

METHOD

Target-group backgrounds prove effective at correcting sampling bias in Maxent models

Robert A. Barber^{1,2}  | Stuart G. Ball³ | Roger K. A. Morris³ | Francis Gilbert²

¹Silwood Park, Imperial College London, Ascot, UK

²School of Life Sciences, University of Nottingham, Nottingham, UK

³Dipterists Forum, Hoverfly Recording Scheme, c/o Biological Records Centre, Centre for Ecology & Hydrology, Wallingford, Oxfordshire, UK

Correspondence

Robert A. Barber, Silwood Park, Imperial College London, Ascot, UK.

Email: r.barber19@imperial.ac.uk

Editor: Boris Leroy

Abstract

Aim: Accounting for sampling bias is the greatest challenge facing presence-only and presence-background species distribution models; no matter what type of model is chosen, using biased data will mask the true relationship between occurrences and environmental predictors. To address this issue, we review four established bias correction techniques, using empirical occurrences with known sampling effort, and virtual species with known distributions.

Innovation: Occurrence data come from a national recording scheme of hoverflies (*Syrphidae*) in Great Britain, spanning 1983–2002. Target-group backgrounds, distance-restricted backgrounds, travel time to cities and human population density were used to account for sampling bias in 58 species of hoverfly. Distributions generated by bias correction techniques were compared in geographical space to the distribution produced accounting for known sampling effort, using Schoener's distance, centroid shifts and range size changes. To validate our results, we performed the same comparisons using 50 randomly generated virtual species. We used sampling effort from the hoverfly recording scheme to structure our biased sampling regime, emulating complex real-life sampling bias.

Main conclusions: Models made without any correction typically produced distributions that mapped sampling effort rather than the underlying habitat suitability. Target-group backgrounds performed the best at emulating sampling effort and unbiased virtual occurrences, but also showed signs of overcompensation in places. Other methods performed better than no-correction, but often differences were difficult to visually detect. In line with previous studies, when sampling effort is unknown, target-group backgrounds provide a useful tool for reducing the effect of sampling bias. Models should be visually inspected for biological realism to identify any areas of potential overcompensation. Given the disparity between corrected and un-corrected models, sampling bias constitutes a major source of error in species distribution modelling, and more research is needed to confidently address the issue.

KEYWORDS

citizen science, Maxent, presence-only data, sampling bias, species distribution modelling, *Syrphidae*, target-group background

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *Diversity and Distributions* published by John Wiley & Sons Ltd.

1 | INTRODUCTION

Understanding macro-geographic processes is becoming ever more important in the challenges facing biodiversity conservation. The threats of climate change, species invasions and continued habitat degradation mean we need accurate methods for understanding distributions and their responses to changing scenarios (Jeschke & Strayer, 2008; Pimentel et al., 2001). Species distribution modelling (SDM) aims to relate the occurrence records of a species to abiotic variables such as climate and land use (Elith & Leathwick, 2009; Phillips, 2004). By effectively describing the environmental niche that a species occupies, models can be used to predict geographic distributions (Soberón & Peterson, 2005).

Many SDMs used today utilize presence-only data, most notably using Maximum Entropy statistical models (Phillips, 2004; Phillips et al., 2006). Because it is simple to use and performs well with small sample sizes, Maxent has driven the use of SDMs in conservation. An important assumption of modelling with presence-only data is that the modelled region has been sampled randomly, with environmental conditions represented in terms of their availability (Phillips et al., 2009); due to the haphazard sampling of museum records and citizen science databases, this is rarely the case (Newbold, 2010; Ponder et al., 2001). Often the locations that are sampled the most are in the most accessible areas, in protected areas, or where the greatest number of species can be observed (Kadmon et al., 2004; Reddy & Dávalos, 2003). This in turn can cause distribution models to resemble sampling effort more closely than to the true distribution, complicating interpretation (Elith et al., 2011; Soberon & Nakamura, 2009). This phenomenon is typically referred to as sampling bias.

While sampling bias may be a result of geographic features, models are typically fitted in environmental space, and so sampling bias becomes a problem if the highly sampled regions correlate with specific climatic conditions (Phillips et al., 2009). This may be true in many instances, such as the proximity of road networks to ridge lines, or the association between cities and rivers. In these cases, the realized niche of a species is often disguised by the climatic variables of over-represented areas. Despite this, sampling bias is still often unaccounted for in distribution modelling (Yackulic et al., 2013).

There are two main methods commonly used to correct for sampling bias. The first is to manipulate the selection of background points used to determine environmental variation so that these share the same bias; this in effect increases the contribution of environmental variation from highly sampled areas (Ponder et al., 2001). The second option is to remove presences from highly clustered areas, known as spatial filtering (Boria et al., 2014; Veloz, 2009). Both methods have been shown to reduce the effect of sampling bias. We chose to focus on background manipulation rather than spatial filtering for two reasons. First, occurrences may be clustered for ecological reasons such as population structure (Dormann et al., 2007), and so removing records from highly clustered areas may

disguise the true patterns; second, presence records may often be sparse, and thus removing records can reduce the sample size unacceptably.

Previous evaluations of methods to account for sampling bias either have used simulated datasets with virtual species (e.g. Barbet-Massin et al., 2012; Moua et al., 2020; Ranc et al., 2017; Stolar & Nielson, 2015), estimations or simulations of sampling bias (e.g. Fourcade et al., 2014; Kramer-Schadt et al., 2013), or presence-only methods using presence-absence data (Fithian et al., 2015; Syfert et al., 2013). By having complete knowledge of the underlying population distribution, evaluating correction methods can be done confidently. However, artificially generated sampling bias can be simplistic and is often represented as gradients of sampling intensity across modelled regions. This does not reflect the complex patterns of bias observed in real life.

In contrast to the above studies, we have a dataset of field-collected data where the sampling bias is known. We compare bias correction methods here using the presence-only data of the Hoverfly Recording Scheme (Ball & Morris, 2012), which has a complete record of sampling effort for each 1-km square of the UK. The data come from a long-standing National Recording Scheme, with records contributed by skilled professionals, amateurs and increasingly by citizen scientists; the data therefore capture the true complexity of bias in sampling, affected by all manner of factors, such as accessibility and the site preferences of surveyors.

We used both empirical and virtual species to test three methods of bias correction via background manipulation: the use of records of similar taxa (the 'target-group' method: Phillips et al., 2009; Ponder et al., 2001), restricting the available background to a specified distance from records (Anderson & Raza, 2010; Phillips, 2008) and the use of potential covariates of sampling effort, such as human population density or the positions of roads (Dubos et al., 2021; El-Gabbas & Dormann, 2018; Monsarrat et al., 2019). We used the explicit knowledge of sampling effort to factor out bias in real distributions as a standard for comparison, and to create biased and unbiased occurrences for virtual species. To the best of our knowledge, this is the first attempt to evaluate bias correction techniques using real data with explicit knowledge of sampling effort and therefore presents a unique opportunity to test them against the complexity of real-life sampling bias.

2 | METHODS

2.1 | Empirical data

Our data came from the Hoverfly Recording Scheme (HRS) (Ball & Morris, 2012), which now contains more than 1 million records. The scheme aims to collate information on the distribution and ecology of hoverflies (Diptera, Syrphidae) throughout the UK. We chose to model 58 individual species with over 1000 records, spanning 1983–2002. The use of higher numbers of records leads to more accurate

models (Kadmon et al., 2003; Wisz et al., 2008), and by using multiple species, we aimed to reduce the effect of species-specific relationships with sampling bias (Fourcade et al., 2014; Stolar & Nielsen, 2015; Warton et al., 2013). The exact number of occurrences for each species is presented in Table 1.

Climatic variables for modelling were sourced from the WorldClim database (Version 2.1) (Flick & Hijmans, 2017); all 19 bioclim variables were initially considered, alongside elevation. Land use data were also incorporated, using the Land Cover Map 2000 (1-km dominant target class), as produced by the Centre for Ecology and Hydrology (Fuller et al., 2002). Target classes for each 1-km raster cell show 25 land use types representing different habitats throughout the UK. We assessed predictor variables for collinearity using the R package *virtualspecies* (Leroy et al., 2015), selecting nine independent variables with a Spearman's rank correlation less than 0.7 (Appendix S1: Figure S1.1) (Braunisch et al., 2013). Information on the environmental predictors is available in Appendix S1: Table S1.1. All processing of environmental variables and occurrences

were coded in R version 4.0.5 (R Core Team, 2021) and projected in the OSGB 1936 coordinate system.

2.2 | Maxent settings

All species distribution modelling was performed using MAXENT (version 3.4.1) in the R package *dismo* (Hijmans et al., 2017). Because we modelled multiple virtual and empirical species across different methods, we chose to use Maxent's default regularization and feature settings to maintain consistency. Tuning individual models can improve their predictive ability (Merow et al., 2013; Radosavljevic & Anderson, 2014), including their response to sampling bias (Anderson & Gonzalez, 2011), which may in our case mask the effect of each bias correction method. However, for practical application we recommend that Maxent users should always carefully consider model settings and visually assess models for biological realism (Fourcade et al., 2018).

