

CHAPTER 6

Ecological niche modelling

Damaris Zurell and Jan O. Engler

6.1 Context

In recent decades, ecology and biogeography have become increasingly quantitative disciplines. Spurred by technological advances, the amount of digital information is growing rapidly. Remote sensing and geographic information systems facilitate the mapping of environmental variables at high spatial and temporal resolution, and global positioning systems and citizen science platforms allow collecting biodiversity data in unprecedented amounts and precision. Many applications, however, require further processing and synthesis of these raw data. For example, conservation planning, reserve design, risk assessment, and invasive species management often require maps of species' actual or potential distributions rather than a limited number of samples of known occurrences (Franklin 2010). Ecological niche models (ENMs) provide an efficient way to interpolate (or sometimes extrapolate) species' distributions. They achieve this by relating species' observations to environmental predictor variables using statistical and machine-learning methods. If digital maps of environmental variables are available, the potential distribution of the species can then be predicted based on the estimated species–environment relationship (Franklin 2010; Guisan et al. 2017).

Applications of ENMs have exploded in recent years as more data are becoming available and the methodological tools are constantly improving. Also, in global change research, ENMs constitute the most widely applied modelling framework to

project potential future range shifts of species (IPBES 2016). Among all organisms, birds have received tremendous attention in ENM applications (Engler et al. 2017). This may in part relate to the deep knowledge of birds (Newton 2003), their high taxonomic coverage (Jarvis 2016), and huge public interest leading to myriads of opportunistic and systematic observations (cf. Chapter 4).

In this chapter, we first explain the concept behind ENMs and their underlying assumptions (section 6.2). We then describe the basic modelling steps and illustrate them using a simple real-world example (section 6.3), and provide an overview of potential sources of uncertainty in underlying data and in the models (section 6.4). After this conceptual and methodological overview, we discuss potential limitations of ecological niche models in a global change context (section 6.5) and outline the latest developments and future perspectives (section 6.6). As ENMs are a vast field under constant development, we will not be able to cover all aspects in depth but aim to provide a concise overview. For more detailed insights, we refer readers to excellent reviews (Elith and Leathwick 2009; Guisan and Thuiller 2005; Guisan and Zimmermann 2000) and books (Franklin 2010; Guisan et al. 2017) that summarize ENM research over the last few decades.

6.2 Theory and concept

What limits species' distributions? ENMs help us understand why species occur in some places but

not in others, and where else we could expect to find the species. The central conceptual framework underlying ENMs is the niche theory as formulated by Joseph Grinnell and later enhanced by G. Evelyn Hutchinson (see Soberón 2007, for an overview). Hutchinson distinguished two main factors that limit species' ranges: (i) the abiotic environment that determines the fundamental niche of a species, and (ii) the biotic environment that comprises all interactions with other species. First, Hutchinson described a species' fundamental niche as an n -dimensional hyperspace comprising all environmental conditions where a species has a positive population growth rate and can persist indefinitely. Second, as species rarely live in solitude but interact with other species, the biotic environment determines if a species can prevail in the presence of other species. The intersection between abiotic and biotic environment is usually referred to as the realized niche of a species.

Later, Pulliam (1988) and Hanski (1999) added source-sink and metapopulation dynamics as important determinants of species' distributions. According to these theories, a species could also be present in unsuitable 'sink' habitats that do not contribute to population growth or be absent from suitable habitats because of dispersal limitations. For example, stochastic events could lead to local extinction of species. Dispersal constraints will then determine whether the species is able to recolonize those areas or not. Figure 6.1 illustrates this heuristic in geographic space (Soberón 2007). A species can only persist in habitats where both the abiotic (A) and biotic environmental conditions (B) allow positive population growth. The intersection between A and B thus marks the potential distribution of the species. The movement or dispersal capacity of the species (M) will then determine which source habitats (allowing positive population growth) are occupied and which sink habitats (having negative population growth) are within reach of the species.

How to quantify species' ecological niches? ENMs provide a framework for quantifying the species–environment relationship and predicting a species' distribution in space and time. They are comparably easy to build because many software packages are available and they often have low data requirements (Franklin 2010; Guisan et al. 2017). As input, ENMs

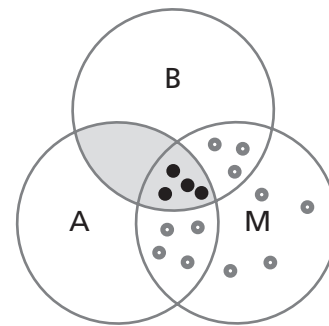


Figure 6.1 Heuristic representation of the factors limiting species' distributions in geographic space (adapted from Soberón 2007). A represents the geographic area where the species has positive population growth. The environmental conditions found within this area correspond to the fundamental niche. B represents the geographic area where a species can coexist with competitors. M represents the geographic area that is accessible to the species within the time period of interest and limited by the species' dispersal abilities. The intersection of A, B, and M is the occupied area that contains source populations (closed circles). The grey shaded area represents the region where abiotic and biotic environments are potentially suitable for the species and thus correspond to the classic view of the realized niche (Soberón 2007). Open circles represent sink populations with negative population growth rates due to suboptimal abiotic conditions or due to competitive exclusion.

use data on the geographic distribution of species' observations (occurrence, presence–absence, or abundance; the response or dependent variable) and environmental variables (the predictor or independent variables), which are now widely available in digital format (Chapters 3–4). Using adequate statistical algorithms, species' observations are then related to prevailing environmental conditions at the same locations. Apart from some purely spatial approaches such as Kriging, the statistical model is fitted in environmental rather than geographic space (Figure 6.2). Last, the species' potential distribution is predicted by calculating the occurrence probability of the species for any location based on the estimated species–environment relationship and the local environmental conditions as provided by digital maps of the current environment. If digital environmental layers for future (or past) climate scenarios are available, the potential distribution of the species can be forecasted (or hindcasted).

