Potential distribution of invasive boxwood blight pathogen (*Calonectria pseudonaviculata*) as predicted by process-based and correlative models

Brittany Barker, Leonard Coop, Chuanxue Hong
Running Head: Potential distribution of invasive boxwood blight

**Potential distribution of invasive boxwood blight pathogen (*Calonectria pseudonaviculata*) as predicted by process-based and correlative models**

Brittany S. Barker\(^1,2\)*, Leonard Coop\(^1,2\), Chuanxue Hong\(^3\)

\(^1\)Oregon Integrated Pest Management Center, 4575 Research Way, Oregon State University, Corvallis, OR 97331 USA

\(^2\)Department of Horticulture, Oregon State University, 4017 Agriculture and Life Sciences Building, Oregon State University, Corvallis, OR 97331 USA

\(^3\)Hampton Roads Agricultural Research and Extension Center, Virginia Polytechnic Institute and State University, 1444 Diamond Springs Road, Virginia Beach, VA 23455 USA

*Corresponding author: Brittany Barker

E-mail: brittany.barker@oregonstate.edu

**Key words:** *Buxus*; plant disease; invasion; climatic suitability; CLIMEX; ensemble

**Open Research Statement:** The data, metadata, code, and derived products to reproduce the analysis and figures are currently available for reviewer access at GitHub repository (https://github.com/bbarker505/Cps-climSuit-modeling). Please note that site information for seven records from Oregon used for modeling are not provided due to confidentiality concerns. The final version of this repository will be archived at Zenodo.org or protocols.io prior to publication of this manuscript.
Abstract

Boxwood blight, caused by the ascomycete fungi *Calonectria pseudonaviculata* and *C. henricotiae*, is an emerging plant disease of boxwood (*Buxus* spp.) that has had devastating impacts on the health and productivity of boxwood in both the horticultural sector and native ecosystems. In this study, we predicted the potential distribution of *C. pseudonaviculata* at regional and global scales and explored how climatic factors shape its known range limits. Our workflow combined multiple modeling algorithms to enhance the reliability and robustness of predictions. We produced a process-based climatic suitability model in the CLIMEX program and combined outputs of six different correlative modeling algorithms to generate an ensemble correlative model. All models were fit and validated using an occurrence record dataset (N = 292 records from 24 countries) comprised of positive detections of *C. pseudonaviculata* from across its entire known invaded range. Evaluations of model performance provided validation of good model fit for all models. A consensus map of CLIMEX and ensemble correlative model predictions indicated that not-yet-invaded areas in eastern and southern Europe and in the southeastern, midwestern, and Pacific coast regions of North America are climatically suitable for establishment. Most regions of the world where *Buxus* and its congeners are native are also at risk of establishment, which suggests that *C. pseudonaviculata* should be able to significantly expand its range globally if susceptible hosts exist. Our findings provide the first insight into the global invasion threat of boxwood blight, and are valuable to stakeholders who need to know where to focus surveillance efforts for early detection and rapid response measures to prevent or slow the spread of the disease.
Introduction

Invasive plant pathogens are a global threat to the health, productivity, and diversity of plants in both agricultural and native ecosystems (Fisher et al. 2012, Lovett et al. 2016, Paini et al. 2016, Thakur et al. 2019). Plant pathogens including viruses, bacteria, oomycetes and fungi have been dispersing at unprecedented levels owing to increasing global trade and human travel, often remaining undetected or unidentified until they have spread and created visible impacts on hosts and recipient ecosystems (Fisher et al. 2012, Ricciardi et al. 2017, Thakur et al. 2019). In forest ecosystems, anthropogenic introductions of fungal and fungal-like pathogens are the main cause of emerging infectious diseases in trees, such as the well-known examples of chestnut blight and Dutch elm disease in North America (Lovett et al. 2016, Thakur et al. 2019). Ascomycete plant pathogens that can infect multiple host species in cultivated (e.g., parks, gardens, orchards, or nurseries) and native ecosystems tend to be particularly invasive, and include some of the most destructive pests of forest trees in countries with high levels of live plant trade (Santini et al. 2013, Lovett et al. 2016, Nahrung and Carnegie 2020).

Boxwood blight, also known as box blight, is an emerging disease of species in the genus *Buxus*, many of which are major evergreen shrub crops and iconic landscape plants (Batdorf 2005, Daughtrey 2019, Hong 2019b), as well as a keystone forest species (Kolganikhina 2014, Matsiakh 2016, Mitchell et al. 2018, Şimşek et al. 2019). This disease is caused by two invasive ascomycete fungi, *Calonectria pseudonaviculata* (*Cps*) (Lombard et al. 2010) and *C. henricotiae* (Gehesquière et al. 2016). Both pathogen species can infect and blight boxwood foliage, resulting in rapid plant death. *Calonectria henricotiae* is only known to occur in Europe, whereas *Cps* has a wider distribution that presently spans 24 countries primarily in Europe, Asia, and North America (Gehesquière et al. 2016, Daughtrey 2019, Castroagudín et al. 2020a, EPPO 2020). *Cps* typically disperses long distances through human-mediated transport of diseased liners (young plants) and nursery stock (Gehesquière 2014, Daughtrey 2019), often going undetected because plants can be asymptomatic until exposed to weather patterns favoring infection and subsequent symptom development (Gehesquière et al. 2013, LeBlanc et al. 2018). After the initial detection of *Cps* in the United Kingdom in 1994 (Henricot et al. 2000), the pathogen was reported from New Zealand in 1998 (Crous et al. 2002) and had spread to at least eight countries in continental Europe by 2013 (Palmer and Shishkoff 2014). *Cps* was first
detected in western Asia in 2010, and has since become widespread throughout native *Buxus* forests in the Black Sea region of Turkey and the Caucasus (Gorgiladze et al. 2011, Akili et al. 2012, Gasich et al. 2013) up to the Caspian Hyrcanian forests of northern Iran (Mirabolfathy 2013, Rezaee et al. 2013, Khazaeli et al. 2018). Initial reports of *Cps* in North America in 2011 and 2012 were from the east coast of the United States (Douglas 2012, Ivors et al. 2012) and in Oregon (Anonymous 2012) and British Columbia (Elmhirst et al. 2013), but the pathogen has now been documented in at least 30 U.S. states throughout the Southeast, Northeast, Midwest, and Pacific coast (Castroagudín et al. 2020b, Hall et al. 2021).

Boxwood blight caused by *Cps* poses a serious threat to the horticultural industry, local economies, and ecosystem integrity (LeBlanc et al. 2018, Mitchell et al. 2018, Daughtrey 2019). In the United States, the ornamental horticulture industry has sustained significant financial losses because boxwood is the number one evergreen shrub sold, with an annual wholesale value greater than $140 million (USDA National Agricultural Statistics Service 2020). Boxwood blight increases the cost of producing boxwood because infected plants are unsellable and must be destroyed, and controlling the disease with chemical treatments is expensive (LaMondia 2015, Daughtrey 2019, Hall et al. 2021). Total economic losses resulting from boxwood blight in Connecticut alone amounted to more than $3 million within the first year of detection (LaMondia 2015). Additionally, the disease has caused declines in native *Buxus* forests in western Asia, which has reduced habitat for *Buxus*-associated biodiversity and negatively impacted ecosystem services (Mirabolfathy et al. 2013, Matsuakh 2016, Lehtijärvi et al. 2017, Mitchell et al. 2018).

The full host range of *Cps* is unknown; however, none of 11 tested *Buxus* species were immune to boxwood blight (Henricot et al. 2008, Shishkoff et al. 2015, LaMondia and Shishkoff 2017), and certain Buxaceae plants in the genera *Sarcococca* Lindl. (Henricot et al. 2008, Malapi-Wight et al. 2016, Ryan et al. 2018) and *Pachysandra* Michx. (LaMondia et al. 2012, LaMondia and Li 2013, Kong et al. 2017) are also vulnerable to infection. Artificial inoculations demonstrated that the host range may even include plants in other taxonomic families (Richardson et al. 2020). These findings suggest that *Cps* could be a significant threat to at least some of the *ca.* 100 *Buxus* species, which are primarily distributed in tropical and subtropical zones of the world, and potentially to other Buxaceae and non-Buxaceae species. Despite the rapid and ongoing spread of *Cps*, assessments of establishment risk for areas which have not (yet) been invaded are not well developed. Identifying areas which are conducive for establishment by invasive plant
pathogens can guide surveillance efforts and increase the likelihood that pathogens are detected early, which is the most effective and cost-efficient method to avoid the potential ecological, economic, and societal consequences of their spread (Santini et al. 2013, Lovett et al. 2016, Hong 2019a).

In this study, we used multiple climatic suitability modeling approaches to predict the potential distribution of *Cps* at regional and global scales and explore how climatic factors shape its known range limits. Climatic suitability models, also known as ecological niche models, habitat suitability models, bioclimatic envelope models, or climatic envelope models (Elith and Graham 2009), have become an important tool for assessing establishment risk for invasive plant fungal pathogens because their growth and survival is closely related to climatic conditions, particularly temperature and moisture (Magarey et al. 2007, Lantschner et al. 2019). Indeed, the epidemiology of *Cps* is strongly influenced by longer periods of high relative air humidity combined with warm temperatures (Gehesquière 2014, Avenot et al. 2017, LeBlanc et al. 2018). We used a workflow that combined multiple modeling algorithms to reduce predictive uncertainty of single-models, which should enhance the reliability and robustness of predictions and provide independent perspectives into the potential distribution of invasive species (Capinha and Anastácio 2011, Lantschner et al. 2019). First, we used the CLIMEX program (Sutherst and Maywald 1985, Kriticos et al. 2016) to develop a climatic suitability model for *Cps* based on its predicted response to growth- and survival-limiting temperature and moisture factors. The CLIMEX approach is considered process-based because models are typically parameterized using a combination of eco-physiological data (e.g., temperature thresholds for development and survival) and point observations of occupancy or abundance from the species’ known geographical distribution (Sutherst and Maywald 1985, Kriticos et al. 2016). CLIMEX is one of the most frequently used climatic suitability modeling tools for invasive pest species, including for plant fungal pathogens (Ireland and Kriticos 2019, Lantschner et al. 2019).

