species is expensive (Colautti *et al.*, 2006), and distribution models that incorrectly predict at-risk areas can lead to management of places that may not need it. Alternatively, managers may miss the window for early intervention in areas that are truly at-risk. Early intervention is considered to be the most effective strategy in controlling invasive species (Venette *et al.*, 2021). Therefore, accurately predicting the range expansion of invasive species is key and thus, the evaluation of SDMs to predict current distributions accurately is crucial before projecting these models into the future.

Building accurate SDMs can be difficult for invasive species for several reasons. Foremost, SDMs assume that the species is at equilibrium with its environment. A species is at equilibrium when it occurs in all suitable areas, and is absent from all unsuitable areas (Araújo & Pearson, 2005). This assumption is violated when modeling distributions of newly introduced species, in particular, because they are still in the process of colonizing all suitable environments in the invaded range (Elith, Kearney & Phillips, 2010; Gallien *et al.*, 2012). Models constructed from observations collected before an invading species reaches equilibrium are expected to produce unreliable results, underestimating their potential range (Elith *et al.*, 2006; Bradley, Early & Sorte, 2015).

Despite the acknowledged need to investigate the equilibrium assumption when using SDMs for invasive species (Václavík & Meentemeyer, 2012), there are still relatively few studies which evaluate this assumption, and methods for testing this assumption are still being developed. One approach is to experimentally identify the physiological limits of a species (e.g., lethal temperatures), and use this information to map the geographic distribution of suitable conditions (Monahan, 2009). However, this requires knowing which physiological factor(s) are limiting, and designing experiments that adequately test them. Alternatively, several studies have used data on the distribution of a species in its native and introduced range to define its global environmental equilibrium: if the environment it occupies in its invaded range matches that of its global range, it can be considered to be at equilibrium (e.g., Roura-Pascual *et al.*, 2009; Gallien *et al.,* 2012). This approach assumes that the species is at equilibrium in its home range, and that its equilibrium remains identical in its invaded range, either or both of which may be false (Early & Sax, 2014). Even when these assumptions hold, this approach cannot be used for species which are poorly documented in their home range.

Both of these approaches assess equilibrium as a binary condition: a species is in equilibrium, or it is not. However, neither the species nor its environment are static (Holyoak, Caspi & Redosh, 2020), so the assumption of equilibrium is dependent on the spatial and temporal context in which it is determined. Araújo and Pearson (2005) reframed the critical question as “how distant from equilibrium are current distributions?” In the context of invasions, where a species is actively expanding its range in a novel area, that question can be further refined to “how is the distance from equilibrium changing over time?”

In this light, we predict a species’ distance from equilibrium to decline as it proceeds through the stages of invasion (Theoharides & Dukes, 2007). During initial colonization, the occupied area is small, representing an essentially random subset of the appropriate area; at this point, the distance from equilibrium is large, and SDMs are expected to be inaccurate (Elith *et al.*, 2006). The establishment stage is characterized by the development of self-sustaining local populations, and a slight expansion in geographic range (Theoharides & Dukes, 2007); we expect correspondingly slight improvement in the predictive ability of SDMs. At the spread stage, the new invader is rapidly expanding its range; distance to equilibrium should rapidly decline in this period, with SDM performance improving markedly, accurately predicting spread to new geographic locations. However, if SDMs require a species to be close to equilibrium before they yield accurate results, and equilibrium is defined as occupying “all suitable areas” (Araújo & Pearson, 2005), then there is little value in using them to anticipate the spread of invasive species. By the time the species has occupied all suitable areas, there will be nowhere left for it to spread to.

Resolving this issue requires that we distinguish between the geographic and environmental distribution of a species. These two concepts are related (and often treated synonymously in the literature; Araújo & Pearson, 2005; Gallien *et al.*, 2012), but not identical: different geographic locations may share identical environmental conditions. Consequently, an invading species may reach environmental equilibrium before it reaches its maximum geographic range (or geographic equilibrium). Understanding these distinctions are especially important when projecting the geographic distribution of invasive species into the future under climate change scenarios where the species' environmental tolerances are expected to remain the same (assuming no adaptation occurs) while the geographic distribution they occupy is likely to change (Wiens & Graham, 2005).

Another challenge to building accurate SDMs for invasive species is the source of the data that are used for the evaluation of SDMs. SDMs assume calibration data (i.e., data used to build an SDM) and evaluation data (i.e., data used to evaluate SDM performance data) are independent. However, independently gathered data are rarely available for the evaluation of SDMs (but see McCune 2016). The dearth of independent datasets has led to the common practice of evaluating models using a random subset of occurrence records withheld from the same dataset used to build the model (i.e., cross-validation, Radosavljevic and Anderson 2014). However, this approach can inflate estimates of model accuracy due to spatial autocorrelation (i.e., the tendency for records that are closer together to be more environmentally similar) between the calibration and evaluation data (Hijmans, 2012). For invasive species, the issue of spatial autocorrelation is even more likely to be an issue since the initial expansion from sites of introduction is thought to occur through spatially autocorrelated dispersal, independent of human-mediated dispersal (Andow *et al.*, 1990; With, 2002; Václavík, Kupfer & Meentemeyer, 2012).

There are two main ways that modelers have dealt with the role of spatial autocorrelation for invasive species. First, some modellers have evaluated the predictive ability of SDMs for invasive species by using the most recent, temporally independent occurrence data as the evaluation dataset (Barbet-Massin *et al.*, 2018). This is a step in the right direction; however, temporally independent datasets can still be spatially autocorrelated, so additional means of dealing with spatial autocorrelation are needed. Second, for expanding species, modellers have compared SDM predictions to simple spatial dispersal models – models that infer species distributions based solely on spread from occupied sites (Rodríguez-Rey, Jiménez-Valverde & Acevedo, 2013). If a simple dispersal model can predict a species' future distribution as well as (or better than, as in Rodríguez-Rey *et al.*, 2013) an SDM, it suggests the performance of the SDM is likely to be an artifact of spatial auto-correlation, rather than an accurate reflection of the environmental niche of the species (Bahn & Mcgill, 2007). However, these comparisons remain rare. Studies that explicitly examine the relative performance of spatial models and SDMs are needed, especially in the context of invasive species.

Rigorously testing assumptions of SDMs requires long-term and broad-spatial scale data, yet such data can be difficult to acquire. Here, I test assumptions and explore issues of SDMs that are particularly challenging for invasive species using *Vincetoxicum rossicum* (Kleopow) Barbar. (commonly known as pale swallow-wort or dog-strangling vine). *Vincetoxicum rossicum* is an herbaceous, perennial, twining vine native to southwestern Russia and Ukraine. It was introduced to North America in the late 1800s as an ornamental plant (Pringle, 1973; DiTommaso, Lawlor & Darbyshire, 2005). Since then, it has become a serious threat to invaded ecosystems, especially in New York State and Ontario, where it was first introduced and has since become widespread (Sheeley & Raynal, 1996; DiTommaso *et al.*, 2005). As *V. rossicum* has invaded a wide variety of habitats (DiTommaso *et al.*, 2005), has a wide geographic distribution, and has a detailed history of its invasion with over 100 years of occurrence records including the introduction sites, it is an ideal species to test key assumptions of SDMs. Moreover, *V. rossicum* has so far been resilient to control methods and regenerates following most damage, making it incredibly hard to eradicate from invaded areas (DiTommaso *et al.*, 2005; McKague & Cappuccino, 2005a; Miller & Kricsfalusy, 2008). As a result, there is a need to identify undetected populations and to predict areas that are likely to be invaded to target mitigation efforts.

In this study, I test three hypotheses. First, that invaders will expand in geographic and environmental space from their point(s) of introduction and over time will approach their environmental and geographic equilibrium. To test this hypothesis, I compare *V. rossicum's* expansion in environmental and geographic space between historical (starting from its introduction in 1890 up to 2010) and current time periods (2011–2020). I do not know the true environmental and geographic equilibrium of the species; instead, I will infer it from the rate of change in environmental and geographic space occupied by *V. rossicum.* No expansion in the occupation of either space over several time periods will indicate that the distance to equilibrium has declined to 0 suggesting it has reached equilibrium. I further hypothesize that the performance of SDMs in predicting the current geographic distribution of *V. rossicum* will increase as the species' range approaches equilibrium with the environmental conditions.

Third, I hypothesize that range expansion in the early stages of an invasion is primarily a function of dispersal rather than environmental constraints, whereas range expansion in later stages is primarily a function of environmental constraints rather than dispersal. To test this hypothesis, I compare the performance of SDMs to spatial dispersal models, using only occurrence records, for several time periods over the past 130 years. If range expansion in the early stages of *V. rossium's* invasion is primarily a function of dispersal rather than environmental constraints, then my spatial dispersal models should outperform SDMs in early time periods. Correspondingly, if range expansion in later stages is primarily a function of environmental constraints instead of dispersal, then in later time periods my SDMs will outperform spatial dispersal models. Finally, I also explore the consequences of these processes for model performance by examining how model performance changes over time.

To my knowledge, this study represents one of the first to explore how the distance to equilibrium changes over time in both environmental and geographical space, while also addressing the concerns inherent to modeling invasive species distributions. Additionally, to address many concerns related to distribution modeling for invasive species (e.g., sampling bias, non-independent evaluation data), I use several best practice approaches (e.g., spatial thinning). I address the need for independent evaluation data by testing my models using data that are both temporally independent and derived from an independent source. By comparing SDM performance to simple dispersal models, I also address the role of spatial autocorrelation between the calibration and evaluation datasets.

Methods

1. Study Species

In its native range in southwestern Russia and Ukraine, *V. rossicum* is found in scrubland, on the slopes of ravines (Pobedimova 1952), steppe and forest-steppe (Visulina 1952; Pobedimova 1978; as cited in DiTommaso *et al.,* 2005). Unlike most other invasive species, there is currently very little known about the native distribution. As of March 2021, only 13 specimens are available in Global Biodiversity Information Facility database (GBIF) for the native range (all of them added in 2020), and I could not locate any additional specimens from other collections (FI, K, KW, L, MA, P, PRC, PR, S, acronyms follow Thiers 2021). Therefore, I was not able to robustly model the native range.

In North America, *V. rossicum* can invade environments with a wide variety of soil types, disturbance levels, moisture conditions, and light exposure, from open fields to heavily shaded understories to exposed sheets of limestone (DiTommaso *et al.*, 2005; Kricsfalusy & Miller, 2010). Once established, *V. rossicum* forms dense stands via local seed dispersal (DiTommaso *et al.*, 2005; Casagrande *et al.*, 2011).

In the habitats it has invaded, *V. rossicum* has had direct and indirect effects on native biodiversity. Dense monocultures of this twining vine can smother and displace native plants from habitats (Cappuccino, 2004; DiTommaso *et al.*, 2005). *Vincetoxicum rossicum* is also associated with a reduced variability in the composition of soil bacterial (Bugiel *et al.*, 2018) and fungal (Day, Antunes & Dunfield, 2015) communities in invaded habitats, and anecdotally, its presence has been shown to deter breeding grassland birds (DiTommaso *et al.*, 2005). Some of these direct impacts have also led to indirect negative consequences. For example, native plants may be further impacted if the soil microbial community is altered to be less suitable (Bongard *et al.*, 2013) or if large herbivores increase their herbivory on native plants due to avoiding grazing on toxic *V. rossicum* (DiTommaso *et al.*, 2005; Averill *et al.*, 2016). Furthermore, reductions in native plants can lead to reductions in associated insects. In fact, insect biodiversity in stands of *V. rossicum* is significantly reduced compared to areas populated with native plants (Ernst & Cappuccino, 2005).

