Range-wide habitat use of the Harpy Eagle indicates four major tropical forest gaps in the Key Biodiversity Area network

Sutton, LJ

http://hdl.handle.net/10026.1/19931

10.1093/ornithapp/duac019
Ornithological Applications
Oxford University Press

All content in PEARL is protected by copyright law. Author manuscripts are made available in accordance with publisher policies. Please cite only the published version using the details provided on the item record or document. In the absence of an open licence (e.g. Creative Commons), permissions for further reuse of content should be sought from the publisher or author.
RESEARCH ARTICLE

Range-wide habitat use of the Harpy Eagle indicates four major tropical forest gaps in the Key Biodiversity Area network

ABSTRACT

Quantifying habitat use is important for understanding how animals meet their requirements for survival and provides information for conservation planning. Currently, assessments of range-wide habitat use that delimit species distributions are incomplete for many taxa. The Harpy Eagle (*Harpia harpyja*) is a raptor of conservation concern, widely distributed across Neotropical lowland forests, that currently faces threats from habitat loss and fragmentation. Here, we use penalized logistic regression to identify species-habitat associations and predict habitat suitability based on a new International Union for the Conservation of Nature range metric, termed Area of Habitat. From the species-habitat model, we performed a gap analysis to identify areas of high habitat suitability in regions with limited coverage in the Key Biodiversity Area (KBA) network. Range-wide habitat use indicated that Harpy Eagles prefer areas of 70-75% evergreen forest cover, low elevation, and high vegetation species richness. Conversely, Harpy Eagles avoid areas of >10% cultivated landcover and mosaic forest, and topographically complex areas. Our species-habitat model identified a large continuous area of potential habitat across the pan-Amazonia region, and a habitat corridor from the Chocó-Darién ecoregion of Colombia running north along the Caribbean coast of Central America. Little habitat was predicted across the Atlantic Forest biome, which is now severely degraded. The current KBA network covered 18% of medium to high Harpy Eagle habitat
exceeding a target biodiversity area representation of 10%, based on species range size. Four major areas of high suitability habitat lacking coverage in the KBA network were identified in north and west Colombia, western Guyana, and north-west Brazil. We recommend these multiple gaps of habitat as new KBAs for strengthening the current KBA network. Modelled area of habitat estimates as described here are a useful tool for large-scale conservation planning and can be readily applied to many taxa.

**Keywords:** Area of Habitat, conservation planning, gap analysis, habitat use, *Harpia harpyja*, Harpy Eagle, Key Biodiversity Areas, Species Distribution Models

**LAY SUMMARY**

- Quantifying habitat use is key to understanding animals’ requirements for survival and can inform spatial conservation planning by mapping species range limits
- Species that inhabit remote, hard-to-survey areas lack sufficient location data and there is a need to be able to predict into poorly sampled areas to estimate the potential area of habitat
- Using Species Distribution Models we identified Harpy Eagle range limits, habitat area and Key Biodiversity Area coverage across the species range
- Harpy Eagles prefer areas of 70-75% evergreen forest cover, high vegetation species richness and low elevation
- Key Biodiversity Areas covered 18% of highly suitable Harpy Eagle habitat but with key gaps in coverage in north and west Colombia, western Guyana, and north-west Brazil
Our method of calculating habitat area estimates based on a predictive spatial model is a useful tool for large-scale conservation planning and can be readily applied to many taxa.

**INTRODUCTION**

Determining habitat resource use is a fundamental aspect of wildlife ecology and conservation planning (Manly et al. 2002; Morrison et al. 2006). However, our understanding of range-wide species-habitat associations across continental extents is incomplete, even for well-studied groups such as birds (Gregory and Baillie 1998; Engler et al. 2017; Lees et al. 2020). Currently, many taxa face increasing threats from human-driven habitat loss and fragmentation across their entire range (Powers and Jetz 2019). Therefore, developing a broad spatial quantification of habitat use is an effective starting point for conservation planning (Margules and Pressey 2000; Early et al. 2008). Once habitat use is identified for a focal species, the key variables characterising those habitats can be used to produce a mapped representation of habitat across the species’ range (Hirzel et al. 2006). Management actions can then be directed to guide conservation planning to protect or enhance those areas (Margules and Pressey 2000; Suárez-Seoane et al. 2002).

Recently, the International Union for the Conservation of Nature (IUCN) developed a new range size metric termed Area of Habitat (AOH, Brooks et al. 2019). AOH is defined as the habitat available to a species based on habitat preferences and elevational limits within the mapped distributional range of a focal species. Various approaches have been taken to estimate AOH which all use a similar method of matching and overlaying the known mapped range, landcover and elevation limits of...
a given species (Brooks et al. 2019). While the AOH method is useful and repeatable, IUCN methods may still have limitations by missing areas that have no occurrence data but may still contain preferred habitat (Ramesh et al. 2017).

On the other hand, Species Distribution Models (SDMs) are statistical methods that assess species’ habitat requirements and predict distribution based on correlating environmental covariates with species occurrences (Elith and Leathwick 2009; Matthiopoulos et al. 2020; Valavi et al. 2021). Two example applications for SDMs are the re-evaluation of range sizes (e.g., Herkt et al. 2017), and the identification of gaps in protected or biodiversity area networks (e.g., de Carvalho et al. 2017).

Indeed, SDMs can predict more complex and ecologically realistic geographic ranges compared to IUCN range maps (Breiner et al. 2017; Herkt et al. 2017). Using model-based interpolation based on the AOH guidelines but adapted to a correlative modelling approach like SDMs (Da Silva et al. 2020), may also be more effective for highlighting species-specific gaps in biodiversity area coverage by identifying higher coverage of suitable pixels (Di Marco et al. 2017).