There is an ongoing debate whether the fitted species–environment relationship approximates the fundamental niche (area A in Figure 6.1), the realized niche (intersection of A and B in Figure 6.1), or

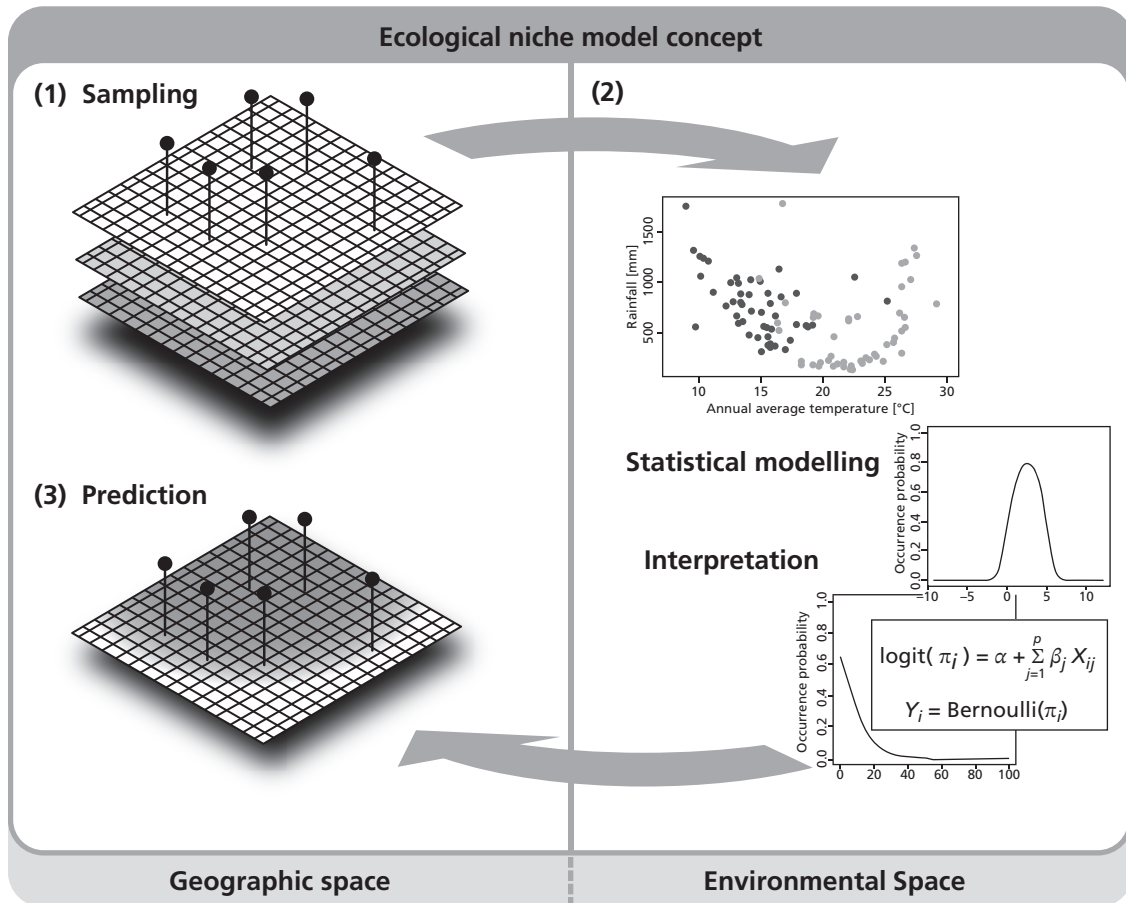


Figure 6.2 Schematic representation of the ecological niche modelling concept. First, data on species distributions and environmental variables are sampled in geographic space. Second, a statistical model (e.g. generalized linear model) is used to relate species observations to prevailing environmental conditions. Third, the potential distribution of the species is mapped by projecting the estimated species–environment relationship back onto geographic space, either to current environmental conditions or to selected climate change scenarios.

the occupied niche (intersection of A , B , and M in Figure 6.1). For example, it has been argued that ENMs based on coarse-scale climatic data describe the fundamental niche, models based on presence-only data the realized niche, and models based on presence–absence data the occupied niche (Franklin 2010). This debate is also reflected in the multitude of names that are used as synonyms for ENMs, for example climate envelopes, habitat model, resource selection functions, species distribution models, to name only a few. Here, we want to take a rather neutral view (similar to Elith and Leathwick 2009) and emphasize that which parts of a species' niche are modelled will depend on the underlying data, the

scale of the study, and its methods. It is important to carefully consider these factors in the model building process (sections 6.3 and 6.4).

Ideally, only ecologically relevant environmental variables should be used to infer species' niches. Thereby, it is useful to distinguish between proximal (resource and direct) and distal (indirect) gradients (Austin 1980; Engler et al. 2017). In the case of birds, resource gradients refer to all materials and energy that can be consumed by the species and are essential for their metabolism such as water, seeds, and insects. Direct gradients are environmental variables, such as temperature, that do not act as resources, but exert a causal (proximal) effect on physiology and

demography of species. By contrast, indirect gradients do not directly affect distributions of species, but are correlated with ecologically more relevant (proximal) predictors. Typical indirect gradients are topographic variables such as elevation or slope, and land cover variables such as habitat types. Indirect variables often provide simple surrogates of more complex combinations of resources and direct gradients (Guisan and Zimmermann 2000). For example, elevation may correlate strongly with temperature, moisture, and solar radiation. We want to stress that when ENMs are used for extrapolation as is the case in climate change research, direct and resource variables are preferable over indirect variables because correlation structures vary across the globe and may change under climate change. For deciding which variables are likely proximal or distal, prior knowledge on species' ecology and ecophysiology is certainly helpful (Guisan et al. 2017).

Ecological processes are highly scale-dependent and, thus, the spatial resolution of environmental and species data plays an important role for understanding and modelling species' niches (Pearson and Dawson 2003). Climate is generally believed to limit ranges at macro-scales, whereas habitat is thought to affect species' distributions at intermediate spatial scales. It is almost impossible to define one single best scale of study. Birds select resources and habitats in a hierarchical manner (Franklin 2010). Also, competition with heterospecifics may happen at a range of scales; for example, birds may compete for food over a large area (e.g., foraging

seabirds), but compete for nest sites at a much finer scale (e.g., on a cliff face for seabirds).

