This is a repository copy of *Climate change vulnerability for species—Assessing the assessments*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/126156/

Version: Published Version

Article:
Wheatley, Christopher J., Beale, Colin M. orcid.org/0000-0002-2960-5666, Bradbury, Richard B. et al. (3 more authors) (2017) Climate change vulnerability for species—Assessing the assessments. Global Change Biology. pp. 3704-3715. ISSN: 1354-1013

https://doi.org/10.1111/gcb.13759

Reuse
Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Received: 16 November 2016 | Revised: 23 February 2017 | Accepted: 28 April 2017

DOI: 10.1111/gcb.13759

PRIMARY RESEARCH ARTICLE

WILEY Global Change Biology

Climate change vulnerability for species—Assessing the assessments

Christopher J. Wheatley[1] | Colin M. Beale[1] | Richard B. Bradbury[2,3,4] | James W. Pearce-Higgins[4,5] | Rob Critchlow[1] | Chris D. Thomas[1]

[1]Department of Biology, University of York, York, UK

[2]RSPB Centre for Conservation Science, RSPB, Sandy, Bedfordshire, UK

[3]RSPB Centre for Conservation Science, Cambridge, UK

[4]Conservation Science Group, Department of Zoology, University of Cambridge, Cambridge, UK

[5]BTO, The Nunnery, Thetford, Norfolk, UK

Correspondence
Christopher J. Wheatley, Department of Biology, University of York, York, UK.
Email: cjw509@york.ac.uk

Funding information
Natural Environment Research Council, Grant/Award Number: NE/L008866/1

Abstract

Climate change vulnerability assessments are commonly used to identify species at risk from global climate change, but the wide range of methodologies available makes it difficult for end users, such as conservation practitioners or policymakers, to decide which method to use as a basis for decision-making. In this study, we evaluate whether different assessments consistently assign species to the same risk categories and whether any of the existing methodologies perform well at identifying climate-threatened species. We compare the outputs of 12 climate change vulnerability assessment methodologies, using both real and simulated species, and validate the methods using historic data for British birds and butterflies (i.e. using historical data to assign risks and more recent data for validation). Our results show that the different vulnerability assessment methods are not consistent with one another; different risk categories are assigned for both the real and simulated sets of species. Validation of the different vulnerability assessments suggests that methods incorporating historic trend data into the assessment perform best at predicting distribution trends in subsequent time periods. This study demonstrates that climate change vulnerability assessments should not be used interchangeably due to the poor overall agreement between methods when considering the same species. The results of our validation provide more support for the use of trend-based rather than purely trait-based approaches, although further validation will be required as data become available.

KEYWORDS

biodiversity, climate change, conservation prioritization, policy, risk assessment, species conservation, vulnerability assessment

1 | INTRODUCTION

Standardized methods of risk assessment are important tools for prioritizing adaptive strategies to counter the impacts of climate change, including conservation action for species most likely to face extinction. The IUCN Red List (De Grammont & Cuarón, 2006; Mace et al., 2008) is globally accepted as the method for assessing the vulnerability of species to extinction. However, it has recently been suggested that this process does not adequately identify potential future risk, such as that posed by climate change, as it focuses more on the symptoms of declines than on the underlying causes (Akçakaya, Butchart, Watson, & Pearson, 2014). Given that global extinction risks are high (Ceballos et al., 2015; Pimm et al., 2014;

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.
© 2017 The Authors. *Global Change Biology* Published by John Wiley & Sons Ltd.

Thomas et al., 2004a) and increasing as a consequence of climate change (Thomas et al., 2004b; Warren, Price, Fischlin, de la Nava Santos, & Midgley, 2011), this could potentially lead to an underestimate of the risk to species. These concerns have led to the parallel development of a number of risk assessment frameworks (Pacifici et al., 2015), each of which aims to quantify the vulnerability or extinction risk of a species due to climate change.

Each framework draws on different input variables and combines them in different ways, so they are not necessarily interchangeable. To allow for meaningful interpretation of the assessments by conservation practitioners and policymakers, it is necessary to evaluate whether the results of different frameworks are in agreement with one another; and this is currently unknown (Wade et al., 2016). If the results of species risk assessments do differ, the choice of framework would affect the perceived vulnerability of different species, hence changing conservation priorities and management actions. It is also unknown whether any of the different assessment frameworks provide a projection of risk that is accurate or realistic. Therefore, it is important that the frameworks should be validated using empirical data on observed changes to the status of species to determine which methods are most appropriate to use, something that has previously been absent from the literature (Wade et al., 2016).

Climate change vulnerability assessment methodologies follow two broad approaches (Pacifici et al., 2015) trait- and trend-based. Ultimately, how the population of a species responds to environmental change is strongly influenced by the unique combination of traits possessed by each species (and those it interacts with), so trait-based vulnerability assessment frameworks have much to commend them (Arribas et al., 2012; Barrows et al., 2014; Chin, Kyne, Walker, & McAuley, 2010; Foden et al., 2013; Gardali, Seavy, DiGaudio, & Comrack, 2012). Typical traits selected by these assessments include life-history information, but they may also incorporate trait data derived from distributional data (e.g. to estimate thermal limits). By contrast, trend-based frameworks (Pearce-Higgins, Ausden, Beale, Oliver, & Crick, 2015; Thomas et al., 2011; Triviño, Cabeza, Thuiller, Hickler, & Araújo, 2013) may recognize the importance of traits in ultimately determining risk, but focus primarily on abundance and distribution changes (observed and projected), supplemented by some trait information to inform assessors of the likelihood that projected trends will be realized. The merit of this approach is that it focuses on the primary cause of conservation concern (population and distribution decline, in the spirit of IUCN Red Listing), and sidesteps the need to identify every causal trait, or how these traits combine to determine population responses to climate change. Some studies have attempted to combine the two types into hybrid frameworks (Garnett et al., 2013; Heikkinen et al., 2010; Moyle, Kiernan, Crain, & Quiñones, 2013; Young et al., 2012), weighting one set of inputs most heavily or including trend-based data as an optional set of inputs. The ease of applying each of these frameworks depends on the availability of trait, trend and modelled input data for the taxon and region under consideration. In this regard, some frameworks have been developed with specific taxa in mind (Arribas et al., 2012; Chin et al., 2010; Gardali et al., 2012; Garnett et al., 2013; Heikkinen et al., 2010; Moyle et al., 2013; Triviño et al., 2013) particularly birds and other vertebrates, while others are generic; and they have been applied at a range of geographic scales (Table 1). However, they are all amenable to being scaled up or applied to different taxonomic groups with little or no adjustment.

