ORIGINAL PAPER

Check for updates

A two-step species distribution modeling to disentangle the effect of habitat and bioclimatic covariates on *Psacothea hilaris*, a potentially invasive species

Enrico Ruzzier · **Daniela Lupi** · **Pietro Tirozzi** · **Olivia Dondina** ·
Valerio Orioli · **Costanza Jucker** · **Luciano Bani**

Abstract Integrating host plants in distribution modeling of phytophagous species and disentangling the effect of habitat and bioclimatic variables are key aspects to produce reliable predictions when the aim is to identify suitable areas outside species' native range. To this aim, we implemented a framework of Species Distribution Model aimed at predicting potential suitable areas of establishment for the beetle *Psacothea hilaris* across the world. Since habitat (including host plants) and bioclimatic variables affect species distribution according to processes acting at different geographical scales, we modeled these variables separately. For the species native range, we fitted a habitat (HSM) and a bioclimatic (BSM) suitability model calibrated on a local and a large scale, respectively; the overall suitability map was obtained as the spatial product of HSM and BSM projection maps. ROC, TSS and Cohen's Kappa obtained in validation confirmed a good predictive performance of modeling framework. Within HSM, host plants played a substantial effect on species presence probability, while among bioclimatic variables, precipitation of the warmer quarter and isothermality were the most important. Native HSM and BSM models were used to realize an overall suitability map at world scale. At global scale, many areas resulted suitable for habitat, some for bioclimate, and few for both conditions; indeed, if the species would not be able to modify its bioclimatic niche, it might not be considered a major invasive species. However, the high rate of range expansion documented for *P. hilaris* in Northern Italy, a poorly suitable bioclimatic area, suggests a plasticity of the species that requires increasing the level of attention to its invasive potential.

Keywords Bioclimatic suitability model · Cerambycidae · Habitat suitability model · Host plants · Species distribution model · Wood borer

E. Ruzzier (✉)
Department of Science, Roma Tre University, viale G. Marconi 446, Rome, Italy
e-mail: enrico.ruzzier@uniroma3.it

E. Ruzzier · O. Dondina · L. Bani
NBFC, National Biodiversity Future Center, Palermo, Italy

D. Lupi · C. Jucker
Department of Food, Environmental and Nutritional Sciences, University of Milan, Milan, Italy

P. Tirozzi · O. Dondina · V. Orioli · L. Bani
Department of Earth and Environmental Sciences, University of Milano-Bicocca, Milan, Italy

Introduction

Species distribution modeling (SDM) is a valuable tool for predicting the potential distribution of

invasive species across space and time (Srivastava et al. 2021). This methodology has often been applied in predicting the potential spatial distribution of invasive insects worldwide, especially through the usage of bioclimatic factors as the main explanatory variables (e.g., Dupin et al. 2011; Iannella et al. 2020; Krishnankutty et al. 2020; Li et al. 2021; Lu et al. 2022). In more recent times, host plants have been integrated as a relevant factor in determining invasive species distribution and introduction patterns, especially in wood-borer beetles (e.g., Dang et al. 2021). Cerambycidae (Coleoptera: Chrysomeloidea), together with Buprestidae (Coleoptera: Buprestoidea) and Curculionidae Scolytinae (Coleoptera: Curculionoidea), are among the most important beetle groups of phytosanitary interest worldwide (Haack et al. 2014; Wu et al. 2017; MacQuarrie et al. 2020; Ruzzier et al. 2023a). Many cerambycid species have been accidentally introduced into various parts of the world through the international trade in plants and wood products (e.g., Di Iorio 2004; Haack 2006; Cocquempot and Lindelöw 2010; Sopow et al. 2015; Rassati et al. 2016; Ruzzier et al. 2020a; Seidel et al. 2021). Some of these have become successfully established in the new environment, becoming a serious threat to both native ecosystems and human activities (e.g., Eyre and Haack 2017; Sarto i Monteys and Torras i Tutusaus 2018; Lee et al. 2021). However, despite the ecological and economic relevance of some of these taxa, SDM has so far been relatively underutilized with the invasive Cerambycidae, and studies were limited to the sole species *Anoplophora glabripennis* (Motschulsky, 1854) (Peterson and Scachetti-Pereira 2004; Shatz et al. 2013; Pedlar et al. 2020; Byeon et al. 2021; Zhou et al. 2021).

Psacothea hilaris (Pascoe, 1858) (Coleoptera, Cerambycidae, Lamiinae), also known as yellow-spotted longhorn beetle, is a species native to Eastern Asia (Danilevsky 2020) that has been repeatedly intercepted or introduced in North America and Europe (Allen and Humble 2002; EPPO 2013). Multiple records of this species in different parts of the world follow common trends of introduction to many other non-native organisms worldwide (Fenn-Moltu et al. 2023) and are the result of extensive detection and interception activities (Turner et al. 2021). Although several individuals were recorded in nature in France, Germany and the UK (EPPO

2008, 2013; INPN 2021), so far *P. hilaris* has become established only in Lombardy (Northern Italy), where it has been a constant presence since 2005 (Jucker et al. 2006; Lupi et al. 2023). The species is one of the several non-native beetles collected in natural environments in Italy over the last two decades (e.g., Conti and Raspi 2007; Faccoli et al. 2009; Nardi et al. 2015; Toma et al. 2017; Ruzzier and Colla 2019; Sparacio et al. 2020; Ruzzier et al. 2020b, 2021, 2022, 2023b, 2023c; Marchioro et al. 2022).

Psacothea hilaris develops primarily on figs (particularly on *Ficus carica*, *F. erecta,* and *F. macrocarpa*), mulberries (particularly on *Morus alba* and *M. indica*) (Moraceae), and Japanese aralia (*Fatsia japonica*, Araliaceae) (Iba 1980), and the trophic activity of the larva, which feeds while tunneling through the wood, causes serious damage to the trunk and branches of the host plant, leading to its dieback and death (Lupi et al. 2013). Infestations by *P. hilaris* are usually difficult to detect and manage due to the cryptic lifestyle of the larvae and are usually visible upon the onset of severe stress on plants or the emergence of adults (Lupi et al. 2013). Due to its host preference and biology, *P. hilaris* is considered a serious pest on wild, ornamental, and cultivated plants, both in its native and non-native range (Fukaya and Honda 1992; Iba 1993; Lupi et al. 2013).

At present, the economic and ecological damage caused by *P. hilaris* are of minor interest in Lombardy (Northern Italy) because the beetle affects plants which are primarily cultivated for ornamental purposes in this area; however, the diffusion of the species into other parts of the Italian peninsula, as well as into other mediterranean countries, might have important repercussions on the productive activities associated with *F. carica* and *M. alba*. *Ficus carica* is a widespread species growing mostly in warm, dry climates and its fruits are an important product for dry and fresh consumption (e.g., Crisosto et al. 2011; Badgujar et al. 2014). This plant and its associated economy are already threatened by multiple beetle pests (e.g., Akşit et al. 2005; Ismail et al. 2009; Tinoco et al. 2010; López-Martínez et al. 2015; Gaaliche et al. 2018), including the invasives *Cryphalus dilutus* Eichhoff 1878 (Faccoli et al. 2016) and *Aclees taiwanensis* Kôno, 1933 (Coleoptera: Curculionidae) (Farina et al. 2020) in the Mediterranean basin. The arrival of another pest (i.e., *P. hilaris*) could bring fig cultivation to a deep crisis as well as jeopardizing the conservation of local historical cultivars and

endemic varieties. Furthermore, figs are a major crop in other countries (e.g., Lianju et al. 2003; Crisosto et al. 2011; Mendoza-Castillo et al. 2017; Shamin-Shazwan et al. 2019) and a possible invasion of *P. hilaris* is a risk not to be underestimated. *Morus alba*, besides its ornamental uses and the production of berries, is mostly used for sericulture both in Italy and abroad (e.g., Urbanek Krajnc et al. 2019; Tzenov et al. 2022). Although no direct damage has yet been observed on *M. alba* in northern Italy, infestations of *P. hilaris* could make it particularly difficult to cultivate mulberry trees for domestic silk moth *Bombyx mori* sustenance, especially because the use of insecticides on mulberries is avoided to prevent silkworm caterpillars' intoxication (Iba 1993).

A recent study by Lupi et al. (2023) has shown that the Italian population of *P. hilaris* is undergoing a range expansion (sensu Richardson et al. 2011); consequently, with a total lack of containment strategies (Lupi et al. 2023), we could thus expect a further spreading of the species into other European and circum-Mediterranean countries. In order to investigate species and environmental relationships useful to identify the possible suitable areas for species populations establishment and range expansion, the present contribution aims to:

(a) Develop a suitability map for *P. hilaris* in its putative native range by means of a two-step SDM approach, evaluating the effects of bioclimatic variables, and those of habitat variables (digital elevation model, land cover, and host plants' distribution);
(b) Predict possible suitable areas for *P. hilaris* establishment at a global scale based on the SDM fitted with data from the native range;
(c) Assess the capability of the SDM developed for the native range to predict the distribution of *P. hilaris* in the invaded range (northern Italy).

Methods

Psacothea hilaris occurrence data

To build a dataset as exhaustive as possible of the actual distribution of *P. hilaris* both in its putative native and non-native range, species records were collected through extensive data mining in online data repositories, namely the Global Biodiversity Information Facility (GBIF, https://www.gbif.org/; doi:https://doi.org/10.15468/dl.yzzrrt) and iNaturalist (https://www.inaturalist.org/). The data deriving from field surveys or direct observation were included in the dataset only if the reliability of the record could be validated through the collection of the specimen, the identification of clear symptoms of infestation, and photographs in which the species was clearly identifiable. With specific regard to Italy, to obtain a representative sample of the current distribution, multiple records of *P. hilaris* were obtained directly through field surveys and direct observations (see Lupi et al. 2023).

The occurrence data used in the analyses are updated to December 31st, 2023. All *Psacothea hilaris* occurrence data are available as Supplementary material (psacothea.hilaris_243101. xlsx). We selected species occurrences with a coordinate uncertainty lower than 500 m to associate the occurrence with the habitat data derived from the land cover cartography of which the resolution is 0.00833° (corresponding to 1.0 km at the equator; see below for details). To reduce possible strong interdependence among occurrence data, we performed a spatial thinning using the *spThin* package (Aiello-Lammens et al. 2015) in R version 4.1.2 (R Core Team 2021). As a thinning parameter, i.e., the maximum distance among occurrence points we wanted to retain, we set 1 km (to conserve the same magnitude of coordinate uncertainty).

Environmental data

The selection of the environmental layers used for building the species distribution model was based on the following assumptions:

(a) Species range is bounded by bioclimatic conditions and ecological barriers (scenopoetic variables) from regional to continental scale (see Hortal et al. 2010; Pearson and Dawson 2003);
(b) Species occurrence from site to local scale (sensu Hortal et al. 2010) is affected by habitat conditions (Pearson and Dawson 2003), where elevation and land cover types can be considered as proxies of bionomic variables. Moreover, for phytophagous insects, the habitat suitability can

be strongly affected by the presence of host plants (bionomic variables by definition); although the distribution of host plants may be influenced by climatic conditions, the effective range of plants is not necessarily definable in bioclimatic terms since ecological barriers can limit it within a rather homogeneous bioclimatic context.

Bioclimatic data (19 strata) and the elevation (i.e., the digital elevation model, DEM) were downloaded from WorldClim 2.1 (available at: http://www.worldclim.com/version2), average for the years 1970–2000, with a 0.00833° resolution.

