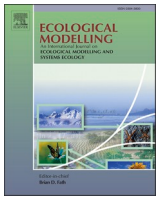
Ecological Modelling 501 (2025) 110982

Contents lists available at ScienceDirect

Ecological Modelling

journal homepage: www.elsevier.com/locate/ecolmodel



Bayesian networks facilitate updating of species distribution and habitat suitability models

Adam Duarte [a,*], Robert S. Spaan [a,b], James T. Peterson [c], Christopher A. Pearl [d], Michael J. Adams [d]

[a] U.S.D.A. Forest Service, Pacific Northwest Research Station, Olympia, WA, USA
[b] Oak Ridge Institute for Science and Education, Oak Ridge, TN, USA
[c] U.S. Geological Survey, Oregon Cooperative Fish and Wildlife Research Unit, Department of Fisheries, Wildlife, and Conservation Sciences, Oregon State University, Corvallis, OR, USA
[d] U.S. Geological Survey, Forest and Rangeland Ecosystem Science Center, Corvallis, OR, USA

ARTICLE INFO

Keywords:
American bullfrog
Decision support
Good modeling practice
Habitat suitability
Lithobates catesbeianus
Model updating
Species distribution

ABSTRACT

Managers often rely on predictions of species distributions and habitat suitability to inform conservation and management decisions. Although numerous approaches are available to develop models to make these predictions, few approaches exist to update existing models as new data accumulate. There is a need for updatable models to ensure good modeling practices in an aim to keep pace with change in the environment and change in data availability to continue to use the best-available science to inform decisions. We demonstrated a workflow to deliver predictive models to user groups within Bayesian networks, allowing models to be used to make predictions across new sites and to be easily updated with new data. To demonstrate this workflow, we focus on species distribution and habitat suitability models given their importance to informing conservation strategies across the globe. In particular, we followed a standard process of collating species encounter data available in online databases and ancillary covariate data to develop a habitat suitability model. We then used this model to parameterize a Bayesian network and updated the model with new data to predict species presence in a new focal ecoregion. We found the network updated relatively quickly as new data were incorporated, and the overall error rate generally decreased with each model update. Our approach allows for the formal incorporation of new data into predictions to help ensure model predictions are based on all relevant data available, regardless of whether they were collected after initial model development. Although our focus is on species distribution and habitat suitability models to inform conservation efforts, the workflow we describe herein can easily be applied to any use case where model uncertainty reduction and increased model prediction accuracy are desired via model updating as new data become available. Thus, our paper describes a generalizable workflow to implement model updating, which is widely recognized as a good modeling practice but is also underutilized in applied ecology.

1. Introduction

We are currently experiencing a global biodiversity crisis as ecosystems degrade under anthropogenic disturbance regimes (Betts et al., 2017; Su et al., 2021). Management interventions are increasingly implemented at local and regional scales to slow or reverse species declines (Society for Ecological Restoration, 2024). Managers often rely on predictions of species distributions and habitat suitability to understand status and trends (Adams et al., 2013; Duarte et al., 2020, 2021; Lorenz et al., 2021; Whitlock et al., 2020), to gauge where management interventions might prove effective (De wan et al., 2009; Freeman et al., 2012; Jacobsen et al., 2020; Sauer et al., 2013), and to help minimize negative effects of human caused disturbances on species and their habitats (Appel et al., 2023; McFarland et al., 2012). Yet, many species of conservation and management concern lack targeted monitoring programs to understand the dynamics of species distributions and habitat suitability over space and time.

Given most species cannot afford to wait until such targeted monitoring programs come online (if ever) to reduce these uncertainties, various tactics have been proposed and implemented to quantify species

* Corresponding author.
E-mail address: adam.duarte@usda.gov (A. Duarte).

https://doi.org/10.1016/j.ecolmodel.2024.110982
Received 27 August 2024; Received in revised form 26 November 2024; Accepted 3 December 2024
0304-3800/Published by Elsevier B.V.

distributions and habitat suitability for data-limited species. Some notable examples include using mental models of ecosystems from expert experience and knowledge (Duarte et al., 2013; O'Leary et al., 2009; Pearman-Gillman et al., 2020); conducting meta-analyses to combine disparate data/estimates collected across regions, time periods, and/or similar species (Diaz et al., 2024; Gilligan-Lunda et al., 2024; Hurtado and Burton, 2022); repurposing data from nonprobability-based monitoring efforts (i.e., 'found' data, Overton et al., 1993) to make quantitative inferences (Betts et al., 2020; Ramesh et al., 2022); making predictions using existing models developed with data from other regions (Helmstetter et al., 2021; Street et al., 2015; Zanini et al., 2009); and implementing data fusion techniques to integrate disparate data sources (Pacifici et al., 2017; Miller et al., 2019). While each of these approaches has their own set of limitations, they also each have advantages and represent a systematic approach to filling a critical information need when more targeted data are lacking.

Although numerous approaches are available to help ameliorate data limitations to develop species distribution and habitat suitability models, few approaches exist to update existing models as new data become available. Indeed, these models, particularly when packaged into maps that graphically depict the predictions, are often developed using a "loading dock" approach (*sensu* Cash et al., 2006). Here, science products are delivered to user groups (i.e., managers and policymakers) with little-to-no feedback loops to update or refine the models. This approach is particularly limiting for model application purposes given coproduction of models and mechanisms to monitor, maintain, and update models represent good modeling practices that help ensure models remain relevant and used (Herman et al., 2019; Jakeman et al., 2024; Peterson and Duarte, 2020). Furthermore, this traditional approach to science delivery may be too static to effectively inform current natural resource management decision making given the ongoing rapid change in ecosystems, coupled with the increasingly rapid accumulation of species encounter data. For example, Species Status Assessments (SSAs) conducted by the U.S. Fish and Wildlife Service often must be completed over relatively short periods of time. During this process, managers are often tasked with quantifying the status of species while continuing to accumulate new and existing species encounter data until formal evaluations are completed, which can require frequent reanalysis of data. As another example, land managers, such as the U.S.D.A. Forest Service, often use species distribution and habitat suitability models to help determine potential impact(s) of land management actions on fish and wildlife species. This process is typically supported with field surveys within proposed project areas to validate models and assess the confidence in the model predictions for specific management units. Yet, these newly collected data are often not incorporated into future predictions, and they therefore represent a missed learning opportunity to improve predictions in future use cases. Developing robust frameworks to avoid stationarity in models and their outputs is widely recognized as a good modeling practice within the field of ecological modeling (reviewed in Jakeman et al., 2024). Yet, this practice is rarely implemented in conservation and management programs due to uncertainty in how to accomplish this in a timeframe that meets the needs of user groups and imperiled species without commissioning more advanced ecological modelers.

One way to formally incorporate new data into existing models is by linking the predictive model and survey data within a Bayesian network. Bayesian networks are probabilistic graphical models that depict and quantify the relationships among variables and outcomes while explicitly considering the uncertainty in these relationships (reviewed in Pourret et al., 2008). Bayesian networks have a long history in the natural resource field (Marcot et al., 2001; Nyberg et al., 2006; Peterson and Evans, 2003; and many more). These models are frequently relied upon for decision analysis within structured decision making processes for at least four reasons: (1) Bayesian network software packages offer a user-friendly graphical user interface that makes them relatively easy to understand and use; (2) they can incorporate complex interactions among variables to capture ecosystem processes using a series of unconditional, marginal, and prior probabilities; (3) Bayesian networks tend to do relatively well when faced with missing, few, and variable quality data; and (4) these networks also can be readily updated to facilitate adaptive resource management. Given these advantages, it is not surprising that Bayesian networks have been used to predict species distributions and habitat suitability (Hamilton et al., 2015; Havron et al., 2017; McNay et al., 2006; Rieman et al., 2001; Zhang et al., 2021). However, to our knowledge there are no examples in the literature on how to update these species distribution and habitat suitability models as new data become available.

We demonstrated a workflow to deliver predictive models to user groups that are packaged in a Bayesian network that can be easily used to make predictions to new environmental conditions and updated as new data become available. By delivering predictive models within a Bayesian network this approach empowers user groups to be able to effectively use and update the model in near real time. Thus, this approach not only promotes the active engagement of user groups into the model refining process (i.e., coproduction), but it also may circumvent the need to commission more advanced ecological modelers that is often required to maintain model relevancy and use. We focused our work on species distribution and habitat suitability models given their importance to informing conservation strategies across the globe. Specifically, we followed a standard process of collating species encounter data from online databases (i.e., found data) and ancillary covariate data to develop a habitat suitability model. We then developed a Bayesian network using this model for user groups to be able to make site-level predictions of habitat suitability. Finally, we updated the model predictions with new data when using the model to predict to a new focal ecoregion. Although our focus is on species distribution and habitat suitability models to inform conservation efforts, the workflow we described can easily be applied to any use case where model uncertainty reduction and increased model prediction accuracy are desired via model updating as new data become available.