Designation of biodiversity areas is a fundamental tool for conservation (IUCN 2016) and has been successful in reducing habitat loss and fragmentation for many taxa (Brooks et al. 2009). However, despite wide coverage in the global biodiversity area network, gaps in biodiversity area coverage still exist with new areas being continually added (KBA Standards and Appeals Committee 2019). Additionally, not all biodiversity areas are located in places deemed effective for conservation but are often designated by human socio-economic factors (Pringle 2017; Morán-Ordóñez 2020; Rodrigues and Cazalis 2020). Key Biodiversity Areas (KBAs, BirdLife
International 2020) are sites of international significance for the global persistence of biodiversity. KBAs also protect areas important for biodiversity and aim to overlap with the entire global protected area network (The World Database on Protected Areas, UNEP-WCMC & IUCN 2021; Donald et al. 2019). The KBA concept is largely based on Important Bird and Biodiversity Areas (IBAs), a template for KBAs which aims to identify and conserve sites of global importance for bird species (Donald et al. 2019). Indeed, the majority of terrestrial KBAs are designated based on birds and contain either: (1) populations of globally threatened species, (2) populations and communities of range- or biome-restricted species, or (3) substantial congregations of specific avian taxa.

Information on where to establish new KBAs identifies where the current biodiversity area networks miss key bird species and where these gaps need filling. Gap analysis is an established method to identify discontinuities in protected or biodiversity area networks (Scott et al. 1993) and has been effective in setting conservation planning priorities across a range of taxa (Margules and Pressey 2000). In particular, gap analysis has identified priority conservation areas for many taxa across the highly biodiverse Neotropics (e.g., de Carvalho et al. 2017; Bax and Francesconi 2019; Perrig et al. 2020). The Harpy Eagle (Harpia harpyja) is a large raptor historically distributed throughout Neotropical lowland tropical forest from southern Mexico to northern Argentina (Miranda et al. 2019; Sutton et al. 2021). The species was recently reclassified from ‘Near-Threatened’ to ‘Vulnerable’ by the IUCN Red List due to continued habitat loss and persecution (Birdlife International 2021). Harpy Eagles are now largely restricted to tropical lowland broadleaf forest but can also
inhabit dry seasonal forest and fragmented habitat (Vargas González et al. 2006; Silva et al. 2013).

Despite this habitat specialization, the Harpy Eagle has a large range due to the extensive distribution of lowland tropical forest across the Neotropics. However, historical and ongoing deforestation has led to extirpations in parts of southern Mexico and Central America, and across the Atlantic Forest of Brazil (Vargas González et al. 2006; Silva et al. 2013; Meller and Guadagnin 2016). Current deforestation rates across the species’ stronghold in Amazonia are also of significant concern for its future persistence (Banhos et al. 2016; Miranda et al. 2019). As an apex predator requiring large tracts of continuous tropical lowland forest for breeding and foraging (Vargas González et al. 2014; Miranda 2015), the Harpy Eagle may also act as a useful trigger species for designating new regional IBAs (BirdLife International 2020), under the assumption that triggering a regional IBA would be justification for inclusion as a KBA. Further, as a threatened species of conservation concern, it fulfils the criteria for designating new regional IBAs based on inferred habitat area (category B1a; BirdLife International 2020), with the assumption that the gap sites identified are predicted to hold significant numbers of a threatened species.

Here, a predictive Species Distribution Model (SDM) was developed to identify species-habitat associations (Matthiopoulos et al. 2020; Valavi et al. 2021) based on penalized logistic regression (Phillips et al. 2017). Estimating Harpy Eagle distribution based solely on habitat predictors at the continental scale should provide the most accurate and reliable estimate of range size due to the Harpy Eagle’s generally high reliance on tropical lowland forest. Specifically, this study sets out a
baseline assessment of large-scale habitat use defining potential Harpy Eagle
distribution. A first estimate of modelled habitat suitability using a spatial framework
based on the Area of Habitat metric was then used to predict areas of highest habitat
suitability for the Harpy Eagle. Using this information, a broad-scale gap analysis
was generated to identify priority areas of highest habitat suitability in regions with
limited KBA network coverage. In short, this study applied statistical modelling to
systematic conservation planning to determine: (1) how effective the current KBA
network is for covering areas of Harpy Eagle habitat, and (2) where gap areas of
highest habitat suitability for the Harpy Eagle are located for inclusion as proposed
KBAs.

METHODS

Occurrence Data

Harpy Eagle occurrences were sourced from the Global Raptor Impact Network
(GRIN, McClure et al. 2021), a data information system for population monitoring of
all raptor species. For the Harpy Eagle, GRIN includes occurrence data from the
Global Biodiversity Information Facility (GBIF 2019) and eBird (Sullivan et al. 2009),
along with two additional occurrence datasets (Vargas González and Vargas 2011;
Miranda et al. 2019). Though it is recommended to apply sampling regime filters to
eBird occurrence data (Johnston et al. 2021), we opted to retain all eBird data points
because the majority of our eBird occurrences did not have sufficient sampling
regime metadata to employ these filters in the analysis (See Supplementary
Material). In doing so we also sought to achieve a large enough sample size to
capture the widest possible range of species-habitat associations needed for robust
predictions (Gaul et al. 2020; Santini et al. 2021).
Duplicate records and those with no geo-referenced location were removed and only occurrences recorded from year 2000 onwards were included to temporally match the timeframe of the habitat covariates. A 5-km spatial filter was applied between each occurrence point, which approximately matches the spatial resolution of the raster data (~4.5-km), resulting in one occurrence per pixel grid cell reducing the effect of biased sampling (Kramer-Schadt et al. 2013). We used this resolution and spatial filter distance because it is an appropriate spatial resolution for identifying environmental variation across lowland tropical regions (Fick & Hjimans 2017), to address continent-scale management issues. A total of 1021 geo-referenced records were compiled after data cleaning. Applying the 5-km spatial filter resulted in a filtered subset of 591 Harpy Eagle occurrence records for use in the calibration models (Fig. 1).