Land cover layers were obtained from the Global 1-km Consensus Land Cover (available at http://www.earthenv.org; Tuanmu and Jetz 2014). The dataset contains 12 layers (evergreen/deciduous needleleaf trees [EDNT], evergreen broadleaf trees [EBT], deciduous broadleaf trees [DBT], mixed/other trees [MOT], shrubs [SHR], herbaceous vegetation [HERB], cultivated and managed vegetation [CMV], regularly flooded vegetation [RFV], urban/built-up [BU], snow/ice [SI], barren lands [BR], and open water [OW]), each of which provides consensus information on the prevalence of one land-cover class, expressed as a percentage in a pixel having a resolution of 30 arc-second (i.e., 0.00833° resolution).

To construct the range map of each host plant at world scale, we used the species occurrence data downloaded from GBIF, setting the coordinate uncertainty lower than 10 km (since the aim is to identify the range of host species by means of an interpolation procedure, particular accuracy is not required in the georeferencing of the occurrences). First, to obtain a reliable range map for each plant species, we used a kernel density function on occurrence data with an interpolation radius of 1° in QGIS 3.32 (QGIS Development Team 2023), with a resolution of 0.00833°. Before producing the kernel density maps, in order to minimize possible spatial sampling bias across geographic areas, we performed a spatial thinning of host plant data using the *spThin* package (Aiello-Lammens et al. 2015) in R version 4.1.2 (R Core Team 2021). As a thinning parameter, i.e., the maximum distance among plant occurrence points we wanted to retain, we set 10 km (to conserve the same magnitude of coordinate uncertainty). Six plant species, namely *Ficus carica*, *F. erecta*, *F. macrocarpa*, *Morus alba*, *M. indica*,

and *Fatsia japonica*, were considered as the main host plants, intended as those with consolidated records in literature and websites (no sporadic reports). Subsequently, assuming that host plants could be vicariant across the *P. hilaris* distribution range (native and non-native), we combined host plant layers, obtaining a unique layer for all the host plants represented by the sum of the densities of each of the seven individual layers at world scale. Virtually all the considered host plants have a global distribution in terms of native area and introduction areas considered as a whole, except for the *F. erecta*, whose distribution is mostly limited to its native range. The rationale of producing a world scale host plants layer relied on the capability of *P. hilaris* to use as reproductive hosts both plant taxa found in its native range and translocated worldwide, as well as congeneric vicariant plants to those occurring in *P. hilaris* native range. This condition justifies the creation of a global host plants layer needed to obtain a world scale prediction of habitat suitability for *P. hilaris*. As the last step, in order not to overestimate the contribution of the host plant layers, i.e., to consider the effect of the host plants only where they have more chance to be actually present, we opted for multiplying the logarithm of the resulting host plants density layer for the overall cover given by EBT, DBT, MOT (layers of broadleaved or mixed arboreal habit) and SHR (layer of shrub habit) (i.e., logarithm of host plant density multiplied for broadleaved and mixed trees and shrub fractional cover, HP.bl).

Species distribution modeling framework

To identify the contribution of the environmental variables in explaining the distribution of *P. hilaris* within its native range, we built a SDM using an ensemble modeling approach (Araújo and New 2007; Guisan et al. 2017). The ensemble approach is widely used to evaluate the contribution of environmental variables in explaining the spatial distribution of species (Guisan et al. 2017), to assess spatial connectivity (Dondina et al. 2020), to define possible future scenarios for species and habitat (Ahmad et al. 2020), or in systematic conservation planning (Meller et al. 2014). The rationale at the base of ensemble modeling is to combine the output from multiple models to obtain reliable estimates of covariates effects with a significant reduction of

bias and variance, minimizing the effect of underfit/overfit predictions through a penalization of poor-performing models based on evaluation metrics (Araújo and New 2007). This aspect is particularly important when making occurrence predictions outside the calibration area of a model (i.e., outside the area from which the data used for modeling comes) (Guisan et al. 2017). Using the ensemble modeling approach, the combined output should tendentially have higher explanatory and predictive performances compared to a model obtained from a single algorithm (Breiner et al. 2015; Hao et al. 2019; but see Früh et al. 2018; Hao et al. 2020, Valavi et al. 2022).

One of the most important aspects to consider when modeling presence-only data is the creation of Pseudo-Absences (PAs) or background data (Barbet-Massin et al 2012). Selecting PAs strictly within the native range of a species does not allow evaluating the effective contribution of bioclimatic variables to species distribution, as they vary chiefly on a large spatial scale. Consequently, the impossibility to disentangle the effects of these bioclimatic conditions would impede the assessment of the importance of those crucial variables that constrain the extent of species' native range itself, and would hinder the ability to carry out a correct identification of potentially suitable bioclimatic areas outside the native range. For this reason, when working on a large scale (e.g., continental or global), in order to verify whether and quantify how bioclimatic variables play an effective role in limiting/defining the range of the species, it is necessary to collect PAs outside the known species' native range (see Guisan et al. 2017; Adde et al. 2023). On the other hand, land cover, DEM, and host plants, given their narrow-scale variability, are variables acting on a local scale, thus predominantly affecting habitat suitability. In this case, to assess the effect of these covariates, PAs should be identified as close as possible to presence points, i.e., strictly within the known species' range (Fournier et al. 2017; Mateo et al.2019; Adde et al. 2023). This PAs selection procedure limits the possible confounding effect of bioclimatic variables on the estimates of habitat variables effect.

To verify our assumption about the degree of spatial variability of habitat and bioclimatic covariates we calculated the Moran's I for each variable across different spatial extents (0.00833°,

0.0833°, 0.833°) using the *geocmeans* package (Gelb and Apparicio 2021) in R (Table S1). As expected, the Moran's I values resulted always lower for habitat respect to bioclimatic variables.

For the above-mentioned considerations, we developed a Habitat Suitability Model (HSM) over a narrow geographic range restricted to the *P. hilaris* putative native range. In addition, we developed a Bioclimatic Suitability Model (BSM), testing the effect of bioclimatic variables over a large geographical area centered on *P. hilaris'* putative native range. As the last step, we combined the two suitability models to obtain the overall suitability model accounting for both bioclimatic and habitat variables.

Given the above premises, we subjectively set the latitudinal extent between 20° and 45° N and the longitudinal extent between 115° and 150° E for the HSM, and between 0° and 80° N and 90° E and 160° E for the BSM. The ranges for the HSM allowed to select an extent strictly limited to the range of the species; thus, remaining in a tendentially homogeneous bioclimatic context respect to a wider extent, especially across latitude, the effect of habitat variables can be assessed more effectively. Conversely, for the BSM, we selected a larger extent (about double in magnitude respect to the native range) in order to assess the effect of bioclimatic conditions subsisting beyond *P. hilaris* native range (i.e. the conditions the species could experience in an area of potential introduction). Since we assume that bioclimatic conditions vary mostly along the latitude, we set a greater difference in latitudinal than in longitudinal limits passing from HSM to BSM extents.

Species distribution model building

Before building the HSM and BSM for the *P. hilaris* within its native range we produced a check of environmental covariates to avoid collinearity issues. Indeed, since many of the bioclimatic variables are strongly correlated, we performed a preselection of bioclimatic covariates, removing those strongly correlated, i.e., with an absolute value of the Pearson's correlation coefficient higher than 0.7 (Dormann et al. 2013) using the "vifcor" function of the *usdm* package (Naimi et al. 2014) in R. The bioclimatic covariates selection was produced by investigating

the possible collinearity within the BSM extent. We preliminarily log-transformed all the bioclimatic variables related to precipitation, to reduce the right-skewed distribution due to extreme values of rainfall (BIO12 to BIO19). The same procedure aimed at evaluating possible variable correlation problems was carried out for all the covariates included in HSM (land cover, DEM, and host plant layers) within the HSM area. In addition, for land cover data, because of their perfect complementarity, we excluded those with frequency within the native range lower than 5%.

To reduce the spatial autocorrelation of *P. hilaris* occurrence data, we performed a spatial thinning (e.g., Steen et al. 2021) rarifying presence points falling within the same raster cell of the bioclimatic and habitat layers; this operation was performed using the "ensemble.spatialThin" function included in the *BiodiversityR* package (Kindt and Coe 2005), setting 1.0 km as thinning threshold (that fits for a raster resolution of 0.00833° since 1.0 km corresponds to just over 0.00833° at the equator) and performing 10 runs.

The modeling procedure for the HSM and BSM was performed using the *biomod2* package version 4.2–4 (Thuiller et al. 2021). R code is available upon request to the authors.

As modeling algorithms, we used the Generalized Linear Model (GLM), Generalized Boosting Model (GBM), Random Forest (RF), and Maximum Entropy (MAXNET). For each model algorithm, as models' options we used the default settings, except for the GLM where we used the "polynomial" option. Since the different algorithms used in the ensemble model framework require a different number of PAs for their optimal performance (Barbet-Massin et al. 2012), we were forced to use a trade-off between a PAs inflation and PAs paucity, high or low number of PAs with respect to those needed by the algorithm, respectively. We decided to set the number of PAs to 4 times the number of occurrences. However, since the number of PAs was relatively low, and since some algorithms require a higher number of PAs, some models might become ineffective (i.e., fail the validation test and, consequently, be discharged from the analyses; see ahead). As a solution, we opted to use a large number of sets of PAs according to the capability of the HSM and BSM of obtaining a substantial number of effective models (i.e., models passing the selection of validation metrics; see below for details). Thus, we decided to use 20 sets of PAs for HSM and 5 for BSM (see the Results section for detailed information). The different number of sets used in the modeling framework depends on the nature of the considered covariates (habitat *vs* bioclimatic variables). Within the narrower extent used for developing the HSM, we opted for a "disk" strategy to randomly select PAs within a buffer ranging from 2.5 km (to avoid putting pseudo-absences close to occurrences, i.e. in inside possible suitable habitats) to 1000 km (to limit the possible strong effect of bioclimatic covariates) respect species occurrence points. For the larger extent within which we developed the BSM, we opted for a completely PAs random selection in the whole area excluding only seas. Since no external data were available to evaluate the models' accuracy, a 5 cross-validations data-splitting procedure was carried out using 80% of the data as the training set with the remaining 20% as the validation set. Overall, we obtained 400 models (4 algorithms×20 PAs sets×5 cross-validation runs) for the HSM and 100 models (4 algorithms×5 PAs sets×5 cross-validation runs) for the BSM. Each model was validated by three different metrics: the Area Under the Receiver Operating Characteristic (ROC) curve (AUC; Hanley and McNeil 1982), True Skill Statistic (TSS; Allouche et al. 2006), and Cohen's Kappa (Choen 1960). Only those models overcoming the following thresholds were included in the ensemble model (if under the threshold, the model was discharged, i.e., not considered in the ensemble model procedure): 0.9 for ROC, 0.8 for TSS, and 0.8 for Cohen's Kappa (Zhang et al. 2015). The ensemble model was computed using the EMwmean function (probabilities from the selected models weighted according to their evaluation scores obtained when building the model). To account for each single model performance used in building the ensemble model, we weighted its contribution by setting a "proportional" option decay. The importance of each variable included in the ensemble model was estimated using a permutation procedure based on three runs. The same framework was adopted for both HSM and BSM. Finally, the performance of the HSM and BSM ensemble models was evaluated through the Boyce index (Boyce et al. 2002; Hirzel et al. 2006) by means of the *modEvA* package (Barbosa et al. 2013), using Spearman correlation coefficient and bin width of 0.2.

For HSM and BSM, we produced two separate maps that were subsequently used for the realization of the overall suitability map. Since both predictions of HSM and BSM produced two maps of an environmental suitability in terms of probability of species presence, the overall suitability map (i.e., the overall probability of species presence) can be calculated as the spatial product of HSM and BSM maps (see Adde et al. 2023). The rationale of this procedure relies on the respective independence of HSM and BSM probabilities; this condition is realized since PAs sets for HSM and BSM were randomly extracted within two geographical areas of different extents in order to better assess the effect of bioclimatic variables in the BSM and reduce their effect as confounding variables in building HSM (see Adde et al. 2023). The prediction performance of the overall suitability map was evaluated by means of the Boyce index.