2. Methods

2.1. Study area

Our study extent includes the states of Oregon (area 254,806 km^2) and Washington (area 184,827 km^2) in the U.S.A. Pacific Northwest region (Fig. 1). Both states are bounded by Pacific Ocean on the west and state of Idaho on the east. Their coastlines comprise the Coast Range, with the following range to the east being the Cascades and Northern Cascades. Between these ranges from south to north, there are the Klamath Mountains in southern Oregon, the Willamette Valley, and the Puget Sound (Omernik and Griffiths, 2014). To the east of the Cascades and Northern Cascades lies the Eastern Cascades Slopes and Foothills, with the Blue Mountains extending from central into northeastern Oregon and southeastern Washington, separating the Northern Basin and Range in southeast Oregon and the Columbia Plateau in northern Oregon and southern Washington (Omernik and Griffiths, 2014). Lastly, in northeastern Washington, there are the Columbia Mountains/Northern Rockies (Omernik and Griffiths, 2014). Elevation within the study extent varies from sea level along the coastline to 4392 m above sea level at Mount Rainier's peak. The climates in this region are generally considered to be cooler, with wetter winters and warmer, drier summers (Beck et al., 2023). More specifically, areas west of the Cascades and Northern Cascades tend to have more cloud cover, receive significantly more precipitation in the winter, and experience cooler summer temperatures, while areas to the east of the Cascades and Northern Cascades are mostly semi-arid with warmer summers, receiving less rainfall and having colder winters (Beck et al., 2023).

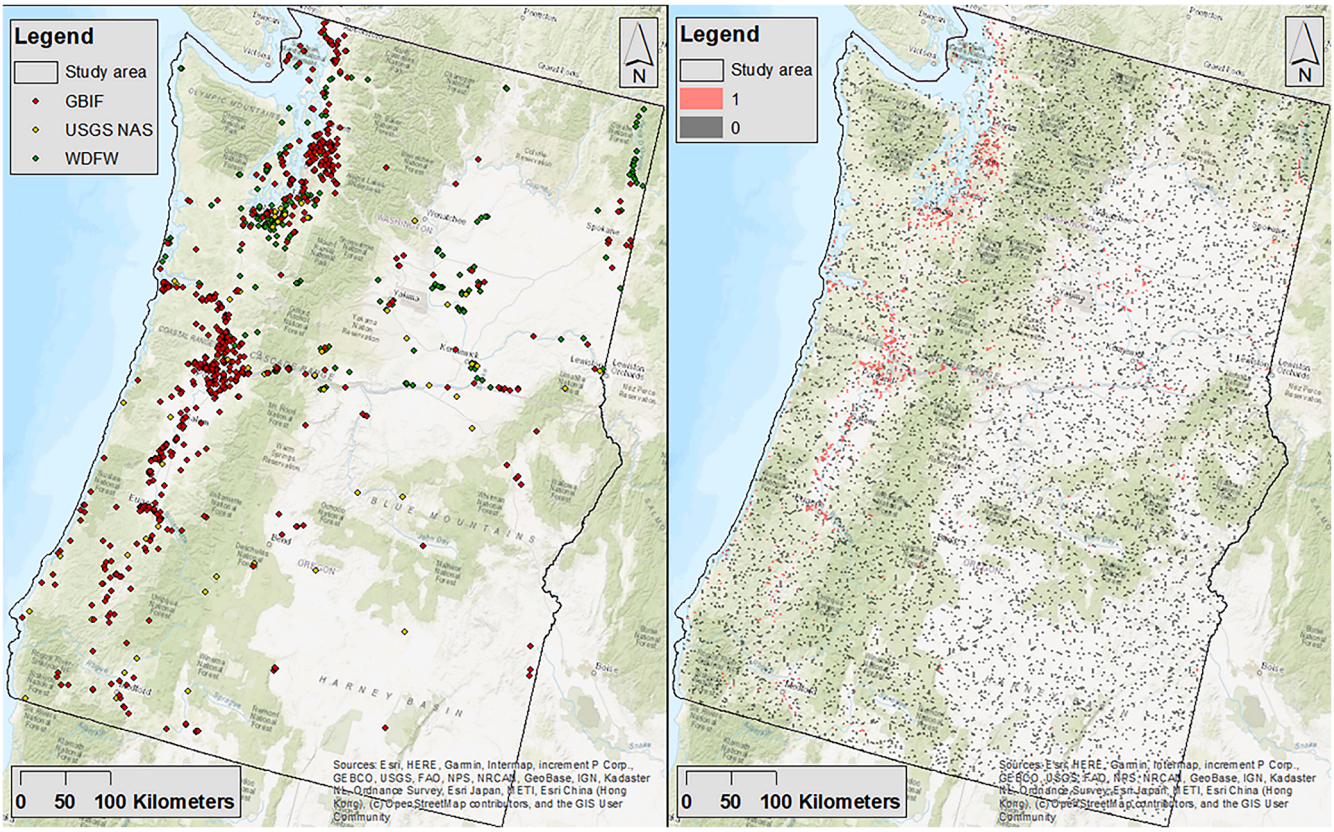


Fig. 1. The left panel shows the American bullfrog (*Lithobates catesbeianus*) point location data ($n = 2167$) from the Global Biodiversity Information Facility (GBIF; $n = 1567$), U.S. Geological Survey Nonindigenous Aquatic Species (USGS–NAS; $n = 141$), and Washington Department of Fish and Wildlife (WDFW; $n = 459$) databases in Oregon and Washington, U.S.A. The right panel shows the study area with thinned data ($n = 794$), indicated by red hexagons, and pseudo-absences ($n = 7940$), indicated by black hexagons.

2.2. Study species

The American bullfrog (*Lithobates catesbeianus*, hereafter bullfrog) were introduced to the region and have been linked to the decline of some native amphibians that co-occur with them in aquatic habitats (Adams et al., 2017; Holgerson et al., 2019; Hossack et al., 2023*a*; Kupferberg, 1997; Pearl et al., 2004; Rowe et al., 2019). Culling bullfrogs is a common management intervention within invaded areas (Hossack et al., 2023*b*). However, resources to survey for bullfrogs prior to implementing management are limited due to the expansive range of bullfrogs and large network of remote waterbodies across the U.S.A. Pacific Northwest region. Thus, our Bayesian network serves as a useful foundation for a tool to evaluate risk (see Carriger and Barron, 2020 for an example on using a Bayesian network to quantify risk) of bullfrog invasion to facilitate early detection, rapid response, and learning (i.e., model updating) for amphibian conservation measures related to bullfrog eradication in the region.

2.3. Data collection

We gathered bullfrog encounter data from the Global Biodiversity Information Facility (GBIF; $n = 1567$), the U.S. Geological Survey Nonindigenous Aquatic Species (USGS-NAS; $n = 141$), and the Washington Department of Fish and Wildlife (WDFW; $n = 459$) databases (Fig. 1). We developed a 5-km^2 hexagon tessellation to serve as our sampling frame. We removed bullfrog encounter data with missing GPS coordinates, missing temporal data, and if the metadata indicated the distance errors were greater than the hexagon resolution. We further thinned the bullfrog encounter data to a single location in each hexagon ($n = 794$) across the study extent (Fig. 1). It is worth noting that we chose this grid size for our sampling frame based on four factors. First, this resolution is relatively small given our large study extent. Second, found data such as the species encounter data made available on online databases often have errors in their locations. Increasing the size of the hexagons helped ensure that our location data were matched to the correct environmental data. Third, this scale is relevant to the assumed dispersal distance of the species, which is thought to be up to ca. 10 km (Kamath et al., 2016; Sepulveda et al., 2015; Suhre, 2010). Fourth, we envision managers in this region will use this model to gauge whether field efforts are warranted within a 5 km^2 area and, if they are, they would have enough local scale knowledge to know where within this area to focus their survey efforts.

We included nine initial topographic, climatic, hydrologic, and anthropologic covariates that are hypothesized to be related to bullfrog occurrence (Fig. 2). All variables were treated as continuous variables (Table 1). For the two topographic variables, we downloaded digital elevation models (DEM) with a resolution of 1 km^2 from Earth Explorer (https://earthexplorer.usgs.gov/) as individual tiles. We then generated a single DEM using the mosaic function and estimated slope in R version 4.3.3 (R Core Team, 2024). For climatic variables, we used remotely sensed minimum temperature, monthly and annual precipitation, and solar radiation data with a resolution of 1 km^2 from the WorldClim dataset (Fick and Hijmans, 2017). We then used the monthly data to derive precipitation totals for the growing season (April–August) and the late summer season (July–October). For waterbodies, we downloaded hydrographic vector data from the U.S. Geological Survey National Map Downloader (https://apps.nationalmap.gov/). We then calculated the distance to water by measuring the distance of the nearest waterbody from the centroid of each hexagon. For anthropologic effects, we downloaded the remotely sensed human footprint index (HFI) with a

