

Title: Environmental resistance predicts the spread of alien species

Authors: Rebecca S L Lovell^{1*}, Tim M Blackburn^{1,2}, Ellie E Dyer¹, Alex L Pigot^{1*}

Affiliations: ¹Centre for Biodiversity and Environment Research, Department of Genetics, Evolution and Environment, University College London, London, United Kingdom, ²Institute of Zoology, Zoological Society of London, London, United Kingdom.

*Correspondence to: Rebecca.Lovell.15@ucl.ac.uk, a.pigot@ucl.ac.uk

The number of alien species introduced outside their native range is increasing unabated with major impacts on biodiversity and human society. Predicting the spread of alien species is therefore an urgent challenge. Most predictions use models of a species' ecological niche to identify climatically suitable areas for invasion, but recent analyses suggest that the reliability of these predictions may be limited. Here, using the global alien avifauna, we demonstrate an alternative approach for predicting the spread of alien species based on the environmental resistance of the landscape. This approach does not require any information on the ecological niche of the invading species, and instead uses gradients of biotic similarity among native communities in the invaded region to predict the most likely routes of spread. We show that environmental resistance predicts current patterns of spread better than a null model of random diffusion or a model based on climate matching to the species' native range. Applying this approach to simulate future patterns of spread reveals major differences in projected invasion risk across regions, shaped both by proximity to existing invasion hotspots as well as broad-scale gradients in the environmental resistance to spread. Our results show how environmental resistance may provide a general and complementary approach for predicting risks of species' invasions that can be rapidly deployed even when information on the niche or the identity of potential invaders is unknown.

Invasive alien species represent a growing threat to biodiversity, food security, human health and economies¹⁻⁴. Reliably predicting the spread of alien species is therefore an urgent challenge. Most predictions rely on estimating the environmental suitability for invasion using models of the invading species' native ecological niche⁵⁻⁸. However, the success of ecological niche models in predicting the spread of alien species is mixed⁹⁻¹¹. While there is some evidence that species' niches may be conserved across their native and alien distributions (e.g.⁵), other studies suggest that predictions of alien spread using ecological niche models are unreliable (e.g.^{12,13}). In particular, alien species often fail to spread into conditions occupied in their native range, while at the same time expanding into novel climates¹²⁻¹⁸. The potential limited accuracy of ecological niche models represents a major challenge for managing risks from biological invasions, and motivates the search for novel approaches for predicting the spread of alien species.

Starting from any point on the Earth's surface, the biotic similarity of ecological communities typically declines with geographic distance, a trend so fundamental it has been referred to as the 'first principle of biogeography'^{19,20}. Patterns of distance decay in compositional similarity can be quantified in various ways, but perhaps most relevant for understanding the spread of invasive alien species is simply the proportion of the native species at a focal site occurring at every other location in the landscape²¹(see Methods). When mapped, this reveals striking geometric patterns, reflecting the environmental gradients and geographical barriers responsible for limiting the spread of species at any given site²²⁻²⁴(Fig. 1).

Forty years ago, Eduardo Rapoport²¹ proposed that taking the complement of biotic similarity would provide an index of 'environmental resistance' (i.e. environmental

resistance = 1 - biotic similarity), that could be used to predict how a newly introduced species would spread through the landscape. Specifically, having successfully established at a focal site, an alien species would have a lower probability of spreading to areas with low biotic similarity to the focal site (i.e. higher environmental resistance) compared to areas with high biotic similarity (i.e. lower environmental resistance). This environmental resistance model disregards all information on species' ecological niches and instead assumes that the biotic similarity of native species communities provides a convenient proxy for suitability. Thus, environmental resistance distills into a single catch-all variable, the potentially many measured and unmeasured environmental factors that interact, in potentially complex ways, to limit the geographic expansion of species at a particular location. Despite its attractive simplicity, the ability of Rapoport's environmental resistance model to predict the spread of alien species has, to our knowledge, never been tested.

Here we apply the environmental resistance model to the global alien avifauna, that includes 339 species distributed across all continents except Antarctica²⁵. Birds provide an ideal study system because comprehensive information is available on (i) the geographic range of each alien species, (ii) the sites where these alien populations established and thus the likely sources of spread^{26,27}, and (iii) the native distributions of all birds ($n = 9993$)²⁸ from which environmental resistance across sites (100km grid cells) can be mapped globally (Fig. 1). Using this dataset, we simulate the spread of alien species from known sites of establishment, with patterns of invasion determined by the environmental resistance of the landscape, and assess the ability of this model accurately to predict the current alien distribution of each species. Because some level of overlap between the observed and simulated alien distribution would be expected simply by chance, we

compare the predictions of the environmental resistance model to a null model in which patterns of invasion are unrelated to environmental resistance and arise purely through random dispersal. Finally, we compare the accuracy of the environmental resistance model to a suite of models where spread is guided by estimates of climatic suitability (Extended Data Table. 1). After demonstrating that environmental resistance more accurately predicts patterns of spread than models based on random dispersal or climate suitability, we apply the environmental resistance model to provide a novel global assessment of the future risk to species assemblages from the spread of alien birds.

Results and Discussion

Predicting patterns of alien spread

We simulate the geographic range of alien species using a spreading dye algorithm^{29,30}. Starting from the known locations where alien populations established, species sequentially invade adjacent cells, stopping when the observed alien range size is reached (Methods). We conduct multiple repeat simulations ($n = 100$) and quantify predictive accuracy as the percentage of the observed range that is overlapped by the simulated distribution. Our first aim is therefore to assess the ability to predict patterns of spread given the observed alien range size attained by each species. We subsequently go on to assess models' ability to predict the extent of spread.

In our null model of random dispersal, alien species invade adjacent cells with equal probability and spread is constrained only by the boundaries of the landmasses³⁰. On average across species, the random dispersal model predicts patterns of alien range spread with an accuracy of 67% (median \pm 15% SD) (Fig. 2a). Most null models assume that in the

absence of any environmental influence species would be randomly scattered across space³¹. However, the relatively high overlap between observed and simulated ranges in our null model highlights the importance of incorporating a realistic process of dispersal when generating null expectations of alien distributions^{32,33}.

We next model the effects of environmental resistance on spread by preferentially selecting for invasion those cells with a lower environmental resistance (i.e. higher biotic similarity) to the site/s of alien population establishment (Fig. 2b). Incorporating this effect of environmental resistance resulted in a consistent and significant increase in the overlap between simulated and observed ranges (Fig. 2e, paired t-test, $t = 11.9$, $p < 0.001$, $n = 283$), with average model predictive accuracy increasing to 78% (median \pm 19% SD). Across our dataset, 65% of species are better fit by the environmental resistance model, compared to 29% of species where model predictive accuracy is tied and only 6% of species where the random dispersal model fits better. At the level of individual species, the null model of random dispersal can be statistically rejected (at 5% significance level) across 21% of species. Many alien species, however, have small geographic ranges and there may be insufficient power to reject the null model. Thus, when we consider only the most widespread alien species (i.e. upper 10% quantile), a model of random dispersal can be rejected in favor of the environmental resistance model across 67% of species. This consistent improvement in predictive accuracy shows that patterns of alien range spread cannot be explained by random dispersal alone. Instead, we find that species preferentially invade sites with lower environmental resistance (i.e. greater native biotic similarity) to sites of population establishment, giving rise to present day alien bird distributions that depart markedly from random expectations.

We next compare the predictions of the environmental resistance model to ‘climate-matching’ models in which the probability of invading adjacent cells is determined directly by climate suitability, estimated as the similarity to the mean conditions occupied in the native species range (Fig. 2c)^{6,27}. We calculate suitability by generating a multidimensional climate space described by four independent climate axes, and simulate alien range spread under all possible combinations of these climate axes ($n = 15$ combinations, Methods). We then select those climate axes that best predict spread (climate-matching_{Optimal}, Extended Data Table 1), allowing the identity of these axes to vary across species and across different fragments of each species’ global alien range (e.g. alien species introduced to multiple continents) (Extended Data Fig. 1).

This climate-matching_{Optimal} model has an average predictive accuracy of 75% (median \pm 14% SD), corresponding to a small but significant decrease in predictive performance compared to the environmental resistance model (paired t-test, $t = 3.22$, $p=0.001$, $n = 283$). Across our dataset, 47% of species are better fit by the environmental resistance model, compared to 31% of species where model predictive accuracy is tied, and 22% of species where the climate-matching_{Optimal} model fits better. The environmental resistance model significantly outperforms the climate-matching_{Optimal} model in 12% of species, compared to only 5% of species where it performs significantly worse (Fig. 2f). Thus, despite the much greater complexity of the climate-matching_{Optimal} model, predictive accuracy is lower than the environmental resistance model in which the spread of all species is governed by a single variable.

In practice the climate-matching_{Optimal} model has limited applicability because the climatic axes limiting alien range expansion are rarely known *a priori*. The standard

approach is instead to identify the climate variables that best predict the native range of each species, and to assume that these variables also limit spread in the alien range³⁴, what we here term the climate-matching_{Native} model (Fig. 2d, Extended Data Table 1). Under these more restrictive assumptions, the predictive accuracy of the climate-matching_{Native} model (median = 65±17% SD) is substantially lower than the environmental resistance model (paired t-test, $t = 13$, $p < 0.001$, $n = 283$). In fact, predictive accuracy of the climate-matching_{Native} model is lower than a null model of random dispersal which assumes that patterns of alien range spread are entirely independent of local climate conditions. Thus, while our climate-matching models demonstrate that climate is a strong determinant of alien range expansion, in practice estimates of climate suitability do not provide accurate predictions of alien spread, because different climatic variables are responsible for limiting the species' native and alien range (Extended Data Fig. 1).

One possible explanation for why the climate-matching models provides less accurate predictions than the environmental resistance model, is that species may undergo niche shifts during the invasion process so that the average conditions occupied in the native range do not accurately reflect climate suitability (e.g.¹³). This cannot explain our results, however, because we obtained a similar predictive accuracy when fitting a 'climate-resistance' model in which we quantify suitability for invasion relative to the climate conditions at the sites of alien population establishment rather than the native range (Extended Data Table 1, Extended Data Fig. 2). Taken together, our results suggest that environmental resistance may provide a more accurate index of invasion risk than estimates of climate suitability (Extended Data Fig. 3).

