

Research Article

Predicting the expansion of invasive species: how much data do we need?

Joana Santana^{1,2,3}, Neftalí Sillero⁴, Joana Ribeiro^{1,2,3}, César Capinha^{5,6}, Ricardo Jorge Lopes^{1,3,7}, Luís Reino^{1,2,3}

1 CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, InBIO Laboratório Associado, Campus de Vairão, Universidade do Porto, 4485-661 Vairão, Portugal

2 CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, InBIO Laboratório Associado, Instituto Superior de Agronomia, Universidade de Lisboa, Tapada da Ajuda, 1349-017 Lisboa, Portugal

3 BIOPOLIS Program in Genomics, Biodiversity and Land Planning, CIBIO, Campus de Vairão, 4485-661 Vairão, Portugal

4 CIBIO-Centro de Investigação em Ciências Geo-Espaciais, Spatial Biology Lab, Faculdade de Ciências da Universidade do Porto, 4430-146 Vila Nova de Gaia, Portugal

5 Centro de Estudos Geográficos, Instituto de Geografia e Ordenamento do Território, Universidade de Lisboa, Rua Branca Edmée Marques, Lisboa, Portugal

6 Laboratório Associado Terra, Lisboa, Portugal

7 cE3c, Center for Ecology, Evolution and Environmental Change & CHANGE, Global Change and Sustainability Institute, Departamento de Biologia Animal, Faculdade de Ciências, Universidade de Lisboa, 1749-016 Lisboa, Portugal

Corresponding author: Joana Santana (joanafsantana@cibio.up.pt)

Abstract

Ecological niche models (ENMs) are a powerful tool to predict the spread of invasive alien species (IAS) and support the implementation of actions aiming to reduce the impact of biological invasions. While calibrating ENMs with distribution data from species' native ranges can underestimate the invasion potential due to possible niche shifts, using distribution data combining species' native and invasive ranges may overestimate the invasion potential due to a reduced fitness and environmental tolerance of species in invaded ranges. An alternative may be using the increasingly available distribution data of IAS as they spread their invaded ranges, to iteratively forecast invasions as they unfold. However, while this approach accounts for possible niche shifts, it may also underestimate the species' potential range, particularly at the early stages of the invasion when the most suitable conditions may not yet be represented in the distribution range data set. Here, we evaluate the capacity of ENMs to forecast the distribution of IAS based on distribution data on invaded ranges as these data become available. We further use dispersion models to assess the expansion process using the predicted potential distributions. Specifically, we used the common waxbill (*Estrilda astrild*) in the Iberian Peninsula as a model system. We built ENMs with 10×10 km grid cells distribution records cumulatively for each decade from 1960 to 2019, and yearly bioclimatic variables, to forecast the species potential range in the coming decades. Then, we assessed the performance of the models for each decade in forecasting the species' observed range expansion in the following decades and evaluated how the number of distribution records determined the quality of the forecasts. Finally, we performed dispersal estimates (based on species traits, topography, climate and land cover) to analyse the prediction capacity of models as their uncertainty may be reduced when projecting them to the next decades. Our results show that invasion-only ENMs successfully forecasted the species' range expansion over three decades after invasion, while dispersion models were not important in forecasting common waxbill expansion. Our study highlights the importance of constantly monitoring alien species, suggesting that iterative updating of ENMs with observed distribution data may accurately forecast the range expansion of alien species.



Academic editor: Helen Sofaer

Received: 13 May 2024

Accepted: 5 August 2024

Published: 4 September 2024

Citation: Santana J, Sillero N, Ribeiro J, Capinha C, Lopes RJ, Reino L (2024) Predicting the expansion of invasive species: how much data do we need? NeoBiota 95: 109–132. <https://doi.org/10.3897/neobiota.95.122335>

Copyright: © Joana Santana et al.

This is an open access article distributed under terms of the Creative Commons Attribution License (Attribution 4.0 International – CC BY 4.0).

Key words: Alien species, common waxbill, dispersal analyses, ecological niche models, *Estrilda astrild*, forecasts

Introduction

Biological invasions are among the most worrisome environmental problems in modern times (Díaz et al. 2019). The spread of invasive alien species (IAS) worldwide has been responsible for population declines of native species, changes in community composition (Bellard et al. 2016, 2021), alterations of ecosystem processes and functioning (Ehrenfeld 2010), disruptions of socio-economic activities (Diagne et al. 2021) and public health issues (e.g. Naeem et al. 2009; Fournier et al. 2019; Ogden et al. 2019). In a globalised world, the number of IAS is expected to increase (Seebens et al. 2021), as well as their potential impacts (Fournier et al. 2019; Essl et al. 2020), promoted by the increasing international wildlife trade and global changes (Scheffers et al. 2019; Naimi et al. 2022). As a response to this urgency, several international regulations and mechanisms have been implemented in recent decades to prevent the introduction and spread of IAS. This includes the establishment of a legal framework with specific legislation, such as the EU Regulation 1143/2014 on IAS (Regulation EU 2014). However, the successful implementation of these mechanisms requires the anticipation of new invasion areas which have been hampered by the lack of monitoring data on species distributions at adequate spatial and temporal resolutions (Santana et al. 2023). There is thus a need for a continuing effort to develop approaches, which may include ecological modelling tools, to accurately predict IAS expansion, in order to reduce both the ecological and socio-economic impacts of IAS.

Modelling and projecting the realised niche of IAS in the geographical space allows the identification of the areas at risk of invasion (Jiménez-Valverde et al. 2011; Guisan et al. 2014). The realised niche is part of the fundamental niche, i.e. the abiotic environmental space where a species can maintain a viable population and persist over time without immigration, which is then further limited by biotic interactions, dispersal capacity, or historical aspects (Soberón and Peterson 2005). This assessment is often done through correlative ecological niche modelling (ENM) (Peterson and Vieglais 2001; Jeschke and Strayer 2008; Jiménez-Valverde et al. 2008; Capinha and Anastácio 2011; Venette 2015; Sillero et al. 2021), which quantify species–environment relationships based on observed patterns of species distributions and environmental predictors (Franklin 2010; Peterson et al. 2011; Guisan et al. 2019; Sillero et al. 2021). A procedure of key practical importance concerns the geographical areas used to calibrate the ENMs. For IAS, these models can be calibrated using distribution data from the species' native range (Peterson et al. 2003), thus assuming that the species native distribution represents the entire suite of suitable environments (i.e., distributional equilibrium; Guisan and Zimmermann 2000; Araújo and Pearson 2005; Araújo et al. 2005), or at least, all suitable habitats where the species can disperse (i.e. pseudo-equilibrium; Anderson and Raza 2010; Sillero et al. 2021). However, species' realised niches may shift in new areas or periods (i.e. niche shift *sensu* Guisan et al. 2014), which implies that IAS will not be necessarily circumscribed to areas that are environmentally analogous to their native ranges (Peterson 2003; Jeschke and Strayer 2008; Elith and Leathwick 2009). This is because, when the environmental conditions change, or the species arrives in a new area, the drivers limiting the species' realised niche can change (e.g. the new area lacks a competing species or the species can now disperse to new habitats), enabling the exploration of new areas inside its fundamental niche (Sillero et al. 2022). Some IAS have shown marked climatic niche shifts during invasion (i.e., a divergence between climatic conditions in native and alien

ranges; *sensu* Broennimann et al. 2009), likely driven by adaptive changes enabling species to endure conditions that were previously unsuitable (Blossey and Notzold et al. 1995), i.e., a shift in its realised climatic niche (Sillero et al. 2022).

Considering the potential for realised niche shifts, previous studies have recommended calibrating ENMs using distribution data of IAS in both native and invasive ranges (Fitzpatrick et al. 2006; Broennimann 2007; Broennimann and Guisan 2007; Urban et al. 2007; Beaumont et al. 2009; Pili et al. 2020). While this approach potentially captures niche shifts as they emerge in invaded areas, the combination of native and invasive distribution data raises relevant practical and conceptual issues. The existence of higher-quality distribution data for the species in one range versus the other (Vanette et al. 2010), may require a reduction in spatial resolution, leading to information loss when merging both data frames (Jarnevich et al. 2022). Although spatial downscaling can be employed to enable modelling at a coarse resolution and projection at a higher resolution, this approach introduces uncertainty due to assumptions regarding the consistent relationships between coarse and fine-resolution data within the area, on the employed methods (Keil et al. 2013). On the other hand, and perhaps more importantly, the use of native distribution data may overestimate invasion ranges. This is because there are circumstances where invasive species may not be able to colonise similar environmental conditions to their native areas. Factors such as novel negative interspecific interactions (e.g., predators, parasites, competitors) (Sih et al. 2010; Dostal et al. 2013; Carthley and Banks 2018), genetic bottlenecks and founder effects, can drive a reduction in the species' environmental tolerances, and species dispersal capacity. These factors cannot be included directly in the native-based model, which will provide the maximum extent of the species distribution in the invasive range if the environmental conditions are the same (Jarnevich et al. 2022).

Invasion monitoring efforts are producing high-quality spatiotemporal data of spread for a large number of IAS in invaded ranges (e.g. Groom et al. 2019; Howard et al. 2022). Hence, given the challenge of reconstructing the invasive process over time for most species, an alternative is to use spatiotemporal invasion data to iteratively forecast invasions as they unfold. The issues raised by using native distribution data are overcome by restricting the calibration of ENMs to the region being invaded. However, any approach relying only on invasive distribution data for calibrating ENMs must acknowledge the likely underestimation of species' potential ranges, particularly at early stages of invasion, when most suitable conditions may not yet be represented in the distribution range data set. In this context, it is pivotal to clarify the data requirements to ensure accurate ENM for IAS, and particularly, the extent to which invasion-only distribution data can be used to accurately predict the expansion of IAS. Modelling the invasion over time will provide information about the routes used by the species during the expansion process.

Here, we evaluate the capacity of iterative calibration of ENM models based on invasion-only distribution data to predict the invasion potential and analyse the expansion process of IAS. We specifically explore the relationship between the number of species distribution records since establishment and the capacity of models to inform about the species' invasion potential. We also assessed the importance of accounting for the dispersal capacity of species to predict their expansion. To achieve these aims, we considered one of the most studied and successful avian invasive alien species, established in different environments and biogeographic regions worldwide: the afro-tropical common waxbill (*Estrilda astrild*). This species was first introduced to

Portugal in 1964, and has spread across much of the country and into part of Spain fed by further introductions (Reino and Silva 1998; Silva et al. 2002). Common waxbill is a small-sized (<12 cm) granivorous finch that forages in low grass vegetation, typically found in open agricultural fields near water bodies (Payne et al. 2020; Ribeiro et al. 2020). The invasive success of this gregarious, non-territorial species may be attributed to its breeding biology and interspecific relations (Ribeiro et al. 2020), including: i) a variable breeding season (Sanz-Aguilar et al. 2015; Payne 2020); ii) the ability to produce several broods a year (Burton and Burton 2002); iii) vagrant movements in search of suitable habitat (flight range < 37 km) (Payne 2020); and iv) the lower prevalence of parasites in non-native ranges compared to native regions (Lopes et al. 2018).