Species	Occurrences	Species	Occurrences
<i>Baccha elongata</i>	1810	<i>Melanostoma mellinum</i>	3976
<i>Cheilosia albitarsis agg</i>	2471	<i>Melanostoma scalare</i>	5312
<i>Cheilosia bergenstammi</i>	1022	<i>Meliscaeva auricollis</i>	1127
<i>Cheilosia illustrata</i>	2125	<i>Meliscaeva cinctella</i>	1461
<i>Cheilosia pagana</i>	2513	<i>Merodon equestris</i>	1075
<i>Cheilosia variabilis</i>	1221	<i>Myathropa florea</i>	3018
<i>Cheilosia vernalis</i>	1101	<i>Neoascia podagrica</i>	2648
<i>Chrysogaster solstitialis</i>	1633	<i>Neoascia tenur</i>	1385
<i>Chrysotoxum bicinctum</i>	1756	<i>Pipizella viduata</i>	1036
<i>Dasysyrphus albostrigatus</i>	1140	<i>Platycheirus albimanus</i>	5925
<i>Dasysyrphus venustus</i>	1288	<i>Platycheirus angustatus</i>	1051
<i>Epistrophe eligans</i>	2021	<i>Platycheirus clypeatus agg</i>	2807
<i>Epistrophe grossulariae</i>	1499	<i>Platycheirus granditarsus</i>	2096
<i>Episyrphus balteatus</i>	7741	<i>Platycheirus manicatus</i>	1635
<i>Eristalinus sepulchralis</i>	1336	<i>Platycheirus peltatus agg</i>	1457
<i>Eristalis arbustorum</i>	4439	<i>Platycheirus rosarum</i>	1269
<i>Eristalis horticola</i>	1940	<i>Platycheirus scutatus agg</i>	2089
<i>Eristalis interruptus</i>	2748	<i>Rhingia campestris</i>	4936
<i>Eristalis intricarius</i>	2280	<i>Scaeva pyrastris</i>	1799
<i>Eristalis pertinax</i>	6787	<i>Sericomyia silentis</i>	2001
<i>Eristalis tenax</i>	5623	<i>Sphaerophoria scripta</i>	2970
<i>Eupeodes corollae</i>	2386	<i>Syrpitta pipiens</i>	6317
<i>Eupeodes luniger agg</i>	2098	<i>Syrphus ribesii</i>	4769
<i>Ferdinandea cuprea</i>	1064	<i>Syrphus torvus</i>	1357
<i>Helophilus pendulus</i>	5702	<i>Syrphus vitripennis</i>	2860
<i>Lejogaster metallina</i>	1105	<i>Volucella bombylans</i>	2145
<i>Leucozona glaucia</i>	1281	<i>Volucella pellucens</i>	3150
<i>Leucozona lucorum</i>	2575	<i>Xylota segnis</i>	3526
<i>Melanogaster hirtella</i>	1716	<i>Xylota sylvorum</i>	1283

TABLE 1 The number of occurrences of each species of Syrphidae in Great Britain used to generate species distribution models. Records were collected during the period 1983–2002

2.3 | Maxent background

Maxent uniformly samples 10,000 background points to estimate environmental variation across the area of interest, which is used to determine the habitat preference of species. When sampling is concentrated in small areas, most of the environmental variation will come from unsurveyed regions, contributing to the effect of sampling bias (Barve et al., 2011). Because many inaccessible regions have extreme environmental values, the consequence is to skew environmental gradients, leading to overestimation where presences occur (Lobo et al., 2010). To prevent this, background point selection can be influenced using a probability surface that mimics sampling bias, so that highly sampled areas receive a greater number of background points.

For all correction techniques, we sampled 10,000 background points without replacement, using the function *randomPoints* from the R package *dismo*. For models with no bias correction, we used a uniform background (Figure 1b), which is the default setting for maxent. To effectively counteract bias, we created a background based on sampling effort of the hoverfly recording scheme. The sampling effort of the HRS is the number of visits to each 1 km². To create a background, these data were first converted into a continuous surface (Elith et al., 2010). This was achieved using 2D kernel density estimation in the R package *ks* (Duong, 2020), with cells weighted by the number of visits over the 20-year period. Resulting rasters were projected to the same resolution, area and coordinate system as environmental variables (Figure 1a). Background points for each species were then sampled using the probability surface and were supplied to Maxent for use when modelling.

2.4 | Target-group background

The target-group approach to sampling bias is the most prevalent in the literature and relies on using the collective records of similar taxa to estimate sampling effort for the focal species (Phillips et al., 2009; Ponder et al., 2001). The reasoning is that surveyors will display the same bias when sampling similar species, and so a larger dataset will contain information on general sampling effort for that taxon. It is important to carefully decide the scope of the target-group, because of changes in locations and seasons, the varying detectability among species, and because observers need to use the same methodology (Phillips et al., 2009; Ponder et al., 2001; Yackulic et al., 2013). Because of this, our aim was to test if target-group methods accurately reflect the distribution of sampling effort, when considering varying detectability and underlying patterns of species richness.

To generate a target-group background, we used the collective records of all the species of hoverfly within the HRS, following established practice (see Elith et al., 2010; Phillips et al., 2009). First, we removed duplicate records for each species, so that each cell value represented the total number of species sampled. We then followed the same processing steps as when generating the sample

effort background, using a 2D kernel density estimation to convert single points into a continuous probability surface that could be used to weight background point selection (Figure 1c).

2.5 | Radius-restricted background

An alternative method of bias correction is to restrict the area that Maxent uses to determine environmental variation, normally as a pre-specified radius from each record (Anderson & Raza, 2010; Phillips, 2008). This means that the model excludes obtaining background points from further away, focussing on available habitat close to sampled locations. This has been shown to improve model predictions (Acevedo et al., 2012; Anderson & Raza, 2010; Barve et al., 2011), but care must be taken to avoid a resulting lack of model accuracy when too few background points are used (Thuiller et al., 2004; VanDerWal et al., 2009).

Similar to choosing minimum distances for spatial filtering, the choice of radius size can be subjective (Aiello-Lammens et al., 2015), which can pose a significant dilemma to conservation practitioners. We selected 10 km as a potential to correct for sampling bias without excessively reducing the available background, but clarify our choice was also subjective. It may have been possible to assess multiple distances, and select the best for each species, but this information is typically unavailable to practitioners, and we wanted our assessment to accurately reflect real-world scenarios. Within the buffer region, each grid cell was equally likely to be selected as a background point. The background was created using the package *raster* (Hijmans, 2020) (Figure 1d).

2.6 | Covariates of sampling effort

In the case when a focal species lacks appropriate target-group data or has too few presence points for background restriction to be viable, another option is the use of potential covariates of sample effort. Maps of human population density and road networks can provide an alternative means of estimating sampling effort, and can be incorporated into SDMs (Guerra et al., 2013; Kadmon et al., 2004; Monsarrat et al., 2019). While accessibility maps may work well in some scenarios (Dubos et al., 2021), there is currently no review of their effectiveness against other bias correction methods.

Human population density was sourced from the Socioeconomic Data and Applications Center (SEDAC). The Gridded Population of the World (GPWv4) (CIESIN, 2016) models the distribution of human population as a smooth continuous surface, based primarily on population census tables. Trials showed that using population density in its raw form resulted in severe overcompensation, and therefore, the data were transformed to logarithms to reduce the scale of change between the most and least densely populated areas. The values of all pixels were shifted to positive values by adding the minimum cell value, to remove the negative logs of the smallest population densities. The resulting probability surface is shown in Figure 1e.