6.3 Modelling steps

The main modelling steps in ENMs are (i) conceptualization, (ii) data preparation, (iii) model fitting, (iv) model evaluation, and (v) spatiotemporal predictions, which we briefly explain below and which are detailed in Franklin (2010) and Guisan et al. (2017). Additionally, we develop a simple, real-world example to illustrate these modelling steps (Box 6.1) using the silvereye (*Zosterops lateralis*) as study species (Figure 6.3). Accompanying code and data are provided as the electronic supplementary material ESM 6.1. (www.oup.co.uk/companion/dunn&moller)

(i) **Conceptualization:** Before modelling, it is important to carefully consider the conceptual setup of the study. Foremost, this involves formulating clear research objectives and summarizing the available ecological knowledge on the species and study system. Based on these considerations, the main underlying assumptions are checked (cf. section 6.5), appropriate environmental predictors are identified and checked for availability, and, if necessary, a sampling strategy is designed for obtaining species' observations. Also, decisions on appropriate modelling algorithms should be taken at this stage and hypotheses about the expected shape of the species–environment relationship (linear or unimodal) should be formulated.

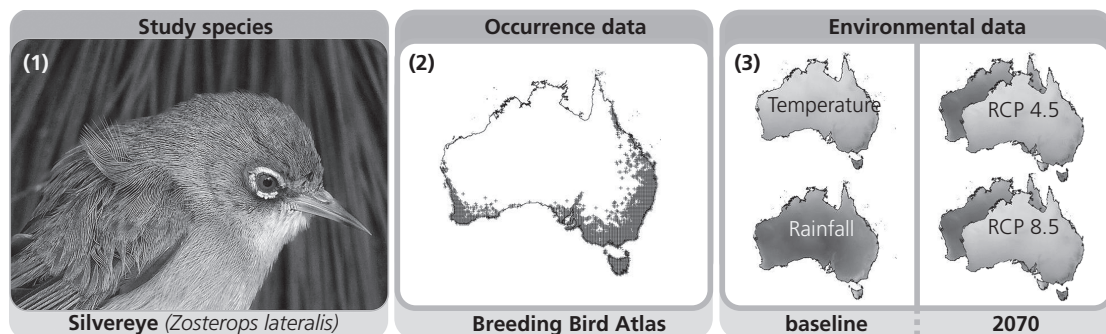


Figure 6.3 (1) The silvereye (*Zosterops lateralis*) is a small passerine bird of the south-west Pacific. (2) In Australia, the species occurs in more humid coastal areas in the south-west and east. (3) Bioclimatic variables such as annual average temperature and annual rainfall sums are available for current climate (baseline) and for climate scenarios under two Representative Concentration Pathways (RCP 4.5 and 8.5) for the year 2070. Picture credit: Rob Davis.

Box 6.1 Case study

Using publicly available data, we describe the main modelling steps for successfully fitting an ENM and projecting species' response to climate change. All data and source codes are available in the Electronic Supplementary Material ESM 6.1. We exemplify all modelling steps using the statistical software environment R (www.r-project.org) that offers different functionalities for ENMs. For simplicity, we use mainly functions contained in the R package *dismo* that provides various functions for data preparation and visualization and integrates different ENM algorithms for modelling and predictions.

(i) Conceptualization

We aim at assessing potential climate change effects on the silvereye (*Zosterops lateralis*, Figure 6.3), a common Australian bird species with stable population trends. We expect its distribution to be at equilibrium with the environment and its range edges to be predominantly determined by climatic conditions. Specifically, the silvereye is widespread along more humid coastal regions, while avoiding the dry and hot interior as well as the northern coast. We, thus, hypothesize that the species is limited mainly by dry and hot conditions on the Australian continent and expect a linearly increasing response to precipitation and a linearly decreasing response to temperature.

(ii) Data preparation

Occurrence information for the period 1977–1981 were available from the first Atlas of Australian Birds (Blakers et al. 1984) at a 10-minute resolution (accessed through www.gbif.org on February 6th 2018) (Figure 6.3). We use these atlas data rather than newer citizen science based occurrence records here to reduce any sampling biases. For subsequent modelling, we randomly selected 500 occurrence records and sampled 5000 random pseudo-absences outside known silvereye occurrences. As predictors, two bioclimatic variables, annual mean temperature (bio1) and sum of annual rainfall (bio12), were obtained from worldclim (v. 1.4; Hijmans et al. 2005) for the period 1960–1990 at the same 10-minute resolution and cropped to the Australian mainland (Figure 6.3). Additionally, we obtained CMIP5 climate change scenarios for the period 2061–2080 from one climate model (ACCESS1-0 GCM) and two representative concentration pathways (RCP 4.5 and 8.5).

(iii) Model fitting

The species–environment relationship was estimated using a generalized linear model (GLM) with a logit link function (Figure 6.4). GLMs constitute a comparably simple, parametric ENM approach and we chose it mainly for its simplicity. In practice, we recommend comparing multiple ENM algorithms (Franklin 2010; Guisan et al. 2017). As described above, and based on the species' distribution in Australia, we only included linear terms of annual temperature and precipitation sums. To avoid overfitting, we weighted pseudo-absences such that the sum of their weights equals the sum of presences.

(iv) Model evaluation

Model behaviour was assessed using inflated response curves (Figure 6.4). As expected, species' occurrence probability decreased with temperature and increased with precipitation. Other response shapes (e.g. unimodal) did not produce a better model fit. The resulting model was evaluated using an 80–20 split-sample approach. For simplicity, we only used a single iteration while, ideally, split samples should be repeated (e.g. 100 times) to reduce possible effects of outliers. Then, we calculated two evaluation statistics for the hold-out test data. Both the AUC (area under the receiver operating characteristic curve) value of 0.915 and the maximum kappa value of 0.58 indicate high predictive performance (Figure 6.4). For transforming continuous ENM output to binary predictions, we chose a threshold that maximizes the sum of sensitivity (true positive rate) and specificity (true negative rate).

(v) Spatiotemporal predictions

Maps of predicted silvereye distribution for current climate closely resemble the observed distribution (Figures 6.3 and 6.5). This strong niche filling confirms our equilibrium assumption and allows a solid climate change impact assessment. Our future projections suggest potential range contractions towards coastal areas in response to climate change with severe range losses especially in Southwest Australia. This could put the subspecies (ssp. *chloronothus*) that is endemic to the Southwest at high risk from climate change. Our uncertainty analysis showed that novel climatic conditions can be mainly expected in Northern Australian regions well outside the species' potential distribution, and will thus not affect our climate impact assessment.

