

ORIGINAL ARTICLE OPEN ACCESS

How Accurately Do Species Distribution Models Predict the Expansion of Invasive Insects, and Does Climate Data Choice Matter? Insights From the Invasion of *Dryocosmus kuriphilus*

José Carlos Pérez-Girón¹  | Fernando Castedo-Dorado²  | María Josefa Lombardero³  | Pedro Álvarez-Álvarez⁴ 

¹Department of Agroforestry Sciences, University of Huelva, Huelva, Spain | ²DRACONES Research Group, Departamento de Ingeniería y Ciencias Agrarias, Universidad de León, Escuela de Ingeniería Agraria y Forestal, León, Spain | ³Unidad de Gestión Ambiental y Forestal Sostenible, Universidad de Santiago de Compostela, Lugo, Spain | ⁴Department of Organisms and Systems Biology, Polytechnic School of Mieres, University of Oviedo, Mieres, Spain

Correspondence: José Carlos Pérez-Girón (josecarlos.perez@dcaf.uhu.es)

Received: 8 December 2025 | **Revised:** 12 March 2026 | **Accepted:** 26 March 2026

Keywords: biological invasions | chestnut gall wasp | ecological modelling | habitat suitability | invasive pest | species distribution models

ABSTRACT

Species distribution models (SDMs) are widely used to predict the spread of invasive species, yet their accuracy over time and the influence of climate data resolution remain unclear. Here, we examine the capacity of SDMs to predict the distribution and short-term expansion of the invasive gall wasp *Dryocosmus kuriphilus*, and compare the performance of locally sourced, annually resolved climate data with global climatic datasets. We modelled the annual distribution and expansion of *D. kuriphilus* using SDMs. Three climate datasets: regional interpolations, temporally explicit CHELSA time series, and long-term CHELSA averages, were compared to test their influence on model accuracy. Habitat suitability was estimated with GLM, MaxEnt, and Random Forest, and model performance was evaluated with the Boyce index. Temporal transferability was assessed by projecting early-year models to subsequent years and analysing accuracy patterns via multifactorial ANOVA. Model accuracy, measured with the Boyce index, improved over time, surpassing 0.7 for most datasets and algorithms. Differences among climate datasets were minor, although regional data slightly enhanced early predictions. MaxEnt consistently achieved the highest and most stable performance, while Random Forest and GLM were more variable. Models projected beyond their calibration years rapidly lost predictive power, showing limited temporal transferability. ANOVA confirmed that neither climate dataset nor algorithm significantly influenced accuracy, though MaxEnt performed best overall. SDMs are valuable tools for invasive species management when used adaptively. Accuracy improves with more occurrence data but declines when projecting future expansion, highlighting the need for regular updates. Incorporating fine-scale, temporally explicit data enhances early detection, monitoring, and rapid-response interventions. Continuously updated SDMs support timely management decisions, including invasion monitoring, prioritisation of actions, and integration of biological control or other mitigation strategies.

1 | Introduction

Biological invasions are one of the leading drivers of global biodiversity loss, posing serious threats to the structure and

functioning of ecosystems (Linders et al. 2019; Pyšek et al. 2020). Invasive species can outcompete native organisms, alter habitat conditions, disrupt trophic interactions, and ultimately lead to declines in species richness and ecosystem resilience. Once

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2026 The Author(s). *Journal of Applied Entomology* published by Wiley-VCH GmbH.

established, their impacts are often exacerbated by their rapid spread and the difficulty of controlling them, leading to irreversible changes in ecosystem function and substantial economic losses to the socio-ecological systems in which they are embedded (Cuthbert et al. 2021; Diagne et al. 2021). Globalisation and climate change are accelerating these processes by increasing the frequency of species introductions and expanding the ecological niches available to invaders (Hulme 2017). Increased global trade and human movement have facilitated the spread of alien species, with invasion rates fluctuating in response to waves of globalisation, while climate change destabilises native ecosystems and creates favourable conditions for non-native species (Walther et al. 2009). Therefore, while globalisation and climate change continue to facilitate the introduction and establishment of non-native species, the early detection, monitoring, and accurate prediction of biological invasions have become critical priorities for pest control, biodiversity conservation and ecosystem management (Barker et al. 2020).

To that end, species distribution models (SDMs) provide useful tools for addressing the complex challenges posed by biological invasions (Battini et al. 2019; Hill et al. 2017; Mainali et al. 2015). By correlating species occurrence data with environmental variables using statistical methods, SDMs can estimate the potential geographic distribution of invasive species under current and future conditions, generating spatially explicit predictions which enable the identification of colonised areas or at risk of colonisation, even before invasions occur (Sillero et al. 2021; Srivastava et al. 2019). Although SDMs are sometimes criticised for assuming equilibrium between a species and its environment, a condition rarely met during the early stages of invasion, their practical value remains widely recognised (Barbet-Massin et al. 2018; Gallien et al. 2012; Nguyen and Leung 2022; Václavík and Meentemeyer 2012; Wen et al. 2024). This is particularly true in supporting early detection, refining monitoring strategies and prioritising management actions. For biological control programs, SDMs are especially useful in pinpointing infested areas where control agents, such as parasitoids, can be most effectively released (Fischbein et al. 2019; Sofaer et al. 2019; Sutton and Martin 2022). However, when dealing with rapidly spreading pests, the need for timely, spatially accurate information becomes critical.

In this context, global climate datasets like CHELSA or WorldClim provide standardised climatic information at broad spatial scales, which has made them widely used in species distribution modelling (Fick and Hijmans 2017; Karger et al. 2017). However, these datasets are primarily designed to represent long-term climatic conditions and their aggregated products (e.g., monthly variables commonly used as SDM predictors) are often released with delays of several years relative to the most recent observations. Although this does not imply gaps in the underlying meteorological records, it limits the temporal currency of the predictors typically used in modelling exercises. This limitation may be particularly relevant in the context of rapidly expanding invasive species, where distributions can change over short time frames and management decisions often require environmental information that reflects current conditions. In such situations, locally available meteorological observations can be rapidly processed and spatialised using relatively simple interpolation techniques, providing timely, accurate and spatially explicit

climatic predictors representative of local climatic conditions. These locally derived datasets may therefore complement global climatologies when near-real-time environmental information is required, for instance to support time-sensitive management actions, such as the targeted release of biological control agents, where precise and current environmental information is key for operational success.

Dryocosmus kuriphilus Yasumatsu, 1951 (Hymenoptera: Cynipidae), commonly known as the Asian chestnut gall wasp, represents an ideal candidate species for validating the predictive capacity of SDMs in the context of biological invasions and to assess the importance of climate data. This is because its invasion history is well documented, its spread has been intensively monitored, and its expansion has occurred largely in the absence of effective biological control agents. Native to China, this invasive insect has rapidly spread across Asia, Europe and North America (Gil-Tapetado et al. 2018, 2020), becoming one of the most damaging pests of chestnut trees (*Castanea* spp.) (EFSA Panel on Plant Health 2010). A key factor behind its successful invasion is the absence of natural predators in newly colonised regions, allowing populations to grow unchecked and spread rapidly, with ecological and economic impacts, including reduced tree vigour, decreased nut production, and heightened vulnerability to secondary pathogens (Battisti et al. 2014; Sartor et al. 2015). In the case of *D. kuriphilus*, several native parasitoid species have been reported attacking this pest, although their effectiveness is generally limited, with parasitism rates typically remaining low in invaded regions (Gil-Tapetado, Cabrero-Sañudo, et al. 2021; Panzavolta et al. 2013; Quacchia et al. 2013). Therefore, biological control strategies have focused on the introduction of *Torymus sinensis* Kamijo, a parasitoid widely used in biological control programs and currently the most effective agent available for controlling *D. kuriphilus* populations (Avtzis et al. 2019; Ferracini et al. 2017; Quacchia et al. 2008). The success of this biological control program depends on the timely and precise identification of infested areas, where *T. sinensis* can be released to parasitize galls and reduce pest densities.