Using a unique, high-quality, database on spatial dispersion of the common waxbill through the Iberian Peninsula over six decades (Reino and Silva 1998; Silva et al. 2002; Reino 2005; Sullivan et al. 2012), we applied a backcasting approach, fitting ENMs using distribution data available until the end of each decade and using the resulting model to project the distribution for the next decade. Then, we analysed how the number of observation records used in each ENM was related to the performance of the forecasts of species dispersal over time. Finally, because ENMs do not account for species' dispersal per se (Sillero et al. 2021; Sillero et al. 2022), we also implemented a species dispersal model over time considering a comprehensive set of species traits and climatic and landscape variables (Engler et al. 2012). We discuss these results in light of the amount of distribution data (i.e. length of the time series since establishment) needed for invasion-only ENMs and dispersal analyses to predict the invasion potential of IAS and iteratively forecast future invasions.

Materials and methods

Study area

The Iberian Peninsula (southwestern Europe), covers an area of 582,860 km² and mainly includes the continental territories of Spain and Portugal (Fig. 1). It is bordered to the southeast and east by the Mediterranean Sea and to the south, north and west by the Atlantic Ocean, and is separated from the rest of Europe by the Pyrenees in the northeast. The Peninsula has a high diversity of climatic conditions, influenced by both the Atlantic Ocean and the Mediterranean Sea, with a longitudinal gradient of precipitation and a latitudinal gradient of precipitation and temperature (Capel 1981). While the Mediterranean climate dominates most of the area, the Oceanic climate occurs primarily in major mountain ranges and isolated southern mountains (Sillero et al. 2009).

Common waxbill distribution data

We gathered historical data on the common waxbill expansion in the Iberian Peninsula since its first introduction in the 1960s. For this, we obtained presence data of the species in the continental territories from Sullivan et al. (2012), including the national and regional breeding bird atlases from Portugal and Spain, updated with all-year-round information from the eBird database (eBird 2019) and authors unpublished data (L. Reino). Presumed new human-mediated introductions were considered whenever an isolated presence of the species was identified or when introduction events were documented through publications or personal

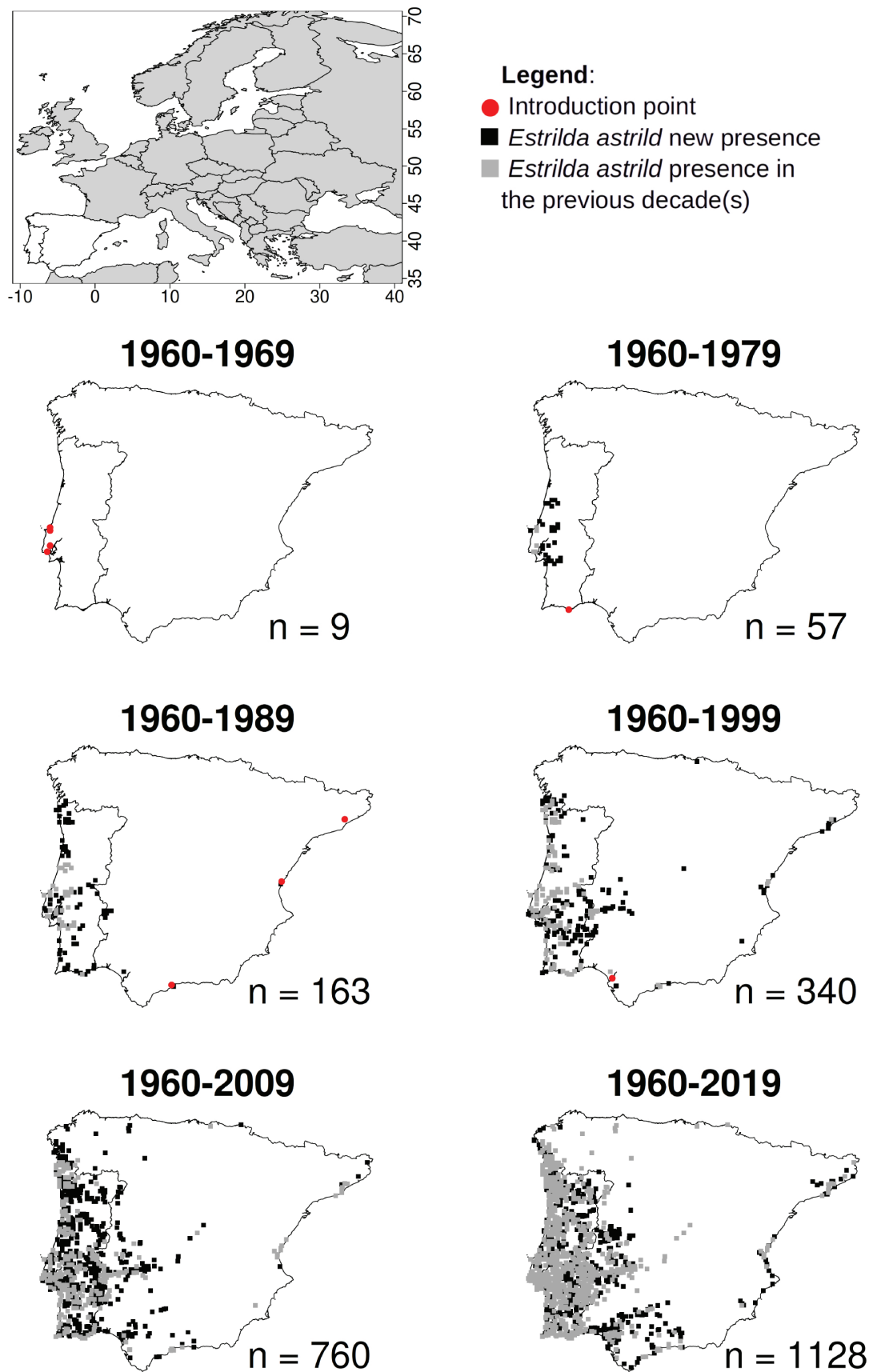


Figure 1. Location of the study area in the Iberian Peninsula, southwest Europe (upper panel) and distribution of the common waxbill *Estrilda astrild* in each decade from 1960 to 2019 (bottom panel) and the accumulated number of presences for each decade (n = #, bottom right of each map).

communications. Following data compilation, we mapped all records onto a 10×10 UTM km grid of Portugal and Spain (Fig. 1). We then aggregated the data by decade, from 1960 to 2019. Therefore, we represented the species expansion over six decades: 1st - 1960–1969 (9 UTM grid cells), 2nd - 1970–1979 (57 cells), 3rd - 1980–1989 (163 cells), 4th - 1990–1999 (340 cells), 5th - 2000–2009 (760 cells) and 6th - 2010–2019 (1128 cells) (Fig. 1, Suppl. material 1).

Environmental data

We obtained yearly climate data for the temporal period covered by the distribution data from the EuMedClim Database (<http://gentree.data.inra.fr/climate/>; Fréjaville and Garzón 2018), which provides yearly climate data between 1901–2014 at 1 km resolution for Europe and the Mediterranean Basin. We considered the seven bioclimatic variables available from this source: bio1 - annual mean temperature; bio2 - mean diurnal temperature range; bio5 - Maximal temperature of the warmest month; bio6 - minimal temperature of the coldest month; bio12 - annual precipitation; bio13 - precipitation of the wettest month; bio14 - precipitation of the driest month). From these variables, to minimise cross-correlation between variables, we kept four variables that had an absolute value of Pearson correlation coefficient below 0.7 (Suppl. material 2: table S1, fig. S1): mean diurnal temperature range (bio2), minimal temperature of the coldest month (bio6), precipitation of the wettest month (bio13) and precipitation of the driest month (bio14). As our species occurrences were available at 10×10 UTM km grid squares, we aggregated the values of these variables to match this resolution.

Ecological niche models

We estimated the realised niche of the species (*sensu* Sillero 2011) every decade, considering, for each model, the species distribution data cumulatively since the first introduction in 1960 (Fig. 1, Suppl. material 2: fig. S2). For this, we assumed that there were no extinctions, which is a reasonable expectation given the 10×10 km resolution, allowing us to use the climate data from the corresponding decade. Then, we projected each model to the following decades (Suppl. material 2: fig. S2). For example, we used the model calibrated with data from the first decade (1960–1969) to project the species' potential distribution in each of the following decades, i.e., 1970–1979, 1980–1989, 1990–1999, 2000–2009 and 2010–2019.

We calculated realised niche models using Maxent v.3.4.4 (Phillips et al. 2006, 2017) following standard procedures (Sillero and Barbosa 2021; Sillero et al. 2021). All Maxent models used the same parameterization. Specifically, we used presence/background data representing the spectrum of environmental conditions available to the species as dependent variables (Phillips et al. 2009; Guillera-Arroita et al. 2014). Background data consisted of 9196 randomly generated points distributed throughout the entire study area, including pixels where the species occurred. Thus, background points are not equivalent to pseudo-absences (Phillips et al. 2009; Guillera-Arroita et al. 2014). Maxent output represents habitat suitability, ranging from 0.0 (not suitable) to 1.0 (suitable), in Cloglog format (Phillips, et al. 2017). Because Maxent includes stochasticity in the training data random selection, different model runs can lead to slightly different outcomes (Phillips et al. 2006, 2017). For this reason, we used the average of 10 distinct modelling events to obtain the final suitability

values for each decade, randomly selecting 70% of the occurrence records as training data and 30% as test data. The models used auto features, where different distribution functions are used depending on the sample size (Phillips et al. 2006, 2017).

We measured model discrimination performance using the area under the curve (AUC) of the receiver operating characteristics (ROC) plots (Liu et al. 2005) and True Skill Statistics (TSS; Allouche et al. 2006). The ROC plot was calculated by representing the sensitivity against 1-specificity for all possible thresholds. It measures the proportion of true positives against the proportion of false positives, i.e., the likelihood that the model will rank a randomly chosen presence higher than a randomly chosen absence. The AUC is an integral of the ROC curve. The AUC discriminates a species' model from a random model, with a value equal to or close to 0.5 corresponding to an accuracy similar to that of a random model and a value of 1 corresponding to a perfect discrimination accuracy. TSS is equal to sensitivity + specificity - 1 (Allouche et al. 2006). TSS ranges from - 1 to 1, with 0 corresponding to random classification. AUC and TSS increase with the extent of the study area and are correlated (Leroy et al. 2018). Additionally, we calculated a set of null models following the methodology by Raes and ter Steege (2007). For this, we generated 100 different datasets with the same number of random points as each dataset following a Poisson distribution. We calculated a Maxent model for each of these random datasets and obtained the AUC values of the ROC plots. Then, we compared the training AUC values of the species models with the ones calculated for the null models using the non-parametric Wilcoxon test. We calculated the null models in R 3.4.4 (R Core Team 2020) with the 'dismo' package (Hijmans et al. 2017). As AUC and TSS are correlated (Leroy et al. 2018), we calculated the null models using only the AUC.