2. Overview

This study is composed of three parts: i) equilibrium analysis: assessing change in the distribution of *V. rossicum* in environmental and geographic space over time; and modelling the historical and current geographic distribution of *V. rossicum* using ii) species distribution models (SDMs), and iii) spatial dispersal models. The equilibrium analysis was based on 10-year time intervals, whereas the modeling was based on 30-year time intervals. I evaluated the performance of the SDMs and spatial dispersal models based on their ability to accurately predict occurrences from its current range (2011–2020) using independent observation records from iNaturalist.

2. Data

2.1 Occurrence data

I created two datasets of *V. rossicum* occurrence records: a calibration and evaluation dataset. First, to create a calibration dataset, I compiled occurrence records in North America from 1889 to 2010 using the GBIF database (excluding iNaturalist records, accessed Feb 2020) as the primary source, and supplementary records were provided by Dr. Jenny McCune (McCune, 2016), Dr. Richard Dickinson (Dickinson *et al.*, 2021), the Ontario Natural Heritage

Information Centre (Mike Oldham, unpublished data), and survey data from the Nature Conservancy of Canada (Mhairi McFarlane, unpublished data), Paul O'Hara (unpublished data), and the following herbaria: DAO, TRT, TRTE, MT, PO, CAN, QUE, UAC, UWO, ACAD, QK, QFA, LKHD, UNB, HAM, OAC, UBS (acronyms follow Thiers 2021). Some, but not all, of the herbarium records I collected were also present in the GBIF data. These duplicates were removed prior to analysis. I included all GBIF records that had lat-lon coordinates, and georeferenced 205 additional records which had location data adequate for georeferencing. To reduce sampling bias, I thinned records based on a 10 km^2 grid to match the environmental variables (see below), such that there was only one occurrence record per 10 km^2 pixel. After this filtering, there were 20 occurrence records from 1890–1920, 42 from 1921–1950, 92 from 1951–1980 and 218 from 1981–2010.

As an independent evaluation dataset used for testing the models, I used research-grade occurrence records available from iNaturalist, a citizen science database. Research-grade records are those that have been verified by two separate reviewers and have GPS coordinates provided. Further validation of these records was provided by visual inspection of all images in the iNaturalist database (1608 records) by botanists with expertise identifying *Vincetoxicum* (Dr. Subbaiah Mechanda and Jean-François Beaulnes). To maintain temporal independence from all of my calibration datasets, I limited my evaluation dataset to include only records from 2011 to 2020. Observations were thinned using the same procedure as for herbarium records, resulting in a final evaluation dataset of 473 occurrence records.

2.2 Background extent

I defined my background extent for all analyses based on ecoregions currently occupied by *V. rossicum.* Given my original motivation to forecast the present and future distribution of *V. rossicum* based on historical records, I included the ecoregions where *V. rossicum* occurs, and extended this to all adjacent ecoregions to allow for further invasion (Merow, Smith & Silander, 2013). The ecoregions I used are defined by the United States Environmental Protection Agency's mid-level (level II) detail for Ecoregions of North America (Omernik & Griffith, 2014).

All environmental variables were clipped to match this extent. To avoid excluding occurrences along shorelines, all raster cells that overlapped with land to any degree were included. The resolution of all rasters included in the models was 10 km^2.

2.3 Environmental data

I created 30-year climate normals for northeastern North America using historical monthly climate grids derived from the Canadian Forest Service of Natural Resources Canada (McKenney *et al.*, 2006) from 1890-2010 (1890–1920, 1920–1950, 1950–1980, 1980–2010). I averaged monthly minimum temperature, maximum temperature, and precipitation data across the 30-year period. From these, I created 19 bioclimatic variables modelled after the Worldclim variables using the *biovars* function in R (package 'dismo' v1.1.4 by Hijmans *et al.,* 2017; Table 1). These rasters were available at a resolution of about 10 km^2.

To characterize the relationship between the distribution of *V. rossicum* and soil properties, I included edaphic variables from the Regridded Harmonized World Soil Database (HWSD v1.2; Wieder *et al.,* 2014). Since *V. rossicum* can survive in very shallow soils (Weston, Barney & DiTommaso, 2005) and the bulk of root mass occurs in the top 15 cm of soil (personal observation), I focused on soil variables found in the topsoil layer. The HWSD describes the topsoil layer as the top 30 cm. To capture topsoil textures preferences of *V. rossicum*, I included the percent clay, silt and sand content. I also included pH of the topsoil layer to see if *V. rossicum* is limited by any pH levels represented in my study's extent. These variables were available at a resolution of about 8 km^2 and were rescaled to match the resolution of the climate and anthropogenic variables using bilinear interpolation using the *projectRaster* function in R (package 'raster' v 3.1.5 by Hijmans *et al.,* 2018).

To capture anthropogenic influences on *V. rossicum's* distribution, I considered the role of human population density in all of the models. Human population density in this context is a proxy for the level of anthropogenic disturbance, since higher human population densities tend to be associated with more traffic, roadways, pollution, construction, etc. Population density, measured as inhabitants/km^2 per raster cell, was available at 10-year increments through the HYDE v3.2.1 database (Goldewijk *et al.*, 2017), for the time period I investigated. I used population density estimates for the third decade of each time period (i.e., 1910 for the 1890–1920 period). These data were available at a 10 km^2 resolution, so they did not require rescaling.

3. Equilibrium Analysis

3.1 Overview

To determine how *V. rossicum*'s distribution in environmental and geographic space has been changing since its initial introduction in the late 1800s, I used the 'ecospat' package (v 3.1 by Broennimann *et al.,* 2020) in R. 'ecospat' calculates a species' "expansion", "stability" and "unfilling" (i.e., contraction) in environmental or geographic space by comparing the density of occurrence records of a species between two contexts, typically geographic areas. For invasive species, this analysis is conventionally used to compare the environment occupied in native and invaded ranges. This comparison is used to evaluate whether the species has maintained its native environmental niche in its introduced range (Eckert *et al.*, 2020; Manzoor *et al.*, 2020; Srivastava *et al.*, 2020). Here, I alter this analysis by comparing the density of occurrence records between two time periods (historic and current). This allowed me to evaluate the change in distribution in environmental and geographic space over time for *V. rossicum* in North America.

I compared the historical distribution in environmental and geographic space in 10-year time intervals from 1890-2010, to its current distribution (i.e., the evaluation dataset from 2011–2020). This allowed me to evaluate the amount of expansion, stability and unfilling in environmental and geographic space across its invasion history in northeastern North America. Each 10-year subset of records was thinned to a 10 km^2 resolution (i.e., single record per pixel) to match the environmental data. Records were then accumulated across the 10-year time intervals, such that the total number of records for the historical period of interest was a sum of the records from each 10-year interval up to the final year (presented in Fig. 2-5).

3.1.1 Environmental space

The environmental space was represented by the first two components of a principal components analysis (PCA) of environmental variables, refer to Table 1 for a full list of variables. I used all 24 environmental variables from each 30-year climate interval to ensure that the environmental space was defined by the same variables across all time periods and to provide a thorough description of the range of potential environmental conditions. Locality and time-period specific environmental data was extracted for each occurrence record. Therefore, environmental values for the first three 10-year time intervals of records (i.e., 1890–1900, 1901–1910, 1911–1920) were each extracted from the 1890–1920 environmental raster stack. Environmental values for the next three record intervals were extracted from the 1921–1950 climate interval, and so on. Therefore, the historical environmental space used for an analysis included the complete environment data from each 30-year period that was associated with the range of historical occurrence records.

The environmental values for the current environmental space (associated with 2011–2020 occurrence records) were extracted from the most recent 30-year climate interval (1980–2010), since environmental data for the current time period could not be acquired from the same data sources. Thus, for example, the PCA describing the total environmental space for a comparison between historical records from 1890–1930 and the 2011–2020 records was a single PCA of environmental values from the first two climate intervals (1890–1920 and 1921–1950) and the 1980–2010 interval (Fig. 2, 1930 panel).

3.1.2 Geographic space

Geographic space is defined by the longitude and latitude coordinates of the study extent, northeastern North America, and is therefore constant. The longitude and latitude of the historic and current occurrence records for *V. rossicum* determine the geographic space occupied by *V. rossicum*.

3.1.3 Analysis details

I used ‘ecospat’ to measure *V. rossicum*’s stability, expansion, and unfilling in environmental and geographic space between each historic time period and the current time period (n=12 comparisons). Here, ‘stability’ represented the space that was occupied in both the historic and current time periods. It was calculated as the number of occupied cells in environmental or geographic space in both historical and current time periods divided by the total number of cells occupied in the current space (including expansion cells). ‘Expansion’ was the space that was only occupied in the current time period; it was calculated as the number of cells occupied only in the current time period divided by the total number of cells occupied in the current time period (including stability cells). These values were then multiplied by 100 to give percentages, therefore expansion and stability values add to 100. ‘Unfilling’ was space that was only occupied in the historic time period and was calculated as the number of cells occupied only in the historic time period, divided by the total number of cells occupied in the historic time period (including those shared with the current time period). The unfilling index was also multiplied by 100. Stability, expansion and unfilling values were weighted by the occupancy density estimate for cells in the historic time period.

4. Modelling

4.1 Species distribution models

SDMs were built using the species distribution modelling software, Maxent (version 3.4.4), through the R package ‘dismo’ (version 1.1.4 by Hijmans *et al.,* 2017). Maxent uses a maximum entropy approach to model species distributions (Phillips, Anderson & Schapire, 2006). Maxent relates the species’ presences and pseudo-absences (Table 3) in a pre-defined area—in this case, the background extent—to the available environmental variables to model environmental suitability across geographic space. It is a commonly used approach for presence-only data (Renner & Warton, 2013) and has been shown to out-perform other presence-only models (Elith *et al.*, 2006). It also has built-in methods of handling multicollinearity (Elith *et al.*, 2010; Feng *et al.*, 2019) and can handle complex, non-linear relationships between the predictors and response (Elith *et al.*, 2011). It also performs well with small sample sizes (Hernandez *et al.*, 2006; Wisz *et al.*, 2008), and the output is intuitive.

Given the large number of potential variables (n=25; Table 1) for each time period, I took a proactive approach to reduce potential overfitting and multicollinearity. Multicollinearity, which can make it difficult to tease apart the influence of correlated variables on the dependent variables, may be less of a concern for Maxent model performance (Feng *et al.*, 2019). However, it is still generally recommended to exclude strongly correlated variables from models, otherwise more caution should be taken when interpreting variable importance (i.e., the percent contribution of a variable to model fit; Phillips *et al.,* 2006). Therefore, I conducted a correlation analysis of all 24 variables for each time period based on the variance inflation factor (package 'usdm' v 1.1.18, Naimi 2017) and excluded variables that were collinear (VIF > 5). There were nine final variables (5 bioclimatic variables, 3 edaphic variables, and population density) for each time period (refer to Table 2 for list of variables included in each model).