In Galicia (northwest Spain), where the Asian chestnut gall wasp was first detected in 2014 (Pérez Otero and Mansilla Vázquez 2014), its spread has been systematically monitored through annual field surveys (Gil-Tapetado et al. 2020; Pérez Otero et al. 2017; Xunta de Galicia 2020). This has resulted in a rich and temporally consistent dataset for model calibration and validation, particularly valuable given the absence of an effective control by introduced agents until 2019. Although experimental releases of *T. sinensis* began in 2015, its establishment in Galicia remained very limited for several years. By 2017, only 12 individuals had been recovered from over 220,000 released, and even after a massive release campaign in 2018 (over 700,000 individuals), post-release monitoring still showed extremely low recovery rates, with only one specimen detected in field surveys (Nieves-Aldrey et al. 2019). As observed in other regions, the establishment and population growth of *T. sinensis* typically require several years after its release, with noticeable increases in parasitoid populations often occurring after 5–7 years and effective control sometimes taking even longer (Ferracini et al. 2018; Murakami and Gyoutoku 1995). Therefore, meaningful parasitism levels in Galicia may not be expected until after 2020. This delayed

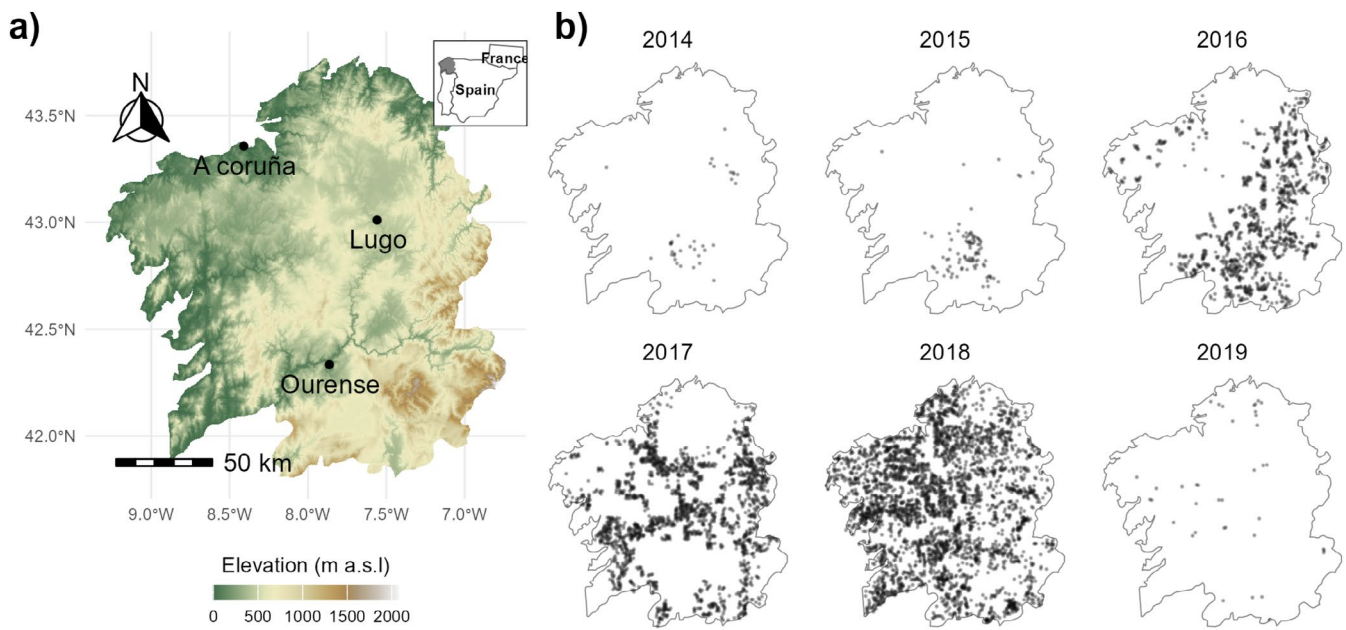


FIGURE 1 | (a) Study area and (b) annual records of *Dryocosmus kuriphilus* invasion from 2014 to 2019. Source: Gil-Tapetado et al. (2020). [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/doi/10.1111/jen.12013)]

establishment highlights the importance of SDMs for guiding biological control strategies, particularly in identifying suitable release sites and optimal timing under changing environmental conditions.

The years preceding the establishment of the parasitoid provide an ideal opportunity to assess the capacity of SDMs to predict the expansion of the invasive species, *D. kuriphilus*. They also allow for the evaluation of whether annually resolved, locally sourced meteorological data could have improved predictive performance compared to traditional macroclimatic datasets, which are typically released with a delay and reflect long-term averages rather than real-time conditions. In this context, the aim of this study is twofold: (i) to evaluate the ability of species distribution models (SDMs) to predict current distribution and the short-term expansion of rapidly spreading insect species across multiple years, using *D. kuriphilus* as a model organism; (ii) to examine whether locally sourced meteorological data, spatialised through relatively simple interpolation techniques, can outperform or provide comparable accuracy to widely used global climatic datasets, which are typically processed using complex methodologies and provided with a certain delay. This will help inform both the theoretical understanding of the limitations of the SDM for insect invasion ecology and practical guidelines for planning future biological pest control.

2 | Material and Methods

2.1 | Study Area

The study was conducted in Galicia, a region in northwest Spain covering approximately 29,575 km² (Figure 1a). Climate is predominantly temperate oceanic, with regional variations driven by proximity to the Atlantic Ocean and elevation. Particularly,

the coastal areas experience mild temperatures year-round, with abundant and evenly distributed precipitation, while the inland zones display a more pronounced continental influence, with greater seasonal temperature variability. Annual precipitation averages around 1200 mm, although it varies greatly across the region, reaching up to 2900 mm in some areas and dropping to much lower values in interior valleys. Rainfall is distributed fairly evenly throughout the year, with a pronounced winter peak and a summer minimum, characteristic of an oceanic climate with abundant precipitation.

The interplay between climate, topography, and human activity shapes land use and vegetation distribution across the region. Thus, one of the characteristic tree species, the European chestnut (*Castanea sativa* Mill.), officially occupies over 490 km². However, its presence extends far beyond this estimate, as it is closely associated with rural societies and appears throughout the landscape as scattered individuals, small stands and traditional agroforestry systems, highlighting its multifaceted ecological, economic, and social importance.

2.2 | *D. kuriphilus* Occurrences

Annual records of *D. kuriphilus* invasion from 2014 to 2019 were obtained from surveys conducted by Galician forestry rangers and technicians and provided by the regional government of Galicia (Figure 1b) (Gil-Tapetado et al. 2020; Xunta de Galicia 2020). Monitoring was carried out during the spring and summer months through systematic sampling along transects, either on foot or by car, to identify the presence of galls. Additionally, chestnut growers frequently reported gall infestations in their orchards, supplementing the official records. When an infestation was detected in a chestnut stand, a single tree was recorded as representative of the stand. For isolated trees, the precise position of the affected tree was recorded.

The geographic coordinates of infested trees were obtained using handheld GPS receivers or smartphones, with an average positional error of $\pm 3\text{--}5$ m. A detailed description of the monitoring protocol, including sampling periods and gall detection procedures, can be found in the original sources (Gil-Tapetado et al. 2020; Xunta de Galicia 2020).

Since the primary objective of monitoring was to delineate the geographical extent of the invasion and inform management strategies rather than systematically survey all chestnut stands, potential spatial biases in the dataset were addressed. For that, occurrences were randomly subsetting at a minimum distance of 1 km from each other using the 'ecoinfo' package (Carlisle and Albeke 2016) in R 4.3.3 (R Core Team 2025). This preprocessing step helped reduce overrepresentation of closely located occurrences, improving the robustness of subsequent spatial analyses. After filtering, 6676 affected chestnut trees were retained from an initial total of 17,517 geolocated records. The number of presences increased markedly over time, starting with 37 in 2014, rising to 92 in 2015, and then sharply expanding to 1065 in 2016 and 1965 in 2017. The invasion peaked in 2018 with 3479 presences, before declining to 38 in 2019. This sharp decrease most likely reflects limited data availability for that year, as the records compiled in the source dataset are incomplete. Although monitoring activities continued, the full set of observations collected by the regional administration was not made publicly available (Gil-Tapetado et al. 2020; Xunta de Galicia 2020).