Next, we developed climatic suitability models for *Cps* using multiple correlative modeling algorithms and combined their predictions into an ensemble model to potentially increase predictive performance (Marmion et al. 2009, Shabani et al. 2016, Hao et al. 2020). Correlative climatic suitability models (hereafter correlative models) involve statistically linking spatial climatic data to species location records to estimate the probability of other locations being part of the species distribution (Elith and Graham 2009, Dormann et al. 2012). Correlative
models are thought to be less reliable in predicting a species’ potential distribution in novel climates than process-based models, but their advantages include their lower input data needs and generally lower number of parameters (Kearney and Porter 2009, Dormann et al. 2012, Peterson et al. 2015). For example, correlative models only require known distribution data as an input whereas CLIMEX models requires a more extensive baseline knowledge of the species. By joining process-based and correlative approaches in a combined workflow, we strive to incorporate advantages of each approach (Lantschner et al. 2019). Our specific objectives were to identify range-limiting climatic factors for Cps using each modeling approach, and to compare the models’ predictions of climatic suitability and the overall potential distribution of the pathogen at both regional and global scales. The models developed for this study may help with identifying locations for surveillance to detect Cps before it establishes, and they may provide insight into its potential native range, which is hypothesized to be in a host center of diversity for Buxus in East Asia, the Caribbean, or Madagascar (LeBlanc et al. 2018, Daughtrey 2019).

Methods

Boxwood blight occurrence records
To fit and validate CLIMEX and correlative models, we compiled 292 occurrence records for Cps from 24 countries, which spans the entire known distribution of the pathogen (Europe, Asia, New Zealand, and North America; Appendix 1, Supporting information). Occurrence records were derived from peer-reviewed literature, theses, reports, media sources (e.g., online news articles), the Global Biodiversity Information Facility (2nd April 2021; GBIF Occurrence Download https://doi.org/10.15468/dl.44z8yr), CERIS Pest Tracker (https://pest.ceris.purdue.edu/), the Agricultural Research Service Fungal Database (https://nt.ars-grin.gov/fungaldatabases/), and personal communications. We excluded any record collected from garden centers and/or newly established plantings with boxwood plant stocks originating from another state. Ideally, positive confirmations of Cps should be based on both morphological and laboratory-collected data (e.g., genetic and physiological characterization) because some symptoms of boxwood blight overlap somewhat with those of other boxwood diseases (Daughtrey 2019). For several Cps records, confirmations were based only on morphological data, or the source did not provide any information on the confirmation process.
However, these records were within or close to areas where \( Cps \) is known to occur, which suggests the species was correctly identified. Most records for the United States (101/156 = 65\%) were spatially resolved only to the county level due to confidentiality concerns, whereas all other records were resolved to at least the city level.

**CLIMEX model**

The CLIMEX model for \( Cps \) was generated using CLIMEX version 4.0 (Kriticos et al. 2016). CLIMEX uses a re-formatted version of the CliMond dataset (Kriticos et al. 2012), which is comprised of 35 Bioclim variables for the 1961–1990 time period (https://www.climond.org/). CLIMEX data have a 10’ resolution (ca. 55 km\(^2\) at the equator) spatial resolution, which is appropriate given that most records from the United States were spatially resolved only to county level. Eco-physiological information for parameterizing a CLIMEX model for \( Cps \) was derived from published studies on the impacts of temperature and moisture on the development and survival of the vegetative growing stage as well as the more stress-tolerant microsclerotia stage, which can remain dormant on soil surfaces for months or even years (Henricot et al. 2008, Dart et al. 2015). We fine-tuned CLIMEX parameters by fitting the model to occurrence records from Europe and western Asia (\( N = 125 \)), where the species may have had more time to fill its climatic niche compared to more recently invaded regions. Only one parameter was adjusted at a time during this process. We then validated the CLIMEX model by verifying that records from North America (\( N = 159 \)) and New Zealand (\( N = 8 \)) fell within climatically suitable areas as defined by the ecoclimatic index, which ranges from 0 to 100 and describes the overall suitability of a location for long-term persistence by a species (Sutherst 2014, Kriticos et al. 2016). The ecoclimatic index integrates the annual growth index, which describes the potential for population growth (also ranging from 0 to 100), with annual stresses that limit survival during unfavorable intervals (cold, heat, dry, and wet stress) and potentially other limiting factors such as diapause. A species is considered to be excluded from locations which have an ecoclimatic index of zero, whereas increasing ecoclimatic index values signify higher potential for growth and survival (Kriticos et al. 2016). We report CLIMEX model parameters for \( Cps \) in Table 1 and describe how we derived each parameter value in the next two subsections.

**Temperature and moisture index parameters**
Four temperature index parameters in CLIMEX describe the ability for temperature-driven population growth: DV0 (limiting low temperature), DV1 (lower optimal temperature), DV2 (upper optimal temperature), and DV3 (limiting high temperature). *Cps* may develop at temperatures as low as 5 °C (Henricot and Culham 2002, Gehesquière 2014, Gehesquière et al. 2016), but we set DV0 to 9 °C to avoid potential biases resulting from canopy temperatures being lower than estimates from weather stations, which can produce errors in plant disease models (Pfender et al. 2012). We set DV1 and DV2 to 21 and 25 °C, respectively, because this temperature range is associated with optimal growing conditions in both field and laboratory settings (Henricot and Culham 2002, Gehesquière 2014, Gehesquière et al. 2016, Avenot et al. 2017, Lehtijärvi et al. 2017). We used an upper threshold of 29 °C because *Cps* colonies exhibit a low growth rate and have irregular and sclerotized morphologies at temperatures ≥ 28 °C (Gehesquiere 2014, Gehesquière et al. 2016, Avenot et al. 2017). Our unpublished re-analysis of Gehesquiere (2014) data indicated that 500 degree-hours during continuous leaf wetness would cause between ca. 10-50% infection for *B. sempervirens* and *B. s.* var ‘Suffruticosa’, which is equivalent to 20 degree-days. However, CLIMEX has no way to integrate moisture with degree-day calculations, so we used a 10× higher value of 200 as a rough stand-in for the degree-days per generation parameter (PDD). The PDD value therefore has no true meaning with regard to actual infection conditions because it accounts only for favorable temperatures.

CLIMEX describes the overall moisture characteristic of a location using estimates of soil moisture, which combine the interactions of temperature, rainfall and evapotranspiration. While precipitation and dew point are the primary moisture drivers of *Cps* growth (Shishkoff and Camp 2016, Avenot et al. 2017, LeBlanc et al. 2018), the use of soil moisture in CLIMEX should capture the species’ response to its moisture environment in a broad sense. Four soil moisture (SM) index parameters describe the influence of moisture on population growth: SM0 (limiting low moisture), SM1 (lower optimal moisture), SM2 (upper optimal moisture), and SM3 (limiting high moisture). For each SM parameter, a value of 0 indicates no soil moisture, a value of 0.5 indicates soil moisture content is 50% of capacity, a value of 1 indicates that soil moisture content is 100% of capacity, and a value > 1 indicates a water content greater than the soil holding capacity (Kriticos et al. 2016). We set SM0 to 0.2, which is higher than the permanent wilting point of plants in CLIMEX (SM0 = 0.1), because pathogens including *Cps* require free water for parts of their lifecycles. We set SM1 to 0.7 because using higher values resulted in...
certain occurrence records from more inland areas of the Black Sea and Caspian Sea regions being excluded (i.e., ecoclimatic index = 0). The upper optimal value (SM2) was set to 1.7 to ensure that wet conditions were suitable, and the upper threshold (SM3) was set to 3 to remove any constraints on growth related to very high rainfall.

Temperature and moisture stress parameters

The cold and heat stress thresholds (TTCS and TTHS, respectively) in CLIMEX define the temperature below (TTCS) or above (TTHS) which stress begins to accumulate according to a weekly rate (Kriticos et al. 2016). For example, if the average weekly maximum temperature ($T_{\text{max}}$) exceeds TTHS, then heat stress = ($T_{\text{max}}$ – TTHS) × THHS, where THHS is described by the slope of the relationship between weekly heat stress and average weekly $T_{\text{max}}$. The threshold temperature function in CLIMEX has a multiplicative factor (referred to as “week number”) that causes stress to accumulate exponentially during consecutive weeks. To help identify appropriate TTCS and THHS values, we extracted minimum temperature of the coldest week data (bio6) and maximum temperature of the warmest week (bio6) data from the CliMond dataset for Cps occurrence records from Europe and western Asia. According to this analysis, all localities in the coldest parts of Cps’s distribution, which occur in northern Europe and high-elevation parts of Georgia, were in areas where weekly minimum temperatures were $\geq -8$ °C. This finding is consistent with temperature limits of the most cold tolerant boxwood varieties, which are almost impossible to grow in areas where temperatures drop below $-10$ °C (United States Department of Agriculture 1976), and with laboratory studies of Cps microsclerotia survival (Shishkoff and Camp 2016, Yang and Hong 2018). We set TTCS to $-10$ °C and adjusted the cold stress rate (THCS) to ensure that the coldest localities for Cps in Europe and western Asia fell within areas where the ecoclimatic index exceeded zero. Additionally, we considered maps of the northern range limit for European boxwood B. sempervirens (Pojark.) in Norway, which is largely confined to districts south of 62° N (Salvesen and Kanz 2009).

We set TTHS to 32 °C and adjusted the heat stress accumulation rate (HDS) so that records from the hottest part of Cps’s distribution, which occur in northern Iran along the Caspian Sea (Mirabolfathy 2013, Khazaeli et al. 2015), had ecoclimatic index values exceeding zero. Microsclerotia have been shown to survive at 40° C for at least 24 hours (Yang and Hong 2018); however, other data sources suggest that heat stress accumulates at lower temperatures.
An upper lethal temperature of 33 °C has been suggested by Henricot and Culham (2002) based on a laboratory study of conidial growth (Henricot and Culham 2002), and by Hagan and Conner based on field reports. Additionally, microsclerotia died after two to five months at 30 °C under laboratory conditions (Shishkoff and Camp 2016), which, if translated to field conditions, would be slightly cooler in the soil under a canopy than in weather shelters. All but a single locality record for *Cps* in Europe and western Asia occurred in areas where weekly maximum temperatures fell below 32 °C, which provides further evidence that this temperature is an appropriate heat stress threshold.

Whereas extremely low soil moisture reduces survival of *Cps* (Shishkoff and Camp 2016, Avenot et al. 2017), excessive moisture is not known to be detrimental to survival of the pathogen. We set the dry stress threshold (SMDS) to 0.2 and weekly dry stress rate (HDS) to −0.001 because this contributed to the exclusion of the species (ecoclimatic index = 0) from relatively arid areas beyond the Black Sea and Caspian Sea regions, where boxwood does not occur (Hūšang 1989, Lehtijärvi et al. 2017). Conversely, we used a relatively high wet stress threshold (SMWS) of 3.0 and set the rate of wet stress accumulation (HWS) to 0.005. We did not apply the hot-dry (interaction) stress parameter in CLIMEX because preliminary analyses indicated that it did not assist in modeling the potential distribution.

**Irrigation**

To explore how supplemental irrigation may influence climatic suitability and the potential distribution of *Cps*, we ran the CLIMEX model both with and without an option to apply ‘top-up’ amounts irrigation (rainfall) of 2.5 mm day⁻¹ during the summer (Kriticos et al. 2016). Summer irrigation is regularly used in horticultural settings where boxwood is grown (United States Department of Agriculture 1976), and it can play a key role in *Cps* growth and survival by increasing the humidity to levels conducive for sporulation and infection (Gehesquière 2014, Bartíková et al. 2020b). Henceforth the model which did not include irrigation is simply referred to as the “CLIMEX model.”