HSM and BSM for the native range were used to forecast the suitable areas for *P. hilaris* at a global scale by means of the same approach used for the native range, obtaining an overall global suitability map. Finally, we produced a suitability map tailored for *P. hilaris* in the areas of introduction, i.e., Northern Italy, for which we performed a validation of the prediction performance applying the Boyce index on presence records available for the invaded range.

For the global and Italian-scale projection, a validation was made using the Multivariate Environmental Similarity Surfaces index (MESS; Elith et al. 2010), implemented in the *dismo* package (Hijmas et al. 2023).

Results

Data mining and data collection from online resources, considering the uncertainty of occurrence coordinates lower than 500 m, produced 535 occurrences of *P. hilaris*, 266 of which were in the putative native area. Thinning procedures on native occurrences reduced the number of occurrences to be used in the modeling framework to 194.

After performing covariate selection to overcome collinearity issues, for the HSM, we had all the variables pertaining to the land cover layers (except CMV), DEM, and host plants layer. The covariate selection for the BSM retained 5 over the 19 initial variables, namely BIO2 (mean diurnal range), BIO3 (isothermality), BIO5 (max temperature of warmest month), BIO8 (mean temperature of wettest quarter), BIO15 (precipitation seasonality), and BIO18 (precipitation of warmest quarter). This selection can be considered satisfactory in term of representativeness of bioclimatic condition, for both scales of analysis since the set of variables includes three variables related to temperatures, and two to precipitations.

Among the 400 models (runs) developed for the HSM, 294 (73.5%) passed the ROC, 62 (15.5%) the TSS, and 53 (13.3%) the Cohen's Kappa validation thresholds. Among the 100 for the BSM, all passed the ROC and TSS validation thresholds, while 66 (66.0%) passed the Cohen's Kappa thresholds. The models passed the above-mentioned thresholds contributed to the final HSM and BSM ensemble models developed for the species native range. The Boyce index obtained in validation for HSM and BSM ensemble models returned 0.985 and 0.966, respectively (See Figs. S1e and S2 for Boyce graphs and Tables S4 and S5 for Boyce index values obtained from HSM and BSM maps, respectively). The ensemble modeling for the HSM and the BSM allowed us to establish the covariates' importance (Tables S6 and S7) and their effect in affecting *P. hilaris* presence probability, separately. For HSM the most important variables were host plants (variable importance: 0.552), and BU (0.438). Host plants and BU showed a positive effect on species presence (Fig. S3). BSM was substantially affected by BIO18 (variable importance: 0.618) and to a lesser extent by BIO3 (0.283) (Fig. S4). The outputs of the ensemble models allowed the spatial prediction for HSM (Figs. 1, S5 map of HSM confidence intervals) and BSM (Figs. 2, S6 map of BSM confidence intervals) for the *P. hilaris* putative native range, separately.

The spatial product of maps of the species' presence probabilities of HSM and BSM allowed us to generate the overall suitability map for the native range (Fig. 3). The value of Boyce index obtained for the overall suitability amounted to 0.991 (Fig. 4; see Table S8 for the corresponding Boyce index values).

The output of both HSM and BSM ensemble models was used to forecast the maps of species presence probabilities for HSM and BSM at the world scale, separately (Figs. 5 and 6). The spatial product of

Fig. 1 Habitat suitability map for *P. hilaris* in its native range obtained projecting the ensemble HSM. In red the occurrence points of the species within native range
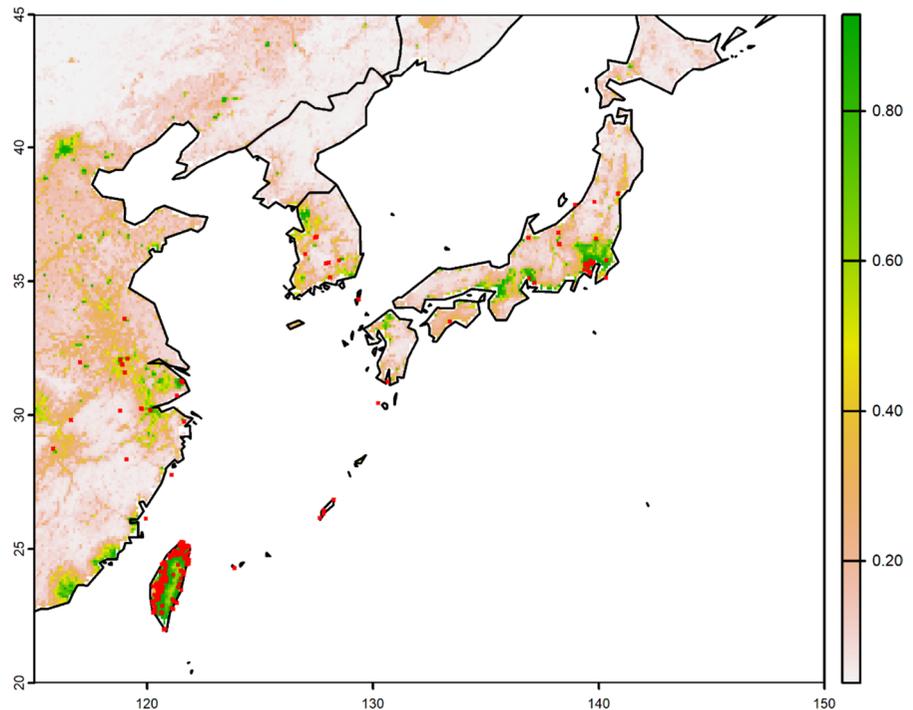
these two maps produced the overall suitability map at a global scale (Fig. 7).

The HSM projection in northern Italy showed a large number of areas highly suitable for *P. hilaris* because of the wide distribution of host plants (Fig. 8). However, the area appears to be rather unsuitable as regards for bioclimatic conditions (Fig. 9). This resulted in a rather low overall suitability, possibly due to a low tolerance of the species to changes in its climatic niche (Fig. 10).

The prediction performance of the overall suitability map in the introduction area in Northern Italy returned a Boyce of 0.909. The graph of Boyce index showed a decreasing tendence for presence probabilities higher than 0.32 (Fig. 11).

The MESS index map showed how the HSM on a global scale was a reliable model. Indeed, the map showed large areas without extrapolation (index > 0) or with moderate extrapolation (0 < index < −50) (Fig. S7). Similarly, the BSM at global scale was also reliable according to the MESS index (Fig. S8). Specifically referring to the invaded area, MESS index highlighted the reliability of the projection of both HSM and BSM (very few areas showed a index < −50).

Discussion

Explanatory models: HSM and BSM built on calibration area (native range)

The host plants represent the fundamental variable to explain the presence of the species, which in the native area is mainly represented by autochthonous plants, i.e., all those considered with the exception of *F. carica*, which is however locally introduced in the native area of *P. hilaris*; this condition emphasized the importance of presence of host plants in determining the distribution of wood-boring beetles. Secondly, a certain degree of urbanization would seem to favor the presence of the species, probably in relation to its ability to exploit the presence of host species used in urban forestry. Furthermore, the species would seem to be able to take advantage of conditions in which natural or semi-natural areas penetrate highly urbanized areas; this condition occurs especially on the island of Taiwan. Altitude is a variable with a modest contribution in explaining the distribution of the species, affecting the local microclimate and indicating a general preference of the species for low altitudes.

Fig. 2 Bioclimatic suitability map for *P. hilaris* in its native range obtained projecting the ensemble BSM; see methods section for explanation of using maps with different calibration extent between HSM and BSM in the native range. In red the occurrence points of the species within native range
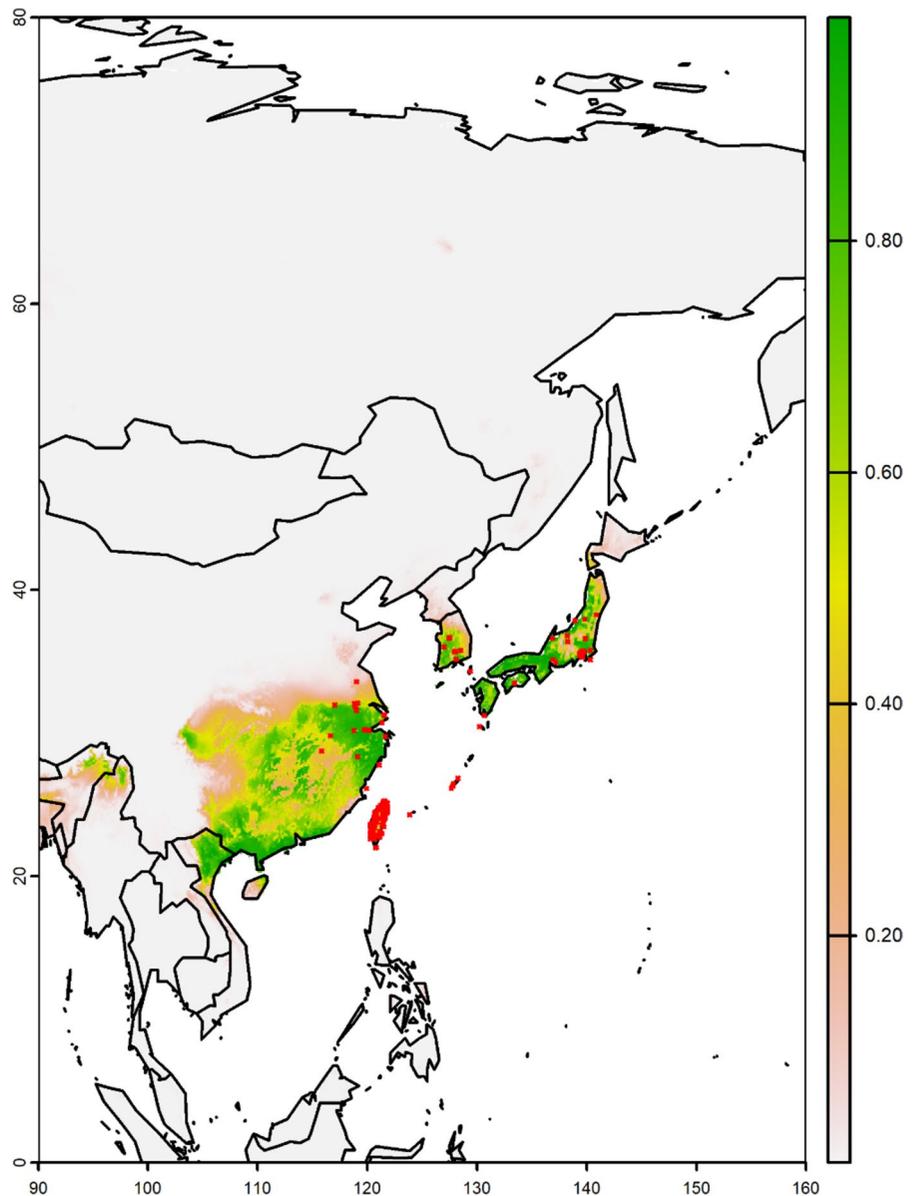
The areas of high suitability are to a large extent located within what is currently the known range of the species, while it is observed that in neighboring areas outside of this, the environmental suitability drops drastically due to the lack of suitable habitats. Most of the eligible areas are concentrated in Taiwan, which is fully eligible except for the summit areas. Elsewhere, in the presumed native range, the areas suitable for *P. hilaris* have a fragmentary distribution and are located in insular (Japan), peninsular (South Korea), or coastal continental (SE China) contexts.