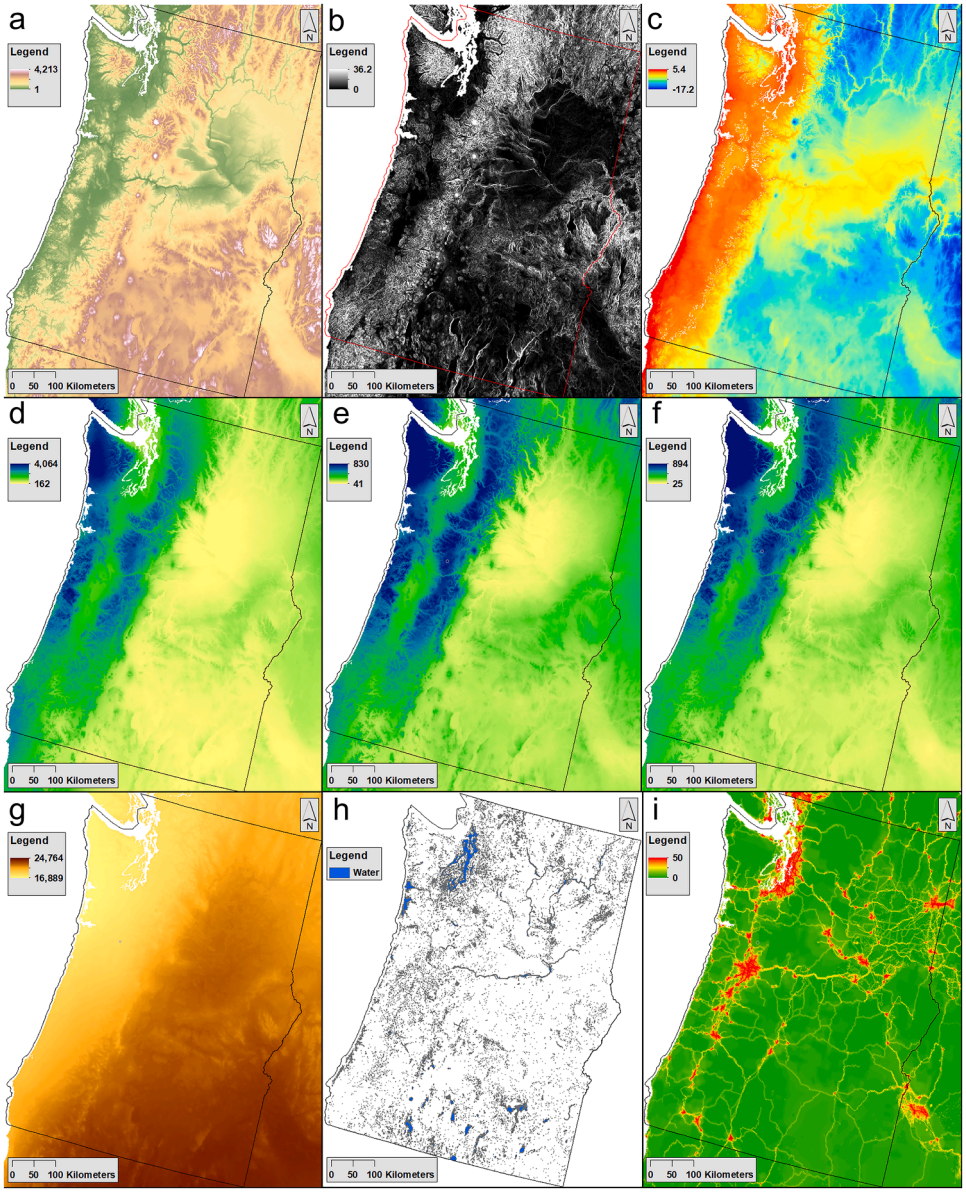


Fig. 2. The figure panel shows all covariates initially included in the modeling process: a. elevation (m), b. slope (°), c. minimum temperature (°C), d. annual precipitation (mm), e. growth season precipitation (mm), f. late summer precipitation (mm), g. solar radiation (MJ m^{-2}), h. waterbodies, and i. human footprint index, within Oregon and Washington, U.S.A.

Table 1
Summary of covariates used to assess the habitat suitability of the American bullfrog (*Lithobates catesbeianus*) in Oregon and Washington, U.S.A. with the hypothesized direction of effect of each covariate on habitat suitability.

Covariates (units)	Type	Image resolution (km^2)	Value range	$\bar{x}$ / σ	Prediction (+/-)	Source
Elevation (m)	Continuous	~1	0 – 2733.2	855.46/563.55	–	Earth Explorer
Slope (°)	Continuous	~1	0 – 23.96	3.95/3.75	–	Derived from DEM
Minimum temperature (°C)	Continuous	~1	−13.8 – 4.6	−4.52/3.93	–	WorldClim
Precipitation - annual (mm)	Continuous	~1	165 – 3359.4	906.35/716.64	+	WorldClim
Precipitation - growth (mm)	Continuous	~1	42.1 – 648.8	205.37/124.86	+	WorldClim
Precipitation – late summer (mm)	Continuous	~1	26 – 621.5	153.60/111.69	+	WorldClim
Solar radiation (MJ m^{-2})	Continuous	~1	17,457 – 24,068	21,151.51/1706.66	–	WorldClim
Distance to water (km)	Continuous	NA	0 – 20.78	1.71/1.79	–	USGS
Human footprint index	Continuous	~1	0 – 47.96	6.59/8.87	+	NASA

resolution of 1 km^2 from Dryad (https://datadryad.org). The HFI incorporates built environments, population density, electrical power infrastructure, crop lands, pasture lands, roads, railways, and navigable waterways (Venter et al., 2018). However, the HFI contained some missing data. In these instances, we assigned a value of "1" as the missing data occurred mainly in agricultural areas and reservations,

which tend to have very low HFI values. We extracted all remotely sensed data layers, except distance to water, to the hexagon tessellation using the zonal statistics function in QGIS version 3.36.1 (QGIS.org, 2024).

2.4. Habitat suitability modeling

We fit logistic regression models in R version 4.3.3 (R Core Team, 2024) to model species habitat use using maximum likelihood estimation. Because found data (hereafter referred to as "use data") often do not include absence locations, we randomly selected 10 pseudo-absence hexagons (hereafter referred to as "available data") per occupied hexagon ($n = 7940$) across the study area. We extracted all covariate data at the hexagon level. We standardized all continuous covariates (z-score) used in the models to facilitate estimation and to allow for comparison of effect sizes across variables (Schielzeth, 2010). We then examined Pearson's correlation coefficient between all nine variables and omitted highly correlated variables, $|r| \geq 0.7$ (Moore et al., 2016). We omitted elevation from the model due to collinearity ($r = -0.82$) with minimum temperature and because we hypothesized minimum temperature was a stronger limiting factor on the distribution of bullfrogs based on their physiology. We also omitted annual precipitation ($r = 0.96$), growth-season precipitation ($r = 0.99$), and solar radiation ($r = -0.83$) due to collinearity with late summer precipitation. Notably, we kept late summer precipitation in our analysis because late summer season is associated with the lowest water conditions in this region, which we hypothesized would influence bullfrog dispersal and by extension limit the distribution of bullfrogs. To assess model performance, we used area under the curve (AUC) to compare models. We used k-fold ($k = 5$) cross-validation to generate the mean AUC scores and the associated 95 % confidence intervals, representing the AUC score's uncertainty. The AUC statistical range is 0–1, where higher values indicate increased model predictive ability, and values > 0.5 indicate models are better predictive classifiers than random classification (Phillips and Dudik, 2008; Jiminez-Valverde, 2012). We also calculated the overall error rate of the best predicting model as the proportion of observations that were incorrectly classified. As we were interested in model prediction, we focused our interpretations on the top-ranking model.

2.5. Bayesian networks

Bayesian networks consist of model components, defined as nodes, with each node representing environmental states that are connected by arcs that represent causal relationships among nodes. A node that has no other node pointing to it is defined as a root node, representing a model component independent of other model components. All nodes are random variables consisting of discrete states (e.g., discretized continuous values) that are mutually exclusive and collectively exhaustive (i.e., the total probability sums to unity).

A Bayesian network is a directed acyclic graph, meaning that nodes are connected by arcs that do not form cycles (i.e., the graph does not loop). Root nodes are parameterized using statistical probability distributions (e.g., normal, Poisson, log-normal, uniform, etc.) or by incorporating data, often referred to as cases. The relationships between nodes connected via arcs are modeled using conditional probabilities. These conditional probabilities, which define the relationships between nodes, can be estimated using data (i.e., cases), expert judgment based on prior knowledge, or equations such as a linear regression model. The number of discrete states for each node should scale to the complexity of the problem at hand. Small increases in the number of states can result in larger conditional probability tables, potentially causing computational memory issues. Therefore, an objective during network development is often to minimize the number of discrete states while still adequately characterizing each node.

Bayesian networks use Bayes' theorem and user-defined relationships among nodes to facilitate forward and backward (i.e., posterior) prediction and to incorporate new data (cases) to update probability distributions. Similar to other types of models, the accuracy of a Bayesian network can be assessed by evaluating its ability to predict out-of-sample data, often using metrics like AUC when predicting observable outcomes. The sensitivity of the Bayesian network to uncertainty can be evaluated through one-way sensitivity analysis (Clemen, 1996), where root nodes are varied from their minimum to maximum values to determine their influence on the variable of interest (e.g., predicted probability of occurrence) or to assess entropy reduction in the response.

