Science Synthesis, Analysis and Research Program

Prepared in cooperation with North Carolina State University, North Carolina Cooperative Fish and Wildlife Research Unit, Department of Applied Ecology

Methods for Evaluating Gap Analysis Project Habitat Distribution Maps with Species Occurrence Data

Techniques and Methods 2-A19

U.S. Department of the Interior
U.S. Geological Survey

Methods for Evaluating Gap Analysis Project Habitat Distribution Maps with Species Occurrence Data

By Matthew J. Rubino, Alexa J. McKerrow, Nathan M. Tarr, and
Steven G. Williams

Science Synthesis, Analysis and Research Program

Prepared in cooperation with North Carolina State University, North Carolina Cooperative Fish and Wildlife Research Unit, Department of Applied Ecology

Techniques and Methods 2-A19

U.S. Department of the Interior
U.S. Geological Survey

U.S. Geological Survey, Reston, Virginia: 2022

Suggested citation:
Rubino, M.J., McKerrow, A.J., Tarr, N.M., and Williams, S.G., 2022, Methods for evaluating Gap Analysis Project habitat distribution maps with species occurrence data: U.S. Geological Survey Techniques and Methods 2-A19, 13 p., https://doi.org/10.3133/tm2A19.

Associated data:
U.S. Geological Survey [USGS] – Gap Analysis Project [GAP], 2018a, Black-bellied Salamander (*Desmognathus quadramaculatus*) aBESAx_CONUS_2001v1 Habitat Map: U.S. Geological Survey data release, https://doi.org/10.5066/F7H1308B.

ISSN 2328-7055 (online)

Acknowledgments

We would like to acknowledge the work of the Global Biodiversity Information Facility, Biodiversity Serving Our Nation, and eBird for their commitment to citizen science and making species occurrence data accessible.

Contents

Figures

Tables

Conversion Factors

International System of Units to U.S. customary units

Multiply	By	To obtain
	Length	
kilometer (km)	0.6214	mile (mi)
meter (m)	3.281	foot (ft)
	Area	
square kilometer (km^2)	247.1	acre
square kilometer (km^2)	0.3861	square mile (mi^2)
square meter (m^2)	0.0002471	acre

Datum

Vertical coordinate information is referenced to the North American Vertical Datum of 1988 (NAVD 88).

Horizontal coordinate information is referenced to the North American Datum of 1983 (NAD 83).

Altitude, as used in this report, refers to distance above the vertical datum.

Abbreviations

BiSON	Biodiversity Serving Our Nation
BPA	buffer proportion assessment
CONUS	conterminous United States
GAP	Gap Analysis Project
GBIF	Global Biodiversity Information Facility
USGS	U.S. Geological Survey

Methods for Evaluating Gap Analysis Project Habitat Distribution Maps with Species Occurrence Data

By Matthew J. Rubino,[1] Alexa J. McKerrow,[2] Nathan M. Tarr,[1] and Steven G. Williams[1]

Abstract

The National Gap Analysis Project created species habitat distribution models for all terrestrial vertebrates in the United States to support conservation assessments and explore patterns of species richness. Those models link species to specific habitats throughout the range of each species. For most vertebrates, there are not enough occurrence data to drive inductive, range-wide species habitat distribution models at high spatial and thematic resolution. However, it is possible to use occurrence data for model evaluation. The combination of citizen science, formal species survey work, and digitized specimen archives are making millions of observations available to the scientific community. Our challenge is to combine the mostly unstructured data into metrics that help us characterize and understand patterns of biodiversity. In this work, we propose two model-evaluation metrics. The first, a buffer proportion assessment, is based on the proportion of habitat in the range relative to the mean proportion of habitat around each of the species' occurrence records. The second is a measure of the sensitivity (proportion of true presence) to buffer distances around occurrence records. The buffer proportion is a modification of model prevalence versus point prevalence metric, whereby comparison to a null model allows us to determine if the model performs better or worse than random.

In this report, we describe the workflow used to compile and filter the species occurrence records from online resources (for example, the Global Biodiversity Information Facility) and show results for a single species, *Desmognathus quadramaculatus* (black-bellied salamander). For the salamander, 222 occurrence points met our criteria for inclusion in the evaluation. We found the model performed better than random with a buffer proportion index of 1.745, indicating about 5 times as much habitat was found adjacent to known occurrence records than would be expected from randomly located sites throughout the range. Sensitivity increased with larger buffer distances and leveled off to around 0.7 between 1,000- and 2,000-meter buffer distances, indicating the model is likely best suited for scales exceeding 1,000 meters. We plan to report the buffer proportion assessment and sensitivity metrics along with the full species model reports to increase understanding of the model's performance and to use the metrics to help prioritize revisions to the models.

Introduction

The National Gap Analysis Project (GAP) works to quantify habitat availability for terrestrial vertebrate species and the extent of those habitats that are protected from conversion. Specifically, the GAP models species' habitats to identify geographic areas of under-representation, or in other words, where species' habitats are not well represented in the conservation network. To address the question of representation, the GAP uses deductive modeling to map habitat affinities for terrestrial vertebrate species (Gergely and others, 2019). The approach is founded on the fact that habitat is a primary limiting factor for species and that habitat loss is a major threat to species (Díaz and others, 2019). GAP data can identify areas for conservation and restoration to help ensure habitats remain available.

For the conterminous United States (CONUS), 1,590 vertebrate species models (282 amphibians, 621 birds, 365 mammals, and 322 reptiles) have been created. We constructed subspecies models when a geographically distinct range for that subspecies could be mapped. A total of 129 subspecies (2 amphibians, 28 birds, 94 mammals, and 5 reptiles) have been modeled to date. The models are based on a literature review to support the development of a habitat affinities database. Additional ancillary parameters are used to constrain the model when supported in the literature (table 1). The output of the modeling process is a map of potentially suitable habitat across a landscape restricted to the species' known range within CONUS, or the geographical extent of occurrence (Loiselle and others, 2003; Graham and Hijmans, 2006) (fig. 1). Given our focus on habitat, the GAP has consciously chosen to stay with a deductive habitat modeling approach, which allows us to create models that are consistent across all four terrestrial vertebrate taxa—amphibians, birds, mammals, and reptiles.

---

[1] North Carolina State University, North Carolina Cooperative Fish and Wildlife Research Unit, Department of Applied Ecology.

[2] U.S. Geological Survey.

Table 1. Sources of data for ancillary parameters used in amphibian species models.

[See Gergely and others (2019, appendix 1) or the uniform resource locators (URLs) for a complete description of each ancillary dataset. GAP, Gap Analysis Project]

Ancillary data	URL for the source data
National GAP land cover ver. 1.0 (2001)	https://doi.org/10.5066/F7959GF5
Hydrologic unit codes (HUCs)	https://doi.org/10.5066/F7DZ0754
Modeling regions	https://doi.org/10.5066/F77H1HGT
Human impact avoidance	https://doi.org/10.5066/F7PC318R
Forest edge	https://doi.org/10.5066/F7XW4HPN
Forest/open and woodland/shrubland	https://doi.org/10.5066/F7T43RZ7
Elevation	https://doi.org/10.5066/F72N515B
Slope	https://doi.org/10.5066/F75D8QQF
Hydrography	https://doi.org/10.5066/F7JM28J1