In general, the frameworks attempt to quantify three major components (or some combination thereof) of risk: sensitivity, exposure

TABLE 1 Summary vulnerability framework information. Overall vulnerability equation used by each framework, broad methodology type, taxonomic group(s) used to test the framework, and geographic scale at which the framework was tested. The Pearce-Higgins et al., 2015 framework is a simplified version of the Thomas et al., 2011 framework, excluding exacerbating factors and including only trend data

General vulnerability equation	Framework	Methodology type	Taxon	Locality
Exposure × sensitivity	Gardali et al., 2012;	Trait	Birds	California State
	Young et al., 2012;	Hybrid	Molluscs, Fish, Amphibians, Birds, Mammals	Nevada State
	Moyle et al., 2013;	Hybrid	Freshwater fish	California State
	Garnett et al., 2013;	Hybrid	Birds	Australia
	Thomas et al., 2011;	Trend	Birds, Plants, Invertebrates	Great Britain
	Pearce-Higgins et al., 2015;	Trend	Birds, Plants, Invertebrates	Great Britain
Exposure × sensitivity × conservation status	Triviño et al., 2013;	Trend	Birds	Iberian Peninsula
Exposure × sensitivity × adaptive capacity	Chin et al., 2010;	Trait	Chondrichthyan fish	Great Barrier Reef
	Foden et al., 2013;	Trait	Birds, Amphibians and Corals	Global
Exposure + sensitivity	Barrows et al., 2014;	Trait	Plants, Mammals, Reptiles, Birds	Joshua Tree National Park, California
	Heikkinen et al., 2010;	Hybrid	Butterflies	Europe
Exposure + sensitivity + adaptive capacity	Arribas et al., 2012	Trait	Water beetles	Iberian Peninsula

and adaptive capacity (Dawson, Jackson, House, Prentice, & Mace, 2011; Williams, Shoo, Isaac, Hoffmann, & Langham, 2008). All approaches, whether trait- or trend-based, explicitly incorporate measures that are intended to represent both species exposure and species sensitivity to climate change (Table 1) but, beyond this, there is little agreement across the frameworks on exactly which input variables to use. This arises, in part, because there is limited evidence to identify which traits are most important in determining the sensitivity of a species to climate change (Pearson et al., 2014) or exactly how climate exposure should be quantified. A range of different inputs are therefore used to assess vulnerability, using a combination of projections from distribution models, population dynamics and life-history traits. These amount to 117 specific input variables across the 12 frameworks considered here, of which three-quarters are unique to a single framework; and only five of the 117 variables are represented in more than two frameworks (Table S1). Ideally, these differences would not matter and each framework would identify the same species as vulnerable, but this should be tested, not assumed. In addition to the variation in input variables used by different frameworks, there is inconsistency in whether inputs are considered measures of sensitivity, exposure or adaptive capacity. For example, metrics of dispersal are treated as indicating sensitivity (Barrows et al., 2014; Gardali et al., 2012; Heikkinen et al., 2010; Thomas et al., 2011; Young et al., 2012), or adaptive capacity (Arribas et al., 2012; Chin et al., 2010; Foden et al., 2013) depending on the framework used.

Here, we assess the utility of 12 published frameworks, using some of the best biodiversity data available. Initially, we consider whether the 12 frameworks generate consistent results, that is whether the frameworks "agree" on which species are at risk from climate change. We also consider the current Red List assessment approach, without incorporating any future projected declines using bioclimate envelope modelling, and compare the outputs against those from each of the 12 frameworks. We then validate the performance of the 12 different frameworks by carrying out an assessment based on historic species data and compare the outcomes to subsequent, observed changes in distribution and population. For frameworks that perform well in validation, species that are classified as at risk using historical data are expected to be most likely to have declined since then.

2 | MATERIALS AND METHODS

2.1 | Exemplar and real species comparisons

The assessments of 18 species (11 birds and 7 butterflies; hereafter "exemplar species," Table 2) and additional British bird and butterfly species (Table S4) were carried out based on trait and distribution data within Great Britain. These species were chosen due to the quality and availability of data for the taxa considered within this region. The exemplar species were chosen because they were the only species of any taxonomic group with both comprehensive distribution (in two or more time periods) and traits data and a northern or southern range margin lying within Great Britain (Gillingham et al., 2015). Species with range boundaries in a region are likely to be of interest when running climate change vulnerability assessments—in this case, species with a southern range edge in this temperate northern hemisphere situation should be more likely to be predicted to be at high climate risk than species with a northern range edge. All common British breeding bird and butterfly species were considered for the additional assessment, the 234 species selected being the ones for which future distributions could be modelled based on data availability.

Trait data for the real species were collected from the scientific literature and species atlas data (Asher et al., 2001; Balmer et al., 2013). Projected distribution changes were generated by applying a Bayesian, spatially explicit (Conditional Autoregressive) GAM to the bird and butterfly distribution data (Beale, Brewer, & Lennon, 2014). We used only a single climate modelling approach rather than an ensemble as our aim was to test framework performance rather than produce a definitive risk assessment of the species, and including projections from multiple models would have increased uncertainty and made comparison of framework outputs more difficult. Climate data for two emissions scenarios, low (UKCP09 B1) and medium (UKCP09 A1B), corresponding to a 2°C and 4°C increase in average temperature relative to a pre-industrial baseline by 2080 were used, as limiting the global rise above baseline temperatures to 2°C is widely considered key to avoiding the worst impact of climate change on species, while current estimates suggest 4°C may be a more realistic potential change (Mora et al., 2013). Both emissions scenarios show similar patterns of climate change of increasing mean temperature and total annual precipitation, with only the expected magnitude of change different between them, and are very close to recent observed changes. This pattern of change for the key climate variables in our model is consistent across the majority of global climate models, so using alternative future climate projections would likely yield a similar pattern of relative risk across species.

For each emissions scenario, we modelled species distributions using 11 different spatially coherent projections (SCPs), allowing us to incorporate uncertainty within each emissions scenario into the model outputs and giving us projected changes based on 22 future climate datasets per species. The change in distribution for a species was then calculated under each emissions scenario by averaging across the 11 different SCPs.

2.2 | Simulated species comparisons

To compare the outputs of the 12 risk assessment frameworks using simulated species, we generated ranges of values for 117 unique input variables (Table S1), covering characteristics such as species traits and population trends. We then drew values for each input variable to generate 10,000 combinations of "trait sets" that were used as simulated species in the assessments, in lieu of real world data for many species.

Where possible to do so, we applied constraints on the input variables to ensure logical consistency. For example, in the case of interspecific interactions, some frameworks ask broadly whether there is a dependence of a species on any interspecific interaction, while other frameworks require inputs relating to multiple, clearly

TABLE 2 Risk assessment output for exemplar real species. Low (white), medium (grey) and high (black) risk category outputs for the 18 exemplar species assessed using all 12 climate change vulnerability assessment frameworks. Assessments were carried out at the Great Britain scale, based upon contemporary data, with modelled future distributions based upon a medium emission scenario (A1B projection for 2070–2099). Northern (N, with a southern range margin) or southern (S, with a northern range margin) distributed species are identified in the distribution column

	Distribution	Chin	Gardali	Foden	Barrows	Arribas	Young	Moyle	Heikkinen	Garnett	Thomas	Pearce-Higgins	Triviño
Birds													
Black grouse (*Tetrao tetrix*)	N												
Capercaillie (*Tetrao urogallus*)	N												
Black-throated diver (*Gavia arctica*)	N												
Common scoter (*Melanitta nigra*)	N												
Red-throated diver (*Gavia stellata*)	N												
Slavonian grebe (*Podiceps auritus*)	N												
Bittern (*Botaurus stellaris*)	S												
Dartford warbler (*Sylvia undata*)	S												
Nightjar (*Caprimulgus europaeus*)	S												
Stone curlew (*Burhinus oedicnemus*)	S												
Woodlark (*Lullula arborea*)	S												
Butterflies													
Large heath (*Coenonympha tullia*)	N												
Mountain ringlet (*Erebia epiphron*)	N												
Northern brown argus (*Aricia artaxerxes*)	N												
Scotch argus (*Erebia aethiops*)	N												
Adonis blue (*Polyommatus bellargus*)	S												
Large blue (*Maculina arion*)	S												
Silver-spotted skipper (*Hesperia comma*)	S												

defined interspecific interactions. In this situation, it would not make sense for the broad interaction to be scored as absent while specific interactions are scored as present. In this case, the broad interaction is generated first and the scores of more specific interaction variables are influenced by that, to ensure consistent inputs across frameworks.