**Habitat Covariates**

To predict occurrence, habitat covariates representing landcover, topography and vegetation heterogeneity were downloaded from the EarthEnv (www.earthenv.org) and ENVIREM (Title and Bemmels 2018) repositories. Six continuous covariates were used at a spatial resolution of 2.5 arc-minutes (~4.5-km resolution): cultivated landcover, elevation, evergreen forest, habitat homogeneity (i.e., vegetation species richness, structure, composition and diversity), mosaic forest (i.e, a mosaic of mixed forest, shrubland and woody savanna) and Terrain Roughness Index (Table S1; See Supplementary Material). Covariates were selected a priori based on the IUCN Area of Habitat criteria from landcover and topographic factors related empirically to Harpy Eagle distribution and tropical forest raptor abundance in previous studies (Robinson...
Raster layers were cropped to a background region using a delimited polygon consisting of all known range countries (including Formosa, Jujuy, Misiones and Salta provinces in northern Argentina, and Chiapas, Oaxaca, and Tabasco states in southern Mexico).

**Species Distribution Model**

We fitted an SDM using penalized elastic net logistic regression (Fithian and Hastie 2013), via maximum penalized likelihood estimation (Hefley and Hooten 2015) in the R package `maxnet` (Phillips et al. 2017). Elastic net logistic regression imposes a regularization penalty on the model coefficients, shrinking towards zero the coefficients of covariates that contribute the least to the model, reducing model complexity (Gastón and García-Viñas 2011; Helmstetter et al. 2020). The `maxnet` package uses penalized logistic regression to fit the SDM based on the maximum entropy algorithm, MAXENT (Phillips et al. 2017), which is mathematically equivalent to estimating the parameters for an inhomogeneous Poisson process (IPP; Fithian and Hastie 2013; Renner and Warton 2013; Hefley and Hooten 2015; Renner et al. 2015). In its original implementation MAXENT imposed a ‘lasso’ (least absolute shrinkage and selection operator) regularization penalty, where only the most significant covariates are retained, with uninformative covariates set at zero. Instead, the `maxnet` package uses an elastic net (via the `glmnet` package, Friedman et al. 2010) to perform automatic covariate selection (lasso) and continuous shrinkage (ridge regression) simultaneously (Zou and Hastie 2005; Phillips et al. 2017), evaluating the contribution of all covariates and shrinking low-contribution coefficients towards zero. Elastic net regularization improves predictive accuracy.
compared to the lasso, in both simulated and real data examples (Zou and Hastie 2005) and may be viewed as a generalization of the lasso. We parametrized the penalized logistic regression model using infinite weighting within the IPP framework because this is the most effective method to model presence-background data as used here (Warton and Shepherd 2010; Hefley and Hooten 2015). Within the maxnet package the complementary log-log (cloglog) link function was selected as a continuous index of habitat suitability, with 0 = low suitability and 1 = high suitability. Phillips et al. (2017) demonstrated the cloglog link is equivalent to an IPP and can be interpreted as a measure of relative occurrence probability proportional to a species potential abundance. We used a tuned penalized logistic regression algorithm because this approach outperforms other SDM algorithms (Valavi et al. 2021), including ensemble averaged methods (Hao et al. 2020).

We used a random sample of 10,000 background points as pseudo-absences recommended for regression-based modelling (Barbet-Massin et al. 2012) and to sufficiently sample the background calibration environment (Guevara et al. 2018; Figure S1). Optimal-model selection was based on Akaike’s Information Criterion (Akaike 1974) corrected for small sample sizes (AICc; Hurvich and Tsai 1989), to determine the most parsimonious model from two key maxnet parameters: regularization beta multiplier (β; level of coefficient penalty) and feature classes (response functions, Warren and Seifert 2011; Phillips et al. 2017). Eighteen candidate models of varying complexity were built by conducting a grid search using a range of regularization multipliers from 1 to 5 in 0.5 increments, and two feature classes (response functions: Linear, Quadratic) in all possible combinations using the ‘trainMaxNet’ function in the R package enmSdm (Smith 2019). We considered
all models with a $\Delta AIC_c < 2$ as having strong support (Burnham and Anderson 2004),
and the model with the lowest $\beta$ was selected to avoid overfitting. We used response
curves and parameter estimates to measure variable performance in the optimal
calibration model.

We used Continuous Boyce index (CBI; Hirzel et al. 2006) as a threshold-independent metric of how predictions differ from a random distribution of observed
presences (Boyce et al. 2002). CBI is consistent with a Spearman correlation ($r_s$) and
ranges from -1 to +1. Positive values indicate predictions consistent with observed
presences, values close to zero suggest no difference from a random model, and
negative values indicate areas with frequent presences having low environmental
suitability. Mean CBI was calculated using five-fold cross-validation on 20 % test
data with a moving window for threshold-independence and 101 defined bins in the
R package *enmSdm* (Smith 2019). The optimal model was tested against random
expectations using partial Receiver Operating Characteristic ratios (pROC), which
estimate model performance by giving precedence to omission errors over
commission errors (Peterson et al. 2008). Partial ROC ratios range from 0 to 2 with 1
indicating a random model. Function parameters were set with a 10% omission error
rate, and 1000 bootstrap replicates on 50% test data to determine significant ($\alpha =
0.05$) pROC values $>1.0$ in the R package *ENMGadgets* (Barve and Barve, 2013).

**Range Size and Gap Analysis**

To calculate Area of Habitat in suitable pixels and assess the effectiveness of the
KBA network, we reclassified the continuous prediction to a binary threshold
prediction. All pixels equal to or greater than the median pixel value of 0.345 from the
continuous model were used as a suitable threshold for conservation planning (Liu et al. 2005; Rodríguez-Soto et al. 2011; Portugal et al. 2019). We selected the median because this threshold is not reliant on measuring predictive ability based on unknown pseudo-absences (Merow et al. 2013), unlike measures that use specificity (Liu et al. 2013). The KBA network polygons (as of September 2020; BirdLife International 2020) were then clipped to the reclassified area, establishing those KBAs covering pixels of habitat suitability \( \geq 0.345 \) threshold. To visualise KBA network coverage, we reclassified the continuous prediction into four discrete quantile habitat classes (No habitat: 0.0 - 0.067; Low: 0.068 - 0.344; Medium: 0.345 - 0.701; High: 0.702 - 1.000).