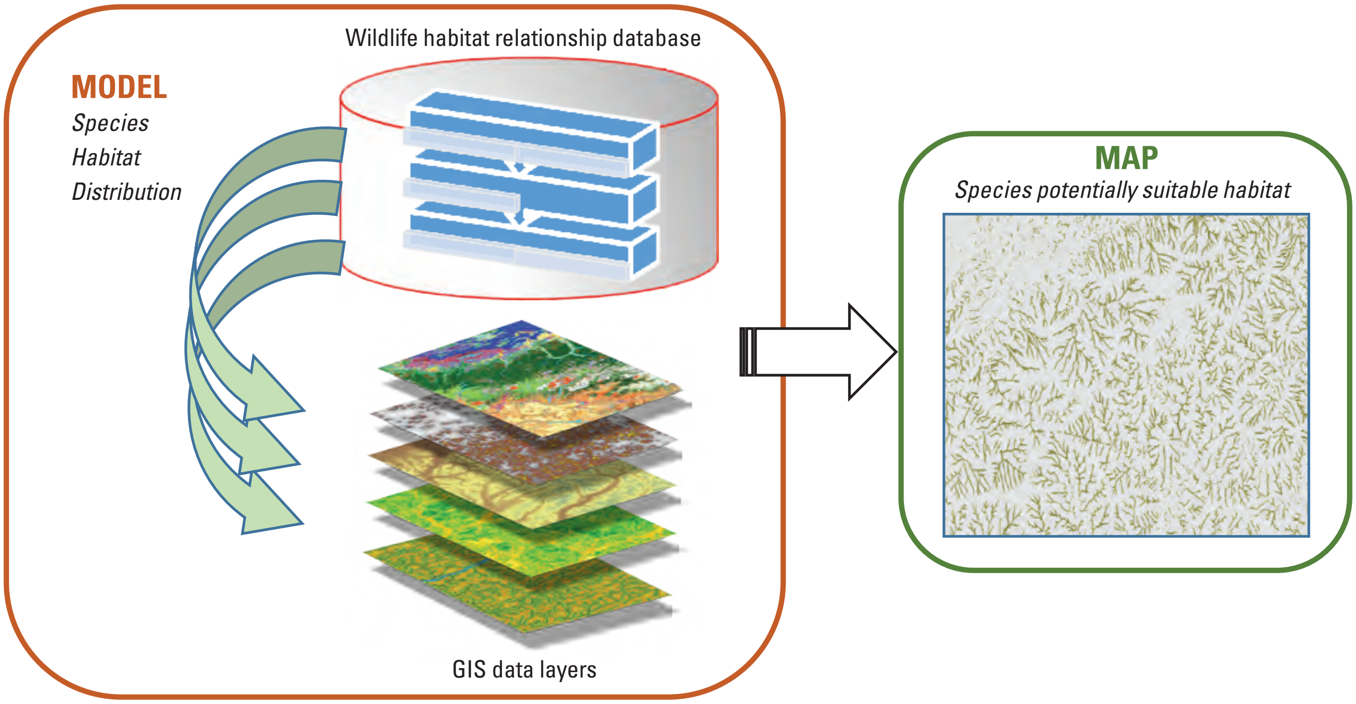


Figure 1. Graphical representation of Gap Analysis Project species distribution modeling process and the resulting output map of potentially suitable habitat. (GIS, geographic information system)

The two major uses of the habitat distribution maps include gap representation analysis and studying patterns of species richness (Gergely and others, 2019). Species richness and representation analyses are sensitive to the spatial patterns and extent of the species distribution maps (Rondinini and others, 2006; McKerrow and others, 2018), so evaluation is critical to improving our understanding of terrestrial vertebrate biodiversity in the United States.

Background on Model/Map Evaluation

Modeling species habitat is an iterative process of gathering relevant data, selecting process methods, determining appropriate parameters, and generating useable outputs (for example, geographic maps). The interrelated nature of this workflow means that evaluation of outputs is, by extension, an evaluation of the procedures used to derive those outputs. Therefore, throughout this document, we refer to evaluation and assessment of models and maps interchangeably.

Previous Evaluation Approaches in State and Regional Gap Analysis Projects

Initially undertaken at the State level, Gap Analysis Projects' data development varied and was inconsistent. Recognizing the need for consistent datasets that span administrative boundaries and taking the lessons learned in State projects, the GAP implemented a series of regional projects and eventually a nationwide effort, taking advantage of the burgeoning availability of national data (for example, the National Land Cover Dataset, National Elevation Dataset, National Hydrography Dataset, and more) and increases in computing power. As the mapping extent expanded from state to regional to national, implementation of meaningful evaluations of species' habitat distribution maps became increasingly challenging, partly from the poor quality and limited quantity of occurrence data and that expert knowledge varied greatly across species and geographies. As a core data product of the GAP mission, meaningful evaluations of species' habitat distribution maps are critical to the data's use and implementation.

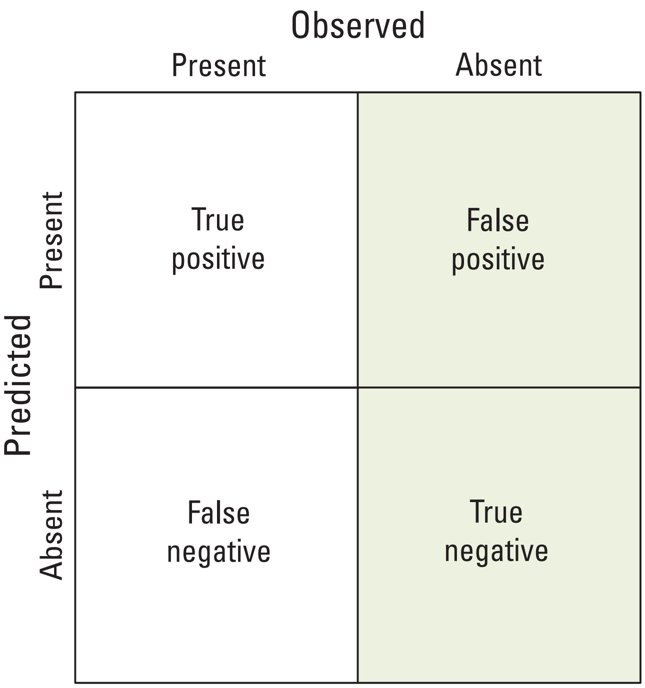
A common approach among the early state-based projects included comparing published checklists with the GAP species habitat distribution maps (Edwards and others 1996; U.S. Geological Survey [USGS] 1996, 2000, 2001a, 2003, 2004, 2005b, 2006b). The Arkansas GAP used the Breeding Bird Survey data to assess maps for birds that are well surveyed by that effort (USGS, 1998), and the Oklahoma GAP used the Breeding Bird Atlas data to assess maps for a subset of the birds modeled (USGS, 2001b). Most States relied on internal and external expert review of the habitat affinities, whereas a few relied on occurrence data from sources other than the Breeding Bird Survey or the Breeding Bird Atlas. For example, in Hawaii, forest bird survey data provided a robust dataset to evaluate forest bird models (USGS, 2006a). The Nebraska GAP used occurrence records to model species habitat using a recursive partitioning model to develop habitat distribution maps and used independent species occurrence records for evaluating the models for species with at least 10 occurrences (USGS, 2005a). Not surprisingly, data availability made it possible to assess bird species most often, whereas data for mammals, amphibians and reptiles was often lacking.

The Southwest Regional GAP effort relied primarily on expert review workshops to evaluate species models. In addition to the workshops, the physical features used in constructing the species models was assessed by visiting 383 sites predicted to be habitat for at least 1 of 30 species (Boykin and others, 2007, 2010). The species were selected to represent habitat specialists and generalists across the four taxonomic groups. For each feature, they quantified the level of concordance between their field observations and the spatial datasets used in the modeling. Features that could be directly measured (for example, slope and soil composition) tended to have higher concordance than features requiring more subjective judgements (for example, ecological system classification and landforms). While that assessment does not tell us about the accuracy of the assignment of habitat features to an individual species, it does provide important information on sources of uncertainty in the spatial datasets used in the deductive modeling approach.

The Pacific Northwest and Alaska Gap Analysis Projects used both expert review and accuracy assessment with presence-only occurrence data (Aycrigg and others, 2015). The two projects implemented a hybrid modeling approach combining a deductive model of habitat affinities and ancillary data to identify suitable habitat, along with an inductive model utilizing occurrence data. Specifically, they used the Maxent modeling algorithm (Maxent version 3.3; Phillips and others, 2006; Phillips and Dudik, 2008) to develop the inductive models and developed a deductive model based on habitats and parameters identified in the literature. The results of those two models were then intersected to create a combined hybrid model for that species. Occurrence records were carefully screened for quality and duplicate records, and spatially imprecise records and records with uncertain taxonomy were removed. A subset of the filtered dataset was used in training the Maxent model (76–80 percent) and the remaining records were set aside for assessment. For each species modeled, the expert review and percent correct classification (the percent of the occurrence that fell on modeled habitat) were reported. For the inductive models, the area-under-curve metric derived by the Maxent software was reported. For the Pacific Northwest GAP the final model was selected from either the deductive or combined model (the intersection of the deductive and inductive models) based on the classification success (how often the occurrence points intersected modeled habitat). For species where greater than (>) 75 percent of the known range did not have modeled habitat in the inductive model, the deductive model was chosen. For the Alaska Gap, a series of workshops were held, where experts chose the final model based on their review. For species without an expert review, the classification success was used to select the final model between the deductive, inductive, and combined models.