(ii) **Data preparation:** Next, the data are prepared at the appropriate scale and preliminary analyses are performed. If new observations of species need to be obtained, this step will also involve consideration of the sampling design. GIS layers of the environmental predictor variables are prepared at the desired spatial and temporal grain (size of map unit) and extent (number of maps units). Potentially, some scale mismatches occur between different data layers that need to be dealt with by upscaling or downscaling grain and extent. ENMs are a classic application of spatial ecology. Nevertheless, we deliberately also mention the temporal scale here, because it is important to also match the temporal grain and extent of the environmental data with that of the species data and the assumptions underlying the study system. For example, we need to ensure that the considered climatic, land cover, and remote sensing data match the timeframe of the species data. If the species distribution data contain only presence records, then adequate background data (pseudo-absences) need to be selected for the ENM. The best strategy for deriving pseudo-absences will depend on the research question and on the ENM algorithm (Guisan et al. 2017).

(iii) **Model fitting:** A number of different ENM algorithms are available (cf. section 6.4.2), from which one or several have been selected during the conceptualization phase. Important aspects to consider during model fitting are how to deal with multicollinearity in environmental predictor variables, with spatial autocorrelation, how many predictors can be included in the model (without overfitting), how to select relevant predictors, and how to select from or average different candidate models.

(iv) **Model evaluation:** The fitted model should be analysed in depth (Figure 6.4). First, the realism of the fitted response curves and the residuals need to be checked. Second, to assess predictive performance, the model should ideally be validated against independent data. Truly independent data are rarely available, in which case resampling methods can be useful. The predictive accuracy is then assessed based on different measures that evaluate discrimination or classification ability as well as goodness-of-fit. Most ENM algorithms will yield continuous output, the probability of species' occurrence. For some validation measures, this

continuous output needs to be transformed into binary predictions, for which an appropriate thresholding approach needs to be selected.

(v) **Spatiotemporal predictions:** Once the ENM has been fitted, the model can be used for making spatial predictions. In most cases, these will be continuous maps of predicted occurrence probability or binary maps of potential species' distribution (Figure 6.5). Predictions to new times and places, for example under different climate change scenarios, are usually referred to as projections to make explicit that uncertainty is involved. When attempting projections, it is imperative to check for novel environmental conditions in the extrapolation data. If the future climatic range has no contemporary analogue and the ENM is extrapolating beyond the environmental range it was calibrated on, these projections should be clearly marked as uncertain (Zurell et al. 2012a). Also, different ENM algorithms will likely show different extrapolation behaviours and, thus, the fitted response curves (cf. model evaluation above) might need re-investigation to assess potential consequences. Projections also require paying special attention to the assumptions underlying ENMs, foremost the equilibrium assumption that is likely to be violated under climate change (cf. section 6.5).

Model building is an iterative process. Each step in the model building process may yield new insights on the species and the study system. Subsequently, single steps could be improved, for example, by adjusting the sampling design, the selection of environmental data or the model fitting, which will re-initiate the modelling cycle. Although many software packages are available, successfully fitting and evaluating predictive ENMs is far from being trivial and we highly recommend studying advanced books on ENMs (Franklin 2010; Guisan et al. 2017).

6.4 Sources of uncertainty

Different sources of uncertainty can affect the quality of ENMs and their projections for climate change scenarios. Here, we provide a brief overview of potential sources of uncertainty in underlying data, in the modelling approaches and in future scenarios. Thus, this section deals with uncertainty related to rather technical issues while we will provide more

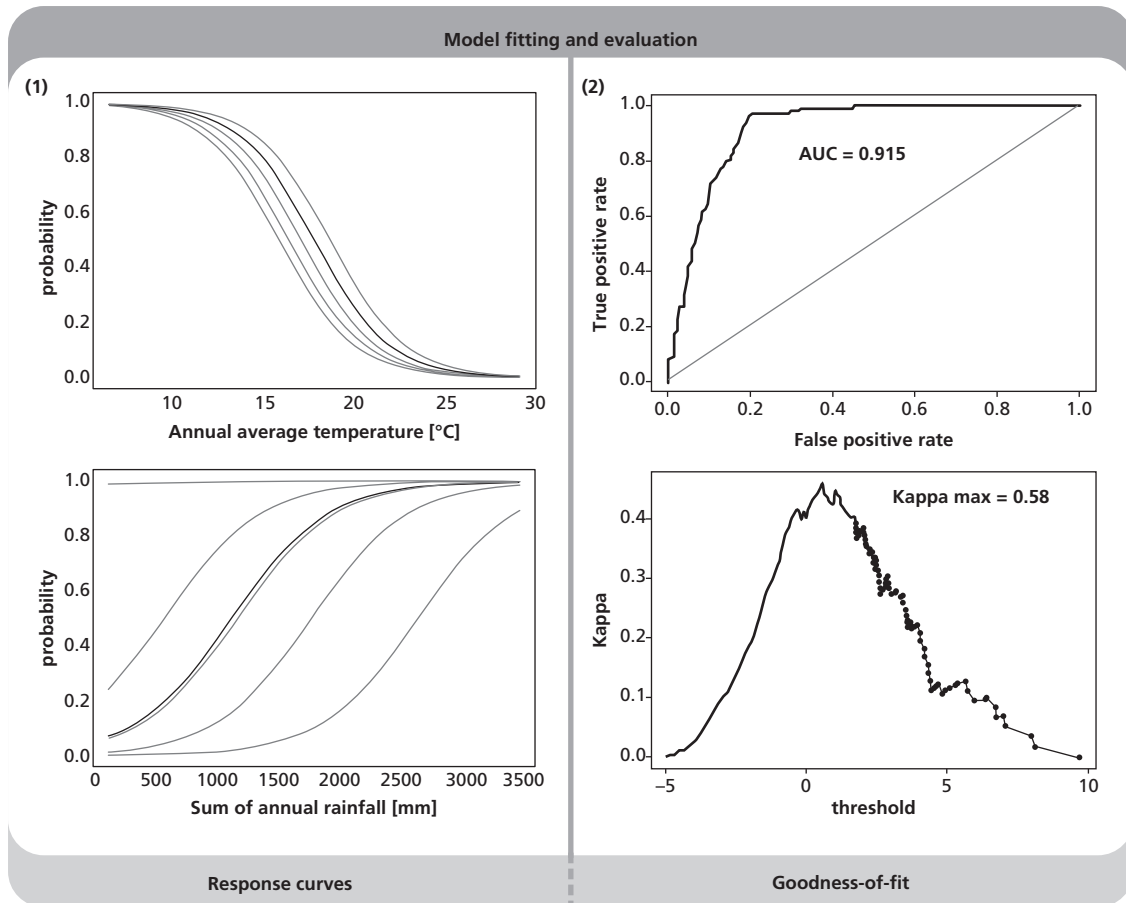


Figure 6.4 Model fit and model evaluation for the silvereye. (1) Response curves illustrate the fitted species' response to different environmental gradients and facilitate interpretation of ENMs as well as comparison to initial hypotheses about species' biology. (2) Evaluation statistics allow judging the model's predictive accuracy. The silvereye model shows overall high predictive performance.

critical thoughts on underlying assumptions and bird biology in section 6.5.