2.3 | Climate Data

Previous studies have identified climatic factors as key drivers of the distribution and dispersal of *D. kuriphilus* (Gil-Tapetado, Castedo-Dorado, et al. 2021; Gil-Tapetado et al. 2020; Quinto et al. 2021), with wind playing a particularly important role in facilitating its movement across regions (Mapes et al. 2020). To account for these influences, we computed 19 bioclimatic variables using three different approaches, incorporating multiple data sources and temporal aggregation scales. The first approach provides a fast, straightforward, yet reliable method for obtaining annual climate data. The second requires minimal expert processing while offering a standardised, high-resolution dataset. The third relies on long-term climatic averages, serving as a baseline for assessing climatic variability and trends.

First, we obtained monthly records of maximum and minimum temperature, precipitation, and wind speed from 140 meteorological stations managed by the Meteorological Observation and Prediction Unit of Galicia (MeteoGalicia) for the period 2014–2018. To generate continuous spatial coverage, we applied Universal Kriging interpolation, using station elevation as a covariate to improve accuracy (see root mean square error (RMSE) values for each variable in Table S1). Elevation was aggregated to a pixel size of 1 km using the mean. Subsequently, we computed the 19 bioclimatic variables on an annual basis using the interpolated temperature and precipitation data at 1 km of spatial resolution. However, instead of following the conventional calendar year, we adopted a biologically relevant time frame aligned with the *D. kuriphilus* life cycle. Specifically, we aggregated monthly climate data from June of year $n-1$ to May of year n , ensuring that the climatic conditions preceding and during

key phenological stages of the pest were adequately captured. The wind speed layers were used as computed for the months of June to August of each year, corresponding to the period when the insect is actively flying (Gil-Tapetado et al. 2020), although the activity period may extend into September in some regions (Lobo Santos et al. 2024).

Second, we obtained preprocessed monthly climate time series from the CHELSA dataset (Karger et al. 2017) for the period 2014–2018 at a spatial resolution of 1 km. Using the same methodology as in the first approach, we computed the 19 bioclimatic variables on an annual basis. Specifically, each year's bioclimatic variables were derived from climate data spanning June of year $n-1$ to May of year n , ensuring alignment with the *D. kuriphilus* life cycle. Wind speed layers were used as provided for the months of June to August of each year.

To assess the degree of divergence between climatic layers derived from MeteoGalicia and CHELSA, we computed spatial differences between the corresponding variables for each year of the study period. These comparisons allowed us to visually evaluate the magnitude and spatial structure of discrepancies between both datasets. Spatial patterns of differences were broadly consistent across years; therefore, a representative example for the year 2014 is shown in Figure S1.

Third, we directly obtained the 19 bioclimatic variables from the CHELSA v2.1 dataset, which provides long-term climatic averages for the period 1981–2010 at a spatial resolution of 1 km (Karger et al. 2017). All bioclimatic variables were computed using the 'dismo' R package (Hijmans et al. 2024).

All environmental predictors were screened for multicollinearity using Spearman-based correlation ($|r| \geq 0.7$) and Variance Inflation Factors ($VIF \geq 5$). Predictors exceeding these thresholds were removed and only the subset of non-collinear variables was retained for model calibration.

2.4 | Species Distribution Modelling

Habitat suitability for *D. kuriphilus* was modelled annually. For each modelling year, the occurrence dataset included all presences recorded in that year and those detected in previous years, assuming that once established the species remains detectable in the invaded locations. For each set of climate predictors described above we used three different modelling techniques: Generalised Linear Models (GLM), Maximum Entropy (MaxEnt), and Random Forest (RF). These models were implemented within the 'biomod2' package (version 4.2–6–2) (Thuiller et al. 2023) in R (R Core Team 2025). GLM is a commonly used parametric approach that provides an interpretable framework for assessing the influence of environmental predictors. MaxEnt is a deterministic, presence-only modelling method (i.e., results remain consistent across runs) that is particularly well-suited for handling small sample sizes with a strong predictive performance while maintaining computational efficiency (Phillips et al. 2006; Sillero et al. 2021; Valavi et al. 2022). Lastly, a Random Forest (RF) was employed as a machine learning technique capable of modelling complex and nonlinear relationships with a high predictive performance (Valavi et al. 2022).

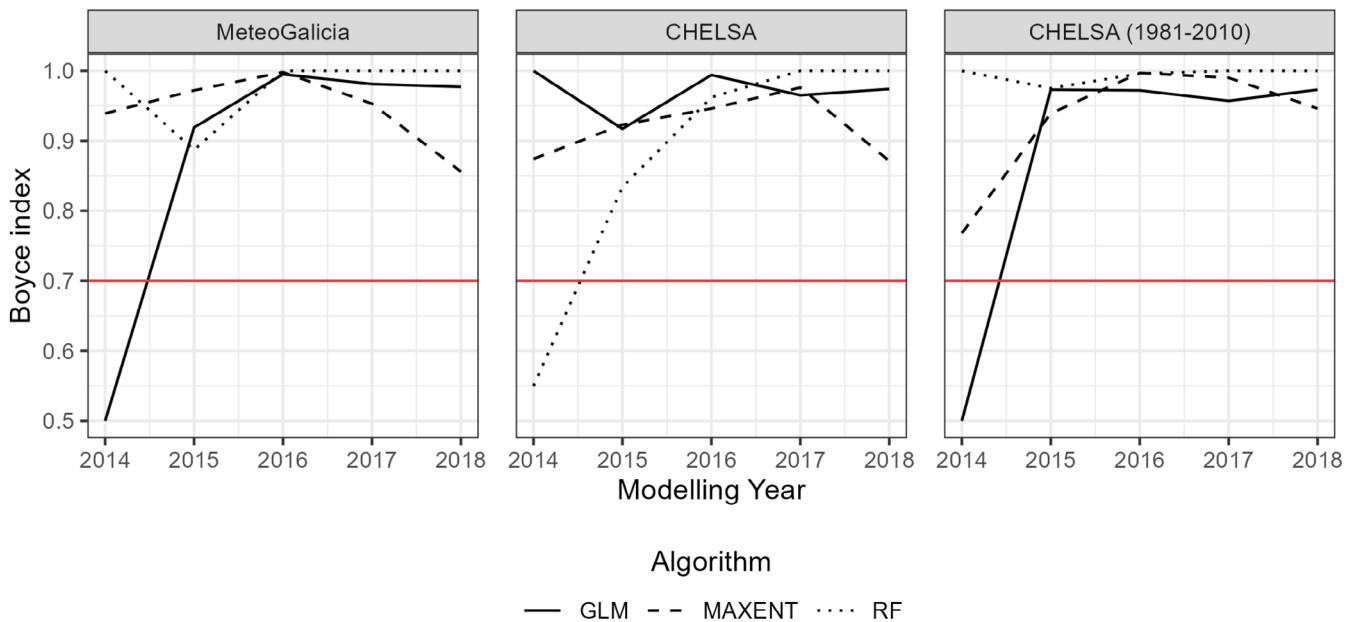


FIGURE 2 | Internal evaluation of model performance used to select model runs and construct the ensemble models. Model performance measured by the Boyce index for each year, climate dataset and modelling algorithm. The horizontal red line represents the threshold of 0.7, which was used to select models for median ensemble construction. [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/doi/10.1111/jen.20103)]

Since only presence data were available for each year, we generated 10 sets of randomly distributed pseudo-absences, each consisting of 10,000 points placed at a minimum distance of 5 km from known presences, aligning with the initial radial spread rate of *D. kuriphilus* (Gil-Tapetado et al. 2020). This approach follows commonly used pseudo-absence generation strategies that avoid placing pseudo-absences too close to presence locations (Barbet-Massin et al. 2012; VanDerWal et al. 2009). These pseudo-absences were used for the GLM and RF models. For MaxEnt, which operates using background data rather than pseudo-absences, the same points were treated as background locations within the ‘biomod2’ framework. To mitigate biases associated with imbalanced data, we then applied a prevalence of 0.5 to balance the majority class (Barbet-Massin et al. 2012).