Evolving Access to Species Occurrence Data

Recently, tremendous progress has been made to mobilize species occurrence data from a wide range of sources (Nelson and Ellis, 2018). At the same time, participation in citizen science has increased at a fantastic pace along with methods to improve and enhance data collection, access, and consistency (Guralnick and Constable, 2010; Theobald and others, 2015). Challenges in using these data sources are that they generally represent uneven sampling across space and time, spatial precision of occurrence record locations can be low or nonexistent, and they represent presence-only data. These issues necessitate careful filtering and quality assurance checks to identify a useful subset of records (Aycrigg and others, 2015; Jin and Yang, 2020). Most of the species occurrence datasets are the result of opportunistic collection (for example, museum records or citizen science efforts) and therefore provide data only on species presence and not information on true absence (fig. 2).



		Observed: Present	Observed: Absent
Predicted	Present	True positive	False positive
Predicted	Absent	False negative	True negative

Figure 2. Illustrates the terminology applied in traditional assessments, when the observations of a species' occurrence are compared to the predictions. Information derived from the left column relates to sensitivity, the proportion of true positives. In other words, the number of times the species was observed in locations where habitat was predicted to occur relative to the total number of presence observations. The left column is also necessary to assess omission error and the number of false negatives (a species was present where no habitat was predicted to occur). Information from the right column relates to specificity, the proportion of true negatives and commission error – the number of times a species was not found but habitat was predicted (false positive). Without absence data, the ability to identify areas where the species does not occur (right column) cannot be quantified.

Presence-Only Model Evaluations

We found several studies where the application of presence-only data for model evaluation has been tested (Ottaviani and others, 2004; Hirzel and others, 2006; Pearson, 2007; Rondinini and others, 2011). In each case, the presence data were compared against a null model to estimate the likelihood of the points occurring in a habitat by chance. Rondinini and others (2011) compared model prevalence (the proportion of the landscape modeled as suitable) to the point prevalence (the number of correctly predicted points). Hirzel and others (2006) found that the three presence-only evaluators they tested correlated well to presence/absence data evaluators ($\rho > 0.7$). In that work, they recommend using the Boyce index (Boyce and others, 2002) in cases where the model output has multiple levels of suitability as opposed to a binary suitable/unsuitable output. Two frequencies are calculated: a predicted frequency, which is the number of evaluation points the model predicts to fall in a given habitat suitability class; and expected frequency, which is a frequency expected by random. The index is calculated for each habitat suitability class as the ratio of predicted to expected frequencies, where higher suitability classes should have Boyce indices greater than one. Ottaviani and others (2004) proposed two approaches: (1) compositional analysis for models where the presence data have spatial and temporal variability; and (2) multinomial analyses where the presence-only data have high precision. Compositional analysis summarizes the proportion of each of the categories of habitat suitability in an area around the presence points and compares that composition to the null model.

Methods

Here we detail a method for evaluating binary (suitable or unsuitable) habitat distribution maps using presence-only species occurrence data. We built on the concepts of evaluating maps against the null model (Pearson, 2007; Rondinini and others, 2011) and developed two metrics: Buffer Proportion Assessment (BPA), and sensitivity at varying buffer distances around occurrence locations. In this chapter, we present the metrics, workflow, and code to support the analyses for a single species, *Desmognathus quadramaculatus* (Holbrook, 1840) (black-bellied salamander). Ultimately, we plan to report these metrics for each of the GAP species habitat distribution maps where there is sufficient occurrence data and to update the evaluation metrics as the models are updated.

Species Habitat Distribution Modeling

We created 284 habitat distribution models (282 species and 2 subspecies) for amphibians across CONUS (Gergely and others, 2019). The species list for the project was based on the Society for the Study of Amphibians and Reptiles authoritative list (Crother, 2008). To demonstrate our approach to map evaluation, we chose to focus on the black-bellied salamander because it represented a widespread species with known range throughout the mountains from northern Georgia to central West Virginia. The species is a member of the family Plethodontidae and according to the Integrated Taxonomic Information System (ITIS, 2020) the taxonomy has not changed since our national list was compiled based on Crother (2008).

Based on the literature review, the black-bellied salamander was attributed to 24 detailed land cover types within its known range; specifically, montane forests and riverine forests (for example, the Southern Appalachian Low Mountain Pine Forest or the Southern Piedmont Small Floodplain and Riparian Forest). The known range for the species was delineated using a modified version of the 12-digit Hydrologic Unit Dataset (USGS and USDA NRCS, 2009), where hydrologic units were selected to represent the approximate geographic distribution of the species based on a variety of sources (for example, Conant and Collins, 1998; Petranka, 1998; Lannoo, 2005). Attribution for each hydrologic unit within the range included values for origin (native, introduced, reintroduced, or vagrant), occurrence (extant, possibly present, potentially present, or extirpated), reproductive use (breeding, nonbreeding, or both), and season (year-round, summer, winter, migratory, or vagrant). In addition to the land cover affinities, distance to flowing water (30 meters [m]) and an elevation range (0–1,646 m) were applied to constrain the mapping of suitable habitat. The model parameters, literature sources, and map of range and modeled habitat are provided in the species model report, available on the GAP data download page (https://gapanalysis.usgs.gov/apps/species-data-download/). The data developed for the species range (USGS, 2018b) and habitat distribution model (USGS, 2018a) and the species model report are publicly available.

Occurrence Records

We chose to use occurrence records from the Global Biodiversity Information Facility (GBIF) for the initial evaluation of our species habitat distribution models. GBIF mobilizes data from a wide range of data providers globally, and the U.S. Geological Survey (USGS) serves as the host to the U.S. node for GBIF. We used the Wildlife Wrangler, a framework for managing occurrence data and documenting the decisions used to include and filter individual records (Tarr and others, 2021). A search of GBIF occurrence records was conducted through the GBIF Application Programming Interface (API; Chamberlain, 2020) and the results were brought into the Wildlife Wrangler for processing (fig. 3). A Jupyter Notebook (Kluyver and others, 2016) is used to document the criteria used in filtering the dataset, and a SQLite database is created with the final subset of occurrence data.

In this study, the criteria for filtering the GBIF occurrence records include the following:

1. The date must match the taxonomic concept that was modeled. If the taxonomy of the species had changed after the initial species model was completed, only records before the taxonomic change occurred were included in the evaluation.

2. Remove duplicate records based on date, location, and taxon concept. To account for times when the data aggregator receives the same occurrence record from multiple providers, the records are summarized, and where date, location, and the species are the same, a single record is captured for that observation.

3. Record must be georeferenced and have a coordinate uncertainty. All three fields—latitude, longitude, and locational uncertainty—are required for a record to be included.

4. Coordinate uncertainty must be less than or equal to (≤) 10,000 m. Records with greater than 10 kilometers (km) spatial uncertainty were excluded.

5. A filtered dataset must have greater than or equal to (≥) 20 records per species. The species habitat distribution model was only evaluated if there were at least 20 occurrence records for a species after the initial four criteria were applied.