We developed a relatively simple model using Netica® (Norsys Software Corp, 2007) that linked the covariates included in the top-ranked logistic regression model (i.e., the root nodes) to the predicted site-level habitat suitability. We then parameterized the model using the built-in equation functionality in Netica®. This allowed us to sample from the full posterior probability distributions, which is not limited by sampling iterations such as simulated data are. We specified normal distributions for the root nodes using the mean and standard deviation of our observed covariate data (Table 1). We then calculated the predicted habitat suitability using the parameter estimates provided in Table 3. We discretized the covariate and predicted habitat suitability nodes into five and six, respectively, roughly equal states that spanned the range of the posterior distribution values. We included sampling uncertainty in our predicted habitat suitability using the built-in functionality in Netica® to account for the program potentially missing any narrow spikes or ridges when sampling values to quantify the posterior distributions, which may happen due to the discretization of nodes. Importantly, we chose to incorporate covariate information on the real scale within the Bayesian network to improve interpretation by users. However, covariate data were scaled in the Bayesian network within the habitat suitability equation based on the summaries provided in Table 1. This ensured predictions were calculated correctly. Finally, we used a one-way sensitivity analysis to evaluate how sensitive the predicted habitat suitability was to uncertainty in the covariates.

To illustrate how to update the Bayesian network, we refit the top-ranked logistic regression model to the original use/availability dataset after omitting locations that fell within the Puget Lowland ecoregion of Washington (total of 499 locations) (Fig. 4). Notably, we chose to perform this exercise for the Puget Lowland ecoregion because the region had a large enough number of locations to demonstrate multiple iterations of model updating. We then parameterized the model using the same methods described above but modified the values (both covariate summaries and the equation used to predict habitat suitability) to match the full dataset after omitting Puget Lowland locations. Given probabilities are not observed in the field, we added an additional discrete node to our Bayesian network that represented the realized outcome (i.e., species presence/absence) and assigned the conditional probabilities based on the midpoint of the habitat suitability for each respective bin. For example, a habitat suitability score of 0–0.17 was assigned a probability of Present equal to 0.085, a habitat suitability score of 0.17–0.33 was assigned a probability of Present equal to 0.25, and so on. We then absorbed the habitat suitability node. Absorbing this node accomplished at least two things: 1) it simplified the conditional probability tables used by the Bayesian network; and 2) more importantly, it allowed the Bayesian network to be updated with new data. Bayesian networks require that all nodes in the network are present in the new data to update the model and (again) habitat suitability is not typically directly measured in the field. Similar to our model selection procedure described above, we conducted k-fold ($k = 5$) cross validation to evaluate the change in predictive accuracy of the Bayesian network after updating with new sample data. We randomly separated the Puget Lowland data into roughly five equal datasets with one set used as the testing data and the remaining four as new case files. We then updated the Bayesian network by incorporating new case files sequentially one at a time. After each update, we estimated the response (i.e., present or

absent) of the test data and calculated the overall error rate and the category wise prediction error rates. Within Netica®, the user has the option to specify how influential the new case files are relative to the samples used to parameterize the network based on 'experience', which is the effective sample size of the sample data. The greater the sample size the greater the influence of the data on the posterior estimates after updating the model. When using equations to parameterize the nodes, the experience for each combination of discrete conditions within the conditional probability table defaults to one. We can then specify the weight (or experience) of each new data point used to update the Bayesian network. We were interested in our ability to learn quickly. Therefore, we conducted one cross validation exercise while weighing the Bayesian network and case files equally and another where we weighted the Bayesian network 100 times more. Each cross-validation exercise was replicated 10,000 times in R version 4.3.3 (R Core Team, 2024) using the RNetica package (Almond, 2024). We summarized the error rates across all iterations using the mean to make our inferences.

3. Results

3.1. Habitat suitability

The top model predicting bullfrog use had an AUC score of 0.925 (95 % CI: [0.912; 0.938]) (Table 2). Assuming a threshold of 0.5 for predicting presence, the overall error rate of the model was 0.08 with a presence prediction error rate of 0.32. Variables included in the top model included slope, minimum temperature, late summer precipitation, distance to water, HFI, and the interaction between minimum temperature and late summer precipitation (Table 2; Fig. 3). The top model strongly supported slope, minimum temperature, distance to water, HFI, and the interaction between minimum temperature and late summer precipitation. There was relatively little support for late summer precipitation. Increasing slope was associated with decreased habitat suitability. Increasing minimum temperature was associated with increased habitat suitability. Increasing distances to water from the hexagon centroid were associated with decreased habitat suitability. An increase in HFI was associated with increased habitat suitability. Finally, wetter late summer seasons with warmer minimum temperatures were associated with increased habitat suitability. Habitat suitability was predicted to be higher in the lowland areas west of the Cascade Range, with relatively few areas (other than larger waterbodies, e.g., the Columbia River, and a few urban areas, e.g., Klamath Falls, Spokane, and Walla) predicted to have high habitat suitability east of the Cascade Range (Fig. 4).

3.2. Bayesian network

Our Bayesian network (Fig. 5A) was able to predict site-level habitat suitability for sites with a range of environmental conditions (Fig. 5B) and was able to predict site-level habitat suitability when covariate data were lacking (Fig. 5C). Overall, predicted habitat suitability was most sensitive to HFI, distance to water, and slope (Fig. 6). As expected, learning was slower when the network was weighted 100 times more than each new data point (Fig. 7). In both cases, the average predicted habitat use quickly increased, which better matched the data that came from this ecoregion based on the error rates and visual inspection of Figs. 1 and 4. Overall error rates were similar between the weighing schemes (Fig. 8). However, errors in predicting absence were lower when the network was weighted 100 times more than each new data point. Yet, errors in predicting presence were higher when the network was weighted 100 times more than each new data point. We found the errors in predicting absence were largely invariant following the first model update for both weighting schemes. Similarly, both weighting schemes seemed to result in errors in predicting presence that leveled off after three updates. However, there was a slight increase in the overall prediction error rate after four updates when the network was weighted equally to each new data point.

Table 2
Model selection results from the top four logistic regression models and the null model predicting American bullfrog (*Lithobates catesbeianus*) habitat suitability in Oregon and Washington, U.S.A. The most supported model is in bold. Models were evaluated using the area under the curve (AUC), where higher AUC values indicate models with better predictive power; lower (LB) and upper (UB) bounds of the 95 % CI are also presented.

		95 % AUC	
Model	AUC	LB	UB
Slope + Min. temp. + Late summer precip. + Dist. to water + HFI + Min. temp. × Late summer precip.	**0.925**	**0.912**	**0.938**
Slope + Min. temp. + Late summer precip. + Dist. to water + HFI	0.924	0.913	0.937
Min. temp. + Late summer precip. + Dist. to water + HFI	0.924	0.914	0.938
Slope + Min. temp + Dist. to water + HFI	0.922	0.910	0.935
Null	0.500	0.500	0.500

Table 3
Model intercept and coefficient estimates from the top logistic regression model predicting the habitat suitability of American bullfrogs (*Lithobates catesbeianus*) in Oregon and Washington, U.S.A.

			95 % CI	
Parameter	Mean	SE	LB	UB
Intercept	−3.85	0.12	−4.09	−3.61
Slope	−0.46	0.10	−0.65	−0.27
Minimum temperature	0.86	0.08	0.70	1.02
Late summer precipitation	−0.03	0.12	−0.26	0.21
Distance to water	−1.68	0.14	−1.96	−1.41
Human footprint index	0.66	0.04	0.59	0.74
Minimum temperature × Late summer precipitation	−0.29	0.10	−0.48	−0.10

4. Discussion

Managers and policymakers need updatable products that keep pace with change in the environment and change in data availability to continue to use the best-available science to inform decisions as new data are collected. Using species distribution and habitat suitability models as an example, we demonstrated how to develop Bayesian networks based on current predictive models and how to update the model as new data become available. This approach allows for the formal incorporation of new data into predictions to help ensure model predictions are based on all relevant data available, regardless of whether they were collected after initial model development. We believe this approach represents a good, albeit underutilized, practice when modeling the distribution of species and their habitats to inform conservation and management efforts.

