

Habitat suitability estimated by niche models is largely unrelated to species abundance

Tad A. Dallas^{1,2,3}  | Alan Hastings^{1,2}

¹Center for Population Biology, University of California, Davis, California

²Department of Environmental Science and Policy, University of California, Davis, California

³Centre for Ecological Change, University of Helsinki, Helsinki, Finland

Correspondence

Tad A. Dallas, University of Helsinki, Centre for Ecological Change, Viikinkaari 1 PO Box 65, FI-00014 Helsinki, Finland.
Email: tad.a.dallas@gmail.com

Funding information

National Science Foundation, Grant/Award Number: DEB-1457652

Abstract

Aim: Data on species occurrences are far more common than data on species abundances. However, a central goal of large-scale ecology is to understand the spatial distribution of abundance. It has been proposed that species distribution models trained on species occurrence records may capture variation in species abundance. Here, we gauge support for relationships between species abundance and predicted climatic suitability from species distribution models, and relate the slope of this relationship to species traits, evolutionary relationships and sampling completeness.

Location: USA.

Time period: 1658–2017.

Major taxa studied: Mammal and tree species.

Methods, Results: To explore the generality of abundance–suitability relationships, we trained species distribution models on species occurrence and species abundance data for 246 mammal species and 158 tree species, and related model-predicted occurrence probabilities to population abundance predictions. Further, we related the resulting abundance–suitability relationship coefficients to species traits, geographic range sizes, evolutionary relationships and the number of occurrence records to investigate a potential trait or sampling basis for abundance–suitability relationship detectability. We found little evidence for consistent abundance–suitability relationships in mammal ($\bar{r} = .045$) or tree ($\bar{r} = -.005$) species, finding nearly as many negative and positive relationships. These relationships had little explanatory power, and coefficients were unrelated to species traits, range size or evolutionary relationships.

Main conclusions: Our findings suggest that species climatic suitability based on occurrence data may not be reflected in species abundances, suggesting a need to investigate nonclimatic sources of species abundance variation.

KEYWORDS

abundance–suitability relationship, climatic niche, GBIF, occurrence probability, species abundance, species distribution model

1 | INTRODUCTION

Understanding the climatic conditions that enable species persistence is a central goal in ecology, and numerous statistical techniques have been developed to understand and forecast species

geographic distributions based on climatic conditions (Drake, 2015; Elith & Leathwick, 2009; Phillips & Dudík, 2008). These models – commonly referred to as species distribution models or niche models (Peterson & Soberón, 2012) – attempt to estimate species occurrence probabilities or habitat suitability based on presence (and

sometimes absence) records given spatial or climatic data (Hijmans & Graham, 2006). These values are a reflection of climatic suitability, such that higher values may correspond to more favourable environments. A logical, and quite common, assumption is that these favourable environments will support higher species abundances. This suggests the possibility that climatic suitability values generated from models trained on binary data may yield information on species abundance given a set of climatic conditions.

These so-called *abundance–suitability* relationships have been examined previously, yielding mixed evidence, with some studies finding strong support (Gutiérrez, Harcourt, Díez, Illán, & Wilson, 2013; VanDerWal, Shoo, Johnson, & Williams, 2009), and others failing to detect an effect (Filz, Schmitt, & Engler, 2013; Nielsen, Johnson, Heard, & Boyce, 2005). Numerous reasons exist for the mixed support of abundance–suitability (Estrada & Arroyo, 2012). For one, the assumption that species are most abundant in the centre of their climatic niches may simply be unsupported (Dallas, Decker, & Hastings, 2017; Sagarin & Gaines, 2002), while habitat suitability estimates from niche models are typically higher in the niche interior.

Second, if climatic tolerances influence population growth rates, but not carrying capacities, equilibrium species abundance in a given site will be largely independent of climatic conditions. Third, variability in climatic conditions may be equally or more important relative to mean conditions, which is not commonly incorporated into species niche estimation and abundance–suitability relationships. Lastly, climatic variables may not influence species abundances as much as other forces such as dispersal limitation, resource availability, and species interactions such as competition and parasitism. Specifically, species interactions – either competitive or trophic – can strongly influence species local abundance in a community (Schoener, 1983), suggesting that the mixed support for abundance–suitability relationships may be related to local demographic processes (e.g., growth rates), as well as the community context (presence of predators, competitors and parasites). Evidence for variability in the relationship between environmental suitability and species abundance comes from a synthesis on marine protected areas (Lester et al., 2009) that demonstrates that reduced mortality through reduced fishing pressures can result in unchanged or decreased species density for a subset of species.

Despite mixed evidence, a recent meta-analysis has suggested abundance–suitability relationships may be quite general (Weber, Stevens, Diniz-Filho, & Grelle, 2017). However, two of the studies (Elmendorf & Moore, 2008; Pearce & Ferrier, 2001) used in the meta-analysis – which provided more than a third of the data used – contained inconsistencies that may influence the general findings. First, the abundance–suitability relationships from Pearce and Ferrier (2001) included in Weber et al. (2017) were rank correlations between suitability and abundance for all sites, including those for which the observed abundance was zero. Rank correlations on predicted climatic suitability and species abundances would be highly sensitive to zero abundance values, as this reduces the problem to more closely match predicting presence absence. That is, predicting low climatic suitability for zero abundance sites will

inflate rank correlation coefficients. Evidence for this comes from the greatly reduced support for abundance–suitability relationships when zero abundance sites were not considered (Pearce & Ferrier, 2001). Second, Elmendorf and Moore (2008) quantified abundance as the summed number of plots in which a species was recorded for 7 years, a measure perhaps more aligned with occupancy than abundance. However, removing data from these two studies does not strongly affect the overall findings of Weber et al. (2017) (see Supporting Information), suggesting a certain level of meta-analytical support for positive abundance–suitability relationships (Weber et al., 2017). While meta-analysis is one available method to quantify support, we examine an alternative here; using species occurrence databases paired with estimates of species abundance for a large number of species. These databases often contain a greater amount of data and are at a larger spatial scale than typical studies, allowing perhaps a more direct assessment of macroecological hypotheses.