The clipped KBA network polygons were then overlaid onto the discrete class map identifying those pixels of medium to high habitat \( \geq 0.345 \) threshold which were within the clipped KBA network polygons. We used the threshold range size to calculate a protected area ‘representation target’, quantifying how much protected area representation is needed for a species dependent on its range size following the formulation of Rodrigues et al. (2004),

\[
\text{Target} = \max(0.1, \min(1, -0.375 \times \log_{10}(\text{range size}) + 2.126)) \quad (1)
\]

where ‘Target’ is equal to the percentage of protected target representation required for the species ‘range size’, as used in subsequent applications of the formula (Butchart et al. 2015; Di Marco et al. 2017). As can be verified by inserting different range size values, this formula yields a target of 10 % for species with a range size \( >250,000 \text{ km}^2 \) and increasing proportional representation for smaller range sizes up
to a target of 100 % if range size <1000 km². We used the current KBA coverage to
calculate the difference between the current level of KBA coverage compared to the
target level representation.

Lastly, we calculated two IUCN range metrics from our modelled AOH binary
prediction. First, Area of Occupancy (AOO) was calculated as the number of raster
pixels predicted to be occupied scaled to a 2x2 km grid following IUCN guidelines
(IUCN 2018) in the R package redlistr (Lee et al. 2019). Second, we converted our
modelled AOH binary raster to a polygon using an 8-neighbour patch rule and
applied a smoothing function using the Chaikin algorithm (Chaikin 1974) in the R
package smoothr (Strimas-Mackey 2021). Extent of Occurrence (EOO) was
calculated by fitting a minimum convex polygon (MCP) around the furthest
boundaries of the projected habitat of the AOH polygon following IUCN guidelines
(IUCN 2018). We calculated both a maximum EOO, including all the area with the
MCP, and a minimum EOO, masking out the area within the MCP that could not be
occupied over the ocean. All range metric calculations were performed using an
Equatorial Lambert Azimuth Equal-Area projection. General model development and
geospatial analysis were performed in R (v3.5.1; R Core Team, 2018) using the
dismo (Hijmans et al. 2017), raster (Hijmans 2017), rgdal (Bivand et al. 2019), rgeos
(Bivand and Rundle 2019) and sp (Bivand et al. 2013) packages.

RESULTS

Species Distribution Model

Six candidate models had an ΔAICc ≤ 2, and the model with the lowest regularization
multiplier (β) was selected (Model 6 in Table S2, see Supplementary Material). The
best-fit SDM (ΔAICc = 1.19) had linear and quadratic terms and β = 2.5 as model parameters, with high calibration accuracy (mean CBI = 0.960), and was robust against random expectations (pROC = 1.431, SD± 0.055, range: 1.244 – 1.594).

From the penalized linear beta coefficients, Harpy Eagles were positively associated with evergreen forest (0.065) and most negatively associated with habitat homogeneity (-3.849), followed by mosaic forest (-0.026), Terrain Roughness Index (-0.023) cultivated land (-0.010) and elevation (-0.001).

The largest continuous area of habitat extended across Amazonia and the Guiana Shield (Figure 2). A habitat corridor was identified through Central America along the Caribbean coast, extending south into the Chocó-Darién ecoregion along the Pacific coast of Colombia (Figure S2). Little habitat was predicted across the largely deforested Atlantic Forest region in Brazil. From the SDM response functions, evergreen forest had peak suitability at 70-75 % forest cover, with highest suitability for topographic areas of both low elevation and terrain ruggedness (Figure 3). Habitat suitability was highest in areas of low homogeneity < 0.2 (i.e., highly heterogenous species-rich vegetation), areas with < 10 % human cultivated landcover, and zero or low percentage of mosaic forest.

**Range Size and Gap Analysis**

The reclassified binary model (median threshold = 0.345) calculated an Area of Habitat equalling 7,479,752 km² (Figure 4). The current KBA network covered 18.1 % (1,352,879 km²) of this habitat area in the medium to high discrete quantile classes (Figure 5), 8.1 % greater than the target representation (10 %). Four major gaps (Figure 5, blue circles/ellipses) for high class habitat without extensive KBA
coverage were identified in: (1) the Chocó-Darién ecoregion in western Colombia (Fig. 6), (2) the Magdalena-Urabá moist forests of northern Colombia (Fig. 6), (3) north-east Amazonas state in Brazil, and (4) north and west Guyana. From our AOH model, maximum Extent of Occurrence (EOO) was 18,130,602 km² and minimum EOO 14,738,408 km², with an AOO of 708,697 occupied cells.

**DISCUSSION**

Our results indicate that Harpy Eagle populations are more likely to be associated with dense (70-75%) evergreen forest cover, low elevation, and high vegetation species richness across their range. Conversely, Harpy Eagles seem to avoid extensive areas of cultivated land, mosaic forest, and high terrain complexity. Using the AOH parameters as the basis for the habitat model predicted a large area of habitat across the pan-Amazónia region, and a habitat corridor extending from the Pacific coast of Colombia, north along the Caribbean coast of Central America. Almost no habitat was predicted across the Atlantic Forest region, which is now severely degraded. The current KBA network coverage exceeded the target biodiversity area representation (10 %), covering 18 % of medium to high Harpy Eagle habitat. Considering the large range of the Harpy Eagle, the current KBA extent is encouraging but misses key areas of potentially important habitat. Four areas of high suitability habitat were identified as gaps in the KBA network for north and west Colombia, western Guyana, and north-west Brazil. We recommend establishing new KBAs in these four areas, further strengthening the current KBA network across the region.
Despite the high predictive performance of our continuous model and the ability of the reclassified discrete model to identify previously unprotected areas of key habitat, we recognise there are limitations to our approach. Thresholding continuous SDMs is common practice but not always appropriate (Guillera-Arroita et al. 2015; Santini et al. 2021). However, in this context thresholding was justifiable to achieve our aim of calculating discrete habitat classes for use in spatial conservation planning (Guillera-Arroita et al. 2015). Using a Bayesian approach with a range of continuous probabilities would be a useful future step forward to account for any uncertainty in model outputs (Carlson 2020). The use of presence-background data in SDMs is widespread but has been fraught with statistical issues related to sampling bias since their inception (Ranc et al. 2017). However, recent advances implementing the unifying inhomogeneous Poisson process framework which models points as a log-linear intensity function of the covariates, as used here, can effectively account for sampling bias that may skew model predictions (Renner et al. 2015; Isaac et al. 2019).