Predicting the extent of alien spread

To test whether the environmental resistance model can also predict the extent of spread, we estimate the area each alien species would be expected to occupy, assuming that range expansion is limited by a given threshold of environmental resistance (Fig. 3a). Under this model, variation in alien range size is explained by variation in the rate at which biotic similarity declines with distance from the sites where each species was established (Fig. 1). We first assess the correspondence between predicted and observed alien range size by quantifying the error around the ordinary least squares (OLS) regression line, systematically repeating this assessment for different thresholds of environmental resistance. Using the threshold of environmental resistance that minimises error (ER = 0.05, Extended Data Fig. 4a), predicted and observed alien range sizes are positively associated (Fig. 3b). Thus, as expected, alien species have attained larger range sizes where biotic similarity declines gradually with distance from sites of establishment, compared to places where biotic similarity declines steeply with distance. However, model explanatory power is relatively low ($r^2 = 0.19$) and across 59% of alien species the extent of spread is underpredicted (Fig. 3b). Indeed, while the observed range limits of many species correspond to a low environmental resistance threshold (e.g. $ER \leq 0.05$ for 45% of species), many other species have spread into cells with a much higher environmental resistance relative to the sites where they first established (Fig. 3a, Extended Data Fig. 4b). Clearly, any single environmental resistance threshold can only partially predict the extent of alien range expansion.

Given that the threshold of environmental resistance corresponding to the limits of alien range expansion appears to vary across species, for managing invasion risks it may be

more appropriate to identify a threshold of environmental resistance beyond which few, if any, species are likely to spread. Using a quantile regression (99% percentile) to define the upper boundary of the relationship between predicted and observed alien range size, we identify an environmental resistance threshold of $ER = 0.49$ as one possible 'safe limit' (Fig. 3c, Extended Data Fig. 4c). This threshold corresponds to the boundary beyond which native biotic similarity declines by approximately 50% (Fig. 1) and accurately predicts the maximum range size that alien species have attained, with few species (2%) expanding beyond this threshold (Fig. 3c). While it is possible that some species may transgress this threshold in the future, our analysis across all alien birds—some of which have been established for centuries—suggests that such risks are probably small over the coming decades. Thus, while a given threshold of environmental resistance provides only limited accuracy in predicting the current range size of any alien species, it may nevertheless be used to define a safe limit beyond which risks of invasion are likely to be minimal.

Predicting global risks of invasion

Using this proposed safe limit, we map the risk of invasion from the continued spread of currently established alien birds. Present day patterns of alien species richness are primarily driven by historical patterns of species' introductions, with hotspots of alien richness occurring throughout North America, Europe, Australia, New Zealand and Japan, as well as on numerous islands, major urban areas and trade hubs (Fig. 4a)²⁵. Patterns of risk from the potential future spread of currently established aliens partially reflect proximity to these existing hotspots of alien species richness (Fig. 4b). However, risk is also strongly shaped by gradients in environmental resistance, resulting in patterns of spread

that differ strongly from patterns expected under random dispersal. For instance, Sub-Saharan Africa emerges as a hotspot of risk from potential future spread, but the Kalahari and Congo Basin remain relative coldspots due to the high environmental resistance of these ecosystems relative to the sites where aliens have established. In the Middle East and Central Asia, deserts and high-altitude plateaus remain at low risk, while adjacent mountain ranges (e.g. Himalayas) with low environmental resistance relative to the sites where aliens have established, emerge as corridors for potential invasion (Fig. 4b).

The spatial patterns of risk that we identify from the continued spread of established aliens differ markedly from recent risk assessments focussing primarily on earlier stages in the invasion pathway, involving the initial transport and establishment of species^{35,36}. Integrating our environmental resistance model with predictions from these earlier stages in the invasion pathway is a critical next step (e.g.³⁷), as the regions most at risk may change over time as more alien bird populations are introduced and established^{2,25}. It is also important to note that our projected patterns of risk are based on current gradients of environmental resistance and may change radically as species distributions shift in response to future climate change.

Prospects and opportunities of the environmental resistance model

Here, we have shown that spatial patterns in community similarity provide a novel metric of environmental resistance that can be used to predict how alien species are likely to spread through the landscape, finally confirming Eduardo Rapoport's hypothesis that had remained untested for four decades²¹. The ability of environmental resistance to predict the spread of alien species in the complete absence of information on the ecological niche

of the species involved is striking, and suggests that this could provide a valuable tool in predicting invasion risk. Environmental resistance may be particularly useful for predicting the spread of species where information on their distribution is insufficient to estimate their ecological niche, including many tropical insects with poorly known native distributions but that have the potential to become major pests³. In fact, our results suggest that as long as the sites where species were introduced can be reliably identified (often major trade hubs and cities), then the environmental resistance model can be used to predict the locations most at risk from invasion, even when the identity of potential invaders are unknown. The ability to provide rapid and general assessments of invasion risk across regions given such minimal information could be highly valuable, complementing existing niche modelling approaches which require detailed geographic or physiological data for individual species.

What though is the explanation for the high predictive ability of the environmental resistance model? In theory, communities with similar compositions to sites of alien establishment could causally promote invasion, perhaps because certain native species facilitate both the establishment and spread of aliens. While this scenario is possible, we think it more likely that the association is indirect, and that relative biotic similarity predicts spread simply because it provides a more comprehensive metric of environmental suitability than direct estimates based on the climate. In particular, species' ranges may be constrained by a potentially large number of abiotic, biotic and historical factors, and this complexity may not be captured by the small set of climatic variables typically used to model species' distributions¹². In contrast, patterns of decay in biotic similarity effectively integrate these many factors into a single, comprehensive 'catch-all' variable. The relatively

small but consistent improvement in predictive accuracy of the environmental resistance model compared to the climate-matching_{Optimal} model provides some support for this explanation. But our results suggest that a more fundamental reason is that the identity of the key climate variables predicting species' distributions often varies between regions, making it difficult to transfer models of climate suitability based on a species' native range¹⁰. This explains why the climate-matching_{Native} model—which assumes the same climate axes limit both the native and alien range—performs so poorly compared to the climate-matching_{Optimal} model, which allows the identity of these axes to vary.

Unfortunately, the identity of the climate axes limiting alien spread may only become identifiable once invasion has already occurred, thereby limiting the practical applicability of such climate matching models. The environmental resistance model side-steps this problem by using the patterns of decay in biotic similarity in the invaded region to predict how alien species are likely to spread.

We found that the relationship between environmental resistance and the extent of range spread is triangular (Fig. 3c). Alien species occurring in regions with high environmental resistance are correctly predicted to have small geographic ranges, whereas species occurring in regions with low environmental resistance can have large or small geographic ranges. These differences in range size could reflect a number of factors, including differences in residence time, propagule pressure, dispersal ability and niche breadth, all of which may lead to variation in the rate or time available for spread^{6,38,39}. Analysis of how these, as well as other factors, may modulate the ability of the environmental resistance model to predict spread is an important avenue for further exploration. While additional information beyond environmental resistance will clearly be

necessary to predict the range size that an alien species will attain at any particular point in time, our results show that environmental resistance can be used to identify those regions at potential risk of invasion, as well as safe limits beyond which spread is unlikely (Fig. 3c).

To our knowledge, this is the first study to test the ability of environmental resistance to predict the spread of alien species, and our analysis raises a number of questions. While environmental resistance provides a more accurate predictor of alien spread than a climate-matching model or assuming random dispersal, the accuracy of predictions varies across species (Fig. 2c). What then are the conditions under which environmental resistance performs well or less well in predicting alien spread? Here we have predicted patterns of spread in the alien avifauna based on the native distribution of all birds, but what about organism groups for which native distributions are less well known? For instance, could the spread of invasive plants be predicted from patterns of biotic similarity of birds? Could the spread of invasive insect pests (e.g.³) be predicted from patterns of biotic similarity of their vertebrate or plant hosts? Would incorporating information on functional traits or phylogenetic relatedness into metrics of biotic similarity further improve predictions of spread compared to models that consider only species identity? Finally, how does environmental resistance perform at the finer spatial resolutions that are more relevant to conservation practitioners and land managers? With the number of alien species increasing unabated², addressing these questions are priorities for further research. Reliably predicting risks of future species invasions would benefit from a diverse set of tools that use different assumptions, data types and methodologies to model how species spread through the environment. Our results suggest that environmental resistance may provide an important addition to the ecologist's toolbox.