The BSM map highlights rather continuous and homogeneous areas of high suitability within the longitudinal band that includes the native range of the species, including all of Japan south of Hokkaido, South Korea, Taiwan, south-eastern China, and north Vietnam.

The suitability map indicates the following as areas with higher probability of *P. hilaris* presence: (1) Japan, south-eastern belts facing the Pacific, regions of Kantō, Chūbu, Kansai, as well as in the northern part of Kyūshū; (2) South Korea,

Fig. 3 Overall suitability map for *P. hilaris* in its native range obtained combining the projection maps of both habitat suitability model (HSM) and bioclimatic suitability model (BSM). In red the occurrence points of the species within native range
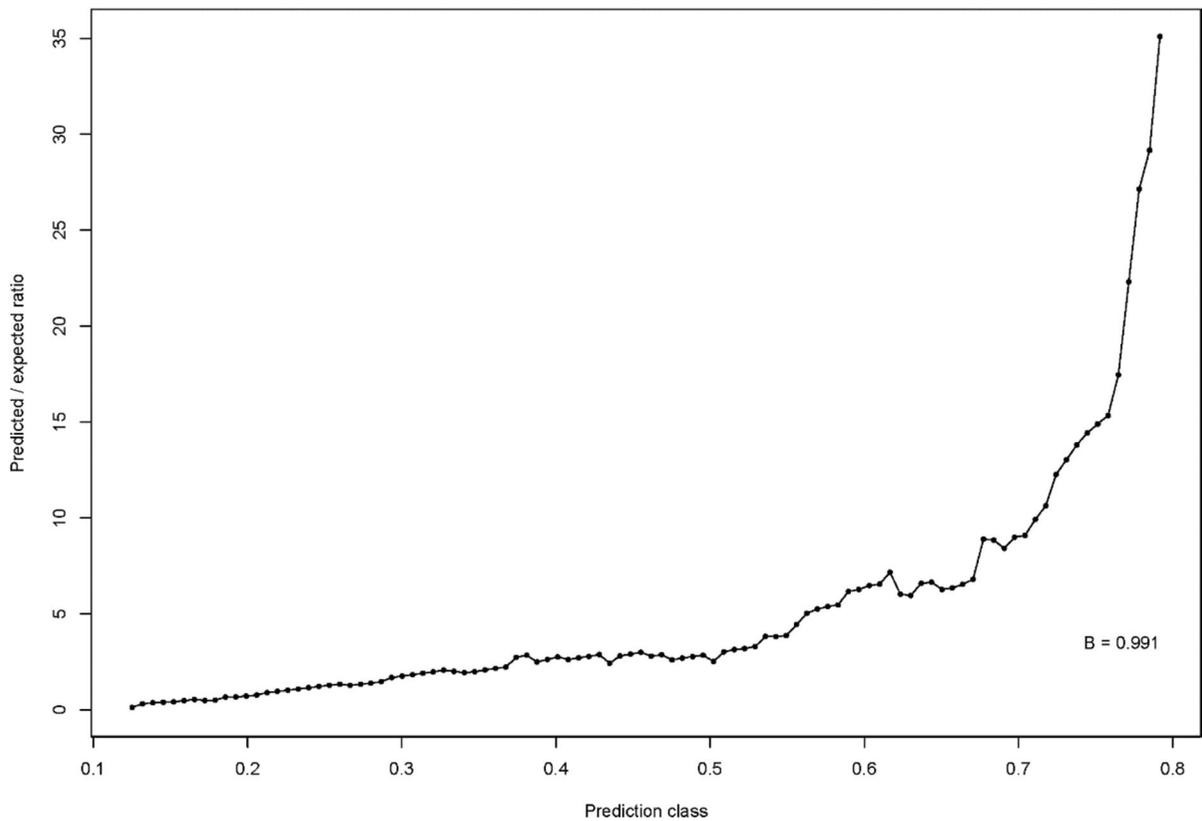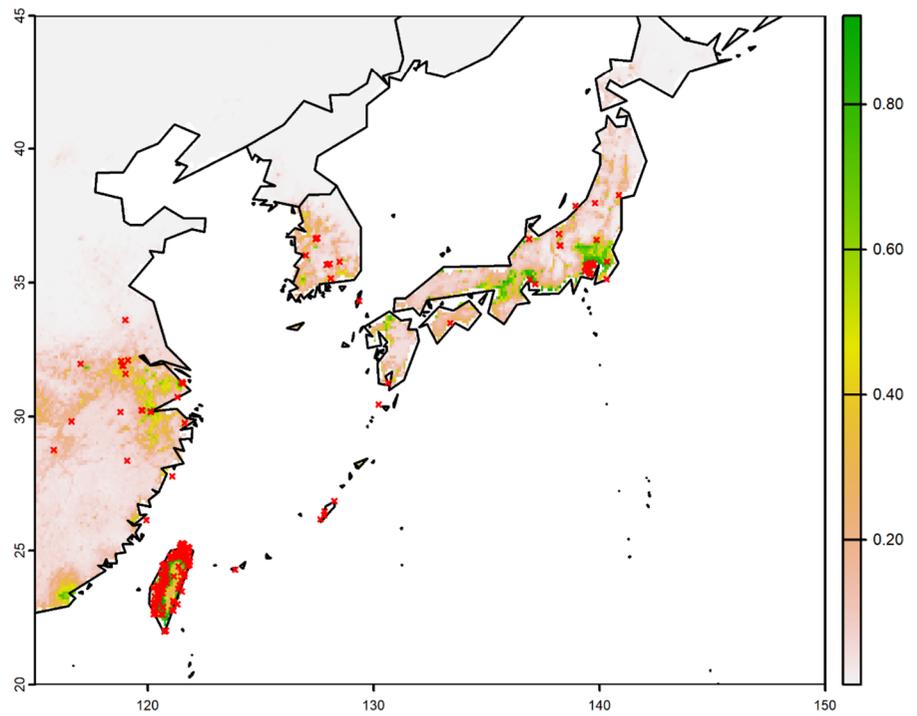

Fig. 4 Boyce of the overall suitability map for *P. hilaris* native range (see Table S8 for the corresponding Boyce index values)
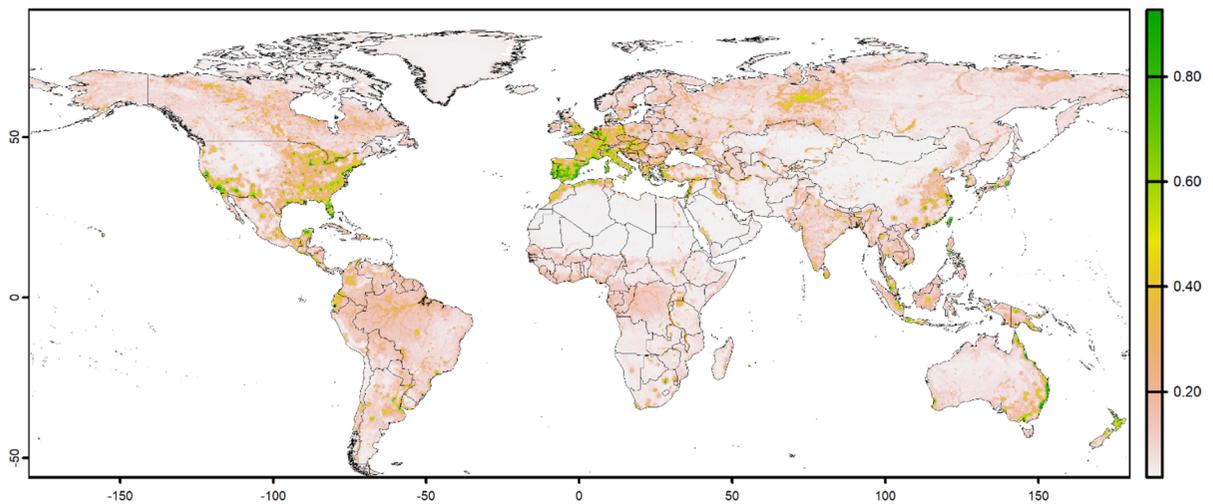
Fig. 5 Habitat suitability map for *P. hilaris* at world scale obtained projecting the ensemble HSM calibrated on the species native range
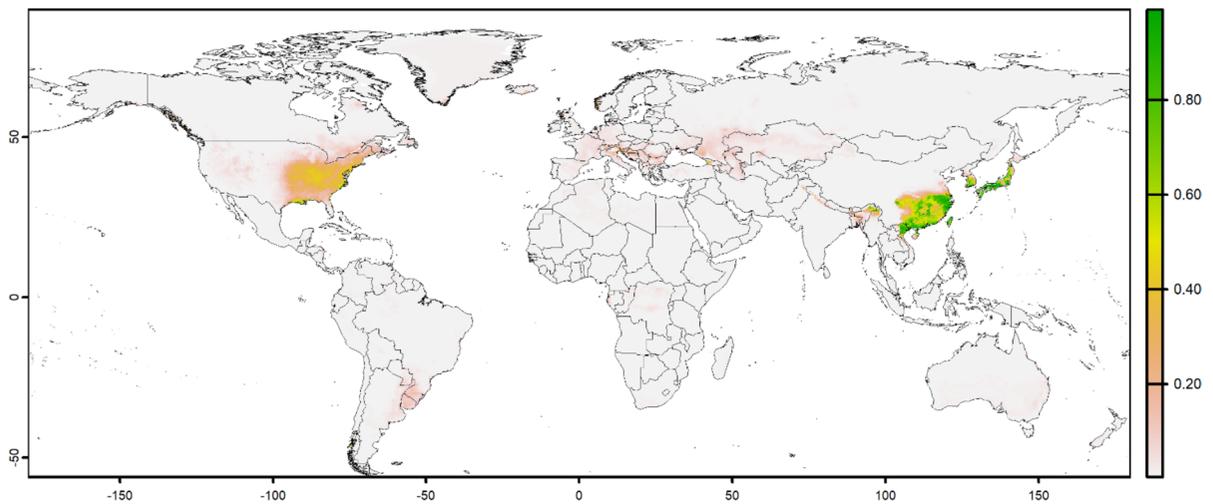
Fig. 6 Bioclimatic suitability map for *P. hilaris* at world scale obtained projecting the ensemble BSM calibrated on the species native range

administrative sections of Seoul, Incheon, and Gyeonggi; (3) China, coastal areas especially in the municipality of Shanghai, and in the provinces of Fujian and Zhejiang; (4) Taiwan almost in its entirety. It is plausible that Taiwan Island, both because of the number of occurrences and the high probability values of occurrence provided by the overall suitability map, may constitute the core area of the range of *P. hilaris.*

Predictive models: *P. hilaris* suitable areas at global scale based on HSM and BSM native model outputs