4.2.1 Maxent model parameters:

I used *ENMevaluate* ('ENMeval' package version 0.3.0 by Muscarella *et al.,* 2014) to parameterize the models as adjusting the parameters according to the goals of the study can improve model performance (Morales, Fernández & Baca-González, 2017). Specifically, I adjusted the regularization parameter, feature classes and combinations, data partitioning method, background locality selection, and type of output (Table 3). Model parameterization resulted in a set of 96 candidate models. From this set of candidate models, the optimal model for each time period was selected based on the lowest Aikake information criterion (AIC). In some cases, there were two models with the same AIC, in this case, the model with fewer feature classes was chosen. I then projected the optimal models for each time period into the most recent environmental conditions (1981-2010 environment) to make prediction maps of estimated environmental suitability. Estimated environmental suitability represents a probability that the characteristics in a locality match a species' environmental tolerances and ranges from 0 (unsuitable environments) to 1 (highly suitable environments).

4.3 Dispersal model

To generate a simple spatial dispersal model for each time period, I followed the procedure of Rodríguez-Rey et al. (2013). I used the calibration occurrence records for each time period as the points of origin. I calculated the distance from each surrounding cell to the nearest origin point, and standardized these distance values between 0 and 1 [(distance-minimum distance)/(maximum distance-minimum distance)]. I then inverted the standardized distances such that a value close to one indicates the highest likelihood of occupancy (i.e., regions closest to any origin point).

4.4 Testing the predictive ability of the models

I evaluated the performance of the Maxent and dispersal models by their ability to predict the current geographic distribution of *V. rossicum* according to research-grade iNaturalist records from 2011–2020. I used the continuous Boyce index (‘ecospat’ package; (Broennimann, Di Cola & Guisan, 2020) as a measure of predictive accuracy.

The continuous Boyce index is a Spearman correlation between ranked classes of model predictions and the observed proportion of presences located in cells in those classes. Rather than looking at fixed classes—whose boundaries and width can easily influence the Boyce index (Hirzel *et al.*, 2006)—it implements a moving window of ranked classes as follows: the cells with the lowest predicted suitability values (typically the bottom 0–10%) are identified, and the number of presences in these locations is recorded. This is repeated for subsequent windows (i.e., 1–11%, 2–12% up to 90–100%). When finished, the correlation between window rank and the number of actual presences in that window is calculated. The Boyce index varies between -1 and +1, where good model performance is reflected in a positive correlation (i.e., the number of presences increases from the lowest-ranked window to the highest-ranked window). A correlation near zero indicates presences are distributed randomly with respect to model predictions and negative values indicate that presences are more frequent in cells with low predicted suitability. The continuous Boyce index is a presence-only evaluation metric, which is insensitive to prevalence (Hirzel *et al.*, 2006), making it one of the most appropriate metrics for evaluating Maxent model performance (Di Cola *et al.*, 2017).

With the Boyce index, I made two types of comparisons. First, for each modeling approach, I compared this metric across time periods. Second, for each time period, I compared the Maxent model to the spatial dispersal model.

Results

Equilibrium analysis

i) Environmental space

The ‘ecospat’ analysis showed rapid expansion of *V. rossicum’s* distribution in environmental space in its invasive range from 1900–1940, and again 1950–1970 (Fig. 2, 4), after which it plateaued close to zero from 1970 to 2010 (Fig. 2, 4). This suggests it has reached environmental equilibrium. Even the small areas of expansion between 1970 and 2010 (pink cells, Fig. 2) decreased over those 40 years. Expansion by 1970 was less than 3%, decreasing to and stabilizing around 0% by 2000. Overall, expansion in environmental space decreased over time.

ii) Geographic space

Except in the first time period, the expansion of *V. rossicum*’s distribution in geographic space followed a similar, but slower, progression over the past century. Expansion steeply decreased immediately after introduction, and by 1970 expansion was less than 5%, plateauing at 1% by 1990 (Fig. 3, 5). The low levels of expansion in recent years of *V. rossicum*’s distribution in geographic space suggests it is also stabilizing.

Maxent and spatial dispersal model performance

i) Maxent

The ability of Maxent models to accurately predict the current geographic distribution of *V. rossicum*, as defined by 2011–2020 iNaturalist records, increased over time (Fig. 6; Boyce_{1920} =0.74, Boyce_{2010} = 0.99).

The earliest model (1920) estimated that a much broader geographic range was suitable compared to later models (Fig. 7). For example, it incorrectly predicted suitable environment in Louisiana, Mississippi, Alabama, and Florida, which the 1980 and 2010 models excluded as suitable environment. Conversely, the 1920 model failed to identify Sudbury and records on the east coast of Canada as having suitable environment; *V. rossicum* has recently colonized both areas. The most recent Maxent model (2010) outperforms all other models, and the prediction map includes Sudbury and areas of New Brunswick close to the most recently colonized locations (Fig. 7).

i.i) Predictors of V. rossicum geographic distribution Occurrence data

The best predictor of *V. rossicum's* geographic distribution across all time periods based on the Maxent models was human population density, which contributed 35-43% to model fit, depending on time period (Table 4). Across all time periods except the first one, precipitation seasonality was the second most important variable, where low precipitation variability was positively associated with *V. rossicum* occurrence. In 1920, the second most important variable was mean diurnal temperature range. Edaphic variables contributed the least to model predictions for any time period (Table 4).

ii) Spatial dispersal

The performance of the spatial models increased from 1920 to 1950 (Fig. 6, Boyce_{1920}=0.82, Boyce_{1950}=0.89), then declined in all subsequent time periods. The most recent model (2010) had the worst performance (Fig. 6, Boyce_{2010}=0.75), although it was still better than random (i.e., Boyce Index > 0).

iii) Comparison between Maxent and spatial dispersal

Maxent model performance relative to spatial model performance depended on the time period investigated. In the earliest time period (1920), the spatial dispersal model outperformed the Maxent model (Fig. 6). In the next time period (1950), the Maxent and spatial models performed similarly (Maxent slightly higher than spatial dispersal model; Fig. 6). In the two last time periods (1980, 2010), Maxent models outperformed the spatial dispersal models (Fig. 6).

iv) Equilibrium and SDM performance

SDM performance increased as *V. rossicum* distribution in environmental space approached equilibrium. There were similarities in the early increase of both performance and stability for the 1920–1980 models and both began to plateau around 1980 (Fig. 4 versus 6).

Discussion

Here, I quantify the expansion of an invasive plant species in environmental and geographic space since its introduction and determine if its current geographic distribution in North America is sufficiently stable to build a reliable SDM. First, my results show that, following 80 years of rapid expansion, the environmental space occupied by *V. rossicum* has remained relatively stable since 1970, suggesting that it has reached its environmental equilibrium. The trend in geographic space is similar, although most of the expansion occurred in the two decades after introduction, and a plateau was not reached until 1990. Second, increases in SDM performance were consistent with *V. rossicum*'s approach to equilibrium in geographic and environmental space. Third, in the earliest time period, Maxent models were outperformed by spatial dispersal models, however, the reverse was true for all subsequent time periods. These results support my hypothesis that *V. rossicum's* range expansion is primarily a function of dispersal in the earliest stages of its invasion, but in later stages is primarily a function of environmental constraints.

Equilibrium Analysis

According to my analysis, it took 80 years after its introduction for *V. rossicum* to approach its environmental equilibrium. Its distribution in environmental space has been largely unchanged over the past 50 years (including the current time period up to 2020), despite small increases in its geographic distribution over the same period. This suggests that *V. rossicum* has reached its environmental equilibrium, and now occupies the full range of suitable environments.

It is difficult to compare this result to other species given how rarely equilibrium has been quantified and the differences in how it is measured. In one example, Gallien et al. (2012) studied a suite of invasive species in the French Alps. They determined that witchgrass (*Panicum capillare* L.), which had been present for 250 years, had reached its equilibrium and was unlikely to continue spreading; however, they did not assess how long it took to do so. In contrast, Tree of Heaven (*Ailanthus altissima* (Mill.) Swingle), with an even longer history in the area (300 years), had yet to reach equilibrium. They defined equilibrium based on comparisons of occupancy in the invaded range relative to predictions based on native and invaded range SDMs. In contrast to my study, this definition combines both environmental and geographic space, and does not incorporate temporal trends. My results suggest that a long temporal perspective is needed to be able to test the assumption that introduced plant species are at equilibrium with their environment and that through-time niche dynamics need to be measured. Future studies should investigate invasive plant species with long invasion timelines to see when (or if) they have reached environmental equilibrium.

My results revealed differences in the dynamics between the expansion in environmental and geographical space. I found that *V. rossicum* reached a plateau in environmental space before it did so in geographic space (1970 vs 1990, Fig. 2 & 3), and it has been more stable in environmental space than in geographic space over the past 50 years (expansion in environmental space ranges from 0-3, while expansion in geographic space ranges from 1-5% since 1970, Fig. 2 & 3). My results suggest that in recent decades, *V. rossicum* expanded its geographic range by establishing in new areas that share environmental conditions with locations it already occupied. Thus, its recent extension of its geographic range into central Ontario and eastern Canada is not associated with an increase in its distribution in environmental space. This highlights the need to consider environmental and geographic equilibrium separately.

Although my results suggest that the distribution of *V. rossicum* in environmental space has stabilized and that its distribution in geographic space is stabilizing, there are caveats to these conclusions because of the period of reference and spatial scale I used. First, environmental stability and expansion (as defined in the 'ecospat' package) were calculated relative to *V. rossicum*'s current geographic distribution. Taking this approach with any invasive species that is actively expanding its geographic distribution will produce a stability value that increases to a maximum value of 100, even if the species will continue to expand its distribution in the future. Thus, demonstrating that a species has reached equilibrium requires that: i) the stability value reaches 100 and ii) that there is consistently no expansion over several time periods. My data show this: for the last 50 years, expansion in environmental space has been at low, stable levels (Fig. 4).

Instead of being indicative of equilibrium, the observed plateau in environmental space could also be the consequence of dispersal barriers or limitations that prevent *V. rossicum* from reaching locations with novel, but suitable, environments that would delay reaching environmental equilibrium. *V. rossicum*'s ability to cross large areas of potentially unsuitable environments may be limited by intrinsic challenges to seed establishment. For example, smaller *V. rossicum* seeds are capable of dispersing longer distances (Ladd & Cappuccino, 2005; DiTommaso *et al.*, 2018), but at the cost of reduced viability; although, in this case, the trade-off between seed weight and quality is potentially negligible (Ladd & Cappuccino, 2005). Additionally, *V. rossicum*'s strong association with human populations may indicate that areas with less anthropogenic disturbance provide less suitable environments. These unsuitable areas may act to separate patches of suitable environments and inhibit colonization. Areas with little human activity are also less likely to create opportunities for human-mediated dispersal events, further limiting spread. In addition, evidence suggests that disturbance facilitates *V. rossicum*'s ability to invade vegetation communities (Averill *et al.*, 2010). While *V. rossicum* was able to invade patches subject to a range of disturbance regimes (Averill *et al.*, 2010), its growth rate was lower in relatively intact locations. This reduction may have been sufficient to reduce the rate of spread enough to temporarily curtail its expansion in environmental space. If such barriers to dispersal are constraining the spread of *V. rossicum*, changing land use that disrupts native plant communities, or increases anthropogenic disturbance, may release the constraint. This highlights the need for detailed study of invasion histories, to better link the spread and impact of invasive species to historic changes in land use patterns.