**Habitat Use**

Broad and fine scale species-habitat assessments often result in different variables emerging as important, potentially leading to contrasting recommendations for conservation (Gregory and Baillie 1998). However, our results show general similarities to habitat models from previous studies at both broad and fine scales. The SDM was consistent with predicted Harpy Eagle habitat from an earlier broad-scale SDM (Miranda et al. 2019). This was expected because both SDMs used measures of forest cover as landcover predictors but different modelling methodologies. This reinforces the consistency in SDM outputs for the Harpy Eagle.
from a range of algorithms and gives confidence in SDM predictions that have been
criticised for lacking ecological realism (Fourcade et al. 2017). Building on the
Miranda et al. (2019) model, the SDM here also predicted a distinct corridor of
habitat extending from the Chocó-Darién ecoregion of west Colombia north through
Central America along the Caribbean coast (Figure 6). This suggests that including a
habitat heterogeneity covariate, along with topographic and landcover predictors,
was able to identify key areas of habitat undetectable from other texture measures
used in that study.

Habitat heterogeneity is a key landscape characteristic, here representing vegetation
species richness, important for determining general biodiversity patterns (Stein et al.
2014), including for lowland tropical forest raptors (Jullien and Thiollay 1996;
Anderson 2001). Areas of high species-rich vegetation provide more diverse niche
space, promoting greater species coexistence and thus increased species diversity
(Tews et al. 2004). For the Harpy Eagle, areas of higher habitat heterogeneity may
be preferred over more homogenenous areas because they contain a greater density
and diversity of prey species (Miranda 2018). Further, a diverse forest canopy
structure may also facilitate aerial attacks on canopy prey, by providing more hunting
perches (Vargas González et al. 2014). Moreover, the SDM confirmed the restricted
elevational distribution for the Harpy Eagle, consistent with a landscape-level SDM
(Vargas González et al. 2020). This may be similarly linked to the Harpy Eagles’
preference for nesting in large, canopy-emergent trees, and the abundance of its
main prey of arboreal mammals, both of which occur in greater abundance at lower
elevations (Miranda 2015; Miranda et al. 2020).
Harpy Eagles are dependent on large tracts of lowland tropical forest for breeding and foraging (Vargas González et al. 2014; Miranda et al. 2019). Indeed, breeding success was higher in areas with > 70 % forest cover in northern Mato Grosso, Brazil (Miranda et al. 2021), consistent with the range-wide response to evergreen forest cover here. Perhaps as important, strong negative associations were identified with >10 % cultivated landcover and mosaic forest, showing that Harpy Eagles avoid areas of high human impact and sporadic forest cover. This implies that, as deforestation increases across the species’ range, the Harpy Eagle may struggle to adapt to large areas of human disturbance and heavily fragmented landscapes (Miranda et al. 2021).

Area of Habitat

Our method of calculating the Area of Habitat metric refines previous range size estimates (Birdlife International 2021; Sutton et al. 2021) and provides a baseline area of habitat map for the Harpy Eagle. There was 4.6 % less area in our modelled AOH range polygon (7,479,752 km²), than in the current IUCN range map (7,838,093 km²; Fig. 4). Therefore, we recommend this new AOH estimate be incorporated into future IUCN assessments for the species. Our modelled AOH polygon also had 24 % less area compared to a binary SDM map using solely climatic and topographic predictors (9,844,399 km²; Sutton et al. 2021). If we assume that the SDM from Sutton et al. (2021) based on climate and terrain is representative of the Harpy Eagle pre-industrial range (in the absence of satellite-derived landcover not available for pre-industrial times), then the species’ habitat range has shrunk by nearly a quarter during the industrial period to the present.
One limitation of the analyses was the timeframe of the remote-sensing data used for the covariates. Both the landcover and vegetation covariates are a consensus product collected between the years 1992-2005, with land use having changed in parts of Neotropics since then (Powers and Jetz 2019). Therefore, the Area of Habitat prediction should be viewed as a conservative baseline assessment, knowing that landcover can change rapidly. Processing large areas of current remote-sensed landcover data at continental-scales can be challenging due to the high computing power required; the EarthEnv habitat variables are recommended as a readily available dataset to use for first estimates of modelled AOH at large scales (Tuanmu and Jetz 2014, 2015).

Current and predicted future habitat loss may lead inevitably to declines in populations of some species, increasing their extinction risk (Powers and Jetz 2019). Continued habitat loss and fragmentation is likely to have a negative impact on the future persistence of many birds across the highly biodiverse Neotropics (Bird et al. 2011). The Harpy Eagle is a good example, despite its large range precluding high extinction risk (Gaston and Fuller 2009). Continued habitat loss and fragmentation through agricultural development and logging across its geographic range (Vargas González et al. 2006; Miranda et al. 2020) should raise the alarm about the species’ future (Krüger and Radford 2008; Miranda et al. 2019). The declining range of the Harpy Eagle is demonstrated by the few breeding and sighting records in the largely deforested Atlantic Forest (Meller and Guadagnin 2016; Suscke et al. 2017), and parts of southern Mexico and Central America (Vargas González et al. 2006), reflected in the results from the SDM. Our results should therefore serve as a forewarning of what could happen across parts of the core habitat area in Amazonia.
where deforestation has steadily increased since 2000 (Hansen et al. 2008), with current deforestation rates across the Brazilian Amazon increasing since 2013 (Silva Junior et al. 2021).