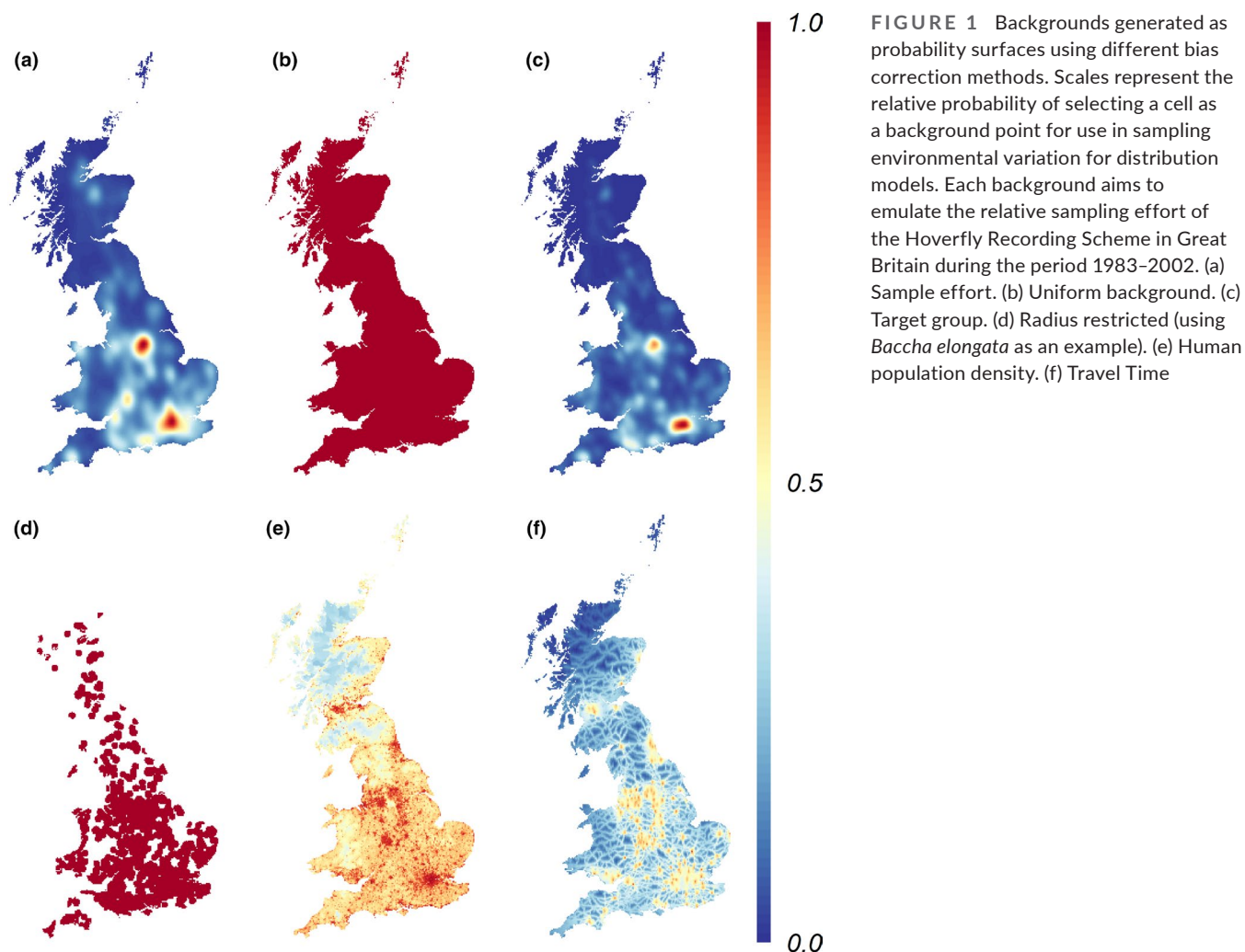


FIGURE 1 Backgrounds generated as probability surfaces using different bias correction methods. Scales represent the relative probability of selecting a cell as a background point for use in sampling environmental variation for distribution models. Each background aims to emulate the relative sampling effort of the Hoverfly Recording Scheme in Great Britain during the period 1983–2002. (a) Sample effort. (b) Uniform background. (c) Target group. (d) Radius restricted (using *Baccha elongata* as an example). (e) Human population density. (f) Travel Time

Travel time from the major cities was used as an alternative proxy for sampling effort. This is based on similar assumptions of sampling bias being shaped by population density, but also including accessibility via road and river networks. A gridded surface of accessibility to cities with more than 50,000 people in the year 2000 was created by the European Commission (Nelson, 2008), defining accessibility as travel time to a location using land or water. As with population density, using travel time in its raw format led to similar overcompensation, and hence, values were transformed by logarithm, multiplied by -1 (to make remote locations less likely to be sampled) and shifted to positive values. The resulting spatial pattern of accessibility is shown in Figure 1f.

2.7 | Virtual species

To validate the results of bias correction on empirical species, we also modelled virtual species where the true distribution is known. We first created 50 random species using the *generateRandomSp* function from the *virtualspecies* R package (Leroy et al., 2015), which used the same predictor variables to create unique habitat suitability

maps (Appendix S1: Table S1.1). The habitat suitability of each random species was converted to a probability of occurrence and binary presence-absence raster using the logistic conversion method of *virtualspecies*. We generated threshold values by allowing beta to vary randomly between 0.3 and 0.7, while keeping alpha constant at -0.1 and setting niche breadth to 'wide'. This generated a random species prevalence for each virtual species, so bias corrections were tested over a wide range of possible scenarios. The exact beta and prevalence values for each virtual species are available with generated distributions in Appendix S3.

To create a set of biased and unbiased occurrences, we sampled each virtual species using two sampling regimes. For our unbiased dataset, we simply sampled each presence-absence raster randomly, taking 1000 presence points. To create a biased dataset, we used the sample effort from the hoverfly recording scheme (Figure 1a) to weight our sampling regime. This allowed us to test models with a realistic sampling bias pattern. To test bias correction, we used the same backgrounds generated with empirical species for target-group, population density and travel time (Figure 1). For distance-restricted backgrounds, we created a unique background for each virtual species, using the same 10-km radius.

2.8 | Statistical analysis

As with colour patterns, comparing maps is a difficult challenge (Fourcade et al., 2014; Warren et al., 2008); to test for bias removal, most analyses compare the ability to predict occurrences, a method that is counter-intuitive when the occurrences contain the sampling bias that we hope to remove. When using discrimination metrics such as area under the receiver-operating curve (AUC), the Kappa statistic and the true skill statistic, biased models can even perform better than their bias-corrected counterparts (Fourcade et al., 2014; Leroy et al., 2018; Veloz, 2009).

For this reason, we chose to rely on similarity metrics between reference distributions and different bias corrections for both empirical and virtual species. For empirical data, these were species-specific distributions created using known sampling effort to factor out sampling bias (Figure 2a). Because each sample effort corrected distribution is only a best estimate at correcting bias, we repeated the same comparisons for virtual data, by testing the similarity between each correction and a model built using unbiased occurrences (Figure 3a). This meant we could confidently assess over- and underprediction rates. Before assessing corrections, we validated the effectiveness of background manipulation to remove bias, using

virtual data. We found that sample effort corrected models performed as well as unbiased models (Appendix S2), and therefore, we were confident that our sample effort distributions were an acceptable standard for comparison.

Distributions were compared by looking at three indices (niche overlap, centroid shift, and range size changes). Schoener's D (Distance) (Schoener, 1968) was chosen as a measure of niche overlap because it has been suggested as the best measure of similarity between niches in reviews (Rödger & Engler, 2011; Warren et al., 2008) and has been an effective metric in studies of sampling bias (Fourcade et al., 2014; Ranc et al., 2017; Stolar & Nielsen, 2015). Centroid shift was computed as the Euclidean distance from the suitability-weighted centroid of the sample effort corrected distributions. Centroids were calculated using the R package *SpatialEco* (Evans, 2020), and Schoener's Distance using the R package *ENMtools* (Warren et al., 2019).

Range size changes were calculated using the R package *Biomod2* (Thuiller et al., 2020), which allowed us to look at both range gains and losses in different parts of the same distribution. We converted all distributions to binary maps using a threshold of 0.5. We chose a constant of 0.5 because trialling thresholds that minimized gains and losses generated extreme thresholds. For instance, if a distribution

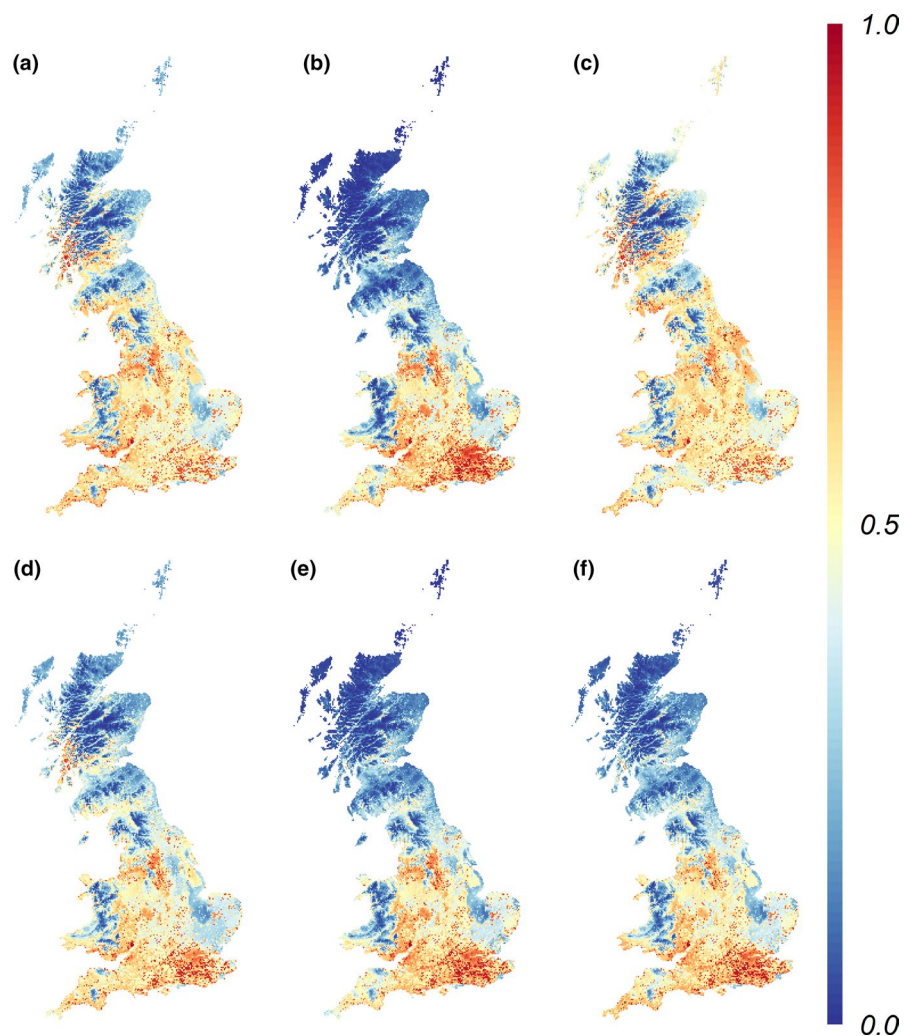


FIGURE 2 Predicted distributions of *Baccha elongata* in Great Britain, generated using occurrences during the period 1983–2002. Maps were generated using 6 different bias correction methods. (a) Sample effort. (b) No correction. (c) Target group. (d) Radius restricted. (e) Human population density. (f) Travel Time. Values range from 0 (no probability of occurrence) to 1 (certain probability of occurrence)