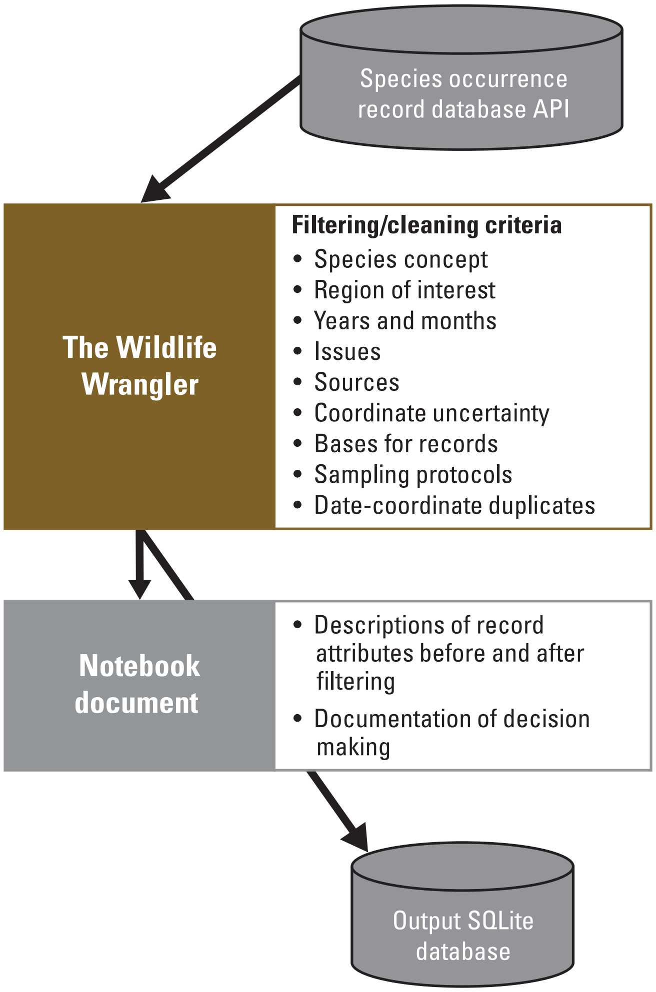


Figure 3. Wildlife Wrangler framework for documenting the filtering and storing of species occurrence records. (API, application programming interface)

Evaluation Metrics

We used the final set of species occurrence records for each species to develop two metrics, the Buffer Proportion Assessment (BPA) and Sensitivity. BPA provides a measure of the amount (proportion) of habitat that is modeled near the known occurrence records for the species. Sensitivity is a measure of how often the occurrence records coincide with modeled habitat.

Buffer Proportion Assessment (BPA)

Rondinini and others (2011) created deductive habitat suitability models (high, medium, unsuitable) for more than 5,000 terrestrial mammals globally at a spatial grain of 300 meters to examine patterns of richness and protection. They conducted model evaluation on a subset of only 263 species because detailed occurrence data are lacking for most mammal taxa. The occurrence data used contained information only on species' presence, limiting assessment methodologies. Therefore, Rondinini and others (2011) chose to compare what they termed "model prevalence" and "point prevalence" to species' landscapes of randomly distributed habitat at a spatial grain of 1 square kilometer (km^2) (the assumed accuracy of their point locality data). Point prevalence is defined as the proportion of 1 km^2 cells containing one or more occurrences, and one or more 300 square meters (m^2) cells with high or medium suitability (the proportion of correctly predicted occurrences). Model prevalence is defined as the proportion of 1 km^2 cells with high or medium suitability within the species' range. As Rondinini and others (2011, p. 2,635) describe: "If predicted suitable cells were distributed randomly with respect to species occurrences, the expectation would be that the point prevalence equals the model prevalence." Hence, the comparison assumes a random habitat landscape is one in which any given location contains habitat at proportions equal to the modeled, predicted habitat across the entire species' range. When point prevalence is greater than model prevalence, the model predicts habitat at species locations better than the range "because it tends to correctly flag as unsuitable the cells that do not contain occurrences" (Rondinini and others, 2011, p. 2,635).

GAP species maps include geographic range along with distribution of potential habitat generated at a spatial resolution of 30 m^2 (fig. 4). Therefore, each species has a measurement of model prevalence, which is simply the areal proportion of predicted habitat within the delineated range. However, to calculate point prevalence at a 30 m^2 spatial grain, point occurrences need to have an equivalent spatial accuracy. Point prevalence calculations at a 30 m^2 spatial grain will be available only for a very limited number of species and with very few data points because most point occurrence datasets (for example, GBIF, 2020; BiSON, 2020; eBird – see Sullivan and others, 2009) have either no information regarding individual location accuracy or vary widely in coordinate uncertainty (for example, meters to province-level centroids). Buffering point occurrences by the degree of uncertainty of presence locations and calculating the proportion of habitat within those buffers allow for comparisons to random habitat landscapes based on the range extent. As with point prevalence versus model prevalence, models that predict habitat better than random would have a mean proportion of habitat in point buffers greater than the proportion of habitat throughout the species' range. Figure 5 shows how the proportion of habitat in the range (model prevalence) compares to the proportion of habitat in the buffered occurrence points (point prevalence). The null model is represented by the 1:1 line where the proportions are equal.

We used the filtered occurrence points and their locational uncertainty to quantify the proportion of habitat in the buffer around each point (fig. 6). We compared the mean of all occurrence location habitat proportions with the proportion of habitat within the entire range (prevalence ratio). Figure 4 shows the known range map and buffered occurrence data for the black-bellied salamander.

We calculated and graphed the prevalence ratio for the black-bellied salamander model based on the approach described above. In addition, we transformed the ratio into a single numerical metric by calculating an index as follows:

$$Index_{BPA} = \ln\left(\frac{mean\ proportion\ habitat\ in\ point\ buffers}{proportion\ habitat\ in\ range}\right) \quad (1)$$

and

$$Index_{BPA} > 0: model\ is\ better\ than\ random$$

$$Index_{BPA} < 0: model\ is\ worse\ than\ random$$

Positive index values indicate model predictions above the 1:1 line (better than random) while negative values are below the 1:1 line (worse than random). If changes are made to the model, index comparisons can provide a quantitative measure of model improvement.

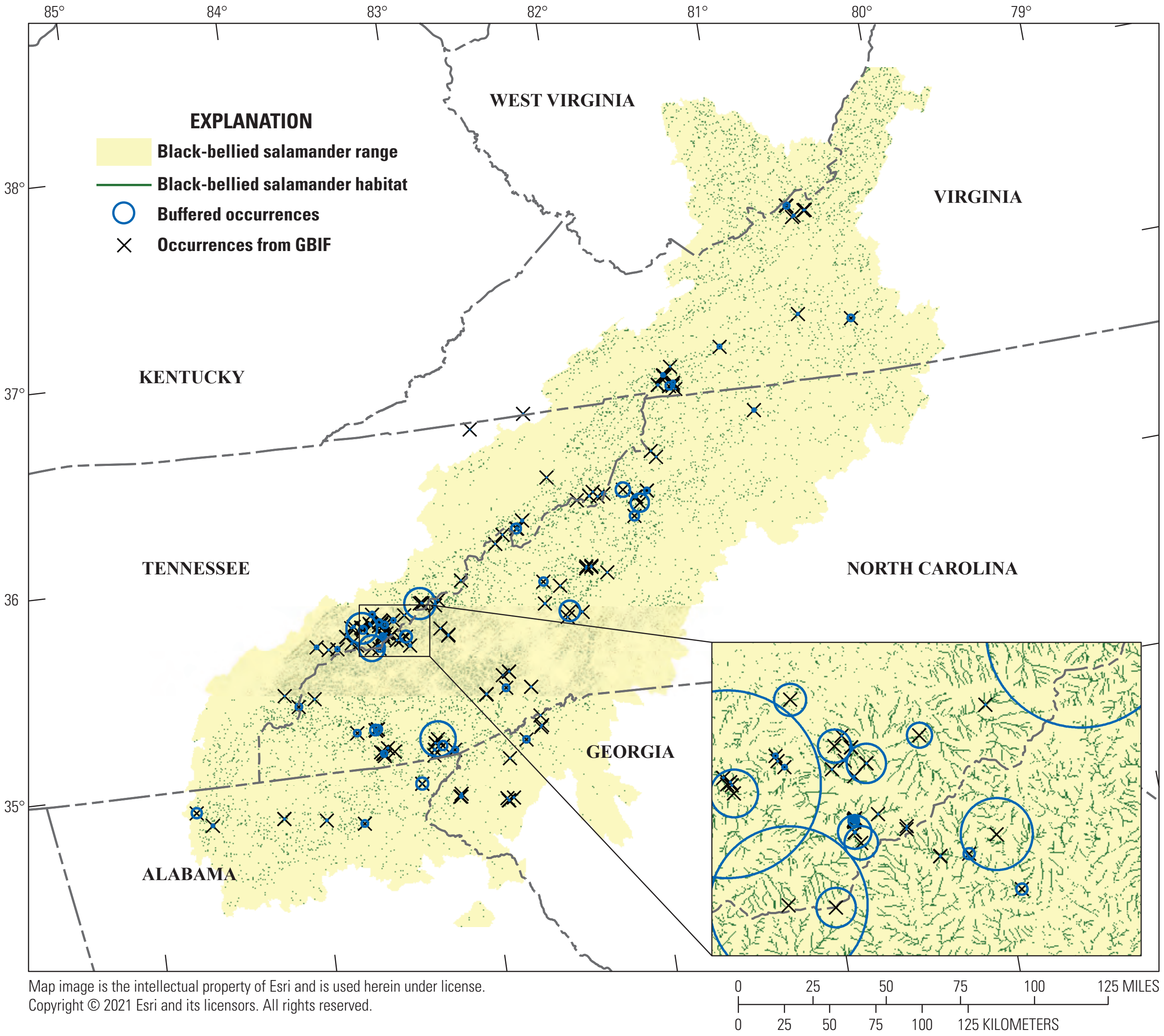
Sensitivity

The accuracy assessment measure, or sensitivity, is also referred to as the true positive fraction in model validation. For habitat map assessment, it is a measure of locations correctly predicted to be habitat. Sensitivity is calculated as the count of true positives divided by the sum of true positive plus false negative counts:

$$Sensitivity = \frac{True\ positive}{True\ positive + False\ negative} \quad (2)$$