Here, we analysed large-scale databases of species occurrences and abundances for 246 mammal and 158 tree species to investigate the generality of abundance–suitability relationships. We further explored the potential for abundance–suitability relationships to be influenced by species traits (body size), distributions (geographic range size), sample size (number of occurrence records) and species evolutionary relationships. Correlations between model-predicted occurrence suitabilities and independent estimates of species abundance were typically small, and largely non-significant, suggesting that occurrence suitabilities may not be as strongly related to abundance as believed. Further, species body mass, phylogenetic relationships, sample size and geographic range size were unrelated to the abundance–suitability correlation coefficient. Together, our failure to detect abundance–suitability relationships – as well as the lack of evidence for a trait, sampling, range size or phylogenetic basis – suggests that models meant to estimate species occurrence suitabilities are unreliable surrogates for species abundances.

2 | METHODS

2.1 | Species data

Mammal species occurrence data were obtained from the Global Biodiversity Information Facility (GBIF), using R package *rgbif* (Chamberlain, Ram, Barve, & Mcglinn, 2016). The *scrubr* package (Chamberlain, 2016) was used to reduce the set of species occurrences to a complete (no missing latitude and longitude values), possible (bounded by latitudinal and longitudinal limits) and likely (0°N, 0°W observations removed) set of species observations. Mammal species abundance data were obtained from the Mammal Community Database (MCDB; Thibault, Supp, Giffin, White, & Ernest 2011) a collection of mammal abundance data across 940 localities. The MCDB contains abundance information for a total of 308 species.

However, we considered only those species with greater than 10 unique occurrences that were identified to the species level,

resulting in a total of 246 mammal species. Species abundance data were standardized by the number of trap nights each study used; a common method to account for sampling effort. While information from the original studies on the area sampled was unavailable in the current database, the number of trap nights is likely strongly associated with study area, assuming the density of mammal traps (which were most commonly Sherman small mammal traps) is relatively similar among studies. Further, while community scale data may be used to calculate relative abundance, this would potentially risk conflating species position in a community with species density. That is, species distribution models provide estimates of climatic suitability for a single species, while relative abundance depends, by definition, on the abundances of other species (also discussed in Weber et al., 2017). Species occurrence data obtained from GBIF for these 246 species were used to train niche models and estimate climatic suitability (described in detail below).

Tree species data were obtained from the United States Department of Agriculture (USDA) Forest Inventory and Analysis (Bechtold & Patterson, 2005), a long-term effort consisting of a fixed-radius design of over 120,000 one-acre plots, each composed of four subplots with nested microplots. Sites were chosen to maximize habitat type and land area covered at the state level [see Bechtold et al. (2005) and www.fia.fs.fed.us/ for more information]. Data were gathered across a wide range of years (1976–2015), and many plots were systematically revisited, especially in more recent years. To get a composite measure of abundance at the site level, we took the mean abundance of each species at each unique latitude and longitude coordinate pair. We used data on seedling distributions over much of the USA, where seedlings were defined as individual trees at least 6 inches tall for softwoods or 12 inches for hardwoods. Seedlings were sampled from the microplots as discussed above. Seedlings were chosen for their relative sensitivity to immediate climatic conditions. That is, adult trees may have established in a location with previously suitable, but presently unsuitable, climate, leaving a legacy effect potentially biasing modelling efforts.

For training species distribution models, the abundance data from standardized plots were treated as binary, that is, species with a non-zero abundance were considered as species occurrence records. This is different from our examination of mammal species distributions, as mammal occurrence data were obtained from the Global Biodiversity Information Facility. On the other hand, tree seedlings were sampled systematically across the USA according to established procedures as part of the USDA Forest Inventory and Analysis effort. This ensures that the range of climates used in species distribution models for trees is the same range for which abundance data are available, providing perhaps a more direct test of abundance–suitability relationships than we were able to perform for mammal species.

2.2 | Climate data

Bioclimatic data at 2.5 arc-minute resolution consisted of the 20 BioClim covariates (including altitude), which capture annual mean and variation in temperature and precipitation (Hijmans, Cameron,

Parra, Jones, & Jarvis, 2005). These data, which have been cited over 8,500 times to date, represent the current gold standard for biogeographic studies, despite being over a decade old, and capture both mean bioclimatic conditions and inter-seasonal climatic variation. These data represent interpolated climate values over space, and likely capture the relevant climatic variation experienced by mammal species in the MCDB data. However, we explored the possibility of scale effects of the climatic data (Pearson & Dawson, 2003) using 5 arc-minute resolution data in the Supporting Information, and providing code to use the 30 arc-second resolution in the associated figshare repository (<https://doi.org/10.6084/m9.figshare.6026777>). Our findings were not influenced by spatial resolution of climatic data (see Supporting Information). We further explore how the spatial resolution of abundance estimation could influence abundance–suitability relationships in the Supporting Information.