**Correlative models**

We generated correlative models for *Cps* in the ENMTML R package v. 1.0.0 (de Andrade et al. 2020) in R version 4.0.5 (R Development Core Team 2021). ENMTML provides a suite of
functions to preprocess locality and environmental input data, fit models using a variety of algorithms, evaluate model performance for each algorithm, and combine model outputs to produce an ensemble model (de Andrade et al. 2020). We fit models using occurrence records from Europe, western Asia, and North America because prediction accuracy of correlative models is often higher when a larger proportion of the realized climatic niche is sampled (Beaumont et al. 2009, Taylor and Kumar 2012, Pili et al. 2020).

Two filtering steps were taken on records to reduce biased geographic sampling, which can strongly affect the predictive performance of correlative climatic suitability models that use presence-only data (Veloz 2009, Kramer-Schadt et al. 2013). First, we reduced the effects of clustered sampling by implementing the “pp.subsample” function in the spatialEco R package v. 1.3.7 (Evans 2021), which created a subsample of 80% of Cps records for both regions based on the expected spatial intensity function of the observed data. Second, we thinned records within ENMTML using the “CELLSIZE” method of the “thin_occ” function in the spThin R package v. 0.2.0 (Aiello-Lammens et al. 2015), which removes records that occur within a distance of two cells. This process resulted in 67 records for Europe and western Asia (67/163 = 41%), and 96 records for North America (96/163 = 59%). Maps presenting the full and subsampled occurrence records for both regions are presented in Fig. S1 (Supporting information).

Twenty-seven bioclimatic variables from the CliMond dataset were used to generate correlative SDMs (Kriticos et al. 2012). The first 19 bioclimatic variables (bio1-bio19) are derived from the WorldClim data set and represent annual, weekly (interpolated from monthly), and seasonal trends and extremes in temperature and precipitation (Hijmans et al. 2005). Eight bioclimatic variables that describe weekly, quarterly and annual indices of soil moisture (bio28-bio35) were also included because considering soil moisture should increase comparability of model predictions between CLIMEX and correlative models. We cropped bioclimatic layers to areas where Cps could reasonably disperse to because restricting the theoretically accessible area used for model fitting can significantly improve model performance (Cooper and Soberón 2018). This included areas between 25.5 °N and 25.5 °S in both regions, between 170 °W and 51 °W in North America (conterminous United States and southern Canada), and between 12 °W and 61.9 °E in Eurasia (western Europe to the eastern border of Iran).

A principal component analysis (PCA) was conducted based on the correlation matrix of the 27 cropped climate variables to reduce variable collinearity, which can reduce uncertainty of
correlative models and increase performance of models projections into new regions (Veloz 2009, Dormann et al. 2013, Petitpierre et al. 2017, De Marco and Nóbrega 2018). We produced a dataset comprised of six principal components (PCs) that explained at least 95% of the total variance (De Marco and Nóbrega 2018) using the “rasterPCA” function in the RSToolbox R package v. 0.2.6 (Leutner and Horning 2017). The first and second PC axes explained the highest proportion of the total variance (52.2% + 27.3% = 79.5%) and had strongest contributions from moisture and temperature variables, respectively (Table 2). The first PC axis (PC1) had a strong positive loading for soil moisture seasonality (bio31) and strong negative loadings for precipitation and soil moisture during warm seasons (bio18 and bio34, respectively), reflecting lower warm season moisture and higher moisture seasonality at positive PC1 scores (Table 2 and Fig. S2, Supporting information). The second PC axis (PC2) had strong positive loadings for temperatures during cold seasons (bio6 and bio11) and strong negative loadings for temperature seasonality and annual range (bio4 and bio7, respectively), reflecting lower winter temperatures and higher temperature seasonality at positive PC2 scores. PC axes 3 through 6 explained the remaining 16.3% of total variance and were primarily related to temperatures during warm and wet seasons (PC3), wet season precipitation (PC4), diurnal cold and wet season temperature range and soil moisture (PC5), and dry season precipitation and precipitation seasonality (PC6).

Six different algorithms were used to fit correlative models in ENMTML and assess variable importance. These included: boosted regression tree (Elith et al. 2008), generalized additive models (Guisan et al. 2002), Gaussian process usage (Golding and Purse 2016), Maxent with applied linear and quadratic features (“Maxent simple”) (Phillips et al. 2006, 2017), random forests (Prasad et al. 2006), and support vector machine (Guo et al. 2005). ENMTML sources modeling algorithm functions from multiple different R packages and uses default settings (Table S1, Supporting information) unless the user manually edits the program. We applied default settings for all algorithms except for Maxent, in which we increased the regularization multiplier parameter from one (default) to four to avoid model overfitting (Phillips et al. 2006). We used the three-step pseudo-absence selection method of Senay et al. (2013) to allocate pseudo-absences, wherein a sample of environmentally dissimilar locations within a 400 km buffer around occurrence records were identified and then sampled using k-means clustering. Information on the individual R package repositories and settings used for each modeling algorithm is presented in Table S1, Supporting information.
In the post-processing stage of ENMTML, we evaluated model performance for each algorithm and produced an ensemble model from single-model outputs. Model performance was assessed using 50 bootstrapped replicates for each algorithm, with a random 70% subset of records used to train the model and 30% reserved for validation. Model replicates were then projected at a global scale using the same climatic PC predictors. We evaluated SDMs using the area under the receiving operating characteristic curve (AUC), true skill statistics (TSS), Kappa, Jaccard, Sørensen, Boyce, and F-measure on presence-background data (F_{pb}) metrics (Boyce et al. 2002, Allouche et al. 2006, Li and Guo 2013, Leroy et al. 2018). Similarity indices from community ecology (Jaccard, Sørensen and F_{pb}) may provide better estimations of model discrimination capacity than metrics which depend on prevalence (the proportion of sites where the species is present) including AUC and Kappa (Allouche et al. 2006, Li and Guo 2013, Leroy et al. 2018).

An ensemble model was produced by calculating a weighted mean of suitability predictions (probability of occurrence) of the best models across all algorithms, defined as those which had an F_{pb} metric exceeding the average for all models (Thuiller 2004). Additionally, we overlaid predictions of presence-absence produced by the six algorithms to compare delimitations of the potential distribution. The maximum TSS threshold, which maximizes the sum of sensitivity (proportion of correctly predicted observations of species presence) and specificity (proportion of correctly predicted observations of species absence), was used to produce presence-absence predictions because it may have higher accuracy than other threshold methods (Liu et al. 2005, França and Cabral 2019). We tested the climatic similarity between the model calibration and global projection areas using a mobility-oriented parity (MOP) analysis (Owens et al. 2013) to identify potential regions where strict extrapolation occurred, wherein climatic conditions are outside of the range of conditions in the calibration area. Model extrapolation into new regions or climate change scenarios may change the correlation structure between parameters and thus lead to unreliable predictions when projected outside the environmental space (Dormann et al. 2013, Owens et al. 2013, Petitpierre et al. 2017). The MOP analysis sampled 10% of reference points from the environmental space of the calibration area and was conducted within ENMTML using the “MOP” function in kuenm R package v. 1.1.7 (Cobos et al. 2019). Finally, we produced consensus maps depicting areas of overlap in the potential distribution as estimated by presence-absence predictions of the ensemble correlative
model and the CLIMEX model. In theory, an ecoclimatic index which exceeds 0 indicates a potential for establishment (Kriticos et al. 2016); however, we defined the potential distribution in CLIMEX as areas which had an ecoclimatic index of at least 10 because most occurrence records (286/292 = 98%) met this criterion.

Results

Model evaluation and variable importance in correlative models
Validation analyses indicated very good performance of the CLIMEX model and correlative models. Of the 124 occurrence records from Europe and western Asia used for fine-tuning CLIMEX parameters, only one fell within an unsuitable location (i.e., ecoclimatic index = 0). All occurrence records from North America and New Zealand were in areas that CLIMEX predicted to be climatically suitable (average ecoclimatic index = 24, range = 10 to 60), which provided validation of good model fit. Evaluation metrics for the final ensemble correlative model were very high (Table 3): the AUC, TSS, Kappa, Jaccard, Sørensen, and Boyce metrics exceeded 0.99 (values > 0.90 are considered excellent performance) and $F_{pb}$ was 1.99 ($F_{pb} = 2 \times \text{Jaccard}$). Metric values for single models across 50 repetitions were also high, with an average of 0.998 for AUC (range = 0.996–1), 0.981 for Kappa (range = 0.974–0.992), 0.981 for TSS (range = 0.975–0.992), 0.981 for Jaccard (range = 0.975–0.992), 0.991 for Sørensen (range = 0.987–0.996), and 1.96 for $F_{pb}$ (range = 1.95–1.98). The PC2 variable contributed most strongly (average = 47%) to correlative models (Table 4), indicating an important role for cold temperatures and temperature seasonality in shaping the distribution of Cps. The PC1 variable provided the next highest contribution to correlative models (average = 24.3%), indicating that warm season moisture and moisture seasonality (PC1) are also important range-limiting factors. On average, the remaining PC variables had relatively low contributions (PC3 = 8.5%, PC4 = 4.8%, PC5 = 9.8%, PC6 = 5.6%).

Climatic suitability for and potential distribution of Cps in Europe and western Asia
CLIMEX and the ensemble correlative model predictions of climatic suitability and the potential distribution for Cps in Europe and western Asia were mostly concordant (Fig. 1). CLIMEX predicted the highest ecoclimatic and population growth index values in the Atlantic region of
western Europe, coastal areas of southern Europe, and the Black and Caspian Sea regions of western Asia (Figs 1A and 2A, respectively), which is consistent with the ensemble correlative model’s predictions of high climatic suitability for these areas (Fig. 1C). Consequently, these regions were included in the potential distribution according to both models (Fig. 1D). In general, the models predicted lower climatic suitability throughout most of central and eastern Europe, but these areas were nonetheless included in the potential distribution. CLIMEX’s estimate of the potential distribution extended farther east than that of the ensemble correlative model to include all of the Baltic states and Belarus, a greater area of Ukraine, and the border region of Russia. However, predictions of suitability and the potential distribution in these regions varied among the six correlative modeling algorithms, which indicates model uncertainty (Figs S3 and S4, Supporting information). For example, the Gaussian process and support vector machine algorithms predicted a larger extent of climatically suitable area in eastern Europe than other algorithms.