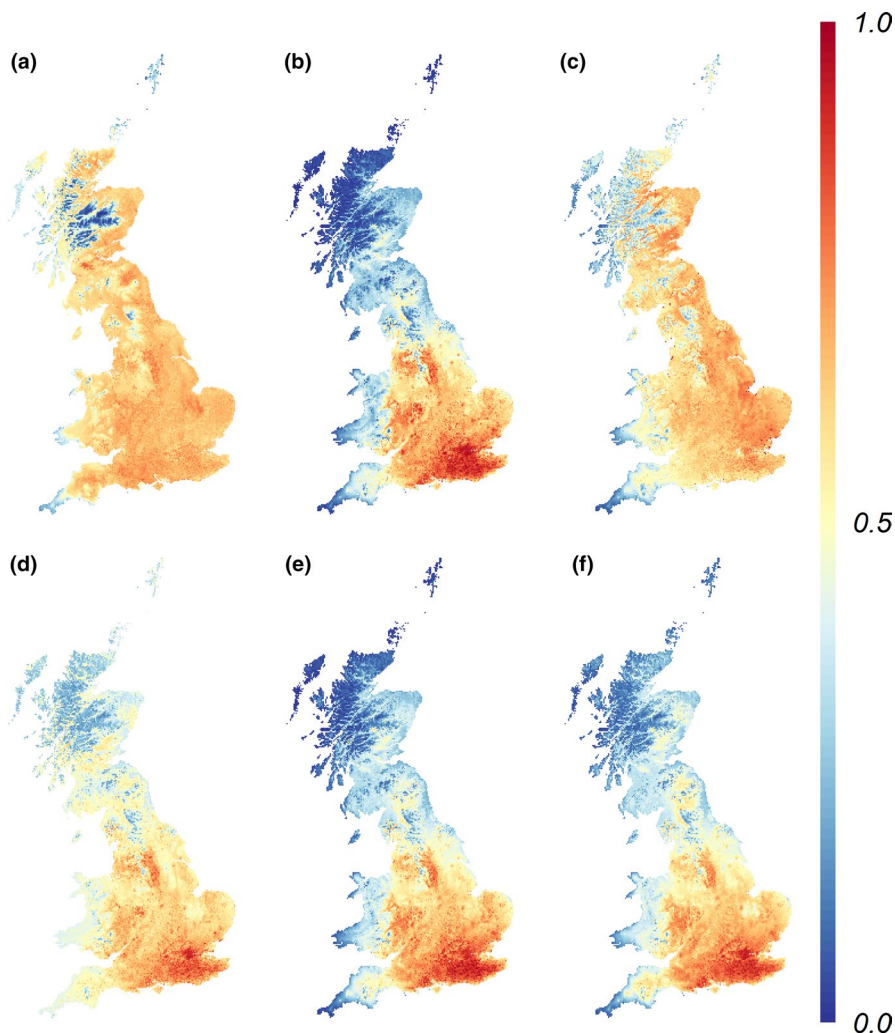


FIGURE 3 Predicted distributions of an example virtual species in Great Britain, generated using both unbiased (a) and biased (b-f) occurrences. Maps were generated using 6 different bias correction methods. (a) Unbiased occurrences. (b) No correction. (c) Target group. (d) Radius restricted. (e) Human population density. (f) Travel Time. Values range from 0 (no probability of occurrence) to 1 (certain probability of occurrence)

severely under-predicted suitability compared to the reference, thresholds would be typically around 0.8–0.9 to minimize losses, even if the two maps were vastly different in reality.

After generating comparison metrics, results were grouped by the method of bias correction, and by species, to account for species-specific effects. A Friedman test was performed to test for significant differences among correction methods, and *post-hoc* Wilcoxon signed-ranks tests were performed on pairwise comparisons applying a Bonferroni adjustment to p-values to account for the multiple tests. We understand that using traditional statistical tests can be inappropriate for virtual species (Meynard et al., 2019); however, we decided to report test statistics alongside empirical species and focus on clear differences in results rather than specific p-values. In addition, we visually inspected the predicted distributions of each species, using expert knowledge to critically assess hoverfly distributions.

3 | Results

While we used all species for our data analysis, for demonstrative purposes we have shown the first species *Baccha elongata*, and the

first random species, which should provide a typical representation of how sampling bias changed predicted distributions. The distributions of each empirical and virtual species is available in Appendices S4-S5.

3.1 | The effect of sampling bias

Without applying any bias correction, the cells with high habitat suitability are densely clustered in central and southern England for both empirical and virtual species (Figure 2b; Figure 3b). This is highest in and around London, but other cities in central England are also hotspots, closely reflecting the distribution of sampling effort (Figure 1b). In contrast, Scotland has very low levels of predicted habitat suitability across nearly the whole region. This is particularly apparent for virtual species, as unbiased models predict high suitability in northern regions (Figure 3a), but predictions without correction are only concentrated in the south (Figure 3b).

When true sampling effort was used to sample background points, there was a clear and strong effect on the predicted habitat suitability (Figure 2a); distributions no longer had a strong bias for population centres and were instead distributed more evenly.

The greatest change was around London and regions directly south, which were predicted to have more isolated areas of highly suitable habitat, rather than one continuous block. Habitat suitability showed moderate increases in suitability in Scotland, but these were generally quite small, with one exception of the western coastline, which showed a much higher predicted suitability in line with previous knowledge of hoverfly habitat preference.

3.2 | Similarity metrics

As calculated using Schoener's D, there was a significant difference between distributions created using each bias correction technique when compared against the reference distribution for both empirical and virtual species (Table 2). For empirical species, pairwise tests showed that all correction methods had higher overlap values than no correction (Figure 4a). For virtual species, all corrections were significantly higher than no correction and showed the same pattern between correction methods (Figure 5a). For both empirical and virtual species, target-group backgrounds generated the highest Schoener's D values.

There was a significant difference in the magnitude of the centroid shift of each bias correction technique, for both empirical and virtual species (Table 2). Pairwise tests showed there were significant differences between no correction and other all bias correction methods (Figure 4b; Figure 5b). Target-group backgrounds generated the smallest centroid shifts, and no correction generated the largest shifts, with the same pattern for both empirical and virtual species.

There was a significant difference between range gains and losses for both empirical and virtual species (Table 2). All correction methods had similar range losses and gains, except for models made with target-group backgrounds (Figure 4c,d; Figure 5c,d). Target-group models had the highest range gains, which is a sign of over-compensation. In comparison, target-group models had the lowest rates of range loss, whereas other correction methods had similar levels of underprediction to no correction.

4 | DISCUSSION

4.1 | The effect of sampling bias

There was a clear and substantial effect of sampling bias in generating models (Figure 2; Figure 3). Without correcting for sampling bias, highly suitable areas were concentrated in central England, where the most sampling took place, while much of the rest of Great Britain was considered poor habitat. The same pattern of high suitability was prevalent in virtual species, despite the fact that virtual species had randomly generated habitat suitability that should differ from empirical data, demonstrating that the same sampling bias is present in both datasets.

When true sampling effort was used as part of the modelling process, distributions were typically more evenly spread, with higher habitat suitability in the south and coastal regions. This reflects a realistic distribution of hoverflies in the UK, which are more abundant at lower latitudes and altitudes. Strong biases towards central England were down weighted, leading to a smaller potential area of occupation in this region. Some areas of Scotland showed high habitat suitability, particularly the western coastline, which matches previous knowledge that hoverflies favour habitats with higher moisture.

