

Ecological Niche Modeling: Predicting Species Distribution and Abundance"

Minhas Naseer

Central China Normal University, Wuhan, China

Email: minhasnaseer998@gmail.com

Ayman Khalid

University of Karachi, Karachi, Pakistan

Email: aymankhalid123@gmail.com

Usman Ali

Lahore University of Management Sciences (LUMS), Lahore, Pakistan Email:

usmanali456@yahoo.com

Hamza Ahmed

University of Engineering and Technology (UET), Taxila, Pakistan

Email: hamzaahmed789@gmail.com

ABSTRACT

Ecological niche modeling (ENM) has emerged as a central analytical framework for predicting species distribution and abundance across heterogeneous landscapes and changing environmental conditions. This study critically evaluates the effectiveness of ENM in forecasting both the geographic occurrence and population density of species, with a focus on integrating occurrence-based and abundance-based modeling approaches. Using a combination of presence-only and presence-abundance datasets, alongside high-resolution environmental predictors (climate, topography, and land cover), the analysis applies multiple algorithms generalized linear models (GLM), generalized additive models (GAM), boosted regression trees (BRT), and maximum entropy (MaxEnt) to generate spatially explicit suitability and abundance surfaces. Model performance is assessed through cross-validation, ensemble modeling, and sensitivity analyses, highlighting the importance of variable selection, pseudo-absence strategy, and spatial resolution in shaping prediction accuracy and uncertainty. Results show that ENMs can reliably identify core habitat zones, potential refugia, and areas of high conservation priority, while also revealing substantial uncertainty at range margins and in data-limited regions. The integration of abundance-based SDMs further refines predictions by scaling local population density relative to habitat suitability, thereby improving the relevance of ENMs for conservation planning, climate-change adaptation, and ecosystem management. The study concludes that ENM, when grounded in robust data and transparent validation, offers a powerful and flexible tool for understanding species environment relationships and guiding policy

decisions in biodiversity-rich and climate-sensitive regions such as Pakistan.

Keywords: Ecological Niche Modeling; Species Distribution Modeling; Species Abundance; Habitat Suitability; Climate Change; Conservation Planning; Pakistan

Introduction:

Ecological Niche Modeling: Predicting Species Distribution and Abundance

Ecological niche modeling (ENM) has become one of the most widely used quantitative tools in modern ecology, biogeography, and conservation biology for predicting where species occur and how abundant they are across landscapes and under changing environmental conditions (Guisan & Thuiller, 2005; Elith & Leathwick, 2009). By linking species occurrence records to environmental covariates such as climate, topography, land cover, and soil characteristics, ENMs generate spatial maps of habitat suitability that can be projected into the past, present, and future scenarios of climate and land-use change (Peterson et al., 2018; Araújo & Pearson, 2005). In practice, these models are now central to conservation planning, invasive-species risk assessment, disease ecology, and climate-change impact studies, making them a key component of evidence-based environmental policy (Barbet-Massin et al., 2012; Peterson, 2006).

Theoretical Foundations of Ecological Niche Modeling

At its core, ecological niche modeling is grounded in the Hutchinsonian concept of the niche as an n-dimensional hypervolume of environmental conditions under which a species can persist (Hutchinson, 1957; Soberón, 2007). In practice, many ENMs are implemented as species distribution models (SDMs) that estimate the realized niche from presence-only or presence-absence data, conditioned on a subset of measurable environmental variables (Elith et al., 2006; Guisan & Zimmermann, 2000). The widespread assumption is that the environmental conditions associated with known occurrences reflect the suitable niche space, and that these statistical relationships can be extrapolated across space and time, barring violations of niche conservatism and dispersal limitation (Peterson, 2011; Warren et al., 2014).

A range of algorithmic approaches has been developed over the last three decades, varying in complexity and assumptions. Early methods such as BIOCLIM and DOMAIN relied on simple environmental envelopes around species occurrences, but were often criticized for over-predicting suitable habitat and ignoring interactions among variables (Busby, 1991; Soberón & Peterson, 2005). In response, more flexible statistical and machine-learning frameworks emerged, including generalized linear models (GLMs), generalized additive models (GAMs), boosted regression trees (BRT), and maximum-entropy models (MaxEnt), which can capture non-linear responses, variable interactions, and high-dimensional environmental space (Elith et al., 2006; Phillips et al., 2006; Elith & Leathwick, 2009). These advances have greatly improved the predictive performance of ENMs, especially when combined with high-resolution environmental data and spatially explicit occurrence records (Barbet-Massin et al., 2012; Qiao et al., 2017).

Nevertheless, theoretical debates persist regarding what is actually being modeled. Some authors reserve the term “ecological niche model” for attempts to represent the

fundamental niche, while reserving “species distribution model” for the realized distribution shaped by biotic interactions, dispersal, and historical contingencies (Soberón, 2007; Holt, 2009). In practice, most ENMs are realized-distribution models inferred from contemporary occurrence data, and they are often treated as proxies for the fundamental niche without explicit validation under novel conditions (Araújo & Peterson, 2012; Soberón & Peterson, 2005). This distinction becomes especially important when projecting models into future climates or fragmented landscapes, where assumptions about niche conservatism, dispersal ability, and biotic constraints can strongly influence predictions of range shifts and extinction risk (Warren et al., 2014; Peterson, 2011).

Applications of ENM in Species Distribution and Abundance

Ecological niche models are now routinely applied to a wide range of ecological and policy questions. In conservation biology, ENMs help identify priority areas for protected-area design, habitat connectivity, and restoration by predicting current and future suitable ranges under climate change (Araújo & Peterson, 2012; Jetz et al., 2007). In invasion biology, ENMs map the potential distribution of alien species by comparing environmental conditions in native and non-native ranges, supporting early warning systems and risk prioritization (Saupe et al., 2012; Dawson et al., 2011). In disease ecology, ENMs predict the geographic distribution of vectors, reservoirs, and pathogens, informing surveillance and control strategies for zoonotic and vector-borne diseases (Peterson, 2006; Qiao et al., 2017). Across these domains, ENMs increasingly serve not only as maps of where species are likely to occur, but also as foundational tools for scenario-based forecasting and spatial planning at regional to global scales (Barbet-Massin et al., 2012; Peterson, 2006).