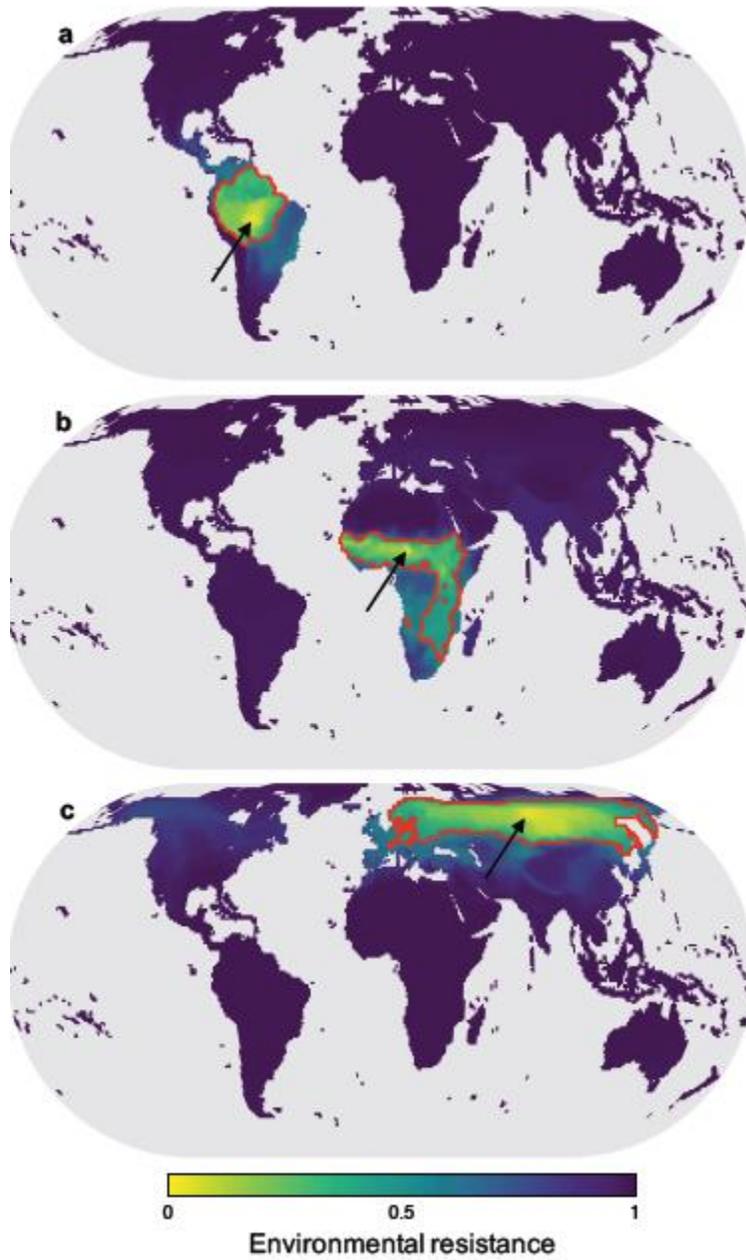


Fig. 1. Patterns of environmental resistance. Environmental resistance (1 - biotic similarity) is shown for selected sites in (a) the Amazon, (b) the Sahel and (c) Siberia (indicated by arrows) based on the native distributions of all birds ($n = 9,993$ species). From each of these sites, the pairwise compositional similarity to all other cells was calculated (see Methods). Hotter colours indicate a higher similarity to the focal location (i.e. a higher proportion of focal cell species present at a site) and lower environmental

resistance to spread. The red contour indicates an environmental resistance of 0.5 i.e. 50% loss of biotic similarity from the focal site.

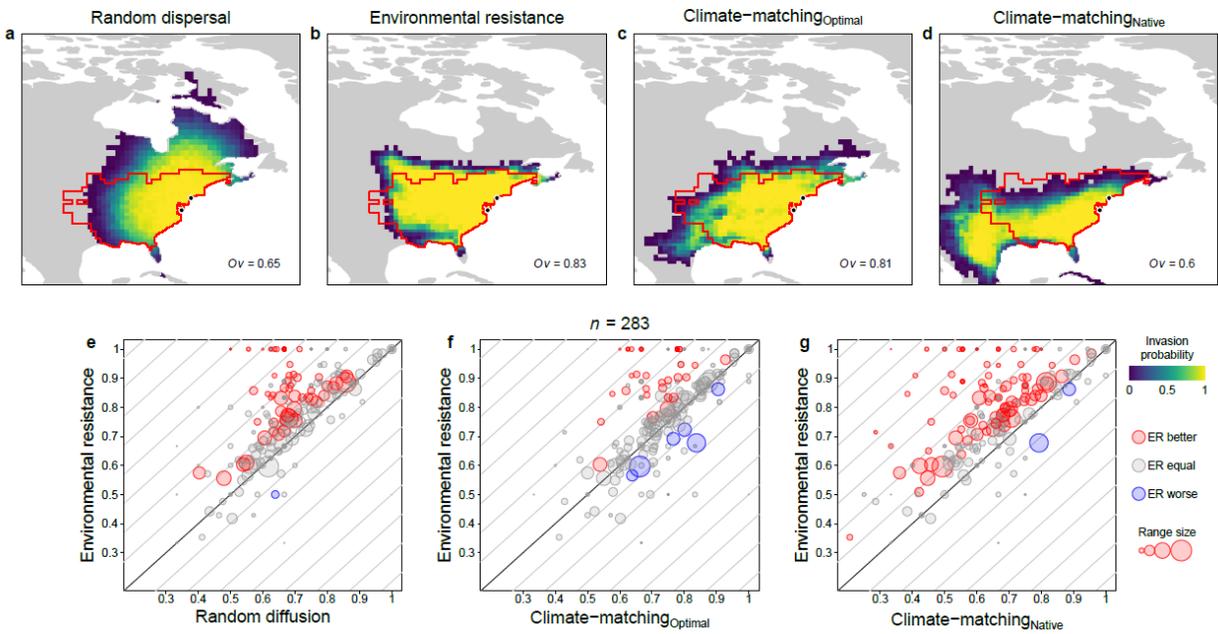


Fig. 2. Predicting the spread of alien birds. (a-d) Observed (red) and predicted spread for a single alien species (*Carpodacus mexicanus*). Hotter colors indicate a higher probability of invasion across $n = 100$ simulations under the (a) random dispersal, (b) environmental resistance and (c-d) climate-matching models. Climate-matching was calculated using the optimal combination of climate axes identified for the (c) alien (climate-matching_{Optimal}) and (d) native (climate-matching_{Native}) species range. *Ov* indicates the median predictive accuracy (i.e. % of invaded cells correctly predicted) across $n=100$ replicate simulations. (e-g) Predictive accuracy across species under the environmental resistance and competing models. Points above the 1:1 line indicate a higher predictive accuracy of the environmental resistance model. Colors indicate whether the environmental resistance model predicts spread significantly better (red) or worse (blue) than the competing model based on the 95% confidence interval in *Ov* scores of the competing model. Point size is proportional to alien range size. Of all established alien

species ($n = 339$), only those that have spread beyond cells of initial establishment were simulated ($n = 283$).

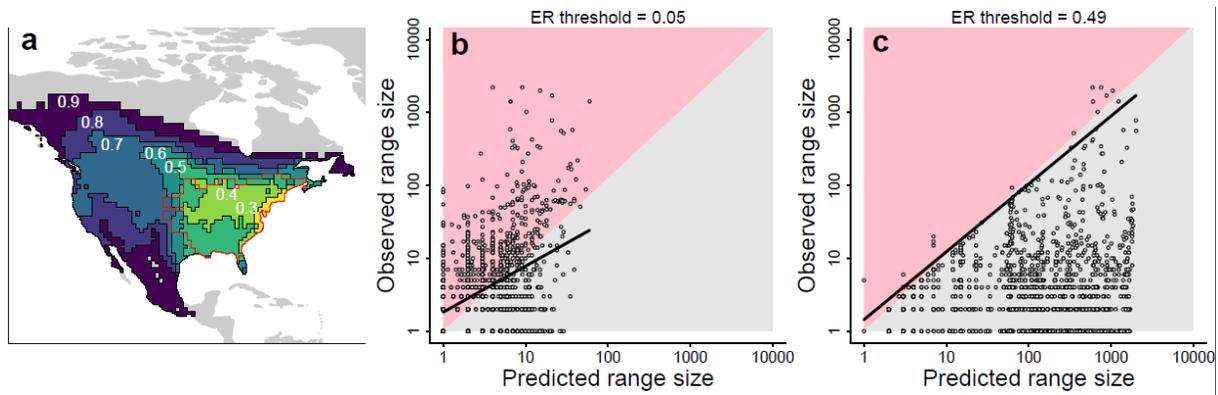


Fig. 3. Predicting the extent of spread across alien ranges. (a) Observed (red) and predicted spread for a single alien species (*Carpodacus mexicanus*) under different environmental resistance (ER) thresholds. ER thresholds < 0.3 are not visible at this scale. (b-c) The relationship between predicted and observed range size (number of 100km grid cells) assuming a threshold of (b) ER = 0.05 and (c) ER = 0.49 (see Extended Data Fig. 4a,c). Fitted lines indicate the (b) least squares and (c) quantile regression (0.99 percentile) of predicted against observed range size. Red regions indicate ranges that are under-predicted. For all $n = 339$ alien species, range size was calculated for individual fragments of species ranges separately ($n = 1704$) because variation in total species range size is dominated by variation in the number of introduced populations (Extended Data Fig. 5).