2.3 | Niche models and climatic suitability estimation

We determined suitability for occurrence by training boosted regression tree models (Friedman, 2001), a 'presence-background approach' to modelling species distributions (Elith, Leathwick, & Hastie, 2008). Models were fitted and evaluated using the *gbm* R package (Ridgeway, 2015). For each species, 5,000 background points were sampled from terrestrial environments within the bounding box encompassing all species occurrence points and including a 1-degree buffer on all sides. Boosted regression tree models were trained on a random subset of 70% of the data, using a maximum of 10,000 trees, an interaction depth of 3, and fivefold cross validation to avoid overfitting and determine the optimal number of trees. Model accuracy was calculated on the remaining 30% of data not used for training using the area under the receiver operating characteristic curve (AUC; Bradley, 1997), which is the area under the receiver operating characteristic (ROC).

2.4 | Abundance–suitability relationships

Population abundance estimates were analysed in two ways. First, we related the model-predicted occurrence probability to the abundance values from the empirical data. Second, we trained boosted regression tree models to capture species abundance based on the same climatic inputs as species occurrence data. This was performed due to the potential uncertainty in abundance estimates given snapshot data (for the mammal data) potentially not reflecting true species abundance at the locality. By training predictive models on the abundance data, we estimate a smooth surface that may be less 'noisy'. Unlike models of occurrence data, regression models were trained without background points, and model accuracy was assessed by comparing sampled abundance to model-predicted abundance. Due to data limitations, we do not have independent data to assess abundance model performance. However, the abundance models were internally cross-validated, reducing the likelihood of overfitting. Model accuracy was quantified as R^2 values between predicted and actual species abundances. Further, some species did not have enough abundance

estimates to train models, resulting in a reduced number of mammal ($n = 60$) and tree ($n = 158$) species that could be analysed. We present the raw abundance estimate analysis in the main text, and the modelled abundance relationships in the Supporting Information, where we discuss the potential shortcomings of relating predictions from models that share the entire set of covariates.

Trained species distribution models were used to estimate climatic suitability for species occurrence based on the climate at coordinates for which abundance data were available. The sign and significance of species abundance–suitability relationships were determined by relating model-predicted climatic suitability to model-predicted species abundance estimates using Pearson's correlation coefficients. We relax the assumption of a linear association between climatic suitability and abundance in the Supporting Information, using Spearman's rank correlation coefficient to quantify abundance–suitability relationships.

To estimate predictive capacity of abundance–suitability relationships, we calculated the coefficient of determination, which is equal to the squared Pearson's correlation coefficient. This measure provides an estimate of variation explained between abundance and suitability. Lastly, previous studies have suggested that abundance–suitability relationships may be difficult to detect if variance in abundance estimates is too low (Ashcroft et al., 2017). That is, low variance and background noise in population abundance estimation could lead to type II errors. We examine this in the Supporting Information by examining how the coefficient of variation in abundance estimates is related to model accuracy (Figure S13) and abundance–suitability correlation coefficients.

2.5 | The effects of species traits, sample size, distribution and evolutionary relationships

The detectability and strength of abundance–suitability relationships may be related to species traits, geographic distribution, sample size or evolutionary history (i.e., phylogenetic or taxonomic relationships). This could explain the mixed support for abundance–suitability relationships among different species groups or among different studies. To examine this, we related abundance–suitability correlation coefficients to species body size, geographic range size, sample size and evolutionary history (phylogeny for mammal and taxonomy for tree species). Species body size was quantified as body mass for mammals (obtained from the Pantheria trait database: Jones et al., 2009), and average plant height for trees (obtained from the TRY plant trait database: Kattge et al., 2011). Species geographic range size was determined by calculating the area of the minimum convex polygon that encompassed all species occurrence records. Sample size – which relates to the commonness and the amount of available data on species occurrences – was quantified as the number of species occurrence records for each species. Phylogenetic (Bininda-Emonds et al., 2007) and taxonomic (Chamberlain & Szocs, 2013) data were obtained for mammals and trees, respectively. We used Moran's I statistic to examine possible phylogenetic/taxonomic signals in abundance–suitability relationships.

3 | RESULTS

3.1 | Abundance–suitability relationships

Trained species distribution models were quite accurate for mammals ($AUC \pm SD = 0.92 \pm 0.085$) and tree seedlings ($AUC \pm SD = 0.96 \pm 0.04$), suggesting that models were able to accurately capture species occurrences as a function of climate variables. Independent of model accuracy, the predicted climatic suitability values were largely unrelated – or very weakly related – to species abundance (Figure 1). This did not change when species abundance was estimated using a statistical model (see Supporting Information).

Of the 246 mammal species examined, 14 had significantly positive abundance–suitability relationships and six had significantly negative abundance–suitability relationships. Similarly, of the 158 tree species examined, 38 had significantly positive abundance–suitability relationships and 35 had significantly negative abundance–suitability relationships. Nonlinear rank correlations found qualitatively similar results, although slightly more significantly positive relationships (20 for mammal species, and 45 for tree species) and more significantly negative relationships for tree species (51 species; see Supporting Information).