Interestingly, for virtual species, the true probability of occurrence is still visible in the biased predictions (Figure 3b; Appendix S5), denoted by slightly lighter areas in the northernmost regions where there should be high predicted occurrence. However, the high suitability in the south caused by sampling bias completely overshadows this pattern, creating distributions that appear vastly different. In general, we were surprised by the stark difference between biased and unbiased models. It may be that the hoverfly recording scheme is particularly biased in sampling; however, conservation practitioners need to be careful that their distributions do not typically map sampling effort as ours did.

When assessing maps using similarity metrics, no-correction maps consistently performed poorly for both empirical and virtual species. Moreover, all metrics displayed a very large range of values,

TABLE 2 Results of Friedman tests between correction methods for different similarity metrics, controlling for the effect of repeated measures for each species. All tests were highly significant

Data	Similarity metric	n	χ^2	df	p-value	Effect size	Magnitude
Hoverfly	Schoener's Distance	58	216.58	4	1.02E-45	0.93	Large
Hoverfly	Centroid Shift (km)	58	207.35	4	9.86E-44	0.89	Large
Hoverfly	Range Gain (%)	58	123.05	4	1.19E-25	0.53	Large
Hoverfly	Range Loss (%)	58	164.43	4	1.64E-34	0.71	Large
Virtual	Schoener's Distance	50	174.70	4	1.02E-36	0.87	Large
Virtual	Centroid Shift (km)	50	179.31	4	1.05E-37	0.90	Large
Virtual	Range Gain (%)	50	97.49	4	3.36E-20	0.49	Moderate
Virtual	Range Loss (%)	50	167.73	4	3.21E-35	0.84	Large

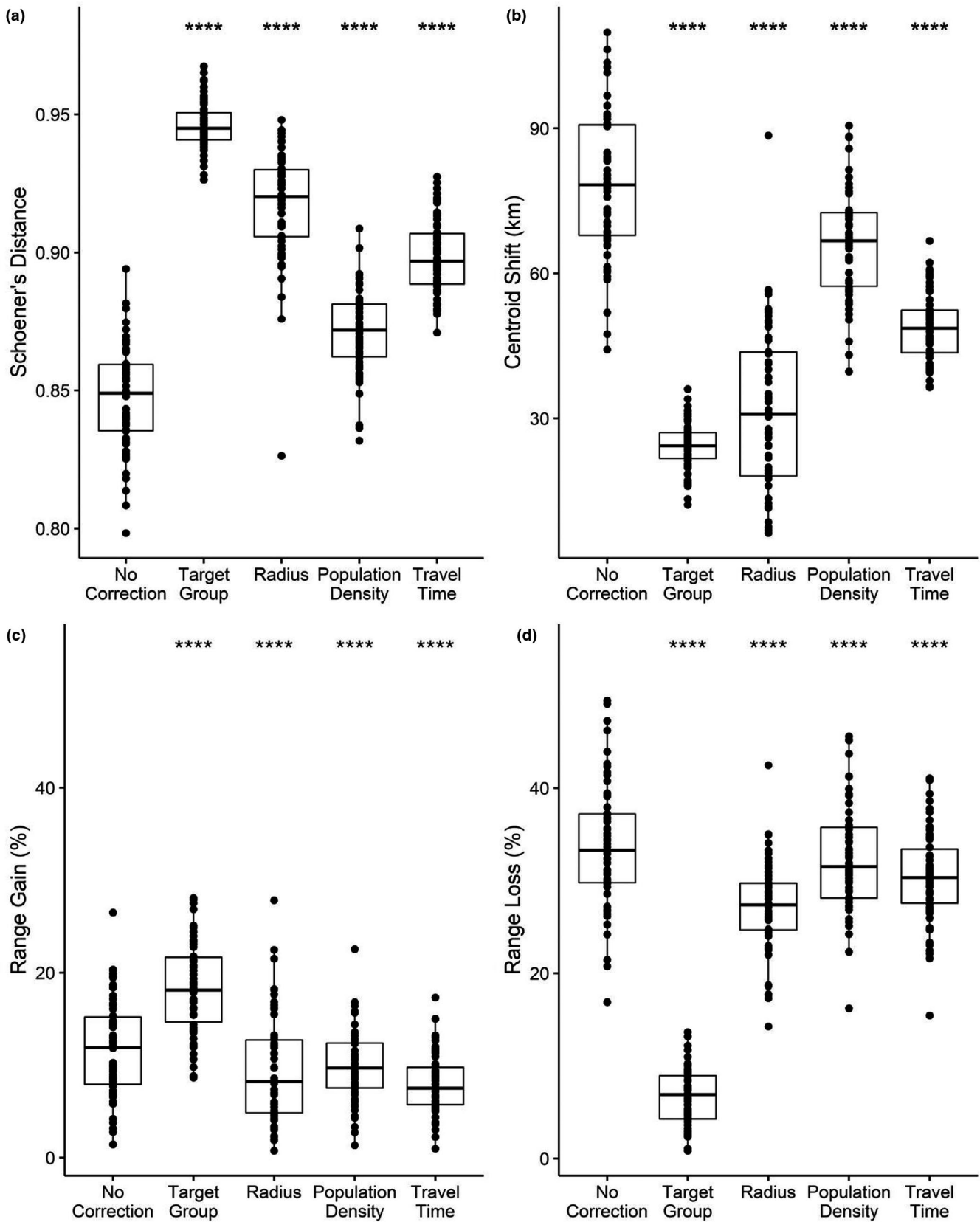


FIGURE 4 Similarity metrics for each bias correction compared to sample effort models for 58 species of hoverfly (*Syrphidae*). (a) Schoener's Distance, a similarity measure of overall spatial correlation. (b) The distance between centroids of each distribution. (c,d) Range size changes as a percentage, after converting maps to a binary raster using a threshold of 0.5. Changes do not represent total change, but changes by pixel, so single distributions may simultaneously display high gains and losses. Significant results of pairwise Wilcoxon tests with a Bonferroni correction are denoted as: * <math><0.05</math>; ** <math><0.01</math>; *** <math><0.001</math>. Only pairwise comparisons against no correction are shown

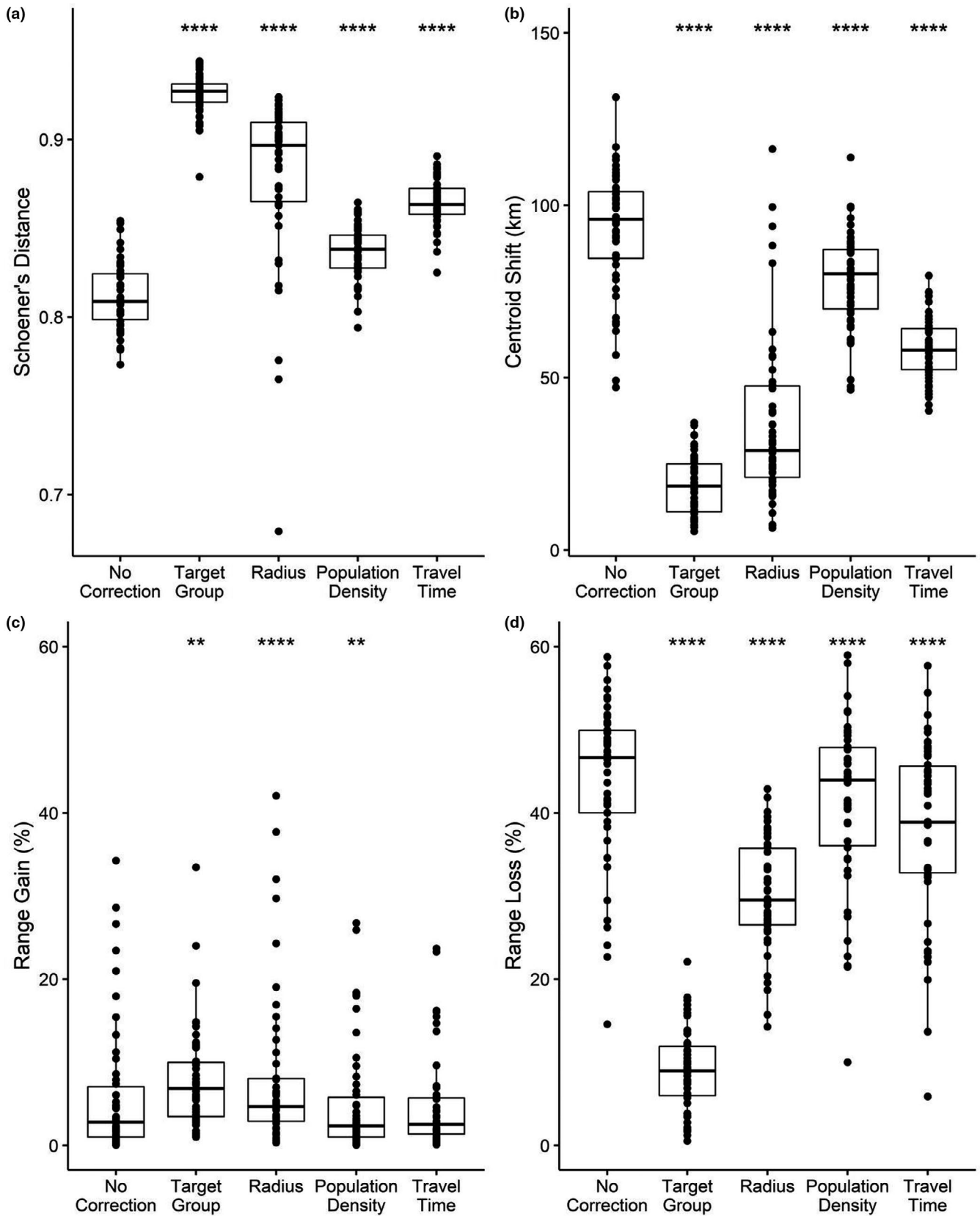


FIGURE 5 Similarity metrics for each bias correction compared to unbiased models for 50 virtual species. (a) Schoener's Distance, a similarity measure of overall spatial correlation. (b) The distance between centroids of each distribution. (c,d) Range size changes as a percentage, after converting maps to a binary raster using a threshold of 0.5. Changes do not represent total change, but changes by pixel, so single distributions may simultaneously display high gains and losses. Significant results of pairwise Wilcoxon tests with a Bonferroni correction are denoted as: * < 0.05; ** < 0.01; *** < 0.001. Only pairwise comparisons against no correction are shown