For continuously distributed input variables, upper and lower bounds were set based on reported values from the literature (e.g. body size, generation time) or theoretical minimum and maximum values. A value for the variable for each simulated species was then drawn from a uniform distribution bounded by those upper and lower limits. Species current distributions were simulated using the same approach, sampling a value for area occupied (in km^2) from a uniform distribution with an upper limited based on known real world distribution limits. For projected changes to species distributions under climate change, a future distribution was generated using the same process as for current distributions, and the percentage change in area between the two calculated.

The uniform distribution was chosen for all variables (equal probability for binary and categorical variables) because, for many input variables, there was little or no data available on how they might be distributed or the covariance between traits in reality (and they differ greatly between taxonomic groups), so an arbitrary selection of distribution would have been needed. Nonetheless, where there was an a priori expectation of the distribution of a trait based on the literature (e.g. logarithmic scaling of dispersal distance), the uniform draw was taken from between the transformed trait limits. The uniform distribution also allows for generation of traits covering the full range of the potential parameter space for the input variables, which was one of the main advantages of generated trait sets rather than a larger sample of real species data. The results therefore test consistency in framework performances, rather than the "true" frequencies of risk (which we do not know, given the differences between framework methods).

Many of the input variables are categorical, typically scored as low/medium/high or a variation thereof. In some cases, it is possible to base these on a continuous variable which is then split into the different categories (e.g. dispersal distance <1 km scored as low, dispersal distance >1 km and <10 km scored as medium, dispersal distance >10 km scored as high). Where it has not been possible to generate a continuous variable to base the categorical split on (e.g. impact of climate mitigation measures—scored as low to high), the category was instead assigned randomly to one of the possible options, with an equal probability of assignment to each. IUCN Red List conservation status was required as an input to one of the frameworks and was generated using IUCN criteria A to D based on simulated traits, with no projected future changes considered. This conservation status for each simulated species was also used in comparisons of Red List risk category against risk category for each framework and therefore informs us of the relationship between climatic and nonclimatic risks rather than whether the Red List could adequately take climate change into account.

2.3 | Validation

Given the large variation in the risk categories assigned to each real and simulated species, validation is required to assess whether any of the vulnerability frameworks has any predictive power. To examine how well the different climate vulnerability assessments performed at projecting future risk we used the results of assessments based on historic species data to compare against observed recent trends in species distribution/abundance. For validation of the frameworks to produce robust results they need to be tested using reliable input data, poor quality input data will always lead to poor assessments of risk regardless of the method used for the assessment. We therefore utilized some of the best quality data available globally by selecting British birds and butterflies for the analysis.

Validations were carried out using historically available data to assign species to low-, medium- and high-risk categories (for each of the 12 risk assessment frameworks), as though the assessments were carried out in the past (i.e. excluding more contemporary information not available during the first time period), and then, we compared recent distribution and population changes for species that had been assigned to each risk category. Assessments for British birds were based on the time period 1988–1991, to match the breeding bird atlas data (Gibbons, Reid, & Chapman, 1993). Assessment inputs based on the "then-current" distribution/population were calculated from these Atlas data, with historic changes in distribution calculated from the 1968–1972 Atlas to the 1988–1991 Atlas (Gibbons et al., 1993). Projected changes in distribution were modelled using the 1988–1991 Atlas distribution data and future climate projections for 2080 under the medium (UKCP09 A1B) emissions scenario. Historic assessments for British butterflies were performed using the same approach, based on the 1995–1999 Millennium Butterfly Atlas (Asher et al., 2001) and historic trends calculated from the previous 1970–1982 national survey. Future projected distributions were modelled using the same methodology as for the bird species. A total of 181 British bird species and 53 British butterfly species were assessed based on this historic data.

In addition to the risk categorization outputs of the assessments, observed recent trend data for distribution and population change as the assessment time period was required. For bird distribution trends, data from the 2007–2011 Atlas was used, giving the percentage change in occupied 10 km grid squares between 1988–1991 and 2007–2011. Observed changes in population for birds were obtained from the State of the UK Birds report (Hayhow et al., 2015), as a percentage change in population from 1995 to 2013. Butterfly population change data were obtained from the State of the UK Butterflies report (Fox et al., 2011), giving a percentage change in population from 1995 to 2005. Although these dates partly overlap with the Millennium Butterfly Atlas (Asher et al., 2001), the population data are collected on fixed transects that are separate from the millions of independent distribution records that give rise to the Atlas maps. Distribution change data for the butterflies was not used in the analysis due to a large increase in observer effort in the latter time period, which resulted in increases in

distribution that are likely to reflect increased effort rather than true changes in distribution.

2.4 | Statistical analysis

The risk category outputs from each of the frameworks were converted to a set of standardized categories: low/medium/high risk (Table S2). Broad agreement between the frameworks was tested on a pairwise basis using Spearman's rank correlation, to establish how consistently species were assigned to the same low/medium/high risk categories by the different frameworks.

Rank correlation allows for a comparison of how well the different frameworks correspond across all levels of risk assignment, but a potentially more useful comparison is of how well they agree in identifying a species as high risk, based on the assumption that assessments will primarily be run to identify the species most vulnerable to climate change. To compare agreement on just high-risk species, the risk categories were further simplified to a binary, "low and medium" versus "high" categorization. Cohen's kappa, a measure of inter-rater reliability, was calculated to compare agreement between frameworks. The prevalence and bias-adjusted Cohen's kappa (PABAK) (Byrt, Bishop, & Carlin, 1993) was used due to the relatively low frequency of species scoring as high risk.

Principal component analysis (PCA) was used to examine how much of the variation in risk assignment was influenced by certain frameworks and to identify whether frameworks of the same general type (trait, trend) showed similar patterns in risk category assignment. Risk category outputs from each framework for the 10,000 simulated species were used in this analysis.

For the validation analysis, we predicted that most species at high risk due to climate change are more likely to have seen population/distribution decreases than species identified at low risk, and are unlikely to have seen increases over the period of the validation analysis. Species identified as low risk under climate change may still have declined due to nonclimatic factors, but should also include increases in population/distribution over the validation period. We used quantile regression to validate framework performance, with change in distribution or abundance as the response variable and framework risk categorization (low/medium/high) as the predictive factor (Cade & Noon, 2003). This allowed us to consider trends in the 0.50 and 0.75 quantiles of distribution/population change instead of just the mean, which would identify if the majority of high-risk species are declining as we would expect if a framework is performing well. The models were tested for significance against a null model using an ANOVA.

3 | RESULTS

3.1 | Consistency between the results of different vulnerability frameworks

We first assessed risk to the 18 exemplar species using each of the 12 frameworks and a medium emissions scenario. The results of the assessments were highly variable, with no single exemplar species assigned to the same risk category by all frameworks (Table 2). The majority of species were classified as high risk by at least one assessment (14/18 species); yet only one species (Black Grouse) was classified as high risk by at least half of the frameworks (Table 2). Pairwise Spearman's rank correlations between frameworks showed poor overall agreement in risk assignment (r_s mean = .17 ± .03, r_s median = .21). The ten "northern" species, with a southern range margin in Great Britain, were classified as higher risk on average than the eight "southern" species with a northern range limit in Great Britain, with average risk values of 1.7 and 1.4, respectively (scoring low/medium/high categories as 1/2/3), although only three of the eight southern distributed species were not classified as high risk by any of the frameworks (*Botaurus stellaris*, *Sylvia undata*, *Caprimulgus europaeus*) and one northern distributed species was not classified as high risk by any (*Tetrao urogallus*).