Temperature and aridity were both important range-limiting factors for Cps in Europe and western Asia. According to CLIMEX, cold stress is predicted to constrain Cps to latitudes below ca. 60° N in Europe, and it would exclude the species from western Russia except for the southernmost regions (Fig. 2B). Conversely, a combination of heat and dry stress in Iran and countries on the eastern edge of the Caspian Sea (e.g., Turkmenistan, Kazakhstan) is predicted to limit the species to predominantly southwestern areas of the Caspian Sea region (Figs 2C, D). Heat and dry stress are also predicted to exclude Cps from most of southern Spain and surrounding non-coastal areas of northwestern Africa. Employing the irrigation option in CLIMEX resulted in increases in climatic suitability along the species’ predicted range throughout Europe and western Asia (Fig. 1B), which subsequently resulted in an expansion in the potential distribution in northern Europe (Sweden and Finland), southern Europe (e.g., Greece), eastern Europe (Ukraine), western Russia, Turkey, and the Caucasus (e.g., Armenia and Azerbaijan).

Climatic suitability for and potential distribution of Cps in North America

Overall, predictions of climatic suitability and the potential distribution estimated by CLIMEX and the ensemble correlative model were concordant for North America (Fig. 3). Both models predicted climatically suitable conditions throughout most of eastern United States, whereas
suitable conditions in the western United States were almost entirely limited to the Pacific coast region (Fig. 3A–C). CLIMEX predicted the highest ecoclimatic and population growth index values in the eastern United States particularly in states along the east coast and Gulf coast (Figs 3A, 4A). Several southeastern and midwestern states where Cps is not known to be established were predicted to be climatically suitable for this pathogen. These include Arkansas, Missouri, Illinois, and Indiana. In the western United States and southern British Columbia, the potential distribution included western Oregon and Washington, coastal areas of California and southern British Columbia, the Sierra Nevada Mountain range (California), and a small area of the northern Rocky Mountains in Idaho and British Columbia. Major portions of the Great Plains, Intermountain West and Southwest were excluded from the potential distribution even in the CLIMEX model that included summer irrigation (Fig. 3B, D). The northernmost parts of the potential distribution were limited to coastal areas of the Pacific (British Columbia) and Atlantic (Quebec, Novia Scotia, and New Brunswick).

The ensemble correlative model predicted a somewhat larger potential distribution for Cps in eastern North America than the CLIMEX model. Specifically, it included more inland parts of the Southeast and higher latitude parts of the Northeast and southern Canada (Fig. 3D). For example, the potential distribution extended farther west in the southeastern United States and included eastern Texas and Oklahoma, and it extended farther north in eastern North America and included all northeastern states and southern Quebec and Ontario. Predictions of climatic suitability and presence for these areas were mostly consistent across individual correlative modeling algorithms (Figs S4 and S5, Supporting information).

According to CLIMEX, cold stress was the primary range-limiting factor for Cps in North America (Fig. 4B), although arid conditions in the Intermountain West and hot temperatures in the South limited the pathogen’s distribution in those areas (Figs 4C, D). Cold stress excluded the species from high-elevation areas in the Intermountain West (most of the Rocky Mountains), from northern parts of the Northeast (northern New York and most of Vermont, New Hampshire, and Maine) and the Midwest (most of Wisconsin and all of North Dakota, South Dakota, and Minnesota), and from Canada and Alaska except for some coastal areas of the Pacific. Estimates of population growth for North America (Fig. 4A) indicate that Cps populations could grow in several areas which were excluded by cold stress, a finding which suggests that populations could at least temporarily establish during favorable seasons. For
example, population growth was high in Wisconsin, New England, and southern parts of Ontario and Quebec; however, cold stress is predicted to prevent overwintering survival throughout most of these areas. Similarly, heat stress contributed to the exclusion of *Cps* in eastern Texas despite high population growth rates. Population growth and survival were both low across the Intermountain West primarily due to insufficient moisture, and heat stress and dry stress contributed to exclusion of the pathogen from much of the Southwest including western Texas.

Employing the irrigation option in CLIMEX resulted in increases in climatic suitability throughout much of the western United States (Fig. 3B); however, increases were insufficient for the inclusion of the Intermountain West and Southwest in the potential distribution.

**Global climatic suitability for and potential distribution of Cps**

Ensemble correlative model projections at a global scale were only partially consistent with the CLIMEX model (Figs 5 and 6). Both modeling approaches predicted highly suitable conditions in New Zealand, where *Cps* has been reported on both the North and South Island (Appendix 1, Supporting information), and throughout southeastern China, Japan, southern Australia, South Africa (coastal areas), Uruguay, and parts of Brazil, Argentina, Paraguay, and southern Chile (Fig. 5). Additionally, the ensemble correlative model predicted unsuitable conditions in most of the same areas where CLIMEX predicted unsuitability due to high levels of cold stress (e.g., in northern Asia; Fig. S6, Supporting information). While concordant predictions of climatic suitability for these regions translated to broad overlap in estimates of the potential distribution (Fig. 6), CLIMEX predicted higher climatic suitability and a larger potential distribution in other regions of the world. Model predictions were particularly discordant in equatorial (tropical) regions of South Asia, Africa, and South and Central America. For example, most high elevation areas of Africa and South Asia that were included the potential distribution according to CLIMEX had low or zero climatic suitability in the ensemble correlative model, whereas lower elevation regions that were included in the potential distribution by the ensemble model such as in India and the Indochina peninsula were unsuitable in the CLIMEX model (Fig. 5) due to heat stress (Fig. S6, Supporting information). The MOP analysis of climatic PC predictors used for the ensemble correlative model revealed high levels of similarity in climate between the calibration and projection area in temperate regions of the world (MOP index ≥ 0.9), but dissimilarity was higher in equatorial regions, particularly in parts of Southeast Asia and South America.
America (Fig. 7). This finding indicates that portions of environmental space in equatorial regions may be within the range of individual variables but they represent new combinations of predictors, which suggests that predictions there may be unreliable (Zurell et al. 2012, Owens et al. 2013). Employing the irrigation option in CLIMEX resulted in only marginal increases in the global potential distribution, mostly in coastal areas of arid parts of South America (e.g., Chile and Peru), Africa (e.g., Morocco and Namibia), and southern Australia.

**Discussion**

This study used both process-based CLIMEX and correlative models to assess the risk for *Cps*, a highly invasive plant pathogen, to establish at local, regional and global scales. This assessment can help guide the development of local and regional phytosanitary protocols for preventing further spread of the pathogen, prioritizing global surveillance efforts for more effective early detection, and planning for eradication, containment and management where accidental introductions do occur. These three steps are critical to preventing accidental introductions of *Cps* to and becoming established in predicted high risk areas where it is not yet present (Daughtrey 2019, Hong 2019b). They are also crucial to preventing boxwood blight from becoming rampant in areas where this invasive pathogen is at its early stages of establishment (Henricot 2006). *Cps* has spread rapidly, as evidenced by its invasion of 24 countries across three distant regions (Europe and western Asia, New Zealand, and North America) in less than 30 years (Palmer and Shishkoff 2014, LeBlanc et al. 2018, Daughtrey 2019). Preventing its accidental introduction to and establishment in new areas and mitigating its local spread are both pivotal to safeguarding global boxwood crops, plantings, and forests (Daughtrey 2019, Hong 2019b).

All models performed very well and were mostly consistent in their predictions for the calibration area (i.e., Europe, western Asia, and North America). The process-based CLIMEX model correctly predicted climatically suitable conditions at validation localities for *Cps* in North America and New Zealand, and the six individual correlative models and ensemble model had very high evaluation metrics for the calibration area. Cold temperatures were a major range-limitation at higher latitudes and elevations, as evidenced by the absence of the species from northern areas which have high levels of cold stress in the CLIMEX model, and by the strong
contribution of the cold-temperature related PC predictor (PC2) to correlative models. Moisture
during warm seasons was also a major range limiting factor, as demonstrated by increases in
climatic suitability and the potential distribution which occurred when implementing the summer
irrigation option in CLIMEX, and by the strong contribution of the PC predictor (PC1) related to
warm season moisture and moisture seasonality to correlative models. Hot temperatures, often in
combination with arid conditions, play a range-limiting role for Cps predominantly in the
southern regions of western Asia (e.g., northern Iran) and in the United States particularly in the
Southwest.

Climatic suitability for and potential distribution of Cps in Europe, western Asia, and North
America

Some of the highest levels of climatic suitability according to the CLIMEX and ensemble
correlative model occurred in western Europe, western Asia (Black and Caspian Sea regions),
and the east coast of the United States, a finding which is consistent with the widespread
presence of Cps in these regions. Oceanic climates in these areas has probably facilitated the
pathogen’s invasion because few gaps in precipitation and high humidity over the year combined
with warm-to-hot summer temperatures creates conducive conditions for infections (Fig. 8)
(Gehesquière 2014, Daughtrey 2019). In the eastern United States, Cps is particularly prevalent
in the Mid-Atlantic and northern parts of the Southeast; however, there are relatively few reports
of the pathogen from Florida and the Deep South (southernmost states in the Southeast), despite
the inclusion of most of these regions in the potential distribution. For example, boxwood blight
has not been reported beyond two locations in the Tallahassee area of northern Florida in 2016
where contaminated stock plants were received and then eradicated in 2016 (Iriarte et al. 2016),
and to date there have been no positive reports for Texas, Louisiana, and Mississippi (Hall 2021).
According to CLIMEX, hot temperatures reduced climatic suitability throughout Florida except
for along coastlines and from the Deep South except for northern parts of some states (e.g.,
northern Alabama and Georgia), which may explain the paucity of reports from these areas.
Hagan and Conner (2013) posited that disease development on container or field stock in
Alabama would most likely occur during extended periods of wet weather in mid-fall into mid-
spring because temperatures would be more ideal for growth than during the summer.
Additionally, shade can reduce temperatures and create humid conditions that may create more
favorable conditions for infections in hot environments (Bush et al. 2016, Daughtrey 2019).

Additional data on the pathogen’s ability to survive prolonged heat, particularly in the more heat-resistant microsclerotia form (Shishkoff and Camp 2016, Miller et al. 2018, Yang and Hong 2018), could help resolve whether it may establish in parts of the Deep South which may have ideal growing conditions during cool seasons.