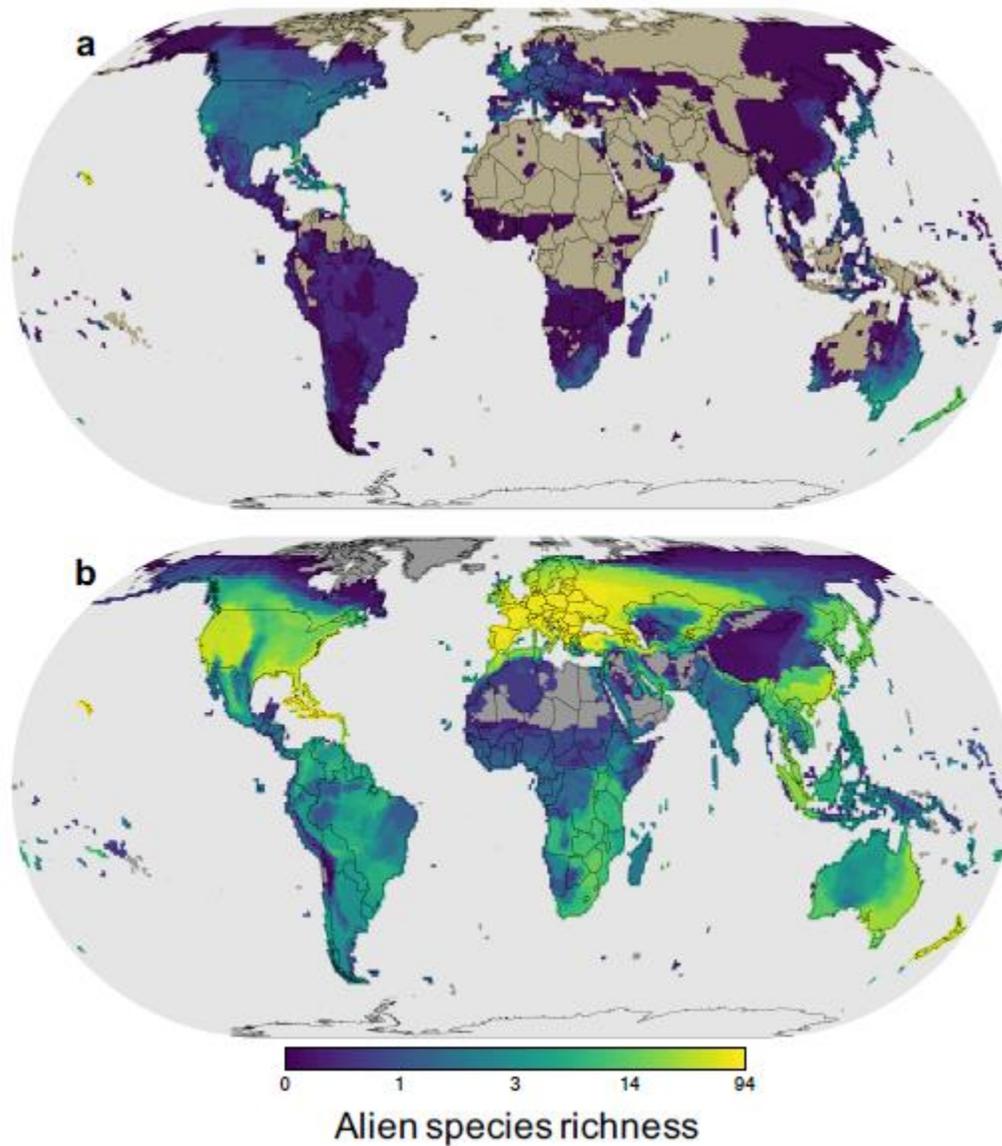


Fig. 4. Global risks of invasion from the spread of alien birds. (a) current richness of alien birds ($n = 339$ species) and (b) risk of invasion from the spread of current alien birds assuming an environmental resistance threshold of $ER = 0.49$ (see Extended Data Fig. 4c), equivalent to an approximately 50% decline in native biotic similarity (see Fig. 1). Grey regions represent areas where no aliens are (a) currently present or (b) predicted to spread into.

Methods

Data

Expert-drawn extent of occurrence range maps of current alien distributions were obtained from the Global Avian Invasions Atlas (GAVIA)^{25,26}. GAVIA represents the most comprehensive dataset on the distribution and introduction history of alien species for any major group of organisms, including information on the introduction of $n = 971$ bird species, $n = 362$ of which have established alien populations. For each alien species with a mapped distribution, we extracted from GAVIA the latitude and longitude coordinates of each introduction event that is known to have successfully established. Expert-drawn extent of occurrence maps of native breeding ranges for all bird species ($n = 9,993$) were obtained from BirdLife International²⁸. These expert range maps represent the most comprehensive data on species' distributions currently available, which is essential for a globally complete and unbiased characterization of compositional similarity among communities⁴⁰.

To quantify environmental-resistance and to allow the use of cell-based simulations, we extracted species' polygon range maps onto an equal-area grid (Behrmann projection) with a cell resolution of 96.486 km⁴¹. We used a grid resolution of ~100km as this is the finest resolution at which expert polygon range maps can be analysed without incurring substantial false presences⁴². We identified the grid cells within each alien species' range where alien populations were introduced and established. Where multiple establishment events occurred in a single grid cell these were lumped.

To generate environmental resistance surfaces for modelling range spread, we generated a matrix of the biotic similarity between each pair of cells based on species'

native distributions. For focal cell *i*, the biotic similarity of cell *j* was calculated as the proportion of species present in cell *i* that are also present in cell *j*^{21,24,43}. This asymmetric measure is more appropriate than symmetric beta-diversity metrics which also take into account the species present in cell *j* but absent from cell *i* (e.g. Jaccard or Sørensen index)²⁰, because here we are only interested in the potential of species to spread from cell *i* to *j* rather than the reverse. Environmental resistance was then calculated as: environmental resistance = 1 - biotic similarity. Thus, for cell *j*, an environmental resistance value of 0 with respect to focal cell *i*, indicates that all species present in cell *i* are also present in cell *j*. An environmental resistance value of 1 indicates that none of the species present in cell *i* are present in cell *j*. We do not include the distribution of alien species when calculating environmental resistance because this would be testing a different hypothesis, namely that the spread of an alien species can be predicted by the spread of other, earlier arriving alien species.

In addition to environmental resistance, we tested the extent to which range spread is predictable based on the similarity between local climate conditions and (i) the mean conditions occupied in a species' native range or (ii) at the site of alien population establishment. We downloaded data for a number of climatic variables at 1km resolution from WorldClim version v1.4⁴⁴ and then calculated the mean conditions within each 100km grid cell. We used a standard set of variables previously identified as important correlates of species' range limits in birds^{30,45}: mean annual temperature (°C) (Bio1), annual precipitation (mm) (Bio12), mean diurnal temperature range (°C) (Bio2), isothermality (°C) (Bio3), temperature seasonality (°C) (Bio4), temperature annual range (°C) (Bio7) and precipitation seasonality (mm) (Bio15). To derive a more limited set of spatially-

independent climate axes we performed a principal component (PC) analysis and retained the first four axes which collectively described 97.5% of the variation in climate (Extended Data Table 2). These PC climate axes were rescaled to between 0 and 1, thereby assuming that a proportional change in each climate axis would have an equivalent effect on invasion probability in our model. Cells lacking environmental data (usually small islands) were excluded from the grid. Excluding these cells resulted in a total of $n = 339$ alien species with distributions represented on our grid and that therefore could be modelled. All data processing and analyses were conducted in the R statistical programming language v4⁴⁶.

Simulating range spread

Of the $n = 339$ alien species, $n = 283$ species have spread beyond the initial grid cells where they established. For this subset of species, we simulated range expansion using a spreading dye algorithm. Starting from the cell/s of known population establishment events, each species could sequentially invade unoccupied adjacent cells in any of the four cardinal directions (i.e. sharing an edge with an occupied cell). Where there were multiple sites of establishment, spread was simultaneously seeded from each of these cells. The process of range spread was repeated until the observed range size was reached. We note that species have been spreading for different periods of time and our simulations thus aim to predict the spatial patterns of spread rather than the rate of spread. Many species have alien ranges divided into multiple, geographically disjunct fragments (e.g. where they have been introduced to different continents). We identified separate range fragments by building a graph object describing the connectivity between grid cells using the 'graph' package v1.66⁴⁷ and then identifying the disconnected subgraphs in the species range (i.e.

those sharing no adjacent cells) using the ‘components’ function in the ‘igraph’ package v1.2.5⁴⁸. For species consisting of multiple fragments, we preserved the observed size and number of fragments by stopping spread when the observed size of each fragment was reached and by preventing fragments from merging together (i.e. in the simulation, cells adjacent to those occupied by one simulated fragment could not be invaded by another simulated fragment of the same species)³⁰. Where information on the location of initial establishment was unavailable, range expansion was seeded from a single cell selected at random from within the observed alien range or for each fragment of the range.

For each of the n cells available for invasion, the probability of invasion $P(I)$ for cell i was calculated as,

$$P(I)_i = (ER_i + 1)^{-\mu_e}$$

Equation 1

Where ER_i is the environmental resistance value of cell i with respect to the cell where the alien population was first established. A constant of value 1 is added to ER_i to ensure a positive number when $ER_i = 0$. The parameter μ is a positive number indicating the degree of determinism in range spread (Extended Data Fig. 6a). When μ is large, cells with low environmental resistance are much more likely to be selected for invasion. When μ is small, cells are selected for invasion with a similar probability regardless of their environmental resistance. Here, we use a value of $\mu = 30$, representing a scenario with strong determinism (Extended Data Fig. 6a)³⁰. When $\mu = 0$, this model reduces to a null model of random diffusion, in which cells are selected for invasion with an identical probability (Extended Data Fig. 6a).

To ensure our results were not contingent on our choice of $\mu = 30$, we re-ran our environmental resistance model using different values of μ from 0.2 to 1,000 (Extended Data Fig. 6b)³⁰. As expected, decreasing values of μ below 30 (i.e. models with weaker determinism), causes the predictive power of the model to decrease, with model performance converging towards that of the random dispersal model (i.e. $\mu = 0$) (Extended Data Fig. 6b). By contrast, increasing μ above 30 only resulted in only a small increase in predictive accuracy (Extended Data Fig. 6b). Thus, our results are robust to the choice of μ and confirm that patterns of alien spread are highly deterministic and strongly shaped by environmental resistance.

We extended the above model to consider scenarios in which the probability of invasion is determined by climate suitability. We simulated spread where the probability of invasion, $P(I)$, for cell i was calculated as,

$$P(I)_i = 1 - \prod_{y=1}^4 1 - (|C_y - L_{iy}| + 1)^{-\mu_y} \quad \text{Equation 2}$$

the product of invasion probabilities across each climate axis y . For each climate axis ($n = 4$ axes), we assume that the invasion probability declines with the absolute distance between the local conditions in cell i (L_{iy}) and the mean climate conditions (C_y) across the grid cells occupied by the species' native range. This climate matching model has previously been shown to be a strong predictor of establishment success and the spread extent of alien birds^{6,27}. For each climate axis, we set $\mu_y = 0$ (no effect of that axis) or $\mu_y = 30$ (strong effect of that axis), thus generating $n = 15$ unique combinations (i.e. excluding the case were $\mu_y =$

0 for all climate axes which corresponds to the model of random dispersal). From these 15 combinations we generated two 'climate-matching models'.

First, for each alien species we identified the combination of climate axes best predicting alien spread (climate-matching_{Optimal}). For alien species with multiple disjunct range fragments, we used the optimal combination of climate axes for each fragment. This model represents the optimal, but probably unrealistic scenario, in which the identity of the climate variables limiting the spread of an alien species at a site is known without error and does not need to be inferred from the native geographic range.

Second, for each alien species we selected the combination of climate axes best predicting spread of the native geographic range (climate-matching_{Native}). To identify the climate axes best predicting the native range, we used an identical simulation procedure to that used to model alien ranges, simulating the species' native range using each combination of climate axes outlined above ($n = 15$ combinations). We seeded the native range simulations using a randomly selected cell from the native range (or each native range fragment), stopping spread when the observed native range size was reached. Having identified the combination of climate axes with the highest predictive power for a species' native range, we then used this same combination of climate axes to simulate that species' alien range. The climate-matching_{Native} model thus assumes that the same set of climate variables are responsible for limiting both the native and alien geographic range.