To assess model performance, we implemented a 10-fold cross-validation for each pseudo-absence set, resulting in a total of 100 model runs per modelling technique. Model performance was evaluated using the Boyce index (Hirzel et al. 2006), a metric specifically designed for presence-only models that quantifies how well predictions deviate from a random distribution of observed presences (Boyce et al. 2002). The Boyce index was preferred over discrimination metrics such as AUC or TSS due to their known limitations related to the spatial extent of the study area and the spatial distribution of the species (Bracho-Estévez et al. 2024; Sillero et al. 2021; Leroy et al. 2018; Lobo et al. 2008). The Boyce index ranges from -1 to 1 , where positive values indicate that predicted suitability aligns well with observed presence locations, suggesting good model performance. Values near zero imply that predictions do not significantly differ from a random distribution, while negative values indicate counter predictions, meaning the model assigns lower suitability to areas where presences are more frequent. This metric was also used to build median ensembles for each algorithm, retaining only model runs with Boyce index scores greater than 0.7.

2.5 | Modelling Post-Processing

The median ensemble for each year, along with species occurrence data from subsequent years, was used to calculate the Boyce index related to the *D. kuriphilus* expansion using the ‘ecospat’ R package (Broennimann et al. 2025). In that case, the Boyce index allows us to assess how well the model for a given year predicts the species’ expansion in the following years.

Furthermore, to statistically evaluate the influence of climate data sources and algorithm selection on model performance, we conducted a post hoc analysis using a linear model. Specifically, we fitted a multifactorial ANOVA using a linear model, testing whether different climate datasets and modelling algorithms significantly influenced the Boyce index (dependent variable), which we used as an evaluation metric of model accuracy. Two separate models were constructed. The first model examined the influence of climate datasets and algorithms on the Boyce index values obtained from the presence data used to train each model, assessing their effects on current habitat suitability predictions. The second model focused exclusively on the Boyce index values calculated for species expansion, evaluating how well models projected future occurrences.

3 | Results

3.1 | Predictive Accuracy and Temporal Stability of SDMs

Model performance, as measured by the Boyce index, generally improved over time across all climate datasets and modelling algorithms, indicating increased model accuracy with more recent data (Figures 2 and 3; see Figure S2 for variables importance). With the exception of 2014, all models achieved Boyce index values above the 0.7 threshold (red line), indicating a good fit. No

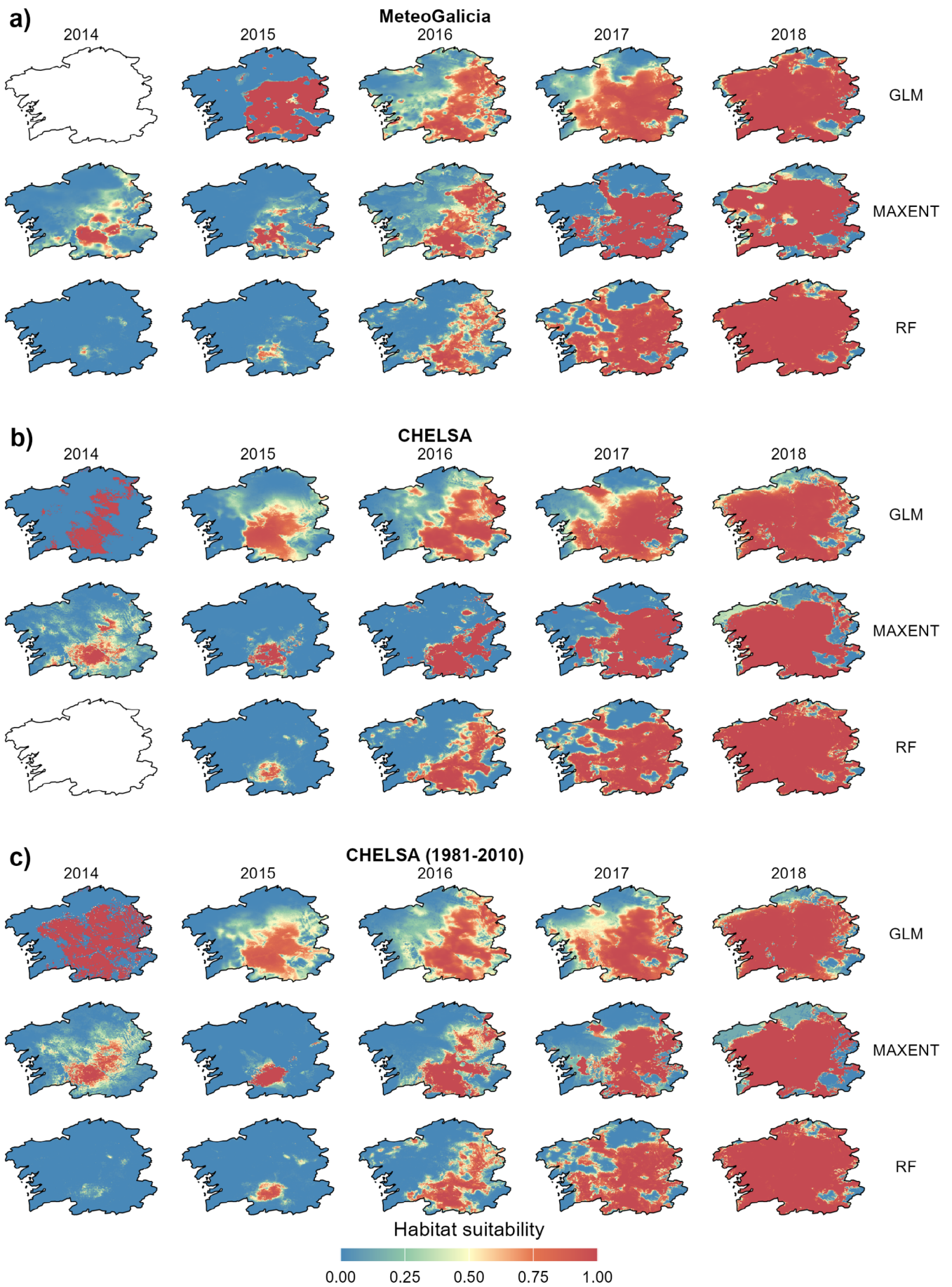


FIGURE 3 | Legend on next page.

FIGURE 3 | Habitat suitability models (median ensemble) for *Dryocosmus kuriphilus* in Galicia for each calibration year (2014–2018), climate dataset, and modelling algorithm. Panels show predicted habitat suitability (0–1) derived from ensemble models built from individual SDM runs that passed the internal evaluation threshold (Boyce index > 0.7). Rows correspond to modelling algorithms (GLM, MaxEnt and Random Forest), while columns represent the calibration year. Results are shown separately for the three climate datasets used in this study: (a) MeteoGalicia, (b) CHELSA and (c) CHELSA climatology (1981–2010). [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com)]