Despite significant relationships, the explanatory power of climatic suitability in predicting abundance was quite low (Figure 2). However, abundance–suitability relationships may be nonlinear, which would not be captured by Pearson's correlations. We examined the potential for nonlinear abundance–suitability relationships using rank correlations, finding agreement with our main text findings (see Supporting Information). Using the raw abundance estimates instead of conditioning abundance on climatic variables also provided agreement with these results (see Supporting Information).

3.2 | The effects of species traits, sample size, distribution and taxonomy

Geographic range size, the number of species occurrence records and species body size were unrelated to abundance–suitability coefficients for mammals (Figure 3) and trees (Figure 4), suggesting that these covariates are unlikely to drive the variation observed in abundance–suitability relationships. Further, while variation existed, there was no detectable difference in abundance–suitability relationships among species' taxonomic families for mammal (Figure S4) or tree (Figure S5) species. Further, we found no evidence for a phylogenetic signal in the abundance–suitability relationships for mammal ($I_o = -.014$, $I_e = -.005$, $p = .22$) or tree ($I_o = -.013$, $I_e = -.006$, $p = .10$) species.

4 | DISCUSSION

We failed to detect consistent abundance–suitability relationships, either in terms of sign – as positive and negative relationships were about equally common – or strength, as climatic suitability generated

from species distribution models did not explain an appreciable portion of variation in species abundances. This was not influenced by whether we used sampled abundance estimates or related these abundance estimates to climatic variables in an attempt to reduce intrinsic noise in estimates of abundance. Further, we failed to detect any association between abundance–suitability relationships

and species body size, geographic range area, sample size or species' taxonomic family. Our failure to detect abundance–suitability relationships may suggest that model estimates of climatic suitability are largely unrelated to species abundance. Our supplemental analyses using rank correlations further demonstrate that climatic suitability values cannot be used to estimate species relative abundance among

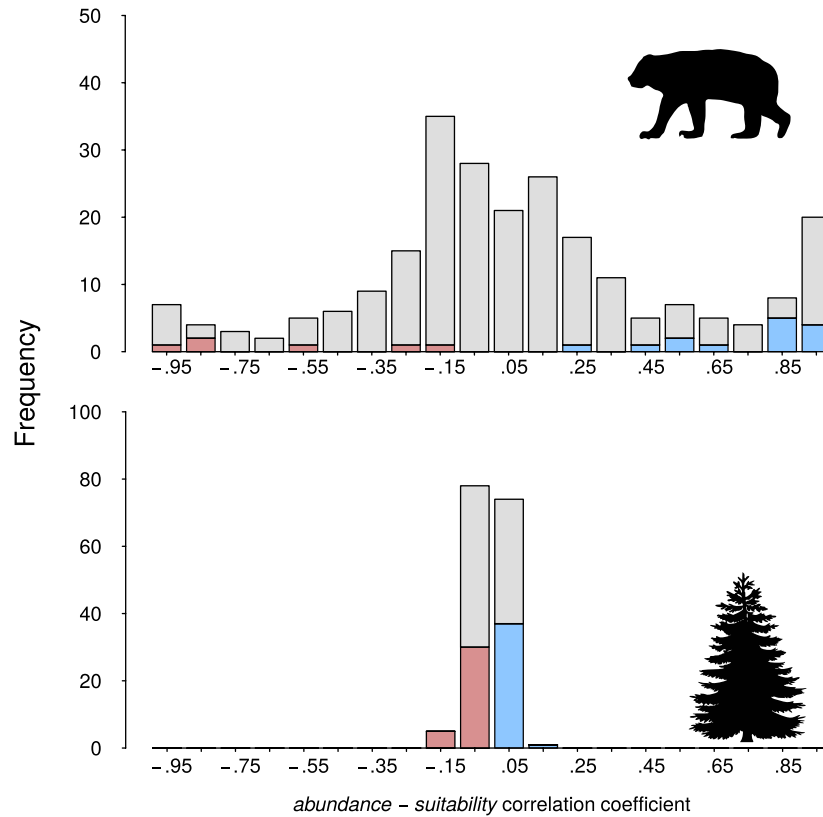


FIGURE 1 The majority of abundance–suitability relationships were weak, with coefficients largely around zero. Significantly positive (blue) and negative (red) were present for some species, although significant correlation coefficients do not correspond to predictive ability

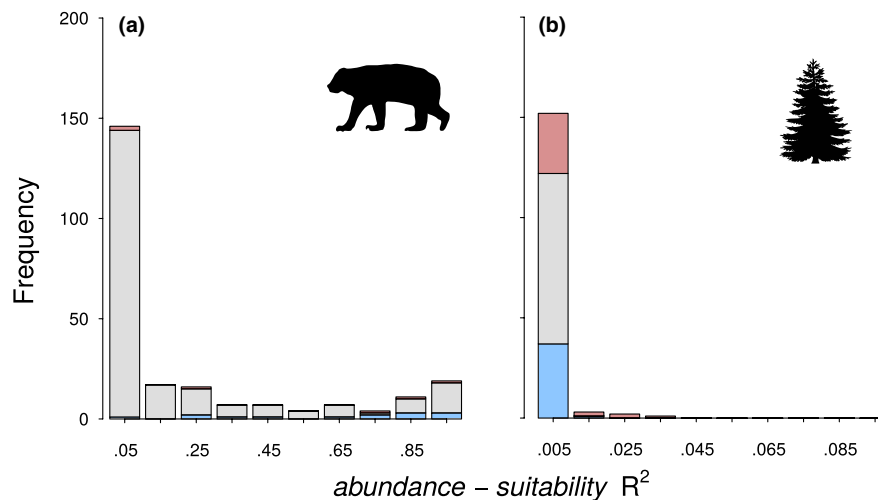


FIGURE 2 When correlation coefficients were significant – indicative of significant positive or negative abundance–suitability relationships – the explanatory power (R^2) tended to be quite low, suggesting that predictions of abundance made from climatic suitability would largely be inaccurate