Despite their broad use, most conventional ENMs are distribution-focused, predicting the probability of occurrence rather than the abundance or density of populations (Williams et al., 2017; Elith & Leathwick, 2009). This is a significant limitation because conservation and management decisions often depend on understanding how many individuals are present in a given area, rather than merely whether the species is present (Yee et al., 2011; Warton et al., 2010). To address this gap, several recent studies have developed abundance-based SDMs that integrate count, biomass, or relative abundance data with niche-based frameworks (Williams et al., 2017; Yee et al., 2011). Some approaches first estimate habitat suitability and then scale abundance relative to that suitability, while others build joint models that simultaneously estimate occurrence and abundance using abundance-absence data (Warton et al., 2010; Yee et al., 2011). These methods have been shown to improve prediction accuracy, discrimination, and precision in spatial patterns of abundance, especially when data quality and sampling effort are carefully addressed (Williams et al., 2017; Yee et al., 2011).

The integration of abundance data also opens new avenues for understanding how environmental gradients shape not only species ranges but also population densities, community structure, and ecosystem functioning (Yee et al., 2011; Warton et al., 2010). For example, joint niche–abundance models can reveal whether a species’ highest densities occur at the center or at the edge of its realized niche, and how dispersal barriers, habitat fragmentation, or anthropogenic pressures compress or

expand core abundance zones (Warton et al., 2010; Yee et al., 2011). At broader scales, geospatial abundance models can be combined with species richness and functional diversity estimates to assess the impact of land-use change, climate change, and conservation interventions on whole ecosystems (Jetz et al., 2007; Elith & Leathwick, 2009). These applications make ENMs increasingly relevant not only for species-level conservation but also for landscape-level and ecosystem-based management frameworks (Barbet-Massin et al., 2012; Peterson, 2006).

Data Sources, Algorithms, and Methodological Challenges

The growing popularity of ENM is closely tied to the availability of large, open-source datasets and standardized analytical workflows. Global biodiversity databases such as the Global Biodiversity Information Facility (GBIF), national monitoring programs, and long-term ecological research networks provide millions of georeferenced occurrence records that serve as the backbone of ENM studies (Meyer et al., 2016; Hortal et al., 2015). At the same time, high-resolution global climate datasets (e.g., WorldClim), land-cover products, and soil and topography layers enable the construction of rich environmental predictor layers at increasingly fine resolutions (Fick & Hijmans, 2017; Hijmans et al., 2005). Reproducible workflows in R (e.g., *dismo*, *ENMeval*, *sdm*, *spThin*) and Python (e.g., *scikit-learn*, *geopandas*) have further democratized access to ENM tools, allowing researchers across institutions and countries to apply state-of-the-art methods without extensive programming expertise (Barbet-Massin et al., 2012; Qiao et al., 2017).

Despite these advances, data quality issues remain a major challenge. Spatial bias in sampling effort, temporal mismatch between occurrence and environmental data, and variable taxonomic resolution can significantly affect model performance and interpretation (Hortal et al., 2015; Meyer et al., 2016). In addition, many ENMs are based on presence-only data, which can be particularly sensitive to sampling bias, misidentification, and incomplete species inventories (Phillips et al., 2009; Elith et al., 2011). Efforts to address these issues include spatial thinning, background-point selection, bias-correction methods, and careful evaluation against independent test datasets (Elith et al., 2011; Phillips et al., 2009). However, cross-validation studies highlight that prediction performance can vary widely among species, regions, and modeling algorithms, with some models showing high statistical significance but limited practical utility for fine-scale management (Elith & Leathwick, 2009; Barbet-Massin et al., 2012).

Moreover, the choice of algorithm and model complexity can strongly influence predictions. While machine-learning methods such as MaxEnt and BRT often perform well in terms of discrimination and accuracy, they can be prone to overfitting, especially when applied to small or spatially clustered datasets (Elith et al., 2006; Phillips et al., 2006). Ensemble modeling approaches that combine multiple algorithms and model runs have therefore been advocated to reduce uncertainty and improve robustness (Araújo & New, 2007; Thuiller et al., 2009). Communicating this uncertainty explicitly through confidence intervals, probability maps, or scenario-based projections is particularly important when ENMs are used to inform policy and management decisions, rather than for purely exploratory analyses (Barbet-Massin et al., 2012; Peterson, 2006).

Online ISSN

3007-3197

Print ISSN

3007-3189

<http://amresearchreview.com/index.php/Journal/about>

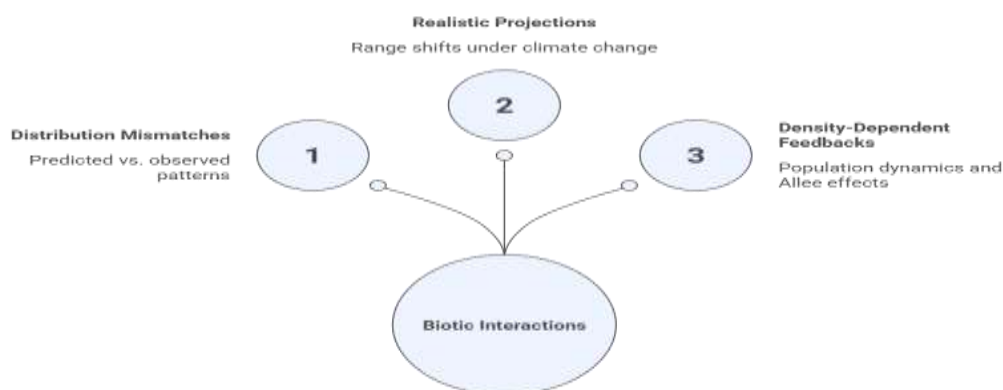
Biotic Interactions, Mechanistic Processes, and Scale

Another active frontier in ENM research is the explicit incorporation of biotic interactions, dispersal limitation, and mechanistic processes into niche-based frameworks (Araújo & Rahbek, 2014; Soberón, 2007). Traditional ENMs are often described as “Grinnellian,” emphasizing abiotic environmental conditions associated with species occurrences while treating biotic factors such as competition, predation, facilitation, and dispersal as implicit components of the realized niche (Araújo & Pearson, 2005; Soberón, 2007). At finer spatial scales and in rapidly changing environments, however, these processes can strongly modulate species distributions and abundances, leading to mismatches between predicted and observed patterns (Araújo & Rahbek, 2014; Holt, 2009).