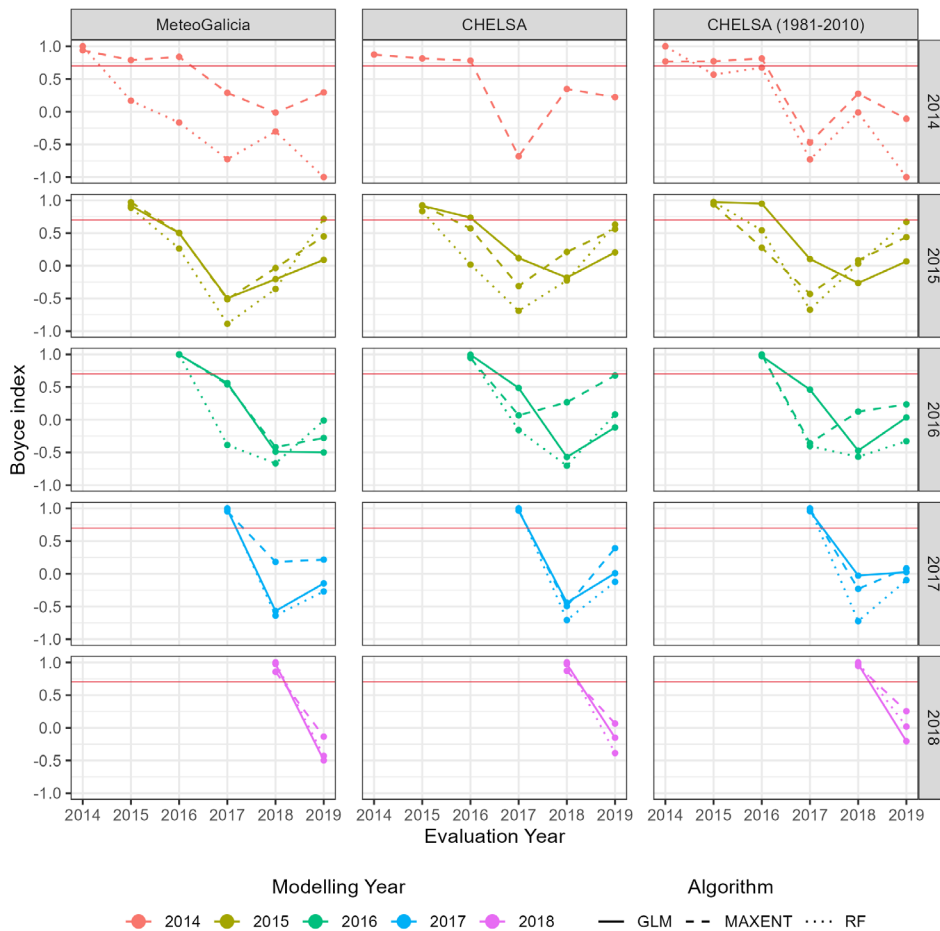


FIGURE 4 | Temporal transferability of the models evaluated using the Boyce index. Models calibrated in each training year (indicated by colours and labels on the right y-axis) were evaluated against the observed expansion of *Dryocosmus kuriphilus* in subsequent years (labels on the x-axis) across climate datasets and modelling algorithms. The horizontal red line indicates the 0.7 threshold used for model selection in the ensemble construction and is shown here only as a reference. Some algorithm-dataset combinations are not displayed for specific calibration years because the corresponding models either did not reach the minimum threshold (Boyce index > 0.7) or produced predictions with insufficient variability to compute the Boyce index during the temporal evaluation step (see Figure 3). [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com)]

substantial differences were observed between climate datasets, although MeteoGalicia appeared to yield the most robust results in the first year. The apparent differences between algorithms in the CHELSA panel for 2014 are mainly driven by model instability during the first invasion year. In terms of modelling techniques, GLM exhibited the lowest performance in 2014, with Boyce index values falling below the threshold for MeteoGalicia and CHELSA (1981–2010). Particularly, for the CHELSA dataset in 2014, GLM produced a Boyce index value of 1 but failed in spatial prediction, indicating an artefact in the results. However, GLM performance improved in subsequent years, reaching high and stable Boyce index values. MaxEnt consistently performed above the threshold across all years, though a slight decline was observed in 2018. Finally, RF displayed higher variability in the

early years, falling below the threshold in 2014 for CHELSA but achieving the highest Boyce index values from 2016 onward. Models that did not meet the 0.7 threshold were excluded from further analysis.

Across all climate datasets and modelling algorithms, the Boyce index generally declined as models were applied to future evaluation years (Figure 4). Models trained in earlier years (e.g., 2014, 2015) performed well in predicting near-future expansion but exhibited a sharp decline in accuracy as the time gap between the model training year and evaluation year increased. In contrast, models trained from 2015 onwards were unable to accurately predict the species' expansion in the following year. These trends were consistent across all climate datasets. Furthermore,

TABLE 1 | Summary of multifactorial ANOVA fitted to assess the effects of climate database and algorithm over the current model performance measured by the Boyce index.

	Variable	Coefficient	SE	<i>p</i>	<i>R</i> ²	<i>R</i> ² _{adj}
Intercept	MeteoGalicia—GLM	0.973	0.018	<0.001	0.180	0.089
Climate dataset	CHELSA	−0.020	0.020	0.314		
	CHELSA (1981–2010)	0.001	0.019	0.974		
Algorithm	MaxEnt	−0.037	0.020	0.072		
	RF	0.008	0.020	0.696		

Note: Estimated coefficients, standard errors (SE), and *p*-values are reported.

TABLE 2 | Summary of multifactorial ANOVA fitted to assess the effects of climate database and algorithm over the model performance measured by the Boyce index for *Dryocosmus kuriphilus* expansion.

	Variable	Coefficient	SE	<i>p</i>	<i>R</i> ²	<i>R</i> ² _{adj}
Intercept	MeteoGalicia—GLM	−0.102	0.098	0.302	0.149	0.118
Climate dataset	CHELSA	0.103	0.102	0.313		
	CHELSA (1981–2010)	0.103	0.098	0.293		
Algorithm	MaxEnt	0.210	0.103	<0.05		
	RF	−0.187	0.106	0.080		

Note: Estimated coefficients, standard errors (SE) and *p*-values are reported.

in early years, when predictions remained reliable for subsequent years, GLM and MaxEnt emerged as the most robust modelling approaches. RF demonstrated higher volatility, with performance declining rapidly over time.

3.2 | The Role of Modelling Algorithms and Climate Data in SDM Predictive Performance

Based on multifactorial ANOVA results, using CHELSA instead of MeteoGalicia resulted in a slight decrease in current model performance, while CHELSA (1981–2010) led to a slight increase. However, these effects were not statistically significant (Table 1). Similarly, the choice of modelling algorithm did not have a significant impact on performance. Therefore, neither the climate dataset nor the algorithm selection significantly influenced model accuracy.

For *D. kuriphilus* expansion, neither the choice of climate dataset nor the selection of modelling algorithms had a statistically significant impact on model accuracy. However, MaxEnt generally showed higher Boyce index values than GLM and RF across the evaluated models (Table 2).

4 | Discussion

This study provides an empirical assessment of how accurately SDMs can predict the short-term expansion of invasive insects and evaluates whether the choice of climate data meaningfully influences those predictions. Using *D. kuriphilus* as a model organism, our findings reveal that SDMs are capable of reliably capturing the realised distribution of a rapidly spreading pest,

particularly once sufficient occurrence data become available, but their predictive accuracy declines markedly when extrapolated beyond the calibration period. The choice between global and locally sourced climatic datasets did not significantly affect model performance, suggesting that, at least for regional applications in relatively homogeneous environments, macroclimatic datasets can provide predictions comparable to those derived from high-resolution local data. These results highlight both the strengths and limitations of SDMs in invasion ecology: while they remain valuable tools for early detection and management prioritisation, their temporal stability is constrained, underscoring the need for iterative updating and the integration of near-real-time environmental information for operational use in pest control programs.

4.1 | Predictive Accuracy and Temporal Stability of SDMs

Consistent with previous research (Barbet-Massin et al. 2018; Nguyen and Leung 2022; Wen et al. 2024), our models achieved moderate to good predictive accuracy, with Boyce index values generally above 0.7 across most datasets and algorithms. This indicates that SDMs can effectively capture the realised distribution of an invasive insect even during its early expansion phase. However, model accuracy declined sharply when projections were extended beyond the calibration period, highlighting the limited temporal transferability of SDMs in dynamic invasion contexts (Nguyen and Leung 2022). This temporal decay likely arises from a combination of ecological and methodological factors. First, *D. kuriphilus* exhibits high dispersal capacity and an ability to exploit new climatic niches (Battisti et al. 2014; Gil-Tapetado, Castedo-Dorado, et al. 2021; Gil-Tapetado et al. 2018,

2020), which may cause models calibrated during early establishment to underestimate subsequent range expansion into new environments. Second, the assumption of species-environmental equilibrium inherent to correlative SDMs (Barbet-Massin et al. 2018; Gallien et al. 2012; Václavík and Meentemeyer 2012) is violated during periods of rapid invasion, leading to potential overfitting to initial conditions and reduced predictive reliability through time. This pattern is consistent with previous studies showing that SDMs calibrated during the early stages of an invasion tend to be less accurate and often underestimate the potential range compared to models developed once the species approaches environmental equilibrium (Václavík and Meentemeyer 2012).

In this sense, model performance improved progressively across the study period, suggesting that as the species' range approached environmental equilibrium within Galicia, SDMs became more robust in capturing its realised niche. During the first years of invasion, *D. kuriphilus* rapidly expanded across the region in the absence of effective natural enemies (Gil-Tapetado, Cabrero-Sañudo, et al. 2021; Gil-Tapetado et al. 2020), benefiting from the lack of biotic resistance. This phase of largely unrestricted spread continued until 2019, when biological control by *T. sinensis* had not yet become effective despite earlier experimental releases. By that time, *D. kuriphilus* had already colonised nearly the entire study area, reaching a broad distribution. Consequently, the improvement in model accuracy over time may reflect the species' progressive stabilisation within available habitats and a closer match between its realised and potential niche, conditions under which correlative SDMs perform more reliably. This aligns with findings from invasive plant and insect studies reporting higher accuracy when models are calibrated with more extensive, post-establishment occurrence data (Barbet-Massin et al. 2018; Nguyen and Leung 2022).