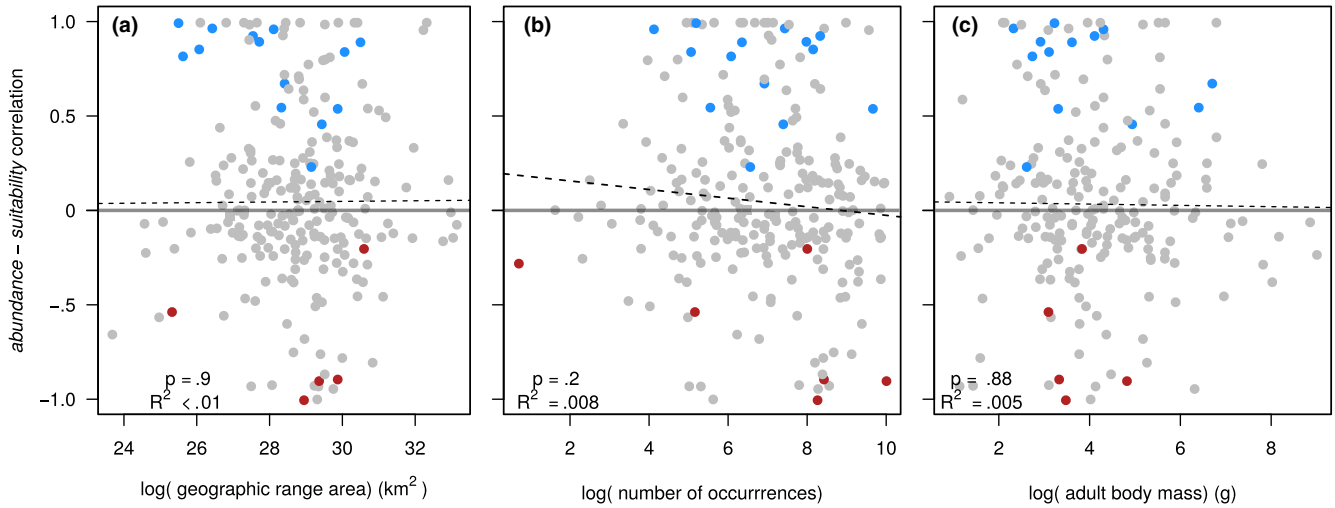


FIGURE 3 Mammal abundance–suitability correlation coefficients were unrelated to log-transformed (a) species geographic range size, (b) sample size and (c) body mass. *p*-values for each coefficient and coefficients of determination from linear models are provided in each panel, with dashed lines corresponding to linear model fits. Points are coloured based on the significance of abundance–suitability relationships, as we observed both positive (blue) and negative (red) relationships

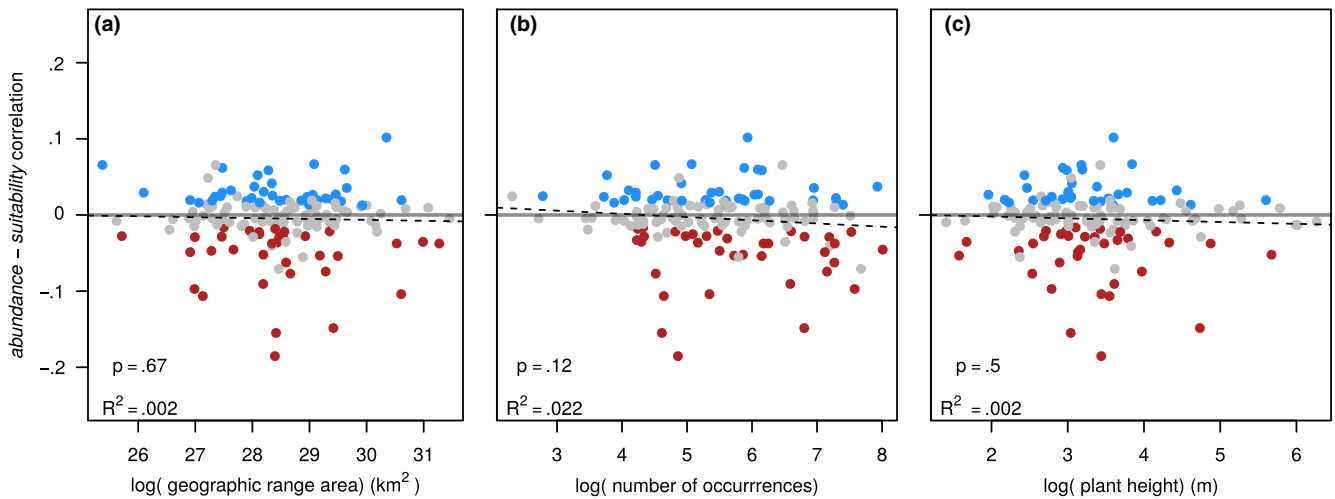


FIGURE 4 Tree abundance–suitability correlation coefficients were unrelated to log-transformed (a) species geographic range size, (b) sample size and (c) tree height. *p*-values for each coefficient and coefficients of determination from linear models are provided in each panel, with dashed lines corresponding to linear model fits. Points are coloured based on the significance of abundance–suitability relationships, as we observed both positive (blue) and negative (red) relationships

sites. Therefore, it seems unlikely that climatic suitability obtained from species distribution models would provide a suitable surrogate for species abundance to inform conservation efforts, guide sampling strategies, or provide insight into species demographic processes.