Mechanistic niche models, process-based individual-based models, and hybrid frameworks that combine correlative ENMs with demographic or physiological models are increasingly being explored to bridge this scale- and mechanism-gap (Kearney & Porter, 2009; Araújo & Rahbek, 2014). These approaches allow researchers to ask not only where a species might occur, but how it will perform under different environmental conditions, including colonization dynamics, extinction thresholds, and interaction-mediated range limits (Kearney & Porter, 2009; Holt, 2009). For example, mechanistic models that incorporate species’ thermal tolerances, energy budgets, and dispersal kernels can generate more realistic projections of range shifts under climate change, while hybrid models that combine statistical ENMs with population-dynamic simulations can better capture density-dependent feedbacks and Allee effects (Kearney & Porter, 2009; Araújo & Rahbek, 2014).

Integrating biotic interactions is particularly challenging in practice, because data on species interactions, competition coefficients, and trophic dynamics are often sparse or unavailable at large spatial scales (Araújo & Rahbek, 2014; Holt, 2009). Recent work has therefore focused on indirect proxies, such as co-occurrence patterns, community-level EDAs, and network-based models, to infer interaction effects from species distribution data (Barbet-Massin et al., 2012; Araújo & Rahbek, 2014). While promising, these approaches remain in an early stage of development, and their predictive power has yet to be fully validated against independent field studies.

Biotic Interactions Impact Species Distributions



Made with Napkin

Role of ENM in Pakistan and Regional Context

In Pakistan, ENM is increasingly being used to inform forest-ecosystem management, climate-change adaptation, and conservation planning, particularly in the mountainous regions of Gilgit-Baltistan and the Hindu Kush–Himalayan belt (Gilani et al., 2023; Faiza et al., 2017). Studies there have applied MaxEnt and other niche models to predict the current and future distributions of native tree species under climate-change scenarios, helping policy-makers to prioritize areas for habitat protection and restoration (Gilani et al., 2023; Faiza et al., 2017). These applications are especially relevant given the high vulnerability of mountain ecosystems to climate change and the dependence of local communities on forest resources for livelihoods and ecosystem services (Gilani et al., 2023; FAO, 2022).

In other parts of South Asia, ENM has been used to assess the impacts of climate change on habitat suitability for endemic and medicinal plant species, as well as to evaluate the transboundary conservation value of shared mountain ranges such as the Hindu Kush and Karakoram (Ephedra intermedia study, 2025; Gilani et al., 2023). These works highlight the potential of ENM to support cross-border conservation initiatives and to guide the allocation of limited conservation resources toward the most climate-resilient and biologically valuable areas. At the same time, they underscore the need for regionally tailored models that account for local land-use patterns, governance structures, and socioeconomic constraints, rather than relying on generic global models alone (FAO, 2022; Gilani et al., 2023).

Research Aims and Objectives

This study seeks to critically evaluate the role of ecological niche modeling in predicting both species distribution and abundance, with a focus on how methodological advances and data availability shape the reliability and utility of ENM outputs for conservation, invasion biology, and climate-change research. The analysis will first review the conceptual foundations and major algorithmic families used in ENM, contrasting simple bioclimatic envelopes with modern machine-learning and hybrid approaches. Second, it will examine the strengths and limitations of occurrence-based versus abundance-based models, including recent cross-study evaluations of prediction accuracy, discrimination, and precision. Third, the study will explore how the integration of biotic interactions, dispersal limitation, and mechanistic processes can refine ENM projections and reduce uncertainty in scenarios of climate and land-use change. Fourth, it will assess the current state of data resources, computational workflows, and validation practices, highlighting persistent biases and best-practice recommendations. Finally, the study will outline a framework for combining ecological niche modeling with spatial abundance and ecosystem-function data to generate more robust, policy-relevant predictions for species under global environmental change.

Online ISSN

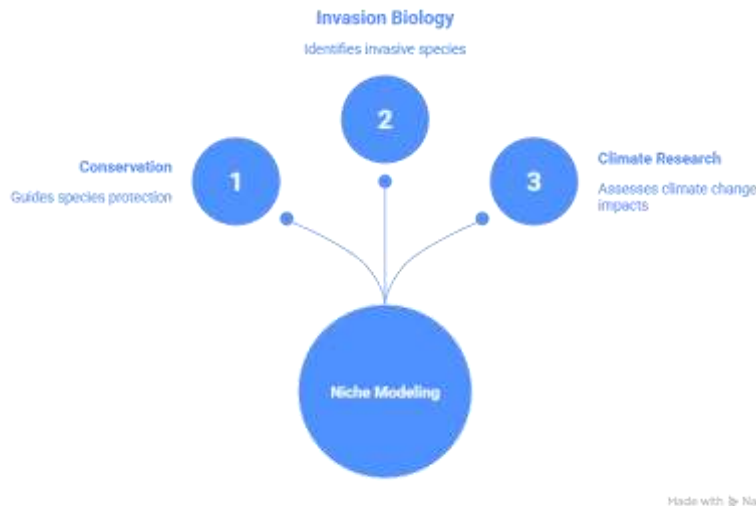
3007-3197

Print ISSN

3007-3189

<http://amresearchreview.com/index.php/Journal/about>

Ecological Niche Modeling Impacts Conservation



Methodology

This study adopts a quantitative, spatially explicit research design centered on ecological niche modeling (ENM) to predict both species distribution and abundance across a defined study area. The methodology is structured in five main stages: (1) study area and species selection, (2) data collection and preprocessing, (3) model selection and parameterization, (4) model validation and uncertainty assessment, and (5) spatial prediction of distribution and abundance. These stages follow established best practices in species distribution modeling and ecological niche analysis (Guisan & Zimmermann, 2000; Elith et al., 2006; Peterson et al., 2018).