4.2 | The Role of Modelling Algorithms and Climate Data in SDM Predictive Performance

Among modelling algorithms, MaxEnt consistently yielded the most stable and accurate predictions for both current distribution and short-term expansion, corroborating earlier findings that MaxEnt performs well under small-sample or presence-only conditions typical of invasion monitoring (Battini et al. 2019; Sutton and Martin 2022; Valavi et al. 2022). In contrast, RF showed greater variability in early years and lower temporal consistency, possibly due to its sensitivity to sampling bias and data imbalance (Liu et al. 2009). GLM, while initially performing poorly, improved over time as more data became available, supporting the view that simpler statistical models can approximate machine-learning approaches when sufficient occurrence data exist.

Contrary to assumptions and previous findings suggesting that the choice of climatic data source can markedly influence SDM outcomes (Peterson and Nakazawa 2008; Wen et al. 2024), our analyses revealed no statistically significant differences between global (CHELSA) and locally derived (MeteoGalicia) datasets. Although the comparison of the climatic layers revealed noticeable spatial differences between both datasets (Figure S1), these discrepancies did not translate into meaningful differences in

model predictions. While the locally interpolated data slightly improved predictions during the early invasion years, particularly in 2014, these advantages diminished as the species expanded and its distribution became more stable. This result suggests that, at least at the spatial and temporal scales considered here, SDM predictions for *D. kuriphilus* are relatively robust to differences among climatic datasets. More broadly, our findings indicate that moderate discrepancies among climate data sources may not necessarily compromise the reliability of short-term invasion forecasts when the underlying environmental gradients are consistently captured.

However, from a management perspective, the operational value of high-resolution, locally sourced climate data should not be overlooked. Although global datasets such as CHELSA provide standardised and widely comparable climatic layers, they reflect long-term averages that may fail to capture short-term fluctuations or atypical seasons—factors that can strongly influence insect phenology and establishment success. In contrast, local datasets offer near-real-time climatic information, which is especially valuable for designing and implementing biological control strategies. In the case of *D. kuriphilus* and its parasitoid *T. sinensis*, incorporating temporally precise climatic data could improve the timing and spatial targeting of release programs, ensuring that interventions align with favourable environmental conditions for parasitoid establishment and dispersal. Thus, even in contexts where overall predictive performance remains statistically similar across data sources, the integration of dynamic, locally resolved climate information can substantially enhance the practical applicability of SDMs for invasive species management.

5 | Conclusion

This study demonstrates that SDMs can effectively describe the realised distribution and spatial expansion of invasive species, but their predictive accuracy tends to decline when extrapolated through time. As invasive expansion progresses towards environmental equilibrium within a region, model performance generally improves, reflecting the dynamic relationship between invasion stage and model reliability. Although noticeable spatial differences existed between the climatic layers derived from global (CHELSA) and locally interpolated (MeteoGalicia) datasets, these discrepancies did not translate into meaningful differences in model predictions, suggesting that SDM outputs were relatively robust to the choice of climatic data source at the spatial and temporal scales considered here. However, the inclusion of temporally explicit and fine-scale climatic information remains important for enhancing the operational value of SDMs in real-time monitoring and rapid-response management.

From an applied standpoint, these findings underscore the need for adaptive modelling frameworks that evolve in parallel with invasion progression. For rapidly expanding pests such as *D. kuriphilus*, regular model updates, ideally on an annual basis, are essential to maintain predictive relevance and guide monitoring, containment, and biological control strategies in near real time. To further enhance model responsiveness to ongoing invasion, as *T. sinensis* establishes and begins to regulate *D. kuriphilus* populations, biotic interactions should be explicitly integrated

into future models to account for potential range contractions or spatial refugia driven by parasitoid dynamics.

Author Contributions

José Carlos Pérez-Girón: conceptualisation, methodology, data curation, formal analysis, writing – original draft, writing – review and editing. **Fernando Castedo-Dorado:** conceptualisation, data curation, writing – review and editing, formal analysis. **María Josefa Lombardero:** conceptualisation, data curation, writing – review and editing, formal analysis. **Pedro Álvarez-Álvarez:** conceptualisation, data curation, writing – review and editing, formal analysis.

Acknowledgements

Funding for open access charge: Universidad de Huelva / CBUA.

Funding

The authors have nothing to report.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The occurrence data of *Dryocosmus kuriphilus* analysed in this study are freely available at <https://doi.org/10.1111/jen.12836>. Meteorological station data (temperature, precipitation, and wind speed) for 2014–2018 were obtained from MeteoGalicia and are publicly accessible through the MeteoGalicia data portal (<https://www.meteogalicia.gal>), subject to their data-use conditions. The CHELSA climate datasets used in this study (monthly climate data for 2014–2018 and long-term bioclimatic variables for 1981–2010) are freely available at <https://chelsa-climate.org>. All scripts used for data preprocessing, species distribution modelling and post-processing analyses are publicly archived in Zenodo and accessible at: <https://doi.org/10.5281/zenodo.1895969>.

References

- Avtzis, D. N., G. Melika, D. Matošević, and D. R. Coyle. 2019. “The Asian Chestnut Gall Wasp *Dryocosmus kuriphilus*: A Global Invader and a Successful Case of Classical Biological Control.” *Journal of Pest Science* 92, no. 1: 107–115. <https://doi.org/10.1007/s10340-018-1046-1>.
- Barbet-Massin, M., F. Jiguet, C. H. Albert, and W. Thuiller. 2012. “Selecting Pseudo-Absences for Species Distribution Models: How, Where and How Many?” *Methods in Ecology and Evolution* 3, no. 2: 327–338. <https://doi.org/10.1111/j.2041-210X.2011.00172.x>.
- Barbet-Massin, M., Q. Rome, C. Villemant, and F. Courchamp. 2018. “Can Species Distribution Models Really Predict the Expansion of Invasive Species?” *PLoS One* 13, no. 3: e0193085. <https://doi.org/10.1371/journal.pone.0193085>.
- Barker, B. S., L. Coop, T. Wepprich, F. Grevstad, and G. Cook. 2020. “DDRP: Real-Time Phenology and Climatic Suitability Modeling of Invasive Insects.” *PLoS One* 15, no. 12: e0244005. <https://doi.org/10.1371/journal.pone.0244005>.
- Battini, N., N. Farias, C. B. Giachetti, E. Schwindt, and A. Bortolus. 2019. “Staying Ahead of Invaders: Using Species Distribution Modeling to Predict Alien Species’ Potential Niche Shifts.” *Marine Ecology Progress Series* 612: 127–140. <https://doi.org/10.3354/meps12878>.
- Battisti, A., I. Benvegnù, F. Colombari, and R. A. Haack. 2014. “Invasion by the Chestnut Gall Wasp in Italy Causes Significant Yield Loss in *Castanea sativa* Nut Production.” *Agricultural and Forest Entomology* 16, no. 1: 75–79. <https://doi.org/10.1111/afe.12036>.

- Boyce, M. S., P. R. Vernier, S. E. Nielsen, and F. K. A. Schmiegelow. 2002. “Evaluating Resource Selection Functions.” *Ecological Modelling* 157, no. 2: 281–300. [https://doi.org/10.1016/S0304-3800\(02\)00200-4](https://doi.org/10.1016/S0304-3800(02)00200-4).

- Bracho-Estévez, C. A., S. Arenas-Castro, J. P. González-Varo, and P. González-Moreno. 2024. “Spatially Explicit Metrics Improve the Evaluation of Species Distribution Models Facing Sampling Biases.” *Ecological Informatics* 84: 102916. <https://doi.org/10.1016/j.ecoinf.2024.102916>.

- Broennimann, O., V. D. Cola, and A. Guisan. 2025. ecospat: Spatial Ecology Miscellaneous Methods. <https://CRAN.R-project.org/package=ecospat>.

- Carlisle, J., and S. Albeke. 2016. “Ecoinfo: Assorted tools for the management and analysis of ecological information.” *R package version 0.9.3*.