The abundance–suitability relationship is a logical expectation if species are most abundant toward the interior of their climatic niches, a common assumption with limited support (Dallas et al., 2017; Sagarin & Gaines, 2002). That is, provided that niche models predict high occupancy probabilities in those locations closer to the centre of the species climatic niche. Without this premise fulfilled, there is little reason to assume that climatic suitability or species occurrence probabilities would be related to species abundance. Further, significantly positive abundance–suitability relationships

would emerge only if climatic suitability enhanced species carrying capacities, instead of acting on species population growth rates. That is, variation in growth rates with climatic suitability would only be detectable through repeated sampling of nonequilibrium populations. Evidence in tree species suggests that climatic suitability values are weakly negative correlated with population growth rates, and weakly positively associated with carrying capacities (Thuiller et al., 2014), suggesting that it is unlikely to accurately infer demographic parameters from species occurrence suitability.

Perhaps it should not be surprising that climatic suitability does not translate to higher species abundance. Other ecological processes such as competition (Greiner La Peyre, Grace, Hahn, & Mendelsohn, 2001) or parasitism (Dobson & Hudson, 1992) may

limit population abundance independent of climate. However, our findings are not without limitation. While the Global Biodiversity Information Facility represents one of the best sources of species occurrence data (Edwards, 2004), it is subject to geographic biases and data quality issues (Beck, Böller, Erhardt, & Schwanghart, 2014). Our programmatic data cleaning procedures may reduce some of this bias, but data quality and bias issues still persist. The USDA Forest Inventory and Analysis seedling data are not influenced by this potential issue, as occurrence data were simply transformed mean abundance data from a systematic repeated survey of forests across the USA. A second concern is that abundance–suitability relationships may only manifest at smaller spatial scales. However, Weber et al. (2017) found little evidence for this, and we detected no effect of geographic range size or sample size on abundance–suitability relationships. A third concern is that the climatic data are not measured at the same resolution as species abundance estimates, which could influence the detection of abundance–suitability relationships. However, we found no evidence for this in the interpolated climatic data used in the current study (see Supporting Information).

Apart from the influence of spatial resolution, we might expect that geography and species demographic traits could influence abundance–suitability relationships. For instance, physical boundaries to dispersal may promote higher abundance in climatically unsuitable environments. This effect, in isolation or combination with dispersal limitation or slow-growing species, could produce near zero abundance–suitability correlations. Negative correlations, as observed for some mammal species in our study, could be the result of climate-dependent demographic processes. For instance, negative abundance–suitability relationships would be observed if a species shifted life history strategy in restrictive climates to have more, although potentially less fit, offspring (Monro, Sinclair-Taylor, & Marshall, 2010). Further, source sink dynamics could displace less dominant individuals to restrictive habitats (Pulliam & Danielson, 1991), which may promote near-zero or negative abundance–suitability relationships.

The lack of observed support for abundance–suitability relationships in the current study relative to previous meta-analytical support in Weber et al. (2017) suggests the potential presence of publication bias, and that species vary in the extent to which climatic suitability determines abundance. One interesting note is that Weber et al. (2017) found that incorporating environmental covariates (e.g., patch size) provided the strongest support of abundance–suitability relationships. This support may be due in part to covariates influencing species abundance more so than species occurrence potential. That is, covariates such as habitat patch size may strongly influence species abundance through the relationship between patch size and carrying capacity. This would suggest that any influence of patch size on species occurrence suitability prediction is driven not by climate but by factors directly related to the capacity of a habitat to support more individuals. Further, a positive relationship between species abundance and detection probability would further conflate species occurrence records with species abundance.

Whereas niche theory predicts that species occurrence probability or climatic suitability should be related to species abundance – as

abundance is predicted to be highest in the niche interior where the environment is most suitable (Brown, Mehlman, & Stevens, 1995) – there is surprisingly little evidence to support this idea. We find that abundance–suitability relationships tended to be weak and captured only small portions of variation in species' abundances. One of the motivating forces for examining abundance–suitability relationships is that abundance data are difficult and costly to collect, while occurrence data are plentiful. Perhaps a more fruitful path forward would come from mechanistic demographic models (Buckley et al., 2010; Cabral, Valente, & Hartig, 2017; Maurer & Taper, 2002) that could forecast the spatial distribution of species abundance over time. Predictions from these models could then be tested either using existing abundance data, or through targeted data collection efforts. Finally, continued collection and curation of scientific data from temporally replicated and standardized sampling efforts of populations and communities are needed to effectively gauge support for large-scale diversity patterns.