We used the Boyce Index (Boyce et al. 2002; Hirzel et al. 2006) to measure the level of agreement between predictions generated by the model and actual observations (proportion of presences). The Boyce Index is specifically designed for presence-background algorithms and measures the extent to which model predictions differ from a random distribution of the observed presences across the prediction gradients (Hirzel et al. 2006). The index calculates the proportion between the number of presences predicted by the model that fall into a particular class and the number of presences expected under a random classifier, plotted against the corresponding suitability category. The habitat suitability range is divided into b classes (or bins) to calculate a predicted-to-expected (P/E) ratio (F_i), where P is the predicted frequency of evaluation points and E is the expected frequency of evaluation points. If the model is well calibrated, low suitability classes should contain fewer evaluation presences than expected by chance ($F_i < 1$) and high suitability classes should contain more evaluation presences than expected by chance ($F_i > 1$). A good model should have a monotonically increasing curve when P/E is plotted against the mean habitat suitability of each class. The index ranges between -1 and +1: positive values indicate a model where predicted presences are consistent with the distribution of presences in the evaluation dataset; values close to zero mean that the model is not statistically different from a random model; negative values indicate counter predictions, i.e., predicting lower suitability in areas where presences are more frequent. We calculated the Boyce index with ModEvA 3.10 R package (Barbosa et al. 2013).

We determined the contribution of each climatic variable in explaining the species' distribution using a jackknife resampling based on: (1) values of the training and test gain; and (2) of AUC values. The jackknife resampling comprises two

steps: (1) the generation of a model with all climatic variables except one; and (2) the generation of univariate models, each using only one climatic variable. In each step, the jackknife analysis measures the change in training and test gain, and the AUC determines the importance of each variable. Using the results from each of these procedures, Maxent calculated an average percentage contribution of each climatic variable. We also calculated the permutation importance: for each environmental variable in turn, the values of that variable on training presence and background data are randomly permuted. The model is re-evaluated on the permuted data, and the resulting drop in training AUC is calculated, and normalised to percentages (Phillips et al. 2006). When variables interact, the variable contributions and the permutation importances are not equally ordered, preventing individual responses for each variable.

Ecological niche models validation over time

We further validated the ENMs for each decade and their respective projections by counting the number of presences classified as presences or as absences. For this, we categorised the continuous models into two categories by applying the threshold ‘Maximum training sensitivity plus specificity’ for the Cloglog output. We used the presence records of each decade and previous decades, i.e., cumulatively. The total number of presences used to validate the projections of each model was the same (9, 114, 160, 334, 752, 1120, see Suppl. material 2: table S4). We also validated the projections of all the models using the presences from the last decade (i.e. 2010’s; 1120 presences, see in Suppl. material 2: table S5).

Dispersal analyses

Accounting for dispersal barriers/capacity has been pointed out as important to reduce uncertainty in future projections of species distribution (Engler et al. 2012). We estimated dispersal movements over time with the R package ‘MigClim’ (Engler et al. 2012), a cellular automaton model that simulates the dispersal of species in the landscape. MigClim uses ENMs as indicators of landscape permeability: the higher the habitat suitability index, the higher the permeability. We applied MigClim to each decadal Maxent model and respective projections (Suppl. material 2: fig. S2). Therefore, Migclim modelled how the species dispersed between Maxent models, i.e. from the first model to the next projections. The MigClim model considers short-distance and long-distance dispersal events, the type of dispersion through the landscape (using a continuous or a categorical Maxent model), propagule production probability, initial maturity age, and the presence of barriers. We considered three possible scenarios: i) no dispersal barriers; ii) weak barriers (i.e. barriers that can be transposed), and iii) strong barriers (i.e. barriers that cannot be transposed) (Table 1). We assessed the significance of dispersion models in forecasting the expansion of the common waxbill by visually inspecting the maps produced for each scenario.

Dispersal barriers were represented by elevation, land cover, and hydrological factors (Fig. 2). We obtained elevation data from the Shuttle Radar Topography Mission (<https://www.earthdata.nasa.gov/sensors/srtm>; Farr et al. 2007) at 90 m and aggregated it to 10×10 km cell resolution. Following the results obtained in previous studies in Iberia (e.g. Silva et al. 2002), we considered elevations higher

than 800 m as barriers to dispersal. We obtained land cover from the Global Land Cover 2000 dataset with 250 m of spatial resolution from the European Environmental Agency (<https://www.eea.europa.eu/data-and-maps/figures/global-land-cover-2000-250m>). We considered the land cover classes including tree cover as barriers to the dispersal, as *E. astrild* is mostly associated with open habitats (Payne et al. 2020; Ribeiro et al. 2020) (Suppl. material 2: table S2). To represent hydrological configuration, we considered the vectorial dataset of free-flowing rivers in Europe from the European Environmental Agency (<https://www.eea.europa.eu/data-and-maps/figures/free-flowing-rivers-in-europe>). River hierarchy ranged from 1 to 8, with a value of 8 corresponding to large watercourses such as main rivers and 1 to small, often intermittent streams. We calculated the average hierarchy of watercourses in each grid cell of 10×10 km. For this variable, we considered as barriers those grid cells with a hierarchy of 1, reflecting the species' association with permanent water courses (Ribeiro et al. 2020). The final layer of barriers corresponded to the combination ((elevation + land cover) x river hierarchy) of all barrier layers, resulting in a layer classified as 0 (without dispersal barriers) or 1 (with barriers) (Fig. 2). The parameters used in MigClim are shown in Table 1.

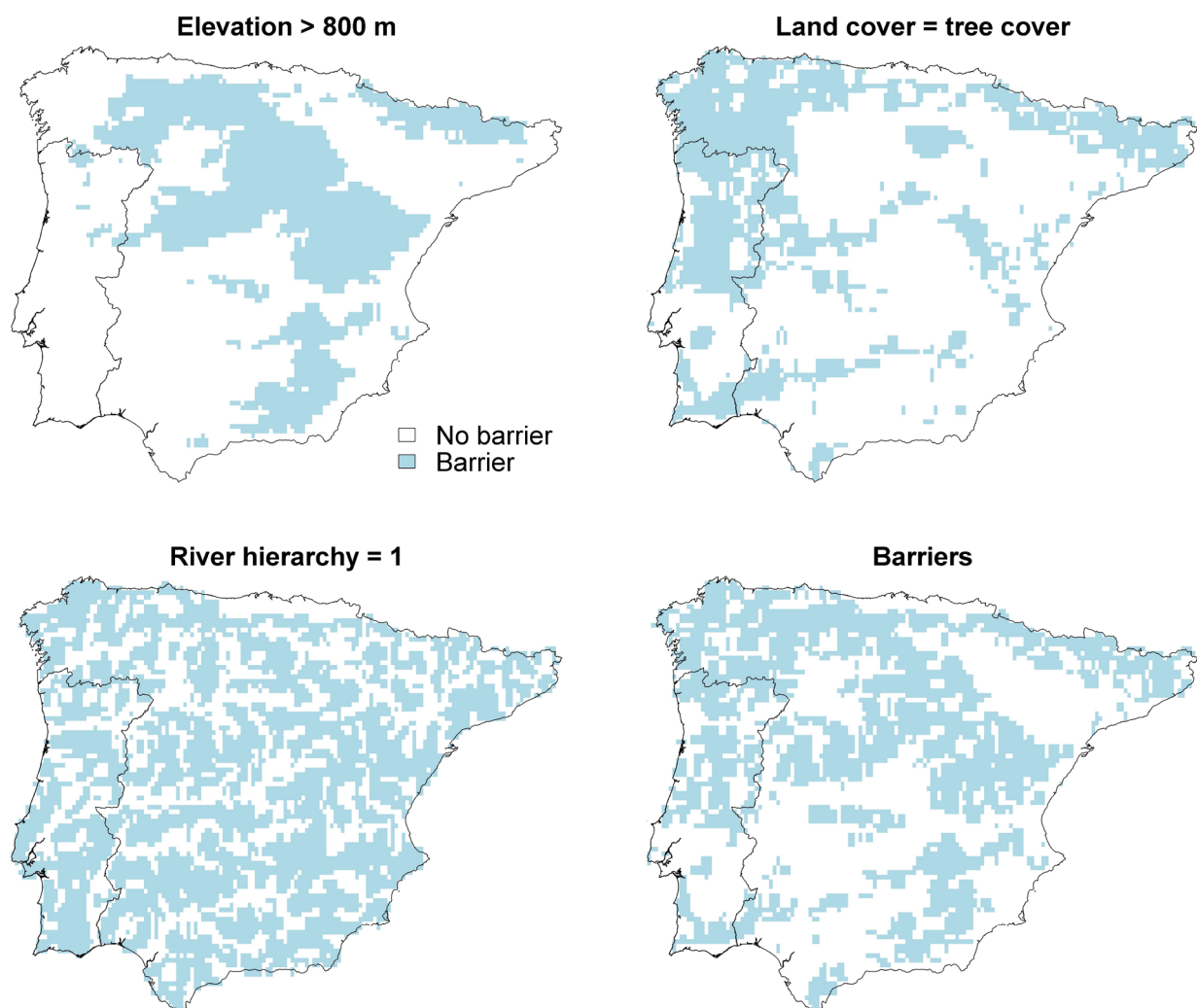


Figure 2. Variables (elevation, land cover, and river hierarchy) used to define the barriers to dispersal through the landscape, and the barriers used in MigClim to measure the dispersion across the landscape. Blue means barrier; white means no barrier.

Table 1. Parameters and values used in dispersal models.

Parameter	Value
Continuous mode	0
Number of environmental change steps to perform	6
Number of dispersal steps to perform within each environmental change step	10
Dispersal kernel: probability of colonising a directly adjacent cell	1
Long-distance dispersal frequency	0.0001
Minimum distance for long-distance dispersal in pixels	2
Maximum distance for long-distance dispersal in pixels	4
Initial maturity age of newly colonized cells	1
Propagule production probability as a function of cell age	1
Number of replicates	100
Barriers	No barriers / Weak / Strong

Results

Common waxbill expansion patterns

The current geographic distribution of the common waxbill in the Iberian Peninsula spans most of the Iberian Atlantic coast, but also through large areas in Southern and Eastern Iberia extending to the Mediterranean coast, as far as Catalonia (Fig. 1). Although this species was initially introduced around the Lisbon region (Central Portugal) during the 60's, it rapidly spread in all directions during the following decade. Its geographic expansion was also enhanced by more recent and independent introductions, as in the Algarve (southern Portugal), during the 70's, and Andalusia (southern Spain) and later in other regions in eastern Spain, during the 80's (Fig. 1).