Our stepwise approach matched how species distribution and habitat suitability models are often developed: using use/availability data and relating them to environmental covariates using regression models with model selection. Thus, the Bayesian network we ultimately developed was relatively simple in that it only considered the covariates included in the top-ranked regression model. It should be noted, however, that Bayesian networks can accommodate correlated variables fairly well and more connected networks with more predictor variables to represent system dynamics are certainly possible (as an example, see Havron et al., 2017). We chose a simpler Bayesian network to match a more common use case for developing species distribution and habitat suitability models. We also stress that simpler models that avoid unnecessary complexity can greatly increase the interpretability and use of model results in real world management decision making (Phillips, 1984). Still, a likely useful extension to our bullfrog Bayesian network in future efforts would be to include the predicted decision to survey a site

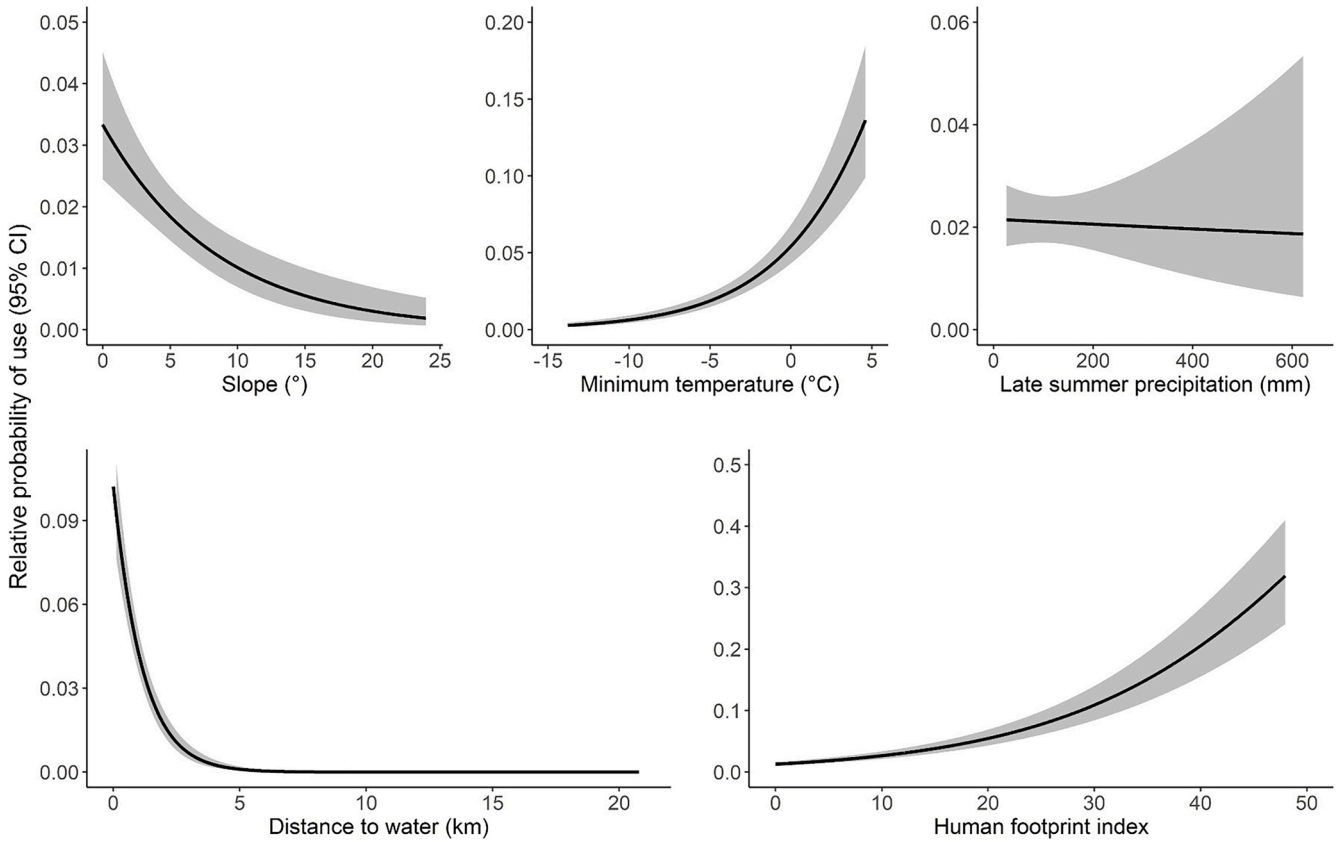


Fig. 3. Predicted habitat suitability (with 95 % confidence intervals) for American bullfrog (*Lithobates catesbeianus*) in Oregon and Washington, U.S.A. across the range of covariates included in the top logistic regression model. Note: We visualized the effect of each covariate using the marginal predictions across the range of the covariate and plots have different scales on the y-axes.

or not. We envision that this decision-support tool would likely include tradeoffs associated with survey effort required based on species detection probabilities and alternative survey methods, resources to invest in monitoring, and predicted bullfrog habitat use.

We used Netica® to develop our Bayesian network, which offers the opportunity to weight the importance of new data when updating the model. As noted by Marcot et al. (2006), lower weights can be used when updating with sparse data to avoid overfitting, where an overfit model has poor performance when predicting across scenarios outside the data. We found that weighting new data the same as the network resulted in faster learning and predictions that better matched the data in the Puget Lowland region. However, we were surprised that the overall error rate seemed to increase (albeit only by a small amount) after only four model updates when weighting the new data equally. Perhaps we were beginning to experience an overfit model. Then again, our updating process assumed our network included all environmental covariates needed to accurately predict bullfrog habitat use. Perhaps after four updates there was enough new sample data that a reanalysis was warranted to better match the specific ecoregion. This may be particularly relevant in our case study example given the Puget Lowland region had substantially less variability in the environmental variables used to predict habitat use at the spatial resolution of 5 km^2 (see the distribution of environmental variables in Fig. 7). The reduced variability in these environmental conditions may have resulted in these covariates not being well suited to differentiate used versus unused sites. In practice, as ecosystems change or enough location-specific data become available, new environmental covariates can be included by conducting a reanalysis or by simply adding a new root node to the network with uninformative priors or a prior distribution based on expert elicitation.

Our habitat suitability model had high predictive accuracy across the region. Yet, our Bayesian network accuracy assessment highlights how the network had relatively low predictive accuracy when examined within the Puget Lowland ecoregion. This is not surprising and is a common outcome when regional or landscape-scale models are used to predict to more local scales. Some error in the predictions should always be expected, and it should be noted that the impact of error(s) depends on the context in which the predictions are being used (reviewed in Conroy and Peterson, 2013 and Marcot, 2020). We found that updating the model resulted in much lower prediction error when attempting to identify sites where bullfrogs were present. This means the Bayesian network provided more reliable predictions concerning where bullfrogs occurred, which was a fundamental objective of our case study. However, these updates came at the cost of under predicting unoccupied sites at this scale. Thus, the Bayesian network became more inclusive, where the network progressively indicated bullfrogs occurred in areas where they were absent. Notably, our sampling unit spatial resolution of 5 km^2 was partially motivated by the extent of our study (i.e., the states of Oregon and Washington in the U.S.A. Pacific Northwest region). As we updated our Bayesian network and refocused its predictions toward the Puget Lowland ecoregion, our results likely began to illustrate some of the tradeoffs between spatial grain and extent in ecological studies (Wiens, 1989). Specifically, as the spatial extent of our Bayesian network reduced to a focal ecoregion, our Bayesian network would likely have benefited from a finer spatial grain (i.e., <5 km^2) to better capture the heterogeneity in environmental conditions at the more local scale. Certainly, if this network was updated using more fine scale data (i.e., use/availability data and associated covariate information summarized at a finer spatial grain) the network would eventually conform to local scale predictions (i.e., the scale of the new data used to update the network). However, doing so would require collecting use/availability data with more accurate locations.

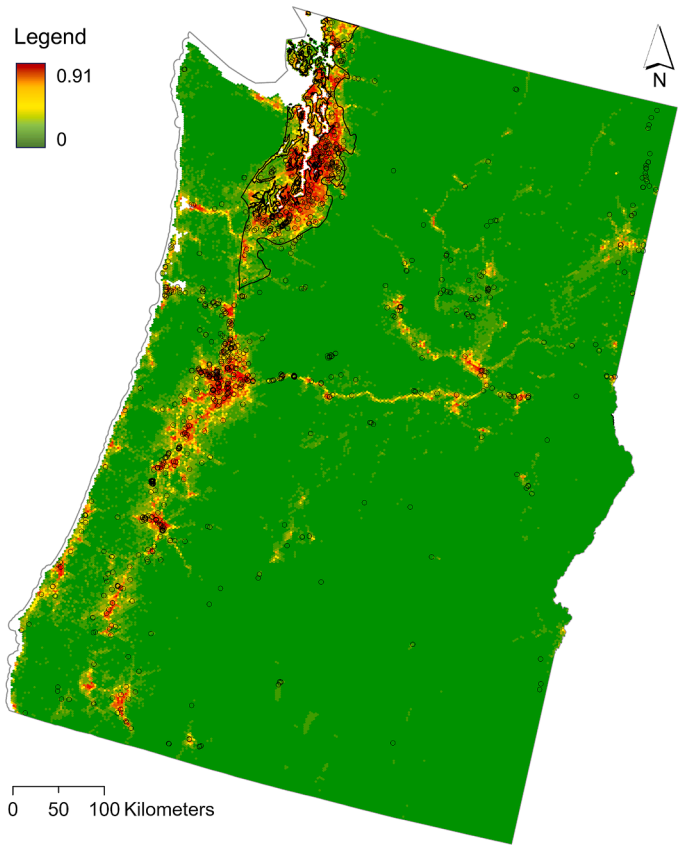


Fig. 4. Predicted American bullfrog (*Lithobates catesbeianus*) habitat suitability derived from the top logistic regression model. Point locations show the publicly available bullfrog detections included in the study. Black outline shows the boundary of the Puget Lowland ecoregion in Washington, U.S.A.

