









CONTRIBUTED PAPER

Using comparative extinction risk analysis to prioritize the IUCN Red List reassessments of amphibians

Pablo Miguel Lucas^{1,2}  | Moreno Di Marco¹  | Victor Cazalis^{3,4}  |
 Jennifer Luedtke^{5,6}  | Kelsey Neam^{5,6}  | Mary H. Brown⁶  |
 Penny F. Langhammer⁶  | Giordano Mancini¹  | Luca Santini¹ 

¹Department of Biology and Biotechnologies "Charles Darwin", Sapienza University of Rome, Rome, Italy

²Departamento de Biología Vegetal y Ecología, Universidad de Sevilla, Sevilla, Spain

³German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig, Germany

⁴Leipzig University, Leipzig, Germany

⁵IUCN SSC Amphibian Specialist Group, Toronto, Ontario, Canada

⁶Re:wild, Austin, Texas, USA

Correspondence

Pablo M. Lucas, Departamento de Biología Vegetal y Ecología, Universidad de Sevilla, Avenida Reina Mercedes, s/n 41012 Sevilla, Spain. Email: plucas1@us.es

Article impact statement: Ensemble modeling of extinction risk is a promising tool for prioritizing species for reassessment and updating the IUCN Red List.

Funding information

Deutsches Zentrum für integrative Biodiversitätsforschung Halle-Jena-Leipzig, Grant/Award Numbers: FZT 118, 202548816; Microsoft, Grant/Award Number: Grant number: 5313.006-AI4E; Re:wild, Grant/Award Number: Grant number: 5313.006-AI4E; MICIU/AEI/10.13039/501100011033 and European Union NextGenerationEU/PRTR, Grant/Award Number: CNS2022-136018; MUR Rita Levi Montalcini program; Junta de Andalucía, Grant/Award Number: EMERGIA20_00135; Sapienza Università di Roma, Grant/Award Number: AR22117A859F7D58; Italian Ministry of Education, University and Research, Programma Operativo Nazionale, Grant/Award Number: CUPB85F21005360001

Abstract

Assessing the extinction risk of species based on the International Union for Conservation of Nature (IUCN) Red List (RL) is key to guiding conservation policies and reducing biodiversity loss. This process is resource demanding, however, and requires continuous updating, which becomes increasingly difficult as new species are added to the RL. Automatic methods, such as comparative analyses used to predict species RL category, can be an efficient alternative to keep assessments up to date. Using amphibians as a study group, we predicted which species are more likely to change their RL category and thus should be prioritized for reassessment. We used species biological traits, environmental variables, and proxies of climate and land-use change as predictors of RL category. We produced an ensemble prediction of IUCN RL category for each species by combining 4 different model algorithms: cumulative link models, phylogenetic generalized least squares, random forests, and neural networks. By comparing RL categories with the ensemble prediction and accounting for uncertainty among model algorithms, we identified species that should be prioritized for future reassessment based on the mismatch between predicted and observed values. The most important predicting variables across models were species' range size and spatial configuration of the range, biological traits, climate change, and land-use change. We compared our proposed prioritization index and the predicted RL changes with independent IUCN RL reassessments and found high performance of both the prioritization and the predicted directionality of changes in RL categories. Ensemble modeling of RL category is a promising tool for prioritizing species for reassessment while accounting for models' uncertainty. This approach is broadly applicable to all taxa on the IUCN RL and to regional and national assessments and may improve allocation of the limited human and economic resources available to maintain an up-to-date IUCN RL.

KEYWORDS

amphibians, climate change, comparative analysis, ensemble model, extinction risk, IUCN, IUCN Red List, prioritize

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2024 The Author(s). *Conservation Biology* published by Wiley Periodicals LLC on behalf of Society for *Conservation Biology*.

INTRODUCTION

Understanding how global change is affecting the extinction risk of species is key to guiding conservation planning and action (Pereira et al., 2010; Urban et al., 2016). The International Union for Conservation of Nature (IUCN) Red List of Threatened Species (RL) is used to monitor the extinction risk of all identified species. Currently, over 150,000 species have been assessed globally (IUCN, 2023), with experts assigning each species a category of extinction risk (RL category) based on a set of quantitative criteria and thresholds (IUCN, 2012). The RL assessments require regular updates, at least once every 10 years, to ensure the information is sufficiently recent to inform and catalyze conservation action (IUCN, 2012). However, resources are insufficient to complete regular reassessments for most species, and, as a result, their RL assessments can become outdated (Cazalis et al., 2022; Rondinini et al., 2014). The long-term sustainability of the RL depends on cost-efficient reassessment strategies (Rondinini et al., 2014), such as the identification of priority species for reassessment based on modeled probability of RL category.

Among terrestrial vertebrates, amphibians have the highest proportion of threatened species and represent a particular challenge for extinction risk monitoring because the number of newly described species grows considerably every year (Ceballos et al., 2017; Luedtke et al., 2023; Tapley et al., 2018) and many species are in rapid decline (Scheele et al., 2019; Sodhi et al., 2008; Wake & Vredenburg, 2008). Climate change and habitat loss are among the major drivers of drastic changes in amphibian conservation status over relatively short periods (Luedtke et al., 2023). Examples are the Atacama toad (*Rhinella atacamensis*), which was listed as least concern in 2010 and vulnerable in 2015, and the southern leopard frog (*Litlobates miadis*), which was classified as vulnerable in 2004 and critically endangered in 2020.

Amphibians' traits, such as small geographic ranges, limited dispersal ability, dependence on water bodies, sensitivity to evaporation, and limited thermoregulatory abilities, render them more susceptible to adverse effects of global changes than other vertebrates (Duellman & Trueb, 1994; Ficetola et al., 2015). Although previous studies show the key role of biological traits, such as body size, geographic range size, and brood size, in predicting amphibian RL category (Cardillo, 2021; Cooper et al., 2008; Fontana et al., 2021; Pincheira-Donoso, Harvey, Cotter, et al., 2021; Sodhi et al., 2008), the effects of climate change on RL category are less known. Developing robust tools that support the identification of species most in need of reassessment is highly valuable for informing global extinction risk monitoring strategies (Cazalis et al., 2022).

Comparative extinction risk models that relate RL categories to extrinsic or intrinsic drivers can be used to predict the status of species in the categories not evaluated and Data Deficient (e.g., Bland & Bohm 2015; Bland et al., 2017; Borgelt et al., 2022; Gonzalez-del-Pliego et al., 2019; Pelletier et al., 2018). They are also important tools in RL assessments with limited resources (Cazalis et al., 2022, 2023) and can help maintain updated IUCN RL assessments (Rondinini et al., 2014). In gen-

eral, comparative analyses of extinction risk have relied on single model algorithms (Bland et al., 2015; Di Marco et al., 2014; Wieringa, 2022; Zizka et al., 2021, 2022). When several models are used, the best model is generally selected based on individual predictive performance (Bland et al., 2015). A single best model may not result in the best predictions, however. Thus, relying on several good performing models and determining predictive uncertainty may be a better strategy (Araujo & New, 2007).