The expansion process was faster in the Central and Northern regions of Portugal (70's), whereas the spread in the south seemed to have been boosted by an additional introduction event in the Algarve in the same decade that enabled the colonisation of western Andalusia during the following decade (Fig. 1). This event appears to underlie the colonisation of almost all Iberian southwestern coasts, unifying the northern population - which had also started to colonise Alentejo (Portugal) from the Sado River valley in the 80's - with the southern population (Fig. 1). Spread eastwards was slower, and several areas remain uncolonized to the present, namely the mountainous regions of Northern and Central Spain (Fig. 1). Spanish populations have arisen from independent introductions and the expansion of Portuguese populations directly from the extensive area bordered by the Minho River in Galicia in the 80's, and through the Tagus and Guadiana River valleys in the Southern and Central regions by the 90's (Fig. 1).

Temporal changes in ecological niche projections

ENMs had test AUC values higher than 0.8 and significantly differed from random (Table 2, Suppl. material 2: table S3). All test TSS values were higher than 0.6 (Table 2). The Boyce index exceeded 0.85 for all decades, except the first (0.02). This indicates a consistent prediction of presences by the models, aligning closely with the distribution of presences in the evaluation dataset for all but the first decade, to which the model is not statistically different from random.

The variable with the highest contribution for the first decade (1960–1969), was the mean diurnal temperature range (bio2), whereas the minimum temperature of

the coldest month (bio6) was the variable with the highest contribution in the following decades (Table 3). The permutation importance of variables maintained the same order as the variable contributions (Table 3).

The areas identified as suitable widened over time, from the coastal areas towards the interior of the Iberian Peninsula (Fig. 3). Suitable areas based on data from the first decade (1960–1969) were enclosed within the vicinity of the introduction area around Lisbon (Portugal, Fig. 3). There was an abrupt change in the extent of suitable areas between the third and the fourth decades (1980–1989 and 1990–1999, Fig. 3). The increment in suitable areas stabilised in the fourth decade after introduction (1990–1999, Fig. 3).

Ecological Niche Models validation

Validation of the ENMs of each decade projected to the remaining periods (Fig. 4, Suppl. material 2: table S4) indicated the proportion of the presences correctly classified is >80% for the projections of all models, except for the projections for the last decade based on the models using distribution data of the first and second decades (44% and 50%, respectively). The same pattern was observed in the models using data from the past decade, so the results are presented only in the Suppl. material 2: fig. S3, table S5.

Table 2. Results of AUC, TSS, and Boyce index for models considering the accumulated presences since the first introduction.

Decade	Training n	Test n	AUC	TSS	Boyce index
1960–1969	4	1	0.99 ± 0.010	0.83 ± 0.09	0.06
1960–1979	24	10	0.98 ± 0.010	0.91 ± 0.03	0.86
1960–1989	73	31	0.95 ± 0.010	0.80 ± 0.03	0.98
1960–1999	171	73	0.92 ± 0.010	0.71 ± 0.02	0.99
1960–2009	432	184	0.88 ± 0.010	0.63 ± 0.01	0.99
1960–2019	639	273	0.86 ± 0.010	0.60 ± 0.01	0.99

Table 3. Contributions and permutation importance of the bioclimatic variables (mean diurnal temperature range (bio02), minimal temperature of the coldest month (bio06), precipitation of the wettest month (bio13) and precipitation of the driest month (bio14) of the Maxent models. Highest values of variable contribution and permutation importance for each model are highlighted in bold. Temperature, and particularly the minimal temperature, was the most important variable affecting the distribution of common waxbill for all models.

Decade	Variable contribution				Permutation importance			
	bio02	bio06	bio13	bio14	bio02	bio06	bio13	bio14
1960–1969	50.97	19.75	0.06	29.22	39.97	3.74	0.14	56.15
1960–1979	4.59	72.24	6.58	16.60	0.90	84.85	4.54	9.710
1960–1989	2.76	79.67	7.62	9.95	1.18	80.53	5.82	12.47
1960–1999	6.06	76.36	8.40	9.18	5.74	78.30	10.46	5.501
1960–2009	4.90	68.36	15.33	11.41	5.37	64.04	21.40	9.19
1960–2019	5.14	67.70	15.10	12.06	5.69	62.53	17.63	14.16

Dispersal analyses

The species' potential range accounting for dispersal capacity increased over time, driven by the results of ENM projections. In the first decade, the range deemed susceptible to colonisation was narrow, and almost all of the Iberian Peninsula was beyond reach. On the other hand, for the last decade, these areas were much wider (Fig. 5). Similarly

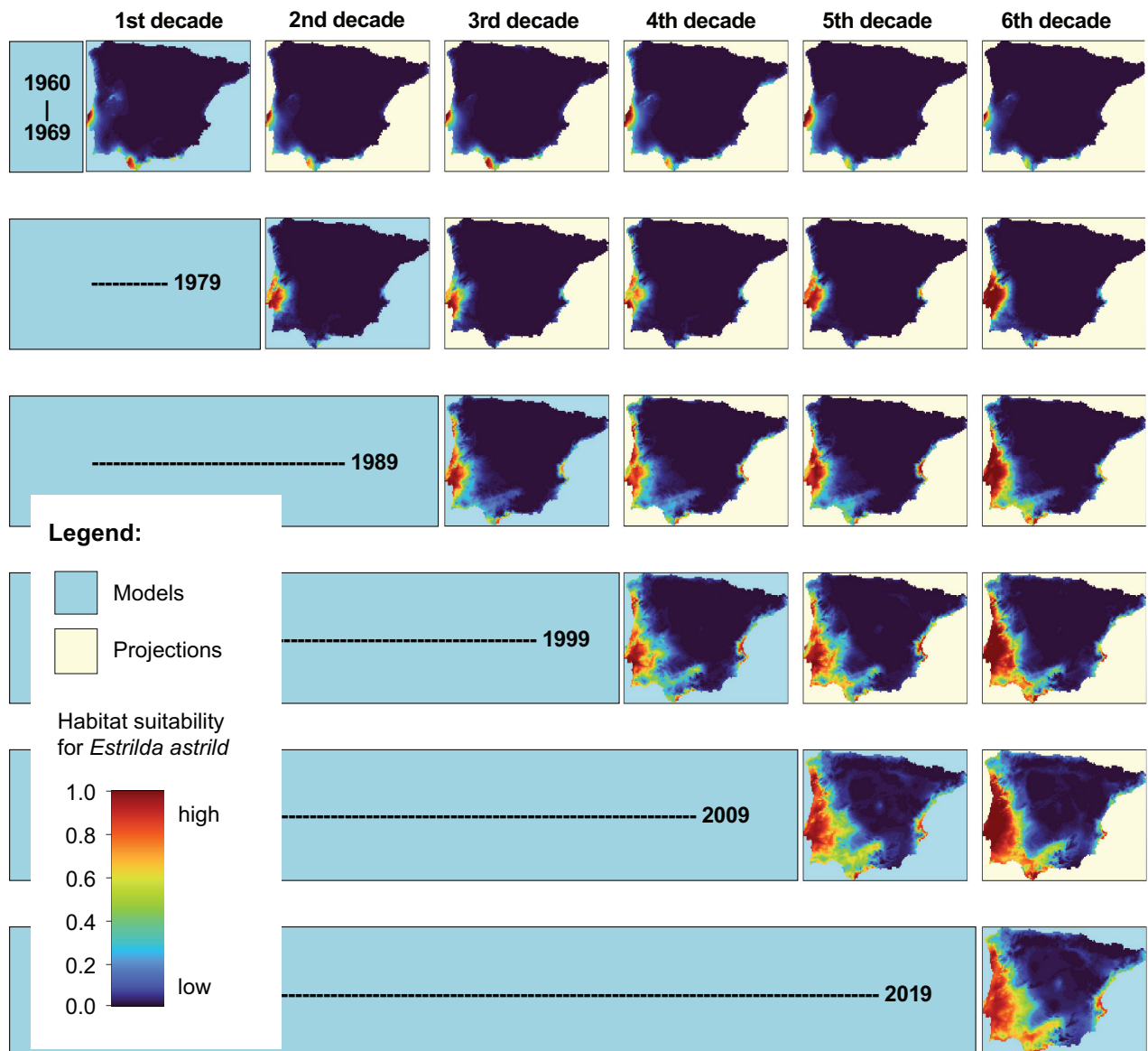


Figure 3. Results of cumulative ecological niche models projecting habitat suitability to the next decades. The suitability maps are organised as in Suppl. material 2: fig. S2: each decade is a row; models (light blue background) and projections (yellow background) are placed in columns. Habitat suitability ranges from dark blue (low suitability) to red (high suitability), following the turbo palette.

to what was verified for the ENM, the spatial patterns of the dispersal models have remained quite stable since the fourth decade (1990–1999). Visual inspection of the maps reveals that there were some differences when barriers were introduced in the dispersal models compared to when they were not (Fig. 5), but the results with weak or strong barriers are very similar. Dispersal models thus confirmed that the species was able to disperse over time following suitable areas identified by ENM.

Discussion

This study evaluates the use of increasingly available spatiotemporal data on IAS spread to iteratively forecast invasions as they unfold. The backbone of these forecasts were ENMs using detailed distribution data of the common waxbill expansion through the Iberian Peninsula over six decades. Our projections, based on invasion-range data, were