Many areas with Mediterranean climates including those in southern Europe and the Pacific coast region of the United States were included in the potential distribution according to both modeling approaches, but Cps has a limited presence in these regions to date. In southern Europe, Cps has been reported on B. sempervirens ‘Suffruticosa’ in nurseries or gardens from only a handful of localities in northwestern Spain (Pintos Varela et al. 2009), southern France (Saurat et al. 2012), northern Italy (Saracchi et al. 2008), and Croatia (Cech et al. 2010). The pathogen has seemingly had opportunities to invade southern Europe given its rapid expansion throughout other parts of the continent beginning in ca. 1994 (LeBlanc et al. 2018, Daughtrey 2019). Host availability is likely not an issue because boxwood is commonly grown in gardens and landscapes throughout southern Europe, and native populations of B. sempervirens and B. balearica occur in pockets in northern Africa (Morocco and Algeria), central France, the southern European peninsulas (Iberian, Italian and Balkan), certain Mediterranean Islands, and Turkey (Di Domenico et al. 2012, Caudullo et al. 2017). In the western United States, Cps has been documented only in a handful of locations in western Oregon and the San Francisco Bay area despite having a potential distribution which encompasses Mediterranean climates throughout the region, including the entire California coast, parts of the Sierra Nevada Mountain range, and areas west of the Cascade Mountains in Oregon and Washington. Long warm-to-hot dry summers and cool wet winters which characterize Mediterranean climates may hinder long-term establishment of Cps because optimal conditions for growth that transpire during warm and wet weather occur too infrequently (Fig. 8). Applying the summer irrigation option in the CLIMEX model resulted in an increase in suitability throughout southern Europe and areas with Mediterranean climates in the western United States, a finding which is consistent with observations that outbreaks in Oregon and California are often associated with summer irrigation (J. Weiland, pers. comm.) or unusually wet spring and summers (Blomquist et al. 2018). Thus, regions with Mediterranean climate will likely be at higher risk of establishment if boxwood is irrigated during periods of optimal temperatures for Cps development, or during relatively wet
years. Overhead irrigation in particular facilitates boxwood blight outbreaks because it creates higher relative humidity and exposes leaf surfaces to longer periods of wetness (Gehesquièr 2014, Bartíková et al. 2020b, 2020a).

Climatic suitability tended to be lower in regions with humid continental climates compared to those with oceanic climates, despite the inclusion of many of these areas in the potential distribution. In humid continental parts of Europe, which includes most of eastern Europe and parts of Ukraine and Russia, optimal conditions for infections (warm and wet weather) may occur too infrequently owing to long, cold winters and warm-to-hot, dry summers. The only reports of the pathogen from these regions have come from nurseries and gardens in the Czech Republic (Safránková et al. 2012, Bartíková et al. 2020a) and a single nursery in western Ukraine (Matsiakh 2016). The common element of diseased boxwood in gardens in the Czech Republic was the use of irrigation systems or partial-shade conditions, which created higher humidity and exposed leaves longer periods of wetness (Bartíková et al. 2020a). Implementing the summer irrigation option in CLIMEX increased suitability throughout central and eastern Europe and resulted in an expansion of the potential distribution in western Russia, Ukraine, Turkey and the Caucasus, which provides additional evidence that irrigation will likely increase the risk of establishment of Cps in these regions.

As with Europe, climatic suitability tended to be lower in humid continental regions of North America that were included in the potential distribution; however, cool temperatures rather than aridity likely explain this finding. According to CLIMEX, cold stress lowered climatic suitability throughout much of non-coastal New York, New England, and southeastern Canada, which is consistent with an absence of Cps from these areas and with lower suitability predicted by the ensemble correlative model. In the midwestern United States, Cps has a limited presence despite the growing number of reports of the pathogen for this region, including from Missouri (2014), Kansas (2014), Illinois (2016), Indiana (2018), Arkansas (2019), Michigan (2018), and Wisconsin (2018). Nevertheless, economic damages to the horticultural industry in the Midwest could be significant if Cps takes hold because this region is one of the top four regions in inter-regional trade of boxwood (Hall et al. 2021), which supports the need for boxwood producers and users to be vigilant in watching for infections and quickly eradicating the pathogen when it is found (Hong 2019a). Our models indicate that cold temperatures will likely prevent establishment in northern Minnesota, northern Wisconsin, most of Nebraska, North Dakota, and
South Dakota. *Cps* was found in North Dakota in 2019 on contaminated stock plants that were received from Ohio, but it has not been found in landscape settings where it could potentially be exposed to winter conditions (Charles Elhard, pers. comm.). Future outbreak reports from areas which are predicted to be too cold for establishment should be followed closely to assess the ability of *Cps* to overwinter. For example, soil or snow cover may offer protection to overwintering microsclerotia that may allow the pathogen to survive in areas which are predicted to be unsuitable by our models.

Areas of Europe, western Asia, and North America which have arid or semi-arid climates had some of the lowest levels of climatic suitability, and will therefore be at relatively low risk of establishment at least in the absence of supplemental moisture. Range expansion of *Cps* in northern Europe and Russia will likely be prevented by cold temperatures; however, aridity often combined with hot temperatures may play a large role in limiting the pathogen’s expansion at its eastern range edge (Caspian Sea region) and southern range edge (Spain, Turkey and the Caspian Sea region). In North America, cold temperatures were predicted to exclude *Cps* from most of western Canada and the Rocky Mountains region; however, aridity in the Intermountain West and Southwest played the most significant role in restricting the pathogen’s potential distribution in the western United States. Implementing the irrigation option in CLIMEX did not appreciably increase climatic suitability in the Intermountain West or Southwest, which suggests that infections there may only be possible in highly irrigated settings, and potentially in shaded areas during the hot season. With the exception of New Mexico, states in these regions have low rankings for production and total sales of boxwood (Hall et al. 2021), which could further limit the chance for *Cps* to establish there.

**Global climatic suitability for and potential distribution of *Cps***

Maps of climatic suitability and the potential distribution for *Cps* indicate that most regions of the world where *Buxus* and its congers (Didymeles, Haptanthus, Pachysandra, Sarcococca, and Styloceras) are native are at risk of establishment. Most of the Buxaceae species are tropical or subtropical, with native ranges that include western and southern Europe, southwest, southern and eastern Asia, Africa, Madagascar, northernmost South America, Central America, Mexico and the Caribbean (Köhler and Brückner 1989, Balthazar et al. 2000, Köhler 2014). The CLIMEX and ensemble correlative models included much of eastern Asia and the
Himalayas in the potential distribution, which are home to ca. 40 species of *Buxus* (Köhler and Brückner 1989), four species of *Pachysandra*, and 11 species of *Sarcococca* (Balthazar et al. 2000). According to CLIMEX, the potential distribution in the Neotropics included the Andes region, where all five species of *Styloceras* Kunth ex A. Juss. are endemic (Balthazar et al. 2000), and it overlapped with at least some of the ca. 50 species of *Buxus* native to Central America and the Caribbean, such as in Mexico, Guatemala, Cuba, Hispaniola, and Puerto Rico (Köhler and Brückner 1989, Gutiérrez 2014, Köhler 2014). For Africa, both modeling approaches included a narrow band of the South African coast which has endemic *Buxus* (Friss 1989) in the potential distribution, and CLIMEX included additional areas where *Buxus* species occur including in Madagascar (nine species) (Schatz and Lowry 2002) and in western and eastern Africa (e.g., in Ethiopia, Kenya, Tanzania, and Angola) (Friss 1989). An overall lack of comprehensive and current maps that depict the ranges of Buxaceae species hinders making detailed assessments into the extent of overlap with the potential distribution of *Cps*.

Nonetheless, our broad-scale assessment indicates the potential for the pathogen to expand its range globally.

Preventing the establishment of *Cps* in regions with native boxwood is important because the pathogen can clearly cause ecological damage to affected ecosystems. Studies of *Cps* in native stands of *B. sempervirens* subsp. *colchica* in Georgia and *B. sempervirens* subsp. *hyrcana* in the Caspian Hyrcanian forests of northern Iran revealed rapid and intensive defoliation of boxwood plants of different ages, with complete defoliation occurring in up to 90% of some populations in just one year after positive detection of boxwood blight (Mirabolfathy 2013, Matsiakh 2016). Infected plants are also vulnerable to attacks by secondary opportunistic pathogens that can lead to eventual death (Matsiakh 2016). A literature survey showed that a loss of native boxwood in Europe and the Caucasus could lead to reductions in soil stability and subsequent declines in water quality and flood protection, and to declines in *Buxus*-associated biodiversity including at least 63 potentially obligate species of lichens, fungi, chromista and invertebrates (Mitchell et al. 2018). Currently there is no effective control for boxwood blight in forests because removing infected plants or applying fungicides across large areas is infeasible (Matsiakh 2016, Patarkalashvili 2017). Early detection of *Cps* will therefore be the most economical and effective method to prevent additional invasions in areas with susceptible native species.
The invasion of *Cps* could be particularly devastating to species which are vulnerable both in terms of their conservation status and their susceptibility to infection. Many *Buxus* species are already threatened or endangered because of small and isolated distributions resulting from natural causes such as island endemism and post-glacial climate change (Di Domenico et al. 2012, Gutiérrez 2014), anthropogenic disturbances such as deforestation and over-harvesting of wood (Mitchell et al. 2018), and invasions of non-native pests such as the box tree moth *Cydalima perspectalis* (Walker, 1859) in Europe and western Asia (Matsiakh 2016, Patarkalashvili 2017, Matsiakh et al. 2018, Panahi et al. 2021). For example, most of the *Buxus* species native to tropical America are endemic to single islands in the Caribbean (Köhler and Brückner 1989), 37 of which occur in Cuba alone (Gutiérrez 2014, Köhler 2014). None of the Buxaceae species tested to date are completely immune to boxwood blight infections, although severity of disease varies widely across *Buxus* species and cultivars (Henricot et al. 2008, Shishkoff et al. 2015, LaMondia and Shishkoff 2017), and it appears to be low in pachysandra (*Pachysandra*) and sweet box species (*Sarcococca*) species (Ryan et al. 2018, Kong and Hong 2019). Susceptible species which have at least partially overlapping native ranges with the potential distribution of *Cps* include *B. sempervirens* and subspecies (southern Europe and the Black and Caspian Sea regions), *B. balearica* (Mediterranean basin), *B. bodinieri* (China), *B. glomerata* (Cuba and Hispaniola), *B. harlandii* (China to Vietnam), *B. macowanii* (South Africa), *B. riparia* (Japan), *B. wallichiana* (Himalayas from east Afghanistan to Nepal), at least three *Pachysandra* species including the endangered *P. procumbens* (eastern United States), and several *Sarcococca* species (East Asia). More studies on the susceptibility of Buxaceae species to infection are needed to better assess the risk of the pathogen establishing and causing ecological harm.

Our global climatic suitability models for *Cps* provide some of the first insights into the potential geographic origin of the pathogen, which is still unknown (Castroagudín et al. 2020a, LeBlanc et al. 2021). The CLIMEX and ensemble correlative model both included a large part of southeastern China and Japan in the potential distribution, a finding which supports the hypothesis that the pathogen may have arrived to Europe on boxwood plants from East Asia (Daughtrey 2019). A possible origin of *Cps* from China is consistent with reports that most non-European imports of *Buxus* species to Europe come from this country (EPPO 2012), and with a leading hypothesis for the likely origin of invasive box tree moth in Europe (Van der Straten and
Nonetheless, we cannot rule out the possibility that *Cps* is native to another host center of diversity for *Buxus* or other Buxaceae species such as in the Caribbean or Madagascar (Castroagudín et al. 2020a), particularly given that at least one of the modeling approaches included parts of these regions in the potential distribution.