The HSM identified as the main eligible areas outside the native area: (1) the central and southern areas of Western Europe, including England; (2) insular and coastal contexts of the Middle East; (3) The central-eastern United States and parts of California; (4) Yucatan region in Mexico; (5) East
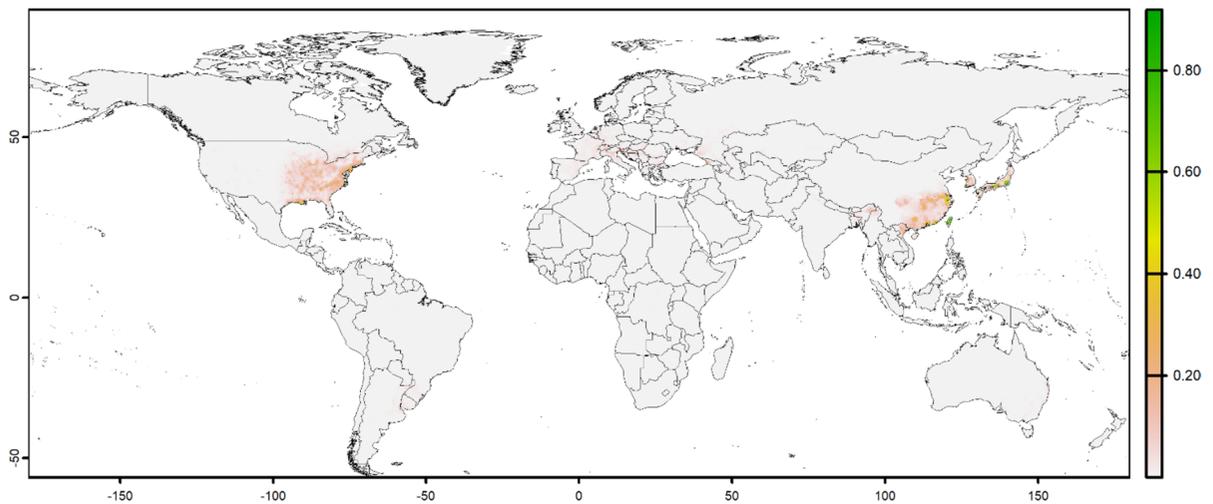
Fig. 7 Overall suitability map for *P. hilaris* at world scale obtained combining the projection maps of both habitat suitability model (HSM) and bioclimatic suitability model (BSM) calibrated on the species native range

Fig. 8 Habitat suitability map for *P. hilaris* in Northern Italy (non-native range) obtained projecting the ensemble HSM calibrated on the species native range. In red the occurrence points of the species within the range of introduction
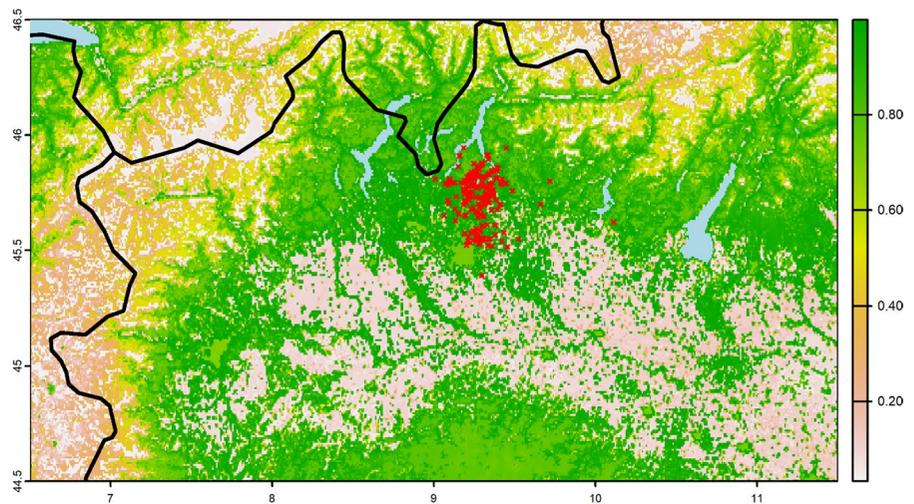
Coast of Australia and North Island in New Zealand; (6) Cape Town and Johannesburg-Pretoria area in South Africa; (7) different coastal contexts of the Malay peninsula and the island of Sumatra. The BSM identified the Central-Eastern United States as the main bioclimatic suitable area outside the native area, while bioclimatically sub-optimal areas were found in: (1) southeastern regions of Brazil and Uruguay; (2) Continental Europe and Central-West Asia regions. The suitability map at global scale indicated how multiple areas with suitable habitats for the species have poorly suited bioclimatic conditions; as a result, the areas suitable for both habitat and bioclimate are decidedly smaller and limited to the Eastern United States. The central and southern areas of Western Europe, while having extensive areas suitable for habitat, have non-optimal bioclimatic conditions. This means that in Europe there are only areas of medium–low overall suitability and are very localized. These areas are located in the lowland and foothill regions of northern Italy, Switzerland, and northern Croatia.

Fig. 9 Bioclimatic suitability map for *P. hilaris* in Northern Italy (non-native range) obtained projecting the ensemble BSM calibrated on the species native range. In red the occurrence points of the species within the range of introduction
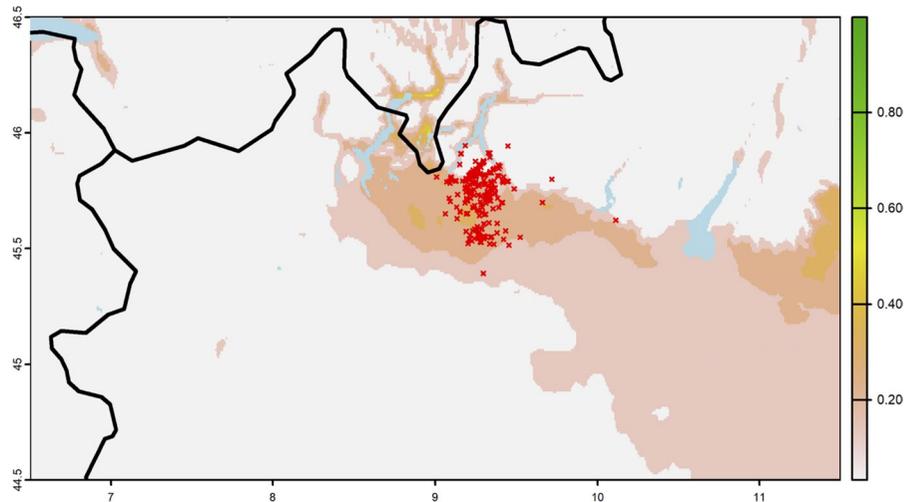
Fig. 10 Overall suitability map for *P. hilaris* in Northern Italy (non-native range) obtained combining the projection maps of both habitat suitability model (HSM) and bioclimatic suitability model (BSM) calibrated on the species native range. In red the occurrence points of the species within the range of introduction
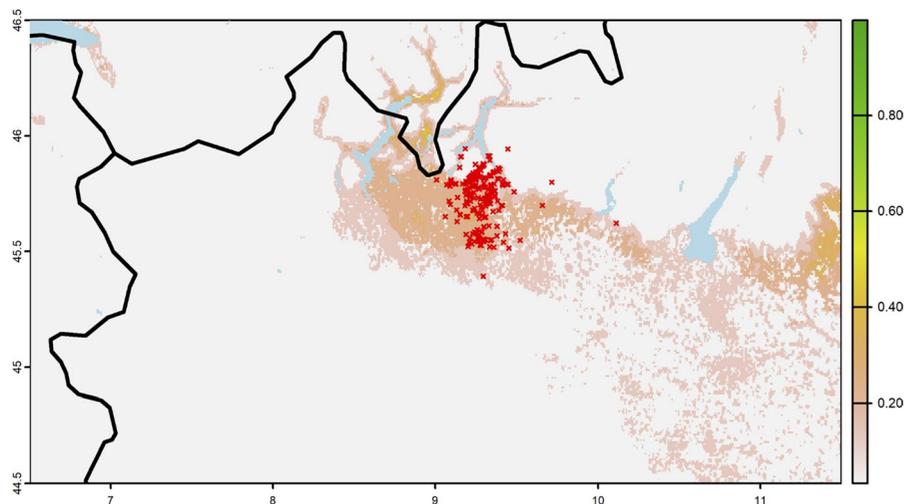
Evaluation of the performance of HSM and BSM in the introduction area

The forecast on a detailed scale regarding the area of introduction of *P. hilaris* in Northern Italy, the only reality of the successful establishment of the species on a global scale, has shown how the species has only partially colonized the suitable areas from the point of view of habitat. These areas extend throughout the Alpine, Pre-Alpine, and Apennine contexts, but also in the lowland areas where patches of residual forest vegetation persist (see Lupi et al. 2023). From a bioclimatic point of view, the species occupies the most bioclimatically favorable areas within the introduction area, although these are suboptimal

conditions according to the native BSM model. This condition is most likely attributable to the presence of the Alpine front which rises abruptly from the Po Valley, causing generally high rainfall in the hottest months of the year, and the presence of lakes which play a buffer role against temperature fluctuations (Shintani and Ishikawa 2002; Claps et al. 2008; Lupi et al. 2023).

The overall model, which takes into account the suitability for both habitat and bioclimatic conditions, once again highlights how *P. hilaris* is found in the most favorable areas, which have average suitability (species presence probability 35–40%). The Boyce index calculated for the invaded area and its graph highlight the good forecasting performance of the
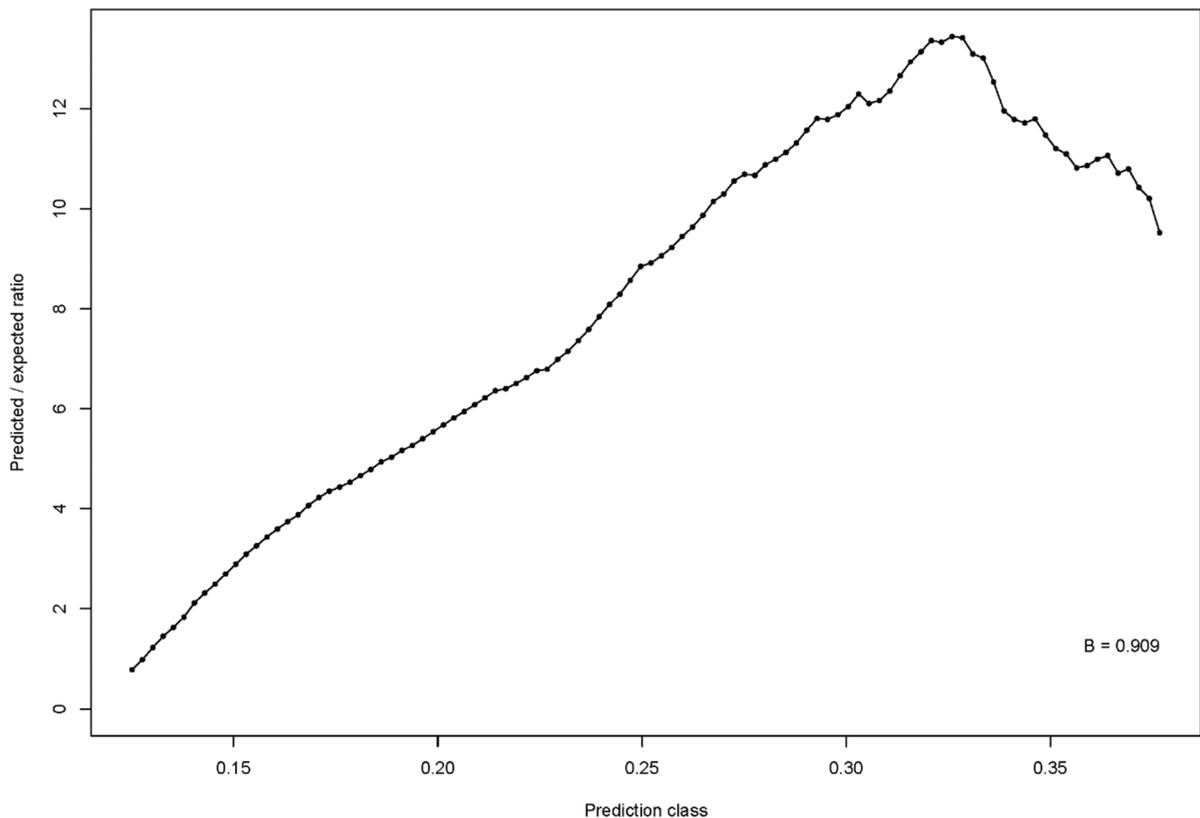
Fig. 11 Boyce index graph of the overall suitability map for *P. hilaris* in Northern Italy (non-native range) (see Table S9 for the corresponding Boyce index values)

overall model, allowing us to exclude a possible niche change by the species in its current phase. On the other hand, the decreasing pattern of the index for the highest presence probability in the invaded area would indicate how the species, although in strong expansion, did not reach the most suitable areas starting from the nuclei of primary introduction (see Lupi et al. 2023). This study indicates that the species is currently widely established and widespread in the provinces of Milan, Lecco, and Como (Lombardy, Northern Italy). The prediction map in the central-western areas of Northern Italy also suggests the potential of *P. hilaris* to expand primarily towards the North (Canton Ticino, CH) given the reduced geographical distance and the presence of more suitable environments for the species in this territory. Furthermore, the spread of *P. hilaris* towards the east is plausible given the presence of suitable areas in the pre-Alpine foothills, such as in the Lessino-Vicentino context.