As a baseline assessment, our SDM should be viewed as a *maximum extent of habitat*, knowing that deforestation is an ongoing process across the pan-Amazonia region (Bird et al. 2011; Hansen et al. 2020). Approximately 18% of tropical forest in Amazonia had been cleared by 2011 (Bird et al. 2011), with predictions of up to 40% of forest cover lost by 2050 (Soares-Filho et al. 2006). Recently, those tropical forests of highest structural integrity most associated with preferred Harpy Eagle habitat (tall, closed canopy forest and low human pressure; Vargas González et al. 2014; Miranda et al. 2020) were identified as largely limited to the Amazon basin (Hansen et al. 2020). These forests generally remain intact due to their remoteness (Soares-Filho et al. 2006), but with the majority having no formal protection.

**Gap Analysis**

Although the current coverage of the KBA network within our modelled AOH range (~18%) exceeded the representative biodiversity area target based on species range size set here (10%), it is substantially lower than the proportion of IBA network coverage for threatened bird species overall in Amazonia (54.9%, Bird et al. 2011). Of the four key gaps identified here only gap 3 in north-west Amazonas state...
in Brazil has any form of current protection as an area of indigenous land (UNEP-WCWC & IUCN 2020). The three remaining gap areas have little formal protection or KBA coverage, despite both the Chocó-Darién ecoregion (gap 1) and Guyana (gap 4) having extensive Harpy Eagle habitat. In the case of Guyana it is likely that most habitat is ‘passively’ protected due to the inaccessibility of the region. However, solely relying on remoteness may be short-sighted and extending the current KBAs east and west of Guyana to cover a larger portion of the Guiana Shield is recommended. To this aim, given that on average ~49 % of the area of each KBA/IBA globally has formal protection (Waliczky et al. 2019), intersecting KBA coverage with nationally protected areas across the Harpy Eagle range would be a useful next step in protected area assessment for the species (Butchart et al. 2012).

The Chocó-Darién ecoregion is one of 25 global biodiversity hotspots prioritized for conservation (Myers et al. 2000). Based on satellite remote-sensing, deforestation for agricultural expansion has steadily increased in the region over the past two decades (Fagua et al. 2019; Fagua and Ramsey 2019). Approximately 42 % of forest remains intact, making this an area of high importance for protection not only for the Harpy Eagle but for all the associated fauna, flora, and crucial ecological processes. Establishing and reinforcing the current KBA network throughout the Chocó-Darién ecoregion could be important for habitat continuity essential to dispersing Harpy Eagles (Urios et al. 2017) between Central and South America.

The Darién region of Panama (in the north of the Chocó-Darién ecoregion) has a high density of breeding Harpy Eagles and is considered the current stronghold of the species in Central America (Vargas González and Vargas 2011). A small population still exists in the highly deforested Chocó humid forest region of north-
west Ecuador in the south of the Chocó-Darién ecoregion (Zhang 2020). Designating new KBAs in the Chocó-Darién ecoregion corridor could thus sustain habitat for fragmented Harpy Eagle populations, maintaining genetic diversity and thus potential adaptation to environmental change (Lerner et al. 2009; Banhos et al. 2016; Maxwell et al. 2020). Indeed, genetic diversity decreased in fragmented Harpy Eagle populations inhabiting deforested regions of the southern Amazon and Atlantic Forest of Brazil (Banhos et al. 2016), reinforcing the need to protect and link habitat patches throughout its whole distribution.

Habitat loss is a principal threat to the long-term survival of the Harpy Eagle and protecting large areas of tropical forest habitat for the species should be a high priority (Banhos et al. 2016). Continued deforestation resulting in habitat loss and fragmentation across the Harpy Eagle range should raise the alarm about the species' future conservation status. Using targeted forest protection through responsible community land use and broad-scale conservation planning is needed to reduce current deforestation rates (Kramer et al. 1997; Bird et al. 2011; Butchart et al. 2015). While the current KBA network coverage for the Harpy Eagle exceeds the representation target, our models identified gaps in the KBA network that ought to be prioritised for enlarging the KBA network estate. As demonstrated here, our method of calculating modelled Area of Habitat estimates based on SDMs are a useful tool for large-scale conservation planning and can be readily applied to many taxa.

LITERATURE CITED


Shortfalls and solutions for meeting national and global conservation area

Butchart, S.H., Scharlemann, J.P., Evans, M.I., Quader, S., Arico, S., Arinaitwe, J.,
... and Woodley, S. (2012). Protecting important sites for biodiversity contributes
to meeting global conservation targets. *PloS one*. 7: e32529.

Carlson, C.J. (2020). Embarcadero: Species distribution modelling with Bayesian


De Carvalho, D.L., Sousa-Neves, T., Cerqueira, P.V., Gonsioroski, G., Silva, S. M.,
Silva, D.P. and Santos, M.P.D. (2017). Delimiting priority areas for the
conservation of endemic and threatened Neotropical birds using a niche-based

Distribution modeling applied to deficient data species assessment: A case study
with Pithecopus nordestinus (Anura, Phyllomedusidae). *Neotropical Biology and
Conservation*. 15: 165-175.

and trade-offs in the use of species distribution maps for protected area

Important Bird and Biodiversity Areas (IBAs): the development and characteristics
of a global inventory of key sites for biodiversity. *Bird Conservation


KBA Standards and Appeals Committee (2019). *Guidelines for using a Global Standard for the Identification of Key Biodiversity Areas*. Version 1.0. Prepared by the KBA Standards and Appeals Committee of the IUCN Species Survival...
Commission and IUCN World Commission on Protected Areas. Gland, Switzerland: IUCN.


Rodríguez-Soto, C., Monroy-Vilchis, O., Maiorano, L., Boitani, L., Faller, J.C., Briones, M.A., Nunez, R., Rosas-Rosas, O., Ceballos, G. and Falcucci, A.


Sutton, L.J. (2022). Data from: Range-wide habitat use of the Harpy Eagle indicates four major tropical forest gaps in the Key Biodiversity Area network. *Ornithological Applications*. https://doi.org/10.5061/dryad.v41ns1rzt


Title, P.O. and Bemmels, J.B. (2018). ENVIREM: An expanded set of bioclimatic and
topographic variables increases flexibility and improves performance of ecological

biodiversity and ecosystem modelling. *Global Ecology and Biogeography*. 23:
1031-1045.

of terrestrial habitat heterogeneity for biodiversity and ecosystem modelling.