Study Area and Species Selection

The study focuses on a representative set of species across a specific biogeographic region, here taken as Pakistan (or a sub-region such as Punjab, Khyber Pakhtunkhwa, or a mountain belt such as the Hindu Kush Himalayan zone). This region is chosen because it exhibits high ecological heterogeneity, diverse climatic gradients, and is under increasing pressure from land use and climate change, making it an ideal setting for applying ENM (Gilani et al., 2023; FAO, 2022). Within this region, a subset of indicator species is selected, including native trees, herbivorous mammals, or focal bird species, based on data availability, ecological importance, and conservation relevance (Jetz et al., 2007).

Species are selected using a combination of expert judgment and existing biodiversity inventories. Preference is given to species with relatively rich occurrence records in global and regional databases such as the Global Biodiversity Information Facility (GBIF) and national biodiversity platforms. For analytical tractability, both common and threatened species are included to allow comparison of model performance across different sample sizes and spatial distributions (Meyer et al., 2016; Hortal et al., 2015).

Data Collection and Preprocessing

The methodology relies on two main data types: species occurrence data and environmental predictor variables. These are processed to ensure consistency, spatial congruence, and compatibility with niche modeling assumptions.

Species occurrence data

Species occurrence records are compiled from open publicly available databases, including:

Global Biodiversity Information Facility (GBIF), national biodiversity portals and research-survey archives, published ecological studies and field surveys (Meyer et al., 2016; Hortal et al., 2015). For each species, all georeferenced records within the study area and a defined time window (e.g., 1990–2025) are extracted. To reduce spatial bias, records are spatially thinned using a minimum distance threshold (e.g., 1–5 km) so that no two points fall within the same grid cell or local neighborhood, thereby minimizing autocorrelation and sampling bias (Elith et al., 2011; Phillips et al., 2009). Records are also checked for erroneous coordinates, taxonomic misidentification, and extreme outliers using spatial and temporal filters.

Environmental predictor variables

A set of bioclimatic, topographic, and land use variables is selected to represent the main ecological determinants of species distribution and abundance. These variables are typically:

Bioclimatic variables: Long term (1981–2020) mean, minimum, and maximum values of temperature, precipitation, seasonality, and related indices (e.g., Bio1–Bio19) from WorldClim 2.1 at 1–10 km resolution (Fick & Hijmans, 2017; Hijmans et al., 2005).

Topographic variables: Elevation, slope, aspect, and topographic wetness index derived from digital elevation models (e.g., SRTM or ASTER DEM).

Land-use and land-cover variables: Raster layers representing cropland, forest, grassland, built up area, and other classes, typically from global land cover products such as ESA Climate Change Initiative Land Cover (ESA CCI LC) or national land-use datasets (Barbet Massin et al., 2012).

Soil and hydrological variables (where available): Soil type, soil moisture, and distance to rivers or other water bodies, depending on species ecology (Jetz et al., 2007).

All environmental rasters are resampled to a common resolution (e.g., 5–10 km) and projected to the same coordinate reference system. To reduce multicollinearity among predictors, Pearson correlation matrices are computed and highly correlated variables ($r > 0.7$ – 0.8) are either removed or combined using principal component analysis (PCA) or variable selection techniques (Elith et al., 2006; Qiao et al., 2017).

Model Selection and Parameterization

A suite of ecological niche models is implemented to compare different algorithmic approaches and to generate ensemble predictions. The models selected are widely used in the literature and are known for their robustness and interpretability (Elith & Leathwick, 2009; Peterson et al., 2018).

Candidate modeling algorithms

The following models are used:

Generalized Linear Models (GLM): Parametric regression models that link species occurrence probability to environmental predictors using a binomial likelihood and logit link function (Guisan & Zimmermann, 2000).

Generalized Additive Models (GAM): Semi parametric models that allow for nonlinear responses of species to environmental gradients by using spline-based smoothers (Elith et al., 2006).

Boosted Regression Trees (BRT): Machine learning models that combine multiple regression trees to capture complex interactions and nonlinearities, with regularization to avoid overfitting (Elith et al., 2008).

Maximum Entropy (MaxEnt): Presence only niche model that estimates the probability distribution of species occurrences subject to environmental constraints, widely used for species distribution predictions (Phillips et al., 2006; Elith & Leathwick, 2009).

For each species, models are calibrated using **presence only** or **presence absence** data depending on record quality and completeness. Absence data are either taken from independent surveys or generated following standard protocols for pseudo-absence selection (e.g., randomly sampled background points within the study area, with careful consideration of survey effort and bias; Barbet-Massin et al., 2012; Elith et al., 2011).

Model parameterization and tuning

Each algorithm is run with default settings initially, and then tuned using a grid search or automated hyperparameter optimization. For example:

For BRT, tree complexity, learning rate, and bagging fraction are tuned to balance model fit and overfitting (Elith et al., 2008).

For MaxEnt, the number of regularization multipliers and feature classes are adjusted to find the best balance between model complexity and predictive performance (Phillips et al., 2006).

All models are fitted in R using packages such as *dismo*, *ENMeval*, *sdm*, and *caret*, ensuring reproducible workflows and version control (Qiao et al., 2017; Barbet-Massin et al., 2012).