In addition to a natural expansion, which amounts to an annual rate of 20% (see Lupi et al. 2023), the colonization of these areas by *P. hilaris* could be favored by accidental translocations due to the movement of materials and vehicles along the main communication routes that run in an east–west direction. Lupi et al. (2023) highlight how the current expansion phase is however oriented towards the south; this trend reflects the presence of more suitable areas in the sub- and peri-urban area of Greater Milan, which is particularly characterized by a rich mosaic of green areas, within an urban matrix, many of which are characterized by arboreal vegetation with the presence of preferred host plants. This condition is similar to what is found in the core area of the native range on the island of Taiwan, where the coastal urban belt is penetrated by arboreal vegetation (Peng et al. 2020; Hsiao 2021).

Considerations on possible changes in the *P. hilaris* niche

From the point of view of the habitat, it would seem difficult to hypothesize a predisposition to a substantial niche change by *P. hilaris*. This hypothesis would seem to be supported by the fact that in Eastern China, the species remains localized only in those areas with a favorable habitat, although suitable bioclimatic conditions exist over a wider area. Therefore, the presence of habitats would seem to be more limiting than the bioclimatic condition. In this context, BSM succeeds very well in explaining the presences (bioclimatically favorable areas with a suitable habitat), but less well in explaining the pseudo-absences (bioclimatically favorable areas, but without a suitable habitat). On the other hand, HSM explains very well both presences and pseudo-absences, a symptom that, in a favorable and relatively homogeneous bioclimatic context, the presence or absence of habitat will condition the overall environmental suitability for *P. hilaris*. The predictions of the model created for the native range show that the areas currently exposed to the greatest risk of colonization is the eastern Nearctic, given its suitability for both bioclimatic and habitat conditions. However, with a view to a possible plasticity of the bioclimatic niche, extensive areas of central-western Europe could be exposed to the species invasion given the high suitability of the habitat, mainly due to the extensive diffusion of the host plants. The potential invasive nature of *P. hilaris* is thus constrained by the ability to quickly modify its ecological niche, as demonstrated by Wiens et al. 2019 for other taxa. Assessing how and whether the ecological niche defined by multiple factors, from habitat features to bioclimatic factors, may change between different geographic areas or time periods (Broennimann et al. 2012; Guisan et al. 2014; Trozzi et al. 2024) is a challenge of particular importance for the study of the invasive potential of species in a world increasingly exposed to accidental translocations due to increased global trade.

Caveats and tricks on the use of data obtained from global facilities in SDM

As expected, HSM was less likely to overcome the metric thresholds in comparison to BSM. Given the high variability of habitat covariates on a local scale in comparison with bioclimatic ones, there is the risk of allocating PAs in areas suitable for the presence of the species but for which there are no occurrence data (which is often the case with data originating from incomplete datasets, e.g., GBIF); this condition makes habitat modeling based on presence-only data relatively ineffective, resulting in a lower number of robust models in explaining species presence probability on the basis of habitat characterized by a narrow-scale variability. This condition would lead to obtaining a reduced number of models to be used in the ensemble modeling for the HSM. To obtain a conspicuous number of HSM models, we thus increased the number of sets of PAs. In this way, we were able to achieve a more robust evaluation of the effect of habitat covariates (land covers, DEM, and host plant layer) in explaining *P. hilaris* occurrences with respect to the background habitats.

To evaluate the reliability of occurrence data obtained from GBIF and used in the modeling framework, other than those with a coordinate uncertainty lower than 500 m, we also tested the use of data without a declared uncertainty. This has made it possible to have more than double the occurrences available for modeling. However, we obtained a marked drop in the robustness of models to be used in building the ensemble model for the HSM, from 15.5 to 0.5% (while for the BSM 100–91%, according to TSS model validation; results not shown). This is due to the difficulty in capturing the true effect of an environmental covariate that has a narrow-scale variability (such as those concerning habitat layers) when an occurrence is georeferenced in an approximate manner. This is the reason why we decided to keep only those data with a declared reduced coordinate uncertainty (<500 m).

Bioclimatic variables show substantial variability on a large geographical scale; thus, to capture the effect of bioclimatic variables, we extended the allocation area of PAs also to include geographical areas not close to the native range, for which a certain degree of deviation from the conditions found in the native range may exist. This condition allows bioclimatic models to have a greater capability in explaining species occurrences since PAs have a higher chance to be true zeros, and some bioclimatic variables may actually have a significant effect in constraining the distribution of the species. This

aspect is of particular importance when one of the aims is to identify potential bioclimatically suitable areas outside the native range. Conversely, bounding the identification of PAs strictly within the native range in building HSM may help to limit the confounding effect of bioclimatic variables in assessing the effect of habitat variables on species presence.

Conclusions

Predicting the establishment and expansion success of an introduced species in a new geographic area remains a major challenge of invasion science (Srivastava et al. 2021; Liu et al. 2022).

The preparation of the data to be used and the fine-tuning of the modeling analyses places the emphasis on the careful usage of georeferenced data according to their accuracy, in relation to the resolution of the cartography used for HSM. This is particularly important since analyses involve the use of data which have high variability on a small scale, such as habitat layers (e.g., land cover). In addition, we stress the importance of adopting a two-step modeling approach to better capture the effect of covariates that have appreciable/substantial variability on different spatial scales (see Adde et al. 2023). In the modeling approach, it is extremely important to consider the host plant when creating a model specifically aimed at explaining and predicting the distribution of a wood-borer, as highlighted by Dang et al. (2021); in fact, the host plant is a key variable in defining the environmental suitability from the point of view of the habitat. The habitat is also particularly binding with respect to the bioclimatic conditions. If *P. hilaris* were unable to change its niche due to bioclimatic constraints, it should in future be considered a modestly invasive species both in Italy and in the circum-Mediterranean and central-western European context. On the other hand, if the species were able to change its bioclimatic niche, it could become a first-level invasive species in the above-mentioned areas given the large presence of suitable habitats primarily defined by the host plants. The latter phenomenon cannot be ruled out at present since

P. hilaris currently presents a high rate of areal expansion in a fairly suitable bioclimatic context.

References

Adde A, Rey PL, Brun P, Külling N, Fopp F, Altermatt F, Petitpierre P, Zimmermann NE, Pellissier L, Guisan A (2023) N-SDM: a high-performance computing pipeline for nested species distribution modelling. Ecography. https://doi.org/10.1111/ecog.06540

Ahmad S, Yang L, Khan TU, Wanghe K, Li M, Luan X (2020) Using an ensemble modelling approach to predict the potential distribution of Himalayan gray goral

(*Naemorhedus goral bedfordi*) in Pakistan. Glob Ecol Conserv. https://doi.org/10.1016/j.gecco.2019.e00845

Aiello-Lammens ME, Boria RA, Radosavljevic A, Vilela B, Anderson RP (2015) spThin: an R package for spatial thinning of species occurrence records for use in ecological niche models. Ecography. https://doi.org/10.1111/ecog.01132

Akşit T, Çakmak İ, Özsemerci F (2005) Some new xylophagous species on fig trees (*Ficus carica* cv. *calymirna* L.) in Aydın. Turkey Turk J Zool 29:211–215

Allen EA, Humble LM (2002) Nonindigenous species introductions: a threat to Canada's forest and forest economy. Can J Plant Pathol. https://doi.org/10.1080/07060660309506983

Allouche O, Tsoar A, Kadmon R (2006) Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). J Appl Ecol. https://doi.org/10.1111/j.1365-2664.2006.01214.x

Araújo MB, New M (2007) Ensemble forecasting of species distributions. Trends Ecol Evol. https://doi.org/10.1016/j.tree.2006.09.010

Badgujar SB, Patel VV, Bandivdekar AH, Mahajan RT (2014) Traditional uses, phytochemistry and pharmacology of *Ficus carica*: A review. Pharm Biol. https://doi.org/10.3109/13880209.2014.892515

Barbet-Massin M, Jiguet F, Albert CH, Thuiller W (2012) Selecting pseudo-absences for species distribution models: How, where and how many? Methods Ecol Evol. https://doi.org/10.1111/j.2041-210X.2011.00172.x

Barbosa AM, Real R, Munoz AR, Brown JA (2013) New measures for assessing model equilibrium and prediction mismatch in species distribution models. Divers Distrib. https://doi.org/10.1111/ddi.12100

Boyce MS, Vernier PR, Nielsen SE, Schmiegelow FKA (2002) Evaluating resource selection functions. Ecol Model. https://doi.org/10.1016/S0304-3800(02)00200-4

Breiner FT, Guisan A, Bergamini A, Nobis MP (2015) Overcoming limitations of modelling rare species by using ensembles of small models. Summary Methods Ecol Evol 6(10):1210–1218. https://doi.org/10.1111/2041-210X.12403

Broennimann O, Fitzpatrick MC, Pearman PB, Petitpierre B, Pellissier L, Yoccoz NG, Thuiller W, Fortin M-J, Randin C, Zimmermann NE, Graham CH, Guisan A (2012) Measuring ecological niche overlap from occurrence and spatial environmental data. Glob Ecol Biogeogr. https://doi.org/10.1111/j.1466-8238.2011.00698.x

Byeon DH, Kim SH, Jung JM, Jung S, Kim KH, Lee WH (2021) Climate-based ensemble modelling to evaluate the global distribution of *Anoplophora glabripennis* (Motschulsky). Agric for Entomol. https://doi.org/10.1111/afe.12462

Claps P, Giordano P, Laguardia G (2008) Spatial distribution of the average air temperatures in Italy: quantitative analysis. J Hydrol Eng 13(4):242–249. https://doi.org/10.1061/(ASCE)1084-0699(2008)13:4(242)

Cocquempot C, Lindelöw Å (2010) Longhorn beetles (Coleoptera, Cerambycidae). BioRisk. https://doi.org/10.3897/biorisk.4.56

Cohen J (1960) A coefficient of agreement for nominal scales. Educ Psychol Meas. https://doi.org/10.1177/001316446002000104

Conti B, Raspi A (2007) First record for Italy of *Luperomorpha nigripennis* Duvivier (Coleoptera Chrysomelidae). Inf Fitopatol 57:51–52

Crisosto H, Ferguson L, Bremer V, Stover E, Colelli G (2011) Fig (*Ficus carica* L.). In: Yahia EM (ed) Postharvest biology and technology of tropical and subtropical fruits. Woodhead Publishing, Sawston, pp 134–160

Dang YQ, Zhang YL, Wang XY, Xin B, Quinn NF, Duan JJ (2021) Retrospective analysis of factors affecting the distribution of an invasive wood-boring insect using native range data: The importance of host plants. J Pest Sci. https://doi.org/10.1007/s10340-020-01308-5

Danilevsky ML (2020) Catalogue of *Palaearctic Coleoptera*, Chrysomeloidea I (Vesperidae, Disteniidae, Cerambycidae). Brill, Leiden, Boston