Another consideration when interpreting the stabilization I documented in environmental space, is the resolution of my analysis. Other studies have shown that equilibrium assessments done at coarser scales, as I have done here (i.e., 10 km^2), may have different conclusions at finer scales (e.g., 30 m^2; Roura-Pascual *et al.* 2009). This could be, in part, because the relative importance of environmental factors can change across spatial scales (Roura-Pascual *et al.*, 2009). For example, climate may significantly influence equilibrium on large spatial scales, but have little impact at local spatial scales, whereas proximity to roadways may only play a role in equilibrium at local scales. This could mean that *V. rossicum* could still be expanding to suitable environments at finer scales (e.g., 2 km from roadways) than I could detect here. Therefore, future studies need to consider how spatial and temporal scale can impact their assessments of equilibrium and influence the conclusions they can make regarding a species' environmental niche.

My analysis also assumes that *V. rossicum* is genetically homogenous and that it is not adapting to environmental conditions in its invaded range (Gallien *et al.*, 2010). Both of these assumptions may be false (Atwater, Ervine & Barney, 2018), and if they are, my efforts to characterize *V. rossicum*'s suitable environmental space as a static characteristic are likely to underestimate its true breadth (e.g., Razgour *et al.* 2019). However, while an extensive genetic survey has not been completed, limited data from New York state suggest there is little inter-population genetic variation in *V. rossicum*, suggesting local adaptation thus far is unlikely (Douglass, 2008). The environmental stability over the past 50 years that I documented further suggests little adaptation has occurred over this period. However, the genetics of *V. rossicum* populations across its invasive North American range should be investigated in future studies to confirm these assumptions.

SDM performance

My results show that increases in SDM performance were consistent with increases in stability in environmental space. Václavík and Meentemeyer (2012) found similar results for an invasive pathogen (*Phytophthora ramorum*). Models constructed with data from the early stages of invasion were less accurate than models constructed with data from later stages, which were closer to equilibrium. In their study, the equilibrium distribution was inferred from comparisons of model performance for a simulated 'pseudo-species' with known fundamental niche (Václavík & Meentemeyer, 2012). When the performance of models constructed using empirical data was similar to that for the simulated pseudo-species, they concluded the biological species was at (or close to) equilibrium.

However, model performance can improve over time even if the species is still relatively far from equilibrium. For example, although my model performance increased from 1920–1950 and the Boyce index for 1950 indicated high model performance, the species was still about 20 years from reaching stability in environmental space and suitable environments were still being overpredicted in the southern US and underpredicted in Canada. I may have found increases in model performance over time given narrower prediction windows between the historic (calibration) and current (evaluation) time periods and/or increased number of occurrences. Indeed, models with high current predictive accuracy can have poor future predictive accuracy (Jones, 2012). Therefore, it is important to note that increases in model performance do not necessarily indicate that the species is close to equilibrium. A more direct method to evaluate distance from equilibrium is needed.

SDMs compared to spatial dispersal models

While the spatial dispersal model outperformed the Maxent model in the earliest time period, its performance decreased over time. These results are unexpected, considering the simple spatial dispersal model used in this study has been shown to outperform Maxent models in a previous study based on a species with an invasion history of a similar length (Rodríguez-Rey *et al.*, 2013). A number of factors may have undermined the performance of my spatial dispersal models. First, the spatial dispersal model in this study was strongly influenced by outliers (e.g., the occurrence record in southern Illinois, represented by a single observation in 1910, Fig. 7) because density of occurrences was not considered. Such historical records, though verified, may represent short-lived populations in environments that ultimately proved to be unsuitable. Second, this test may have been weakened by the study species, extent, and landscape connectivity investigated in this study. An invasive plant species, such as *V. rossicum*, is likely to be influenced by long-range human-mediated dispersal. Human-mediated dispersal can lead to a potentially disconnected range that would not be represented by the spatially-autocorrelated dispersal used in my model. Finally, my simple dispersal model failed to take into account important landscape features that could limit dispersal, such as the Great Lakes. My results suggest that Maxent models can outperform a simple spatial dispersal model, but future studies addressing this issue with more sophisticated spatial models will be needed to validate these conclusions.

There are several methods that could be used to improve spatial models used in future studies. If raster cells that represented the origins of spread were limited to cells in which there were two occurrence records, the most transient populations (i.e., singleton observations in Illinois (1910) and Minnesota (1967) would not have such a strong effect on model performance. Alternatively, kernel density estimates that incorporate density of occurrences in more sophisticated spatial models could be used in place of simple distance models. Like Maxent, this approach down-weights outliers (Worton, 1989), while still maintaining the simplicity of a purely spatial model.

Anthropogenic Influences

In all of the Maxent models, human population density was the strongest predictor of *V. rossicum*'s geographical distribution. Associations with human factors are often found for invasive plant species, which are not only often introduced by humans, but can be spread to new areas within invaded regions via human-mediated dispersal (Horvitz *et al.*, 2017). *Vincetoxicum rossicum* is a strong candidate for this type of dispersal because it has light, tufted seeds which could get attached to clothing or tires or be swept up by air flow of vehicles on highways and may be carried long distances (Von Der Lippe & Kowarik, 2007; von der Lippe *et al.*, 2013). Invasive plant species are also often resilient to various anthropogenic disturbances (e.g., road disturbance, mowing, herbicide treatments), which can make them stronger competitors than native plant species in disturbed areas (Hager, 2004; Averill *et al.*, 2010; Jauni, Gripenberg & Ramula, 2015) and can lead to them colonizing environments that are less suitable to native plant species (Hager, 2004). *Vincetoxicum rossicum* demonstrates a high tolerance to anthropogenic disturbance (Averill *et al.*, 2010), as it is frequently seen along roadsides (personal observation), railroads, and near gardens (DiTommaso *et al.*, 2005), grows back easily from many mowing regimes (McKague & Cappuccino, 2005b; Averill *et al.*, 2010), and has been shown to survive in piles of gravel or concrete (personal observation).

Identifying human population density as the most important factor limiting the large-scale distribution of *V. rossicum* in northeastern North America allows for directed investigations of preventative measures to reduce the risk of spread to at-risk areas. Management efforts geared towards changing human behaviours may prove particularly effective. For example, raising awareness about *V. rossicum* and its human-mediated methods of dispersal, as well as making sure gear, machinery and people are clean of invasive species in at-risk parks or agricultural areas, may aid in limiting its future spread. It is evident that current measures and awareness are not strict enough since human-mediated long-distance jumps appear to be responsible for several new records in *V. rossicum's* current distribution. Whether these records will represent points of future spread remains to be determined.

Conclusion

When modelling distributions of invasive species, previous studies have highlighted the importance of investigating environmental equilibrium, comparing SDM performance to that of simple spatial dispersal models, and demonstrating the predictive ability of SDMs (Václavík & Meentemeyer, 2012; Gallien *et al.*, 2012; Rodríguez-Rey *et al.*, 2013; Barbet-Massin *et al.*, 2018). Despite the increased awareness of the value of these comparisons, it is still not standard practice to interpret model output and performance by conducting even one of these comparisons, let alone all of them. *Vincetoxicum rossicum* provides an interesting case study for invasive species where it was possible to assess equilibrium dynamics and changes in performance of SDMs, and spatial dispersal models over time. Its long history in its invaded range allowed me to see trends in distribution in environmental and geographic space and model performance that would have otherwise not been possible. The methods provided in this paper could be applied to other invasive species to assess whether the species is approaching equilibrium with its environment for the time period modelled. This could also be a useful validation for native and naturalized species, since nature is often in a state of disequilibrium (Svenning & Skov, 2004; Araújo & Pearson, 2005).

References

Andow D.A., Kareiva P.M., Levin S.A. & Okubo A. (1990) Spread of invading organisms. *Landscape Ecology* **4**, 177–188.

Araújo M.B. & Pearson R.G. (2005) Equilibrium of species' distributions with climate. *Ecography* **28**, 693–695.

Atwater D.Z., Ervine C. & Barney J.N. (2018) Climatic niche shifts are common in introduced plants. *Nature Ecology and Evolution* **2**, 34–43.

Averill K.M., DiTommaso A., Mohler C.L. & Milbrath L.R. (2010) Establishment of the invasive perennial Vincetoxicum rossicum across a disturbance gradient in New York State, USA. *Plant Ecology* **211**, 65–77.

Averill K.M., Mortensen D.A., Smithwick E.A.H. & Post E. (2016) Deer feeding selectivity for invasive plants. *Biological Invasions* **18**, 1247–1263.

Bahn V. & Mcgill B.J. (2007) Can niche-based distribution models outperform spatial interpolation? *Global Ecology and Biogeography* **16**, 733–742.

Barbet-Massin M., Rome Q., Villemant C. & Courchamp F. (2018) Can species distribution models really predict the expansion of invasive species? *PLoS ONE* **13**.

Bongard C.L., Navaranjan G., Yan W. & Fulthorpe R.R. (2013) Fungal colonization of the invasive vine Vincetoxicum rossicum and native plants. *Plant Ecology and Evolution* **146**, 45–52.

Bradley B.A., Early R. & Sorte C.J.B. (2015) Space to invade? Comparative range infilling and potential range of invasive and native plants. *Global Ecology and Biogeography* **24**, 348–359.

Broennimann O., Di Cola V. & Guisan A. (2020) ecospat: Spatial Ecology Miscellaneous Methods. R package version 3.1.

Bugiel L.N., Livingstone S.W., Isaac M.E., Fulthorpe R.R. & Martin A.R. (2018) Impacts of invasive plant species on soil biodiversity: A case study of dog-strangling vine (vincetoxicum rossicum) in a canadian national park. *Canadian Journal of Soil Science* **98**, 716–723.

Cappuccino N. (2004) Allee effect in an invasive alien plant, pale swallow-wort Vincetoxicum rossicum (Asclepiadaceae). *Oikos* **106**, 3–8.

Carlson A.M. & Gorchov D.L. (2004) Effects of herbicide on the invasive biennial Alliaria petiolata (garlic mustard) and initial responses of native plants in a southwestern Ohio forest. *Restoration Ecology* **12**, 559–567.

Casagrande R., Weed A., Hazlehurst A., Tewksbury L., Gassmann A. & Bourchier R. (2011) *A petition for experimental open-field release of Hypena opulenta a potential biological control agent of swallow-worts (Vincetoxicum nigrumand V. rossicum) in North America.*

Clewley G.D., Eschen R., Shaw R.H. & Wright D.J. (2012) The effectiveness of classical biological control of invasive plants. *Journal of Applied Ecology* **49**, 1287–1295.

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

Colautti R.I., Bailey S.A., van Overdijk C.D.A., Amundsen K. & MacIsaac H.J. (2006) Characterised and Projected Costs of Nonindigenous Species in Canada. *Biological*

Invasions **8**, 45–59.

Dainese M., Kühn I. & Bragazza L. (2014) Alien plant species distribution in the European Alps: Influence of species' climatic requirements. *Biological Invasions* **16**, 815–831.

Dangremond E.M., Pardini E.A. & Knight T.M. (2010) Apparent competition with an invasive plant hastens the extinction of an endangered lupine. *Ecology* **91**, 2261–2271.

Day N.J., Antunes P.M. & Dunfield K.E. (2015) Changes in arbuscular mycorrhizal fungal communities during invasion by an exotic invasive plant. *Acta Oecologica* **67**, 66–74.

Dickinson R.B., Bourchier R.S., Fulthorpe R.R., Shen S.Y., Jones I.M. & Smith S.M. (2021) Fungal endophytes increase biomass production in pale swallow-wort (Vincetoxicum rossicum (Kleopow) Barbar.). *Botany*, In press.