Since the occurrence points we are using vary in their coordinate uncertainty, we were interested in seeing if the proportion of true positives was sensitive to the size of the buffers used to evaluate the map. Minimum assessment scale for each occurrence is the coordinate uncertainty because the “true” location cannot be resolved below that uncertainty distance. Since the map resolution is constant (30 m^2), map assessment is conducted at varying scales using the occurrence locations uncertainty measure (buffer size). If sensitivity increases at increasing buffer distances (that is, increasing assessment scale), it could indicate the scale at which the maps are useful given a subjectively set minimum sensitivity value (for example, 0.6). Illustrating this variation across several scales (buffer sizes) can help map users determine which scale is most suitable for their intended use.

Using occurrence points in GAP habitat map evaluation, a true positive is recorded when at least one predicted habitat cell occurs within an occurrence point uncertainty buffer. A false negative is recorded when no habitat cells occur within an occurrence point buffer. A sensitivity measure can be calculated for increasing buffer distances where a cumulative count of true positive (at least one habitat cell) and false negative (no habitat cells) are made for buffer sizes less than or equal to one of several size categories (for example 30, 100, 500, 1,000, 2,000, 5,000, and 10,000 meters). High sensitivity is indicative of a low omission rate or false positive fraction because sensitivity and omission rate sum to one.



Map image is the intellectual property of Esri and is used herein under license.
Copyright © 2021 Esri and its licensors. All rights reserved.

Figure 4. Map of the potential habitat (green) within the range (yellow) of *Desmognathus quadramaculatus* (black-bellied salamander) with filtered occurrence records (black crosses) from the Global Biodiversity Information Facility archive shown. The size of the blue circles represents the spatial uncertainty of the occurrence records. (GBIF, Global Biodiversity Information Facility)

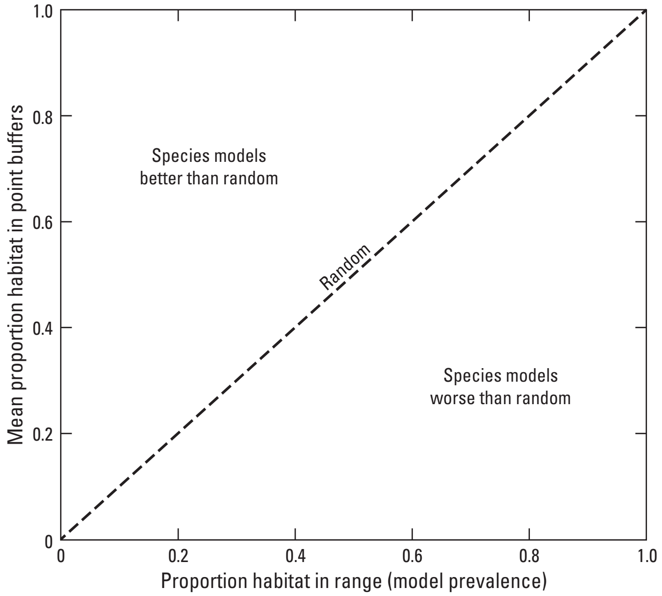


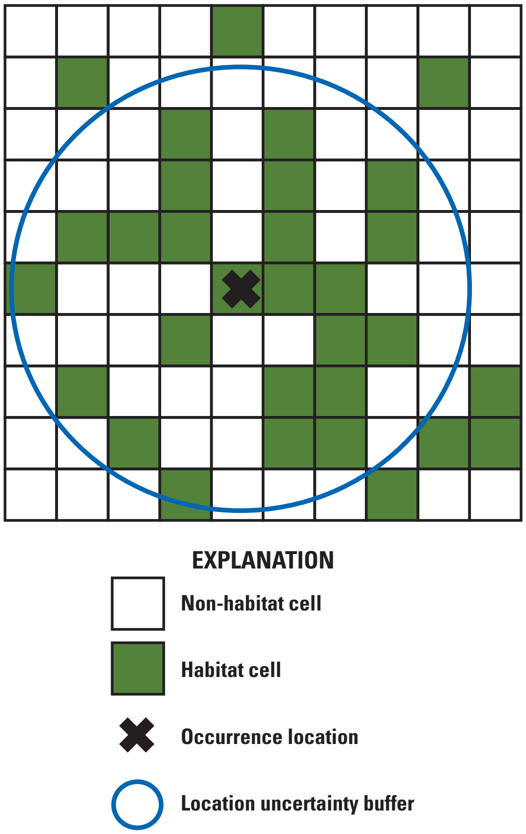
Figure 5. Graphic representation of the buffer proportion assessment (BPA) plot space. For each species, the model-derived map of potential habitat is compared to a map of randomly distributed habitat.

Statistical significance tests

To test whether the map of habitat predictions is greater than would be expected by chance, we need to test for statistical significance. Statistical significance tests can be used to evaluate the relationship between sensitivity and model prevalence (proportion of habitat in the species range) using an exact one-sided binomial test (Anderson and others, 2002; Pearson 2007). Because sensitivity can be calculated for each buffer distance category, a test can be performed for each category.

Data and Code Availability

We developed a code-based approach to assemble the data using a transparent and repeatable process, in keeping with the scientific community's findable, accessible, interoperable, and reusable (FAIR) principles (Wilkinson and others, 2016). Our approach automates the major steps of the filtering and evaluation processes. Specifically, we used the Wildlife Wrangler framework (Tarr and others, 2021) to document the decisions made when the occurrence records are being filtered and created a Jupyter Notebook specific to the occurrence records used for evaluating the species models. Code used to calculate the metrics used in our evaluation are archived as well (https://github.com/mjrubino/GAP-Habitat-Map-Assessment/blob/master/Habitat-Map-Evaluation.ipynb).



$$prevalance\ ratio = \frac{23}{23 + 37} = 0.383$$

Figure 6. Example of the basic approach to calculating prevalence ratio: the proportion of habitat (green) versus non-habitat (white) cells in a buffer surrounding an occurrence location (black cross) defined by coordinate uncertainty (blue circle). In this case, 23 cells are considered habitat and 37 are non-habitat. Cells with less than (<) 50 percent of their area inside the buffer are not used in the calculation.

Results

GBIF occurrences for the black-bellied salamander were downloaded and filtered using the Wildlife Wrangler framework and included the records from the following sources: North Carolina Museum of Natural Sciences Herpetology Collection, Smithsonian Museum of Natural History, and research-grade observations from iNaturalist. When records were filtered to only those collected in the United States, a total of 9,119 observations were available for downloading. Additional filtering per the five-stage Wildlife Wrangler framework flow (fig. 3) outlined in the "Occurrence Records" section decreased the number of acceptable records as detailed in table 2.

There were 222 occurrence points that met the criteria for inclusion in the species model evaluation (GBIF, 2022). The prevalence ratio plot evaluating the proportion of habitat throughout the species range (0.040) relative to the mean proportion of habitat within occurrence point buffers (0.229) indicates that the model is better than random in predicting habitat adjacent to species locations (fig. 7). On average, the model predicts greater than five times more habitat adjacent to species locations than across the entire species' range. The BPA index is 1.745, again indicating the model predicts habitat better than random across the range.