ACKNOWLEDGMENTS

We thank Easton White and two anonymous referees for constructive comments on an earlier manuscript version. We sincerely thank all individuals associated with the collection and curation of the open data on species abundances (Forest Inventory and Analysis and the Mammal Community Database), species occurrences (GBIF), species traits (Pantheria, TRY) and taxonomy (National Center for Biotechnology Information) used in this manuscript. The study has been supported by the TRY initiative on plant traits (<https://www.tryjdb.org>). The project was funded under National Science Foundation (NSF) grant Division of Environmental Biology (DEB)-1457652.

CONFLICT OF INTEREST

The authors have no conflicts of interest to declare.

AUTHOR CONTRIBUTIONS

TAD and AH designed the experiment. TAD analysed the data and wrote the manuscript. Both authors contributed to manuscript edits and provided approval prior to submission.

DATA ACCESSIBILITY

R code is available on figshare at <https://doi.org/10.6084/m9.figshare.6026777>. Data were obtained from the Global Biodiversity Information Facility (Chamberlain et al., 2016), the Mammal Community Database (Thibault et al., 2011), USDA Forest Inventory and Analysis (Bechtold et al., 2005), Pantheria (Jones et al., 2009) and TRY plant trait database (Kattge et al., 2011).

ORCID

Tad A. Dallas  <http://orcid.org/0000-0003-3328-9958>

REFERENCES

- Ashcroft, M. B., King, D. H., Raymond, B., Turnbull, J. D., Wasley, J., & Robinson, S. A. (2017). Moving beyond presence and absence when examining changes in species distributions. *Global Change Biology*, 23(8), 2929–2940.
- Bechtold, W. A., & Patterson, P. L. (2005). The enhanced forest inventory and analysis program-national sampling design and estimation procedures. *General Technical Report SRS-80*. Asheville, NC: US Department of Agriculture, Forest Service, Southern Research Station (Vol. 85). USDA.
- Beck, J., Böller, M., Erhardt, A., & Schwanghart, W. (2014). Spatial bias in the GBIF database and its effect on modeling species' geographic distributions. *Ecological Informatics*, 19, 10–15.
- Bininda-Emonds, O. R., Cardillo, M., Jones, K. E., MacPhee, R. D., Beck, R. M., Grenyer, R., ... Purvis, A. (2007). The delayed rise of present-day mammals. *Nature*, 446(7135), 507.
- Bradley, A. P. (1997). The use of the area under the roc curve in the evaluation of machine learning algorithms. *Pattern Recognition*, 30(7), 1145–1159.
- Brown, J. H., Mehlman, D. W., & Stevens, G. C. (1995). Spatial variation in abundance. *Ecology*, 76(7), 2028–2043.
- Buckley, L. B., Urban, M. C., Angilletta, M. J., Crozier, L. G., Rissler, L. J., & Sears, M. W. (2010). Can mechanism inform species distribution models? *Ecology Letters*, 13(8), 1041–1054. <https://doi.org/10.1111/j.1461-0248.2010.01479.x>
- Cabral, J. S., Valente, L., & Hartig, F. (2017). Mechanistic simulation models in macroecology and biogeography: State-of-art and prospects. *Ecography*, 40(2), 267–280. <https://doi.org/10.1111/ecog.02480>
- Chamberlain, S. (2016). scrubr: Clean Biological Occurrence Records. R package version 0.1.1. <https://cran.r-project.org/web/packages/scrubr/index.html>
- Chamberlain, S., Ram, K., Barve, V., & Mcglinn, D. (2016). *rgbif: Interface to the Global 'Biodiversity' Information Facility 'API'*. R package version 0.9.2.9430.
- Chamberlain, S., & Szocs, E. (2013). Taxize - taxonomic search and retrieval in R. *F1000Research* 2013, 2:191. <https://doi.org/10.12688/f1000research.2-191.v1>
- Dallas, T. A., Decker, R. R., & Hastings, A. (2017). Species are not most abundant in the center of their geographic range or climatic niche. *Ecology Letters*, 20(12), 1526–1533.
- Dobson, A. P., & Hudson, P. J. (1992). Regulation and stability of a free-living host-parasite system: *Trichostrongylus tenuis* in red grouse. II. Population models. *Journal of Animal Ecology*, 61(2), 487–498.
- Drake, J. M. (2015). Range bagging: A new method for ecological niche modelling from presence-only data. *Journal of the Royal Society Interface*, 12(107), 20150086.
- Edwards, J. L. (2004). Research and societal benefits of the global biodiversity information facility. *AIBS Bulletin*, 54(6), 485–486.
- 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.
- Elith, J., Leathwick, J. R., & Hastie, T. (2008). A working guide to boosted regression trees. *Journal of Animal Ecology*, 77(4), 802–813.
- Elmendorf, S. C., & Moore, K. A. (2008). Use of community-composition data to predict the fecundity and abundance of species. *Conservation Biology*, 22(6), 1523–1532.
- Estrada, A., & Arroyo, B. (2012). Occurrence vs abundance models: Differences between species with varying aggregation patterns. *Biological Conservation*, 152, 37–45.
- Filz, K. J., Schmitt, T., & Engler, J. O. (2013). How fine is fine-scale? questioning the use of fine-scale bioclimatic data in species distribution models used for forecasting abundance patterns in butterflies. *European Journal of Entomology*, 110(2), 311.
- Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 29(5), 1189–1232.