DiTommaso A., Lawlor F.M. & Darbyshire S.J. (2005) The Biology of Invasive Alien Plants in Canada. 2. Cynanchum rossicum (Kleopow) Borhidi [= Vincetoxicum rossicum (Kleopow) Barbar.] and Cynanchum louiseae (L.) Kartesz & Gandhi [= Vincetoxicum nigrum (L.) Moench]. *J. Plant Sci*.

DiTommaso A., Stokes C.A., Cordeau S., Milbrath L.R. & Whitlow T.H. (2018) Seed-Dispersal Ability of the Invasive Perennial Vines Vincetoxicum nigrum and Vincetoxicum rossicum. *Invasive Plant Science and Management* **11**, 10–19.

Douglass C.H. (2008) *The role of allelopathy, morphology, and genetic diversity in the invasion of swallow-wort species in New York state*.

Drake J.M. & Bossenbroek J.M. (2004) The potential distribution of zebra mussels in the United States. *BioScience* **54**, 931–941.

Early R. & Sax D.F. (2014) Climatic niche shifts between species' native and naturalized ranges raise concern for ecological forecasts during invasions and climate change. *Global Ecology and Biogeography* **23**, 1356–1365.

Eckert S., Hamad A., Kilawe C.J., Linders T.E.W., Ng W.T., Mbaabu P.R., *et al.* (2020) Niche change analysis as a tool to inform management of two invasive species in Eastern Africa. *Ecosphere* **11**, e02987.

Elith J., Graham C.H., Anderson R.P., Dudik M., Ferrier S., Guisan A., *et al.* (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography* **29**, 129–151.

Elith J., Kearney M. & Phillips S. (2010) The art of modelling range-shifting species. *Methods in Ecology and Evolution* **1**, 330–342.

Elith J., Phillips S.J., Hastie T., Dudík M., Chee Y.E. & Yates C.J. (2011) A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions* **17**, 43–57.

Ernst C.M. & Cappuccino N. (2005) *The effect of an invasive alien vine, Vincetoxicum rossicum (Asclepiadaceae), on arthropod populations in Ontario old fields*.

Feng X., Park D.S., Liang Y., Pandey R. & Papeş M. (2019) Collinearity in ecological niche modeling: Confusions and challenges. *Ecology and Evolution* **9**, 10365–10376.

Gallien L., Douzet R., Pratte S., Zimmermann N.E. & Thuiller W. (2012) Invasive species distribution models - how violating the equilibrium assumption can create new insights. *Global Ecology and Biogeography* **21**, 1126–1136.

Gallien L., Münkemüller T., Albert C.H., Boulangeat I. & Thuiller W. (2010) Predicting potential distributions of invasive species: Where to go from here? *Diversity and Distributions* **16**, 331–342.

Gibbons S.M., Lekberg Y., Mummey D.L., Sangwan N., Ramsey P.W. & Gilbert J.A. (2017) Invasive Plants Rapidly Reshape Soil Properties in a Grassland Ecosystem. *mSystems* **2**.

Goldewijk K.K., Beusen A., Doelman J. & Stehfest E. (2017) Anthropogenic land use estimates for the Holocene - HYDE 3.2. *Earth System Science Data* **9**, 927–953.

Hager H.A. (2004) Competitive effect versus competitive response of invasive and native wetland plant species. *Oecologia* **139**, 140–149.

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

Hijmans R.J. (2012) Cross-validation of species distribution models: Removing spatial sorting bias and calibration with a null model. *Ecology* **93**, 679–688.

Hijmans R.J. (2020) *raster: Geographic Data Analysis and Modeling*.

Hijmans R.J., Phillips S., Leathwick J. & Elith J. (2017) dismo: Species Distribution Modeling. R package version 1.1-4.

Hirzel A.H., Le Lay G., Helfer V., Randin C. & Guisan A. (2006) Evaluating the ability of habitat suitability models to predict species presences.

Holyoak M., Caspi T. & Redosh L.W. (2020) Integrating Disturbance, Seasonality, Multi-Year Temporal Dynamics, and Dormancy Into the Dynamics and Conservation of Metacommunities. *Frontiers in Ecology and Evolution* **8**.

Horvitz N., Wang R., Wan F.H. & Nathan R. (2017) Pervasive human-mediated large-scale invasion: analysis of spread patterns and their underlying mechanisms in 17 of China's worst invasive plants. *Journal of Ecology* **105**, 85–94.

Jauni M., Gripenberg S. & Ramula S. (2015) Non-native plant species benefit from disturbance: A meta-analysis. *Oikos* **124**, 122–129.

Jenkins P.T. (2002) Paying for protection from invasive species. *Issues in Science and Technology* **19**, 67–72.

Jones C.C. (2012) Challenges in predicting the future distributions of invasive plant species.

Kearney M. & Porter W.P. (2004) *Mapping the fundamental niche: physiology, climate, and the distribution of a nocturnal lizard*.

Kricsfalusy V. V & Miller G.C. (2010) Community ecology and invasion of natural vegetation by Cynanchum rossicum (Asclepiadaceae) in the Toronto region, Canada. *Thaiszia Journal of Botany* **20**, 53–70.

Ladd D. & Cappuccino N. (2005) A field study of seed dispersal and seedling performance in the invasive exotic vine Vincetoxicum rossicum. *Canadian Journal of Botany* **83**, 1181–1188.

Lewis J.S., Farnsworth M.L., Burdett C.L., Theobald D.M., Gray M. & Miller R.S. (2017) Biotic and abiotic factors predicting the global distribution and population density of an invasive large mammal. *Scientific Reports* **7**, 1–12.

von der Lippe M., Bullock J.M., Kowarik I., Knopp T. & Wichmann M. (2013) Human-Mediated Dispersal of Seeds by the Airflow of Vehicles. *PLoS ONE* **8**.

Von Der Lippe M. & Kowarik I. (2007) Long-distance dispersal of plants by vehicles as a driver of plant invasions. *Conservation Biology* **21**, 986–996.

Low B.W., Zeng Y., Tan H.H. & Yeo D.C.J. (2020) Predictor complexity and feature selection affect Maxent model transferability: Evidence from global freshwater invasive species. *Diversity and Distributions*, 1–15.

Manzoor S.A., Griffiths G., Obiakara M.C., Esparza-Estrada C.E. & Lukac M. (2020) Evidence of ecological niche shift in Rhododendron ponticum (L.) in Britain: Hybridization as a possible cause of rapid niche expansion. *Ecology and Evolution* **10**, 2040–2050.

McCune J.L. (2016) Species distribution models predict rare species occurrences despite significant effects of landscape context. *Journal of Applied Ecology* **53**, 1871–1879.

McKague C.I. & Cappuccino N. (2005a) Response of Pale Swallow-wort, Vincetoxicum rossicum, following aboveground tissue loss: Implications for the timing of mechanical control. *Canadian Field-Naturalist* **119**, 525–531.

McKague C.I. & Cappuccino N. (2005b) Response of Pale Swallow-wort, Vincetoxicum rossicum, following aboveground tissue loss: Implications for the timing of mechanical control. *Canadian Field-Naturalist* **119**, 525–531.

McKenney D.W., Pedlar J.H., Papadopol P. & Hutchinson M.F. (2006) The development of 1901-2000 historical monthly climate models for Canada and the United States. *Agricultural and Forest Meteorology* **138**, 69–81.

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

Meyerson L.A. & Mooney H.A. (2007) Invasive alien species in an era of globalization. *Frontiers in Ecology and the Environment* **5**, 199–208.

Miller G.C. & Kricsfalusy V. (2008) *Dog-strangling Vine - Cynanchum rossicum (Kleopow) Borhidi: A review of distribution , ecology and control of this invasive exotic plant*.

Monahan W.B. (2009) A mechanistic niche model for measuring species' distributional responses to seasonal temperature gradients. *PLoS ONE* **4**, 7921.

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

Muscarella R., Galante P.J., Soley-Guardia M., Boria R.A., Kass J.M., Uriarte M., *et al.* (2014) ENMeval: An R package for conducting spatially independent evaluations and estimating optimal model complexity for ecological niche models. *Methods in Ecology and Evolution* **5**.

Naimi B. (2017) *usdm: Uncertainty analysis for species distribution models*.

Omernik J.M. & Griffith G.E. (2014) Ecoregions of the Conterminous United States: Evolution of a Hierarchical Spatial Framework. *Environmental Management* **54**, 1249–1266.

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

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

Pobedimova E.G. (1978) Family 132. Ascledpiadaceae R. Br. Pages 50-75 in A.A. Fedorov, ed. Flora partis Europaeae URSS. Volume 3. Nauka Press, Leningrad, Russia [in Russian].

Powell K.I., Chase J.M. & Knight T.M. (2011) A synthesis of plant invasion effects on biodiversity across spatial scales. *American Journal of Botany* **98**, 539–548.

Pringle J.S. (1973) The Spread of Vincetoxicum species (Asclepiadaceae) in Ontario. *The Canadian Field-Naturalist* **87**, 27–33.

Radosavljevic A. & Anderson R.P. (2014) Making better Maxent models of species distributions: Complexity, overfitting and evaluation. *Journal of Biogeography* **41**, 629–643.

Razgour O., Forester B., Taggart J.B., Bekaert M., Juste J., Ibáñez C., *et al.* (2019) Considering adaptive genetic variation in climate change vulnerability assessment reduces species range loss projections. *Proceedings of the National Academy of Sciences of the United States of America* **116**, 10418–10423.

Renner I.W. & Warton D.I. (2013) Equivalence of MAXENT and Poisson Point Process Models for Species Distribution Modeling in Ecology. *Biometrics* **69**, 274–281.

Rodríguez-Rey M., Jiménez-Valverde A. & Acevedo P. (2013) Species distribution models predict range expansion better than chance but not better than a simple dispersal model. *Ecological Modelling* **256**, 1–5.

Roura-Pascual N., Bas J.M., Thuiller W., Hui C., Krug R.M. & Brotons L. (2009) From introduction to equilibrium: Reconstructing the invasive pathways of the Argentine ant in a Mediterranean region. *Global Change Biology* **15**, 2101–2115.

Seebens H., Essl F., Dawson W., Fuentes N., Moser D., Pergl J., *et al.* (2015) Global trade will accelerate plant invasions in emerging economies under climate change. *Global Change Biology* **21**, 4128–4140.

Sheeley S.E. & Raynal D.J. (1996) The distribution and status of species of Vincetoxicum in eastern North America. *Journal of the Torrey Botanical Society* **123**, 148–156.

Srivastava V., Liang W., Keena M.A., Roe A.D., Hamelin R.C. & Griess V.C. (2020) Assessing niche shifts and conservatism by comparing the native and post-invasion niches of major forest invasive species. *Insects* **11**, 1–19.

Svenning J.C. & Skov F. (2004) Limited filling of the potential range in European tree species. *Ecology Letters* **7**, 565–573.

Taylor S. & Kumar L. (2013) Potential distribution of an invasive species under climate change scenarios using CLIMEX and soil drainage: A case study of Lantana camara L. in Queensland, Australia. *Journal of Environmental Management* **114**, 414–422.

Theoharides K.A. & Dukes J.S. (2007) Plant invasion across space and time: Factors affecting nonindigenous species success during four stages of invasion. *New Phytologist* **176**, 256–273.

Thiers B. (2021) Index Herbariorum : A global directory of public herbaria and associated staff. New York Botanical Garden’s Virtual Herbarium. Continuously updated *http://sweetgum.nybg.org.proxy.bib.uottawa.ca/ih*.