- Cuthbert, R. N., A. C. Bartlett, A. J. Turbelin, et al. 2021. “Economic Costs of Biological Invasions in the United Kingdom.” *NeoBiota* 67: 299–328. <https://doi.org/10.3897/neobiota.67.59743>.

- Diagne, C., B. Leroy, A.-C. Vaissière, et al. 2021. “High and Rising Economic Costs of Biological Invasions Worldwide.” *Nature* 592, no. 7855: 571–576. <https://doi.org/10.1038/s41586-021-03405-6>.

- EFSA Panel on Plant Health. 2010. “Risk Assessment of the Oriental Chestnut Gall Wasp, *Dryocosmus kuriphilus* for the EU Territory and Identification and Evaluation of Risk Management Options.” *EFSA Journal* 8, no. 6: 1619. <https://doi.org/10.2903/j.efsa.2010.1619>.

- Ferracini, C., S. Bertolino, U. Bernardo, et al. 2018. “Do *Torymus sinensis* (Hymenoptera: Torymidae) and Agroforestry System Affect Native Parasitoids Associated With the Asian Chestnut Gall Wasp?” *Biological Control* 121: 36–43. <https://doi.org/10.1016/j.biocontrol.2018.01.009>.

- Ferracini, C., E. Ferrari, M. Pontini, L. K. H. Nova, M. A. Saladini, and A. Alma. 2017. “Post-Release Evaluation of Non-Target Effects of *Torymus sinensis*, the Biological Control Agent of *Dryocosmus kuriphilus* in Italy.” *BioControl* 62, no. 4: 445–456. <https://doi.org/10.1007/s10526-017-9803-2>.

- Fick, S. E., and R. J. Hijmans. 2017. “WorldClim 2: New 1-Km Spatial Resolution Climate Surfaces for Global Land Areas.” *International Journal of Climatology* 37, no. 12: 4302–4315. <https://doi.org/10.1002/joc.5086>.

- Fischbein, D., M. V. Lantschner, and J. C. Corley. 2019. “Modelling the Distribution of Forest Pest Natural Enemies Across Invaded Areas: Towards Understanding the Influence of Climate on Parasitoid Establishment Success.” *Biological Control* 132: 177–188. <https://doi.org/10.1016/j.biocontrol.2019.02.016>.

- Gallien, L., R. Douzet, S. Pratte, N. E. Zimmermann, and W. Thuiller. 2012. “Invasive Species Distribution Models—How Violating the Equilibrium Assumption Can Create New Insights.” *Global Ecology and Biogeography* 21, no. 11: 1126–1136. <https://doi.org/10.1111/j.1466-8238.2012.00768.x>.

- Gil-Tapetado, D., F. J. Cabrero-Sañudo, C. Polidori, J. F. Gómez, and J. L. Nieves-Aldrey. 2021. “Climate as a Possible Driver of Gall Morphology in the Chestnut Pest *Dryocosmus kuriphilus* Across Spanish Invaded Areas.” *Bulletin of Entomological Research* 111, no. 2: 160–173. <https://doi.org/10.1017/S0007485320000450>.

- Gil-Tapetado, D., F. Castedo-Dorado, M. J. Lombardero, J. Martel, and P. Álvarez-Álvarez. 2020. “Spatial Propagation and Patterns of Abundance of *Dryocosmus kuriphilus* Throughout an Invaded Region.” *Journal of Applied Entomology* 145: 1–16. <https://doi.org/10.1111/jen.12836>.

- Gil-Tapetado, D., F. Castedo-Dorado, J. L. Nieves-Aldrey, and M. J. Lombardero. 2021. “Gall Size of *Dryocosmus kuriphilus* Limits Down-Regulation by Native Parasitoids.” *Biological Invasions* 23, no. 4: 1157–1174. <https://doi.org/10.1007/s10530-020-02427-x>.

- Gil-Tapetado, D., J. F. Gómez, F. J. Cabrero-Sañudo, and J. L. Nieves-Aldrey. 2018. “Distribution and Dispersal of the Invasive Asian Chestnut