The environmental resistance model assumes that the suitability of a site for invasion increases with the biotic similarity to the site/s where an alien species is established. In contrast, the climate-matching model assumes that the suitability of a site for invasion increases with climatic similarity to the mean conditions in the native range.

To provide a more direct comparison to the environmental resistance model, we implemented a ‘climate-resistance model’ where the suitability of a site for invasion increases with climatic similarity to the conditions at the site/s where an alien species established. In practice, this was done by replacing C_y in Equation 2 with the climate conditions at the site where each alien population was established. Biologically, this climate-resistance model may represent a scenario in which an invading species rapidly shifts its ecological niche to match the conditions at the site of establishment¹³. For the climate-resistance model we conducted two sets of simulations: using either the optimal combination of climate axes for predicting the spread of each alien range fragment (climate-resistance_{Optimal}) or using the combination of climate axes best predicting the spread of the native range (climate-resistance_{Native}).

All models used to predict alien dispersal are summarised in Extended Data Table 1.

Quantifying model predictive accuracy

For each model of range spread (random dispersal, climate-matching_{Optimal}, climate-matching_{Native}, climate-resistance_{Optimal}, climate-resistance_{Native}) we performed $n = 100$ replicate simulations. For each simulation, we quantified predictive accuracy as the proportional spatial overlap (Ov) between the simulated and observed range (i.e. proportion of cells where the species is present that are invaded in the simulation) and then calculated the median and 95% confidence interval (CI) in Ov scores across simulations. We compared the accuracy of models in three ways. First, we used a paired t-test to compare species Ov scores between the environmental resistance model and each competing model. Second, we calculated the number of species where each model performs

best (Extended Data Fig. 3). Third, for each species we tested whether the environmental resistance model performs significantly better or worse than the competing model, based on the 95% confidence in Ov scores of the competing model.

Predicting the extent of range spread

For each site of alien population establishment, we identified all grid cells falling within a given threshold of environmental resistance relative to the focal cell, and then used the 'components' function in the 'igraph' package v1.2.5⁴⁸ to identify the connected subset of cells (i.e. the cells that could be reached from the establishment site via spread through only adjacent cells). The number of connected cells indicates the potential size the alien range may attain if spread is prevented beyond a given environmental resistance threshold. For alien species (or fragments of alien species' ranges) lacking information on sites of population establishment, we repeated this analysis using each cell in the range (or range fragment) as the potential source of spread and calculated the mean expected range size. To reduce computational burden, for alien ranges larger than 100 cells, we selected 100 cells at random without replacement as potential sources of spread. In this analysis, the failure of species to spread beyond sites of establishment is informative and we therefore include all alien range fragments ($n = 1704$) and species ($n = 339$). Failure to include these range restricted species would underestimate the environmental resistance threshold that best predicts the spread of species.

We systematically explored environmental resistance thresholds from 0 to 0.8 in increments of 0.01. We did not consider environmental resistance thresholds greater than 0.8 because this was computationally expensive and our systematic search showed that a

value of 0.8 already led to simulated ranges far larger than those observed. We identified the threshold of environmental resistance across species that minimises prediction error by fitting an ordinary least squares regression of simulated against observed range size and calculating the root-mean-square error ($\text{Error}_{\text{OLS}}$). $\text{Error}_{\text{OLS}}$ penalizes over- and under-prediction of range sizes equally, but for managing risks from species invasions, underpredicting the extent of spread is a more serious concern. We therefore also used the *quantreg* R package to fit a quantile regression (99% percentile), with the fitted line defining the upper boundary of the relationship between simulated and observed range size. We then calculated error as the mean absolute deviation between this fitted line and the line of unity ($\text{Error}_{\text{Quantile}}$).

We assessed the relationship between simulated and observed range size at the level of individual range fragments ($n = 1704$), rather than the total range size for each species ($n = 339$). This is because variation in total range size is dominated by differences in the number of populations that were established (i.e. propagule pressure)³⁹ and thus the number of range fragments⁴¹. Our analysis shows that using total range size would lead to the appearance of a much stronger correspondence between observed and predicted range size (because both are driven by propagule pressure) and a lower estimate for $\text{Error}_{\text{Quantile}}$, thus underestimating the area at risk of invasion compared to when individual range fragments are modelled (Extended Data Fig. 5).

Projecting global risks of invasion

Having identified an environmental resistance threshold of $\text{ER} = 0.49$ as a 'safe limit' beyond which spread is unlikely (Extended Data Fig. 4c), we proceeded to map the locations (i.e. grid

cells) at potential risk of invasion from the spread of existing alien birds. A given cell may be at risk of invasion from the spread of a single species from multiple sources (i.e. when the alien species established at multiple sites and can spread from each of these) or from multiple species. For each grid cell i the total probability of invasion $P(I)$ for a species spreading from n sources was quantified as,

$$P(I) = 1 - \prod_{j=1}^n 1 - p_j$$

Where p_j is the probability of invasion from source j . For each cell we then summed invasion probabilities $P(I)$ across species to quantify the projected alien richness.

Data availability

Native and alien birds' geographic range data is available from <http://www.datazone.birdlife.org/> and The Global Avian Invasions Atlas (DOI: 10.1038/sdata.2017.41) respectively.

Code availability

All custom scripts will be made available with the final version of this article.

References

1. Bellard, C., Cassey, P. & Blackburn, T. M. Alien species as a driver of recent extinctions. *Biol. Lett.* **12**, 20150623 (2016).
2. Seebens, H. *et al.* No saturation in the accumulation of alien species worldwide. *Nat. Commun.* **8**, 14435 (2017).
3. Paini, D. R. *et al.* Global threat to agriculture from invasive species. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 7575–7579 (2016).
4. Bradshaw, C. J. A. *et al.* Massive yet grossly underestimated global costs of invasive insects. *Nat. Commun.* **7**, 12986 (2016).
5. Petitpierre, B. *et al.* Climatic niche shifts are rare among terrestrial plant invaders. *Science* **335**, 1344–1348 (2012).
6. Abellán, P., Tella, J. L., Carrete, M., Cardador, L. & Anadón, J. D. Climate matching drives spread rate but not establishment success in recent unintentional bird introductions. *Proceedings of the National Academy of Sciences* vol. 114 9385–9390 (2017).
7. Strubbe, D., Broennimann, O., Chiron, F. & Matthysen, E. Niche conservatism in non-native birds in Europe: niche unfilling rather than niche expansion. *Global Ecology and Biogeography* vol. 22 962–970 (2013).
8. Jiménez-Valverde, A. *et al.* Use of niche models in invasive species risk assessments. *Biological Invasions* vol. 13 2785–2797 (2011).
9. Guisan, A., Petitpierre, B., Broennimann, O., Daehler, C. & Kueffer, C. Unifying niche shift studies: insights from biological invasions. *Trends Ecol. Evol.* **29**, 260–269 (2014).

10. Qiao, H. *et al.* An evaluation of transferability of ecological niche models. *Ecography* vol. 42 521–534 (2019).
11. Cardador, L. & Blackburn, T. M. A global assessment of human influence on niche shifts and risk predictions of bird invasions. *Global Ecology and Biogeography* (2020) doi:10.1111/geb.13166.
12. Early, R. & Sax, D. F. Climatic niche shifts between species' native and naturalized ranges raise concern for ecological forecasts during invasions and climate change. *Global Ecology and Biogeography* vol. 23 1356–1365 (2014).
13. Atwater, D. Z., Ervine, C. & Barney, J. N. Climatic niche shifts are common in introduced plants. *Nat Ecol Evol* **2**, 34–43 (2018).
14. Parravicini, V., Azzurro, E., Kulbicki, M. & Belmaker, J. Niche shift can impair the ability to predict invasion risk in the marine realm: an illustration using Mediterranean fish invaders. *Ecol. Lett.* **18**, 246–253 (2015).
15. Gallagher, R. V., Beaumont, L. J., Hughes, L. & Leishman, M. R. Evidence for climatic niche and biome shifts between native and novel ranges in plant species introduced to Australia. *Journal of Ecology* vol. 98 790–799 (2010).
16. Su, S., Cassey, P., Dyer, E. E. & Blackburn, T. M. Geographical range expansion of alien birds and environmental matching. *Ibis* vol. 159 193–203 (2017).
17. Mothes, C. C., Stroud, J. T., Clements, S. L. & Searcy, C. A. Evaluating ecological niche model accuracy in predicting biotic invasions using South Florida's exotic lizard community. *Journal of Biogeography* vol. 46 432–441 (2019).
18. Hill, M. P., Gallardo, B. & Terblanche, J. S. A global assessment of climatic niche shifts and human influence in insect invasions. *Global Ecology and Biogeography*

vol. 26 679–689 (2017).

19. Lomolino, M. V., Riddle, B. R., Whittaker, R. J. & Brown, J. H. *Biogeography*. (Sinauer Associates Incorporated, 2010).
20. Nekola, J. C. & White, P. S. The distance decay of similarity in biogeography and ecology. *Journal of Biogeography* vol. 26 867–878 (1999).
21. Rappoport, E. H. *Areography: geographical strategies of species* (Elsevier Science Limited, 1982).
22. Graves, G. R. & Rahbek, C. Source pool geometry and the assembly of continental avifaunas. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 7871–7876 (2005).
23. Ruggiero, A., Lawton, J. H. & Blackburn, T. M. The geographic ranges of mammalian species in South America: spatial patterns in environmental resistance and anisotropy. *Journal of Biogeography* vol. 25 1093–1103 (1998).
24. Borregaard, M. K., Graves, G. R. & Rahbek, C. Dispersion fields reveal the compositional structure of South American vertebrate assemblages. *Nat. Commun.* **11**, 491 (2020).
25. Dyer, E. E. *et al.* The Global Distribution and Drivers of Alien Bird Species Richness. *PLoS Biol.* **15**, e2000942 (2017).
26. Dyer, E. E., Redding, D. W. & Blackburn, T. M. The global avian invasions atlas, a database of alien bird distributions worldwide. *Sci Data* **4**, 170041 (2017).
27. Redding, D. W. *et al.* Location-level processes drive the establishment of alien bird populations worldwide. *Nature* **571**, 103–106 (2019).
28. Birdlife International. Bird species distribution maps of the world. Version 2.0. Cambridge/Arlington, TX: BirdLife International/NatureServe. (2012).