demonstrating that individual species respond differently to sampling bias. This suggests that in most scenarios any attempt at bias correction will bring the result closer to the true estimate of sampling effort, and hence a more accurate representation of the true distribution (Phillips et al., 2009; Ranc et al., 2017; Stolar & Nielson, 2015).

4.2 | Target-group correction

Target-group backgrounds aim to use the occurrences of similar taxa to estimate sampling effort and were a highly successful method of counteracting sampling bias. They consistently performed well in terms of all similarity metrics, whether using empirical or virtual data. Furthermore, when visually inspecting maps, target-group corrections were the only method that produced distributions that appeared similar to sample effort corrections in empirical data, and unbiased models for virtual species. Given that visual inspection is often the primary assessment by conservation practitioners, this was perhaps the most important result.

Concerns that differences in detectability and species richness may lead to a dissociation between sampling effort and target-group occurrence records were unfounded. This demonstrates that when sampling effort is unknown, using presence records of similar taxa can produce results almost as good as the use of true sampling effort, and this supports previous reviews of bias correction, which favour target group as an effective method of bias correction (Phillips et al., 2009; Ranc et al., 2017; Syfert et al., 2013).

The only negative of using target-group as a bias correction is that in many species, there were signs of overcompensation in the most under-sampled areas of the distribution, echoing past studies that observe increases in false positives (Syfert et al., 2013). This was apparent in both empirical and virtual species, as target-group models had the largest range size gains of any method (Figure 4c; Figure 5c). This is most apparent in the northernmost parts of the distribution (Figure 2c; Figure 3c) and demonstrate that target-group backgrounds are not a perfect bias correction. Still, these changes are minor compared to the underprediction of other correction methods, which performed similarly to no correction (Figure 4d; Figure 5d).

4.3 | Background restriction

Restricting background points to within a 10-km radius of occurrences produced distributions that were moderately successful at counteracting sampling bias (Figure 2d; Figure 3d; Figure 4). When considering niche overlap, distributions performed significantly better than unbiased models; however, maps often appeared similar to no-correction models during visual inspection. This is supported when looking at range loss metrics, which were similar to no-correction models, especially for empirical species.

While there have been many reviews on the effectiveness of restricting background areas in terms of increasing model accuracy

(Acevedo et al., 2012; Anderson & Raza, 2010; Barve et al., 2011), there has been limited focus when correcting sampling bias is the main objective. To the best of our knowledge, there has only been one review by Fourcade et al. (2014). While our models were somewhat successful at correcting bias, they found that restricting the background area in geographical space led to some of the worst performing models. This disparity is also apparent in our results, with 10-km buffers generating the greatest variation in bias correction performance. The main reason for this is that restricting background points seemed to perform gradually worse with denser occurrences. This is because limiting the background more severely runs the risk of reducing model accuracy as species appear to utilize all available environment (Thuiller et al., 2004; VanDerWal et al., 2009) and so should only be done when the occurrence records are extensive enough.

It is also worth noting that the variation in performance between species was partially a result of the subjective choice of radius size. This may be the greatest barrier to using distance-restricted backgrounds, as without explicit knowledge of sampling effort it is not possible to easily assess different radius sizes. A similar method for restricting the background was recently proposed by Vollering et al. (2019), using the spatial autocorrelation of environmental variables to determine radius size. A key difference was that they account for denser occurrences by summing overlapping radii, which they refer to as 'background thickening'. When assessing the method, they found thickening performed better than target-group methods, but importantly they used individual target-group occurrences rather than a continuous surface of species richness, so a direct comparison of the two methods is still needed to draw firm conclusions.

4.4 | Population density and travel time

Using a potential covariate of sampling effort is an appealing concept to simplify distribution modelling, especially when records of similar taxa are lacking, or there are too few points for restricting background points. A recent paper by Monsarrat et al. (2019) showed that accessibility maps were able to predict sampling-effort biases in historical data, and as our results showed an improvement over no correction, this also suggests that accessibility is a factor in sampling bias.

However, while human population density and travel time produced models that were significantly better than no correction, the visual differences were often barely noticeable (Figures 2e-f, 3e-f). The inclusion of access in the form of travel time slightly improved models, but not to significant degree, suggesting that a large proportion of accessibility is already accounted for in the distribution of samplers. There was also a large degree of variation between species, similar to previous reviews of accessibility (Dubos et al., 2021). Therefore, we are hesitant to recommend using covariates of sampling effort as a bias correction method when alternatives are available.

5 | CONCLUSION

Choosing a bias correction method is a difficult choice, but one fundamental to the accuracy of species distribution modelling (Merow et al., 2013; Newbold, 2010). While there are many options available for modelling species, if the data have an inherent bias any distributions generated will always be an inaccurate representation. This becomes an issue of great importance when species distribution modelling is frequently used to plan and manage conservation projects around the world (Loiselle et al., 2003).

We found that using target-group sampling to generate backgrounds was the most effective method of estimating bias when the sampling effort is unknown. In this scenario, every care should be taken to make sure that the target-group taxa demonstrate the same sampling bias, and that sampling methods and detectability are consistent among species. This assumption was easily met when using target-group data from the same recording scheme and so may be particularly useful to other large citizen science projects. Furthermore, as target-group data are also determined by species richness, it is possible that sampling bias is being counteracted by a bias for species richness, which may not be appropriate for all taxa or situations (El-Gabbas & Dormann, 2018; Warton et al., 2013). Users should therefore always provide strong evidence of any assumptions made when justifying which taxa to include for background samples (Merow et al., 2013).

Throughout this study, the issue of overcompensation and subjectivity became increasingly apparent. This was most obvious when deciding radius size for distance-restricted backgrounds, but also became apparent when creating backgrounds using travel time; initial backgrounds caused such overcompensation that only the most inaccessible areas were considered suitable habitats. This effect was also apparent to a lesser extent when looking at distributions generated by the target-group approach. It is almost certain that some areas received inaccurately inflated habitat suitability scores as a result of reducing biased regions. Modifying population density and travel time backgrounds to reduce their overcompensation produced more realistic maps but introduced a subjective process that may lead to users to create distributions that they feel are correct, rather than accurately predicting the underlying pattern of suitability. Recent work has shown that overprediction can be addressed by incorporating dispersal constraints into models (Mendes et al., 2020) and should be used in concert with bias corrections to provide the most accurate distributions.

Advances in technology such as remote sensing and machine learning mean we are supplied with increasing amounts of high-quality data, ideal for modelling species distributions. At the same time, climate change and other human stressors are dramatically increasing our need for sophisticated methods of conservation. As a result, it is becoming more and more important to base our decisions on accurate estimations of species distributions, and paramount that we address the issue of sampling bias. Our results show that sampling bias can have a serious effect on predicted distributions. As we, and others, have demonstrated, sampling bias correction can greatly increase the

accuracy of species distribution models, and should be used wherever possible to generate effective tools to aid in conservation.

ACKNOWLEDGEMENTS

We would like to thank the volunteers of the Hoverfly Recording Scheme for supplying us with such useful data for modelling. We would also like to thank Alicja Witwicka for her very helpful comments.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

PEER REVIEW

The peer review history for this article is available at <https://publons.com/publon/10.1111/ddi.13442>.

DATA AVAILABILITY STATEMENT

All occurrence data and code for preparing backgrounds, SDMs, analysis and plotting are available at: https://github.com/Syrph/Sampling_bias. Bioclim data can be sourced from the WorldClim database: <https://www.worldclim.org/>. Land use data are available to download from CEH at: <https://www.ceh.ac.uk/services/land-cover-map-2000>. Human population density is available from SEDAC at: <https://sedac.ciesin.columbia.edu/data/collection/gpw-v4>. Travel time data are available from the European Commission at: <https://forobs.jrc.ec.europa.eu/products/gam/index.php>.