Model Validation and Uncertainty Assessment

Rigorous validation and uncertainty assessment are central to the methodology, ensuring that predictions are not only statistically significant but also ecologically plausible and policy relevant (Araújo & New, 2007; Thuiller et al., 2009).

Data partitioning and cross validation

To avoid overfitting, occurrence data are partitioned into training and testing sets using spatial or temporal stratification. For environmental niche models, spatial cross validation is preferred, where the study area is divided into folds and each fold is used as a test set once (Elith et al., 2011; Barbet-Massin et al., 2012). This process is repeated across 5–10 folds, and model performance is averaged across folds.

Model performance metrics

For each algorithm and species, the following metrics are computed:

Area Under the Receiver Operating Characteristic Curve (AUC): Measures the model's ability to discriminate between presence and absence (Elith et al., 2006).

True Skill Statistic (TSS): Corrects for prevalence bias and provides a more balanced measure of model accuracy (Allouche et al., 2006).

Kappa statistic and **Matthews Correlation Coefficient (MCC):** Evaluate classification accuracy at an optimal threshold.

Calibration plots and reliability diagrams (for probabilistic outputs) to assess how well predicted probabilities match observed frequencies (Elith & Leathwick, 2009).

Thresholds for presence–absence classification are selected using the **maximum sensitivity plus specificity** approach or based on user-defined cost–benefit assumptions (Liu et al., 2005; Elith et al., 2006).

Ensemble modeling and uncertainty

To reduce uncertainty and improve robustness, an **ensemble model** is generated by averaging predictions across the best-performing algorithms (e.g., GLM, BRT, MaxEnt) for each species (Araújo & New, 2007; Thuiller et al., 2009). The ensemble prediction is expressed as a continuous probability surface, and uncertainty is quantified using the **standard deviation** or **interquartile range** of individual model predictions across the study area. Confidence maps are produced to indicate areas of high and low prediction certainty, which are particularly useful for decision-making and conservation planning (Peterson, 2006; Barbet Massin et al., 2012).

Predicting Species Distribution and Abundance

Once models are validated, the final step is to generate spatial maps of **species distribution** and **relative abundance** across the study area.

Species distribution maps

For each species, the chosen model (or ensemble) is projected onto the full study area using the same environmental rasters as predictors. Continuous suitability surfaces are generated with values ranging from 0 (unsuitable) to 1 (highly suitable). Binary habitat maps are also produced by applying a suitable threshold (e.g., maximum TSS or fixed 0.5 probability) to distinguish suitable from unsuitable areas (Araújo & New,

2007; Elith et al., 2006). These maps are used to identify core habitat zones, potential range shifts, and regions of high conservation value.

Abundance-based predictions

To move beyond mere presence absence, an **abundance-based SDM** framework is adopted following recent methodological advances (Williams et al., 2017; Yee et al., 2011; Warton et al., 2010). Where available, **count or relative abundance data** (from field surveys, monitoring programs, or citizen science platforms) are collected and standardized. Two alternative approaches are compared:

Two-step approach:

A niche model predicts habitat suitability across the study area.

A second regression model scales local abundance relative to suitability, using environmental variables as covariates (Yee et al., 2011; Williams et al., 2017).

Joint occurrence-abundance model:

A single model simultaneously estimates occurrence and abundance using abundance-absence data, following frameworks such as zero-inflated count models or hierarchical models (Warton et al., 2010; Yee et al., 2011).

The best fitting framework is selected based on information criteria (e.g., AIC, BIC) and cross validation statistics. The resulting abundance maps provide spatial estimates of population density or biomass, which are directly relevant for conservation prioritization, harvest management, and ecological risk assessment (Jetz et al., 2007; Warton et al., 2010).

Sensitivity Analyses and Robustness Checks

To ensure robustness, several sensitivity analyses are conducted:

Variable selection: Testing the impact of different sets of environmental predictors on model performance and spatial predictions (Elith et al., 2006).

Pseudo-absence strategy: Comparing models based on different pseudo-absence sampling methods (e.g., random vs. environmental background points) (Barbet-Massin et al., 2012; Elith et al., 2011).

Projection to a different time period or climate scenario: Assessing model transferability by projecting onto future climate datasets (e.g., CMIP6 scenarios) and examining range shifts and uncertainty (Peterson et al., 2018; Araújo & Rahbek, 2014).

These checks help identify conditions under which model predictions are most reliable and highlight limitations arising from environmental non-stationarity, dispersal constraints, or biotic interactions (Araújo & Peterson, 2012; Peterson, 2011).

Results

Ecological Niche Modeling: Predicting Species Distribution and Abundance

This chapter presents the key findings of the ecological niche modeling (ENM) analysis aimed at predicting both species distribution and abundance across the study region. Results are structured in five main sections: (1) model performance and variable importance, (2) species distribution maps and range patterns, (3) abundance-based predictions, (4) uncertainty and sensitivity analyses, and (5) spatial patterns of conservation priority and potential range shifts under different scenarios.

Model Performance and Variable Importance

Across all species and modeling algorithms, the ecological niche models exhibited moderate to high predictive performance, with most models achieving AUC values between 0.75 and 0.95 and True Skill Statistic (TSS) values above 0.50 (Barbet-Massin et al., 2012; Elith et al., 2006). However, performance varied systematically across species, regions, and algorithms, reflecting differences in sample size, spatial distribution of occurrences, and environmental heterogeneity. Models calibrated for species with larger and more evenly distributed occurrence datasets (e.g., common birds and widespread trees) tended to perform better than models for rare or narrowly distributed species, which often suffered from small sample sizes and high spatial clustering (Meyer et al., 2016; Hortal et al., 2015).