29. Jetz, W. & Rahbek, C. Geometric constraints explain much of the species richness pattern in African birds. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 5661–5666 (2001).
30. Pigot, A. L., Owens, I. P. F. & Orme, C. D. L. The environmental limits to geographic range expansion in birds. *Ecology Letters* vol. 13 705–715 (2010).
31. Warren, D. L., Glor, R. E. & Turelli, M. Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. *Evolution* **62**, 2868–2883 (2008).
32. Bahn, V. & McGill, B. J. Testing the predictive performance of distribution models. *Oikos* vol. 122 321–331 (2013).
33. Beale, C. M., Lennon, J. J. & Gimona, A. Opening the climate envelope reveals no macroscale associations with climate in European birds. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 14908–14912 (2008).
34. Petitpierre, B., Broennimann, O., Kueffer, C., Daehler, C. & Guisan, A. Selecting predictors to maximize the transferability of species distribution models: lessons from cross-continental plant invasions. *Global Ecology and Biogeography* vol. 26 275–287 (2017).
35. Early, R. *et al.* Global threats from invasive alien species in the twenty-first century and national response capacities. *Nat. Commun.* **7**, 12485 (2016).
36. Bellard, C., Leroy, B., Thuiller, W., -F. Rysman, J. & Courchamp, F. Major drivers of invasion risks throughout the world. *Ecosphere* vol. 7 (2016).
37. Fournier, A., Penone, C., Pennino, M. G. & Courchamp, F. Predicting future invaders and future invasions. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 7905–7910 (2019).
38. Byers, J. E. *et al.* Invasion Expansion: Time since introduction best predicts

- global ranges of marine invaders. *Sci. Rep.* **5**, 12436 (2015).
39. Dyer, E. E. *et al.* A global analysis of the determinants of alien geographical range size in birds. *Global Ecology and Biogeography* vol. 25 1346–1355 (2016).
 40. Meyer, C., Kreft, H., Guralnick, R. & Jetz, W. Global priorities for an effective information basis of biodiversity distributions. *Nat. Commun.* **6**, 8221 (2015).
 41. Pigot, A. L. *et al.* Species invasions and the phylogenetic signal in geographical range size. *Global Ecology and Biogeography* vol. 27 1080–1092 (2018).
 42. Hurlbert, A. H. & White, E. P. Disparity between range map- and survey-based analyses of species richness: patterns, processes and implications. *Ecology Letters* vol. 8 319–327 (2005).
 43. Ruggiero, A., Lawton, J. H. & Blackburn, T. M. The geographic ranges of mammalian species in South America: spatial patterns in environmental resistance and anisotropy. *Journal of Biogeography* vol. 25 1093–1103 (1998).
 44. Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G. & Jarvis, A. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* vol. 25 1965–1978 (2005).
 45. Jiménez-Valverde, A. *et al.* Dominant climate influences on North American bird distributions. *Global Ecology and Biogeography* vol. 20 114–118 (2011).
 46. *R Core Team*: *A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. (2020).
 47. Gentleman R, Whalen E, Huber W, Falcon S. *igraph: A package to handle graph data structures*. R package. (2020).
 48. Csardi G, N. T. The igraph software package for complex network research.

International Journal of Complex Systems **1695**, (2006).

Extended Data Table. 1. Models used to predict alien dispersal and their predictive accuracy across species.

| Model | Description | Median predictive accuracy |
|---------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------|
| Random dispersal | Null model of spread, species invade adjacent cells with equal probability. | 67% ($\pm 15\%$ SD) |
| Environmental resistance | Invasion probability increases with biotic similarity to the site of establishment. | 78% ($\pm 19\%$ SD) |
| Climate matching_{Optimal} | Invasion probability increases with climatic similarity to mean conditions occupied in the native range. Alien ranges are simulated using the combination of climatic axes that best predict alien spread. The identity of these axes can vary across species and across different fragments of each species range. | 75% ($\pm 14\%$ SD) |
| Climate matching_{Native} | Invasion probability increases with climatic similarity to mean conditions occupied in the native range. Alien ranges are simulated using the combination of climatic axes that best predict the native range of each species spread. The identity of these axes can thus vary across alien species but are held constant across the different fragments of an alien species range. | 65% ($\pm 17\%$ SD) |
| Climate resistance_{Optimal} | Invasion probability increases with climatic similarity to the conditions occupied at the site of establishment. Alien ranges are simulated using the combination of climatic axes that best predict alien spread. The identity of these axes can vary across species and across different fragments of each species range. | 75% ($\pm 15\%$ SD) |
| Climate resistance_{Native} | Invasion probability increases with climatic similarity to the conditions occupied at the site of establishment. Alien ranges are simulated using the combination of climatic axes that best predict the native range of each species spread. The identity of these axes can thus vary across alien species but are held constant across the different fragments of an alien species range. | 67% ($\pm 16\%$ SD) |

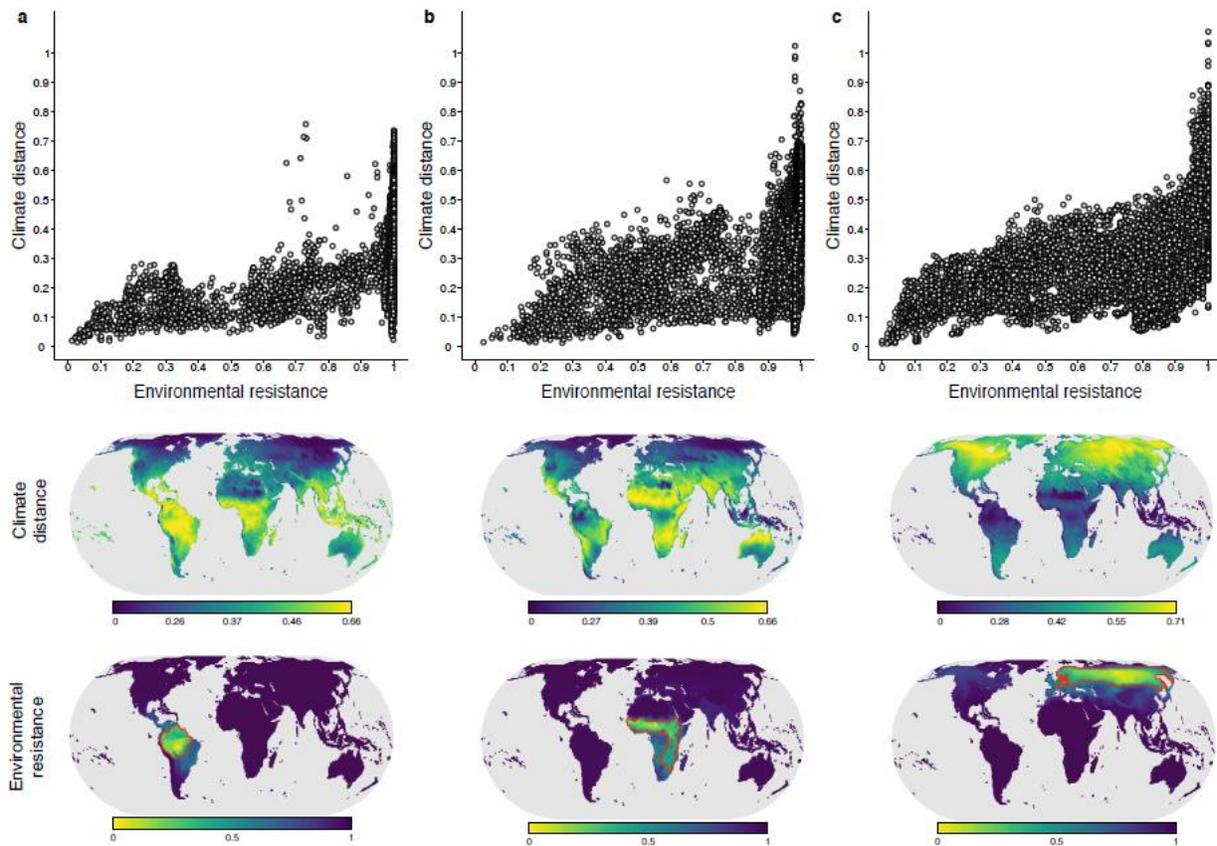
Extended Data Table. 2. Trait loadings for bioclimatic space and proportion of variance accounted for by each principal component axis.

| | PC1 | PC2 | PC3 | PC4 | PC5 | PC6 | PC7 |
|------------------------|--------|--------|--------|--------|--------|--------|--------|
| Bio1 | 0.437 | -0.26 | 0.08 | 0.187 | -0.836 | -0.033 | -0.033 |
| Bio2 | -0.022 | -0.634 | 0.499 | -0.445 | 0.116 | 0.354 | 0.105 |
| Bio3 | 0.477 | -0.163 | 0.196 | -0.04 | 0.332 | -0.751 | 0.18 |
| Bio4 | -0.497 | -0.009 | -0.025 | -0.102 | -0.302 | -0.266 | 0.761 |
| Bio7 | -0.48 | -0.148 | 0.054 | -0.293 | -0.222 | -0.487 | -0.61 |
| Bio12 | 0.319 | 0.404 | -0.202 | -0.814 | -0.164 | 0.045 | 0.048 |
| Bio15 | 0.048 | -0.564 | -0.813 | -0.079 | 0.104 | -0.002 | 0.028 |
| Proportion of Variance | 0.564 | 0.276 | 0.081 | 0.055 | 0.02 | 0.003 | 0.001 |
| Cumulative Proportion | 0.564 | 0.84 | 0.92 | 0.975 | 0.996 | 0.999 | 1 |