We used an ensemble of comparative extinction risk models to identify amphibian species that should be prioritized for reassessment based on their potential RL category, with the ultimate goal of keeping the RL up to date. Specifically, we evaluated the effects of species' biological traits, environmental variables, and global changes (including climate and land-use change) on species' RL category. In contrast to previous research on amphibians in which similar models were used to predict the status of Data-Deficient species (González-del-Pliego et al., 2019), we used models to identify data-sufficient species whose assessments may require updating. Moreover, we departed from previous studies (e.g., Bland & Bohm 2015; Bland et al., 2017; Borgelt et al., 2022; Gonzalez-del-Pliego et al., 2019; Pelletier et al., 2018) by jointly considering the relative predictive power and overall consistency of an ensemble of predictive models. By combining the mismatch between official and predicted RL categories and uncertainty due to the inconsistency among model predictions, we suggest a prioritization approach for RL reassessments.

METHODS

Biological trait data

We compiled biological trait data from existing data sets (Table 1; Figure 1). In particular, we collected data on body size (snout-vent length [SVL]), brood size, breeding strategy, microhabitat, and habitat generalism. Body size is an important predictor of RL category in amphibians (Cardillo, 2021) and other vertebrate classes (Cardillo, 2003; Cardillo et al., 2008). Brood size can be related to species' recovery ability (Pincheira-Donoso, Harvey, Cotter, et al., 2021). Breeding strategy determines different modes of life development (indirect development, direct development, viviparity) and can be used to determine species' flexibility and adaptability to different environments and could be related to the vulnerability of amphibians to global changes (Brooks & Kindsvater, 2022). Finally, habitat generalism determines species adaptability to land-use change (Carilo Filho et al., 2021).

Using the list of habitat types per species in the RL database (IUCN, 2022), we placed species in 1 of 4 categories based on the number of habitat types a species uses (1, specialist; >1, generalist) and whether the species uses forest (Table 1). We used microhabitat classification from Oliveira et al. (2017) and classified amphibians as semiaquatic, generalist, or other. The generalist and specialist characteristics were described at

TABLE 1 Variables used for comparative analysis of species' extinction risk.

Variable group	Description	Source	Justification
Variable			
Land use and land-use change			
Urbanization	Percentage of urban areas in the species range	C3S CDS, 2019	McKinney, 2002; Newbold et al., 2015
Agriculture	Percentage of agriculture areas in the species range	C3S CDS, 2019	Newbold et al., 2015
Urbanization change	Change in urban in 10 years (urban t_0 – urban t_1)	C3S CDS, 2019	Newbold et al., 2015
Agriculture change	Change in agriculture in 10 years (agriculture t_0 – agriculture t_1)	C3S CDS, 2019	Newbold et al., 2015
Human density	Human population density (humans/km ²)	NASA, 2018	Newbold et al., 2015
Accessibility	Travel time to cities (higher values correspond to more inaccessible areas)	Weiss et al., 2018	Benítez-López et al., 2019
Climate			
Annual temperature	Annual mean air temperature, BIO1	Karger et al., 2017, 2018	Hof et al., 2011; Silvano & Segalla, 2005; Sonn et al., 2019; Thuiller et al., 2019
Temperature seasonality	Temperature seasonality, BIO4	Karger et al., 2017, 2018	Hof et al., 2011; Silvano & Segalla, 2005; Sonn et al., 2019; Thuiller et al., 2019
Annual precipitation	Annual precipitation, BIO12	Karger et al., 2017, 2018	Hof et al., 2011; Silvano & Segalla, 2005; Sonn et al., 2019; Thuiller et al., 2019
Precipitation seasonality	Precipitation seasonality, BIO15	Karger et al., 2017, 2018	Hof et al., 2011; Silvano & Segalla, 2005; Sonn et al., 2019; Thuiller et al., 2019
Climate change			
Change annual temperature	Change in annual mean air temperature, BIO1	Karger et al., 2017, 2018	Hof et al., 2011; Silvano & Segalla, 2005; Urban, 2015; Urban et al., 2016
Change temperature seasonality	Change in temperature seasonality, BIO4	Karger et al., 2017, 2018	Hof et al., 2011; Silvano & Segalla, 2005; Urban, 2015; Urban et al., 2016
Change annual precipitation	Change in annual precipitation, BIO12	Karger et al., 2017, 2018	Hof et al., 2011; Silvano & Segalla, 2005; Urban, 2015; Urban et al., 2016
Change precipitation seasonality	Change in precipitation seasonality, BIO15	Karger et al., 2017, 2018	Hof et al., 2011; Silvano & Segalla, 2005; Urban, 2015; Urban et al., 2016
Context			
Realm	Most representative realm where the species lives, with a numerical code ranging from 1 to 8 representing Australasian, Antarctica, Afrotropical, Indomalayan, Neartic, Neotropical, Oceanian, and Palearctic, respectively	Olson et al., 2001	Sonn et al., 2019; Yackulic et al., 2011
Range and spatial configuration			
Range area	km ² of the species range log transformed and then truncated	IUCN, 2021	Lucas et al., 2019
Range fragments	Number of fragments in which the species range is divided	IUCN, 2021	Lucas et al., 2019
Range circularity	Shape ratio: PH/PO, where PH is total perimeter of idealized fragments with the same area as the observed fragments but with circular shape and PO is actual total perimeter of the fragments; ratio ranges from 0 (most irregular shapes) to 1 (completely circular shapes)	IUCN, 2021	Lucas et al., 2016, 2019

(Continues)

TABLE 1 (Continued)