Although we used the built-in equation functionality in Netica® to parameterize our Bayesian network, as noted earlier we could have also parameterized the network using data. We opted to use the equations, as opposed to simulated or field data case files, to help represent every combination of variables within the conditional probabilities. We found this to be a large issue when directly using the use/availability data as case files in our initial efforts. Furthermore, in a preliminary analysis our Bayesian network still had multiple combinations that were not represented in the simulated case files even after simulating this process 1000,000 times using stochastic processes. In such scenarios, the network will provide a unform distribution on the predicted habitat suitability, which indicates maximum uncertainty (see Zhang et al., 2021 as an example of this). Using the built-in equations overcomes this issue and is equivalent to the standard approach of using logistic regression models to predict habitat suitability across landscapes. Thus, we suggest others take advantage of this built-in functionality from the onset in future efforts.

5. Conclusion

In the face of the current global biodiversity crisis, managers and policymakers need access to the best-available science to help inform decisions. Although models that predict species distributions and habitat suitability will undoubtedly continue to help in this endeavor, predictions using static models assumes stationarity in how animals use habitats across space and time. As ecosystems continue to shift, animals may modify how they use their environment. Thus, newly collected monitoring data should not be excluded from subsequent predictions unless there is some fundamental flaw in how these new data were collected, particularly for data-limited species and/or regions. The workflow we demonstrated herein provides a transparent and repeatable approach to continue to incorporate the best-available science into conservation and management decisions based on the distribution of species and their habitats. Furthermore, we stress that this workflow can easily be applied to any use case where model uncertainty reduction and increased model prediction accuracy are desired via model updating as new data become available.

CRediT authorship contribution statement

Adam Duarte: Writing – original draft, Visualization, Validation, Supervision, Software, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Robert S. Spaan:** Writing – original draft, Visualization,

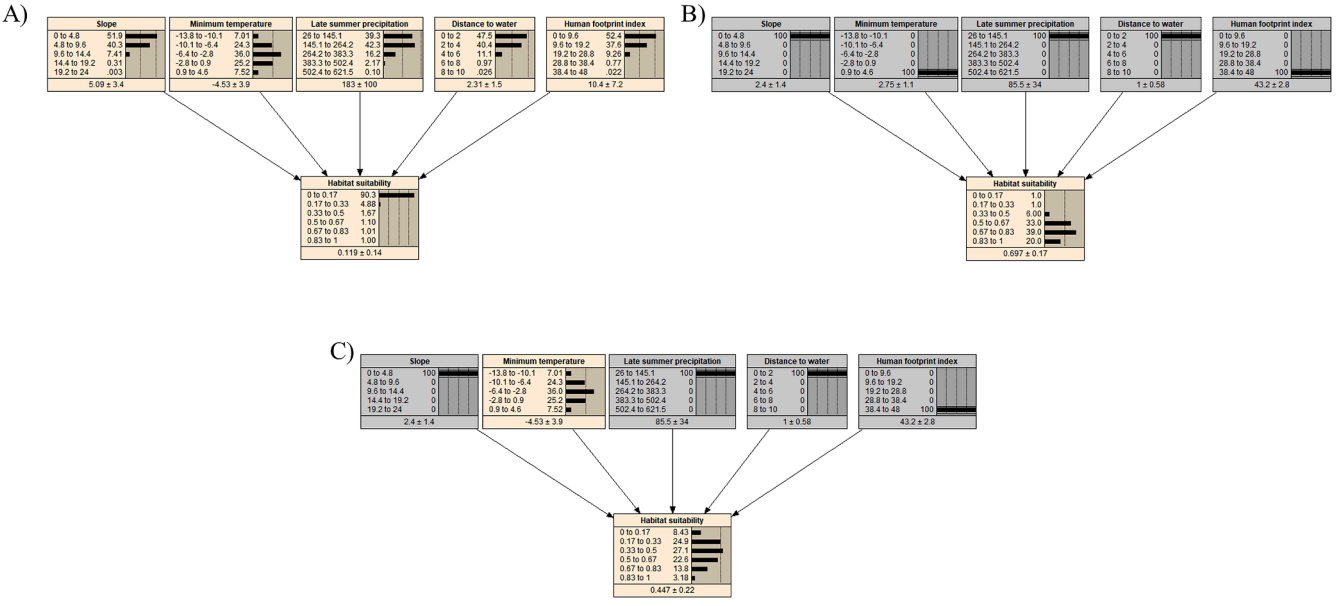


Fig. 5. Bayesian network predicting American bullfrog (*Lithobates catesbeianus*) habitat suitability in Oregon and Washington, U.S.A. using full dataset (A) with example site-level habitat suitability predictions with all data (B) and missing covariate data (C).

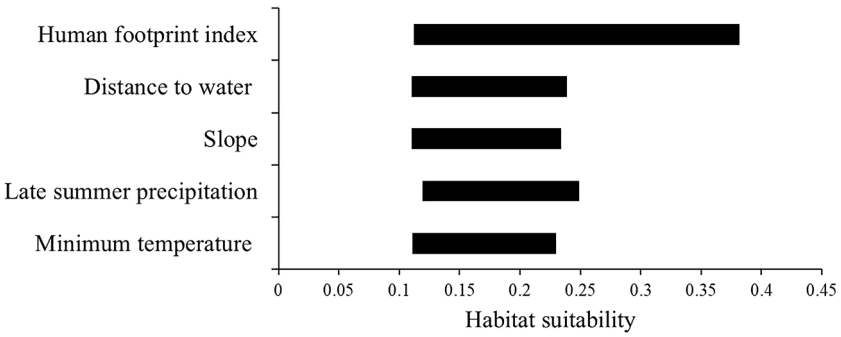


Fig. 6. Tornado diagram for one-way sensitivity analysis with model components listed from greatest (top) to least (bottom) influential for predicted American bullfrog (*Lithobates catesbeianus*) habitat suitability based on the Bayesian network with the full dataset.

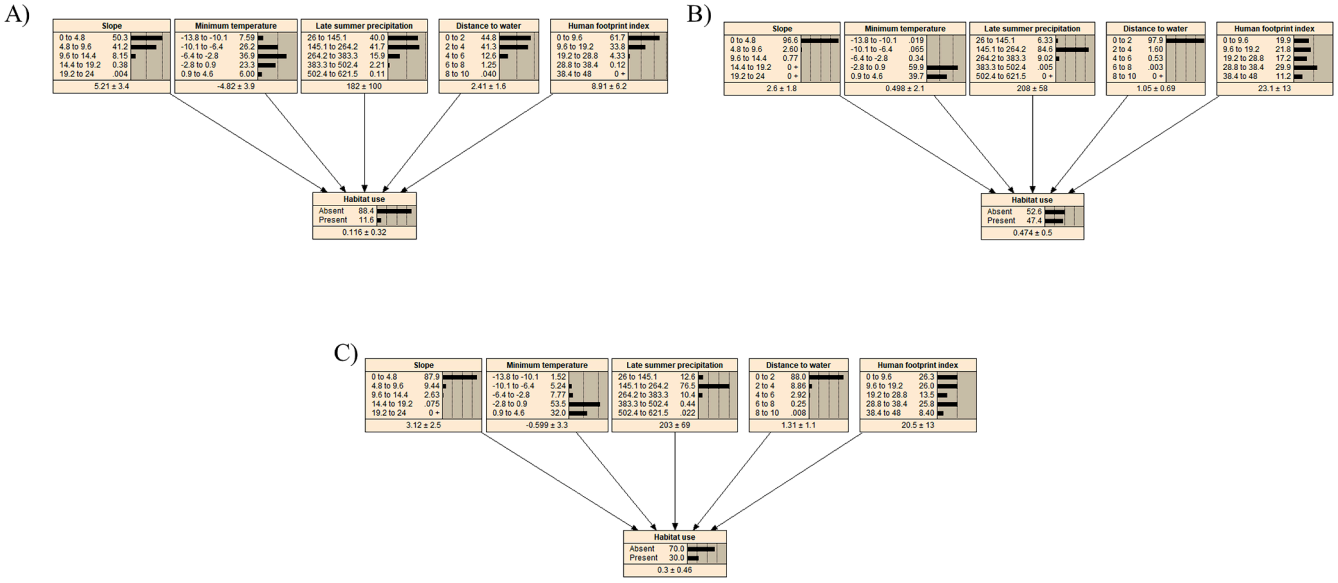


Fig. 7. Bayesian network predicting American bullfrog (*Lithobates catesbeianus*) habitat use after omitting the Puget Lowland ecoregion in Washington, U.S.A. Depicted is the original network (A) followed by updates, where we weighted each new data point equally with (B) or 100 times less than (C) each data point used to parameterize the original network.