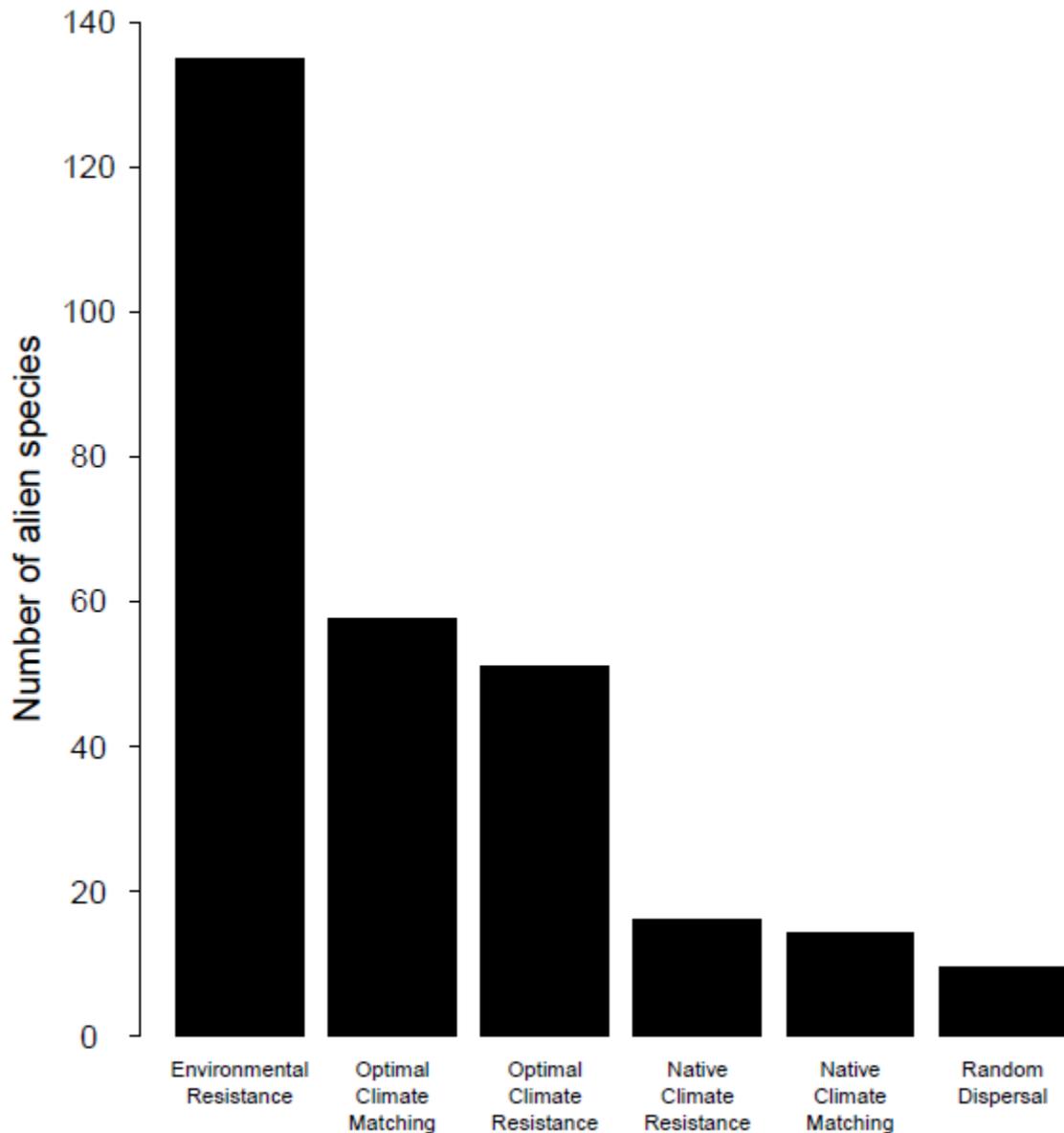
| | | Alien | | | | | | | | | | | | | | |
|--------|---------|-------|------|-------|------|-------|-------|--------|------|-------|-------|--------|-------|--------|--------|---------|
| | | PC 1 | PC 2 | PC 12 | PC 3 | PC 13 | PC 23 | PC 123 | PC 4 | PC 14 | PC 24 | PC 124 | PC 34 | PC 134 | PC 234 | PC 1234 |
| Native | PC 1 | 101 | 16 | 8 | 18 | 8 | 8 | 2 | 24 | 6 | 1 | 2 | 2 | 1 | 0 | 1 |
| | PC 2 | 33 | 6 | 2 | 2 | 0 | 0 | 0 | 4 | 1 | 0 | 0 | 4 | 0 | 0 | 0 |
| | PC 12 | 78 | 11 | 6 | 13 | 2 | 2 | 2 | 7 | 3 | 2 | 5 | 2 | 0 | 1 | 0 |
| | PC 3 | 43 | 7 | 4 | 8 | 2 | 0 | 0 | 4 | 1 | 0 | 1 | 0 | 2 | 0 | 0 |
| | PC 13 | 12 | 3 | 2 | 4 | 1 | 2 | 0 | 3 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| | PC 23 | 32 | 4 | 1 | 5 | 0 | 1 | 0 | 7 | 1 | 1 | 0 | 0 | 1 | 0 | 0 |
| | PC 123 | 91 | 11 | 10 | 11 | 5 | 1 | 2 | 12 | 1 | 2 | 1 | 0 | 0 | 0 | 1 |
| | PC 4 | 3 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| | PC 14 | 32 | 5 | 5 | 5 | 3 | 0 | 1 | 5 | 1 | 0 | 1 | 0 | 1 | 1 | 0 |
| | PC 24 | 66 | 8 | 2 | 7 | 2 | 6 | 1 | 9 | 0 | 0 | 1 | 1 | 1 | 0 | 1 |
| | PC 124 | 163 | 30 | 15 | 22 | 2 | 7 | 1 | 26 | 8 | 2 | 1 | 6 | 2 | 3 | 2 |
| | PC 34 | 14 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 |
| | PC 134 | 15 | 10 | 0 | 4 | 3 | 2 | 0 | 2 | 1 | 1 | 2 | 1 | 1 | 0 | 0 |
| | PC 234 | 9 | 1 | 0 | 2 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | PC 1234 | 35 | 3 | 1 | 5 | 2 | 0 | 1 | 3 | 1 | 1 | 0 | 2 | 1 | 0 | 1 |

Extended Data Fig. 1. The optimal combination of principal component climate axes predicting spread for native (rows) and alien (columns) ranges. For each alien species' range fragment ($n = 1278$ fragments) the set of climate axes best predicting spread was identified, as well as the set of climate axes best predicting the native range. Each cell in the matrix defined by position (x,y) , indicates the number of alien range fragments where x is the best model and y is the best model in the native range. Diagonal elements indicate cases

where the same set of climate axes limit spread in the native and alien range of a species. In cases where multiple sets of climate axes predicted spread equally well, fragments were assigned to multiple cells but weighted by the inverse of the number of ties. Thus, the sum of the entries in the matrix is equal to the total number of range fragments.

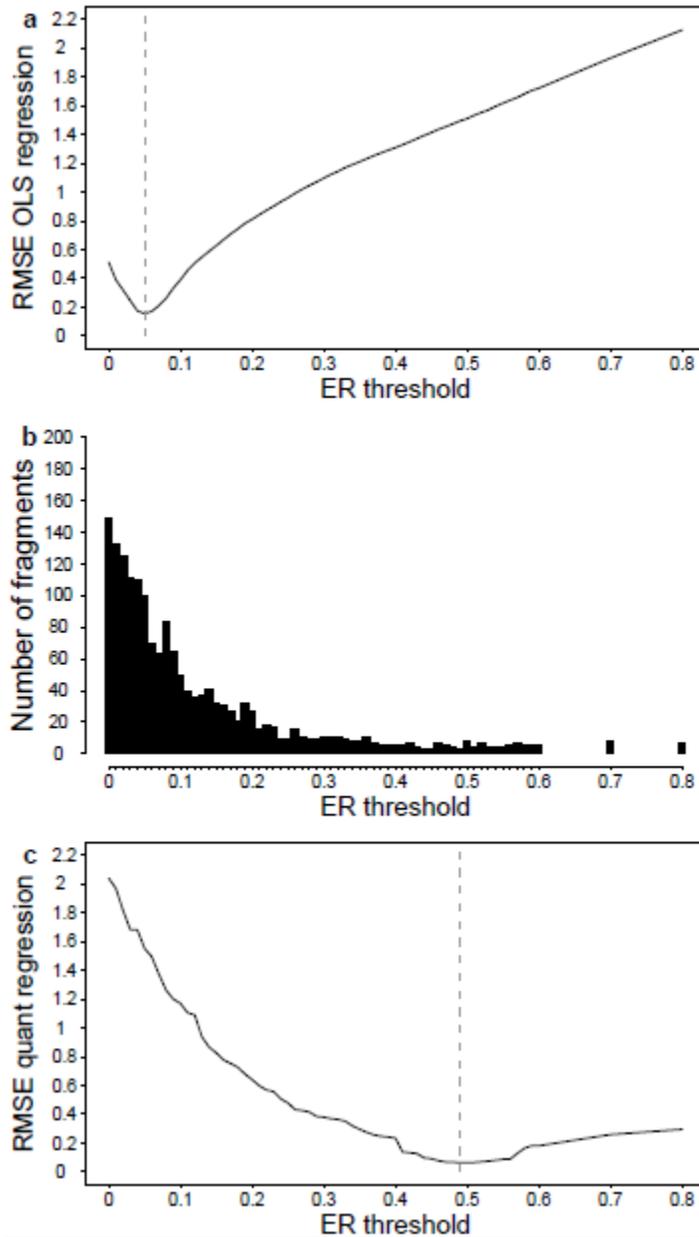


Extended Data Fig. 2. Patterns of environmental and climate distance. For three exemplar focal cells in the (a) Amazon, (b) Sahel and (c) Siberia (Fig. 1), environmental resistance and climate distance were calculated to all other cells. Environmental resistance is $1 - \text{biotic similarity}$, where biotic similarity is the proportion of species present in the focal cell that is also present in each other cell. Climate distance is the Euclidean distance in four-dimensional climate space between the conditions in the focal cell and each other cell. Bivariate plots show the relationship between the climatic distance and environmental resistance of each cell relative to the focal cell, and these patterns are shown on the maps below.