Focussing only on classification of species in the highest risk category, inter-rater reliability analysis (for high risk versus low or medium risk) produced a similar pattern to the rank correlation results, with "weak" (McHugh, 2012) agreement across frameworks (mean κ_{PABAK} = .51 ± .03, median κ_{PABAK} = .55). Almost exactly the same pattern was observed for the exemplar taxa when using a low emissions climate scenario: the average proportion of agreement between frameworks for the two scenarios was 95%, with only 11 changes in risk category across the two projected futures. The results for the low emissions scenario are given in Table S3.

The frameworks also showed poor overall agreement with the Red List assessment (r_s mean = −.28 ± .03, r_s median = −.25), and this agreement was not improved when we considered trait-based and trend-based frameworks separately (trait-based: r_s mean = −.39 ± .02, trend-based: r_s mean = .01 ± .01).

We further tested the frameworks with an additional 181 British bird and 53 British butterfly species (Table S4) for which data were available to model GB distribution changes, under a medium emissions climate change scenario. Of these 234 species, 131 were classified as high risk by at least one framework (56%) (Figure 1b), with only 12 species (2 bird and 10 butterfly species) classified into the same risk category by every framework. Pairwise rank correlations showed poor overall agreement (r_s mean = .18 ± .03, r_s median = .17), confirming that even with a larger sample of real species with strong correlations between traits, there was little consistency across the frameworks. In addition, inter-rater reliability analysis indicated weak (McHugh, 2012) agreement across frameworks when classifying species as high risk (mean κ_{PABAK} = .43 ± .03, median κ_{PABAK} = .61). We also ran the assessments for the 234 species using a low emissions climate change scenario, which produced the same overall pattern in risk and similar levels of agreement as for the medium emissions scenario.

All 10,000 simulated species were assessed individually using each of the 12 risk assessments. The frameworks showed broadly similar patterns in the overall assignment of risk to the real species, classifying the majority of species as low risk and relatively few as high risk (Fig. S1). However, over 75% of the 10,000 simulated

species were classified as high risk by at least one framework considered, and only 135 were assessed as high risk by more than half of the frameworks (Figure 1a). Overall, we found poor agreement across the frameworks in assigning risk (Figure 2, r_s mean = .07 ± .01, r_s median = .04). Pairwise correlations within broad framework types were stronger than the overall pairwise correlations (between trait-based frameworks: r_s mean = .13 ± .04, r_s median = .08; between trend-based frameworks: r_s mean = .29 ± .12, r_s median = .18), but still relatively poor. There was also little difference between frameworks designed for a single taxonomic group and more generic frameworks (between taxon-specific frameworks: r_s mean = .09 ± .05, r_s median = .04 and between generic frameworks: r_s mean = .11 ± .03, r_s median = .04). Using inter-rater reliability analysis to compare agreement between frameworks in their classification of simulated species in the highest risk category only, we again found weak overall agreement (mean κ_{PABAK} = .55 ± .02, median κ_{PABAK} = .52). This inconsistency suggests against using a consensus of contrasting methods as the basis for prioritization.

Comparing the outputs of the frameworks to Red List outputs also produced poor correlations (Figure 2: Spearman's rank correlation r_s mean = .04 ± .01, r_s median = .01), with trait-based assessments showing weaker correlation with Red List outputs than trend-based approach types (trait-based: r_s mean = .02 ± .01, r_s median = .01, trend-based: r_s mean = .11 ± .01, r_s median = .13).

To investigate similarities between the risk assignments of different frameworks further, we used Principal Components Analysis (PCA) on the risk category outputs. We found distinct clusters for trait-only frameworks (Arribas et al., 2012; Barrows et al., 2014; Chin et al., 2010; Foden et al., 2013; Gardali et al., 2012) and trend-based frameworks (Pearce-Higgins et al., 2015; Thomas et al., 2011; Triviño et al., 2013) with hybrid assessments falling between the two (Moyle et al., 2013; Young et al., 2012) (Figure 3, Table S5). This pattern is the same for the pairwise correlations between frameworks, with weak agreement overall, but stronger correlations within the five purely trait-based frameworks and within the three trend-based frameworks.

3.2 | Validation of different vulnerability frameworks

Overall, none of the frameworks showed strong predictive power (Table 3), with only two of the frameworks (Pearce-Higgins et al.,

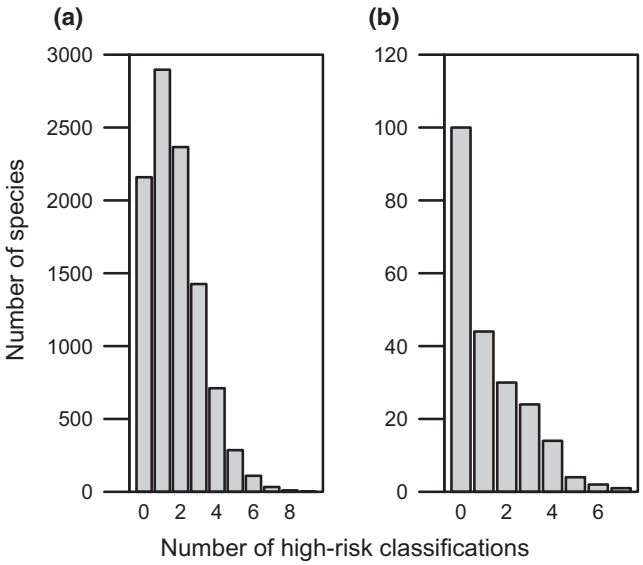

FIGURE 1 Frequency distribution of high-risk classifications for a) simulated species and b) real species assessed with historic data. The number of risk assessment frameworks under which each simulated or real species was classified as high risk

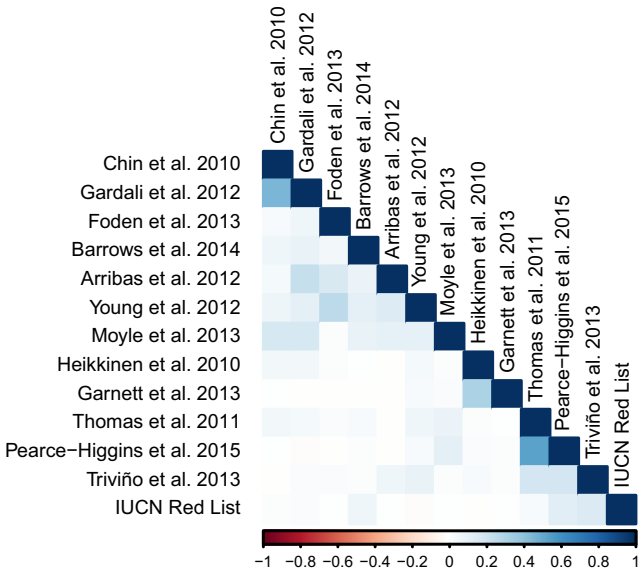

FIGURE 2 Correlation matrix showing spearman rank correlation coefficients (r_s) for each of the 12 frameworks, pairwise against the others and the Red List outputs for the simulated species. The matrix is a visual representation of the r_s value (see *x*-axis for range), with darker blue indicating a stronger positive correlation; using output data for the 10,000 simulated species