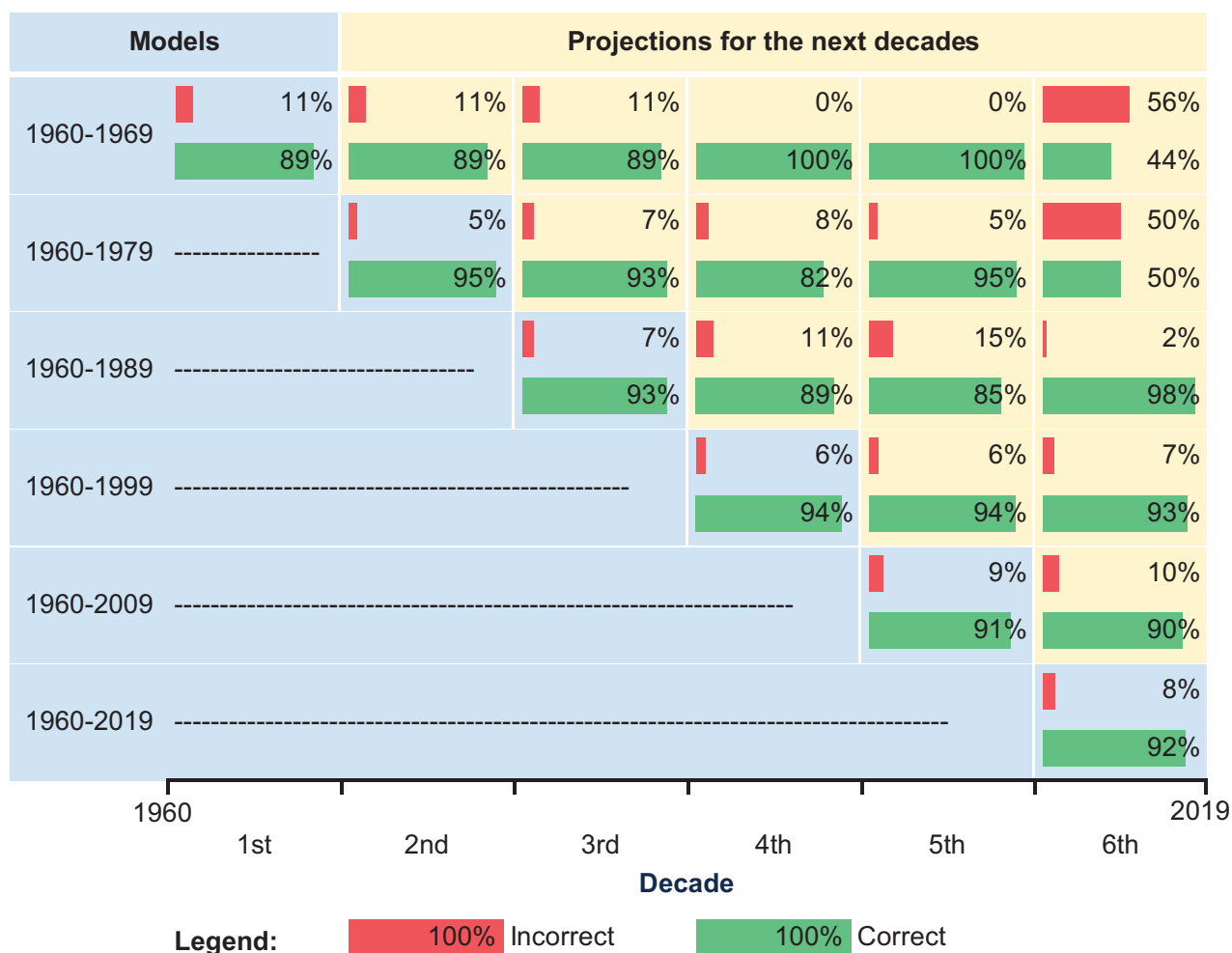


Figure 4. Percentage of presences incorrectly (red) and correctly (green) classified over time, for each model (blue background) and projection (yellow background). Validation for each model was conducted using the number of cumulative presences from the previous decade(s) (for details see Suppl. material 2: table S4). Our backcasting approach showed a high forecast capacity of EMNs after the 3rd decade following the common waxbill establishment (high % of correctly classified presences).

successful in forecasting the species’ current distribution after three decades following its introduction. These results support the idea that ENMs can successfully forecast the species’ range expansion, although they may have limited utility in the early stages of invasion, supporting the use of an iterative approach (Dietze et al. 2018), where models are recurrently updated with the species’ most recent distribution data. This implies the need to constantly monitor IAS, which despite receiving long-standing support in invasion science and policy (Genovesi and Shine 2004; Büyüktaktın and Haight 2018), depends strongly on the availability of resources (Groomet al. 2019).

Our results are in line with previous studies arguing that ENMs may underestimate the species’ potential ranges (Liu et al. 2020), particularly at the early stages of the invasion when the most suitable conditions may not yet be represented in the distribution range dataset. The common waxbill is a tropical species sensitive to low temperatures (<15 °C) and wet weather conditions and inhabits open fields with tall grasses, often near water (Ribeiro et al. 2020). While detailed information on the distribution data in the native area is not available for analysis, existing data suggest that the distribution range of common waxbill in the invaded area generally falls within the climatic variation of its native African range, with

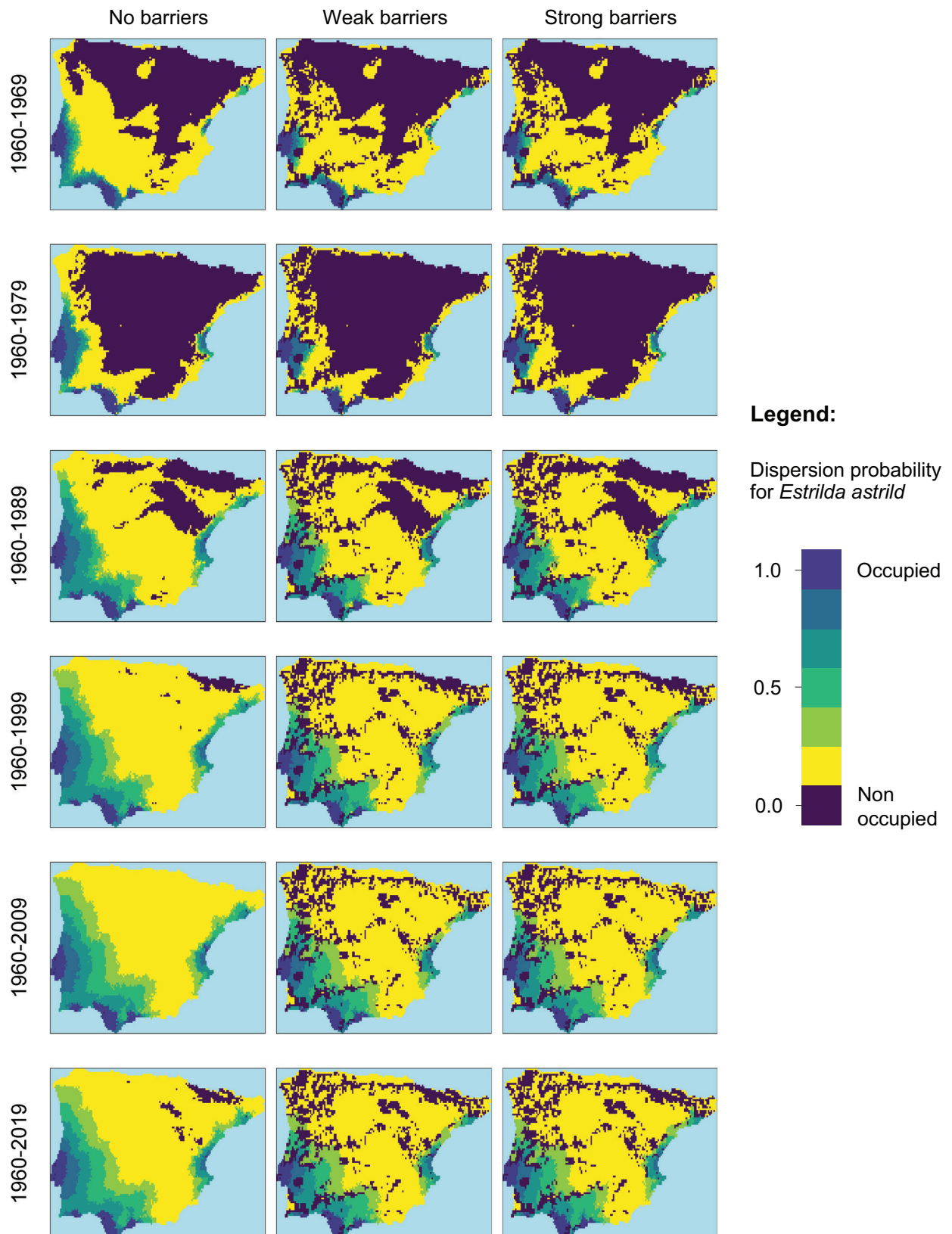


Figure 5. Results of dispersal models per decade and type of barriers (no barriers, weak barriers, and strong barriers). The colour sequence Blue -> Green -> Light Green -> Yellow indicates the dispersal of species over time in each decade. Yellow indicates areas where the species did not have time to arrive. Purple indicates areas where the species cannot occur because habitat suitability predicted by ENMs was low.

expansions observed into colder and rainier areas in Northern Iberia in recent decades (Ribeiro et al. 2020). Our projections based on invasion-only data failed to forecast the current species distribution using data from the first two decades after species introduction, likely because the species range was still not representative of the species' suitable environmental conditions (Araújo and Pearson 2005). The actual niche overlap of the species is probably contributing to the model accuracy of the last decades. However, the AUC has slightly decreased in the models for the last three decades probably due to the increment in the presences: as the species distribution range increases, the species has a more generalist character, making it harder to predict (Guisan and Thuiller 2005; Sillero et al. 2021).

Models disregarding the species' global distribution provide worse results than full distribution models (Barbet-Massin et al. 2010; Capinha et al. 2011; Jarnevich et al. 2022). This is because ENM algorithms assume that the species distribution data used is a good representation of the species' environmental requirements (Sillero et al. 2021). In other words, the algorithm assumes that the data used represents a species in equilibrium with the environment, i.e. the species occupies all available suitable habitats where it can disperse (Guisan and Thuiller 2005; Anderson and Raza 2010; Sillero et al. 2021). Distributional data representing only a portion of a species' global range may fail to capture all the suitable conditions where the species can thrive, potentially leading to an underestimation of its potential range. Therefore, modelling the realized niche of an expanding alien species presents significant challenges (Ficetola et al. 2005). It is expected that the ENM for a particular period will fail to forecast imminent range expansion stages, although this does not mean that the ENM is inaccurate (Barbet-Massin et al. 2018). The ENM for that particular period can be accurate, but the ENM does not have enough information to predict the upcoming dispersion process. This was our case: the increment of new suitable areas in ENM projections stabilised from the third decade after introduction (1990–1999), i.e., only after three decades of dispersion, the species' occurrence data were representative of the species' environmental requirements. The abrupt change between the third and fourth decades might be strongly influenced by new introduction events in the third decade that occurred in regions environmentally different from previous ones, particularly in the East of the Iberian Peninsula. These results thus suggest that modelling expansion based on the early stages of introduction may provide limited results, demanding the interactive recalibration of models as new distribution data becomes available.

Contrary to expectations, our results suggest that barriers to dispersion were not insuperable by the common waxbill, although they might be important for other species with lower dispersal capacity. MigClim considers long dispersal events where the species reaches new locations without human intervention. In that case, the species only needs to arrive at a pixel with enough suitability. The few differences found in the projections using the ENM-only and dispersion models (with strong and weak barriers) indicate that the species was able to disperse over time following the suitable areas predicted by ENMs. In line with the commonly observed lag period during the invasion process (generally attributed to the exponential growth process, stochastic extinction of propagules, or an evolutionary modification of species following establishment, Sakai et al. 2001), previous studies have suggested that the dispersion capacity of the common waxbill across Iberia was very low in the first decades after the

first introduction (e.g., Silva et al. 2002; Reino 2005). Justifications for this lag period for the common waxbill relied on the stochastic extinction of propagules due to the absence of favourable habitat conditions (i.e. agricultural fields near water bodies, Ribeiro et al. 2020) out of the areas where the species was first introduced. This was based on the slower colonisation process in the southern regions of Portugal, where initial populations were very small and limited to the Tagus valley around Lisbon and the westernmost region of Iberia (Portugal) and an acceleration after the 80s. According to these studies, the additional introductions across Iberia might have jointly fostered higher dispersion rates to new areas, suggesting that the dispersal capacity of the common waxbill in Iberia is a combination of both habitat suitability and propagule pressure. However, our results suggest that this is likely to be a consequence of insufficient data to capture the species' environmental requirements during the first decades, as they projected a potential for expansion lower than the real one.