Di Iorio OR (2004) Exotic species of Cerambycidae (Coleoptera) introduced in Argentina. Part 2. New records, host plants, emergence periods, and current status. Agrociencia 38:663–678

Dondina O, Orioli V, Torretta E, Merli F, Bani L, Meriggi A (2020) Combining ensemble models and connectivity analyses to predict wolf expected dispersal routes through a lowland corridor. PLoS ONE. https://doi.org/10.1371/journal.pone.0229261

Dormann CF, Elith J, Bacher S, Buchmann C, Carl G, Carré G, Marquéz JR, Gruber B, Lafourcade B, Leitão BJ, Münkemüller T, Mcclean C, Osborne PE, Reineking B, Schröder B, Skidmore AK, Zurell D, Lautenbach S (2013) Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography. https://doi.org/10.1111/j.1600-0587.2012.07348.x

Dupin M, Reynaud P, Jarošík V, Baker R, Brunel S, Eyre D, Pergl J, Makowski D (2011) Effects of the training dataset characteristics on the performance of nine species distribution models: application to *Diabrotica virgifera virgifera*. PLoS ONE. https://doi.org/10.1371/journal.pone.0020957

Elith J, Kearney M, Phillips S (2010) The art of modelling range-shifting species. Methods Ecol Evol. https://doi.org/10.1111/j.2041-210X.2010.00036.x

EPPO (2008) *Psacothea hilaris* detected in the United Kingdom: addition to the EPPO Alert List. EPPO (2008) Reporting Service, no. 10-2008/201. https://gd.eppo.int/reporting/article-824. Accessed on 15 March 2022

EPPO (2013) New data on quarantine pests and pests of the EPPO Alert List. EPPO Reporting Service, no. 11-2013/245. https://gd.eppo.int/reporting/article-2707. Accessed on 15 March 2022

Eyre D, Haack RA (2017) Invasive cerambycid pests and biosecurity measures. In: Wang Q (ed) Cerambycidae of the world: biology and pest management. CRC Press, Boca Raton, pp 563–618

Faccoli M, Frigimelica G, Mori N, Petrucco Toffolo E, Vettorazzo M, Simonato M (2009) First record of *Ambrosiodmus* (Hopkins, 1915) (Coleoptera: Curculionidae, Scolytinae) in Europe. Zootaxa. https://doi.org/10.11646/zootaxa.2303.1.4

Faccoli M, Campo G., Perrotta G, Rassati D (2016) Two newly introduced tropical bark and ambrosia beetles (Coleoptera: Curculionidae, Scolytinae) damaging figs (*Ficus carica*) in southern Italy. Zootaxa. https://doi.org/10.11646/zootaxa.4138.1.10

Farina P, Mazza G, Benvenuti C, Cutino I, Giannotti P, Conti B, Bedini S, Gargani E (2020) Biological notes and distribution in Southern Europe of *Aclees taiwanensis* Kôno, 1933 (Coleoptera: Curculionidae): a new pest of the fig tree. Insects. https://doi.org/10.3390/insects12010005

Fenn-Moltu G, Ollier S, Bates OK, Liebhold AM, Nahrung HF, Pureswaran DS, Yamanaka T, Bertelsmeier C (2023) Global flows of insect transport and establishment: The role of biogeography, trade and regulations. Divers Distrib. https://doi.org/10.1111/ddi.13772

Fournier A, Barbet-Massin M, Rome Q, Courchamp F (2017) Predicting species distribution combining multi-scale drivers. Glob Ecol Conserv. https://doi.org/10.1016/j.gecco.2017.11.002

Früh L, Kampen H, Kerkow A, Schaub GA, Walther D, Wieland R (2018) Modelling the potential distribution of an invasive mosquito species: comparative evaluation of four machine learning methods and their combinations. Ecol Model. https://doi.org/10.1016/j.ecolmodel.2018.08.011

Fukaya M, Honda H (1992) Reproductive biology of the yellow-spotted longicorn beetle, *Psacothea hilaris* (Pascoe) (Coleoptera: Cerambycidae): I. Male mating behaviors and female sex pheromones. Appl Entomol Zool. https://doi.org/10.1007/BF02055097

Gaaliche B, Ben Abdelaali N, Mouttet R, Ben Halima-Kamel M, Hajlaoui MR (2018) Nouveau signalement de *Hypocryphalus scabricollis* (Eichhoff, 1878) en Tunisie, un ravageur émergent sur figuier (*Ficus carica* L.). EPPO Bull. https://doi.org/10.1111/epp.12459

Gelb J, Apparicio P (2021) Contribution of the spatial c-means fuzzy classification in geography: a socio-residential and environmental taxonomy in Lyon. Cybergeo. https://doi.org/10.4000/cybergeo.36414

Guisan A, Petitpierre B, Broennimann O, Daehler C, Kueffer C (2014) Unifying niche shift studies: insights from biological invasions. Trends Ecol Evol. https://doi.org/10.1016/j.tree.2014.02.009

Guisan A, Thuiller W, Zimmermann NE (2017) Habitat suitability and distribution models: with applications in R. Cambridge University Press, Cambridge

Haack RA (2006) Exotic bark-and wood-boring Coleoptera in the United States: recent establishments and interceptions. Can J For Res 36(2): 269–288

Haack RA, Britton KO, Brockerhoff EG, Cavey JF, Garrett LJ, Kimberley M, Lowenstein F, Nuding A, Olson LJ, Turner J, Vasilaky KN (2014) Effectiveness of the international phytosanitary standard ISPM no.15 on reducing wood borer infestation rates in wood packaging material entering the United States. PLoS One. https://doi.org/10.1371/journal.pone.0096611

Hanley JA, McNeil BJ (1982) The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology. https://doi.org/10.1148/radiology.143.1.7063747

Hao T, Elith J, Guillera-Arroita G, Lahoz-Monfort JJ (2019) A review of evidence about use and performance of species distribution modelling ensembles like BIOMOD. Divers Distrib. https://doi.org/10.1111/ddi.12892

Hao T, Elith J, Lahoz-Monfort JJ, Guillera-Arroita G (2020) Testing whether ensemble modelling is advantageous for maximising predictive performance of species distribution models. Ecography. https://doi.org/10.1111/ecog.04890

Hijmans RJ, Phillips S, Leathwick J, Elith J (2023) dismo: Species Distribution Modeling. R package version 1.3–14, <https://CRAN.R-project.org/package=dismo>.

Hirzel AH, Le Lay G, Helfer V, Randin C, A. Guisan A, (2006) Evaluating the ability of habitat suitability models to predict species presences. Ecol Model. https://doi.org/10.1016/j.ecolmodel.2006.05.017

Hortal J, Roura-Pascual N, Sanders NJ, Rahbek C (2010) Understanding (insect) species distributions across spatial scales. Ecography. https://doi.org/10.1111/j.1600-0587.2009.06428.x

Hsiao H (2021) Characteristics of urban gardens and their accessibility to locals and non-locals in Taipei City Taiwan. Landsc Ecol Eng. https://doi.org/10.1007/s11355-020-00430-x

Iannella M, D'Alessandro P, Biondi M (2020) Forecasting the spread associated with climate change in Eastern Europe of the invasive Asiatic flea beetle *Luperomorpha xanthodera* (Coleoptera: Chrysomelidae). Eur J Entomol. https://doi.org/10.14411/eje.2020.015

Iba M (1980) Ecological studies on the yellow-spotted longicorn beetle, *Psacothea hilaris* Pascoe. 376 IV. The geographic difference of the pattern on the pronotum of adult. J Sericult Sci Jpn 49:429–433 (**in Japanese**)

Iba M (1993) Studies on the ecology and control of the yellow spotted longicorn beetle, *Psacothea hilaris* Pascoe (Coleoptera; Cerambycidae), in the mulberry fields. Bulletin of the National Institute of Sericultural and Entomological Science (japan) 81:1–119

INPN (2021) Sheet of *Psacothea hilaris* (Pascoe, 1857). Inventaire national du patrimoine naturel 379 (INPN). https://inpn.mnhn.fr/espece/cd_nom/943208. Accessed on 15 Dec. 2021

Ismail AI, El-Hawary FM, Abdel-Moniem ASH (2009) Control of the fig long-horned beetle, *Hesperophanes griseus* (Fabricius) (Coleoptera: Cerambycidae) on the fig trees, *Ficus carica* L. Egypt J Biol Pest Control 19:65–70

Jucker C, Tantardini A, Colombo M (2006) First record of *Psacothea hilaris* (Pascoe) in Europe (Coleoptera Cerambycidae Lamiinae Lamiini). Bollettino Di Zoologia Agraria e Di Bachicoltura 38:187–191

Kindt R, Coe R (2005) Tree diversity analysis. A manual and software for common statistical methods for ecological and biodiversity studies. World Agroforestry Centre, Nairobi

Krishnankutty SM, Bigsby K, Hastings J, Takeuchi Y, Wu Y, Lingafelter SW, Nadel H, Myers SW, Ray AM (2020) Predicting establishment potential of an invasive wood-boring beetle, *Trichoferus campestris* (Coleoptera: Cerambycidae) in the United States. Ann Entomol Soc Am. https://doi.org/10.1093/aesa/saz051

Lee S, Cha D, Nam Y, Jung J (2021) Genetic diversity of a rising invasive pest in the native range: Population genetic structure of *Aromia bungii* (Coleoptera: Cerambycidae) in South Korea. Diversity. https://doi.org/10.3390/d13110582

Li Y, Johnson AJ, Gao L, Wu C, Hulcr J (2021) Two new invasive *Ips* bark beetles (Coleoptera: Curculionidae) in mainland China and their potential distribution in Asia. Pest Manag Sci. https://doi.org/10.1002/ps.6423

Lianju W, Weibin J, Kai M, Zhifeng L, Yelin W (2003) The production and research of fig (*Ficus carica* L.) in China. Acta Hortic. https://doi.org/10.17660/ActaHortic.2003.605.28

Liu C, Wolter C, Courchamp F, Roura-Pascual N, Jeschke JM (2022) Biological invasions reveal how niche change affects the transferability of species distribution models. Ecology. https://doi.org/10.1002/ecy.3719

López-Martínez V, Vargas OR, Alia-Tejacal I, Toledo-Hernández VH, Corona-López AM, Delfín-González H, Guillen-Sánchez D, Jiménez-García D (2015) Xylophagous beetles (Coleoptera: Buprestidae and Cerambycidae) from *Ficus carica* L. (Moraceae) in Morelos, Mexico. Coleopt Bull. https://doi.org/10.1649/0010-065X-69.4.780

Lu Z, Liu X, Wang T, Zhang P, Wang Z, Zhang Y, Kriticos DJ, Zalucki MP (2022) Malice at the Gates of Eden: current and future distribution of *Agrilus mali* threatening wild and domestic apples. Bull Entomol Res. https://doi.org/10.1017/S000748532200013X

Lupi D, Jucker C, Colombo M (2013) Distribution and biology of the yellow-spotted longicorn 405 beetle *Psacothea hilaris hilaris* (Pascoe) in Italy. EPPO Bull. https://doi.org/10.1111/epp.12045

Lupi D, Malabusini S, de Milato S, Heinzl AL, Ruzzier E, Bani L, Savoldelli S, Jucker C (2023) Exploring the range expansion of the yellow-spotted longhorn beetle *Psacothea hilaris hilaris* in northern Italy. Agric for Entomol. https://doi.org/10.1111/afe.12570

MacQuarrie CJ, Gray M, Lavallée R, Noseworthy MK, Savard M, Humble LM (2020) Assessment of the systems approach for the phytosanitary treatment of wood infested with wood-boring insects. J Econ Entomol. https://doi.org/10.1093/jee/toz331