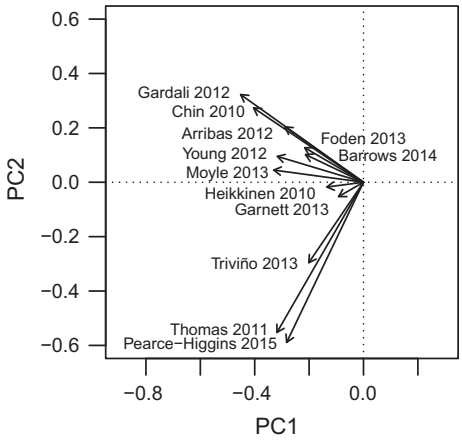

FIGURE 3 Principal component biplot. The first two principal components obtained by applying principal components analysis to the risk category outputs from the 12 frameworks for the 10,000 simulated species

2015; Thomas et al., 2011) producing significantly better-than-random risk assessments (one significant for the 0.50 and 0.75 quantiles, and one for the 0.75 quantile, Figure 4). Both of these frameworks are trend-based approaches, which would suggest incorporating this type of data into the assessment process produces more robust risk outputs. The results of validation for both birds and butterflies when using population change, rather than distribution change as the response variable, also suggested limited framework effectiveness. When considering changes in bird populations, there were no significant trends in the 0.50 quantile for any of the frameworks and only a single framework showed a significant trend for the 0.75 quantile (Fig. S2), although this was in the opposite direction to what we would expect for a framework performing well. There were no significant trends in either the 0.50 or 0.75 quantile for any of the 12 frameworks when assessing population change for butterflies (Fig. S3), although overall performance appeared to be better than for the bird population analysis.

3.3 | Validation using an ensemble approach

In addition to the individual framework validation, we also consider the effectiveness of using an ensemble approach to climate vulnerability assessment. We compared the modal risk category assigned to a species by the 12 frameworks against the same change in distribution/population value used in the individual framework validations. For the 181 bird species, only two had a modal risk classification of high risk, with both showing positive changes in distribution (Figure 5a) and population (Figure 5b), measured over the validation period. The 53 butterfly species also had just two species with a modal high-risk classification, with one increasing its population over the validation period and the other showing little change in its population (Figure 5c). Therefore, the ensemble approach did not identify high-risk species that subsequently declined—and across all species, there was no link between the consensus risk category and subsequent distribution trends in quantile regressions. We also considered the maximum risk category assigned by an ensemble approach (Fig. S4), which was also not significant and would be impractical to use to set conservation priorities because maximum risk identified over half the bird and butterfly species as high risk (Figure 1b). When considering average risk score (again scoring low/medium/high categories as 1/2/3), the values for both birds and butterflies ranged from 1.0 to 2.3, with a median score of 1.3 for birds and 1.4 for butterflies. There is relatively little variation across the average risk scores, which would make prioritization based on this measure difficult due to the difficulty involved in trying to differentiate between the scores.

4 | DISCUSSION

4.1 | Assessment comparisons and validation

Risk assessments for both real and simulated species showed poor overall agreement across the 12 frameworks, particularly between trend- and trait-based approaches. These inconsistencies between methods hold, regardless of whether we take into account the correlated traits that exist for real species within a given taxonomic group or if we minimize correlations between traits in simulated species (given that different higher taxa possess dissimilar

TABLE 3 Summary validation trends. Directions of trends in either distribution or abundance change for birds and butterflies from low risk species to high-risk species. A negative trend indicates the framework is performing as expected, and a positive trend indicates poor framework performance. Significant trends are denoted with *. The frameworks are ranked first by number of significant negative trends and then by number of non-significant negative trends

		Bird distribution trend direction		Bird population trend direction		Butterfly population trend direction				
Framework	Methodology Type	0.50 quantile	0.75 quantile	0.50 quantile	0.75 quantile	0.50 quantile	0.75 quantile	Correct significant trends	Correct non-significant trends	Rank
Thomas et al., 2011;	Trend	−*	−*	–	–	–	–	2	4	1
Pearce-Higgins et al., 2015;	Trend	−*	−*	–	+	–	+	2	2	2
Young et al., 2012;	Hybrid	–	–	–	–	–	–	0	6	3.5
Barrows et al., 2014;	Trait	–	–	–	–	–	–	0	6	3.5
Garnett et al., 2013;	Hybrid	–	–	–	+	–	–	0	5	5
Arribas et al., 2012;	Trait	–	–	+	+	–	–	0	4	6.5
Triviño et al., 2013;	Trend	–	+	–	–	–	+	0	4	6.5
Gardali et al., 2012;	Trait	–	–	+	–	+	+	0	3	8.5
Chin et al., 2010;	Trait	–	–	+	–	+	+	0	3	8.5
Moyle et al., 2013;	Hybrid	+	+	+	+	–	–	0	2	10
Foden et al., 2013;	Trait	+	+*	+	+	–	–	0	2	11
Heikkinen et al., 2010	Hybrid	+	+*	+	+*	+	+	0	0	12

trait correlations), which we might expect to have caused greater inconsistencies between frameworks, depending upon the degree of similarity between the traits considered. The similarities between our results for simulated and real species suggest that the inconsistencies arise from differences between the risk framework methods themselves (i.e. which variables are included in an assessment, and how they are combined to place each species in a risk category) rather than from the test data that we used. Uncertainty in the assessments is likely to be increased if projections of future distributions from multiple modelling approaches are considered, rather than the single approach we have utilized here, suggesting that the results if used for definitive risk assessments of species could be even more variable than we have demonstrated.