While the common waxbill has been a highly successful avian invader across various continents and islands, its spatial spread appears more limited in comparison to other, predominantly older, Palearctic invaders worldwide, such as the European starling (*Sturnus vulgaris*) or the house sparrow (*Passer domesticus*). However, it seems to show a rather eclectic adaptation as these last two species colonise a great diversity of open and semi-open habitats, but not limited to human-made habitats (e.g., agricultural habitats, gardens), but also to wetlands. Probably, its expansion is more comparable with a more recent invader: red-billed leiothrix (*Leiothrix lutea*) in Europe. Though this species is more associated with forest habitats and is likely to be more limited to tree-based habitats, it has been a very successful and established invader with populations in several European countries and regions (Pereira et al. 2019). It is worth noting that both the European starling and the house sparrow have much older introductions, with established populations in regions like North America and Australia. For instance, the European starling was successfully introduced to multiple areas during the same period, as in Australia (Stuart et al. 2023). Variances in propagule pressures may offer a partial explanation for differences in range expansion.

Conclusions

Accurately anticipating the expansion of IAS is key to ensuring the successful implementation of preventive and mitigation actions. Forecasting invasions by means of different quantitative methods and modelling strategies have been used in the last decades, and new approaches are constantly emerging (Peterson 2003; Reino et al. 2009; Jiménez-Valverde et al. 2011). However, predictions may be severely compromised by different methodological options and their specific limitations. ENMs are powerful tools for predicting the spread of IAS and guide management. Although ENM enables predicting and evaluating biological invasions, it is often compromised by the amount (time-series length), quality (spatial and temporal resolution) and availability (data access) of distribution data in both native and invaded ranges. Our study evaluates the capacity of ENMs based on spatiotemporal data of invaded ranges only to forecast the potential distribution of IAS using common waxbill dispersion analysis as a case study. We demonstrate that invaded range-only data may be used to accurately project the expansion of alien species in novel regions if enough time (at least three decades in our study model) is given to allow the species to expand and occupy the most suitable conditions. These results indicate that the invaded ranges-only ENMs are of limited utility in the early stages of invasion, while they clarify the need

for using an iterative approach where models are recurrently updated with the most recent distribution data of the species since establishment. This approach will contribute to a better understanding of climatic niche changes during the expansion process of alien species, and offers a solution to managers and scientists dealing with the scarcity and asymmetry of distribution data available for alien species worldwide, in their native and invaded ranges. Our study helps solve a much-discussed conundrum and offers a practical solution to better guide management actions and significantly improve stakeholders' ability to halt biological invasions.

Acknowledgments

We thank Victor Encarnação from CEMPA/ICNF for providing information on the common waxbill ringing recaptures in Portugal. We are also thankful to Martin Sullivan for his help with data and initial discussions, and to the subject editor and three anonymous reviewers for their valuable comments and suggestions that helped improve the manuscript.

Additional information

Conflict of interest

The authors have declared that no competing interests exist.

Ethical statement

No ethical statement was reported.

Funding

JS was financed by the FEDER Funds through the Operational Competitiveness Factors Program-COMPETE and by National Funds through FCT-Foundation for Science and Technology within the scope of the project "PTDC/ BIA-EVL/30931/2017-POCI-01-0145-FEDER-030931" and the project EuropaBON (Grant agreement No. 101003553, EU Horizon 2020 Coordination and Support Action). NS and LR are supported by FCT - Fundação para a Ciência e a Tecnologia public institute (IP), under the Stimulus of Scientific Employment: Individual Support contract no. CEECIND/02213/2017 and CEECIND/00445/2017, respectively. CC acknowledges support from FCT through support to CEG/IGOT Research Unit (UIDB/00295/2020 and UIDP/00295/2020). RJL through an FCT - Transitory Norm contract [DL57/2016/CP1440/CT0006]. This work was supported by the European Union's Horizon 2020 Research and Innovation Programme under the Grant Agreement Number 857251. The authors also acknowledge research support via national funds from the Portuguese Foundation for Science and Technology (FCT) in the scope of the project PTDC/BIA-ECO/0207/2020 (DOI: 10.54499/PTDC/BIA-ECO/0207/2020).

Author contributions

Conceptualization: Joana Santana, Luís Reino; Data curation: Joana Ribeiro, Luís Reino, César Capinha, Neftalí Sillero; Formal analysis: Neftalí Sillero; Funding acquisition: Luís Reino, Joana Ribeiro, Joana Santana; Investigation: Joana Santana, Luís Reino, César Capinha, Neftalí Sillero, Joana Ribeiro, Ricardo Jorge Lopes; Methodology: Neftalí Sillero, Joana Santana, César Capinha, Luís Reino, Joana Ribeiro, Ricardo Jorge Lopes; Project administration: Luís Reino; Validation: Joana Santana, Neftalí Sillero; Visualization: Joana Santana, Neftalí Sillero; Writing - original draft: Joana Santana, Neftalí Sillero, Luís Reino; Writing - review and editing: Joana Santana, Neftalí Sillero, César Capinha, Joana Ribeiro, Luís Reino, Ricardo Jorge Lopes.

Author ORCIDs

Joana Santana  <https://orcid.org/0000-0002-4100-8012>

Neftalí Sillero  <https://orcid.org/0000-0002-3490-3780>

Joana Ribeiro  <https://orcid.org/0000-0002-3755-5785>

César Capinha  <https://orcid.org/0000-0002-0666-9755>

Ricardo Jorge Lopes  <https://orcid.org/0000-0003-2193-5107>

Luís Reino  <https://orcid.org/0000-0002-9768-1097>

Data availability

EuMedClim Database: <http://gentree.data.inra.fr/climate/>. Shuttle Radar Topography Mission: <https://www.earthdata.nasa.gov/sensors/srtm>. Global Land Cover 2000 dataset: <https://www.eea.europa.eu/data-and-maps/figures/global-land-cover-2000-250m>. Free-flowing rivers in Europe: <https://www.eea.europa.eu/data-and-maps/figures/free-flowing-rivers-in-europe>. Estrilda astrild occurrence data: Suppl. material 1. R Code used to run the analyses: Suppl. material 3.

References

- Allouche O, Tsoar A, Kadmon R (2006) Assessing the accuracy of species distribution models: Prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology* 43(6): 1223–1232. <https://doi.org/10.1111/j.1365-2664.2006.01214.x>
- Araújo MB, Pearson RG (2005) Equilibrium of species' distributions with climate. *Ecography* 28(5): 693–695. <https://doi.org/10.1111/j.2005.0906-7590.04253.x>
- Araújo MB, Pearson RG, Thuiller W, Erhard M (2005) Validation of species–climate impact models under climate change. *Global Change Biology* 11(9): 1504–1513. <https://doi.org/10.1111/j.1365-2486.2005.01000.x>
- Atlas E (2008) Atlas das Aves Nidificantes em Portugal (1999–2005). Instituto da Conservação da Natureza e da Biodiversidade, Sociedade Portuguesa para o Estudo das Aves, Parque Natural da Madeira e Secretaria Regional do Ambiente e do Mar. Assírio & Alvim, Lisboa.
- Barbet-Massin M, Thuiller W, Jiguet F (2010) How much do we overestimate future local extinction rates when restricting the range of occurrence data in climate suitability models? *Ecography* 33(5): 878–886. <https://doi.org/10.1111/j.1600-0587.2010.06181.x>
- Barbet-Massin M, Rome Q, Villemant C, Courchamp F (2018) Can species distribution models really predict the expansion of invasive species? *PLoS ONE* 13(3): 1–14. <https://doi.org/10.1371/journal.pone.0193085>
- Barbosa AM, Real R, Muñoz AR, Brown JA (2013) New measures for assessing model equilibrium and prediction mismatch in species distribution models. *Diversity & Distributions* 19(10): 1333–1338. <https://doi.org/10.1111/ddi.12100>
- Beaumont LJ, Gallagher RV, Thuiller W, Downey PO, Leishman MR, Hughes L (2009) Different climatic envelopes among invasive populations may lead to underestimations of current and future biological invasions. *Diversity & Distributions* 15(3): 409–420. <https://doi.org/10.1111/j.1472-4642.2008.00547.x>
- Bellard C, Jeschke JM (2016) A spatial mismatch between invader impacts and research publications. *Conservation Biology* 30(1): 230–232. <https://doi.org/10.1111/cobi.12611>
- Bellard C, Cassey P, Blackburn TM (2016) Alien species as a driver of recent extinctions. *Biology Letters* 12(2): 20150623. <https://doi.org/10.1098/rsbl.2015.0623>
- Bellard C, Bernery C, Leclerc C (2021) Looming extinctions due to invasive species: Irreversible loss of ecological strategy and evolutionary history. *Global Change Biology* 27(20): 4967–4979. <https://doi.org/10.1111/gcb.15771>