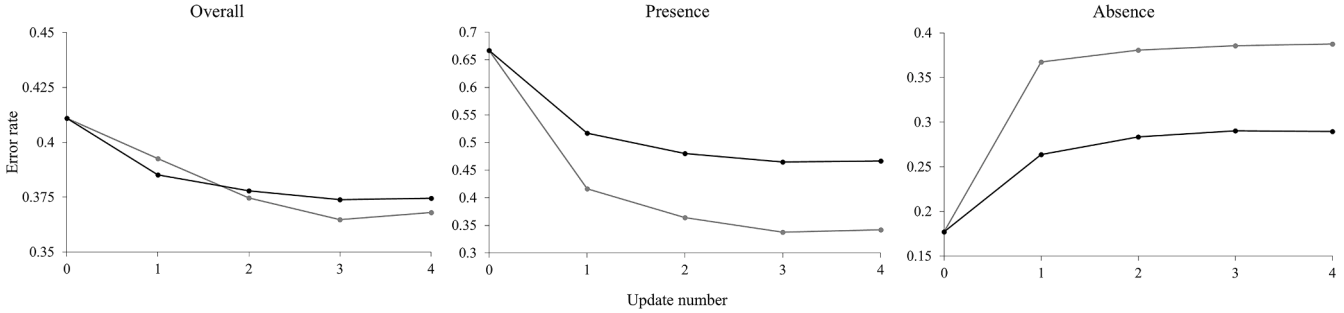


Fig. 8. Error rates for Bayesian network predicting American bullfrog (*Lithobates catesbeianus*) habitat use after omitting the Puget Lowland ecoregion in Washington, U.S.A. Depicted is the original network (update number 0) followed by updates, where we weighted each new data point equally with (gray line) or 100 times less than (black line) the original network. Note the different scales on the y-axes.

Validation, Methodology, Investigation, Formal analysis, Data curation. **James T. Peterson:** Writing – review & editing, Validation, Software, Methodology. **Christopher A. Pearl:** Writing – review & editing, Validation, Project administration, Funding acquisition. **Michael J. Adams:** Writing – review & editing, Validation, Project administration, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This project was funded by the U.S. Geological Survey Science Support Partnership Program (SSP) in collaboration with the U.S. Fish and Wildlife Service's Washington Fish and Wildlife Office in Lacey, WA (F. Waterstrat). Additional support was provided by the U.S.D.A. Forest Service, Pacific Northwest Research Station. The Oregon Cooperative Fish and Wildlife Research Unit is jointly sponsored by the U.S. Geological Survey, the U.S. Fish and Wildlife Service, the Oregon Department of Fish and Wildlife, Oregon State University, and the Wildlife Management Institute. The use of trade or firm names in this publication is for reader information and does not imply endorsement by the U.S. Government of any product or service. This is product number 931 of the U.S. Geological Survey's Amphibian Research and Monitoring Initiative.

Data availability

All data used are freely available on public domain as indicated in the text.

References

Adams, A.J., Kupferberg, S.J., Wilber, M.Q., Pessier, A.P., Grefsrud, M., Bobzien, S., Vredenburg, V.T., Briggs, C.J., 2017. Extreme drought, host density, sex, and bullfrogs influence fungal pathogen infection in a declining lotic amphibian. Ecosphere 8, e01740.

Adams, M.J., Miller, D.A.W., Muths, E., Corn, P.S., Grant, E.H.C., Bailey, L.L., Fellers, G. M., Fisher, R.N., Sadinski, W.J., Waddle, H., Walls, S.C., 2013. Trends in amphibian occupancy in the United States. PLoS. One 8, e64347.

Almond, R.. R interface to Netica® Bayesian Network Engin. Version 0.10-3. https://ralmond.r-universe.dev.

Appel, C.L., Lesmeister, D.B., Duarte, A., Davis, R.J., Weldy, M.J., Levi, T., 2023. Using passive acoustic monitoring to estimate northern spotted owl landscape use and pair occupancy. Ecosphere 14, e4421.

Beck, H.E., McVicar, T.R., Vergopolan, N., Berg, A., Lutsko, N.J., Dufour, A., Zeng, Z., Jiang, X., Van Dijk, A.I.J.M., Miralles, D.G., 2023. High-resolution (1 km) Köppen-Geiger maps for 1901–2099 based on constrained CMIP6 projections. Sci. Data 10, 724.

Betts, M.G., Wolf, C., Ripple, W.J., Phalan, B., Millers, K.A., Duarte, A., Butchart, S.H.M., Levi, T., 2017. Global forest loss disproportionately erodes biodiversity in intact landscapes. Nature 547, 441–444.

Betts, M.G., Northrup, J.M., Guerrero, J.A.B., Adrean, L.J., Nelson, S.K., Fisher, J.L., Gerber, B.D., Garcia-Heras, M.-S., Yang, Z., Roby, D.D., Rivers, J.W., 2020. Squeezed by a habitat split: warm ocean conditions and old-forest loss interact to reduce long-term occupancy of a threatened seabird. Conserv. Lett. 13, e12745.

Carriger, J.F., Barron, M.G., 2020. A Bayesian network approach to refining ecological risk assessments: mercury and Florida panther (*Puma concolor coryi*). Ecol. Modell. 418, 108911.

Cash, D.W., Borck, J.C., Patt, A.G., 2006. Countering the loading-dock approach to linking science and decision making: comparative analysis of El Niño/Southern Oscillation (ENSO) forecasting systems. Sci. Technol. Hum. Values 31, 465–494.

Clemen, R.T., 1996. Making Hard Decisions: an Introduction to Decision. Analysis. Vol. 2.

Conroy, M.J., Peterson, J.T., 2013. Decision Making in Natural Resource Management: A Structured, Adaptive Approach. John Wiley and Sons, Hoboken, NJ.

De Wan, A.A., Sullivan, P.J., Lembo, A.J., Smith, C.R., Maerz, J.C., Lassoie, J.P., Richmond, M.E., 2009. Using occupancy models of forest breeding birds to prioritize conservation planning. Biol. Conserv. 142, 982–991.

Diaz, L., Duarte, A., Beakes, M., Peterson, J.T., 2024. Ontogenetic niche partitioning in a facultatively anadromous salmonid: implications for population dynamics. Glob. Ecol. Conserv. 49, e02770.

Duarte, A., Jensen, J.L.R., Hatfield, J.S., Weckerly, F.W., 2013. Spatiotemporal variation in range-wide golden-cheeked warbler breeding habitat. Ecosphere 4, 152.

Duarte, A., Peterson, J.T., Pearl, C.A., Rowe, J.C., McCreary, B., Galvan, S.K., Adams, M. J., 2020. Estimation of metademographic rates and landscape connectivity for a conservation-reliant anuran. Landsc. Ecol. 35, 1459–1479.

Duarte, A., Pearl, C.A., McCreary, B., Rowe, J.C., Adams, M.J., 2021. An updated assessment of status and trend in the distribution of the Cascades frog (*Rana cascadae*) in Oregon, USA. Herpetol. Conserv. Biol. 16, 361–373.

Fick, S.E., Hijmans, R.J., 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. Int. J. Climatol. 37, 4302–4315.

Freeman, M.C., Buell, G.R., Hay, L.E., Hughes, W.B., Jacobson, R.B., Jones, J.W., Jones, S.A., Lafontaine, J.H., Odom, K.R., Peterson, J.T., Riley, J.W., Schindler, J.S., Shea, C., Weaver, J.D., 2012. Linking river management to species conservation using dynamic landscape models. River. Res. Appl. 29, 906–918.

Gilligan-Lunda, E.K., Duarte, A., Peterson, J.T., 2024. Habitat use of anadromous and amphidromous sturgeons in North America: a systematic review. Can. J. Fish. Aquat. Sci. 81, 508–524.

Hamilton, S.H., Pollino, C.A., Jakeman, A.J., 2015. Habitat suitability modelling of rare species using Bayesian networks: model evaluation under limited data. Ecol. Modell. 299, 64–78.

Havron, A., Goldfinger, C., Henkel, S., Marcot, B.G., Romsos, C., Gilbane, L., 2017. Mapping marine habitat suitability and uncertainty of Bayesian networks: a case study using Pacific benthic macrofauna. Ecosphere 8, e01859.

Helmstetter, N.A., Conway, C.J., Stevens, B.S., Goldberg, A.R., 2021. Balancing transferability and complexity of species distribution models for rare species conservation. Divers. distrib. 27, 95–108.

Herman, B.D., McKay, S.K., Altman, S., Richards, N.S., Reif, M., Piercy, C.D., Swannack, T.M., 2019. Unpacking the black box: demystifying ecological models through interactive workshops and hands-on learning. Front. Environ. Sci. 7, 122.

Holgerson, M.A., Duarte, A., Hayes, M.P., Adams, M.J., Tyson, J.A., Douville, K.A., Strecker, A.L., 2019. Floodplains provide important amphibian habitat despite multiple ecological threats. Ecosphere 10, e02853.

Hossack, B.R., Oja, E.B., Owens, A.K., Hall, D., Cobos, C., Crawford, C.L., Goldberg, C.S., Hedwall, S., Howell, P.E., Lemos-Espinal, J.A., MacVean, S.K., McCaffery, M., Mosley, C., Muths, E., Sigafus, B.H., Sredl, M.J., Rorabaugh, J.C., 2023a. Empirical evidence for effects of invasive American bullfrogs on occurrence of native amphibians and emerging pathogens. Ecol. Appl. 33, e2785.

Hossack, B.R., Hall, D., Crawford, C.L., Goldberg, C.S., Muths, E., Sigafus, B.H., Chambert, T., 2023b. Successful eradication of invasive American bullfrog leads to coextirpation of emerging pathogens. Conserv. Lett. 16, e12970.

Hurtado, C.M., Burton, A.C., 2022. Using habitat suitability information from across a species range to inform understudied ecosystems. J. Mammal. 103, 814–825.