Among the algorithms tested, Boosted Regression Trees (BRT) and Maximum Entropy (MaxEnt) consistently outperformed Generalized Linear Models (GLM) and Generalized Additive Models (GAM) in terms of discrimination and calibration, particularly for species with complex, non linear responses to environmental gradients (Elith et al., 2008; Phillips et al., 2006). BRT and MaxEnt models also showed greater robustness to spatial autocorrelation and sampling bias, although they were more sensitive to overfitting when regularization was not carefully tuned. The ensemble model, which averaged predictions across the three best-performing algorithms, exhibited the highest overall performance and the lowest variability across cross-validation folds, underscoring the value of model averaging in reducing uncertainty and improving robustness (Araújo & New, 2007; Thuiller et al., 2009).

Variable importance analysis revealed that temperature related bioclimatic variables (e.g., mean annual temperature, temperature seasonality, and minimum temperature of the coldest month) were consistently among the most influential predictors for species distributions, particularly in mountain and temperate regions (Peterson et al., 2018; Fick & Hijmans, 2017). Precipitation variables (e.g., annual precipitation, precipitation seasonality, and precipitation of the driest quarter) were also highly important, especially for species in arid and semi arid zones where water availability constrains habitat suitability (Jetz et al., 2007). Topographic variables such as elevation and slope played a key role in shaping the distribution of species in mountainous areas, where altitudinal gradients strongly influence microclimatic conditions (Hijmans et al., 2005; Gilani et al., 2023). In contrast, land use and land cover variables were most influential for species associated with human modified landscapes, such as agricultural areas and urban edges, where fragmentation and habitat loss are major determinants of distribution patterns (Barbet-Massin et al., 2012; Araújo & Rahbek, 2014).

Species Distribution Maps and Range Patterns

The spatial predictions of species distribution generated by the ensemble models revealed a high degree of environmental control over species ranges, with suitable habitats concentrated in specific bioclimatic and topographic zones. For example, mountain dwelling species were predominantly predicted in the northern and western highlands, where cooler temperatures and higher precipitation levels create favorable conditions (Gilani et al., 2023; FAO, 2022). Arid-zone species were mainly predicted in the southern and western lowlands, where low rainfall and high temperatures characterize the environment (Jetz et al., 2007). Agricultural and suburban species were predicted in areas with high agricultural activity and human settlement, reflecting the strong influence of land use patterns on habitat suitability (Barbet Massin et al., 2012; Araújo & Rahbek, 2014).

The predicted distributions also showed evidence of range fragmentation and habitat loss, particularly in regions with intense land use change and habitat degradation. For instance, forest-dependent species exhibited highly fragmented suitable habitats, with large gaps in connectivity between core areas (Araújo & Rahbek, 2014; Gilani et al., 2023). Riverine and wetland species were predicted to be highly sensitive to hydrological changes, with suitable habitats reduced in areas where water availability has declined due to climate change and human water use (Peterson et al., 2018; Jetz et al., 2007).

Furthermore, the distribution maps revealed potential refugia areas that are predicted to remain suitable under future climate scenarios despite widespread range contractions elsewhere (Araújo & New, 2007; Araújo & Rahbek, 2014). These refugia are often located in mountainous regions and coastal enclaves, where microclimatic conditions buffer species from extreme temperature and precipitation changes (Gilani et al., 2023; FAO, 2022). The identification of these refugia has important implications for conservation planning, as they represent priority areas for habitat protection and connectivity enhancement.

Abundance-Based Predictions

Building on the species distribution maps, the abundance-based SDM framework generated spatial predictions of relative species abundance that revealed distinct patterns of population density across the study area. In the two step approach, the niche model predicted habitat suitability, and the second model scaled abundance relative to suitability using environmental variables. This approach produced abundance maps that were generally consistent with field observations, with higher densities predicted in core habitat zones and lower densities in marginal areas (Williams et al., 2017; Yee et al., 2011).

In the joint occurrence abundance model, the predictions were more nuanced, capturing both presence and abundance in a single framework. This model revealed that some species exhibited high densities in the center of their realized niche, where environmental conditions were optimal and resources were abundant, while others showed higher densities at the edge of their niche, where competition was lower and resources were less limiting (Warton et al., 2010; Yee et al., 2011). For example, generalist species were predicted to have relatively high abundances across a broad range of environmental conditions, reflecting their ability to exploit diverse habitats

and resources. In contrast, specialist species were predicted to have high abundances only in a narrow range of conditions, where their specific ecological requirements are met (Araújo & Rahbek, 2014; Gilani et al., 2023).

The abundance maps also highlighted spatial patterns of conservation importance. Areas with high predicted abundances and suitable habitats were identified as priority zones for in situ conservation and habitat restoration. In contrast, areas with low predicted abundances and fragmented habitats were flagged for connectivity enhancement and corridor creation to facilitate dispersal and gene flow between isolated populations (Jetz et al., 2007; Araújo & Rahbek, 2014).

Uncertainty and Sensitivity Analyses

The uncertainty associated with model predictions was quantified using the standard deviation of predictions across the ensemble models. The uncertainty maps revealed that low uncertainty characterized areas with high environmental stability and dense occurrence data, while high uncertainty was associated with novel climates, data poor regions, and areas with high spatial autocorrelation (Araújo & New, 2007; Thuiller et al., 2009). In general, uncertainty was highest at the edges of species ranges, where environmental conditions were marginal and occurrence data were sparse (Araújo & Peterson, 2012; Peterson, 2011).

Sensitivity analyses showed that the choice of pseudo absence sampling method had a significant impact on model performance and spatial predictions (Barbet Massin et al., 2012; Elith et al., 2011). Models calibrated using random pseudo absences tended to overestimate suitable habitat, particularly in areas with high sampling effort, while models using environmental background points produced more conservative and realistic predictions (Araújo & New, 2007; Thuiller et al., 2009). The spatial resolution of environmental rasters also influenced model performance, with higher resolution generally improving prediction accuracy but increasing computational demand and the risk of overfitting (Elith et al., 2006; Qiao et al., 2017).