Trethowan P.D., Robertson M.P. & McConnachie A.J. (2011) Ecological niche modelling of an invasive alien plant and its potential biological control agents. *South African Journal of Botany* **77**, 137–146.

Václavík T., Kupfer J.A. & Meentemeyer R.K. (2012) Accounting for multi-scale spatial autocorrelation improves performance of invasive species distribution modelling (iSDM). *Journal of Biogeography* **39**, 42–55.

Václavík T. & Meentemeyer R.K. (2012) Equilibrium or not? Modelling potential distribution of invasive species in different stages of invasion. *Diversity and Distributions* **18**, 73–83.

Venette R.C., Gordon D.R., Juzwik J., Koch F.H., Liebhold A.M., Peterson R.K.D., *et al.* (2021) Early Intervention Strategies for Invasive Species Management: Connections Between Risk Assessment, Prevention Efforts, Eradication, and Other Rapid Responses. In: *Invasive Species in Forests and Rangelands of the United States*. pp. 111–131.

Vilà M., Espinar J.L., Hejda M., Hulme P.E., Jarošík V., Maron J.L., *et al.* (2011) Ecological impacts of invasive alien plants: A meta-analysis of their effects on species, communities and ecosystems. *Ecology Letters* **14**, 702–708.

Visulina O.D. (1952) Family CXVII. AsclepiadaceaeLindl. Pages 270-287 in M. I. Kotov and A. I. Varvarich, eds. Flora of the Ukrainian SSR. Volume 8. Academy of sciences of the Ukrainian SSR., Kyiv, Ukraine. [in Ukrainian].

Weston L.A., Barney J.N. & DiTommaso A. (2005) A review of the biology and ecology of three invasive perennials in New York State: Japanese knotweed (Polygonum cuspidatum), mugwort (Artemisia vulgaris) and pale swallow-wort (Vincetoxicum rossicum). *Plant and Soil* **277**, 53–69.

Wieder W.R., Boehnert J., Bonan G.B. & Langseth M. (2014) Regridded Harmonized World Soil Database v1.2. ORNL DAAC. *ORNL DAAC, Oak Ridge, Tennessee, USA.*

Wiens J.J. & Graham C.H. (2005) Niche conservatism: Integrating evolution, ecology, and conservation biology. *Annual Review of Ecology, Evolution, and Systematics* **36**, 519–539.

Wilson M. V. & Clark D.L. (2001) Controlling invasive Arrhenatherum elatius and promoting native prairie grasses through mowing. *Applied Vegetation Science* **4**, 129–138.

Wisz M.S., Hijmans R.J., Li J., Peterson A.T., Graham C.H., Guisan A., *et al.* (2008) Effects of sample size on the performance of species distribution models. *Diversity and Distributions* **14**, 763–773.

With K.A. (2002) The landscape ecology of invasive spread. *Conservation Biology* **16**, 1192–1203.

Worton B.J. (1989) Kernel methods for estimating the utilization distribution in home- range studies. *Ecology* **70**, 164–168.

Tables

Table 1. Full set of original environmental variables (prior to correlation analysis) used in equilibrium analysis. Climate variables are based on historical monthly temperature and precipitation from Natural Resources Canada (NRCAN). Acronyms are based on Worldclim variables and were downloaded at a resolution of 10 km^2. Edaphic variables are from the Regridded Harmonized World Soil Database (HWSD). Edaphic variables were downloaded at an original resolution 8 km^2 and were upscaled to match 10 km^2 using bilinear interpolation. Human population densities were downloaded from HYDE 3.2, and were available in 10-year intervals at a resolution 10 km^2.

Category (*source*)	**Variables**
Climate (*Natural Resources Canada*)	BIO1 = Annual Mean Temperature BIO2 = Mean Diurnal Temperature Range (Mean of monthly (max temp - min temp)) BIO3 = Isothermality (BIO2/BIO7) (×100) BIO4 = Temperature Seasonality (standard deviation ×100) BIO5 = Max Temperature of Warmest Month BIO6 = Min Temperature of Coldest Month BIO7 = Temperature Annual Range (BIO5-BIO6) BIO8 = Mean Temperature of Wettest Quarter BIO9 = Mean Temperature of Driest Quarter BIO10 = Mean Temperature of Warmest Quarter BIO11 = Mean Temperature of Coldest Quarter BIO12 = Annual Precipitation BIO13 = Precipitation of Wettest Month BIO14 = Precipitation of Driest Month BIO15 = Precipitation Seasonality (Coefficient of Variation) BIO16 = Precipitation of Wettest Quarter BIO17 = Precipitation of Driest Quarter BIO18 = Precipitation of Warmest Quarter BIO19 = Precipitation of Coldest Quarter
Edaphic (for topsoil layer – depth 30 cm; *HWSD*)	% Clay content % Silt content % Sand content pH
Anthropogenic (*HYDE 3.2*)	Human population density (inhabitants/km^2)

Table 2. Final variables included in each time period used in the Maxent SDMs after screening highly-correlated pairs via variance inflation factor analysis. Years indicate the final year up to which calibration occurrence records were collected for model and the final year of the 30-year climate normal for that model. For data sources, raster manipulation and raster resolution, refer to Table 1.

Variables used in Models			
1920	**1950**	**1980**	**2010**
Bio2	Bio2	Bio2	Bio2
Bio5	Bio5	–	–
Bio8	Bio8	Bio8	Bio8
–	–	Bio9	Bio9
Bio15	Bio15	Bio15	Bio15
Bio18	Bio18	Bio18	Bio18
% Clay	% Clay	% Clay	% Clay
% Silt	% Silt	% Silt	% Silt
pH	pH	pH	pH
Human population density in 1910	Human population density in 1940	Human population density in 1970	Human population density in 2000

Table 3. List, explanation, and rationale for choices of Maxent parameters available for model tuning in 'ENMeval' package.

Parameter	Purpose	Default	My choice	Rationale	Reference
Regularization parameter (RM)	Penalizes model complexity, where lower values allow for more complex models	1	a set of RMs to run through from 0.5 to 8 in steps of 0.5.	Testing a range of regularization parameters can optimize the level of complexity on a per-model basis, taking into account sample size and predictor variable redundancy and importance	Morales *et al.*, 2017
Feature classes (FCs) and FC combinations	Determines the type of transformation that can be applied to the predictor variables	Linear-Quadratic-Hinge-Product (LQHP)	Linear (L), Linear-Quadratic (LQ), Hinge (H), Linear-Quadratic-Hinge (LQH), LQHP, and Linear-Quadratic-Product-Threshold (LQHPT).	Maxent will restrict which feature classes will be allowed, based on the number of occurrence records in the calibration dataset. Feature selection through combining different RM values and FCs can lead to less overfitting and improved parsimony compared to default settings	Muscarella *et al.*, 2014; Low *et al.*, 2020
Data partitioning	Iterative partitioning of occurrence data into calibration and evaluation datasets to build model	None	"Checkerboard2": k-fold cross validation method using two checkerboard-like grids based on aggregation factors	Bins created using Checkerboard2 provide an approximately equal sample of geographic and, likely,	Muscarella *et al.*, 2014; Radosavljevic & Anderson, 2014

			(2, 4)	environmental space across bins. Improvement to random k-fold partitioning	
Type of output	Provides a quantitative value to represent environmental suitability according to model predictions	Complementary log-log (clog-log)	clog-log	Accounts for extreme values in environmental variables, which exist in the population data, and provides an easily interpreted value between 0 and 1	Phillips *et al.*, 2017
Background localities	Allows comparison of relationships between predictor variables at background localities and those at presence localities	Random distribution of user-selected number of points	Random distribution of 10,000 points	Sufficiently captures the range of environmental conditions within the extent studied	Phillips *et al.*, 2006

Table 4. The percent contribution of environmental variables to environmental suitability for each time period according to the Maxent algorithm. Shown for each variable is their percent contribution and ranking in terms of contribution. Models designated by the final year up to which occurrence records were accumulated. Variables with over 20% contribution to model predictions are bolded.

Variables	Modeled Time period							
	1920		1950		1980		2010	
	Rank	% Contribution	Rank	% Contribution	Rank	% Contribution	Rank	% Contribution
Human population density	**1**	**34.65**	**1**	**61.62**	**1**	**47.08**	**1**	**43.37**
Precipitation Seasonality	3	18.23	**2**	**28.42**	**2**	**27.02**	**2**	**21.93**
Mean Diurnal Temperature Range	**2**	**23.00**	4	4.50	5	2.06	6	3.78
Precipitation of Warmest Quarter	6	2.82	5	0.52	**3**	**20.25**	3	13.14
Mean Temperature of Wettest Quarter	5	8.96	3	4.58	9	0.01	7	2.29
Max. Temperature of Warmest Month	4	11.99	7	0	–	–	–	–
Mean Temperature of Driest Quarter	–	–	–	–	7	0.04	4	7.72
% Silt	8	0	6	0.35	4	3.19	5	6.40
% Clay	7	0.35	8	0	6	0.34	9	0.05
pH	9	0	9	0	8	0.02	8	1.31

Figures

Figure 1. Current geographic distribution of *Vincetoxicum rossicum* in northeastern North America. Occurrence records (black circles) are from iNaturalist from 2011–2020. These records were used as the evaluation dataset in the analyses. The background extent of the study is shown in white and provincial and state boundaries are indicated in grey.