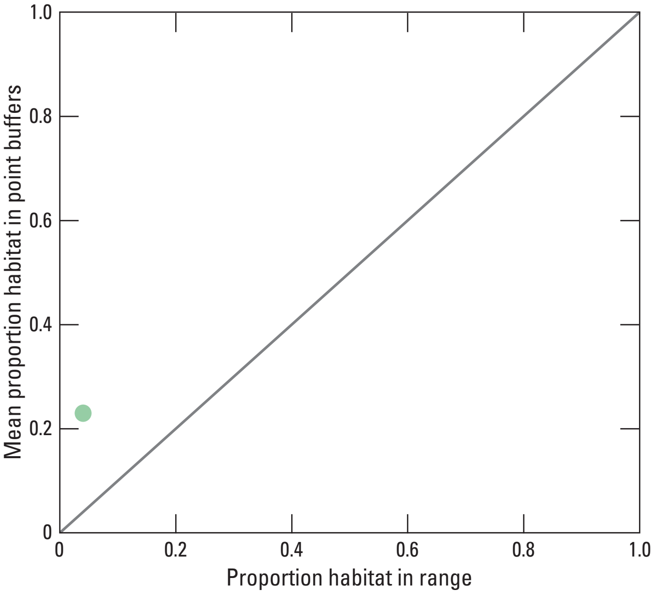


Figure 7. Graph of the prevalence ratio for *Desmognathus quadramaculatus* (black-bellied salamander). The marker falls in the area above the 1:1 line, which indicates that the model performed better than random. For this species, 222 Global Biodiversity Information Facility occurrence records were used in the assessment.

Sensitivity by Buffer Distance

Sensitivity, or true positive fraction, was calculated across seven buffer distances for the black-bellied salamander and used to graph the sensitivity along with the cumulative number of points used to assess the species' model. A test of statistical significance was performed for each buffer distance category to calculate the probability of obtaining the true positive fraction relative to random using an exact one-sided binomial test (Anderson and others, 2002; Pearson 2007). Figure 8 illustrates the buffer sensitivity metric for black-bellied salamander.

Sensitivity (true positive fraction) of point locations at all but one buffer distance is greater than 0.5., except at 100 m where the sensitivity is 0.489. P-values of exact one-sided binomial tests for each buffer distance category indicate that sensitivity values are all significantly different than would be expected from random locations.

Table 2. Filtering criteria, record counts removed, and records remaining for the black-bellied salamander from the Global Biodiversity Information Facility repository collected in the United States (9,119 records available).

[m, meters; ≥, greater than or equal to]

Filter	Number of records removed	Number of records remaining
1 – Date (1999–2020)	8,314	805
2 – Duplicates removed	23	782
3 – Georeferenced	247	535
4 – Coordinate uncertainty (≤ 10,000 m)	313	222

Black-bellied salamander (*Desmognathus quadramaculatus*). Photograph by the U.S. Geological Survey.

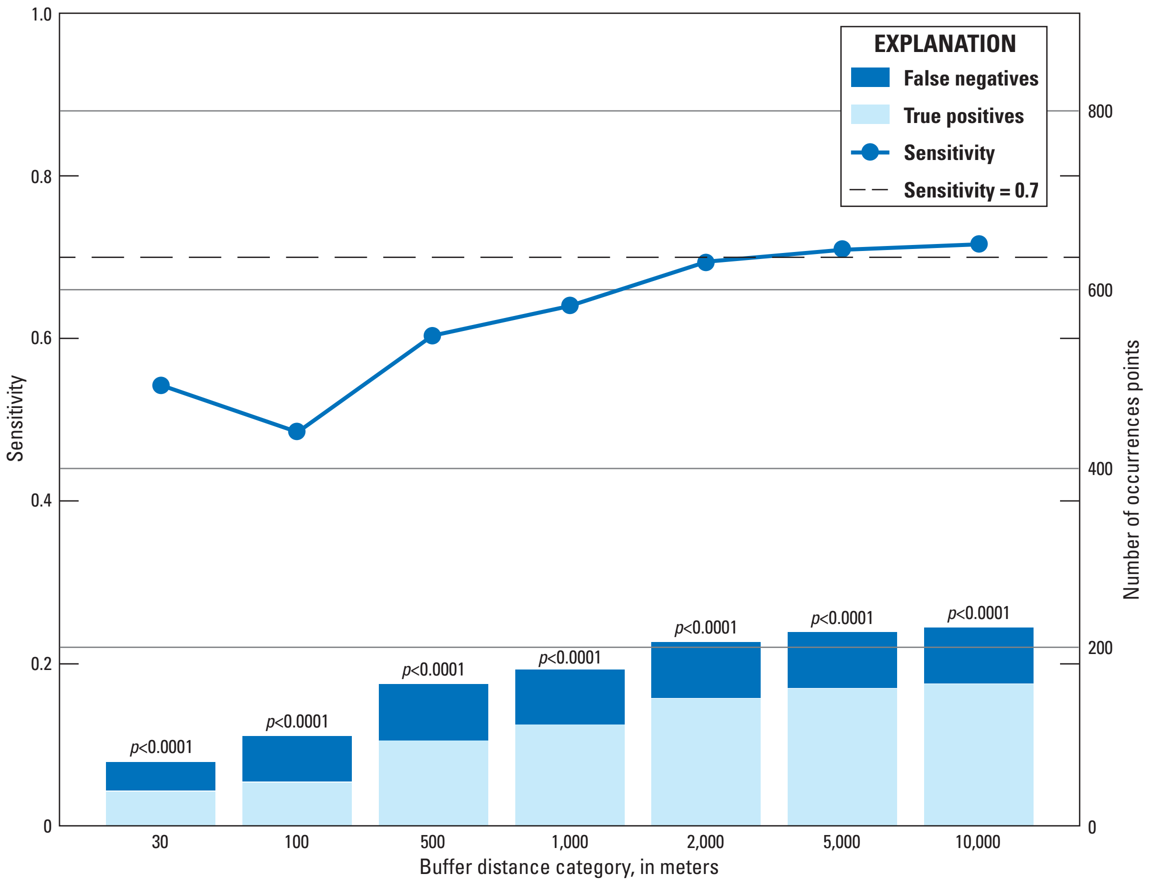


Figure 8. Graph of sensitivity and the ratio of true positives to false negatives at the seven buffer distances for a single species, *Desmognathus quadramaculatus* (black-bellied salamander), illustrating the sensitivity, number of true positives, and number of false negatives at the seven buffer distance categories. P-values above each bar represent significance for an exact one-sided binomial test.

Summary

Given the objectives of the Gap Analysis Project (GAP) to map available habitat for species, assess the conservation status for those species, and explore patterns of species richness, it is important to assess the uncertainty in the models. GAP habitat distribution models purposefully focus on modeling potential habitat knowing that not all habitat is occupied; therefore, the models tend to have high commission rates. Unfortunately, quantifying commission errors without high quality absence data is impossible. In this assessment, we focused on known species occurrences to characterize omission errors, proportion of known occurrence locations not predicted as habitat, and model performance against the null model.

Here we provide a case study for the approach based on the habitat distribution model for *Desmognathus quadramaculatus* (black-bellied salamander). In this case, Global Biodiversity Information Facility data provided a fair number of occurrence records (222) that met our criteria for inclusion, and those records appeared well dispersed throughout the range. More than 90 percent of the occurrence records (8,314 of 9,119 United States records) were filtered out before the evaluation because collection dates were prior to 1999—the earliest year of satellite imagery on which our land cover dataset used in modeling is based. Almost 250 records were removed because they lacked georeferencing, and of those that did have latitude and longitude information, more than 300 had a coordinate uncertainty of greater than 1,0000 meters. The species taxonomy was stable over the period considered, so we avoided losing additional data points to taxonomic uncertainty issues.

Expectedly, as the buffer distance increased, sensitivity increased. In practical model use, sensitivity should be relatively high (> 0.7) because it is inversely related to omission rate (not predicting a species to occur when it was in fact observed at that location). In this example for the black-bellied salamander, model sensitivity approaches 0.7 somewhere between 1,000 and 2,000 meters, indicating this model is most likely best suited to analyses where the spatial grain exceeds 1,000 meters.