Variable group	Description	Source	Justification
Range heterogeneity	Proportion of the total range area represented by the largest fragment; range from close to 0 (similar fragment size) to close to 1 (very different fragment size)	IUCN, 2021	Lucas et al., 2019
Phylogeny			
Phylogeny 1	First principal component from phylogeny	Jetz and Pyron 2018	González-del-Pliego et al., 2019
Phylogeny 2	Second principal component from phylogeny	Jetz and Pyron 2018	González-del-Pliego et al., 2019
Biological traits			
Body size	For anurans and salamanders, maximum snout-vent length (SVL) as measure of body size, given that this is the most widely used proxy for body size in these orders; for caecilians, maximum total body length is traditional measure of size (Pincheira-Donoso et al., 2019) and thus the proxy used; all were log transformed	Amado et al., 2021; Pincheira-Donoso, Harvey, Cotter, et al., 2021; Pincheira-Donoso, Harvey, Grattarola, et al., 2021	Amado et al., 2021; Cardillo, 2021; Cardillo et al., 2008
Brood size	Brood size of species log transformed	Oliveira et al., 2017; Pincheira-Donoso, Harvey, Cotter, et al., 2021; Pincheira-Donoso, Harvey, Grattarola, et al., 2021	Pincheira-Donoso, Harvey, Cotter, et al., 2021; Pincheira-Donoso, Harvey, Grattarola, et al., 2021
Breeding strategy	Factor variable indicating whether the species reproduces via direct, larval development or is viviparous means	Oliveira et al., 2017	
Habitat generalist forest	Factorial variable with 4 levels indicating whether the species is a generalist (lives in more than 1 habitat) forest species or nonforest species: generalist forest, generalist nonforest, specialist forest, specialist nonforest	IUCN, 2021	Carilo Filho et al., 2021
Microhabitat	Factorial variable with 3 levels indicating whether the species is generalist, semiaquatic, or others at the microhabitat scale	Oliveira et al., 2017	Carilo Filho et al., 2021

microhabitat level, which may influence the species RL category (Appendix S1).

Due to the continuous and extensive changes in amphibian taxonomy at the species level, different data sets use different taxonomic nomenclature. To correct for this discrepancy, we used the listed synonyms in the IUCN RL and AmphioNom package (Liedtke, 2019) to combine the different data sets. We obtained all listed synonyms per species from Amphibian Species of the World (i.e., the taxonomic authority for amphibians on the RL) (Frost, 2021).

The compiled data set of biological traits had substantial missing information across different variables (Appendix S2). Because nonrandom missing information can result in biases (Nakagawa & Freckleton, 2008, 2011), we filled data gaps by applying a data imputation procedure (Mancini et al., 2023; Penone et al., 2014; Stewart et al., 2023) in which there was a wider set of variables than used for the modeling. Detailed descriptions of the imputation and the data used are in Appendices S3–S6.

Phylogenetic data

We retrieved the consensus phylogenetic tree from Jetz and Pyron (2018) and used the PVR R package (Santos, 2018) to decompose it into a set of orthogonal eigenvectors. Then, we extracted the first 2 eigenvectors, which explained 87% of the total variance (Appendix S7). For species not included in the phylogeny from Jetz and Pyron (2018), we applied a procedure previously used to fill phylogenetic data gaps (Bland et al., 2015; Diniz-Filho et al., 1998). First, we checked the coefficient of variation ($CV = SD/mean$) in genus and family of eigenvectors. After confirming the high similarity of those values (CV_{genus} and CV_{family} for eigenvectors 1 and 2 < 0.05 [Appendix S8]; $CV < 0.05$ considered very similar [Dormann, 2013]) for species not included in Jetz and Pyron's (2018) phylogeny, we assigned the mean value of the eigenvector from the genus. In cases where there were no data for the genus, we assigned the mean value of the family.

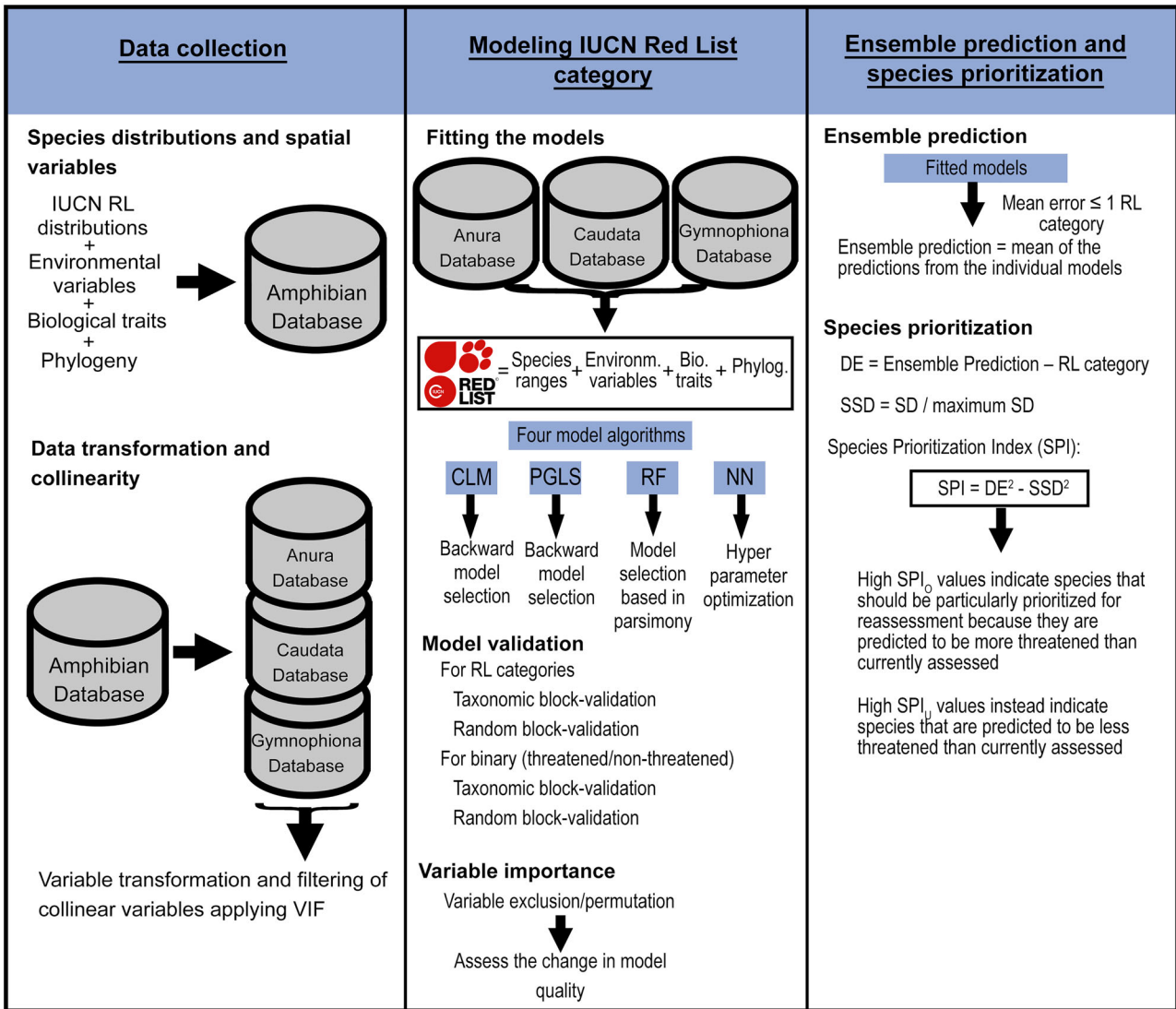


FIGURE 1 Method used to calculate the species prioritization index (SPI) (IUCN, International Union for Conservation of Nature; RL, IUCN Red List; VIF, variance inflation factor; CLM, cumulative link model; PGLS, phylogenetic generalized least squares; RF, random forest; NN, neural network; DE, different in extinction risk; SSD, scaled standard deviation; SPI, species prioritization index; SPI_o , species prioritization index overpredicted; SPI_u , species prioritization index underpredicted).