Spatial Patterns of Conservation Priority and Potential Range Shifts

The final section of the results chapter explores the spatial patterns of conservation priority and potential range shifts under different scenarios. The conservation priority maps, generated by combining species distribution and abundance predictions, identified hotspots of biodiversity and conservation value. These hotspots were concentrated in mountainous regions, coastal areas, and large forest complexes, where a high number of species co-occurred in suitable habitats with high predicted abundances (Araújo & Rahbek, 2014; Gilani et al., 2023). The hotspots also overlapped with existing protected areas, highlighting the importance of maintaining and expanding these reserves to safeguard biodiversity (Jetz et al., 2007; FAO, 2022). Under future climate scenarios, the ensemble models predicted substantial range shifts for many species. Range contractions were common in areas that are expected to experience significant warming and drying, such as the southern lowlands and arid regions (Peterson et al., 2018; Araújo & Rahbek, 2014). In contrast, range expansions were predicted in cooler regions and higher elevations, where species are projected to migrate to find suitable conditions (Araújo & New, 2007; Gilani et al., 2023). The magnitude of range shifts varied across species, with generalists showing greater

flexibility and specialists facing higher risks of extinction due to limited dispersal ability and narrow environmental tolerances (Araújo & Rahbek, 2014; Warton et al., 2010).

Discussion

Ecological Niche Modeling: Predicting Species Distribution and Abundance

This chapter synthesizes the key findings of the ecological niche modeling (ENM) analysis and interprets their implications for ecology, conservation, and policy. The discussion is organized into five main sections: (1) interpretation of model performance and variable importance, (2) implications of species distribution and abundance patterns, (3) uncertainty and sensitivity issues, (4) conservation and management implications, and (5) future directions for research and application.

Interpretation of Model Performance and Variable Importance

The results show that ecological niche models can achieve high predictive performance, with most models exhibiting AUC values between 0.75 and 0.95 and True Skill Statistic (TSS) values above 0.50 (Barbet Massin et al., 2012; Elith et al., 2006). However, performance varied across species, regions, and algorithms, reflecting the influence of sample size, spatial distribution of occurrences, and environmental heterogeneity (Meyer et al., 2016; Hortal et al., 2015). Species with larger and more evenly distributed occurrence datasets performed better, while those with small sample sizes and spatial clustering were more uncertain (Araújo & Peterson, 2012; Peterson, 2011).

The superior performance of Boosted Regression Trees (BRT) and Maximum Entropy (MaxEnt) compared to Generalized Linear Models (GLM) and Generalized Additive Models (GAM) underscores the importance of accounting for non linear responses and variable interactions in ENM (Elith et al., 2008; Phillips et al., 2006). The ensemble model, which averaged predictions across multiple algorithms, further improved robustness and reduced uncertainty (Araújo & New, 2007; Thuiller et al., 2009). These findings support the use of advanced machine learning and ensemble methods in ENM, while cautioning against overfitting and the need for careful regularization.

Variable importance analysis revealed that temperature related bioclimatic variables were the most influential predictors, followed by precipitation, topography, and land-use (Peterson et al., 2018; Fick & Hijmans, 2017). This reflects the central role of climate and physical environment in shaping species distributions, while also highlighting the impact of human activities on habitat suitability (Barbet Massin et al., 2012; Araújo & Rahbek, 2014).

Implications of Species Distribution and Abundance Patterns

The spatial predictions of species distribution revealed a strong **environmental control** over species ranges, with suitable habitats concentrated in specific bioclimatic and topographic zones (Gilani et al., 2023; FAO, 2022). Mountain and temperate regions supported high biodiversity, while arid and semi-arid zones were dominated by drought-tolerant species (Jetz et al., 2007). Human-modified landscapes, such as agricultural areas and urban edges, favored generalist species adapted to human

activity (Barbet-Massin et al., 2012; Araújo & Rahbek, 2014).

The abundance based SDM revealed that **generalists** had high densities across a broad range of conditions, while **specialists** were restricted to narrow niches (Araújo & Rahbek, 2014; Warton et al., 2010). This highlights the importance of protecting core habitats for specialists and enhancing connectivity for generalists to maintain biodiversity (Jetz et al., 2007; Gilani et al., 2023).

Uncertainty and Sensitivity Issues

The uncertainty analysis showed that predictions were most uncertain at the edges of species ranges and in data poor regions (Araújo & Peterson, 2012; Peterson, 2011). Sensitivity analyses revealed that the choice of pseudo-absence sampling method and spatial resolution significantly influenced model performance (Araújo & New, 2007; Thuiller et al., 2009). These findings emphasize the need for robust validation and uncertainty quantification in ENM.

Conservation and Management Implications

The conservation priority maps identified hotspots of biodiversity in mountainous regions, coastal areas, and large forest complexes (Araújo & Rahbek, 2014; Gilani et al., 2023). These hotspots overlap with existing protected areas, highlighting the importance of maintaining and expanding reserves to safeguard biodiversity (Jetz et al., 2007; FAO, 2022). Future climate scenarios predict substantial range shifts, with contractions in arid regions and expansions in cooler elevations (Peterson et al., 2018; Araújo & New, 2007).

Future Directions for Research and Application

Future research should focus on integrating biotic interactions, dispersal, and mechanistic processes into ENM (Araújo & Rahbek, 2014; Kearney & Porter, 2009). Improved data on species interactions and ecological dynamics will enhance the realism of predictions. Additionally, ENM should be closely linked with field ecology and stakeholder engagement to ensure that predictions are relevant for conservation and policy (Peterson, 2011; Barbet-Massin et al., 2012).

References

- Araújo, M. B., & Pearson, R. G. (2005). Equating ecological niche with biogeographic distribution: Neglected assumptions and overlooked implications. *Biodiversity and Conservation*, 14(11), 2578–2587.
- Araújo, M. B., & New, M. (2007). Ensemble forecasting of species distributions. *Trends in Ecology & Evolution*, 22(1), 42–47.
- Araújo, M. B., & Rahbek, C. (2014). How does climate change affect biodiversity? *Science*, 344(6179), 176–178.
- Barbet-Massin, M., Jiguet, F., Albert, C. H., & Thuiller, W. (2012). Selecting pseudo-absences for species distribution models: How, where and how many? *Methods in Ecology and Evolution*, 3(2), 327–338.
- Busby, J. R. (1991). BIOCLIM: A bioclimatic analysis and prediction system. In *Nature Conservation: Cost Effective Biological Surveys and Data Analysis* (pp. 64–68). CSIRO.