- Blossey B, Notzold R (1995) Evolution of Increased Competitive Ability in Invasive Nonindigenous Plants: A Hypothesis. *Journal of Ecology* 83(5): 887. <https://doi.org/10.2307/2261425>
- Boyce MS, Vernier PR, Nielsen SE, Schmiegelow FK (2002) Evaluating resource selection functions. *Ecological Modelling* 157(2–3): 281–300. [https://doi.org/10.1016/S0304-3800\(02\)00200-4](https://doi.org/10.1016/S0304-3800(02)00200-4)
- Broennimann O, Guisan A (2008) Predicting current and future biological invasions: Both native and invaded ranges matter. *Biology Letters* 4(5): 585–589. <https://doi.org/10.1098/rsbl.2008.0254>
- Broennimann O, Treier UA, Müller-Schärer H, Thuiller W, Peterson AT, Guisan A (2007) Evidence of climatic niche shift during biological invasion. *Ecology Letters* 10(8): 701–709. <https://doi.org/10.1111/j.1461-0248.2007.01060.x>
- Büyüktaktakın İE, Haight RG (2018) A review of operations research models in invasive species management: State of the art, challenges, and future directions. *Annals of Operations Research* 271(2): 357–403. <https://doi.org/10.1007/s10479-017-2670-5>
- Capel JJ (1981) *Los climas de España* Oikos-Tau, S.A. Ediciones, Barcelona.
- Capinha C, Anastácio P (2011) Assessing the environmental requirements of invaders using ensembles of distribution models. *Diversity & Distributions* 17(1): 13–24. <https://doi.org/10.1111/j.1472-4642.2010.00727.x>
- Capinha C, Leung B, Anastácio P (2011) Predicting worldwide invasiveness for four major problematic decapods: An evaluation of using different calibration sets. *Ecography* 34(3): 448–459. <https://doi.org/10.1111/j.1600-0587.2010.06369.x>
- Carthey AJ, Banks PB (2018) Naïve, bold, or just hungry? An invasive exotic prey species recognises but does not respond to its predators. *Biological Invasions* 20(12): 3417–3429. <https://doi.org/10.1007/s10530-018-1782-4>
- del Moral JC (2003) *Atlas de las aves reproductoras de España*. Dirección General de Conservación de la Naturaleza, Madrid.
- Diagne C, Leroy B, Vaissière AC, Gozlan RE, Roiz D, Jarić I, Salles J-M, Bradshaw CJA, Courchamp F (2021) High and rising economic costs of biological invasions worldwide. *Nature* 592(7855): 571–576. <https://doi.org/10.1038/s41586-021-03405-6>
- Díaz S, Settele J, Brondízio ES, Ngo HT, Agard J, Arneth A, Balvanera P, Brauman KA, Butchart SHM, Chan KMA, Garibaldi LA, Ichii K, Liu J, Subramanian SM, Midgley GF, Miloslavich P, Molnár Z, Obura D, Pfaff A, Polasky S, Purvis A, Razzaque J, Reyers B, Chowdhury RR, Shin Y-J, Visseren-Hamakers I, Willis KJ, Zayas CN (2019) Pervasive human-driven decline of life on Earth points to the need for transformative change. *Science* 366(6471): eaax3100. <https://doi.org/10.1126/science.aax3100>
- Dietze MC, Fox A, Beck-Johnson LM, Betancourt JL, Hooten MB, Jarnevich CS, Keitt TH, Kenney MA, Laney CM, Larsen LG, Loescher HW, Lunch CK, Pijanowski BC, Randerson JT, Read EK, Tredennick AT, Vargas R, Weathers KC, White EP (2018) Iterative near-term ecological forecasting: Needs, opportunities, and challenges. *Proceedings of the National Academy of Sciences of the United States of America* 115(7): 1424–1432. <https://doi.org/10.1073/pnas.1710231115>
- Dostál P, Müllerová J, Pyšek P, Pergl J, Klinerová T (2013) The impact of an invasive plant changes over time. *Ecology Letters* 16(10): 1277–1284. <https://doi.org/10.1111/ele.12166>
- eBird (2019) eBird: An online database of bird distribution and abundance [web application]. eBird, Cornell Lab of Ornithology, Ithaca, New York. <http://www.ebird.org> [Accessed: 1 July, 2019]
- Ehrenfeld JG (2010) Ecosystem consequences of biological invasions. *Annual Review of Ecology, Evolution, and Systematics* 41(1): 59–80. <https://doi.org/10.1146/annurev-ecolsys-102209-144650>
- Elith J, Leathwick JR (2009) Species Distribution Models: Ecological Explanation and Prediction Across Space and Time. *Annual Review of Ecology, Evolution, and Systematics* 40(1): 677–697. <https://doi.org/10.1146/annurev.ecolsys.110308.120159>

- Engler R, Hordijk W, Guisan A (2012) The MIGCLIM R package - seamless integration of dispersal constraints into projections of species distribution models. *Ecography* 35(10): 872–878. <https://doi.org/10.1111/j.1600-0587.2012.07608.x>
- Engler R, Hordijk W, Pellissier L (2013) MigClim: implementing dispersal into species distribution models. R package version, 1(6). <https://CRAN.R-project.org/package=MigClim>
- Farr TG, Rosen RA, Caro E, Crippen R, Duren R, Hensley S, Kobrick M, Paller M, Rodriguez E, Roth L, Seal D, Shaffer S, Shimada J, Umland J, Werner M, Oskin M, Burbank D, Alsdorf D (2007) The Shuttle Radar Topography Mission. *Reviews of Geophysics* 45(2): RG2004. <https://doi.org/10.1029/2005RG000183>
- Ficetola GF, Thuiller W, Miaud C (2007) Prediction and validation of the potential global distribution of a problematic alien invasive species—The American bullfrog. *Diversity & Distributions* 13(4): 476–485. <https://doi.org/10.1111/j.1472-4642.2007.00377.x>
- Fick SE, Hijmans RJ (2017) WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology* 37(12): 4302–4315. <https://doi.org/10.1002/joc.5086>
- Fitzpatrick MC, Weltzin JE, Sanders NJ, Dunn RR (2006) The biogeography of prediction error: Why does the introduced range of the fire ant over-predict its native range? *Global Ecology and Biogeography* 16(1): 24–33. <https://doi.org/10.1111/j.1466-8238.2006.00258.x>
- Fournier A, Penone C, Pennino MG, Courchamp F (2019) Predicting future invaders and future invasions. *Proceedings of the National Academy of Sciences of the United States of America* 116(16): 7905–7910. <https://doi.org/10.1073/pnas.1803456116>
- Franklin J (2010) Mapping species distributions: spatial inference and prediction. Cambridge University Press. <https://doi.org/10.1017/CBO9780511810602>
- Fréjaville T, Benito Garzón M (2018) The EuMedClim database: Yearly climate data (1901–2014) of 1 km resolution grids for Europe and the Mediterranean Basin. *Frontiers in Ecology and Evolution* 6: 31. <https://doi.org/10.3389/fevo.2018.00031>
- Genovesi P, Shine C (2004). European strategy on invasive alien species: Convention on the Conservation of European Wildlife and Habitats (Bern Convention) (No. 18-137). Council of Europe.
- González-Moreno P, Diez JM, Richardson DM, Vilà M (2015) Beyond climate: Disturbance niche shifts in invasive species. *Global Ecology and Biogeography* 24(3): 360–370. <https://doi.org/10.1111/geb.12271>
- Groom Q, Strubbe D, Adriaens T, Davis AJ, Desmet P, Oldoni D, Reyserhove L, Roy HE, Vanderhoeven S (2019) Empowering citizens to inform decision-making as a way forward to support invasive alien species policy. *Citizen Science: Theory and Practice* 4(1): 33. <https://doi.org/10.5334/cstp.238>
- Guillera-Arroita G, Lahoz-Monfort JJ, Elith J (2014) Maxent is not a presence-absence method: A comment on Thibaud et al. *Methods in Ecology and Evolution* 5(11): 1192–1197. <https://doi.org/10.1111/2041-210X.12252>
- Guisan A, Thuiller W (2005) Predicting species distribution: Offering more than simple habitat models. *Ecology Letters* 8(9): 993–1009. <https://doi.org/10.1111/j.1461-0248.2005.00792.x>
- Guisan A, Zimmermann NE (2000) Predictive habitat distribution models in ecology. *Ecological Modelling* 135(2–3): 147–186. [https://doi.org/10.1016/S0304-3800\(00\)00354-9](https://doi.org/10.1016/S0304-3800(00)00354-9)
- Guisan A, Petitpierre B, Broennimann O, Daehler C, Kueffer C (2014) Unifying niche shift studies: Insights from biological invasions. *Trends in Ecology & Evolution* 29(5): 260–269. <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. <https://doi.org/10.1017/9781139028271>
- Hirzel AH, Le Lay G, Helfer V, Randin C, Guisan A (2006) Evaluating the ability of habitat suitability models to predict species presences. *Ecological Modelling* 199(2): 142–152. <https://doi.org/10.1016/j.ecolmodel.2006.05.017>

- Howard L, van Rees CB, Dahlquist Z, Luikart G, Hand BK (2022) A review of invasive species reporting apps for citizen science and opportunities for innovation. *NeoBiota* 71: 165–188. <https://doi.org/10.3897/neobiota.71.79597>
- Hutchinson G (1957) Concluding remarks. *Cold Spring Harbor Symposia on Quantitative Biology* 22(0): 415–427. <https://doi.org/10.1101/SQB.1957.022.01.039>
- Jarnevich CS, Sofaer HR, Belamaric P, Engelstad P (2022) Regional models do not outperform continental models for invasive species. *NeoBiota* 77: 1–22. <https://doi.org/10.3897/neobiota.77.86364>
- Jeschke JM, Strayer DL (2008) Usefulness of bioclimatic models for studying climate change and invasive species. *Annals of the New York Academy of Sciences* 1134(1): 1–24. <https://doi.org/10.1196/annals.1439.002>
- Jiménez-Valverde A, Lobo JM, Hortal J (2008) Not as good as they seem: The importance of concepts in species distribution modelling. *Diversity & Distributions* 14(6): 885–890. <https://doi.org/10.1111/j.1472-4642.2008.00496.x>
- Jiménez-Valverde A, Peterson AT, Soberón J, Overton JM, Aragón P, Lobo JM (2011) Use of niche models in invasive species risk assessments. *Biological Invasions* 13(12): 2785–2797. <https://doi.org/10.1007/s10530-011-9963-4>
- Keane R (2002) Exotic plant invasions and the enemy release hypothesis. *Trends in Ecology & Evolution* 17(4): 164–170. [https://doi.org/10.1016/S0169-5347\(02\)02499-0](https://doi.org/10.1016/S0169-5347(02)02499-0)
- Keil P, Belmaker J, Wilson AM, Unitt P, Jetz W (2013) Downscaling of species distribution models: A hierarchical approach. *Methods in Ecology and Evolution* 4(1): 82–94. <https://doi.org/10.1111/j.2041-210x.2012.00264.x>
- Leroy B, Delsol R, Hugueny B, Meynard CN, Barhoumi C, Barbet-Massin M, Bellard C (2018) Without quality presence–absence data, discrimination metrics such as TSS can be misleading measures of model performance. *Journal of Biogeography* 45(9): 1994–2002. <https://doi.org/10.1111/jbi.13402>
- Liu C, Berry PM, Dawson TP, Pearson RG (2005) Selecting thresholds of occurrence in the prediction of species distributions. *Ecography* 28(3): 385–393. <https://doi.org/10.1111/j.0906-7590.2005.03957.x>
- Liu C, Wolter W, Xian W, Jeschke JM (2020) Species Distribution Models Have Limited Spatial Transferability for Invasive Species. *Ecology Letters* 23(11): 1682–1692. <https://doi.org/10.1111/ele.13577>
- Naeem S, Bunker DE, Hector A, Loreau M, Perrings C [Eds] (2009) Biodiversity, ecosystem functioning, and human wellbeing: an ecological and economic perspective. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199547951.001.0001>
- Naimi B, Capinha C, Ribeiro J, Rahbek C, Strubbe D, Reino L, Araújo MB (2022) Potential for invasion of traded birds under climate and land-cover change. *Global Change Biology* 28(19): 5654–5666. <https://doi.org/10.1111/gcb.16310>
- Ogden NH, Wilson JR, Richardson DM, Hui C, Davies SJ, Kumschick S, Le Roux JJ, Measey J, Saul W-C, Pulliam JRC (2019) Emerging infectious diseases and biological invasions: A call for a One Health collaboration in science and management. *Royal Society Open Science* 6(3): 181577. <https://doi.org/10.1098/rsos.181577>
- Payne RB, Bonan A, Kirwan GM (2020) Common Waxbill (*Estrilda astrild*), version 1.0. In: del Hoyo J, et al. (Eds) *Birds of the World*. Cornell Lab of Ornithology, Ithaca, NY, USA. <https://doi.org/https://doi.org/10.2173/bow.comwax.01>
- Pereira PF, Barbosa AM, Godinho C, Salgueiro PA, Silva RR, Lourenço R (2020) The spread of the red-billed leiothrix (*Leiothrix lutea*) in Europe: The conquest by an overlooked invader? *Biological Invasions* 22(2): 709–722. <https://doi.org/10.1007/s10530-019-02123-5>
- Peterson AT (2003) Predicting the geography of species' invasions via ecological niche modeling. *The Quarterly Review of Biology* 78(4): 419–433. <https://doi.org/10.1086/378926>