6.4.1 Data sources

Species data are available from many different sources including range maps, national monitoring programmes, atlas projects, and citizen science data (Chapter 4). In the cause of an ENM study, also one's own species data might be sampled. Generally, any source of species data needs to be checked for reliability and potential biases as well as its adequacy for answering the question at hand. This is especially true when downloading data from

large databases such as GBIF (www.gbif.org) that contain records from different sources with different reliability. Typical challenges of species data for ENMs include uncertainty in spatial and temporal resolution, positional uncertainty, biased sampling design or uneven spatial coverage, low sample sizes, and imperfect detection.

The adequate spatial resolution of the species data depends on the research question and on the error, or positional uncertainty, the user is willing to accept in the model (Graham et al. 2008). If point observations are available, ENMs could, in principle, be applied at comparably fine resolution. However, in mobile animals such as birds, it is crucial to think about the

meaning of a sampled point location. Is the observation of an individual at any place representative of a foraging location, a chance passage, or a home range, and which spatial resolution would best represent this? At the very least, the spatial resolution used for the ENM must be at least twice as coarse as the positional error, e.g. the potential error around a GPS location. At the other extreme are expert-drawn range maps. In order to use these as input for an ENM, the range maps need to be rasterized to an arbitrary spatial resolution. As species do not occur at all locations throughout their range, these maps represent species' occurrences only at very coarse resolution while they contain a lot of false positives at fine resolutions (Hurlbert and Jetz 2007).

A good sampling design should systematically cover all major environmental gradients of the study area and should, ideally, encompass the entire niche of the species in order to yield robust ENMs. Most monitoring and atlas programmes follow this principle while museum and citizen science data often suffer from spatial sampling bias, as a consequence of uneven coverage of presence records throughout the species' range. Often, museum and citizen science data are biased towards better-surveyed areas (e.g. popular bird watching areas, in national parks and hiking areas, or near urban areas) while having low sampling density or even data voids in remote areas. Strategies to deal with such spatial sampling bias include spatial filtering of the occurrence records or manipulation of the background (pseudo-absence) data such that the latter contain the same spatial bias as the presence records (Fourcade et al. 2014). Spatial filtering will be especially difficult in data-poor situations. As a rule of thumb, ten presence observations should be available per environmental predictor variable used in the ENM. In general, the minimum amount of presence records is advised to be between 20 and 50 (Guisan et al. 2017). Many species occurring in Asia, Africa, or South America suffer from few available samples (Meyer et al. 2015) as well as spatial sampling bias, making reliable ENM predictions for these regions difficult.

Additional uncertainty in the data may arise from imperfect detection. Low detectability can lead to false absences, meaning the species is not recorded although present. Indirect observations such as aural detection or environmental DNA can easily lead to false presences,

meaning the species is recorded although absent. If data from repeated surveys or additional information on detection distances or detection time are available, then hierarchical frameworks can be used to model detection probability (additional to the species' environmental niche) by explicitly taking into account the observation process (Guillera-Arroita 2017).

Environmental data from remote sensing products (He et al. 2015) and climate databases (Hijmans et al. 2005; see also Chapter 3) are becoming increasingly available at fine spatial resolution for large regions, often the globe. Yet, precision as well as spatial and temporal coverage of the data may vary across regions and across datasets, and it is important to understand underlying assumptions and potential biases in these data.

Climate data are now available at resolution of 30 seconds for the entire globe, for example from the WorldClim database (Hijmans et al. 2005). Here, it is important to note that the climate data were not sampled at a 30-second resolution but were interpolated from available weather stations. The precision of the data thus relies on the density of weather stations, which varies considerably across the globe. Additionally, these datasets differ in their baseline (e.g. the time period used to average climate baseline values) and the downscaling methods used for interpolation. These uncertainties from different data sources could be accounted for by using ensemble approaches (Araújo and New 2007). Ensembles combine multiple alternative model predictions, which may arise from differences in the input data, the modelling algorithms, or the future scenarios. Instead of choosing one best model setup, ensembles provide information on the expected variability in model predictions additional to the main trend.

Choosing appropriate environmental variables to describe species' niches is non-trivial and could also constitute a considerable source of uncertainty. As we have discussed in section 6.2, it is generally advisable to use direct environmental variables that are limiting the species' distribution. For example, temperature could have a direct, physiological effect on the species' niche. However, climate data provide a range of different temperature variables such as mean annual temperature, temperature seasonality, or minimum and maximum temperatures of the coldest and warmest months, respectively. Previous knowledge of

the species' ecology can help guide the variable choice. Another possibility is to separately test the explanatory power of each variable and choose those that best explain the distribution of the species (and are not too highly correlated). Also, uncertainty in model predictions due to the choice of predictor variables could be accounted for in an ensemble framework.

Differences in spatial and temporal coverage of environmental variables may lead to scaling issues as mentioned in section 6.3. For successfully building ENMs, all environmental data layers must fulfil the underlying assumptions and fit the spatial and temporal grain and extent of the species data. For example, serious temporal scale mismatches could arise if the species' presence records cover a longer time period, including records from before and after a major disturbance such as a hurricane, while the land cover data only cover the time period after the hurricane. Climate data are usually averaged over longer time periods. Still it is important to check for potential temporal scale mismatches between the species distribution and climatic data.

6.4.2 Different modelling approaches

A multitude of models have been developed for constructing ENMs, ranging from parametric to machine learning approaches (Franklin 2010; Guisan et al. 2017). All of these approaches make specific assumptions

about the species–environment relationship that can lead to inconsistent projections under climate change. Many non-parametric or machine-learning approaches such as MaxEnt, random forests, and boosted regression trees can fit very complex response surfaces and potentially overfit data, while (semi-)parametric methods fit simpler response surfaces and potentially underfit data (at least when fitted with lower order polynomials and derivatives). The necessary complexity should be determined by the research question and previous knowledge on the study system. If the species is not at equilibrium with its environment, then overly complex responses may lead to spurious predictions (Elith et al. 2010). At the same time, overly simplistic responses may lead to models that are not much better than the null expectation. The use of MaxEnt has been increasing rapidly over the last years. However, we want to stress that there is no evidence that MaxEnt is generally performing better than other approaches. Rather, its popularity may be related to the fact that a stand-alone user interface is available for MaxEnt, which makes it easy to use even without extensive statistical and programming experience. Generally, it is advisable to compare multiple model techniques to account for uncertainty in algorithmic choices (IPBES 2016).