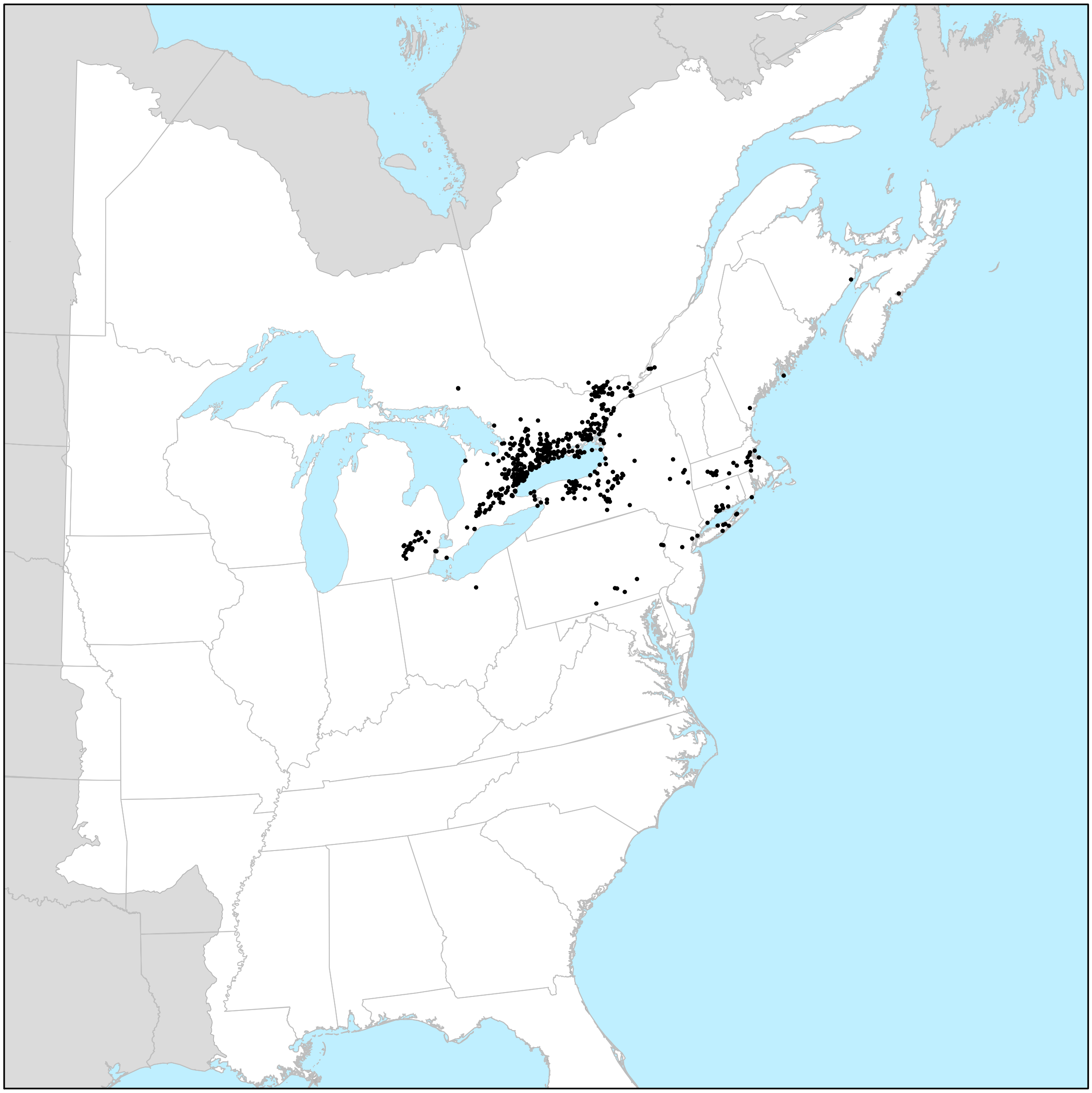

Figure 2. Overlap between historical (1900–2010) and current (2011–2020) environmental space (e-space) occupied by *Vincetoxicum rossicum* in 10-year intervals across its invasion history. E-space is defined as the first two components of a PCA (see Table 1 for variables). Green and red contour lines show the extent of historic and current e-space, respectively. Coloured cells indicate the occupied e-space across the compared time periods. Pink cells represent expansion (i.e., e-space occupied solely in the current time period) and purple cells show stability (i.e., e-space occupied in both time periods). Together they sum to 100% of the current distribution. Green cells show unfilling, e-space occupied solely in the historic time period. Density of historic occurrence records is designated with grey shading. Indices for the density-weighted proportions (%) of records representing expansion (exp), stability (stab) and unfilling (unfill) are provided in each panel and are rounded to the nearest whole number.

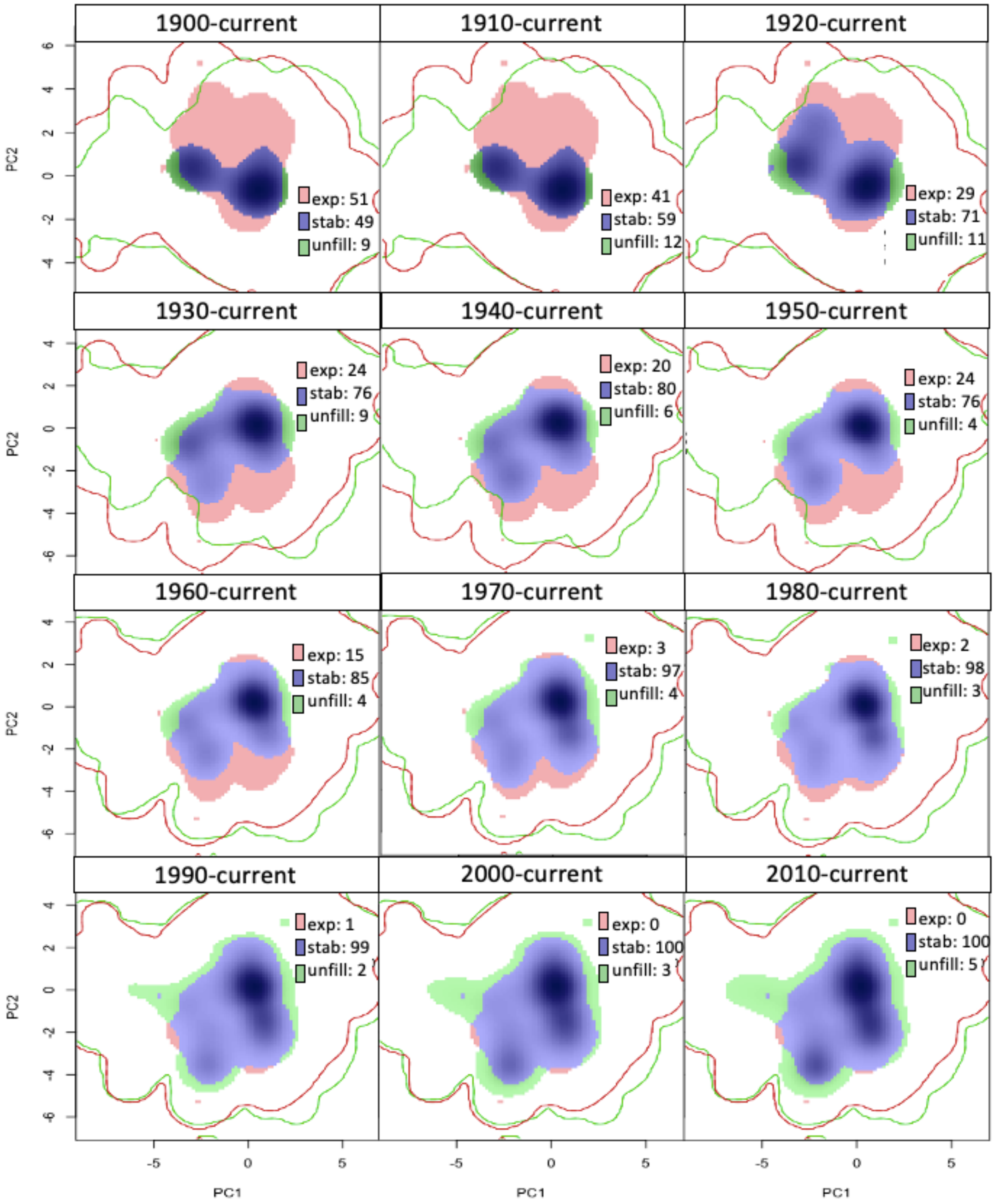

Figure 3. Overlap between historical (1900–2010) and current (2011–2020) geographic space (g-space) occupied by *Vincetoxicum rossicum* in 10-year intervals across its invasion history. G-space is defined by the longitude and latitude coordinate plane. The red contour line shows the Eastern North American coastline and the visible part of the full extent of g-space. Cell colours, shading, and index values as in Figure 2.

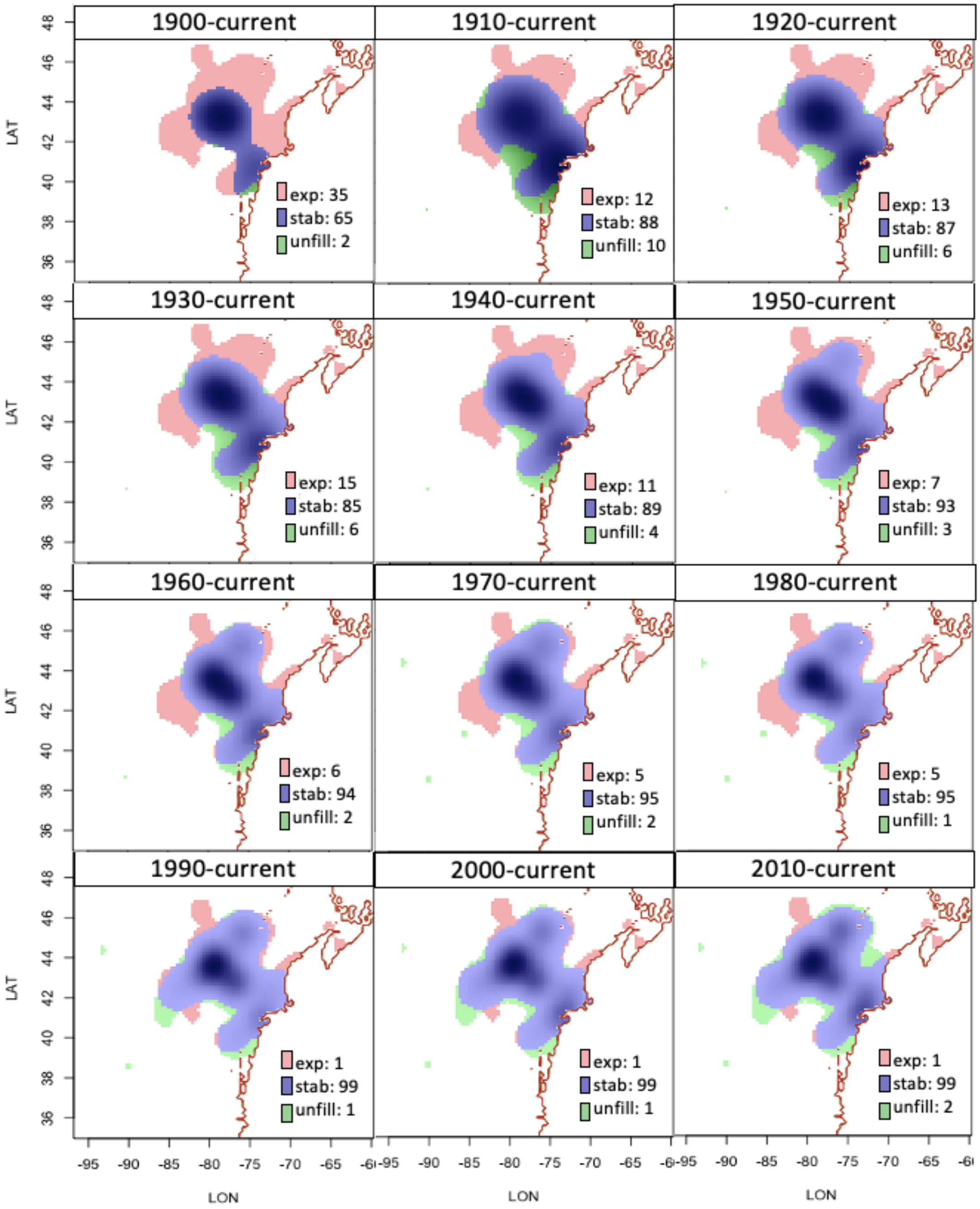

Figure 4. Density-weighted percentages (%) of *Vincetoxicum rossicum* occurrence records showing expansion, stability and unfilling in environmental space (e-space) between historical (1900–2010) and current (2011–2020) time periods.