RL data

We downloaded RL categories from the IUCN RL database in September 2021 (IUCN, 2021). This data set included most assessments from the second Global Amphibian Assessment (Luedtke et al., 2023) ($n = 5647$) and a minority of assessments from the first Global Amphibian Assessment (Stuart et al., 2004) ($n = 1568$) for which the reassessment was not yet publicly available.

Spatial data

Using species’ geographic range, we quantified a series of environmental conditions in the distribution of each species

(Figure 1). We retrieved geographic range polygons for amphibian species from the RL data set, where available (IUCN, 2021). We considered range polygons classified as native or reintroduced in origin and with extant or probably extant presence (IUCN, 2018). We projected the selected polygons with the Lambert cylindrical equal area projection to avoid bias in the calculation of range size and spatial variables when using nonequal area projections (Budic et al., 2016) and then calculated 19 spatial variables in each species’ distribution. We quantified 4 climate variables and recent changes in species’ ranges with CHELSA data sets (Karger et al., 2017, 2018). We selected 4 variables representing climatic parameters relevant during key life-history stages of amphibian species (Cohen et al., 2019; Green, 2017; Lertzman-Lepofsky et al., 2020; Miller et al., 2018; Thuiller et al., 2019): annual mean temperature (BIO1),

temperature seasonality (BIO4), annual precipitation (BIO12), and precipitation seasonality (BIO15). We preferred to select only a limited number of climatic variables to reduce the overall complexity of the models. To estimate recent change in the 4 climatic variables, we calculated the difference between a reference climate calculated for a 30-year period (1965–1994) and the current climate (the closest possible to the time in which the IUCN RL assessment that we used was done [IUCN, 2021]), that is, 2005–2014. We also estimated 2 land-use change variables, urbanization and agriculture, and their recent change in the range of each species (C3S CDS, 2019). Change in these variables was estimated for the 10 years previous to each species' assessment. We calculated human density (NASA, 2018) and accessibility (Weiss et al., 2018) in each species' range because these factors are directly related to overexploitation, habitat loss, and fragmentation (Cardillo et al., 2004; Chen et al., 2019; Sodhi et al., 2008).

We also classified species into different biogeographic realms (Olson et al., 2001) because realm can carry latent information on species exposure to different conditions and threats. Finally, we estimated geographic range attributes, including area, circularity, number of range fragments, and the proportion of area in the largest fragment. The spatial configuration of a species' geographic range is a relevant predictor of extinction risk and RL category (Lucas et al., 2016, 2019). A complete list of variables and their sources is in Table 1. Even though emerging infectious diseases are considered one of the most pressing threats to amphibians worldwide (Luedtke et al., 2023), predicting disease risk remains challenging and is plagued with uncertainty, so we did not consider this threat. However, recent methods have been proposed and could be incorporated in the future (Akçakaya et al., 2023).

We excluded species without published RL range maps. Because our purpose was to prioritize species for reassessment and because we needed the IUCN RL category of the species for our priority index (see “Ensemble prediction and species prioritization” below), we excluded all data deficient species. Thus, we had 5684 species of amphibians in our database.

Data transformation and collinearity

We organized the data into one data set per taxonomic order (5004 Anura, 592 Caudata, 88 Gymnophiona). To meet homoscedasticity and normality assumptions, we transformed the variables with a cumulative link model (CLM) and phylogenetic generalized least squares (PGLS) as reported in Table 1. Predictor variable distribution does not affect random forest (RF) and neural network (NN) models. Finally, we filtered out highly collinear variables with variance inflation factor (VIF) >4 (Figure 1; Appendix S11).

Fitting the models

We generated separate models for anurans, caudates, and caecilians (Figure 1). Differences in ecological and life-history traits of

these 3 orders can make them differentially sensitive to distinct threats (González-del-Pliego et al., 2019; Pincheira-Donoso, Harvey, Cotter, et al., 2021), which in turn leads to different relationships between intrinsic and extrinsic drivers and extinction risk. Not all variables are comparable across these orders, which can lead to different expected relationships (e.g., SVL differs among orders due to different body plans) (Santini et al., 2018).

Following Cazalis et al. (2022), who reviewed recent efforts in comparative extinction risk models, we applied 4 commonly used algorithms: CLM, PGLS, RF, and NN. Each of these model algorithms have different requirements regarding 3 key elements: type of predictor variables, type of response variables, and procedures for variable selection (Christensen, 2020; LeDell et al., 2022; Liaw & Wiener, 2002; Pinheiro et al., 2021). Thus, we applied different transformations to the predictor and response variables and variable selection procedures in line with previous comparative analyses of extinction risk in which these modeling techniques were used (González-Suárez et al., 2012; Lucas et al., 2019; Zizka et al., 2021).

We used the RL categories as response variable: least concern (LC), near threatened (NT), vulnerable (VU), endangered (EN), critically endangered (CR). Although RL categories inherently possess an ordinal nature, only CLM allows the use of ordinal factor variables (Henry et al., 2024; Lucas et al., 2019), so we followed previous studies to adapt the response variable to each model algorithm (Bland et al., 2015; Borgelt et al., 2022; González-Suárez et al., 2012; Mancini et al., 2023; Silva et al., 2022; Soto-Saravia et al., 2021; Zizka et al., 2021, 2022). In RF models, we used RL categories as a factor variable. For NN and PGLS, RL categories were transformed to a numerical variable: 1, LC; 2, NT; 3, VU; 4, EN; 5, CR.

For model selection and the selection of parameters and hyperparameters, we applied different procedures for the 4 model algorithms. Thus, different models may include different variables and different types of relationships (Appendix S11). For CLM, we fitted a full model with all variables with the `clm()` function from the R package `ordinal` (Christensen, 2020), which applies a probit link function. To account for unbalanced distribution among categories, we used weights inversely proportional to the number of species in each category. Over this full model, we performed a stepwise backward model selection procedure with Bayesian information criterion (BIC) with the MASS R package (Venables & Ripley, 2002).