Model uncertainty

Discordance between CLIMEX and ensemble correlative model predictions for *Cps* in Europe, western Asia, and North America primarily occurred at the predicted range edges. The potential distribution according to the ensemble correlative model extended somewhat farther north in Europe and the eastern United States, and farther east in the southeastern United States. Conversely, the potential distribution according to CLIMEX extended slightly farther east in Europe and included the border region of Russia. However, ensemble correlative model predictions for areas of discordance with CLIMEX should be interpreted with caution because predictions also varied among the six different correlative models, which indicates uncertainty both across and within modeling approaches (i.e. process-based vs. correlative models and correlative vs. correlative models). These findings are consistent with studies showing that model type is a primary source of uncertainty when predicting species distributions, and that uncertainty is often greater at range margins compared with range cores (Marmion et al. 2009, Capinha and Anastácio 2011, Vale et al. 2014, Watling et al. 2015, Shabani et al. 2016).

Global projections of the ensemble correlative model were particularly discordant with the CLIMEX model for equatorial regions. We focused more on interpreting CLIMEX model predictions for equatorial regions because the MOP analysis indicated that dissimilarity in climate for the calibration and projection area was highest for equatorial areas, which suggests that predictions there may be unreliable (Zurell et al. 2012, Owens et al. 2013, Higgins et al. 2020). Process-based models such as CLIMEX are thought to be more reliable in predicting a species’ potential distribution in novel climates than correlative models because they rely on proximate constraints limiting distributions, rather than on model extrapolations (Kearney and Porter 2009, Evans et al. 2016, Higgins et al. 2020). For example, most equatorial areas that were included in the potential distribution by the ensemble correlative model were predicted to be excluded by heat stress in the CLIMEX model, including those in central Africa, India, and mainland Southeast Asia (Indochina and Malay peninsulas). Heat stress is measured using
thresholds and rates that were calibrated using ecophysiological information and records for the
pathogen in the hottest parts of its known distribution, and its predicted role in shaping the
potential distribution of Cps seems realistic given present-day knowledge of the species. As the
occurrence record dataset for model fitting influences projections into new areas, future work
should investigate whether Cps has persisted at localities used to fit correlative models for this
study (Appendix S1, Supporting information), particularly in newly invaded areas such as those
in the southern, midwestern, and Pacific coast region of North America. Records from newly
invaded areas could potentially represent short-term establishments, such as during a year(s) with
favorable weather, and may therefore be excluded from future presence-only correlative
modeling analyses.

Future climate-based risk mapping studies for Cps which use more recent climate data
and potentially incorporate inter-annual variability into models may provide more robust
estimates of present-day risk. Our models used historical 30-year climate normals for 1961 to
1990 because the current version of CLIMEX has no native ability to import and process other
forms of gridded data, such as climate normals for a more recent time frame (e.g., 1991–2021).
Additionally, CliMond data for more recent time frames have not been developed to our
knowledge, which hinders making a comparison of correlative models based on the same set of
climate predictors for different time frames. Global temperatures and precipitation patterns have
significantly changed even over the past 30 years (USGCRP 2018, IPCC 2021), which suggests
that climatic suitability models for Cps could misrepresent establishment risk in areas where
climates have become more (or less) favorable for the pathogen’s growth and survival. For
example, higher minimum winter temperatures or decreased frequency or intensity of extreme
cold resulting from climate change may increase rates of overwintering survival for invasive
microbial pathogens (Dukes et al. 2009, Thakur et al. 2019), which raises the possibility that
establishment risk at the northernmost range edges of Cps may be higher than our models
predict. Additionally, increasing humidity, precipitation, and rising temperatures in certain
regions such as the midwestern United States (USGCRP 2018, IPCC 2021) could increase risk of
establishment, whereas aridification in regions such as southern Europe, western and central
Asia, and western North America (IPCC 2021) may reduce risk. Climate suitability models
which account for inter-annual variations may increase the accuracy of predictions for Cps under
climate change because biologically relevant climatic variation that can arise from events such as
droughts or heat waves may be obscured in aggregated climate datasets such as 30-year climate normals (Gardner et al. 2021).

Conclusions

In developing species distribution models for Cps and evaluating the role of climatic factors in shaping its known range limits, we have provided some of the first insights into the potential invasive distribution and geographic origin of the most widespread and damaging pathogens of boxwood. Understanding where the pathogen could establish is particularly important in light of evidence for intercontinental dispersal and multiple introductions of Cps in the United States, which suggests that introductions of the pathogen are common and will likely continue to occur (Castroagudín et al. 2020a, LeBlanc et al. 2021). The CLIMEX and ensemble correlative model are consistent in predicting the potential for further spread in Europe (southern and eastern Europe), and in North America (southern, midwestern, and Pacific coast region). While our models can assist with identifying areas to watch for Cps both regionally and globally, an assessment of local climates for a target area may provide greater insight into the likelihood of the establishment. For example, the pathogen’s limited presence in areas of the potential distribution in Europe and North America which have Mediterranean and humid continental climates may suggest that regions of the global potential distribution with similar climates are at lower risk of establishment. Locations which are climatically marginal for Cps, but which have extensive boxwood plantings, may be best able to exclude or eradicate boxwood blight outbreaks by implementing best practices such as using less dense plantings, limiting shade cover, and exclusively make use of underground irrigation (Bush et al. 2016, Dart et al. 2016, Daughtrey 2019). Additionally, the avoidance of highly susceptible cultivars including Buxus sempervirens 'Suffruticosa' and Buxus sempervirens (Shishkoff et al. 2015; LaMondia and Shishkoff 2017; LeBlanc et al. 2018; Kramer et al. 2020) may help reduce the risk of establishment. Surveillance of Cps in regions of the world that fall within the potential distribution of the pathogen and have native Buxaceae species will be key for early detection and rapid responses measures.

Future modeling work that uses current climate data to evaluate risk of establishment may provide better insights into potential range limits for Cps, such as in high-elevation or high-latitude areas where the pathogen may now be capable of survival due to warming temperatures in recent decades. The CLIMEX model developed for this study could be modified to predict the
potential distribution of *C. henricotiae*, a closely related but genetically distinct species that also causes boxwood blight (Ghesquierre et al. 2016, LeBlanc et al. 2021). To date *C. henricotiae* has only been found in five countries in Europe, but further range expansion of this pathogen is expected and would likely influence boxwood blight epidemiology in the landscape because its thermotolerance is greater than *Cps* (Miller et al. 2018, LeBlanc et al. 2021).

**Data availability**

The data, metadata, code, and derived products to reproduce the analysis and figures have been archived at Zenodo (https://doi.org/XX.XXXX/zenodo.XXXXXXXX).

**Funding**

This work was funded by USDA APHIS Cooperative Agreement No. 20-8130-0282-CA.

**Competing interests**

The authors have declared that no competing interests exist.

**Acknowledgements**

We extend our thanks to Thomas Brand, Manus Gammelgard, Iryna Matsiakh, Nicole Ward-Gauthier, Fulya Baysal-Gurel, Funda Oskay, and Jerry Weiland for providing geographic location information for *Cps* occurrence records.
References


climatic envelopes among invasive populations may lead to underestimations of current and
https://doi.org/10.1111/j.1472-4642.2008.00547.x

https://doi.org/https://doi.org/10.1094/PDIS-05-18-0765-PDN

https://doi.org/https://doi.org/10.1016/S0304-3800(02)00200-4

for boxwood blight in the Virginia home landscape. Virginia Cooperative Extension
Publication PPWS-85NP. Virginia Tech, VA. Online:

CABI (2021) Cydalima perspectalis (box tree moth). Invasive Species Copendium. Wallingford,

Capinha C, Anastácio P (2011) Assessing the environmental requirements of invaders using
https://doi.org/10.1111/j.1472-4642.2010.00727.x

Castroagudín VL, Weiland JE, Baysal-Gurel F, Cubeta MA, Daughtrey ML, Gauthier NW,
LaMondia J, Luster DG, Hand FP, Shishkoff N, Williams-Woodward J, Yang X, LeBlanc
N, Crouch JA (2020a) One clonal lineage of Calonectria pseudonaviculata is primarily
responsible for the boxwood blight epidemic in the United States. Phytopathology 110:

Castroagudín VL, Yang X, Daughtrey ML, Luster DG, Pscheidt JW, Weiland JE, Crouch JA
https://doi.org/10.1094/PHP-06-20-0053-DG


Cech T, Diminic D, Heungens K (2010) Cylindrocladium buxicola causes common box blight in


Hall C (2021) Observations regarding the value of boxwood sales from 2014 to 2019. Boxwood


I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change.


Kong P, Hong C (2019) Host responses and impact on the boxwood blight pathogen, *Calonectria*

https://doi.org/10.1094/PDIS-10-16-1513-PDN


https://doi.org/10.1094/PDIS-02-14-0217-RE

LaMondia JA, Li DW (2013) *Calonectria pseudonaviculata* can cause leaf spot and stem blight of *Pachysandra procumbens*. Plant Health Progress 14: Online.

https://doi.org/https://doi.org/10.1094/PHP-2013-0226-01-BR

LaMondia JA, Shishkoff N (2017) Susceptibility of boxwood accessions from the national boxwood collection to boxwood blight and potential for differences between *Calonectria pseudonaviculata* and *C. henricotiae*. HortScience 52: 873–879.

https://doi.org/10.21273/HORTSCI11756-17


https://doi.org/10.1094/PDIS-03-12-0235-PDN


https://doi.org/10.1371/journal.pone.0202403


Mirabolfathy M, Ahangaran Y, Lombard L, Crous PW (2013) Leaf blight of *Buxus sempervirens*
in northern forests of Iran caused by *Calonectria pseudonaviculata*. Plant Disease 97: 1121.

https://doi.org/10.1094/PDIS-03-13-0237-PDN


Shishkoff N, Camp MJ (2016) The effect of different temperatures and moisture levels on survival of Calonectria pseudonaviculata in boxwood leaves and twigs and as
https://doi.org/10.1094/PDIS-09-15-1098-RE
https://doi.org/10.1094/php-rs-14-0033
https://doi.org/10.1371/journal.pone.0040969
https://doi.org/10.1038/s41579-019-0236-z


Tables and Figures

**Table 1.** CLIMEX parameter values for *Calonectria pseudonaviculata.*

**Table 2.** Summary of the principal component analysis of 27 bioclimatic variables used for correlative models. Principal component (PC) axes were selected until the cumulative explanation proportion reached 95% or more of the total variation of the original matrix. Loadings of PCs for each variable are presented, as well as PC's eigenvalues, the proportion of explained variance of each PC, and accumulated proportion of explained variance. The largest loadings (positive or negative) for each component (>0.30) are indicated with bold font.