ORCID

Robert A. Barber  <https://orcid.org/0000-0001-6735-5250>

REFERENCES

- Acevedo, P., Jiménez-Valverde, A., Lobo, J. M., & Real, R. (2012). Delimiting the geographical background in species distribution modelling. *Journal of Biogeography*, *39*, 1383–1390. <https://doi.org/10.1111/j.1365-2699.2012.02713.x>
- Aiello-Lammens, M. E., Boria, R. A., Radosavljevic, A., Vilela, B., & Anderson, R. P. (2015). spThin: An R package for spatial thinning of species occurrence records for use in ecological niche models. *Ecography*, *38*, 541–545. <https://doi.org/10.1111/ecog.01132>
- Anderson, R. P., & Gonzalez, I. (2011). Species-specific tuning increases robustness to sampling bias in models of species distributions: An implementation with Maxent. *Ecological Modelling*, *222*, 2796–2811. <https://doi.org/10.1016/j.ecolmodel.2011.04.011>
- Anderson, R. P., & Raza, A. (2010). The effect of the extent of the study region on GIS models of species geographic distributions and estimates of niche evolution: Preliminary tests with montane rodents (genus *Nephelomys*) in Venezuela. *Journal of Biogeography*, *37*, 1378–1393.
- Ball, S., & Morris, R. (2012). The Hoverfly Recording Scheme – putting Diptera on the map. *Antenna*, *36*, 177–185.
- Barbet-Massin, M., Jiguet, F., Albert, C. H., & Thuiller, W. (2012). Selecting pseudo-absences for species distribution models: how, where and how many? *Methods in Ecology and Evolution*, *3*, 327–338. <https://doi.org/10.1111/j.2041-210X.2011.00172.x>
- Barve, N., Barve, V., Jiménez-Valverde, A., Lira-Noriega, A., Maher, S. P., Peterson, A. T., Soberón, J., & Villalobos, F. (2011). The crucial role of the accessible area in ecological niche modeling and

- species distribution modeling. *Ecological Modelling*, 222, 1810–1819. <https://doi.org/10.1016/j.ecolmodel.2011.02.011>
- Boria, R. A., Olson, L. E., Goodman, S. M., & Anderson, R. P. (2014). Spatial filtering to reduce sampling bias can improve the performance of ecological niche models. *Ecological Modelling*, 275, 73–77. <https://doi.org/10.1016/j.ecolmodel.2013.12.012>
- Braunisch, V., Coppes, J., Arlettaz, R., Suchant, R., Schmid, H., & Bollmann, K. (2013). Selecting from correlated climate variables: A major source of uncertainty for predicting species distributions under climate change. *Ecography*, 36, 971–983. <https://doi.org/10.1111/j.1600-0587.2013.00138.x>
- Center for International Earth Science Information Network - CIESIN - Columbia University. (2016). *Gridded Population of the World, Version 4 (GPWv4): Population density*. NASA Socioeconomic Data and Applications Center (SEDAC).
- Dormann, C. F., McPherson, J. M., B. Araújo, M., Bivand, R., Bolliger, J., Carl, G., G. Davies, R., Hirzel, A., Jetz, W., Daniel Kissling, W., Kühn, I., Ohlemüller, R., R. Peres-Neto, P., Reineking, B., Schröder, B., Schurr, F. M., & Wilson, R. (2007). Methods to account for spatial autocorrelation in the analysis of species distributional data: A review. *Ecography*, 30, 609–628. <https://doi.org/10.1111/j.2007.0906-7590.05171.x>
- Dubos, N., Préau, C., Lenormand, M., Papuga, G., Montsarrat, S., Denelle, P., Le Louarn, M., Heremans, S., Roel, M., Roche, P., & Luque, S. (2021). Assessing the effect of sample bias correction in species distribution models. arXiv preprint arXiv:2103.07107.
- Duong, T. (2020). ks: Kernel Smoothing. R package version 1.11.7. <https://CRAN.R-project.org/package=ks>
- El-Gabbas, A., & Dormann, C. F. (2018). Improved species-occurrence predictions in data-poor regions: Using large-scale data and bias correction with down-weighted Poisson regression and Maxent. *Ecography*, 41, 1161–1172. <https://doi.org/10.1111/ecog.03149>
- Elith, J., Kearney, M., & Phillips, S. (2010). The art of modelling range-shifting species. *Methods in Ecology and Evolution*, 1, 330–342. <https://doi.org/10.1111/j.2041-210X.2010.00036.x>
- Elith, J., & Leathwick, J. R. (2009). Species distribution models: Ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution and Systematics*, 40, 677–697. <https://doi.org/10.1146/annurev.ecolsys.110308.120159>
- Elith, J., Phillips, S. J., Hastie, T., Dudik, M., Chee, Y. E., & Yates, C. J. (2011). A statistical explanation of Maxent for ecologists. *Diversity and Distributions*, 17, 43–57. <https://doi.org/10.1111/j.1472-4642.2010.00725.x>
- Evans, J. S. (2020) spatialEco: R package version 1.3-1. Retrieved from <https://cran.r-project.org/package=spatialEco>
- Fithian, W., Elith, J., Hastie, T., & Keith, D. A. (2015). Bias correction in species distribution models: Pooling survey and collection data for multiple species. *Methods in Ecology and Evolution*, 6, 424–438. <https://doi.org/10.1111/2041-210X.12242>
- Flick, S. E., & Hijmans, R. J. (2017). WorldClim 2: New 1km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, 37, 4302–4315.
- Fourcade, Y., Besnard, A. G., & Secondi, J. (2018). Paintings predict the distribution of species, or the challenge of selecting environmental predictors and evaluation statistics. *Global Ecology and Biogeography*, 27, 245–256. <https://doi.org/10.1111/geb.12684>
- Fourcade, Y., Engler, J. O., Rödder, D., & Secondi, J. (2014). Mapping species distributions with MAXENT using a geographically biased sample of presence data: A performance assessment of methods for correcting sampling bias. *PLoS One*, 9, e97122. <https://doi.org/10.1371/journal.pone.0097122>
- Fuller, R. M., Smith, G. M., Sanderson, J. M., Hill, R. A., Thomson, A. G., Cox, R., Brown, N. J., Clarke, R. T., Rothery, P., & Gerard, F. F. (2002). *Land Cover Map 2000 (1km dominant target class, GB)*. NERC Environmental Information Data Centre. <https://doi.org/10.5285/abff8409-0995-48d2-9303-468e1a9fe3df>
- Guerra, C., Metzger, M. J., Honrado, J., & Alonso, J. (2013). A spatially explicit methodology for a *a priori* estimation of field survey effort in environmental observation networks. *International Journal of Geographical Information Science*, 27, 2077–2098.
- Hijmans, R. J. (2020). raster: Geographic Data Analysis and Modeling. R package version 3.1-5. Retrieved from <https://CRAN.R-project.org/package=raster>
- Hijmans, R. J., Phillips, S. J., Leathwick, J., & Elith, J. (2017). dismo: Species Distribution Modeling. R package version (p. 1.1-4.). Retrieved from <https://CRAN.R-project.org/package=dismo>.
- Jeschke, J., & Strayer, D. (2008). Usefulness of bioclimatic models for studying climate change and invasive species. *Annals of the New York Academy of Sciences*, 1134, 1–24. <https://doi.org/10.1196/annals.1439.002>
- Kadmon, R., Farber, O., & Danin, A. (2003). A systematic analysis of factors affecting the performance of climatic envelope models. *Ecological Applications*, 13, 853–867. [https://doi.org/10.1890/1051-0761\(2003\)013\(0853:ASA0FA\)2.0.CO;2](https://doi.org/10.1890/1051-0761(2003)013(0853:ASA0FA)2.0.CO;2)
- Kadmon, R., Farber, O., & Danin, A. (2004). Effect of roadside bias on the accuracy of predictive maps produced by bioclimatic models. *Ecological Applications*, 14, 401–413. <https://doi.org/10.1890/02-5364>
- Kramer-Schadt, S., Niedballa, J., Pilgrim, J. D., Schröder, B., Lindenborn, J., Reinfelder, V., Stillfried, M., Heckmann, I., Scharf, A. K., Augeri, D. M., Cheyne, S. M., Hearn, A. J., Ross, J., Macdonald, D. W., Mathai, J., Eaton, J., Marshall, A. J., Semiadi, G., Rustam, R., ... Wilting, A. (2013). The importance of correcting for sampling bias in Maxent species distribution models. *Diversity and Distributions*, 19, 1366–1379. <https://doi.org/10.1111/ddi.12096>
- Leroy, B., Delsol, R., Huguency, B., Meynard, C. N., 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, 1994–2002. <https://doi.org/10.1111/jbi.13402>
- Leroy, B., Meynard, C. N., Bellard, C., & Courchamp, F. (2015). Virtualspecies, an R package to generate virtual species distributions. *Ecography*, 39, 599–607. <https://doi.org/10.1111/ecog.01388>
- Lobo, J. M., Jiménez-Valverde, A., & Hortal, J. (2010). The uncertain nature of absences and their importance in species distribution modelling. *Ecography*, 33, 103–114. <https://doi.org/10.1111/j.1600-0587.2009.06039.x>
- Loiselle, B. A., Howell, C. A., Graham, C. H., Goerck, J. M., Brooks, T., Smith, K. G., & Williams, P. H. (2003). Avoiding pitfalls of using species distribution models in conservation planning. *Conservation Biology*, 17, 1591–1600. <https://doi.org/10.1111/j.1523-1739.2003.00233.x>
- Mendes, P., Velazco, S. J. E., Alves de Andrade, A. F., & De Marco, P. (2020). Dealing with overprediction in species distribution models: How adding distance constraints can improve model accuracy. *Ecological Modelling*, 431, 109180. <https://doi.org/10.1016/j.ecolmodel.2020.109180>
- Merow, C., Smith, M. J., & Silander, J. A. (2013). A practical guide to Maxent for modelling species' distributions: What it does, and why inputs and settings matter. *Ecography*, 36, 1058–1069.
- Meynard, C. N., Leroy, B., & Kaplan, D. M. (2019). Testing methods in species distribution modelling using virtual species: What have we learnt and what are we missing? *Ecography*, 42, 2021–2036. <https://doi.org/10.1111/ecog.04385>
- Monsarrat, S., Boshoff, A. F., & Kerley, G. I. H. (2019). Accessibility maps as a tool to predict sampling bias in historical biodiversity occurrence records. *Ecography*, 42, 125–136. <https://doi.org/10.1111/ecog.03944>
- Moua, Y., Roux, E., Seyler, F., & Briolant, S. (2020). Correcting the effect of sampling bias in species distribution modelling – A new method in the case of a low number of presence data. *Ecological Informatics*, 57, 101086.
- Nelson, A. (2008). *Travel time to major cities: A global map of Accessibility*. Global Environment Monitoring Unit - Joint Research Centre of the European Commission.