We fitted the PGLS with `gls()` function in the R package `nlme` (Pinheiro et al., 2021). Instead of using phylogenetic eigenvectors as predictors, the PGLS accounts for phylogeny with the `corBrownian` function from R package `ape` (Paradis & Schliep, 2019) to derive a correlation matrix based on species phylogenetic distances (Revell, 2010; Rohlf, 2001). Using the full generalized least square model, we performed a stepwise backward model selection procedure by BIC in the R package MASS (Venables & Ripley, 2002). We fitted the PGLS only to the species in common to the phylogeny of Jetz and Pyron (2018) and the IUCN RL, but we predicted for all species irrespective of their inclusion in Jetz and Pyron (2018).

We fitted RF models with the randomForest R package (Liaw & Wiener, 2002). Although RF models can operate with large numbers of variables, this can lead to an increase in the correlation of trees, reducing the overall performance of the model (Murphy et al., 2010). To avoid this problem, we applied an RF model selection approach with the function `rf.modelSel()` from the R package `rfUtilities` (Evans & Murphy, 2018; Murphy et al., 2010). We used the `mir` option for scaling importance values, a vector 100 percentiles values to test r , a `mtry` value equal to the square root of the number of variables, and 2000 trees.

We fitted the NN with the `h2o` R package (LeDell et al., 2022). NN models depend on many hyperparameters, and selecting the appropriate set of hyperparameters is critical for model performance (Diaz et al., 2017; Torres et al., 2019). There is no a priori set of best hyperparameters; an alternative consists of checking a range of hyperparameters, evaluating the model performance, and selecting the best combination (Diaz et al., 2017). We applied a process of hyperparameter optimization for multilayer artificial NN models that considered 648 different potential models, which resulted from combinations of 3 activation functions (hyperbolic tangent activation function [Tanh], rectifier with dropout [RectifierWithDropout], and hyperbolic tangent activation function with dropout [TanhWithDropout]), 4 options for hidden layers ([349, 174, 87, 29], [174, 87, 29], [87, 29], [27, 9]), 6 input-dropout ratio options (0.05, 0.1, 0.15, 0.2, 0.3, 0.4), 3 options for Lasso regularization (10^{-3} , 10^{-4} , 10^{-5}), and 3 options for ridge regularization (10^{-3} , 10^{-4} , 10^{-5}). The search criteria in the grid of the potential models was done using a random discrete strategy, which is considerably more efficient at reducing the computation time and finding models that are as good or better than a systematic grid search (Bergstra & Bengio, 2012). Random search samples uniformly from the set of all possible hyperparameter value combinations and specifies a stopping criterion, which controls when the random grid search is completed. We used a combination of 3 criteria to stop the grid search: max runtime seconds = 5000, which specified the maximum runtime in seconds for the entire grid; stopping rounds, which stopped the search after 3 training rounds without improving the stopping metric selected in the model ($\log\text{loss} = 1 \times 10^{-2}$); and stopping tolerance, which stopped the search if the ratio between the best moving average and reference moving average of the last models was $\geq 1 \times 10^{-2}$ (LeDell et al., 2022).

To make predictions comparable among the model algorithms and to allow the use of the same validation measures, all model predictions were transformed to an integer variable. For CLM, we selected the integer value with the highest probability in the prediction. For PGLS and NN, the continuous predictions were rounded to integer values, whereas for RF, factorial predictions were transformed to integer values.

Comparative extinction risk analyses that aim to identify the drivers of extinction risk, measured with RL category or population trend (Lucas et al., 2019), usually exclude species classified under criterion B or D to avoid circularity. Indeed, the extent of occurrence and the area of occupancy used for application of criteria B1 and B2, respectively, are highly correlated with geographic range size, which is typically included as model pre-

dictor. Such circularity can lead to an overestimation of the importance of range size, obscuring the role of other important drivers. Instead, when comparative extinction risk analyses have a predictive goal, species classified under criterion B are not excluded so as to generate the best possible predictions (see, e.g., Zizka et al. [2021] and Caetano et al. [2022]). Our main goal was to obtain the best predictions of RL category and identify potential mismatches with official RL assessments among the same species used in model training, rather than evaluating the importance of range size compared with other variables. For this reason, we retained species assessed under criterion B (36.12% of all species included in our analyses, $n = 2053$) or D (3.20% of all species included in our analyses, $n = 182$).

Variable importance

To calculate the variable importance for CLM and PGLS, we excluded each variable and assessed the resulting change in the models' log likelihood (Breiman, 2001; Lu & Ishwaran, 2021; Williamson et al., 2021). For RF models, we used the mean decrease accuracy (MDA), which is an index indicating how much the accuracy decreases when the variable is excluded (Genuer et al., 2010; Hong et al., 2016). For NN models, we used the method implemented by Gedeon (1997), which considers the weights connecting the input features to the first 2 hidden layers. Then, we calculated variable importance by taxonomic order and for all amphibians, averaging individual variable importance from all algorithms. When a variable was not included in a specific model, we assigned it a value of 0 (Barton, 2009; Burnham & Anderson, 2002; Lukacs et al., 2010). All importance values were rescaled to 100 for comparability.

Model validation

We evaluated each model's performance with taxonomic-block validation, iteratively extracting one family from the data set for testing and then fitting the model on the remaining families (Roberts et al., 2017). This approach is more robust to assess models' predictive performance compared with the random cross-validation used in several previous studies (González-del-Pliego et al., 2019) because random cross-validation tends to produce overoptimistic results due to the autocorrelation between phylogenetically close species. However, to compare our results to previous comparative extinction risk models on amphibians, we also ran a classical random cross-validation in which we extracted 10% of the species in each run and repeated the operation 10 times. For each predicted set, we validated the models at 2 different levels. First, the models were validated at the RL category level by calculating the overall accuracy (rate of correct classification) and the mean classification error (absolute value of the difference between predicted and observed categories) for all categories and the overall accuracy, sensitivity (rate of correct classification of threatened species), specificity (rate of correct classification of nonthreatened species), and true skill statistic ($\text{TSS} = \text{specificity} + \text{sensitivity} - 1$). To

ensure comparability with previous studies of extinction risk (González-del-Pliego et al., 2019), although models were fitted using all categories, we aggregated models' predictions into binary classes of nonthreatened (LC and NT) versus threatened (VU, EN, CR) and calculated the overall accuracy, sensitivity, specificity, and TSS.