Models linking species to individual cover types will always have inaccuracies related to missing habitat features that are important to species in habitat selection, the inability to account for population dynamics, and inherent oversimplification of patterns of habitat use. As mentioned previously, evaluating models based on presence-only data has its limitations, especially when those data are gathered through unstructured surveys. However, evaluations using presence-only data can provide insights into habitat models, including parts of the range that were previously missed, habitat types not included in the habitat affinity relationship database, or where parameters used to constrain the models (for example, distance to water) may be too restrictive.

Previously, GAP projects have used a variety of methods to assess habitat maps/models. Many have relied on sparse and geographically restricted data to evaluate species habitat predictions, including checklists from protected areas such as refuges, parks, and other conservation lands. Here we sought to provide metrics that would allow data users to judge the individual species habitat distribution models. This approach is best suited for species whose habitat distributions are dispersed and represent a small proportion of the range (for example, where model prevalence is low). By contrast, for generalist species where modeled habitat prevalence is high, these metrics will be increasingly less informative as habitat predictions become a larger proportion of the range.

References Cited

Anderson, R.P., Gómez-Laverde, M., and Peterson, A.T., 2002, Geographical distributions of spiny pocket mice in South America—Insights from predictive models: Global Ecology and Biogeography, v. 11, no. 2, p. 131–141, accessed July 30, 2018, at https://doi.org/10.1046/j.1466-822X.2002.00275.x.

Aycrigg, J., Beauvais, G., Gotthardt, T., Huettmann, F., Pyare, S., Andersen, M., Keinath, D., Lonneker, J., Spathelf, M., and Walton, K., 2015, Novel approaches to modeling and mapping terrestrial vertebrate occurrence in the Northwest and Alaska—An evaluation: Northwest Science, v. 89, no. 4, p. 355–381, accessed July 30, 2018, at https://doi.org/10.3955/046.089.0405.

BiSON, 2020, Biodiversity Serving our Nation Home Page website, accessed October 29, 2020, at https://bison.usgs.gov/.

Boyce, M.S., Vernier, P.R., Nielsen, S.E., Schmiegelow, F.K.A., 2002, Evaluating resource selection functions: Ecological Modelling, v. 157, p. 281–300. [Also available at https://doi.org/10.1016/S0304-3800(02)00200-4.]

Boykin, K.G., Thompson, B.C., Deitner, R.A., Schrupp, D., Bradford, D., O’Brien, L., Drost, C., Propeck-Gray, S., Rieth, W., Thomas, K.A., Kepner, W., Lowry, J., Cross, C., Jones, B., Hamer, T., Mettenbrink, C., Oakes, K.J., Prior-Magee, J., Schulz, K., Wynne, J.J., King, C., Puttere, J., Schrader, S., and Schwenke, Z., 2007, Predicted animal habitat distributions and species richness, chap. 3 *of* Prior-Magee, J.S., Boykin, K.G., Bradford, D.F., Kepner, W.G., Lowry, J.H., Schrupp, D.L., Thomas, K.A., and Thompson, B.C., eds., Southwest Regional Gap Analysis—Final report: Moscow, Idaho, U.S. Geological Survey, Gap Analysis Program, 422 p.

Boykin, K.G., Thompson, B.C., and Propeck-Gray, S., 2010, Accuracy of gap analysis habitat models in predicting physical features for wildlife-habitat associations in the southwest U.S: Ecological Modelling, v. 221, no. 23, p. 2769–2775, accessed August 15, 2020, at https://doi.org/10.1016/j.ecolmodel.2010.08.034.

Chamberlain, S., 2020, Pygbif—A python interface to the global biodiversity information facility API, version 0.5.0, accessed October 15, 2020, at https://pygbif.readthedocs.io/en/latest/index.html#.

Conant, R., and Collins, J.T., 1998, A field guide to reptiles and amphibians—Eastern and central North America (third edition, expanded): Houghton Mifflin Co., Boston and New York, 616 p.

Crother, B.I., ed., 2008, Scientific and standard English names of amphibians and reptiles of North America north of Mexico, with comments regarding confidence in our understanding (6th ed.): Society for the Study of Amphibians and Reptiles, Herpetological Circular no. 37, 94 p.

Díaz, S., Settele, J., Brondízio, E.S., Ngo, H.T., Arneth, A., Balvanera, P., Brauman, K.A., Butchart, S.H.M., Chan, K.M.A., Garibaldi, L.A., Ichii, K., Liu, J., Subramanian, S.M., Midgley, G.F., Miloslavich, P., Molnár, Z., Obura, D., Pfaff, A., Polasky, S., Purvis, A., Razzaque, J., Reyers, B., Chowdhury, R.R., Shin, Y., Visseren-Hamakers, I., Willis, K.J., and Zayas, C.N., 2019, Pervasive human-driven decline of life on Earth points to the need for transformative change: Science, v. 366, no. 6471, article eaax3100, 12 p. [Also available at https://doi.org/10.1126/science.aax3100.]

Edwards, T.C., Deshler, E.T., Foster, D., and Moisen, G.G., 1996, Adequacy of Wildlife Habitat Relation Models for Estimating Spatial Distributions of Terrestrial Vertebrates: Conservation Biology, v. 10, no. 1, p. 263–270. [Also available at https://doi.org/10.1046/j.1523-1739.1996.10010263.x.]

Global Biodiversity Information Facility [GBIF], 2022, Global Biodiversity Information Facility (GBIF) Occcurrence Download, accessed June 15, 2022, at https://doi.org/10.15468/dl.3cqd45.

Global Biodiversity Information Facility [GBIF], 2020, Global Biodiversity Information Facility (GBIF) Home Page, accessed October 20, 2020, at https://www.gbif.org.

Gergely, K., Boykin, K., McKerrow, A., Rubino, M., Tarr, N., and Williams, S., 2019, Gap Analysis Project (GAP) terrestrial vertebrate species richness maps for the conterminous U.S.: U.S. Geological Survey Scientific Investigations Report 2019–5034, 99 p., accessed August 20, 2020, at https://doi.org/10.3133/sir20195034.

Graham, C.H., and Hijmans, R.J., 2006, A comparison of methods for mapping species ranges and species richness: Global Ecology and Biogeography, v. 15, no. 6, p. 578–587, accessed July 31, 2018, at https://doi.org/10.1111/j.1466-8238.2006.00257.x.

Guralnick, R., and Constable, H., 2010, VertNet—Creating a data-sharing community: Bioscience, v. 60, no. 4, p. 258–259, accessed July 30, 2018, at https://doi.org/10.1525/bio.2010.60.4.2.

Hirzel, A.H., Le Lay, G., Helfer, V., Randin, C., and Guisan, A., 2006, Evaluating the ability of habitat suitability models to predict species presences: Ecological Modelling, v. 199, p. 142–152.

ITIS, 2020, Integrated Taxonomic Information System website, accessed October 29, 2020, at http://www.itis.gov/.

Jin, J., and Yang, J., 2020, BDcleaner—A workflow for cleaning taxonomic and geographic errors in occurrence data archived in biodiversity databases: Global Ecology and Conservation, v. 21, article e00852, 12 p., accessed July 31, 2018, at https://doi.org/10.1016/j.gecco.2019.e00852.

Kluyver, T., Ragan-Kelley, B., Pérez, F., Granger, B.E., Bussonnier, M., Frederic, J., Kelley, K., Hamrick, J.B., Grout, J., Corlay, S., Ivanov, P., Avila, D., Abdalla, S., and Willing, C., 2016, Jupyter Notebooks—A publishing format for reproducible computational workflows *in* Loizides, F. and Schmidt, B., eds., Positioning and Power in Academic Publishing—Players, Agents and Agendas—Proceedings of the 20th International Conference on Electronic Publishing, Göttingen, Germany, June 7–9, 2016: Amsterdam, IOS Press, p. 87–90.

Lannoo, M., 2005, Amphibian declines—The conservation status of United States species: Berkeley, Calif., University of California Press, 1,115 p.

Loiselle, B.A., Howell, C.A., Graham, C.H., Goerck, J.M., Brooks, T., Smith, K.G., and Williams, P.H., 2003, Avoiding pitfalls of using species distribution models in conservation planning: Conservation Biology, v. 17, no. 6, p. 1591–1600. [Also available at https://doi.org/10.1111/j.1523-1739.2003.00233.x.]