**Table 3.** Mean values of evaluation statistics for individual correlative models and the ensemble model.

**Table 4.** The percent contribution of each principal component (PC) variable to correlative models produced by six algorithms. The climatic relevance of each variable [based on which bioclimatic variables had the largest loadings (positive or negative, Table 2)] and the average and range of contributions across all algorithms is indicated.

**Figure 1.** Maps of climatic suitability and potential distribution for *Calonectria pseudonaviculata* in Europe and western Asia. Climatic suitability is estimated as the ecoclimatic index in the CLIMEX model A with and B without irrigation, and as C the probability of occurrence in the ensemble correlative model. Areas of overlap in the potential distribution (purple shading) according to both CLIMEX models (ecoclimatic index = 10–100) and the ensemble correlative model (presence predictions) are shown in comparison to areas that were included in the potential distribution by only one model [red shading = CLIMEX model (no irrigation); orange shading = CLIMEX model (including irrigation); blue shading = ensemble correlative model]. Black circles represent the approximate locations of occurrence records.

**Figure 2.** Population growth and climate stress accumulation for *Calonectria pseudonaviculata* in Europe and western Asia. Population growth in CLIMEX is measured as the A annual growth
index (annual growth index, range = 0–100). Climate stress indices (range = 0–999) include B cold stress, C heat stress, and D dry stress. Results are for the CLIMEX model which did not include irrigation.

Figure 3. Maps of climatic suitability and the potential distribution for *Calonectria pseudonaviculata* in North America. Climatic suitability is estimated as the ecoclimatic index in the CLIMEX model A with and B without irrigation, and as C the probability of occurrence in the ensemble correlative model. Areas of overlap in the potential distribution (purple shading) according to both CLIMEX models (ecoclimatic index = 10–100) and the ensemble correlative model (presence predictions) are shown in comparison to areas that were included in the potential distribution by only one model [red shading = CLIMEX model (no irrigation); orange shading = CLIMEX model (including irrigation); blue shading = ensemble correlative model]. Black circles represent the approximate locations of occurrence records.

Figure 4. Population growth and climate stress accumulation for *Calonectria pseudonaviculata* in North America. Population growth in CLIMEX is measured as the A annual growth index (annual growth index, range = 0–100). Climate stress indices (range = 0–999) include B cold stress, C heat stress, and D dry stress.

Figure 5. Climatic suitability for *Calonectria pseudonaviculata* globally. Climatic suitability is estimated as A the ecoclimatic index in the CLIMEX model (includes irrigation), and as B the probability of occurrence in the ensemble correlative model. Areas where the ecoclimatic index is zero and the probability of occurrence is less than 0.1 are shown in gray.

Figure 6. Map of the global potential distribution for *Calonectria pseudonaviculata*. Areas of overlap in the potential distribution (purple shading) according to both CLIMEX models (ecoclimatic index = 10–100) and the ensemble correlative model (presence predictions) are shown in comparison to areas that were included in the potential distribution by only one model (red shading = CLIMEX model; orange shading = CLIMEX model that included irrigation; blue shading = ensemble correlative model).
Figure 7. Mobility-oriented parity (MOP) assessment outputs for projections of the ensemble correlative model for Calonectria pseudonaviculata. Areas with MOP metric values close to 1 have highly comparable climatic conditions to the those in the model calibration area. Areas with values approaching 0 indicate higher extrapolation because one or more climatic variables have values outside the range of variable(s) in the calibration area.

Figure 8. Climate comparisons for sites which are expected to differ in favorability for boxwood blight infections. Line plots depict monthly temperature (solid lines) and precipitation (dashed lines) across eight sites in Europe (orange lines) and the United States (blue lines). Sites with a Mediterranean climate (e.g., Cannes, France; Naples, Italy; Seattle, Washington; and Portland, Oregon) are less conducive for infections than sites which have higher humidity, few gaps in precipitation, and ideal temperatures for growth throughout the year, such as those in temperate/coastal climates in western Europe (e.g., Brussels, Belgium and Bordeaux, France) and warm and humid climates in the mid-Atlantic and southeastern regions of the United States (e.g., Virginia Beach, Virginia and Atlanta, Georgia). Data source: 1981-2010 climate normals, World Meteorological Organization (https://climatedata-catalogue.wmo.int; accessed 24 Sep 2021).
Table 1. CLIMEX parameter values for *Calonectria pseudonaviculata*.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>SM0</td>
<td>Limiting low moisture</td>
<td>0.2</td>
</tr>
<tr>
<td>SM1</td>
<td>Lower optimal moisture</td>
<td>0.7</td>
</tr>
<tr>
<td>SM2</td>
<td>Upper optimal moisture</td>
<td>1.7</td>
</tr>
<tr>
<td>SM3</td>
<td>Limiting high moisture</td>
<td>3.0</td>
</tr>
<tr>
<td>DV0</td>
<td>Limiting low temperature (°C)</td>
<td>8</td>
</tr>
<tr>
<td>DV1</td>
<td>Lower optimal temperature (°C)</td>
<td>21</td>
</tr>
<tr>
<td>DV2</td>
<td>Upper optimal temperature (°C)</td>
<td>25</td>
</tr>
<tr>
<td>DV3</td>
<td>Limiting high temperature (°C)</td>
<td>29</td>
</tr>
<tr>
<td>TTCS</td>
<td>Cold stress temperature threshold (°C)</td>
<td>−10</td>
</tr>
<tr>
<td>TCCS</td>
<td>Cold stress temperature rate (week⁻¹)</td>
<td>−0.005</td>
</tr>
<tr>
<td>THHS</td>
<td>Heat stress temperature threshold (°C)</td>
<td>32</td>
</tr>
<tr>
<td>THHS</td>
<td>Heat stress temperature rate (week⁻¹)</td>
<td>0.01</td>
</tr>
<tr>
<td>SMDS</td>
<td>Dry stress threshold</td>
<td>0.2</td>
</tr>
<tr>
<td>HDS</td>
<td>Dry stress rate (week⁻¹)</td>
<td>−0.001</td>
</tr>
<tr>
<td>SMWS</td>
<td>Wet stress threshold</td>
<td>3.0</td>
</tr>
<tr>
<td>HWS</td>
<td>Wet stress rate (week⁻¹)</td>
<td>0.005</td>
</tr>
</tbody>
</table>
Table 2. Summary of the principal component analysis of 27 bioclimatic variables used for correlative models. Principal component (PC) axes were selected until the cumulative explanation proportion reached 95% or more of the total variation of the original matrix. Loadings of PCs for each variable are presented, as well as PC’s eigenvalues, the proportion of explained variance of each PC, and accumulated proportion of explained variance. The largest loadings (positive or negative) for each component (>0.30) are indicated with bold font.

<table>
<thead>
<tr>
<th>Variables and proportion of variance</th>
<th>PC1</th>
<th>PC2</th>
<th>PC3</th>
<th>PC4</th>
<th>PC5</th>
<th>PC6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annual mean temperature (bio1)</td>
<td>0.033</td>
<td>0.212</td>
<td>0.298</td>
<td>-0.001</td>
<td>-0.001</td>
<td>-0.014</td>
</tr>
<tr>
<td>Mean diurnal temperature range (bio2)</td>
<td>0.083</td>
<td>-0.164</td>
<td>0.261</td>
<td>-0.226</td>
<td>-0.369</td>
<td>-0.199</td>
</tr>
<tr>
<td>Isothermality (bio3)</td>
<td>0.012</td>
<td>0.235</td>
<td>0.082</td>
<td>-0.192</td>
<td>-0.271</td>
<td>-0.066</td>
</tr>
<tr>
<td>Temperature seasonality (bio4)</td>
<td>0.06</td>
<td>-0.484</td>
<td>0.06</td>
<td>0.006</td>
<td>0.039</td>
<td>0.023</td>
</tr>
<tr>
<td>Max temperature of warmest week (bio5)</td>
<td>0.12</td>
<td>-0.013</td>
<td>0.404</td>
<td>-0.023</td>
<td>-0.023</td>
<td>-0.068</td>
</tr>
<tr>
<td>Min temperature of coldest week (bio6)</td>
<td>0.023</td>
<td>0.365</td>
<td>0.151</td>
<td>0.027</td>
<td>0.035</td>
<td>-0.003</td>
</tr>
<tr>
<td>Temperature annual range (bio7)</td>
<td>0.095</td>
<td>-0.562</td>
<td>0.208</td>
<td>-0.065</td>
<td>-0.077</td>
<td>-0.069</td>
</tr>
<tr>
<td>Mean temperature of wettest quarter (bio8)</td>
<td>-0.317</td>
<td>-0.024</td>
<td>0.566</td>
<td>0.07</td>
<td>0.118</td>
<td>0.361</td>
</tr>
<tr>
<td>Mean temperature of driest quarter (bio9)</td>
<td>0.204</td>
<td>0.246</td>
<td>0.098</td>
<td>-0.029</td>
<td>-0.04</td>
<td>-0.184</td>
</tr>
<tr>
<td>Mean temperature of warmest quarter (bio10)</td>
<td>0.082</td>
<td>0.051</td>
<td>0.4</td>
<td>0.006</td>
<td>0.019</td>
<td>-0.017</td>
</tr>
<tr>
<td>Mean temperature of coldest quarter (bio11)</td>
<td>0.013</td>
<td>0.321</td>
<td>0.187</td>
<td>-0.004</td>
<td>-0.019</td>
<td>-0.017</td>
</tr>
<tr>
<td>Annual precipitation (bio12)</td>
<td>-0.075</td>
<td>-0.002</td>
<td>0.019</td>
<td>-0.3</td>
<td>0.054</td>
<td>-0.179</td>
</tr>
<tr>
<td>Precipitation of wettest week (bio13)</td>
<td>-0.059</td>
<td>-0.027</td>
<td>0.016</td>
<td>-0.484</td>
<td>0.073</td>
<td>0.066</td>
</tr>
<tr>
<td>Precipitation of driest week (bio14)</td>
<td>-0.142</td>
<td>0.034</td>
<td>-0.009</td>
<td>-0.054</td>
<td>-0.041</td>
<td>-0.415</td>
</tr>
<tr>
<td>Precipitation seasonality (bio15)</td>
<td>0.001</td>
<td>0.095</td>
<td>-0.169</td>
<td>-0.389</td>
<td>-0.258</td>
<td>0.574</td>
</tr>
<tr>
<td>Precipitation of wettest quarter (bio16)</td>
<td>-0.056</td>
<td>-0.025</td>
<td>0.012</td>
<td>-0.464</td>
<td>0.077</td>
<td>0.035</td>
</tr>
<tr>
<td>Precipitation of driest quarter (bio17)</td>
<td>-0.132</td>
<td>0.028</td>
<td>0.003</td>
<td>-0.075</td>
<td>-0.028</td>
<td>-0.405</td>
</tr>
<tr>
<td>Precipitation of warmest quarter (bio18)</td>
<td>-0.386</td>
<td>-0.054</td>
<td>0.141</td>
<td>-0.157</td>
<td>-0.004</td>
<td>-0.01</td>
</tr>
<tr>
<td>Precipitation of coldest quarter (bio19)</td>
<td>0.174</td>
<td>0.033</td>
<td>-0.08</td>
<td>-0.373</td>
<td>0.091</td>
<td>-0.216</td>
</tr>
<tr>
<td>Annual mean moist index (bio28)</td>
<td>-0.138</td>
<td>0.012</td>
<td>-0.011</td>
<td>-0.041</td>
<td>0.27</td>
<td>-0.041</td>
</tr>
<tr>
<td>Highest weekly moist index (bio29)</td>
<td>0.032</td>
<td>-0.005</td>
<td>0.007</td>
<td>-0.112</td>
<td>0.427</td>
<td>0.041</td>
</tr>
<tr>
<td>Lowest weekly moist index (bio30)</td>
<td>-0.317</td>
<td>0.032</td>
<td>-0.075</td>
<td>0.001</td>
<td>0.036</td>
<td>-0.082</td>
</tr>
<tr>
<td>Moist index seasonality (bio31)</td>
<td>0.497</td>
<td>0.052</td>
<td>0.013</td>
<td>-0.101</td>
<td>0.195</td>
<td>0.09</td>
</tr>
<tr>
<td>Mean moist index of wettest quarter (bio32)</td>
<td>0.026</td>
<td>-0.005</td>
<td>0.008</td>
<td>-0.093</td>
<td>0.426</td>
<td>0.024</td>
</tr>
<tr>
<td>Mean moist index of driest quarter (bio33)</td>
<td>-0.31</td>
<td>0.03</td>
<td>-0.061</td>
<td>-0.004</td>
<td>0.055</td>
<td>-0.08</td>
</tr>
<tr>
<td>Mean moist index of warmest quarter (bio34)</td>
<td>-0.345</td>
<td>0.036</td>
<td>-0.093</td>
<td>-0.054</td>
<td>0.02</td>
<td>0.003</td>
</tr>
<tr>
<td>Mean moist index of coldest quarter (bio35)</td>
<td>0.051</td>
<td>-0.003</td>
<td>0.072</td>
<td>-0.036</td>
<td>0.44</td>
<td>-0.077</td>
</tr>
</tbody>
</table>