Marchioro M, Faccoli M, Dal Cortivo M, Branco M, Roques A, Garcia A, Ruzzier E (2022) New species and new records of exotic Scolytinae (Coleoptera, Curculionidae) in Europe. Biodivers Data J. https://doi.org/10.3897/BDJ.10.e93995

Mateo RG, Gastón A, Aroca-Fernández MJ, Broennimann O, Guisan A, Saura S, García-Viñas JI (2019) Hierarchical species distribution models in support of vegetation conservation at the landscape scale. J Veg Sci. https://doi.org/10.1111/jvs.12726

Meller L, Cabeza M, Pironon S, Barbet-Massin M, Maiorano L, Georges D, Thuiller W (2014) Ensemble distribution models in conservation prioritization: from consensus predictions to consensus reserve networks. Divers Distrib. https://doi.org/10.1111/ddi.12162

Mendoza-Castillo VM, Vargas-Canales JM, Calderón-Zavala G, Mendoza-Castillo MDC, Santacruz-Varela A (2017) Intensive production systems of fig (*Ficus carica* L.) under greenhouse conditions. Exp Agric. https://doi.org/10.1017/S0014479716000405

Naimi B, Hamm NAS, Groen TA, Skidmore AK, Toxopeus AG (2014) Where is positional uncertainty a problem for species distribution modelling? Ecography. https://doi.org/10.1111/j.1600-0587.2013.00205.x

Nardi G, Badano D, De Cinti B (2015) First record of *Dinoderus* (*Dinoderastes*) *japonicus* in Italy (Coleoptera: Bostrichidae). Fragm Entomol. https://doi.org/10.13133/2284-4880/143

Pearson RG, Dawson TP (2003) Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? Glob Ecol Biogeogr. https://doi.org/10.1046/j.1466-822X.2003.00042.x

Pedlar JH, McKenney DW, Yemshanov D, Hope ES (2020) Potential economic impacts of the Asian longhorned beetle (Coleoptera: Cerambycidae) in Eastern Canada. J Econ Entomol. https://doi.org/10.1093/jee/toz317

Peng MH, Hung YC, Liu KL, Neoh KB (2020) Landscape configuration and habitat complexity shape arthropod assemblage in urban parks. Sci Rep. https://doi.org/10.1038/s41598-020-73121-0

Peterson AT, Scachetti-Pereira R (2004) Potential geographic distribution of *Anoplophora glabripennis* (Coleoptera: Cerambycidae) in North America. Am Midl Nat. https://doi.org/10.1674/0003-0031(2004)151[0170:PGDOAG]2.0.CO;2

QGIS Development Team (2023) QGIS 3.32. https://www.qgis.org/it/site/index.html

R Core Team (2021) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Rassati D, Lieutier F, Faccoli M (2016) Alien wood-boring beetles in Mediterranean regions. In: Paine TD, Lieutier F (eds) Insects and diseases of Mediterranean forest systems. Springer, Cham, pp 293–327

Richardson DM, Pyšek P, Carlton J (2011) A compendium of essential concepts and terminology in invasion ecology. In: Richardson DM (ed) Fifty years of invasion ecology: the legacy of Charles Elton. John Wiley & Sons Ltd., Oxford, pp 409–420

Ruzzier E, Colla A (2019) *Micromalthus debilis* LeConte, 1878 (Coleoptera: Micromalthidae), an American wood-boring beetle new to Italy. Zootaxa. https://doi.org/10.11646/zootaxa.4623.3.12

Ruzzier E, Morin L, Glerean P, Forbicioni L (2020a) New and Interesting Records of Coleoptera from Northeastern Italy and Slovenia (Alexiidae, Buprestidae, Carabidae, Cerambycidae, Ciidae, Curculionidae, Mordellidae, Silvanidae). Coleopt Bull. https://doi.org/10.1649/0010-065X-74.3.523

Ruzzier E, Tomasi F, Poso M, Martinez-Sanudo I (2020b) *Archophileurus spinosus* Dechambre, 2006 (Coleoptera: Scarabaeidae: Dynastinae) a new exotic scarab possibly acclimatized in Italy, with a compilation of exotic Scarabaeidae found in Europe. Zootaxa. https://doi.org/10.11646/zootaxa.4750.4.8

Ruzzier E, Tomasi F, Platia G, Pulvirenti E (2021) Exotic Elateridae (Coleoptera: Elateroidea) in Italy: an overview. Coleopt Bull. https://doi.org/10.1649/0010-065X-75.3.673

Ruzzier E, Bani L, Cavaletto G, Faccoli M, Rassati D (2022) *Anisandrus maiche* Kurentzov (Curculionidae: Scolytinae), an Asian species recently introduced and now widely established in Northern Italy. BioInvasions Rec. https://doi.org/10.3391/bir.2022.11.3.07

Ruzzier E, Haack RA, Curletti G, Roques A, Volkovitsh MG, Battisti A (2023a) Jewels on the go: exotic buprestids around the world (Coleoptera, Buprestidae). NeoBiota. https://doi.org/10.3897/neobiota.84.90829

Ruzzier E, Morin L, Zugno M, Tapparo A, Bani L, Di Giulio A (2023b) New records of non-native Coleoptera in Italy. Biodivers Data J. https://doi.org/10.3897/BDJ.11.e111487

Ruzzier E, Sañudo IM, Cavaletto G, Faccoli M, Smith SM, Cognato AI, Rassati D (2023c) Detection of native-alien populations of *Anisandrus dispar* (Fabricius, 1792) in Europe. J Asia Pac Entomol. https://doi.org/10.1016/j.aspen.2023.102137

Sarto i Monteys V, Torras i Tutusaus G (2018) A new alien invasive longhorn beetle, *Xylotrechus chinensis* (Cerambycidae), is infesting mulberries in Catalonia (Spain). Insects. https://doi.org/10.3390/insects9020052

Seidel M, Lüttke M, Cocquempot C, Potts K, Heeney WJ, Husemann M (2021) Citizen scientists significantly improve our knowledge on the non-native longhorn beetle *Chlorophorus annularis* (Fabricius, 1787) (Coleoptera, Cerambycidae) in Europe. BioRisk. https://doi.org/10.3897/biorisk.16.61099

Shamin-Shazwan K, Shahari R, Amri CNAC, Tajuddin NSM (2019) Figs (*Ficus carica* L.): cultivation method and production based in Malaysia. Eng Herit J. https://doi.org/10.26480/gwk.02.2019.06.08

Shatz AJ, Rogan J, Sangermano F, Ogneva-Himmelberger Y, Chen H (2013) Characterizing the potential distribution of the invasive Asian longhorned beetle (*Anoplophora glabripennis*) in Worcester county, Massachusetts. Appl Geogr. https://doi.org/10.1016/j.apgeog.2013.10.002

Shintani Y, Ishikawa Y (2002) Geographic variation in cold hardiness of eggs and neonate larvae of the yellow-spotted longicorn beetle *Psacothea hilaris*. Physiol Entomol. https://doi.org/10.1046/j.1365-3032.1999.00126.x

Sopow S, Gresham B, Bain J (2015) Exotic longhorn beetles (Coleoptera: Cerambycidae) established in New Zealand. N Z Entomol. https://doi.org/10.1080/00779962.2014.993798

Sparacio I, Ditta A, Surdo S (2020) On the presence of the alien exotic sap beetle *Phenolia* (*Lasiodites*) *picta* (Macleay, 1825) (Coleoptera Nitidulidae) in Italy. Biodivers J. https://doi.org/10.31396/Biodiv.Jour.2020.11.2.439.442

Srivastava V, Roe AD, Keena MA, Hamelin RC, Griess VC (2021) Oh the places they'll go: improving species distribution modelling for invasive forest pests in an uncertain world. Biol Invasions. https://doi.org/10.1007/s10530-020-02372-9

Steen VA, Tingley MW, Paton PW, Elphick CS (2021) Spatial thinning and class balancing Key choices lead to variation in the performance of species distribution models with citizen science data. Methods Ecol Evol. https://doi.org/10.1111/2041-210X.13525

Thuiller W, Georges D, Gueguen M, Engler R, Breiner F (2021). biomod2: Ensemble Platform for Species Distribution Modeling. R package version 4.2–4. https://CRAN.R-project.org/package=biomod2

Tinoco STJ, Berti Filho E, Pereira RA (2010) Aspects of the biology of *Marshallius bonellii* (Boheman, 1830) (Coleoptera, Curculionidae, Hylobiinae), a pest of fig (*Ficus carica*). Rev Agric (piracicaba) 85:172–178

Tirozzi P, Orioli V, Dondina O, Bani L (2024) Connection between ecological niche changes and population trends in the Eurasian Skylark *Alauda arvensis* breeding in lowland and mountain areas of southern Europe. Ibis (in press). https://doi.org/10.1111/ibi.13322

Toma L, Ramos RY, Severini F, Di Luca M, Mei M, Zampetti MF (2017) First record of *Bruchidius raddianae* in Italy infested seeds of *Vachellia karroo* from Lampedusa Island (Coleoptera: Bruchidae; Fabales: Fabaceae). Fragm Entomol. https://doi.org/10.13133/2284-4880/236

Tuanmu M-N, Jetz W (2014) A global 1-km consensus land-cover product for biodiversity and ecosystem modeling. Glob Ecol Biogeogr. https://doi.org/10.1111/geb.12182

Turner RM, Brockerhoff EG, Bertelsmeier C, Blake RE, Caton B, James A, MacLeod A, Nahrung HF, Pawson SM, Plank MJ, Pureswaran DS, Seebens H, Yamanaka T, Liebhold AM (2021) Worldwide border interceptions provide a window into human-mediated global insect movement. Ecol Appl. https://doi.org/10.1002/eap.2412

Tzenov P, Cappellozza S, Saviane A (2022) Black, Caspian seas and central asia silk association (BACSA) for the future of sericulture in Europe and Central Asia. Insects. https://doi.org/10.3390/insects13010044

Urbanek Krajnc A, Ugulin T, Paušič A, Rabensteiner J, Bukovac V, Mikulič Petkovšek M et al (2019) Morphometric and biochemical screening of old mulberry trees (*Morus alba* L.) in the former sericulture region of Slovenia. Acta Soc Bot Pol. https://doi.org/10.5586/asbp.3614

Valavi R, Guillera-Arroita G, Lahoz-Monfort JJ, Elith J (2022) Predictive performance of presence-only species distribution models: a benchmark study with reproducible code. Ecol Monogr. https://doi.org/10.1002/ecm.1486

Wiens JJ, Litvinenko Y, Harris L, Jezkova T (2019) Rapid niche shifts in introduced species can be a million times faster than changes among native species and ten times faster than climate change. J Biogeogr. https://doi.org/10.1111/jbi.13649

Wu Y, Trepanowski NF, Molongoski JJ, Reagel PF, Lingafelter SW, Nadel H, Mayers SW, Ray AM (2017) Identification of wood-boring beetles (Cerambycidae and Buprestidae) intercepted in trade-associated solid wood packaging material using DNA barcoding and morphology. Sci Rep. https://doi.org/10.1038/srep40316

Zhang L, Liu S, Sun P, Wang T, Wang G, Zhang X, Wang L (2015) Consensus forecasting of species distributions: the effects of niche model performance and niche properties. PLoS ONE. https://doi.org/10.1371/journal.pone.0120056

Zhou Y, Ge X, Zou Y, Guo S, Wang T, Zong S (2021) Prediction of the potential global distribution of the Asian long-horned beetle *Anoplophora glabripennis* (Coleoptera: Cerambycidae) under climate change. Agric for Entomol. https://doi.org/10.1111/afe.12461