McKerrow, A.J., Tarr, N.M., Rubino, M.J., and Williams, S.G., 2018, Patterns of species richness hotspots and estimates of their protection are sensitive to spatial resolution: Diversity and Distributions, v. 24, no. 10, p. 1464–1477. [Also available at https://doi.org/10.1111/ddi.12779.]

Nelson, G., and Ellis, S., 2018, The history and impact of digitization and digital data mobilization on biodiversity research: Philosophical Transactions of the Royal Society B, v. 374, no. 1763, article 20170391, 9 p., accessed July 30, 2018, at https://doi.org/10.1098/rstb.2017.0391.

Ottaviani, D., Lasinio, G.J., and Boitani, L., 2004, Two statistical methods to validate habitat suitability models using presence-only data: Ecological Modelling, v. 179, no. 4, p. 417–443, accessed July 30, 2018, at https://doi.org/10.1016/j.ecolmodel.2004.05.016.

Pearson, R.G., 2007, Species' distribution modeling for conservation educators and practitioners: Lessons in Conservation, v. 3, p. 54–89, accessed January 5, 2022, at http://ncep.amnh.org/linc/.

Petranka, J.W., 1998, Salamanders of the United States and Canada: Washington, D.C., Smithsonian Institution Press, 587 p.

Phillips, S.J., Anderson, R.P., and Schapire, R.E., 2006, Maximum entropy modeling of species geographic distributions: Ecological Modelling, v. 190, nos. 3–4, p. 231–259, accessed January 17, 2022, at https://doi.org/10.1016/j.ecolmodel.2005.03.026.

Phillips, S.J., and Dudik, M., 2008, Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation: Ecography v. 31, p. 161–175 accessed January 17, 2022, at https://doi.org/10.1111/j.0906-7590.2008.5203.x.

Rondinini, C., Di Marco, M., Chiozza, F., Santulli, G., Baisero, D., Visconti, P., Hoffmann, M., Schipper, J., Stuart, S.N., Tognelli, M.F., Amori, G., Falcucci, A., Maiorano, L., and Boitani, L., 2011, Global habitat suitability models of terrestrial mammals: Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, v. 366, no. 1578, p. 2633–2641. [Also available at https://doi.org/10.1098/rstb.2011.0113.]

Rondinini, C., Wilson, K.A., Boitani, L., Grantham, H., and Possingham, H.P., 2006, Tradeoffs of different types of species occurrence data for use in systematic conservation planning: Ecology Letters, v. 9, no. 10, p. 1136–1145. [Also available at https://doi.org/10.1111/j.1461-0248.2006.00970.x.]

Sullivan, B.L., Wood, C.L., Iliff, M.J., Bonney, R.E., Fink, D., and Kelling, S., 2009, eBird—A citizen-based bird observation network in the biological sciences: Biological Conservation, v. 142, no. 10, p. 2,282–2,292. [Also available at https://doi.org/10.1016/j.biocon.2009.05.006.]

Tarr, N.M., McKerrow, A.J., and Rubino, M.J., 2021, The Wildlife Wrangler 0.4.0: U.S. Geological Survey software release, accessed January 6, 2021, at https://github.com/nmtarr/wildlife-wrangler.

Theobald, E.J., Ettinger, A.K., Burgess, H.K., DeBey, B., Schmidt, N.R., Froehlich, H.E., Wagner, C., HilleRisLambers, J., Tewksbury, J., Harsch, M.A., and Parrish, J.K., 2015, Global change and local solutions—Tapping the unrealized potential of citizen science for biodiversity research: Biological Conservation, v. 181, p. 236–244. [Also available at https://doi.org/10.1016/j.biocon.2014.10.021.]

U.S. Geological Survey [USGS], 1996, The Wyoming GAP analysis—Final report: U.S. Geological Survey, Gap Analysis Program, 694 p., 2 CD-ROMs.

U.S. Geological Survey [USGS], 1998, Arkansas Gap Analysis Project—1998 final report: U.S. Geological Survey, Gap Analysis Program, 205 p., 2 CD-ROMs.

U.S. Geological Survey [USGS], 2000, A Colorado Gap Analysis Program—A Geographic Approach to Planning for Biological Diversity—April 2000 final Report: U.S. Geological Survey, Gap Analysis Program, 1,387 p., 3 CD-ROMs.

U.S. Geological Survey [USGS], 2001a, A GAP analysis of Pennsylvania—2001 final report: U.S. Geological Survey, Gap Analysis Program, 773 p., 5 CD-ROMs.

U.S. Geological Survey [USGS], 2001b, The Oklahoma Gap Analysis Project—June 2001 final report: U.S. Geological Survey, Gap Analysis Program, 759 p., 5 CD-ROMs.

U.S. Geological Survey [USGS], 2003, A GAP analysis of Mississippi—2003 final report: U.S. Geological Survey, Gap Analysis Program, 607 p., 2 CD-ROMs.

U.S. Geological Survey [USGS], 2004, A GAP analysis of southern New England—SNE-GAP final report April 2004: U.S. Geological Survey, Gap Analysis Program, 572 p., 2 CD-ROMs.

U.S. Geological Survey [USGS], 2005a, A Gap analysis of Nebraska—June 2006 Nebraska final report: U.S. Geological Survey, Gap Analysis Program, 652 p., 4 CD-ROMs.

U.S. Geological Survey [USGS], 2005b, A GAP analysis of North Dakota—April 2005 final report: U.S. Geological Survey, Gap Analysis Program, 451 p., 2 CD-ROMs.

U.S. Geological Survey [USGS], 2006a, A Gap analysis of Hawaii—February 2006 final report: U.S. Geological Survey, Gap Analysis Program, 163 p., 2 CD-ROMs.

U.S. Geological Survey [USGS], 2006b, A GAP analysis of North Carolina—April 2006 final report: U.S. Geological Survey, Gap Analysis Program, 773 p., 2 DVDs.

U.S. Geological Survey [USGS] – Gap Analysis Project [GAP], 2018a, Black-bellied Salamander (*Desmognathus quadramaculatus*) aBESAx_CONUS_2001v1 Habitat Map: U.S. Geological Survey data release, accessed January 17, 2022, at https://doi.org/10.5066/F7H1308B.

U.S. Geological Survey [USGS] – Gap Analysis Project [GAP], 2018b, Black-bellied Salamander (*Desmognathus quadramaculatus*) aBESAx_CONUS_2001v1 Range Map: U.S. Geological Survey data release, accessed January 17, 2022, at https://doi.org/10.5066/F75M64R5.

U.S. Geological Survey and U.S. Department of Agriculture Natural Resources Conservation Service [USGS and USDA NRCS], 2009, Federal guidelines, requirements, and procedures for the national Watershed Boundary Dataset (1st ed.): U.S. Geological Survey Techniques and Methods 11–A3, 55 p., accessed July 31, 2018, at https://doi.org/10.3133/tm11A34. [Superseded by the 4th edition in 2013.]

Wilkinson, M.D., Dumontier, M., Aalbersberg, I., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L.B., Bourne, P.E., Bouwman, J., Brookes, A.J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C.T., Finkers, R., Gonzalez-Beltran, A., Gray, A.J.G., Groth, P., Goble, C., Grethe, J.S., Heringa, J., Hoen, P.A.C. ’t, Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S.J., Martone, M.E., Mons, A., Packer, A.L., Persson, B., Rocca-Serra, P., Roos, M., Schaik, R. van, Sansone, S., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M.A., Thompson, M., Lei, J. van der, Mulligen, E. van, Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J., and Mons, B., 2016, The FAIR guiding principles for scientific data management and stewardship: Scientific Data, v. 3, article 160018, 9 p., accessed February 12, 2020, at https://doi.org/10.1038/sdata.2016.18.

Publishing support provided by the Science Publishing Network,
Denver Publishing Service Center

For more information concerning the research in this report, contact the
Center Director, USGS Core Science Analytics and Synthesis
P.O. Box 25046, Mail Stop 302
Denver, CO 80225
(303) 202-4774

Or visit the Core Science Analytics, Synthesis, and Libraries website at
https://www.usgs.gov/core-science-systems/science-analytics-and-synthesis/

Rubino and others—**Methods for Evaluating Gap Analysis Project Habitat Distribution Maps with Species Occurrence Data**—TM 2-A19

ISSN 2328-7055 (online)
https://doi.org/10.3133/tm2A19