Proportion of variance

<table>
<thead>
<tr>
<th>Proportion explained by each PC (%)</th>
<th>52.2</th>
<th>27.3</th>
<th>6.4</th>
<th>4.1</th>
<th>3.2</th>
<th>2.6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accumulated proportion explained by the PCs (%)</td>
<td>52.2</td>
<td>79.5</td>
<td>85.9</td>
<td>90</td>
<td>93.2</td>
<td>95.8</td>
</tr>
</tbody>
</table>
Table 3. Mean values of evaluation statistics for individual correlative models and the ensemble model.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>AUC</th>
<th>Kappa</th>
<th>TSS</th>
<th>Jaccard</th>
<th>Sørensen</th>
<th>F_pb</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boosted regression tree</td>
<td>0.998</td>
<td>0.974</td>
<td>0.974</td>
<td>0.975</td>
<td>0.987</td>
<td>1.949</td>
</tr>
<tr>
<td>Generalized additive models</td>
<td>0.996</td>
<td>0.986</td>
<td>0.986</td>
<td>0.986</td>
<td>0.993</td>
<td>1.972</td>
</tr>
<tr>
<td>Bayesian Gaussian process</td>
<td>0.999</td>
<td>0.982</td>
<td>0.982</td>
<td>0.982</td>
<td>0.991</td>
<td>1.965</td>
</tr>
<tr>
<td>Maxent</td>
<td>0.998</td>
<td>0.974</td>
<td>0.974</td>
<td>0.975</td>
<td>0.987</td>
<td>1.949</td>
</tr>
<tr>
<td>Random forests</td>
<td>0.998</td>
<td>0.978</td>
<td>0.978</td>
<td>0.978</td>
<td>0.989</td>
<td>1.956</td>
</tr>
<tr>
<td>Support vector machine</td>
<td>1</td>
<td>0.992</td>
<td>0.992</td>
<td>0.992</td>
<td>0.996</td>
<td>1.984</td>
</tr>
<tr>
<td>Ensemble</td>
<td>1</td>
<td>0.996</td>
<td>0.996</td>
<td>0.996</td>
<td>0.998</td>
<td>1.992</td>
</tr>
</tbody>
</table>

AUC, Area Under the ROC Curve; TSS, True Skill Statistics; F_pb, F-measure on presence-background
Table 4. The percent contribution of each principal component (PC) variable to correlative models produced by six algorithms. The climatic relevance of each variable [based on which bioclimatic variables had the largest loadings (positive or negative, Table 2)] and the average and range of contributions across all algorithms is indicated.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Climatic relevance</th>
<th>BRT</th>
<th>GAM</th>
<th>GAU</th>
<th>MXS</th>
<th>RDF</th>
<th>SVM</th>
<th>Average (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PC1</td>
<td>Warm season precipitation and soil moisture, soil moisture seasonality</td>
<td>29.4</td>
<td>23.9</td>
<td>21.4</td>
<td>22.7</td>
<td>26.9</td>
<td>21.4</td>
<td>24.3 (21.4–29.4)</td>
</tr>
<tr>
<td>PC2</td>
<td>Cold season temperatures, temperature seasonality</td>
<td>60</td>
<td>23.9</td>
<td>47.3</td>
<td>45.2</td>
<td>61.6</td>
<td>44.2</td>
<td>47 (23.9–61.6)</td>
</tr>
<tr>
<td>PC3</td>
<td>Warm and wet season temperatures</td>
<td>0.3</td>
<td>22.2</td>
<td>9.2</td>
<td>8.9</td>
<td>1.1</td>
<td>9.4</td>
<td>8.5 (0.3–22.2)</td>
</tr>
<tr>
<td>PC4</td>
<td>Wet season precipitation</td>
<td>0.7</td>
<td>10</td>
<td>5.3</td>
<td>6.3</td>
<td>0.2</td>
<td>6.4</td>
<td>4.8 (0.2–10)</td>
</tr>
<tr>
<td>PC5</td>
<td>Diurnal temperature range, wet and cold season soil moisture</td>
<td>9.3</td>
<td>10</td>
<td>9</td>
<td>9.5</td>
<td>10</td>
<td>10.9</td>
<td>9.8 (9–10.9)</td>
</tr>
<tr>
<td>PC6</td>
<td>Dry season precipitation, precipitation seasonality</td>
<td>0.4</td>
<td>10</td>
<td>7.8</td>
<td>7.4</td>
<td>0.3</td>
<td>7.7</td>
<td>5.6 (0.3–10)</td>
</tr>
</tbody>
</table>

BRT, boosted regression tree; GAM, generalized additive models; GAU, Gaussian process; MXS, Maxent “simple”; RDF, random forests; SVM, support vector machine.
**Figure 1.** Maps of climatic suitability and potential distribution for *Calonectria pseudonaviculata* in Europe and western Asia. Climatic suitability is estimated as the ecoclimatic index in the CLIMEX model A with and B without irrigation, and as C the probability of occurrence in the ensemble correlative model. Areas of overlap in the potential distribution (purple shading) according to both CLIMEX models (ecoclimatic index = 10–100) and the ensemble correlative model (presence predictions) are shown in comparison to areas that were included in the potential distribution by only one model [red shading = CLIMEX model (no irrigation); orange shading = CLIMEX model (including irrigation); blue shading = ensemble correlative model]. Black circles represent the approximate locations of occurrence records.
Figure 2. Population growth and climate stress accumulation for *Calonectria pseudonaviculata* in Europe and western Asia. Population growth in CLIMEX is measured as the annual growth index (annual growth index, range = 0–100). Climate stress indices (range = 0–999) include cold stress, heat stress, and dry stress. Results are for the CLIMEX model which did not include irrigation.
Figure 3. Maps of climatic suitability and the potential distribution for *Calonectria pseudonaviculata* in North America. Climatic suitability is estimated as the ecoclimatic index in the CLIMEX model A with and B without irrigation, and as C the probability of occurrence in the ensemble correlative model. Areas of overlap in the potential distribution (purple shading) according to both CLIMEX models (ecoclimatic index = 10–100) and the ensemble correlative model (presence predictions) are shown in comparison to areas that were included in the potential distribution by only one model [red shading = CLIMEX model (no irrigation); orange shading = CLIMEX model (including irrigation); blue shading = ensemble correlative model]. Black circles represent the approximate locations of occurrence records.
Figure 4. Population growth and climate stress accumulation for *Calonectria pseudonaviculata* in North America. Population growth in CLIMEX is measured as the **A** annual growth index (annual growth index, range = 0–100). Climate stress indices (range = 0–999) include **B** cold stress, **C** heat stress, and **D** dry stress.
Figure 5. Climatic suitability for *Calonectria pseudonaviculata* globally. Climatic suitability is estimated as A the ecoclimatic index in the CLIMEX model (includes irrigation), and as B the probability of occurrence in the ensemble correlative model. Areas where the ecoclimatic index is zero and the probability of occurrence is less than 0.1 are shown in gray.
Figure 6. Map of the global potential distribution for *Calonectria pseudonaviculata*. Areas of overlap in the potential distribution (purple shading) according to both CLIMEX models (ecoclimatic index = 10–100) and the ensemble correlative model (presence predictions) are shown in comparison to areas that were included in the potential distribution by only one model (red shading = CLIMEX model; orange shading = CLIMEX model that included irrigation; blue shading = ensemble correlative model).
Figure 7. Mobility-oriented parity (MOP) assessment outputs for projections of the ensemble correlative model for *Calonectria pseudonaviculata*. Areas with MOP metric values close to 1 have highly comparable climatic conditions to those in the model calibration area. Areas with values approaching 0 indicate higher extrapolation because one or more climatic variables have values outside the range of variable(s) in the calibration area.
Figure 8. Climate comparisons for sites which are expected to differ in favorability for boxwood blight infections. Line plots depict monthly temperature (solid lines) and precipitation (dashed lines) across eight sites in Europe (orange lines) and the United States (blue lines). Sites with a Mediterranean climate (e.g., Cannes, France; Naples, Italy; Seattle, Washington; and Portland, Oregon) are less conducive for infections than sites which have higher humidity, few gaps in precipitation, and ideal temperatures for growth throughout the year, such as those in temperate/coastal climates in western Europe (e.g., Brussels, Belgium and Bordeaux, France) and warm and humid climates in the mid-Atlantic and southeastern regions of the United States (e.g., Virginia Beach, Virginia and Atlanta, Georgia). Data source: 1981-2010 climate normals, World Meteorological Organization (https://climatedata-catalogue.wmo.int; accessed 24 Sep 2021).