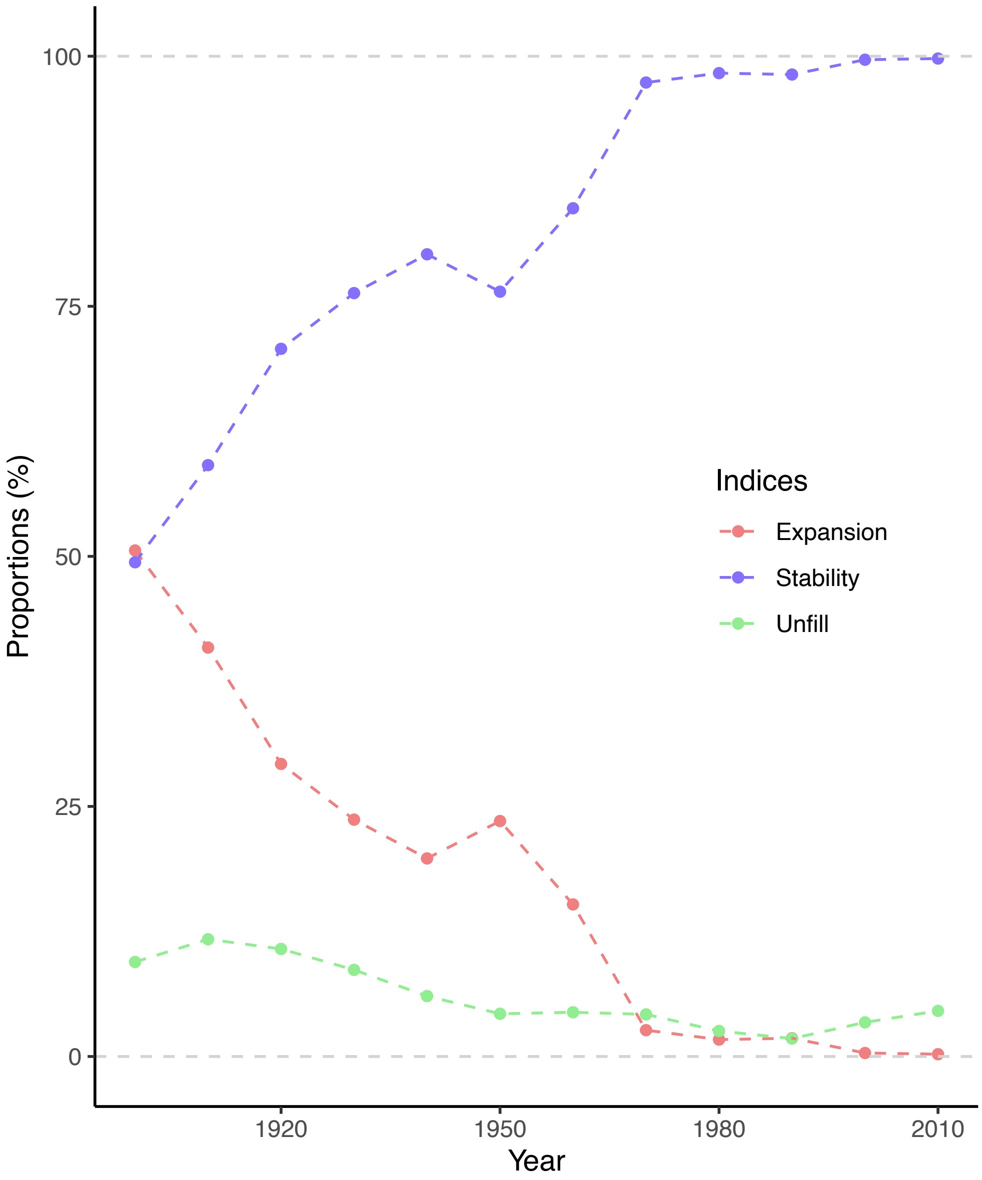

Figure 5. Density-weighted percentages (%) of *Vincetoxicum rossicum* occurrence records showing either expansion, stability or unfilling in geographic space (g-space) between historical (1900–2010) and current (2011–2020) time periods.

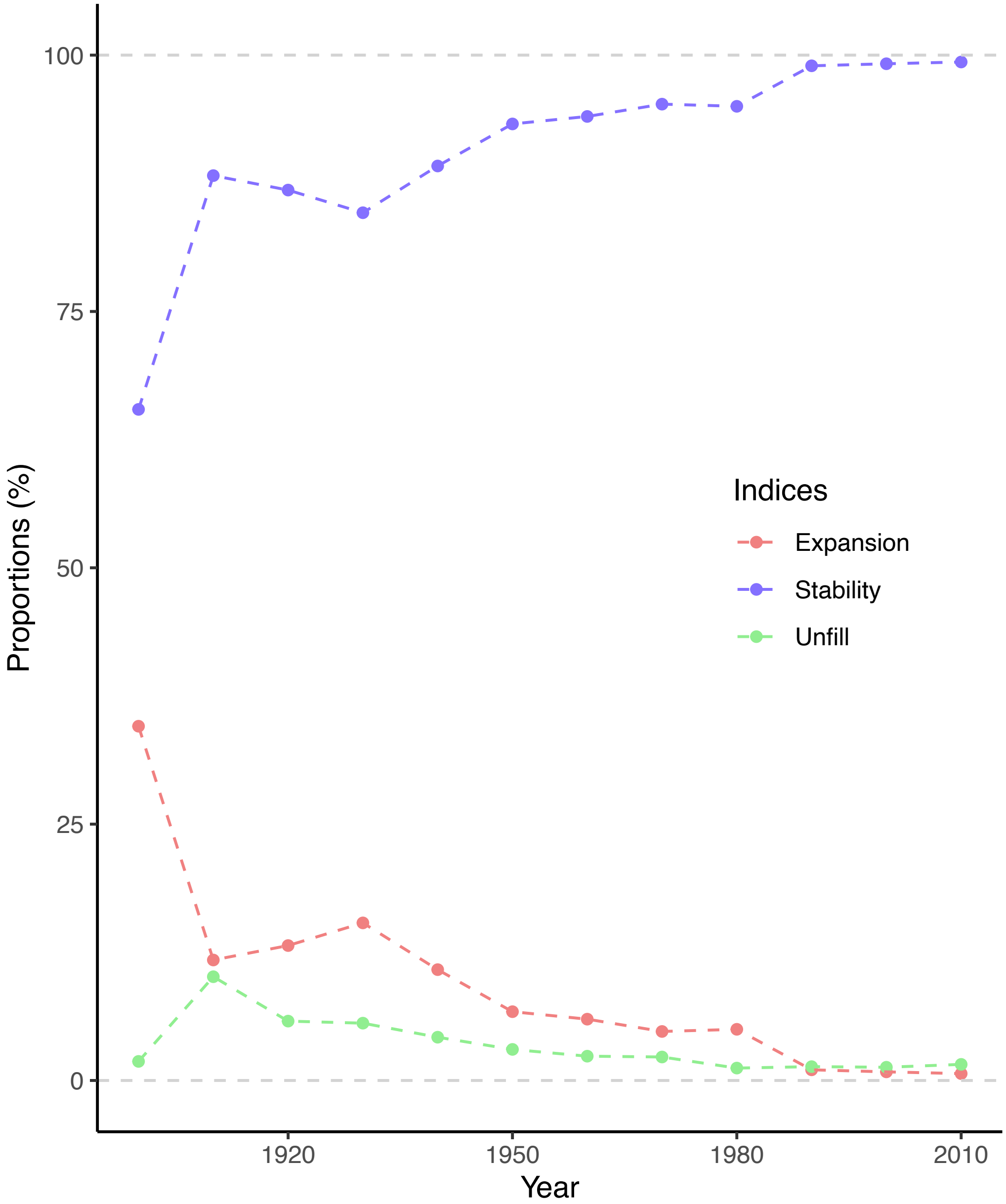

Figure 6. Evaluation of model performance over time using the Boyce index for Maxent and spatial dispersal models.

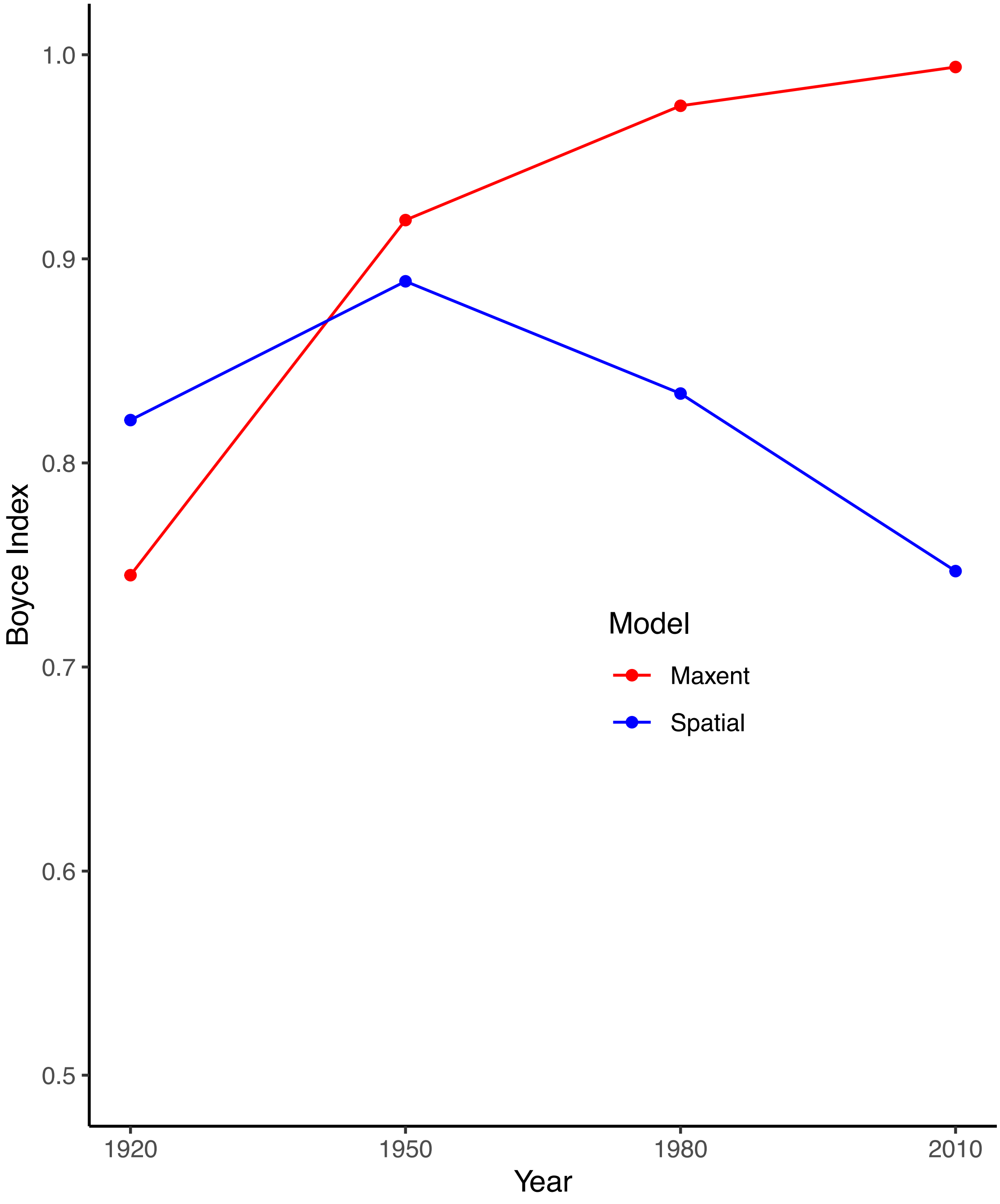

Figure 7. Predicted suitability and likelihood of occurrence maps of *Vincetoxicum rossicum* for Maxent and spatial dispersal models, respectively, of each time period with species calibration occurrence records used in each model. Panels A–D show the Maxent model predictions while panels E–H show the spatial dispersal model predictions. Calibration records for each model are cumulative across time periods and are shown as yellow open circles superimposed on the spatial dispersal models (panels E–H). The extent of all models is shown in white.