- Newbold, T. (2010). Applications and limitations of museum data for conservation and ecology, with particular attention to species distribution models. *Progress in Physical Geography*, 34, 3–22. <https://doi.org/10.1177/0309133309355630>
- Phillips, S. J. (2004). A maximum entropy approach to species distribution modelling. Proceedings of the Twenty-First International Conference on Machine Learning, 655–662.
- Phillips, S. J. (2008). Transferability sample selection bias and background data in presence-only modelling: A response to Peterson et al (2007). *Ecography*, 31(2), 272–278. <https://doi.org/10.1111/j.0906-7590.2008.5378.x>
- Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modelling of species geographic distributions. *Ecological Modelling*, 190, 231–259.
- Phillips, S. J., Dudik, M., Elith, J., Graham, C. H., Lehmann, A., Leathwick, J., & Ferrier, S. (2009). Sample selection bias and presence-only distribution model: Implications for background and pseudo-absence data. *Ecological Society of America*, 19, 181–197.
- Pimentel, D., McNair, S., Janecka, J., Wightman, J., Simmonds, C., O'Connell, C., Wong, E., Russel, L., Zern, J., Aquino, T., & Tsomondo, T. (2001). Economic and environmental threats of alien plant, animal, and microbe invasions. *Agriculture, Ecosystems & Environment*, 84, 1–20. [https://doi.org/10.1016/S0167-8809\(00\)00178-X](https://doi.org/10.1016/S0167-8809(00)00178-X)
- Ponder, W. F., Carter, G. A., Felmons, P., & Chapman, R. R. (2001). Evaluation of museum collection data for use in biodiversity assessment. *Conservation Biology*, 15, 648–657. <https://doi.org/10.1046/j.1523-1739.2001.015003648.x>
- R Core Team. (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing.
- Radosavljevic, A., & Anderson, R. P. (2014). Making better Maxent models of species distributions: Complexity, overfitting and evaluation. *Journal of Biogeography*, 41, 629–643.
- Ranc, N., Santini, L., Rondinini, C., Boitani, L., Poitevin, F., Angerbjörn, A., & Maiorano, L. (2017). Performance tradeoffs in target-group bias correction for species distribution models. *Ecography*, 40, 1076–1087. <https://doi.org/10.1111/ecog.02414>
- Reddy, S., & Dávalos, L. M. (2003). Geographic sampling bias and its implications for conservation priorities in Africa. *Journal of Biogeography*, 30, 1719–1727.
- Rödder, D., & Engler, J. O. (2011). Quantitative metrics of overlaps in Grinnellian niches: Advances and possible drawbacks. *Global Ecology and Biogeography*, 20, 915–927. <https://doi.org/10.1111/j.1466-8238.2011.00659.x>
- Schoener, T. (1968). The Anolis lizards of Bimini: Resource partitioning in a complex fauna. *Ecology*, 49, 704–726. <https://doi.org/10.2307/1935534>
- Soberon, J., & Nakamura, M. (2009). Niches and distributional areas: Concepts, methods, and assumptions. *Proceedings of the National Academy of Sciences of the United States of America*, 106, 19644–19650. <https://doi.org/10.1073/pnas.0901637106>
- Soberón, J., & Peterson, A. T. (2005). Interpretation of models of fundamental ecological niches and species' distributional areas. *Biodiversity Informatics*, 2, 1–10. <https://doi.org/10.17161/bi.v2i0.4>
- Stolar, J., & Nielsen, S. E. (2015). Accounting for spatially biased sampling effort in presence-only species distribution modelling. *Diversity and Distributions*, 21, 595–608. <https://doi.org/10.1111/ddi.12279>
- Syfert, M. M., Smith, M. J., & Coomes, D. A. (2013). The effects of sampling bias and model complexity on the predictive performance of Maxent species distribution models. *PLoS One*, 8, e55158.
- Thuiller, W., Brotons, L., Araújo, M. B., & Lavorel, S. (2004). Effects of restricting environmental range of data to project current and future species distributions. *Ecography*, 27, 165–172. <https://doi.org/10.1111/j.0906-7590.2004.03673.x>
- Thuiller, W., Georges, D., Engler, R., & Breiner, F. (2020). biomod2: Ensemble Platform for Species Distribution Modeling. R package version 3.4.6. Retrieved from <https://CRAN.R-project.org/package=biomod2>
- VanDerWal, J. J., Shoo, L., Graham, C. H., & Williams, S. (2009). Selecting pseudo-absence data for presence-only distribution modelling: How far should you stray from what you know? *Ecological Modelling*, 220, 589–594.
- Velo, S. D. (2009). Spatially autocorrelated sampling falsely inflates measures of accuracy for presence-only niche models. *Journal of Biogeography*, 36, 2290–2299. <https://doi.org/10.1111/j.1365-2699.2009.02174.x>
- Vollering, J., Halvorsen, R., Auestad, I., & Rydgren, K. (2019). Bunching up the background better bias in species distribution models. *Ecography*, 42, 1717–1727. <https://doi.org/10.1111/ecog.04503>
- Warren, D. L., Glor, R. E., & Turelli, M. (2008). Environmental niche equivalency versus conservatism: Quantitative approaches to niche evolution. *Evolution*, 62, 2868–2883.
- Warren, D. L., Matzke, N., Cardillo, M., Baumgartner, J., Beaumont, L., Huron, N., Dinnage, R. (2019). ENMTools (Software Package). Retrieved from <https://github.com/danwarren/ENMTools>
- Warton, D. I., Renner, I. W., & Ramp, D. (2013). Model-based control of observer bias for the analysis of presence-only data in ecology. *PLoS One*, 8, e79168. <https://doi.org/10.1371/journal.pone.0079168>
- Wisn, M. S., Hijmans, R. J., Li, J., Peterson, A. T., Graham, C. H., & Guisan, A. (2008). Effects of sample size on the performance of species distribution models. *Diversity and Distributions*, 14, 763–773. <https://doi.org/10.1111/j.1472-4642.2008.00482.x>
- Yackulic, C. B., Chandler, R., Zipkin, E. F., Royle, J. A., Nichols, J. D., Campbell Grant, E. H., & Veran, S. (2013). Presence-only modelling using MAXENT: When can we trust the inferences? *Methods in Ecology and Evolution*, 4, 236–243. <https://doi.org/10.1111/2041-210x.12004>

BIOSKETCH

Robert Barber is a British ecologist in the 3rd year of his PhD degree at Imperial College London. He is interested in both phylogenetics and biogeography. He currently conducts macro-evolutionary research on the diversity of bird traits. Before his PhD, he worked at the Wellcome Trust Sanger Institute, conducting genomic sequencing for the Human Biobank project and Tree of Life project, and received his Master's degree from the University of Nottingham, where the early stages of this research were conducted.

Author contributions: RB and FG conceived and designed methods. SB and RM collated data. RB conducted modelling, analysis and wrote the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Barber, R. A., Ball, S. G., Morris, R. K. A., & Gilbert, F. (2022). Target-group backgrounds prove effective at correcting sampling bias in Maxent models. *Diversity and Distributions*, 28, 128–141. <https://doi.org/10.1111/ddi.13442>