UNEP-WCMC & IUCN (2020). Protected Planet: Brazil; The World Database on
Protected Areas (WDPA). Cambridge, UK. Available at: www.protectedplanet.net/country/BRA


UNEP-WCMC & IUCN, Cambridge, UK. Available at: www.protectedplanet.net


performance of presence-only species distribution models: a benchmark study

and Vargas, F.H. (2020). Predictive Habitat Model Reveals Specificity in a Broadly
Distributed Forest Raptor, The Harpy Eagle. *Journal of Raptor Research*. 54: 349-
363.


https://digitalcollections.sit.edu/isp_collection/3326
Figure 1. Distribution of spatially filtered Harpy Eagle occurrences (black points) across the study extent, showing the relationship to elevation and evergreen forest cover (brown). White borders define national boundaries within the study extent.
Figure 2. Species Distribution Model for the Harpy Eagle. Map denotes cloglog prediction with darker green areas (values closer to 1) having highest suitability and expected abundance. Gray borders define national boundaries within the study extent and internal state boundaries for Brazil. Black points define Harpy Eagle occurrences using a 5-km spatial filter. See Figure S2 in Supplement for map showing cropped model prediction for Central America without Harpy Eagle occurrences for clarity.
Figure 3. Penalized logistic regression response curves for each habitat covariate from the Harpy Eagle Species Distribution Model. The response curves show the contribution to model prediction (y-axis) as a function of each continuous habitat covariate (x-axis). Maximum values in each response curve define the highest predicted relative suitability. The response curves reflect the partial dependence on predicted suitability for each covariate and the dependencies produced by interactions between the selected covariate and all other covariates.
Figure 4. Reclassified binary Species Distribution Model (threshold = 0.345) for the Harpy Eagle. Dark khaki area is habitat above the 0.345 threshold, white areas below the threshold. Red polygons define current IUCN range map for the Harpy Eagle as a comparison to the SDM prediction. Blue hashed polygon represents the Harpy Eagle Extent of Occurrence (EOO) range metric. Gray borders define national boundaries within the study extent and internal state boundaries for Brazil.
Figure 5. Key Biodiversity Area (KBA) network gap analysis for Harpy Eagle habitat. Map denotes cloglog prediction reclassified into four discrete quantile threshold classes (brown = no habitat; yellow = low, pale green = medium; dark green = high). Black bordered polygons denote current KBA network. Blue ellipses identify priority KBA network coverage gaps: (1) Chocó-Darién ecoregion in Colombia, Ecuador and Panama, (2) Magdalena-Urabá moist forests in northern Colombia, (3) north-east Amazonas state in Brazil, (4) north and west Guyana. Gray borders define national boundaries within the study extent and internal state boundaries for Brazil.
Figure 6. Key Biodiversity Area (KBA) network gap analysis for Harpy Eagle habitat projected into the Chocó-Darién ecoregion. Map denotes cloglog prediction reclassified into four discrete quantile threshold classes (brown = no habitat; yellow = low, pale green = medium; dark green = high). Black bordered transparent polygons denote current KBA network. Hashed blue ellipses identify priority KBA network coverage gaps: (1) Chocó-Darién region in Colombia, Ecuador, and Panama, (2) Magdalena-Urabá moist forests in northern Colombia.
SUPPLEMENTARY MATERIAL

Occurrence data

From the 591 filtered occurrences we had 188 eBird records in total, with 57 of these with sampling regime metadata to define as quality records based on checklists with a sampling duration on >5 mins and <240 mins and a distance effort of <5 km (Johnston et al. 2021). We recognise the potential issues this raises with regard to precisely defining the environmental conditions and resources at occurrence points. However, because of the broad scale of our analysis we opted to retain all eBird occurrence data because using just the quality-controlled eBird occurrences would result in less data to build an appropriate continental-scale model. Further, the majority of our occurrence data were sourced from three other datasets that do not contain these sampling protocol data fields but give precise point localities for nests and sightings, rendering these quality checks across our entire dataset obsolete.

Habitat Covariates

Elevation and Terrain Roughness Index (TRI) are both key topographic variables influencing Harpy Eagle distribution (Vargas González and Vargas 2011; Vargas González et al. 2020; Sutton et al. 2021). Elevation was derived from a digital elevation model (DEM) product from the 250m Global Multi-Resolution Terrain Elevation Data 2010 (GMTED2010, Danielson and Gesch 2011). TRI was derived from the 30 arc-sec resolution Shuttle Radar Topographic Mission (SRTM30, Becker et al. 2009). Homogeneity is a biophysical similarity measure closely related to vegetation species richness (i.e., vegetation structure, composition and diversity) derived from textural features of Enhanced Vegetation Index (EVI) between adjacent pixels; sourced from the Moderate Resolution Imaging Spectroradiometer (MODIS,
Homogeneity varies between zero (zero similarity = maximum heterogeneity) and one (complete similarity) to represent the spatial variability and arrangement of vegetation species richness on a continuous scale (Table S1).

The three measures of percentage landcover (Evergreen Forest, Mosaic Forest, Cultivated) are consensus products integrating GlobCover (v2.2), MODIS land-cover product (v051), GLC2000 (v1.1) and DISCover (v2) at 30 arc-sec (~1km) spatial resolution. Mosaic forest is derived from the EarthEnv variable ‘Mixed trees’ and represents a mosaic of mixed forest, shrubland and woody savanna, with cultivated representing a mix of cropland, tree cover and managed vegetation (Table S1). All landcover layers were resampled to a spatial resolution of 2.5 arc-minutes using bilinear interpolation. Full details on methodology and image processing can be found in Tuanmu and Jetz (2014) for the landcover layers, and Tuanmu and Jetz (2015) for the habitat heterogeneity texture measure. All selected covariates showed low collinearity and thus all six were included as predictors in model calibration (Variance Inflation Factor (VIF) < 5; Table S3). Finally, we summarized the environmental range of all habitat covariates used in our models at the species occurrences, pseudo-absences and background region to account for instances of extrapolation (Table S4).