For climate change projections, it is indispensable to check for novel environmental predictions in the climate change scenario (Figure 6.5) and to check

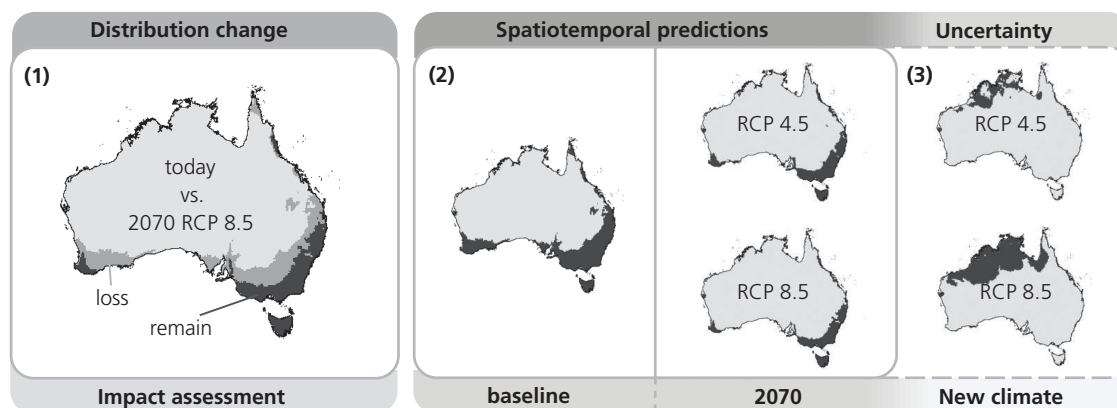


Figure 6.5 Maps of potential future climate impact on the silvereye. (1) The map of projected range changes indicates substantial loss in occupied area until 2070. (2) Distribution change can be easily derived by comparing potential distribution under current (baseline) climate and projected distribution under future climate scenarios (Representative Concentration Pathways, [RCP]). (3) One source of uncertainty in future projections of species distribution could be the emergence of novel climatic conditions. In Australia, novel climatic conditions will mainly occur in the north where the silvereye does not occur.

the extrapolation behaviour of the fitted ENM (cf. section 6.3) (Elith et al. 2010). For example, response surfaces that imply that a species could survive for any range of future temperatures should be judged as implausible (Elith et al. 2010; Zurell et al. 2012a). It is important to note that uncertainty through different modelling approaches cannot be studied in isolation from uncertainty in data sources. Problems with extrapolation will mainly occur in cases where the distribution data do not cover the entire range of environments experienced by the species, which could lead to truncated niches in the data (Zurell et al. 2012a).

Uncertainty due to different modelling approaches can be accounted for by combining these in an ensemble framework (Araújo and New 2007). However, we want to stress that even in ensembles it is of utmost importance to understand the peculiarities of the different ENM algorithms and check their extrapolation behaviour.

6.4.3 Different climate models and scenarios

A prerequisite for projecting species' distributions into the future are information about future climates that are provided by climate scenarios. Climate models simulate the complex interactions between atmosphere, ocean, land surface, snow and ice, ecosystem, and several chemical and biological processes. Naturally, many assumptions are being made in the modelling processes, and currently many different models exist ranging from very simple to more comprehensive climate models to Earth System Models (ESMs) that include an interactive carbon cycle. No single best climate model can be identified but rather the available models differ in their ability to simulate certain climate attributes (IPCC 2013). For simulating future climate, these models additionally rely on socio-economic scenarios about future anthropogenic forcings, including emissions of greenhouse gases and aerosols, which depend on socioeconomic factors and geopolitical agreements. The newest climate simulations currently available have been carried out within the Coupled Model Intercomparison Project Phase 5 (CMIP5) of the World Climate Research Programme. Under this framework, a large number of comprehensive climate models and ESMs have been compared.

Climate simulations from CMIP5 are based on a new set of scenarios of anthropogenic forcings, called the Representative Concentration Pathways (RCPs). Four different RCPs represent different mitigation scenarios and different targets of radiative forcing by 2100.

The IPCC (2013) provides details about the key uncertainties in these climate scenarios. In summary, uncertainty in climate models and scenarios may arise through uncertainty in process representation in the different climate models, spatial and temporal scales, uncertainty in observational data used for model evaluation, and uncertainty in other factors such as anthropogenic forcings. When projecting species' distributions into the future, uncertainty in climate change scenarios should be taken into account, for example by exploring a number of plausible climate models and socio-economic scenarios.

6.5 Robustness and validation against observed distribution and population changes

In the previous sections, we have learned that building and applying ecological niche models requires critical consideration of the necessary modelling steps and of the data used to build and project models. Yet, careful construction and evaluation of the models does not guarantee high transferability in space and time. As ecological niche models are phenomenological, they are strictly not valid beyond the range of variables they are calibrated on. The emergence of novel environments could thus seriously hamper prediction accuracy. Also, several of the assumptions underlying ecological niche models may limit their use in a global change context. Specifically, ENMs typically assume that: 1) species are at equilibrium with the environment, 2) interspecific interactions do not affect the distribution of the species, 3) the selected environmental variables are adequate for describing the niche of species, and 4) the species–environment relationship will stay constant under future climates.

The equilibrium assumption could be violated in two ways. First, if the distribution of a species and, thus, the observed data are affected by transient

(time-delayed) dynamics such as dispersal limitations and adaptation, it may lead to biased niche estimates in ENMs. This could be the case if the species is currently expanding its range, for example during invasion or due to competitor release. Second, when making projections under climate change ENMs assume an instantaneous realization of a new equilibrium situation. This would mean that newly suitable habitat will be colonized immediately without any time lags due to dispersal limitations or due to competition. Projections of historical ranges paired with DNA analyses in willow and rock ptarmigan suggested that these cold-adapted species were able to track changes in their suitable habitat over the last 20 millennia (Lagerholm et al. 2017). However, the authors caution that increasing current-day fragmentation may hamper future range shifts.