- Greiner La Peyre, M., Grace, J., Hahn, E., & Mendelssohn, I. (2001). The importance of competition in regulating plant species abundance along a salinity gradient. *Ecology*, 82(1), 62–69.
- Gutiérrez, D., Harcourt, J., Díez, S. B., Illán, J. G., & Wilson, R. J. (2013). Models of presence-absence estimate abundance as well as (or even better than) models of abundance: The case of the butterfly *Parnassius apollo*. *Landscape Ecology*, 28(3), 401–413.
- Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25(15), 1965–1978.
- Hijmans, R. J., & Graham, C. H. (2006). The ability of climate envelope models to predict the effect of climate change on species distributions. *Global Change Biology*, 12(12), 2272–2281.
- Jones, K. E., Bielby, J., Cardillo, M., Fritz, S. A., O'Dell, J., Orme, C. D. L., ... Connolly, C. (2009). Pantheria: A species-level database of life history, ecology, and geography of extant and recently extinct mammals. *Ecology*, 90(9), 2648–2648.
- Kattge, J., Diaz, S., Lavorel, S., Prentice, I. C., Leadley, P., Bönsch, G., ... Cornelissen, J. H. C. (2011). TRY - a global database of plant traits. *Global Change Biology*, 17(9), 2905–2935.
- Lester, S. E., Halpern, B. S., Grorud-Colvert, K., Lubchenko, J., Ruttenberg, B. I., Gaines, S. D., ... Warner, R. R. (2009). Biological effects within no-take marine reserves: A global synthesis. *Marine Ecology Progress Series*, 384, 33–46.
- Maurer, B. A., & Taper, M. L. (2002). Connecting geographical distributions with population processes. *Ecology Letters*, 5(2), 223–231.
- Monro, K., Sinclair-Taylor, T., & Marshall, D. J. (2010). Selection on offspring size among environments: The roles of environmental quality and variability. *Functional Ecology*, 24(3), 676–684.
- Nielsen, S. E., Johnson, C. J., Heard, D. C., & Boyce, M. S. (2005). Can models of presence-absence be used to scale abundance? Two case studies considering extremes in life history. *Ecography*, 28(2), 197–208. <https://doi.org/10.1111/j.0906-7590.2005.04002.x>
- Pearce, J., & Ferrier, S. (2001). The practical value of modelling relative abundance of species for regional conservation planning: A case study. *Biological Conservation*, 98(1), 33–43.
- Pearson, R. G., & Dawson, T. P. (2003). Predicting the impacts of climate change on the distribution of species: Are bioclimate envelope models useful? *Global Ecology and Biogeography*, 12(5), 361–371. <https://doi.org/10.1046/j.1466-822X.2003.00042.x>
- Peterson, A. T., & Soberón, J. (2012). Species distribution modeling and ecological niche modeling: Getting the concepts right. *Natureza & Conservação*, 10(2), 102–107.
- Phillips, S. J., & Dudík, M. (2008). Modeling of species distributions with maxent: New extensions and a comprehensive evaluation. *Ecography*, 31(2), 161–175.
- Pulliam, H. R., & Danielson, B. J. (1991). Sources, sinks, and habitat selection: A landscape perspective on population dynamics. *The American Naturalist*, 137, S50–S66.
- Ridgeway, G. (2015). gbm: Generalized Boosted Regression Models. R package version 2.1.1. <https://cran.r-project.org/web/packages/gbm/index.html>
- Sagarin, R. D., & Gaines, S. D. (2002). The “abundant centre” distribution: To what extent is it a biogeographical rule? *Ecology Letters*, 5(1), 137–147. <https://doi.org/10.1046/j.1461-0248.2002.00297.x>
- Schoener, T. W. (1983). Field experiments on interspecific competition. *The American Naturalist*, 122(2), 240–285.
- Thibault, K. M., Supp, S. R., Giffin, M., White, E. P., & Ernest, S. (2011). Species composition and abundance of mammalian communities. *Ecology*, 92(12), 2316–2316.
- Thuiller, W., Münkemüller, T., Schifffers, K. H., Georges, D., Dullinger, S., Eckhart, V. M., ... Moore, K. (2014). Does probability of occurrence relate to population dynamics? *Ecography*, 37(12), 1155–1166. <https://doi.org/10.1111/ecog.00836>
- VanDerWal, J., Shoo, L. P., Johnson, C. N., & Williams, S. E. (2009). Abundance and the environmental niche: Environmental suitability

estimated from niche models predicts the upper limit of local abundance. *The American Naturalist*, 174(2), 282–291.

Weber, M. M., Stevens, R. D., Diniz-Filho, J. A. F., & Grelle, C. E. V. (2017).

Is there a correlation between abundance and environmental suitability derived from ecological niche modelling? a meta-analysis. *Ecography*, 40(7), 817–828.

BIOSKETCH

TAD A. DALLAS is a postdoctoral researcher at University of Helsinki in the Centre for Ecological Change. When this work was completed, Tad was a postdoctoral researcher at the University of California-Davis. Tad is interested in population and community ecology at moderate to large spatial scales.

ALAN HASTINGS is a theoretical ecologist at University of California-Davis, a Hawaiian shirt (aloha wear) connoisseur, and a great mentor.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Dallas TA, Hastings A. Habitat suitability estimated by niche models is largely unrelated to species abundance. *Global Ecol Biogeogr.* 2018;00:1–9.
<https://doi.org/10.1111/geb.12820>