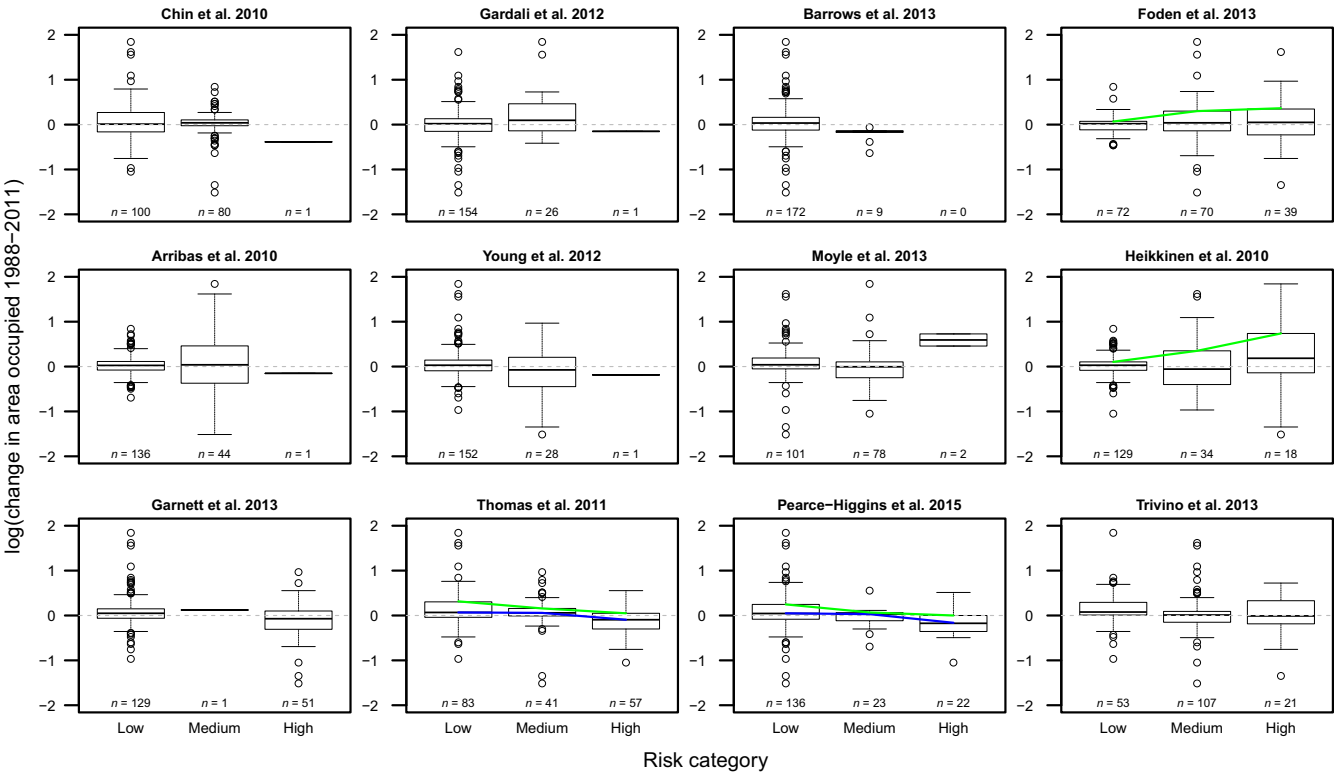

FIGURE 4 Validation boxplots showing logged change in bird distribution against simplified risk category for each of the 12 risk assessment frameworks. Blue lines show a significant trend in the 0.50 quantile, and green lines show a significant trend in the 0.75 quantile. Assessments are for 181 British bird species

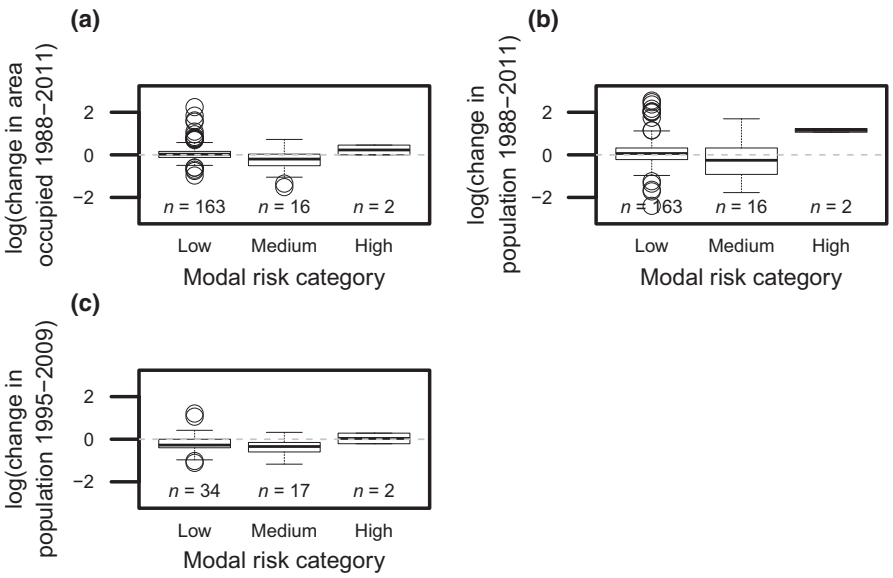

FIGURE 5 Validation boxplots showing a) logged change in bird distribution, b) logged change in bird population and c) logged change in butterfly population, against modal simplified risk category from across all 12 risk assessment frameworks

Given that real and simulated species are assigned different climate-risk categories by different risk assessment frameworks, it is essential that validations are carried out to assess whether none, some or all of the frameworks have predictive power. The validation analysis here revealed that most frameworks perform poorly (Table 3). Only two methods (Pearce-Higgins et al., 2015; Thomas et al., 2011), both of which were trend-based, assigned risk appropriately (i.e. the high-risk species declined more than lower risk species) and significantly (Figure 4); although predictions were only significant when considering change in distribution as the response variable, not change in population (top two rows of Table 3). One of these methods (Thomas et al., 2011) also generated nonsignificant predictions in the expected direction in all of the other tests (top row of Table 3). These two methods are closely related to one another, with both using predicted trends based on climate as the driving force, with one (Thomas et al., 2011) using additional trait/habitat information that modifies the capacity of each species to respond as predicted. These additional constraints apparently increased the predictive power of this framework.

Some of the other frameworks do show a similar overall pattern, but assign such small numbers of species to the high-risk category that it was not possible to detect significant trends (see Figure 4). For example, one trait-based framework (Barrows et al., 2014) failed to assign any species to the high-risk category (and only between 9 and 13 to the medium-risk category) and one hybrid framework (Young et al., 2012) only assigned either one or five species to high risk across the three tests.

Two of the frameworks (Garnett et al., 2013; Heikkinen et al., 2010) classify species into risk categories based on proportions (e.g. top tenth of values assigned high risk) instead of consistently set threshold values, as seen in the other frameworks. The risk outputs from these two frameworks correlate poorly with most others, and they fall close to the origin in the PCA (Figure 3). Another framework (Foden et al., 2013) uses proportional cut-offs for some input data and along with a method that uses proportional risk categories (Heikkinen et al., 2010) performs poorly overall in the validation analysis, with significant trends in the opposite direction to that expected if assigning risk suitably. Proportions of species at risk from climate change are not expected to be the same in different regions (or taxonomic groups), so we recommend avoiding proportional approaches.

4.2 | Consensus assessment approach

As each framework we tested gives markedly different results, it limits the effectiveness of using the assessments to inform conservation responses. A potential alternative is to consider the results from an ensemble of climate vulnerability assessments. The high variability in outputs, however, also limits the effectiveness of taking an ensemble approach. We considered three possible approaches to this. The first was to consider the possibility that there are many different mechanisms of endangerment from climate change, and hence to consider a species as at risk if any of the 12 methods classified it as at high risk. This was not practically useful because the majority of species were identified as high risk using this approach. The second was to assign species to the modal class of vulnerability, which resulted in almost no species being classified as high risk. Neither approach significantly identified declining species in the validation. The third approach considered was the average risk score across the 12 frameworks, which again identified very few species as high risk and with very little variation in scores between them.

None of the outputs from the ensemble approaches offer sufficient improvement over any individual method to justify the time and effort required to collect the data to run all the assessments. Combining the results of different climate vulnerability assessments also has the potential problem of a single input variable appearing in multiple methods, which could lead a single species characteristic having an unduly large influence on the overall risk score.

4.3 | Validation analysis limitations

It should be noted that the time period for the observed changes used in the validation analysis are relatively short for both birds and butterflies (20 and 10 years, respectively), and from a period when a range of other pressures have also affected species' population in the area considered, particularly changes in agricultural management (Burns et al., 2016). There is a possibility that some species considered may be climate-threatened but not yet showing a strong negative response in distribution or population, while others may be limited by other factors, potentially leading to the underestimation of longer term framework performance. In particular, species that might be expected to be most climate threatened by changing patterns of extreme weather events, such as droughts or floods, are unlikely to have experienced the full impact of this over the time period used in our validation.