- Peterson AT, Vieglais DA (2001) Predicting Species Invasions Using Ecological Niche Modeling: New Approaches from Bioinformatics Attack a Pressing Problem: A new approach to ecological niche modeling, based on new tools drawn from biodiversity informatics, is applied to the challenge of predicting potential species' invasions. *Bioscience* 51(5): 363–371. [https://doi.org/10.1641/0006-3568\(2001\)051\[0363:PSIUEN\]2.0.CO;2](https://doi.org/10.1641/0006-3568(2001)051[0363:PSIUEN]2.0.CO;2)
- Peterson AT, Soberón J, Pearson RG, Anderson RP, Martínez-Meyer E, Nakamura M, Araújo MB (2011) *Ecological niches and geographic distributions (MPB-49)* Vol. 56, Princeton University Press. <https://doi.org/10.23943/princeton/9780691136868.001.0001>
- Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling* 190(3–4): 231–259. <https://doi.org/10.1016/j.ecolmodel.2005.03.026>
- Phillips SJ, Dudík M, Elith J, Graham CH, Lehmann A, Leathwick J, Ferrier S (2009) Sample selection bias and presence-only distribution models: Implications for background and pseudo-absence data. *Ecological Applications* 19(1): 181–197. <https://doi.org/10.1890/07-2153.1>
- Phillips SJ, Anderson RP, Dudík M, Schapire RE, Blair ME (2017) Opening the black box: An open-source release of Maxent. *Ecography* 40(7): 887–893. <https://doi.org/10.1111/ecog.03049>
- Pili AN, Tingley R, Sy EY, Diesmos MLL, Diesmos AC (2020) Niche shifts and environmental non-equilibrium undermine the usefulness of ecological niche models for invasion risk assessments. *Scientific Reports* 10(1): 7972. <https://doi.org/10.1038/s41598-020-64568-2>
- Pysek P, Richardson DM, Pergl J, Jarošík V, Sixtová Z, Weber E (2008) Geographical and taxonomic biases in invasion ecology. *Trends in Ecology & Evolution* 23(5): 237–244. <https://doi.org/10.1016/j.tree.2008.02.002>
- R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>
- Raes N, ter Steege H (2007) A null-model for significance testing of presence-only species distribution models. *Ecography* 30(5): 727–736. <https://doi.org/10.1111/j.2007.0906-7590.05041.x>
- Regulation (EU) No 1143/2014 of the European Parliament and of the Council of 22 October 2014 on the prevention and management of the introduction and spread of invasive alien species. <https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1417443504720&curi=CELEX:32014R1143>
- Reino L (2005) Variation partitioning for range expansion of an introduced species: The common waxbill *Estrilda astrild* in Portugal. *Journal of Ornithology* 146(4): 377–382. <https://doi.org/10.1007/s10336-005-0093-6>
- Reino L, Moya-Laraño J, Cláudio Heitor A (2009) Using survival regression to study patterns of expansion of invasive species: Will the common waxbill expand with global warming? *Ecography* 32(2): 237–246. <https://doi.org/10.1111/j.1600-0587.2008.05354.x>
- Ribeiro J, Sillero N, Lopes RJ, Sullivan MJ, Santana J, Capinha C, Reino L (2020) 20 Common Waxbill (*Estrilda astrild* Linnaeus, 1758). *Invasive Birds*, 155. <https://doi.org/10.1079/9781789242065.0155>
- Rufino R (Coord.) (1989). Atlas das aves que nidificam em Portugal Continental. Ministério do Plano e da Administração do Território, Serviço Nacional de Parques, Reservas e Conservação da Natureza, Lisboa.
- Sakai AK, Allendorf FW, Holt JS, Lodge DM, Molofsky J, With KA, Baughman S, Cabin RJ, Cohen JE, Ellstrand NC, McCauley DE, O'Neil P, Parker IM, Thompson JN, Weller SG (2001) The Population Biology of Invasive Species. *Annual Review of Ecology and Systematics* 32(1): 305–332. <https://doi.org/10.1146/annurev.ecolsys.32.081501.114037>
- Santana J, Porto M, Brotons L, Junker J, Kissling WD, Lumbierres M, Moe J, Morán-Ordóñez A, Pereira H, Solheim AL, Villero D, Moreira F, Beja P (2023) D3. 2 Report on gaps and important

- new areas for monitoring in Europe. ARPHA Preprints 4: e103657. <https://preprints.arphahub.com/article/103657/>
- Scheffers BR, Oliveira BF, Lamb I, Edwards DP (2019) Global wildlife trade across the tree of life. *Science* 366(6461): 71–76. <https://doi.org/10.1126/science.aav5327>
- Seebens H, Bacher S, Blackburn TM, Capinha C, Dawson W, Dullinger S, Genovesi P, Hulme PE, van Kleunen M, Kühn I, Jeschke JM, Lenzner B, Liebhold AM, Pattison Z, Pergl J, Pyšek P, Winter M, Essl F (2021) Projecting the continental accumulation of alien species through to 2050. *Global Change Biology* 27(5): 970–982. <https://doi.org/10.1111/gcb.15333>
- Sih A, Bolnick DI, Luttbeg B, Orrock JL, Peacor SD, Pintor LM, Preisser E, Rehage JS, Vonesh JR (2010) Predator–prey naïveté, antipredator behavior, and the ecology of predator invasions. *Oikos* 119(4): 610–621. <https://doi.org/10.1111/j.1600-0706.2009.18039.x>
- Sillero B, Skidmore T (2009) Biogeographical patterns derived from remote sensing variables: the amphibians and reptiles of the Iberian Peninsula. *Amphibia-Reptilia* 30(2): 185–206. <https://doi.org/10.1163/156853809788201207>
- Sillero N (2011) What does ecological modelling model? A proposed classification of ecological niche models based on their underlying methods. *Ecological Modelling* 222(8): 1343–1346. <https://doi.org/10.1016/j.ecolmodel.2011.01.018>
- Sillero N, Ribeiro-Silva J, Arenas-Castro S (2022) Shifts in climatic realised niches of Iberian species. *Oikos* 2022(4): e08505. <https://doi.org/10.1111/oik.08505>
- Silva T, Reino LM, Borralho R (2002) A model for range expansion of an introduced species: The common waxbill *Estrilda astrild* in Portugal. *Diversity & Distributions* 8(6): 319–326. <https://doi.org/10.1046/j.1472-4642.2002.00156.x>
- Soberon J, Peterson AT (2005) Interpretation of models of fundamental ecological niches and species' distributional areas. <https://doi.org/10.17161/bi.v2i0.4>
- Stuart KC, Hofmeister NR, Zichello JM, Rollins LA (2023) Global invasion history and native decline of the common starling: Insights through genetics. *Biological Invasions* 25(5): 1291–1316. <https://doi.org/10.1007/s10530-022-02982-5>
- Sullivan MJ, Davies RG, Reino L, Franco AM (2012) Using dispersal information to model the species–environment relationship of spreading non-native species. *Methods in Ecology and Evolution* 3(5): 870–879. <https://doi.org/10.1111/j.2041-210X.2012.00219.x>
- Urban MC, Phillips BL, Skelly DK, Shine R (2007) The cane toad's (*Rhinophrynus dorsalis*) increasing ability to invade Australia is revealed by a dynamically updated range model. *Proceedings. Biological Sciences* 274(1616): 1413–1419. <https://doi.org/10.1098/rspb.2007.0114>
- van Wilgen NJ, Gillespie MS, Richardson DM, Measey J (2018) A taxonomically and geographically constrained information base limits non-native reptile and amphibian risk assessment: A systematic review. *PeerJ* 6: e5850. <https://doi.org/10.7717/peerj.5850>
- Venette RC [Ed.] (2015) Pest risk modelling and mapping for invasive alien species (Vol. 7). CABI. <https://doi.org/10.1079/9781780643946.0000>
- Venette RC, Kriticos DJ, Magarey RD, Koch FH, Baker RH, Worner SP, Gómez Raboteaux NN, McKenney DW, Dobesberger EJ, Yemshanov D, De Barro PJ, Hutchison WD, Fowler G, Kalaris TM, Pedlar J (2010) Pest risk maps for invasive alien species: A roadmap for improvement. *BioScience* 60(5): 349–362. <https://doi.org/10.1525/bio.2010.60.5.5>
- Wiens JJ, Graham CH (2005) Niche Conservatism: Integrating Evolution, Ecology, and Conservation Biology. *Annual Review of Ecology, Evolution, and Systematics* 36(1): 519–539. <https://doi.org/10.1146/annurev.ecolsys.36.102803.095431>
- Wilson JRU, Richardson DM, Rouget M, Procheş Ş, Amis MA, Henderson L, Thuiller W (2007) Residence time and potential range: Crucial considerations in modelling plant invasions. *Diversity & Distributions* 13(1): 11–22. <https://doi.org/10.1111/j.1366-9516.2006.00302.x>

Supplementary material 1

Oldest records of common waxbill *Estrilda astrild* for each 10×10 UTM km grid cells of Portugal and Spain

Authors: Joana Santana, Neftalí Sillero, Joana Ribeiro, César Capinha, Ricardo Jorge Lopes, Luís Reino

Data type: csv

Copyright notice: This dataset is made available under the Open Database License (<http://opendata-commons.org/licenses/odbl/1.0/>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: <https://doi.org/10.3897/neobiota.95.122335.suppl1>

Supplementary material 2

Supplementary figures and tables

Authors: Joana Santana, Neftalí Sillero, Joana Ribeiro, César Capinha, Ricardo Jorge Lopes, Luís Reino

Data type: docx

Copyright notice: This dataset is made available under the Open Database License (<http://opendata-commons.org/licenses/odbl/1.0/>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: <https://doi.org/10.3897/neobiota.95.122335.suppl2>

Supplementary material 3

R code scripts used to run the analyses

Authors: Joana Santana, Neftalí Sillero, Joana Ribeiro, César Capinha, Ricardo Jorge Lopes, Luís Reino

Data type: zip

Copyright notice: This dataset is made available under the Open Database License (<http://opendata-commons.org/licenses/odbl/1.0/>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: <https://doi.org/10.3897/neobiota.95.122335.suppl3>