- Gall Wasp, *Dryocosmus kuriphilus* (Hymenoptera: Cynipidae), Across the Heterogeneous Landscape of the Iberian Peninsula." *EJE* 115, no. 1: 575–586. <https://doi.org/10.14411/eje.2018.055>.
- Hijmans, R. J., S. Phillips, J. Leathwick, and J. Elith. 2024. *dismo: Species Distribution Modeling*. <https://doi.org/10.32614/CRAN.package.dismo>.
- Hill, M. P., B. Gallardo, and J. S. Terblanche. 2017. "A Global Assessment of Climatic Niche Shifts and Human Influence in Insect Invasions." *Global Ecology and Biogeography* 26, no. 6: 679–689. <https://doi.org/10.1111/geb.12578>.
- Hirzel, A. H., G. Le Lay, V. Helfer, C. Randin, and A. Guisan. 2006. "Evaluating the Ability of Habitat Suitability Models to Predict Species Presences." *Ecological Modelling* 199, no. 2: 142–152. <https://doi.org/10.1016/j.ecolmodel.2006.05.017>.
- Hulme, P. E. 2017. "Climate Change and Biological Invasions: Evidence, Expectations, and Response Options." *Biological Reviews* 92, no. 3: 1297–1313. <https://doi.org/10.1111/brv.12282>.
- Karger, D. N., O. Conrad, J. Böhner, et al. 2017. "Climatologies at High Resolution for the Earth's Land Surface Areas." *Scientific Data* 4, no. 1: 170122. <https://doi.org/10.1038/sdata.2017.122>.
- Leroy, B., R. Delsol, B. Hugueny, et al. 2018. "Without Quality Presence–Absence Data, Discrimination Metrics Such as TSS Can Be Misleading Measures of Model Performance." *Journal of Biogeography* 45, no. 9: 1994–2002. <https://doi.org/10.1111/jbi.13402>.
- Linders, T. E. W., U. Schaffner, R. Eschen, et al. 2019. "Direct and Indirect Effects of Invasive Species: Biodiversity Loss Is a Major Mechanism by Which an Invasive Tree Affects Ecosystem Functioning." *Journal of Ecology* 107, no. 6: 2660–2672. <https://doi.org/10.1111/1365-2745.13268>.
- Liu, X.-Y., J. Wu, and Z.-H. Zhou. 2009. "Exploratory Undersampling for Class-Imbalance Learning." *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 39, no. 2: 539–550. <https://doi.org/10.1109/TSMCB.2008.2007853>.
- Lobo, J. M., A. Jiménez-Valverde, and R. Real. 2008. "AUC: A Misleading Measure of the Performance of Predictive Distribution Models." *Global Ecology and Biogeography* 17, no. 2: 145–151. <https://doi.org/10.1111/j.1466-8238.2007.00358.x>.
- Lobo Santos, A., S. A. P. Santos, P. A. Casquero, and A. Bento. 2024. "Lifecycle of *Dryocosmus kuriphilus* Yasumatsu and Diversity and Importance of the Native Parasitoid Community Recruited in the Northern Region of Portugal." *Insects* 15, no. 1: 22. <https://doi.org/10.3390/insects15010022>.
- Mainali, K. P., D. L. Warren, K. Dhileepan, et al. 2015. "Projecting Future Expansion of Invasive Species: Comparing and Improving Methodologies for Species Distribution Modeling." *Global Change Biology* 21, no. 12: 4464–4480. <https://doi.org/10.1111/gcb.13038>.
- Mapes, C. C., G. P. Setliff, R. S. Courtney, and K. Bothur. 2020. "Range Expansion and Dispersal of the Asian Chestnut Gall Wasp *Dryocosmus kuriphilus* Yasumatsu (Hymenoptera: Cynipidae) in the Mid-Atlantic States." *Transactions of the American Entomological Society* 146, no. 3: 503–519. <https://doi.org/10.3157/061.146.0304>.
- Murakami, Y., and Y. Gyoutoku. 1995. "A Delayed Increase in the Population of an Imported Parasitoid, *Torymus (Syntomaspis) Sinensis* (Hymenoptera: Torymidae) in Kumamoto, Southwestern Japan." *Applied Entomology and Zoology* 30, no. 1: 215–224. <https://doi.org/10.1303/aez.30.215>.
- Nguyen, D., and B. Leung. 2022. "How Well Do Species Distribution Models Predict Occurrences in Exotic Ranges?" *Global Ecology and Biogeography* 31, no. 6: 1051–1065. <https://doi.org/10.1111/geb.13482>.
- Nieves-Aldrey, J. L., D. Gil-Tapetado, O. Gavira, et al. 2019. "*Torymus sinensis* Kamijo, a Biocontrol Agent Against the Invasive Chestnut Gall Wasp *Dryocosmus kuriphilus* Yasumatsu in Spain: Its Natural Dispersal From France and the First Data on Establishment After Experimental Releases." *Forestry Systems* 28, no. 1: e001. <https://doi.org/10.5424/fs/2019281-14361>.
- Panzavolta, T., U. Bernardo, M. Bracalini, et al. 2013. "Native Parasitoids Associated With *Dryocosmus kuriphilus* in Tuscany, Italy." *Bulletin of Insectology* 66: 195–201.
- Pérez Otero, R., D. Crespo, and J. P. Mansilla Vázquez. 2017. "*Dryocosmus kuriphilus* Yasumatsu, 1951 (Hymenoptera: Cynipidae) in Galicia (NW Spain): Pest Dispersion, Associated Parasitoids and First Biological Control Attempts." *Archivos Entomológicos* 17: 439–448.
- Pérez Otero, R., and J. P. Mansilla Vázquez. 2014. "El Cinípido del Castaño *Dryocosmus kuriphilus* Yasumatsu, 1951 llega a Galicia (NO de la Península Ibérica)." *Archivos Entomológicos* 12: 33–36.
- Peterson, A. T., and Y. Nakazawa. 2008. "Environmental Data Sets Matter in Ecological Niche Modelling: An Example With *Solenopsis invicta* and *Solenopsis richteri*." *Global Ecology and Biogeography* 17, no. 1: 135–144. <https://doi.org/10.1111/j.1466-8238.2007.00347.x>.
- Phillips, S. J., R. P. Anderson, and R. E. Schapire. 2006. "Maximum Entropy Modeling of Species Geographic Distributions." *Ecological Modelling* 190, no. 3: 231–259. <https://doi.org/10.1016/j.ecolmodel.2005.03.026>.
- Pyšek, P., P. E. Hulme, D. Simberloff, et al. 2020. "Scientists' Warning on Invasive Alien Species." *Biological Reviews* 95, no. 6: 1511–1534. <https://doi.org/10.1111/brv.12627>.
- Quacchia, A., C. Ferracini, J. A. Nicholls, et al. 2013. "Chalcid Parasitoid Community Associated With the Invading Pest *Dryocosmus kuriphilus* in North-Western Italy." *Insect Conservation and Diversity* 6, no. 2: 114–123. <https://doi.org/10.1111/j.1752-4598.2012.00192.x>.
- Quacchia, A., S. Moriya, G. Bosio, I. Scapin, and A. Alma. 2008. "Rearing, Release and Settlement Prospect in Italy of *Torymus sinensis*, the Biological Control Agent of the Chestnut Gall Wasp *Dryocosmus kuriphilus*." *BioControl* 53, no. 6: 829–839. <https://doi.org/10.1007/s10526-007-9139-4>.
- Quinto, J., M. E. Wong, J. R. Boyero, J. M. Vela, and M. Aguirrebengoa. 2021. "Population Dynamics and Tree Damage of the Invasive Chestnut Gall Wasp, *Dryocosmus kuriphilus*, in Its Southernmost European Distributional Range." *Insects* 12, no. 10: 10. <https://doi.org/10.3390/insects12100900>.
- R Core Team. 2025. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Sartor, C., F. Dini, D. Torello Marinoni, et al. 2015. "Impact of the Asian Wasp *Dryocosmus kuriphilus* (Yasumatsu) on Cultivated Chestnut: Yield Loss and Cultivar Susceptibility." *Scientia Horticulturae* 197: 454–460. <https://doi.org/10.1016/j.scienta.2015.10.004>.
- Sillero, N., S. Arenas-Castro, U. Enriquez-Urzelai, et al. 2021. "Want to Model a Species Niche? A Step-By-Step Guideline on Correlative Ecological Niche Modelling." *Ecological Modelling* 456: 109671. <https://doi.org/10.1016/j.ecolmodel.2021.109671>.
- Sofaer, H. R., C. S. Jarnevich, I. S. Pearse, et al. 2019. "Development and Delivery of Species Distribution Models to Inform Decision-Making." *Bioscience* 69, no. 7: 544–557. <https://doi.org/10.1093/biosci/biz045>.
- Srivastava, V., V. Lafond, and V. C. Griess. 2019. "Species Distribution Models (SDM): Applications, Benefits and Challenges in Invasive Species Management." *CABI Reviews* 2019: 1–13. <https://doi.org/10.1079/PAVSNNR201914020>.
- Sutton, G. F., and G. D. Martin. 2022. "Testing MaxEnt Model Performance in a Novel Geographic Region Using an Intentionally Introduced Insect." *Ecological Modelling* 473: 110139. <https://doi.org/10.1016/j.ecolmodel.2022.110139>.
- Thuiller, W., D. Georges, M. Gueguen, et al. 2023. *biomod2: Ensemble Platform for Species Distribution Modeling*. <https://CRAN.R-project.org/package=biomod2>.
- Václavík, T., and R. K. Meentemeyer. 2012. "Equilibrium or Not? Modelling Potential Distribution of Invasive Species in Different Stages

of Invasion.” *Diversity and Distributions* 18, no. 1: 73–83. <https://doi.org/10.1111/j.1472-4642.2011.00854.x>.

Valavi, R., G. Guillerá-Arroita, J. J. Lahoz-Monfort, and J. Elith. 2022. “Predictive Performance of Presence-Only Species Distribution Models: A Benchmark Study With Reproducible Code.” *Ecological Monographs* 92, no. 1: e01486. <https://doi.org/10.1002/ecm.1486>.

VanDerWal, J., L. P. Shoo, C. Graham, and S. E. Williams. 2009. “Selecting Pseudo-Absence Data for Presence-Only Distribution Modeling: How Far Should You Stray From What You Know?” *Ecological Modelling* 220, no. 4: 589–594. <https://doi.org/10.1016/j.ecolmodel.2008.11.010>.

Walther, G.-R., A. Roques, P. E. Hulme, et al. 2009. “Alien Species in a Warmer World: Risks and Opportunities.” *Trends in Ecology & Evolution* 24, no. 12: 686–693. <https://doi.org/10.1016/j.tree.2009.06.008>.

Wen, X., G. Fang, S. Chai, et al. 2024. “Can Ecological Niche Models Be Used to Accurately Predict the Distribution of Invasive Insects? A Case Study of *Hyphantria cunea* in China.” *Ecology and Evolution* 14, no. 3: e11159. <https://doi.org/10.1002/ece3.11159>.

Xunta de Galicia. 2020. *Plan de Actuacións de Loita Contra as Pragmas e Enfermidades do Castiñeiro no 2020. Solta do Parasitoide (Torymus sinensis) e Actuacións Relativas ó Control Biolóxico de Cryphonectria parasitica*. Consellería de Medio Rural. Xunta de Galicia.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Table S1:** Root mean square error (RMSE) and associated standard deviation (SD) values for each variable spatially interpolated using Universal Kriging, with station elevation included as a covariate. Of the 140 meteorological stations available, 80% were randomly selected for model training, while the remaining 20% were used for independent validation. **Figure S1:** Spatial differences between climatic variables derived from the CHELSA and MeteoGalicia datasets for the year 2014. Each panel represents the difference calculated as CHELSA—MeteoGalicia for one bioclimatic or wind variable. Positive values indicate higher values in CHELSA, whereas negative values indicate higher values in MeteoGalicia. Colour scales are independent for each variable. Spatial patterns of differences were visually consistent across all analysed years (2014–2018); therefore, only one representative year is shown. **Figure S2:** Mean variable importance of environmental predictors used in species distribution models of *Dryocosmus kuriphilus*. Environmental predictors include bioclimatic variables (bio1–bio19) and monthly wind speed (sfcWind). Only models meeting the performance threshold used for ensemble construction (Boyce index > 0.7) were shown.