Ensemble prediction and species prioritization

We compared the models' predictions with the official RL categories (Figure 1). Any mismatch between predictions and official RL categories can pinpoint species that might be more or less threatened than officially reported by the RL. Assuming the models capture genuine drivers of RL category, a species that is predicted to be in a more threatened RL category than officially reported by the RL can have 2 possible interpretations: the species was misclassified due to lack of sufficient data at the time of the assessment or even though the species is not threatened, it is expected to be particularly vulnerable due to either intrinsic traits or extrinsic factors and hence more likely than other species to experience a change in RL category over time. In either case, the species should be prioritized for reassessment. The ranking of priorities, however, should also account for model uncertainty. At equal mismatch value, high model uncertainty should correspond to lower priority.

We used an ensemble forecasting approach to predict RL category (Araujo & New, 2007). To calculate the ensemble prediction, we combined the predictions of all models that achieved an acceptable predictive performance. We considered predictive performance acceptable when mean error was <1.00 during family-block validation (i.e., <1 category mismatch). Using these subsets of models, we calculated the ensemble prediction for each species as the mean of the predictions from the individual models. In addition, we calculated the standard deviation of RL category prediction among the subset of models. Then, we calculated the difference in extinction risk (DE) with the ensemble prediction and the current RL category of the species:

$$DE = \text{ensemble prediction} - \text{RL category.} \quad (1)$$

A positive difference represents species predicted to be more threatened than the published RL category (i.e., overpredicted species). Conversely, a negative difference indicates species predicted to be less threatened than the published RL category (i.e., underpredicted species).

To account for the variability across model predictions, we also calculated the scaled standard deviation (SSD) per species as the SD of the predictions for the species divided by the maximum SD across all species in each taxonomic order, which resulted in a scaled value from 0 to 1. Then, we used the DE and the SSD to calculate a species prioritization index (SPI). To differentiate overpredicted species and underpredicted species, we applied the index separately to these 2 groups obtaining 2 SPIs: species prioritization index overpredicted (SPI_O) for over-

predicted species (species for which their published RL category showed a lower RL category than our predicted category) and species prioritization index underpredicted (SPI_U) for underpredicted species (species for which their published RL category was higher than our predicted category):

$$SPI = DE^2 - SSD^2. \quad (2)$$

High SPI values were represented by species with high DE and low values of SSD among the predictions of the algorithms included. High SPI_O values indicated species that should be particularly prioritized for reassessment because they were predicted to be more threatened than currently assessed (according to the ensemble prediction). High SPI_U values instead indicated species that were predicted to be less threatened than currently assessed.

To visualize the spatial pattern of assessment priorities, we intersected species range maps with a grid of 100×100 km with the Lambert cylindrical equal area projection (Harfoot et al., 2021) in ArcMap 10.3 (ESRI, 2008) and calculated for each grid cell the average SPI_O (the sum of SPI_O corrected by the number of species on the same grid to control for the species richness) and average SPI_U (the sum of SPI_U corrected by the number of species on the same grid to control for the species richness). To determine whether there was a taxonomic bias in the SPI values, we calculated SPI_O and SPI_U for each family.

Validation of the predicted priorities based on recent reassessments

We used 1772 species reassessed in the latest update of the IUCN RL (IUCN, 2023; Appendices S9 & S10) to assess the validity of our SPI for guiding reassessments. We compared the SPI values of these species with their change in RL categories (uplisting to a higher category, no change, downlisting to a lower category). We fitted a multinomial model explaining the change in the RL category of species as a function of the SPI, with 3 response levels: no change ($n = 1395$), uplist ($n = 153$), and downlist ($n = 243$). Because the SPI does not differentiate the directionality of change in RL, we also tested whether the previously calculated DE was different for uplisted species, species with no change, and downlisted species. We also repeated our multinomial model with DE, instead of SPI, as the predictor variable. Multinomial models were fitted using the *nnet* R package (Venables & Ripley, 2002), and we used Wald tests in the *RVAideMemoire* R package (Herve, 2023) to test for significance of coefficients.

RESULTS

Predictors of RL category in amphibians

Using the 4 fitted algorithms, the most important group of variables explaining RL category for all amphibians were variables