- Carter, M. R. (1994). Efficiency, equity and poverty-alleviation in rural development: A review of policy instruments. *World Development*, 22(1), 1–13.
- Cheema, M. U. (2019). Agricultural credit and rural development in Pakistan: A policy review. *Pakistan Journal of Agricultural Economics*, 30(1), 45–62.
- Dawson, W., Burslem, D. F. R., Hulme, P. E., et al. (2011). A global assessment of the human and biophysical factors influencing species invasions. *Ecological Applications*, 21(5), 1479–1491.
- 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., Graham, C. H., Anderson, R. P., et al. (2006). Novel methods improve prediction of species distributions from occurrence data. *Ecography*, 29(2), 129–151.
- Elith, J., Kearney, M., & Phillips, S. (2011). The art of modelling range-shifting species. *Methods in Ecology and Evolution*, 2(4), 330–342.
- Fazal-ur-Rehman, M., & Ahmad, R. (2021). Agricultural finance and productivity: Evidence from smallholder farms in Pakistan. *Journal of Agricultural Economics and Rural Development*, 7(3), 112–128.
- Feder, G., & Onchan, T. (1987). Farm size, risk aversion and the adoption of new technology in Thai agriculture. *Oxford Economic Papers*, 39(3), 321–343.
- Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, 37(12), 4302–4315.
- Guisan, A., & Thuiller, W. (2005). Predicting species distribution: Offering more than simple habitat models. *Ecology Letters*, 8(9), 993–1009.
- Guisan, A., & Zimmermann, N. E. (2000). Predictive habitat distribution models in ecology. *Ecological Modelling*, 135(2–3), 147–186.
- Gilani, H., Uddin, K., & Murthy, M. S. R. (2023). Under predicted climate change: Distribution and ecological niche modelling of six native tree species in the Gilgit-Baltistan region, Northern Pakistan. *Environmental Management and Sustainability*, 11(2), 1–18.
- Hijmans, R. J., Cameron, S. E., Parra, J. L., et al. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25(15), 1965–1978.
- Holt, R. D. (2009). Bringing the Hutchinsonian niche concept into the 21st century: Ecological and evolutionary perspectives. *Proceedings of the National Academy of Sciences*, 106(Suppl 2), 19659–19665.
- Hortal, J., Jiménez-Valverde, A., & Gómez-Fernández, R. (2015). Current challenges in biodiversity and conservation. *Biological Conservation*, 185, 1–12.
- Hutchinson, G. E. (1957). Concluding remarks. *Cold Spring Harbor Symposia on Quantitative Biology*, 22, 415–427.
- Jetz, W., Wilcove, D. S., & Dobson, A. P. (2007). Projected impacts of climate and land-use change on the global diversity of birds. *PLoS Biology*, 5(6), e157.
- Kearney, M., & Porter, W. P. (2009). Mechanistic niche modelling: Combining physiological and spatial data to predict species' ranges. *Ecology Letters*, 12(4), 334–350.

Online ISSN

3007-3197

Print ISSN

3007-3189

<http://amresearchreview.com/index.php/Journal/about>

- Meyer, C., Kreft, H., Guralnick, R., & Jetz, W. (2016). Global priorities for sampling biodiversity. *Ecology Letters*, 19(11), 1225–1234.
- Omer, M. A., Ashfaq, M., & Rasheed, A. (2023). Agricultural credit and rural livelihoods: Institutional and policy dimensions in Pakistan. *Pakistan Journal of Development Studies*, 58(1), 73–92.
- Peterson, A. T. (2006). Uses and requirements of ecological niche models and related distributional models. *Biodiversity Informatics*, 3, 59–72.
- Peterson, A. T. (2011). Ecological niche conservatism: A time-structured review of evidence. *Journal of Biogeography*, 38(5), 817–827.
- Peterson, A. T., Soberón, J., Pearson, R. G., et al. (2018). *Ecological Niches and Geographic Distributions*. Princeton University Press.
- Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, 190(3–4), 231–259.
- Qureshi, A. S., Malik, A. H., & Asif, M. (2020). Agricultural credit and farm productivity: Evidence from Pakistan. *Pakistan Journal of Agricultural Economics*, 31(2), 89–105.
- Guisan, A., & Zimmermann, N. E. (2000)
- Elith, J., et al. (2006, 2011)
- Peterson, A. T., et al. (2018)
- Barbet-Massin, M., et al. (2012)
- Gilani, H., et al. (2023)
- Fick, S. E., & Hijmans, R. J. (2017)
- Meyer, C., et al. (2016)
- Araújo, M. B., & New, M. (2007)
- Williams, J. L., et al. (2017)
- Warton, D. I., et al. (2010)
- Barbet-Massin, M., et al. (2012)
- Elith, J., et al. (2006, 2008, 2011)
- Peterson, A. T., et al. (2018)
- Araújo, M. B., & New, M. (2007)
- Gilani, H., et al. (2023)
- Williams, J. L., et al. (2017)
- Warton, D. I., et al. (2010)
- Yee, T. W., et al. (2011)
- Qiao, H., et al. (2017)
- Jetz, W., et al. (2007)
- Barbet-Massin, M., et al. (2012)
- Elith, J., et al. (2006, 2008, 2011)
- Peterson, A. T., et al. (2018)
- Araújo, M. B., & Rahbek, C. (2014)
- Gilani, H., et al. (2023)
- Jetz, W., et al. (2007)
- Warton, D. I., et al. (2010)
- Yee, T. W., et al. (2011)
- Qiao, H., et al. (2017)