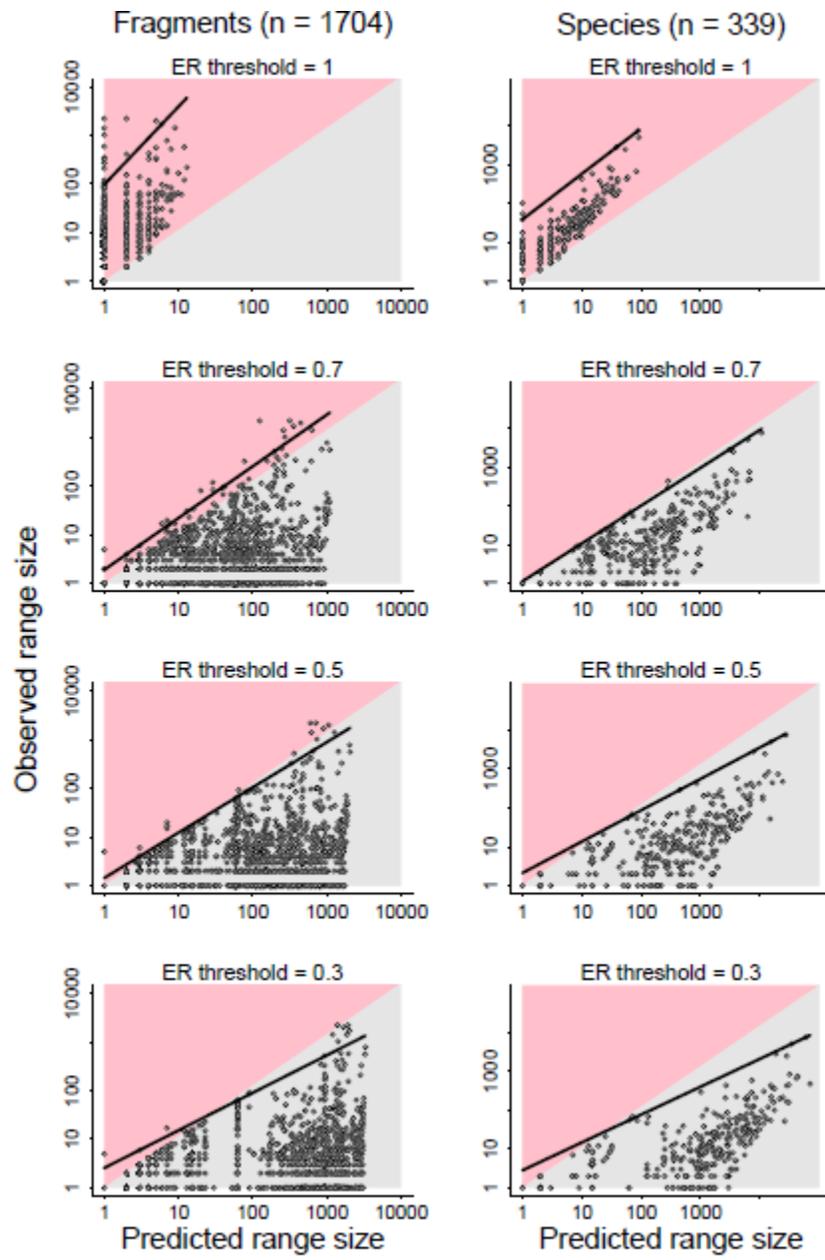
Extended Data Fig. 3. The best fitting model across species. For each species, ($n = 283$), the model with the highest predictive accuracy for the species' alien range was identified from the following models: random dispersal, environmental resistance model, climate matching_{Optimal}, climate matching_{Native}, climate resistance_{Optimal}, climate resistance_{Native} (Extended Data Table 1). Predictive accuracy was rounded to two decimal places before comparing models to prevent detecting marginal differences in accuracy. Where multiple

models had the same predictive accuracy, that species was assigned to multiple models but weighted by the inverse of the number of ties. Thus, the sum of the entries is equal to the total number of species.

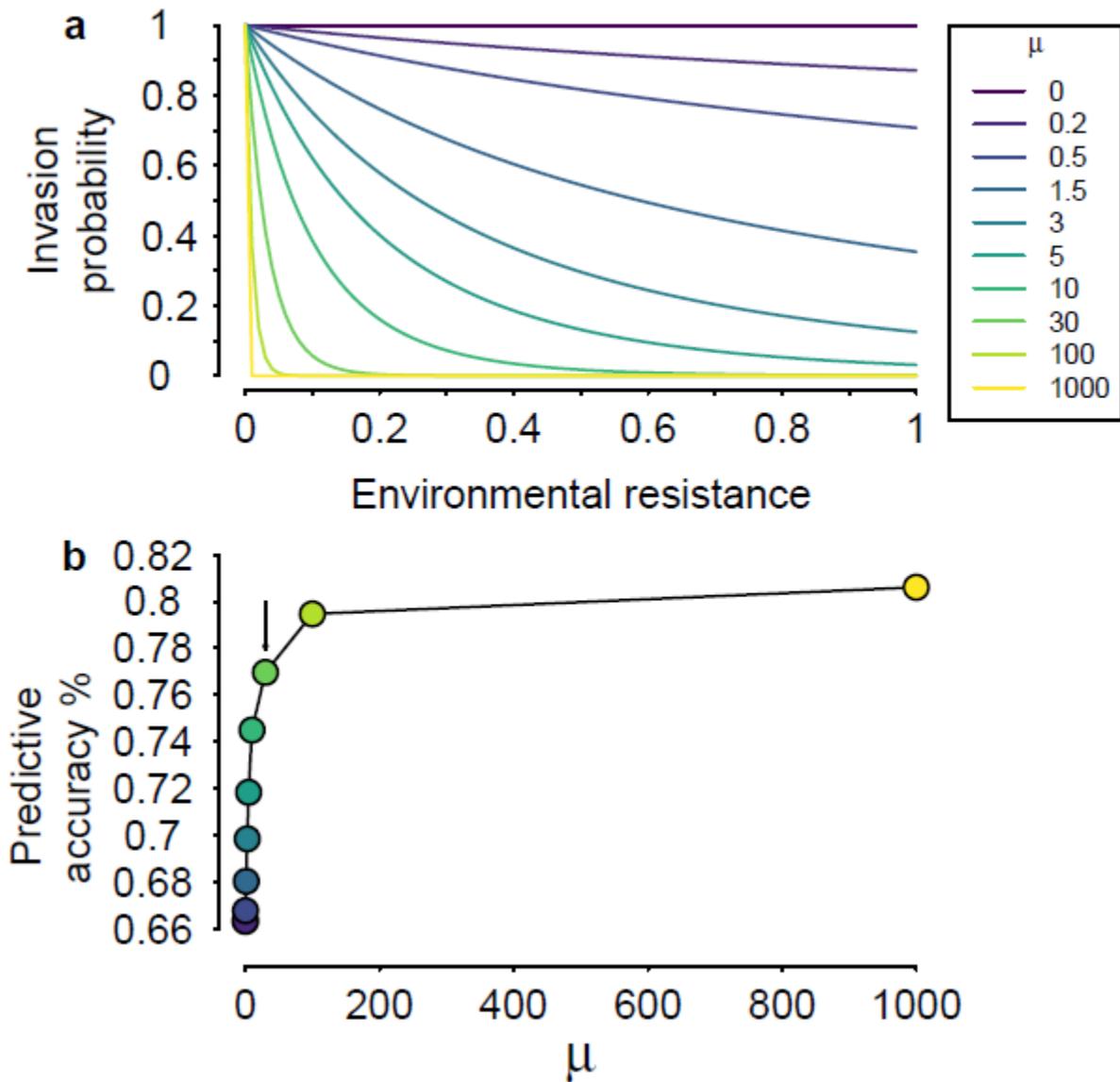


Extended Data Fig. 4. Optimal environmental resistance (ER) thresholds for predicting the extent of alien spread at the level of species range fragments ($n = 1704$). **a**, the root mean square error in mean predicted range size (ordinary least squares regression) under different environmental resistance thresholds; **b**, the optimal environmental resistance threshold for predicting the size of each alien range fragment; **c**, error in the upper boundary of predicted range size (quantile regression, 0.99 percentile)

under different environmental resistance thresholds. The ER threshold minimising error is shown (grey dashed line).



Extended Data Fig. 5. Predicted against observed range size for different environmental resistance thresholds. Red regions indicate range sizes that are under-predicted. Lines indicate the quantile regression (0.99 percentile) of observed against predicted range size. Range size was calculated for individual fragments of species' ranges ($n = 1704$ fragments, left column) or for total species range size ($n = 339$ species, right column).



Extended Data Fig. 6. The effects of μ on invasion and patterns of spread. a, relative probability of invasion as a function of environmental resistance for different values of μ ; **b**, median predictive accuracy across species ($n = 283$) of the environmental resistance model under different values of μ . When μ is high (e.g. $\mu = 30$), patterns of spread are strongly deterministic. When $\mu = 0$, patterns of spread are independent of environmental resistance corresponding to the null model of random dispersal. Increasing μ above 30 only resulted

in a small increase in the median overlap scores across species. Therefore, $\mu = 30$ represents a highly deterministic scenario where further increases in μ result in minimal improvement in model predictive ability.

Acknowledgements:

A.L.P. is supported by a Royal Society University Research Fellowship. Initial funding for data collection was provided by a grant from the Leverhulme Trust (RF/2/RFG/2010/0016) (E.E.D.), with additional support from a UCL IMPACT studentship (10989) (E.E.D.), a Leverhulme Trust grant (RPG-2015-392) (T.M.B., A.L.P. and E.E.D.), and from a King Saud University Distinguished Scientist Research Fellowship (T.M.B. and E.E.D.)

Author Information

Affiliations

Centre for Biodiversity and Environment Research, Department of Genetics, Evolution and Environment, University College London, Gower Street, London, WC1E 6BT, UK.

Rebecca S L Lovell

Tim M Blackburn

Ellie E Dyer

Alex L Pigot

Author Contributions:

R.S.L.L. and A.L.P. conceived the study, performed the analysis, and wrote the first draft of the manuscript. T. M.B. and E.E.D. assembled the alien distribution data. All authors contributed to the writing of the final manuscript.

Competing Interests:

The authors declare no competing financial interests.

Correspondence

Correspondence and requests for information should be addressed to Rebecca Lovell (Rebecca.Lovell.15@ucl.ac.uk) and Alex Pigot (a.pigot@ucl.ac.uk)