However, we would expect frameworks to show some separation between expanding and contracting species, because both bird and butterfly communities have responded to climate change during this period (Davey, Chamberlain, Newson, Noble, & Johnston, 2012; Devictor et al., 2012), for example by polewards range shifts (Gillings, Balmer, & Fuller, 2015; Mason et al., 2015; Massimino, Johnston, & Pearce-Higgins, 2015). The fact we do not see such a pattern for most assessments (and some trends are the reverse of those expected), combined with the results of our comparison between frameworks, does highlight the lack of evidence currently available to support the use of most of these frameworks. As some of the assessments are designed for global assessments of risk, there is a possibility that the poor performance is a consequence of applying them over a regional scale. As data becomes available, it would be valuable to repeat our analysis at the scale of entire species distributions, rather than on regions or subpopulations, to test this. However, these methodologies are being applied at nonglobal scales by researchers and practitioners (Meng et al., 2016), so the results of our validation at a regional scale remain applicable to how the methods are actually being used.

4.4 | Future climate vulnerability assessment use

The science underpinning trend-based approaches is stronger, with increasing evidence that species distribution models, which are used to measure exposure in trend-based approaches, can retrodict recent population and range trends (Green et al., 2008; Gregory et al., 2009; Stephens et al., 2016). There remains uncertainty around identifying the key traits influencing species vulnerability to climate change (Pacifici et al., 2017; Pearson et al., 2014), which may vary widely by taxonomic group and could explain the wide range of inputs across the different trait-based assessments. Recent work (Willis et al., 2015) has advocated the combination of elements of trait-based vulnerability assessments with species distribution modelling to produce more realistic projections of future risk. This approach has already been implemented to different extents by some frameworks considered here (Thomas et al., 2011; Triviño et al., 2013; Young et al., 2012), although the outputs of these show at best weak correlations with purely trait-based assessments, suggesting that trait-only assessments may not adequately capture the exposure component of climate risk. The two general types of assessment (trait and trend) effectively represent different paradigms, with combined approaches representing arbitrarily weighted blends of the two.

We have demonstrated that different vulnerability assessment frameworks should not be used interchangeably when attempting to assess a species' potential future risk to climate change, because assessments made with either real or simulated species produce conflicting results. Our validation results suggest there is currently less evidence to support the use of purely trait-based vulnerability assessments than trend-based approaches, although neither performed very strongly for the species', time period and location tested. Ideally, further tests of these approaches in different circumstances, for different taxa and locations, would be valuable. Trend-based approaches are the only type of methodology shown to assign species to appropriate risk categories, particularly when this information is supplemented with additional species trait data. If this conclusion is supported by other studies in other contexts, it would restrict the assessment options available to practitioners (e.g. without long-term monitoring data, trend-based approaches will not be possible). However, if frameworks not incorporating this type of information produce highly uncertain results, their long-term value remains questionable. Without significant investment in long-term monitoring, to study change as it occurs, and in research to identify exactly what traits make a species' vulnerable to climate change, our ability to identify the species most in need of conservation attention in the face of climate change will remain limited.

ACKNOWLEDGEMENTS

This work was supported by the Natural Environment Research Council (grant NE/L008866/1).

CONFLICT OF INTEREST

The authors have no conflict of interest to declare.

REFERENCES

Akçakaya, H. R., Butchart, S. H. M., Watson, J. E. M., & Pearson, R. G. (2014). Preventing species extinctions resulting from climate change. *Nature Climate Change*, *4*, 1048–1049.

Arribas, P., Abellán, P., Velasco, J., Bilton, D. T., Millán, A., & Sánchez-Fernández, D. (2012). Evaluating drivers of vulnerability to climate change: A guide for insect conservation strategies. *Global Change Biology*, *18*, 2135–2146.

Asher, J., Warren, M., Fox, R., Harding, P., Jeffcoate, G., & Jeffcoate, S. (2001). *The millennium atlas of butterflies in Britain and Ireland*. Oxford: Oxford University Press.

Balmer, D. E., Gillings, S., Caffrey, B., Swann, R. L., Downie, I. S., & Fuller, R. J. (2013). *Bird Atlas 2007–11: The breeding and wintering birds of Britain and Ireland*. Thetford: BTO.

Barrows, C. W., Hoines, J., Fleming, K. D., Vamstad, M. S., Murphy-Mariscal, M., Lalumiere, K., & Harding, M. (2014). Designing a sustainable monitoring framework for assessing impacts of climate change at Joshua Tree National Park, USA. *Biodiversity and Conservation*, *23*, 3263–3285.

Beale, C. M., Brewer, M. J., & Lennon, J. J. (2014). A new statistical framework for the quantification of covariate associations with species distributions. *Methods in Ecology and Evolution*, *5*, 421–432.

Burns, F., Eaton, M. A., Barlow, K. E., Beckmann, B. C., Brereton, T., Brooks, D. R., . . . Gregory, R. D. (2016). Agricultural management and climatic change are the major drivers of biodiversity change in the UK. *PLoS ONE*, *11*, e0151595.

Byrt, T., Bishop, J., & Carlin, J. B. (1993). Bias, prevalence and kappa. *Journal of Clinical Epidemiology*, *46*, 423–429.

Cade, B. S., & Noon, B. R. (2003). A gentle introduction to quantile regression for ecologists. *Frontiers in Ecology and the Environment*, *1*, 412–420.

Ceballos, G., Ehrlich, P. R., Barnosky, A. D., García, A., Pringle, R. M., & Palmer, T. M. (2015). Accelerated modern human-induced species losses: Entering the sixth mass extinction. *Science Advances*, *1*, e1400253.

Chin, A., Kyne, P. M., Walker, T. I., & McAuley, R. B. (2010). An integrated risk assessment for climate change: Analysing the vulnerability of sharks and rays on Australia's Great Barrier Reef. *Global Change Biology*, *16*, 1936–1953.

Davey, C. M., Chamberlain, D. E., Newson, S. E., Noble, D. G., & Johnston, A. (2012). Rise of the generalists: Evidence for climate driven homogenization in avian communities. *Global Ecology and Biogeography*, *21*, 568–578.

Dawson, T. P., Jackson, S. T., House, J. I., Prentice, I. C., & Mace, G. M. (2011). Beyond predictions: Biodiversity conservation in a changing climate. *Science*, *332*, 53–58.

De Grammont, P. C., & Cuarón, A. D. (2006). An evaluation of threatened species categorization systems used on the American continent. *Conservation Biology*, *20*, 14–27.

Devictor, V., van Swaay, C., Brereton, T., Brotons, L., Chamberlain, D., Heliola, J., . . . Jiguet, F. (2012). Differences in the climatic debts of birds and butterflies at a continental scale. *Nature Climate Change*, *2*, 121–124.

Foden, W. B., Butchart, S. H. M., Stuart, S. N., Vié, J-C., Akçakaya, H. R., Angulo, A., . . . Mace, G. M. (2013). Identifying the world's most climate change vulnerable species: A systematic trait-based assessment of all birds, amphibians and corals. *PLoS ONE*, *8*, e65427.

Fox, R., Brereton, T. M., Asher, J., Botham, M. S., Middlebrook, I., Roy, D. B., & Warren, M. S. (2011). *The state of the UK's butterflies 2011*.

Wareham, Dorset: Butterfly Conservation and the Centre for Ecology & Hydrology.

Gardali, T., Seavy, N. E., DiGaudio, R. T., & Comrack, L. A. (2012). A climate change vulnerability assessment of California's at-risk birds. *PLoS ONE*, *7*, e29507.

Garnett, S. T., Franklin, D. C., Ehmke, G., VanDerWal, J. J., Hodgson, L., Pavey, C., ... Williams, S. E. (2013). *Climate change adaptation strategies for Australian birds*. Gold Coast: National Climate Change Adaptation Research Facility.