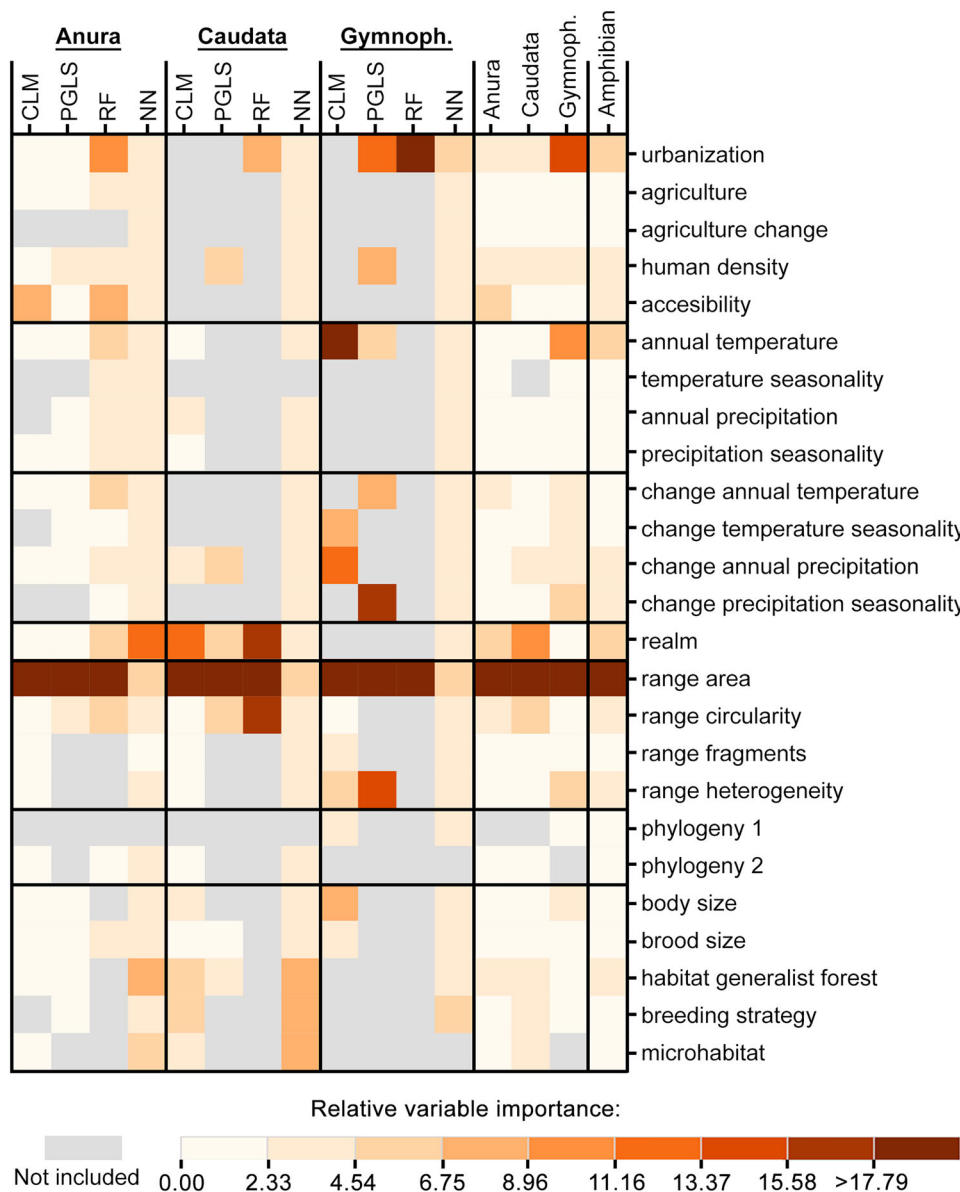


FIGURE 2 Variable importance for 4 model algorithms (cumulative link models [CLM], random forest [RF], phylogenetic generalized least square models [PGLS], neural network [NN]) modeling the International Union for Conservation of Nature Red List categories for each amphibian order, for the averaged importance in each order, and for the averaged importance for the 3 orders and 4 model algorithms (amphibians). Variables (land use and land-use change, climate, climate change, realm, range and spatial configuration, phylogeny, biological traits) are ordered by variable group following Table 1 and separated by black horizontal lines. Values of variable importance for binary variables derived from factor variables are summed in the original factor variable. Because different models could include different variables, when a variable was not included in a specific model (gray), a value of zero was assigned for the calculation of average importance at the level of taxonomic order and the class amphibians.

describing the geographic range area and the spatial configuration (50.71%), followed by land-use and land-use change variables (14.05%), biological traits (10.06%), climate change variables (9.53%), climate variables (8.63%), the context in which the species is present (realm, 5.36%), and phylogeny variables (1.67%) (Figure 2; Appendix S12). We observed that range area, urbanization, and range circularity were consistently important across the 3 amphibian orders, but some variables

were important only for certain taxa, such as accessibility for Anura, realm for Caudata, or annual temperature and change precipitation seasonality for Gymnophiona. As expected, range area was negatively associated with RL category (i.e., species with larger ranges were less likely to be threatened), and lower accessibility (i.e., less travel time to cities, less distance from cities) and higher urbanization were associated with higher RL category (Appendices S12–S24).

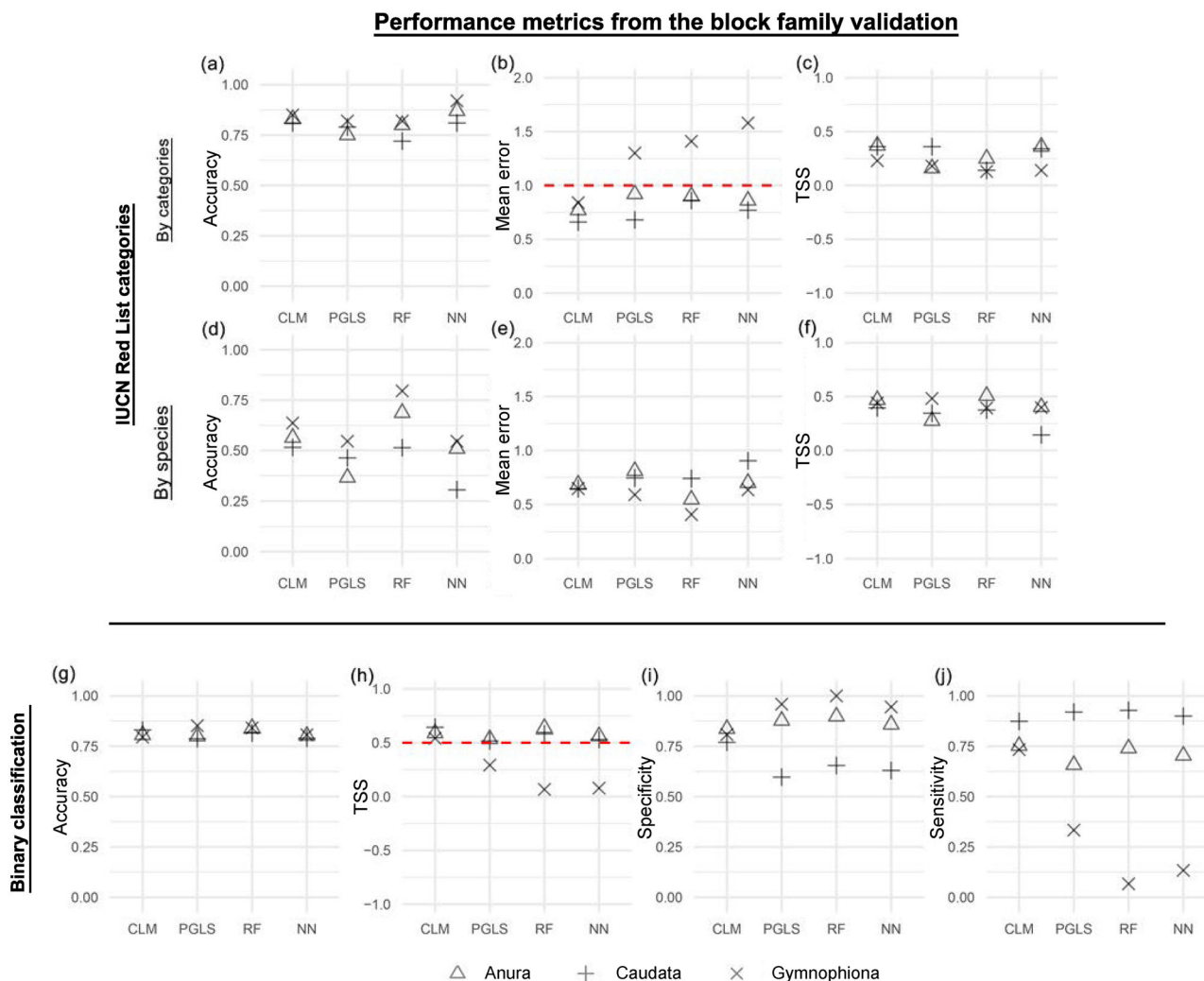
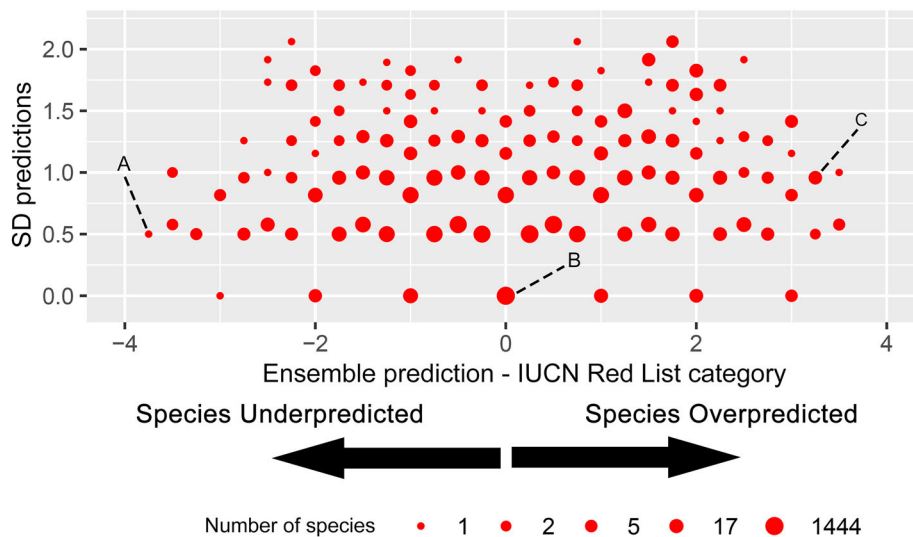