The question whether the species' niche is adequately described by the selected environmental variables is extremely difficult to evaluate. Distinguishing between proximal and distal effects on birds is often non-trivial. Birds are highly mobile organisms that often show seasonal variations in their range boundaries, or even spatially disjunct breeding and overwintering ranges as is the case in many migrants. These seasonal variations are likely driven by variations in seasonal resource availability closely intertwined with high climate seasonality (Eyres et al. 2017). Additionally, birds are endothermic and their distributions may be less tightly bounded by climatic conditions than in ectotherms (for an overview in birds see Engler et al. 2017, and references therein). Recent analyses indicate that climate variables may both have a direct effect on species' ranges, indicative of thermal tolerances (Khaliq et al. 2017), and an indirect effect due to their correlation with resource availability (Buckley et al. 2012). One impressive example is the massive northward shift in winter distributions of many North American bird species through supplemental feeding (Newton 2003). This example indicates that the winter climate is not physiologically (i.e. proximally) limiting northern range edges in these species but is correlated with resource limitations in winter. Similarly, range edges along upper thermal limits can be relaxed through artificial water supplements or behavioural adaptations in arid environments,

allowing birds to compensate for overheating and thermal stress (Pattinson and Smit 2017). However, increased desertification as a consequence of climate change is expected to exacerbate upper thermal limitations in birds, even in desert-adapted species (Albright et al. 2017).

Numerous studies have projected bird species' response to climate change (cf. Engler et al. 2017), but few have confronted projections with observations to test if projections were successful. In a cross-taxon analysis forecasting historic range changes of tetrapods over the last 60 years in Australia, Morán-Ordóñez et al. (2017) found ENMs to produce useful predictions. Similarly, Fordham et al. (2018) predicted range changes of British birds over the last 40 years and found that ENMs can provide reasonable first approximations of the magnitude of potential range shifts. Still more complex models, taking into account information on dispersal and demography, were better at predicting range changes at the grid-cell level. Although not explicitly focusing on climate change, Oliver et al. (2012) showed that ENM predictions were more strongly correlated with occupancy and population density than with population stability, suggesting that ENMs might be poor at predicting long-term population persistence.

6.6 Latest developments and perspectives in a global change context

ENMs have greatly enhanced our understanding of avian species–environment relationships and potential climate change impacts on birds. Nevertheless, because they do not explicitly take into account the effect of biotic interactions, demography, and adaptation, ENMs are of limited use for understanding the interplay of different niche components. Overall, we are still in the infancy of having a functional and holistic understanding of species–environment relationships. Therefore, it is not surprising that this is an active field with many exciting results and much debate.

In recent years, several approaches have been discussed to establish ENM frameworks that integrate demographic processes and interspecific interactions (Zurell 2017). Such a framework would help in disentangling the different niche constraints illustrated

in Figure 6.1. Demographic models help to better understand the complex dependencies between species' life history, their distribution, and environmental processes. To date, few studies have integrated demographic information for projecting bird species' response to climate change (Fordham et al. 2018; Zurell et al. 2012b). In part, this may be due to the higher data requirements needed for parameterizing the demographic model. At the same time, technical challenges remain. Mostly, hybrid models that have been used still rely on ENM predictions and couple these with a population dynamic model. Although these models are able to predict transient dynamics into the future, the hybrid approach could lead to circularity problems. If the current distribution of the species is affected by any transient dynamics such as dispersal constraints, then this is implicitly included in the underlying ENM and the predictions would be biased that way. So-called dynamic range models provide an alternative approach by directly relating demographic rates to environmental gradients and simultaneously fitting all model components (demography, dispersal, and environmental response) to data (Pagel and Schurr 2012). However, these models have not been applied to real-world species yet, and the underlying demographic model is as yet too inflexible to incorporate the complex life histories of birds.

Joint species distribution models (JSDMs) have been introduced as extensions to single-species ENMs. These models simultaneously model the species–environment relationships of multiple species as well as the residual correlation between those species. The residual correlations can be indicative of interspecific interactions but also of missing environmental information. Thus, JSDMs may help in deriving hypotheses about potential range-limiting interactions between species, but they are also highly scale-dependent (Zurell 2017).

Another branch of progress towards a more functional understanding of species–environment relationships is the integration of molecular genetic information below the species level. While ENMs have frequently been combined with phylogenetic data to understand niche evolution, integrating population genetic or genomic information into ENMs is a more recent innovation (see Engler et al.

2017 for an overview and references). Incorporating the information on population genetic structure into ENMs would allow capture of local adaptations that may be particularly important for anticipating species' response to climate change (Valladares et al. 2014). As genomic tools are getting more sophisticated and allow studying gene regulation and expression for many bird species (Kraus and Wink 2015), it will soon be possible to apply ENMs to study genotype–environment relationships in birds. In combination with enhanced ENM frameworks incorporating demography and biotic interactions, this will open up new perspectives for studying climate change impacts on birds at a functional-molecular level and for understanding adaptive responses to sudden environmental changes.

References

- Albright, T.P., Mutiibwa, D., Gerson, A.R., Smith, E.K., Talbot, W.A., O'Neill, J.J., McKechnie, A.E., et al. (2017). Mapping evaporative water loss in desert passerines reveals an expanding threat of lethal dehydration. *Proceedings of the National Academy of Sciences of the United States of America*, 114, 2283–8.
- Araújo, M.B., and New, M. (2007). Ensemble forecasting of species distributions. *Trends in Ecology & Evolution*, 22, 42–7.
- Austin, M.P. (1980). Searching for a model for use in vegetation analysis. *Vegetatio*, 42, 11–21.
- Blakers, M., Reilly, P.N., and Davies, S. (1984). *The atlas of Australian birds*. Melbourne Univ. Press, Melbourne, Australia.
- Buckley, L.B., Hurlbert, A.H., and Jetz, W. (2012). Broad-scale ecological implications of ectothermy and endothermy in changing environments. *Global Ecology and Biogeography*, 21, 873–85.
- Elith, J., Kearney, M., and Phillips, S. (2010). The art of modelling range-shifting species. *Methods in Ecology and Evolution*, 1, 330–42.
- Elith, J., and Leathwick, J.R. (2009). Species distribution models: Ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics*, 40, 677–97.
- Engler, J.O., Stiels, D., Schidelko, K., Strubbe, D., Quillfeldt, P., and Brambilla, M. (2017). Avian SDMs: current state, challenges, and opportunities. *Journal of Avian Biology*, 48, 1483–1504.
- Eyres, A., Böhning-Gaese, K., and Fritz, S.A. (2017). Quantification of climatic niches in birds: adding the