Gibbons, D. W., Reid, J. B., & Chapman, R. A. (1993). *The new atlas of breeding birds in Britain and Ireland: 1988–1991*. London: T. & A.D. Poyser.

Gillingham, P. K., Bradbury, R. B., Roy, D. B., Anderson, B. J., Baxter, J.M., Bourn, N. A. D., ... Thomas, C. D. (2015). The effectiveness of protected areas in the conservation of species with changing geographical ranges. *Biological Journal of the Linnean Society*, *115*, 707–717.

Gillings, S., Balmer, D. E., & Fuller, R. J. (2015). Directionality of recent bird distribution shifts and climate change in Great Britain. *Global Change Biology*, *21*, 2155–2168.

Green, R. E., Collingham, Y. C., Willis, S. G., Gregory, R. D., Smith, K. W., & Huntley, B. (2008). Performance of climate envelope models in retrodicting recent changes in bird population size from observed climatic change. *Biology Letters*, *4*, 599–602.

Gregory, R. D., Willis, S. G., Jiguet, F., Voříšek, P., Klvaňová, A., van Strien, A., ... Green, R. E. (2009). An indicator of the impact of climatic change on European bird populations. *PLoS ONE*, *4*, e4678.

Hayhow, D. B., Bond, A. L., Eaton, M. A., Grice, P. V., Hall, C., Hall, J., ... Wotton, S. (2015). *The state of the UK's birds 2015*. Sandy, Bedfordshire: RSPB, BTO, WWT, JNCC, NE, NIEA, NRW and SNH.

Heikkinen, R. K., Luoto, M., Leikola, N., Pöyry, J., Settele, J., Kudrna, O., ... Thuiller, W. (2010). Assessing the vulnerability of European butterflies to climate change using multiple criteria. *Biodiversity and Conservation*, *19*, 695–723.

Mace, G. M., Collar, N. J., Gaston, K. J., Hilton-Taylor, C., Akçakaya, H. R., Leader-Williams, N., ... Stuart, S. N. (2008). Quantification of extinction risk: IUCN's system for classifying threatened species. *Conservation Biology*, *22*, 1424–1442.

Mason, S. C., Palmer, G., Fox, R., Gillings, S., Hill, J. K., Thomas, C. D., & Oliver, T. H. (2015). Geographical range margins of many taxonomic groups continue to shift polewards. *Biological Journal of the Linnean Society*, *115*, 586–597.

Massimino, D., Johnston, A., & Pearce-Higgins, J. W. (2015). The geographical range of British birds expands during 15 years of warming. *Bird Study*, *62*, 523–534.

McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, *22*, 276–282.

Meng, H., Carr, J., Beraducci, J., Bowles, P., Branch, W. R., Capitani, C., ... Burgess, N. D. (2016). Tanzania's reptile biodiversity: Distribution, threats and climate change vulnerability. *Biological Conservation*, *204*, 72–82.

Mora, C., Frazier, A. G., Longman, R. J., Dacks, R. S., Walton, M. M., Tong, E. J., ... Giambelluca, T. W. (2013). The projected timing of climate departure from recent variability. *Nature*, *502*, 183–187.

Moyle, P. B., Kiernan, J. D., Crain, P. K., & Quiñones, R. M. (2013). Climate change vulnerability of native and alien freshwater fishes of California: A systematic assessment approach. *PLoS ONE*, *8*, e63883.

Pacifici, M., Foden, W. B., Visconti, P., Watson, J. E. M., Butchart, S. H. M., Kovacs, K. M., ... Rondinini, C. (2015). Assessing species vulnerability to climate change. *Nature Climate Change*, *5*, 215–224.

Pacifici, M., Visconti, P., Butchart, S. H. M., Watson, J. E. M., Cassola, F. M., & Rondinini, C. (2017). Species' traits influenced their response to recent climate change. *Nature Climate Change*, *7*, 205–208.

Pearce-Higgins, J. W., Ausden, M. A., Beale, C. M., Oliver, T. H., & Crick, H. Q. P. (2015). *Research on the assessment of risks & opportunities for species in England as a result of climate change*. Natural England Commissioned Reports, Number 175.

Pearson, R. G., Stanton, J. C., Shoemaker, K. T., Aiello-Lammens, M. E., Ersts, P. J., Horning, N., ... Akçakaya, H. R. (2014). Life history and spatial traits predict extinction risk due to climate change. *Nature Climate Change*, *4*, 217–221.

Pimm, S. L., Jenkins, C. N., Abell, R., Brooks, T. M., Gittleman, J. L., Joppa, L. N., ... Sexton, J. O. (2014). The biodiversity of species and their rates of extinction, distribution, and protection. *Science*, *344*, 1246752.

Stephens, P. A., Mason, L. R., Green, R. E., Gregory, R. D., Sauer, J. R., Alison, J., ... Willis, S. G. (2016). Consistent response of bird populations to climate change on two continents. *Science*, *352*, 84–87.

Thomas, C. D., Cameron, A., Green, R. E., Bakkenes, M., Beaumont, L. J., Collingham, Y. C., ... Williams, S. E. (2004a). Extinction risk from climate change. *Nature*, *427*, 145–148.

Thomas, C. D., Hill, J. K., Anderson, B. J., Bailey, S., Beale, C. M., Bradbury, R. B., ... Yardley, T. (2011). A framework for assessing threats and benefits to species responding to climate change. *Methods in Ecology and Evolution*, *2*, 125–142.

Thomas, J. A., Telfer, M. G., Roy, D. B., Preston, C. D., Greenwood, J. J. D., Asher, J., ... Lawton, J. H. (2004b). Comparative losses of British butterflies, birds, and plants and the global extinction crisis. *Science*, *303*, 1879–1881.

Triviño, M., Cabeza, M., Thuiller, W., Hickler, T., & Araújo, M. B. (2013). Risk assessment for Iberian birds under global change. *Biological Conservation*, *168*, 192–200.

Wade, A. A., Hand, B. K., Kovach, R. P., Muhlfeld, C. C., Waples, R. S., & Luikart, G. (2016). Assessments of species' vulnerability to climate change: From pseudo to science. *Biodiversity and Conservation*, *26*, 1–7.

Warren, R., Price, J., Fischlin, A., de la Nava Santos, S., & Midgley, G. (2011). Increasing impacts of climate change upon ecosystems with increasing global mean temperature rise. *Climatic Change*, *106*, 141–177.

Williams, S. E., Shoo, L. P., Isaac, J. L., Hoffmann, A. A., & Langham, G. (2008). Towards an integrated framework for assessing the vulnerability of species to climate change. *PLoS Biology*, *6*, e325.

Willis, S. G., Foden, W., Baker, D. J., Belle, E., Burgess, N. D., Carr, J. A., ... Butchart, S. H. M. (2015). Integrating climate change vulnerability assessments from species distribution models and trait-based approaches. *Biological Conservation*, *190*, 167–178.

Young, B. E., Hall, K. R., Byers, E., Gravuer, K., Hammerson, G. A., Redder, A., & Szabo, K. (2012). Rapid assessment of plant and animal vulnerability to climate change. In J. F. Brodie, E. S. Post, & D. F. Doak (Eds.), *Wildlife conservation in a changing climate* (pp. 129–150). Chicago, IL: University of Chicago Press.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Wheatley CJ, Beale CM, Bradbury RB, Pearce-Higgins JW, Critchlow R, Thomas CD. Climate change vulnerability for species—Assessing the assessments. *Glob Change Biol*. 2017;23:3704–3715. https://doi.org/10.1111/gcb.13759