LITERATURE CITED

bathymetry and elevation data at 30 arc seconds resolution: SRTM30_PLUS.


Johnston, A., Hochachka, W.M., Strimas-Mackey, M.E., Ruiz Gutierrez, V.,
Analytical guidelines to increase the value of community science data: An example using eBird data to estimate species distributions. Diversity and Distributions. 27: 1265-1277.


**Table S1.** Habitat covariates used in all spatial modelling analyses for the Harpy Eagle, with citations for the sources of the environmental data used.

<table>
<thead>
<tr>
<th>Covariate</th>
<th>Source</th>
<th>Citation</th>
<th>Resolution</th>
<th>Year(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cultivated (%)</td>
<td>EarthEnv</td>
<td>Tuanmu &amp; Jetz 2014</td>
<td>30 arc secs</td>
<td>1992-2005</td>
</tr>
<tr>
<td>Elevation (m)</td>
<td>EarthEnv</td>
<td>Amatulli et al. 2018</td>
<td>2.5 arc mins</td>
<td>2010</td>
</tr>
<tr>
<td>Evergreen forest (%)</td>
<td>EarthEnv</td>
<td>Tuanmu &amp; Jetz 2014</td>
<td>30 arc secs</td>
<td>1992-2005</td>
</tr>
<tr>
<td>Homogeneity (0.0-1.0)</td>
<td>EarthEnv</td>
<td>Tuanmu &amp; Jetz 2015</td>
<td>2.5 arc mins</td>
<td>2001-2005</td>
</tr>
<tr>
<td>Mosaic forest (%)</td>
<td>EarthEnv</td>
<td>Tuanmu &amp; Jetz 2014</td>
<td>30 arc secs</td>
<td>1992-2005</td>
</tr>
<tr>
<td>Terrain Roughness Index</td>
<td>ENVIREM</td>
<td>Title &amp; Bemmels 2018</td>
<td>30 arc secs</td>
<td>2000</td>
</tr>
</tbody>
</table>

**Table S2.** Model selection metrics for all six candidate models with $\Delta AIC_c < 2$. RM = regularization multiplier ($\beta$), FC = feature classes, LQ = Linear, Quadratic.

<table>
<thead>
<tr>
<th>Model</th>
<th>RM</th>
<th>FC</th>
<th>$AIC_c$</th>
<th>$\Delta AIC_c$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4.0</td>
<td>LQ</td>
<td>7574.316</td>
<td>0.000</td>
</tr>
<tr>
<td>2</td>
<td>3.5</td>
<td>LQ</td>
<td>7574.389</td>
<td>0.070</td>
</tr>
<tr>
<td>3</td>
<td>4.5</td>
<td>LQ</td>
<td>7574.561</td>
<td>0.245</td>
</tr>
<tr>
<td>4</td>
<td>3.0</td>
<td>LQ</td>
<td>7574.785</td>
<td>0.470</td>
</tr>
<tr>
<td>5</td>
<td>5.0</td>
<td>LQ</td>
<td>7575.125</td>
<td>0.809</td>
</tr>
<tr>
<td>6</td>
<td>2.5</td>
<td>LQ</td>
<td>7575.509</td>
<td>1.193</td>
</tr>
</tbody>
</table>
Table S3. Multi-collinearity test using stepwise elimination Variance Inflation Factor (VIF) analysis.

Variables with VIF < 5 have low correlation with other variables, and thus are suitable for inclusion in calibration models when further evaluated for ecological relevance.

<table>
<thead>
<tr>
<th>Covariate</th>
<th>VIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homogeneity</td>
<td>1.65</td>
</tr>
<tr>
<td>Terrain Ruggedness Index</td>
<td>1.76</td>
</tr>
<tr>
<td>Elevation</td>
<td>2.41</td>
</tr>
<tr>
<td>Mosaic forest</td>
<td>2.54</td>
</tr>
<tr>
<td>Cultivated</td>
<td>2.62</td>
</tr>
<tr>
<td>Evergreen forest</td>
<td>4.64</td>
</tr>
</tbody>
</table>

Table S4. Environmental range of habitat covariates at species occurrences, pseudo-absences and the background region used in Species Distribution Models for the Harpy Eagle. Values are mean (min-max).

<table>
<thead>
<tr>
<th>Covariate</th>
<th>Occurrences</th>
<th>Pseudo-absences</th>
<th>Background region</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cultivated (%)</td>
<td>9 (0-70)</td>
<td>20 (0-99)</td>
<td>20 (0-100)</td>
</tr>
<tr>
<td>Elevation (m)</td>
<td>245 (3-2336)</td>
<td>538 (0-5368)</td>
<td>550 (0-5850)</td>
</tr>
<tr>
<td>Evergreen forest (%)</td>
<td>77 (0-100)</td>
<td>48 (0-100)</td>
<td>48 (0-100)</td>
</tr>
<tr>
<td>Homogeneity (0-1)</td>
<td>0 (0-1)</td>
<td>0 (0-1)</td>
<td>0.4 (0.1-1)</td>
</tr>
<tr>
<td>Mosaic forest (%)</td>
<td>5 (0-54)</td>
<td>13 (0-78)</td>
<td>13 (0-83)</td>
</tr>
<tr>
<td>Terrain Roughness Index (0-Inf)</td>
<td>22 (0-217)</td>
<td>27 (0-586)</td>
<td>27 (0-615)</td>
</tr>
</tbody>
</table>
**Figure S1.** Distribution of random background points ($n = 10,000$, gray points) across the study extent used as pseudo-absences in Species Distribution Models for the Harpy Eagle. Red points denote spatially filtered Harpy Eagle occurrences.
Figure S2. Cropped Species Distribution Model for the Harpy Eagle across Central America. Map denotes cloglog prediction with darker green areas (values closer to 1) having highest suitability and expected abundance. Gray borders define national boundaries.