Jacobsen, C.D., Brown, D.J., Flint, W.D., Pauley, T.K., Buhlmann, K.A., Mitchell, J.C., 2020. Vulnerability of high-elevation endemic salamanders to climate change: a case study with the cow knob salamander (*Plethodon punctatus*). Glob. Ecol. Conserv. 21, e00883.

Jakeman, A.J., Elsawah, S., Wang, H.-H., Hamilton, S.H., Melsen, L., Grimm, V., 2024. Towards normalizing good practice across the whole modeling cycle: its instrumentation and future research topics. Socio-Environ. Syst. Model. 6, 18755.

Jiminez-Valverde, A., 2012. Insights in the area under the receiver operating characteric curve (AUC) as a discrimination measure in species distribution modelling: insights in the AUC. Glob. Ecol. Biogeogr. 21, 498–507.

Kamath, P.L., Sepulveda, A.J., Layhee, M., 2016. Genetic reconstruction of a bullfrog invasion to elucidate vectors of introduction and secondary spread. Ecol. Evol. 6, 5221–5233.

Kupferberg, S.J., 1997. Bullfrog (*Rana catesbeiana*) invasion of a California river: the role of larval competition. Ecology. 78, 1736–1751.

Marcot, B.G., 2020. The Science and Management of Uncertainty, 1st Edition. CRC Press, Boca Raton, FL.

Marcot, B.G., Holthausen, R.S., Raphael, M.G., Rowland, M.M., Wisdom, M.J., 2001. Using Bayesian belief networks to evaluate fish and wildlife population viability under land management alternatives from an environmental impact statement. For. Ecol. Manage. 153, 29–42.

Marcot, B.G., Steventon, J.D., Sutherland, G.D., McCann, R.K., 2006. Guidelines for developing and updating Bayesian belief networks applied to ecological modeling and conservation. Can. J. For. Res. 36, 3063–3074.

McFarland, T.M., Mathewson, H.A., Groce, J.E., Morrison, M.L., Newman, J.C., Snelgrove, R.T., Snow, K.L., Collier, B.A., Wilkins, R.N., 2012. Utilization of a species occupancy model for management and conservation. Wildl. Soc. Bull. 36, 432–439.

McNay, R.S., Marcot, B.G., Brumovsky, V., Ellis, R., 2006. A Bayesian approach to evaluating habitat for woodland caribou in north-central British Columbia. Can. J. For. Res. 36, 3117–3133.

Miller, D.A.W., Pacifici, K., Sanderlin, J.S., Reich, B.J., 2019. The recent past and promising future for data integration methods to estimate species' distributions. Methods Ecol. Evol. 10, 22–37.

Moore, D.S., McCabe, G.P., Craig, B.A., 2016. Introduction to the Practice of Statistics, Ninth Edition. W.H. Freemen, Carlstadt, NJ.

Norsys Software Corp. Netica 3.17 Bayesian Network Software from Norsys. http://www.norsys.com.

Nyberg, J.B., Marcot, B.G., Sulyma, R., 2006. Using Bayesian belief networks in adaptive management. Can. J. For. Res. 36, 3104–3116.

Lorenz, T.J., Raphael, M.G., Young, R.D., Lynch, D., Nelson, S.K., McIver, W.R., 2021. Status and Trend of Marbled Murrelet Populations Under the Northwest Forest Plan, 1993 to 2017. U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station, Portland, OR, p. 64.

O'Leary, R.A., Choy, S.L., Murray, J.V., Kynn, M., Denham, R., Martin, T.G., Mengersen, K., 2009. Comparison of three expert elicitation methods for logistic regression on predicting the presence of the threatened brush-tailed rock-wallaby *Petrogale penicillate*. Environmetrics. 20, 379–398.

Omernik, J.M., Griffith, G.E., 2014. Ecoregions of the conterminous United States: evolution of a hierarchical spatial framework. Environ. Manage 54, 1249–1266.

Overton, J.M., Young, T.C., Overton, W.S., 1993. Using 'found' data to augment a probability sample: procedure and case study. Environ. Monit. Assess. 26, 65–83.

Pacifici, K., Reich, B.J., Miller, D.A.W., Gardner, B., Stauffer, G., Singh, S., McKerrow, A., Collazo, J.A., 2017. Integrating multiple data sources in species distribution modeling: a framework for data fusion. Ecology. 98, 840–850.

Pearl, C.A., Adams, M.J., Bury, R.B., McCreary, B., 2004. Asymmetrical effects of introduced bullfrogs (*Rana catesbeiana*) on native ranid frogs in Oregon. Copeia 11–20, 2004.

Phillips, L., 1984. A theory of requisite decision models. Acta Psychol 56, 29–48.

Phillips, S.J., Dudik, M., 2008. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. Ecography. 31, 161–175.

Pearman-Gillman, S.B., Katz, J.E., Mickey, R.M., Murdoch, J.D., Donovan, T.M., 2020. Predicting wildlife distribution patterns in New England U.S.A. with expert elicitation techniques. Glob. Ecol. Conserv. 21, e00853.
Peterson, J.T., Duarte, A., 2020. Decision analysis for greater insights into the development and evaluation of Chinook salmon restoration strategies in California's Central Valley. Restor. Ecol. 28, 1596–1609.
Peterson, J.T., Evans, J.W., 2003. Quantitative decision analysis for sport fisheries management. Fisheries. (Bethesda) 28, 10–21.
Pourret, O., Naim, P., Marcot, B., 2008. Bayesian Networks: A Practical Guide to Applications. John Wiley & Sons, Hoboken, NJ.
QGIS.org, 2024. QGIS geographic information system. Open Source Geospatial Foundation Project. https://qgis.org.
Ramesh, V., Gupte, P.R., Tingley, M.W., Robin, V.V., DeFries, R., 2022. Using citizen science to parse climatic and land cover influences on bird occupancy in a tropical biodiversity hotspot. Ecography. e06075, 2022.
R Core Team, 2024. R: A language and Environment For Statistical Computing. R Foundation for Statistical Computing. https://www.R-project.org/.
Rieman, B., Peterson, J.T., Clayton, J., Howell, P., Thurow, R., Thompson, W., Lee, D., 2001. Evaluation of potential effects of federal land management on trends of salmonids and their habitats in the interior Columbia River basin. For. Ecol. Manage. 153, 43–62.
Rowe, J.C., Duarte, A., Pearl, C.A., McCreary, B., Galvan, S.K., Peterson, J.T., Adams, M. J., 2019. Disentangling effects of invasive species and habitat while accounting for observer error in a long-term amphibian study. Ecosphere 10, e02674.
Sauer, J.R., Blank, P.J., Zipkin, E.F., Fallon, J.E., Fallon, F.W., 2013. Using multi-species occupancy models in structured decision making on managed lands. J. Wildl. Manag. 77, 117–127.
Sepulveda, A.J., Layhee, M., Stagliano, D., Chaffin, J., Begley, A., Maxell, B., 2015. Invasion of American bullfrogs along the Yellowstone river. Aquat. Invasions. 10, 69–77.
Schielzeth, H., 2010. Simple means to improve the interpretability of regression coefficients: interpretation of regression coefficients. Methods Ecol. Evol. 1, 103–113.
Su, G., Logez, M., Xu, J., Tao, S., Villéger, S., Brosse, S., 2021. Human impacts on global freshwater fish biodiversity. Science (1979) 371, 835–838.
Suhre, D.O., 2010. Dispersal and Demography of the American bullfrog (*Rana catesbeiana*) in a Semi-Arid Grassland. University of Arizona, Tucson. Masters Thesis.
Society for Ecological Restoration, 2024. Restoration Resource Center Project Database. https://ser-rrc.org/restoration-database/.
Street, G.M., Vander Vennen, L.M., Avgar, T., Mosser, A., Anderson, M.L., Rodgers, A.R., Fryxell, J.M., 2015. Habitat selection following recent disturbance: model transferability with implications for management and conservation of moose (*Alces alces*). Can. J. Zool. 93, 813–821.
Venter, O., Sanderson, E.W., Magrach, A., Allan, J.R., Beher, J., Jones, K.R., Possingham, H.P., Laurance, W.F., Wood, P., Fekete, B.M., Levy, M.A., Watson, J.E., 2018. Last of the Wild Project, Version 3 (LWP-3): 2009 Human Footprint, 2018 Release. NATA Socioeconomic Data and Applications Center (SEDAC), Palisades, New York. https://doi.org/10.7927/H46T0JQ4.
Wiens, J.A., 1989. Spatial scaling in ecology. Funct. Ecol. 3, 385–397.
Whitlock, S.L., Lewis, T.M., Peterson, J.T., 2020. Using a Bayesian multistate occupancy model to assess seabird and shorebird status in Glacier Bay. Alaska. Wildl. Soc. Bull. 44, 451–467.
Zanini, F., Pellet, J., Schmidt, B.R., 2009. The transferability of distribution models across regions: an amphibian case study. Divers. distrib. 15, 469–480.
Zhang, Q., Wei, H., Liu, J., Zhao, Z., Ran, Q., Gu, W., 2021. A Bayesian network with fuzzy mathematics for species habitat suitability analysis: a case with limited *Angelica sinensis* (Oliv.) Diels data. Ecol. Modell. 450, 109560.