- temporal dimension. *Journal of Avian Biology*, 48, 1517–31.
- Fordham, D.A., Bertelsmeier, C., Brook, B.W., Early, R., Neto, D., Brown, S.C., Ollier, S., et al. (2018). How complex should models be? Comparing correlative and mechanistic range dynamics models. *Global Change Biology*, 24, 1357–70.
- Fourcade, Y., Engler, J.O., Rödder, D., and Secondi, J. (2014). Mapping species distributions with MAXENT using a geographically biased sample of presence data: a performance assessment of methods for correcting sampling bias. *PLoS One*, 9, e97122.
- Franklin, J. (2010). *Mapping species distributions: Spatial inference and prediction*. Cambridge University Press, Cambridge, UK.
- Graham, C.H., Elith, J., Hijmans, R.J., Guisan, A., Townsend Peterson, A., Loiselle, B.A., and The NCEAS Predicting Species Distributions Working Group (2008). The influence of spatial errors in species occurrence data used in distribution models. *Journal of Applied Ecology*, 45, 239–47.
- Guillera-Arroita, G. (2017). Modelling of species distributions, range dynamics, and communities under imperfect detection: advances, challenges, and opportunities. *Ecography*, 40, 281–95.
- Guisan, A., and Thuiller, W. (2005). Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, 8, 993–1009.
- Guisan, A., Thuiller, W., and Zimmermann, N.E. (2017). *Habitat suitability and distribution models: with applications in R*. Cambridge University Press, Cambridge, UK.
- Guisan, A., and Zimmermann, N. E. (2000). Predictive habitat distribution models in ecology. *Ecological Modelling*, 135, 147–86.
- Hanski, I. (1999). *Metapopulation ecology*. Oxford University Press, Oxford, UK.
- He, K.S., Bradley, B.A., Cord, A.F., Rocchini, D., Tuanmu, M.-N., Schmidtlein, S., Turner, W., et al. (2015). Will remote sensing shape the next generation of species distribution models? *Remote Sensing in Ecology and Conservation*, 1, 4–18.
- Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., and Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25, 1965–78.
- Hurlbert, A.H., and Jetz, W. (2007). Species richness, hotspots, and the scale dependence of range maps in ecology and conservation. *Proceedings of the National Academy of Sciences of the United States of America*, 104, 13384–9.
- IPBES (2016). *The Methodological Assessment Report on Scenarios and Models of Biodiversity and Ecosystem Services*. Secretariat of the Intergovernmental Platform for Biodiversity and Ecosystem Services, Bonn, Germany.
- IPCC (2013). *Climate Change 2013: The Physical Science Basis: Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, UK.
- Jarvis, E.D. (2016). Perspectives from the Avian Phylogenomics Project: Questions that can be answered with sequencing all genomes of a vertebrate class. *Annual Review of Animal Biosciences*, 4, 45–59.
- Khalid, I., Böhning-Gaese, K., Prinzing, R., Pfenninger, M., and Hof, C. (2017). The influence of thermal tolerances on geographical ranges of endotherms. *Global Ecology and Biogeography*, 26, 650–68.
- Kraus, R.H.S., and Wink, M. (2015). Avian genomics: fledging into the wild! *Journal of Ornithology*, 156, 851–65.
- Lagerholm, V.K., Sandoval-Castellanos, E., Vaniscotte, A., Potapova, O.R., Tomek, T., Bochenki, Z.M., Shepherd, P., et al. (2017). Range shifts or extinction? Ancient DNA and distribution modelling reveal past and future responses to climate warming in cold-adapted birds. *Global Change Biology*, 23, 1425–35.
- Meyer, C., Kreft, H., Guralnick, R., and Jetz, W. (2015). Global priorities for an effective information basis of biodiversity distributions. *Nature Communications*, 6, 8221.
- Morán-Ordóñez, A., Lahoz-Monfort, J.J., Elith, J., and Wintle, B.A. (2017). Evaluating 318 continental-scale species distribution models over a 60-year prediction horizon: what factors influence the reliability of predictions? *Global Ecology and Biogeography*, 26, 371–84.
- Newton, I. (2003). *Speciation and Biogeography of Birds*. Academic Press, London, UK.
- Oliver, T.H., Gillings, S., Girardello, M., Rapacciuolo, G., Brereton, T.M., Siriwardena, G. M., Roy, D.B., et al. (2012). Population density but not stability can be predicted from species distribution models. *Journal of Applied Ecology*, 49, 581–90.
- Pagel, J., and Schurr, F.M. (2012). Forecasting species ranges by statistical estimation of ecological niches and spatial population dynamics. *Global Ecology and Biogeography*, 21, 293–304.
- Pattinson, N.B., and Smit, B. (2017). Seasonal behavioral responses of an arid-zone passerine in a hot environment. *Physiology and Behavior*, 179, 268–75.
- Pearson, R.G., and 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, 361–71.
- Pulliam, H.R. (1988). Sources, sinks, and population regulation. *American Naturalist*, 132, 652–61.
- Soberón, J. (2007). Grinnellian and Eltonian niches and geographic distributions of species. *Ecology Letters*, 10, 1115–23.

- Valladares, F., Matesanz, S., Guilhaumon, F., Araújo, M.B., Balaguer, L., Benito-Garzón, M., Cornwell, W., et al. (2014). The effects of phenotypic plasticity and local adaptation on forecasts of species range shifts under climate change. *Ecology Letters*, 17, 1351–64.
- Zurell, D. (2017). Integrating demography, dispersal and interspecific interactions into bird distribution models. *Journal of Avian Biology*, 48, 1505–16.
- Zurell, D., Elith, J., and Schröder, B. (2012a). Predicting to new environments: tools for visualizing model behaviour and impacts on mapped distributions. *Diversity and Distributions*, 18, 628–34.
- Zurell, D., Grimm, V., Rossmannith, E., Zbinden, N., Zimmermann, N.E., and Schröder, B. (2012b). Uncertainty in predictions of range dynamics: black grouse climbing the Swiss Alps. *Ecography*, 35, 590–603.