FIGURE 3 Performance metrics from the family block validation for 4 model algorithms (cumulative link models [CLM], random forest [RF], phylogenetic generalized least square models [PGLS], neural network [NN]) modeling the International Union for Conservation of Nature (IUCN) Red List (RL) categories for each amphibian order (a–f) considering IUCN RL categories and (g–j) considering binary classification (threatened or nonthreatened) risk categories: accuracy (rate of correct classification in all categories), mean error (absolute value of the difference between predicted and current RL categories), and true skill statistic (TSS) ($TSS = specificity + sensitivity - 1$) (a–c) averaged among IUCN RL categories (independently of the number of species in each category) and (d–f) by the number of species (b), mean error averaged among categories was used to exclude models for the ensemble prediction when the value was ≥ 1.00 ; red line, mean error = 1.00; [d–f], values averaged among species provide a comparison with classical performance metrics reported in comparative studies of extinction risk) and (g–j) accuracy, TSS, specificity (rate of correct classification of nonthreatened species), and sensitivity (rate of correct classification of threatened species) reported for each amphibian order and each model algorithm (red line, TSS = 0.5, which indicates an accurate model).

Validation for RL categories

Family-block validation showed an average accuracy across categories (independent of the number of species in each category) and models of 0.81 (SD 0.10) (Figure 3a; Appendices S25 & S26) and an average TSS of 0.25 (0.22) for all models. Caudata models showed the highest TSS values (0.30 [0.21]), followed by Anura (0.28 [0.20]) and Gymnophiona, which showed the lowest TSS values (0.17 [0.24]) among all models. Among the models, CLM models performed best (Figure 4c). The average mean error was 0.96 (0.65), therefore just less than one category on average, although it was highly variable across the 3 orders (Figure 3b).

When we averaged the values by the number of species (instead of by categories), our independent family-block validation showed an average accuracy of 0.54 (SD 0.13) among all models (Figure 3d). RF showed the best results (Appendix S27). The TSS averaged across species showed a mean value of 0.39 (0.10) for all models. Anura models showed the best TSS values, followed by Caudata and Gymnophiona, which showed the lowest TSS values among all models. The CLM models showed the best TSS (Figure 3f). Average mean error among all models was 0.67 (0.13). The RF models showed on average the lowest mean error (0.57 [0.17]), followed by CLM (0.66 [0.02]), PGLS (0.72 [0.11]), and NN (0.75 [0.14]). The mean error was lowest for Gymnophiona (0.57 [0.11]), followed by






<p>(a) <i>Litoria booroolongensis</i></p> 	<p>(b) <i>Dermophis mexicanus</i></p> 	<p>(c) <i>Nototriton matama</i></p> 
<p>IUCN category: CR Predicted category: LC SPI justification: this species should be prioritized for reassessment because of the big mismatch between predicted and IUCN Red List category. IUCN considers that the very small area of occupancy is the main cause to include this species as CR. Our models do not include this measure of range size which may explain the mismatch among predicted and IUCN categories.</p>	<p>IUCN category: LC Predicted category: LC SPI justification: this species should not be prioritized for reassessment as our prediction and the IUCN Red List category agree. LC prediction and category are based on its big range, their probable large population, and tolerance to a range of habitats.</p>	<p>IUCN category: LC Predicted category: EN SPI justification: this species should be prioritized for reassessment because there is a big mismatch between predicted and the IUCN Red List category. The IUCN considers that although the extent of occurrence of its range is very small (10 km²), the population is stable and it occurs in a protected area so it was classified as LC. Instead our models indicate that based on environmental characteristics and species traits of the species should be considered as EN.</p>

FIGURE 4 Distribution of the number of species by difference in extinction risk (DE) value (ensemble prediction—red-list category) and standard deviation (SD) of the predictions. Because DE and SD values stem from integer values of the predicted category and the red-list category, many species overlap, so the number of species per point is represented by different point sizes (SD 0, species with only one model with TSS > 0.4 [Gymnophiona]) (IUCN, International Union for Conservation of Nature; LC, least concern; EN, endangered; CR, critically endangered; SPI, species prioritization index). Photo A by Matthew Clancy, photo B by José G. Martínez-Fonseca, and photo C by Eduardo Boza-Oviedo.

Anura (0.69 [0.11]) and Caudata (0.76 [0.11]) (Figure 3e). Random cross-validation led to substantially better estimates. Mean error by category averaged among categories was 0.83 (0.56) (Appendices S28–S31).

Validation for binary (threatened or nontreated) outcome

When considering family-block validation for binary classification (threatened or nontreated species), the average accuracy was 0.81 (SD 0.02) across all models (Figure 3g–j; Appendix S32). Accuracy values were very similar among the

4 model algorithms (SD 0.01). Average TSS across all models was 0.46 (0.20). There was low variability in TSS among different model algorithms (SD 0.09), but variability was high among orders (SD 0.19). Among all algorithms tested, CLM resulted in the highest TSS (TSS = 0.59 [0.05]). In general, all models showed a good balance for specificity and sensitivity values, except for PGLS, RF, and NN models for gymnophionans. The CLM models showed on average the least absolute difference between specificity and sensitivity in the 3 orders (Figure 3g–j). Random cross-validation showed a TSS for all models of 0.58 (0.19), 0.63 (0.07) for anurans, 0.68 (0.08) for caudates, and 0.44 (0.28) for gymnophionans (Appendices S33 & S34).

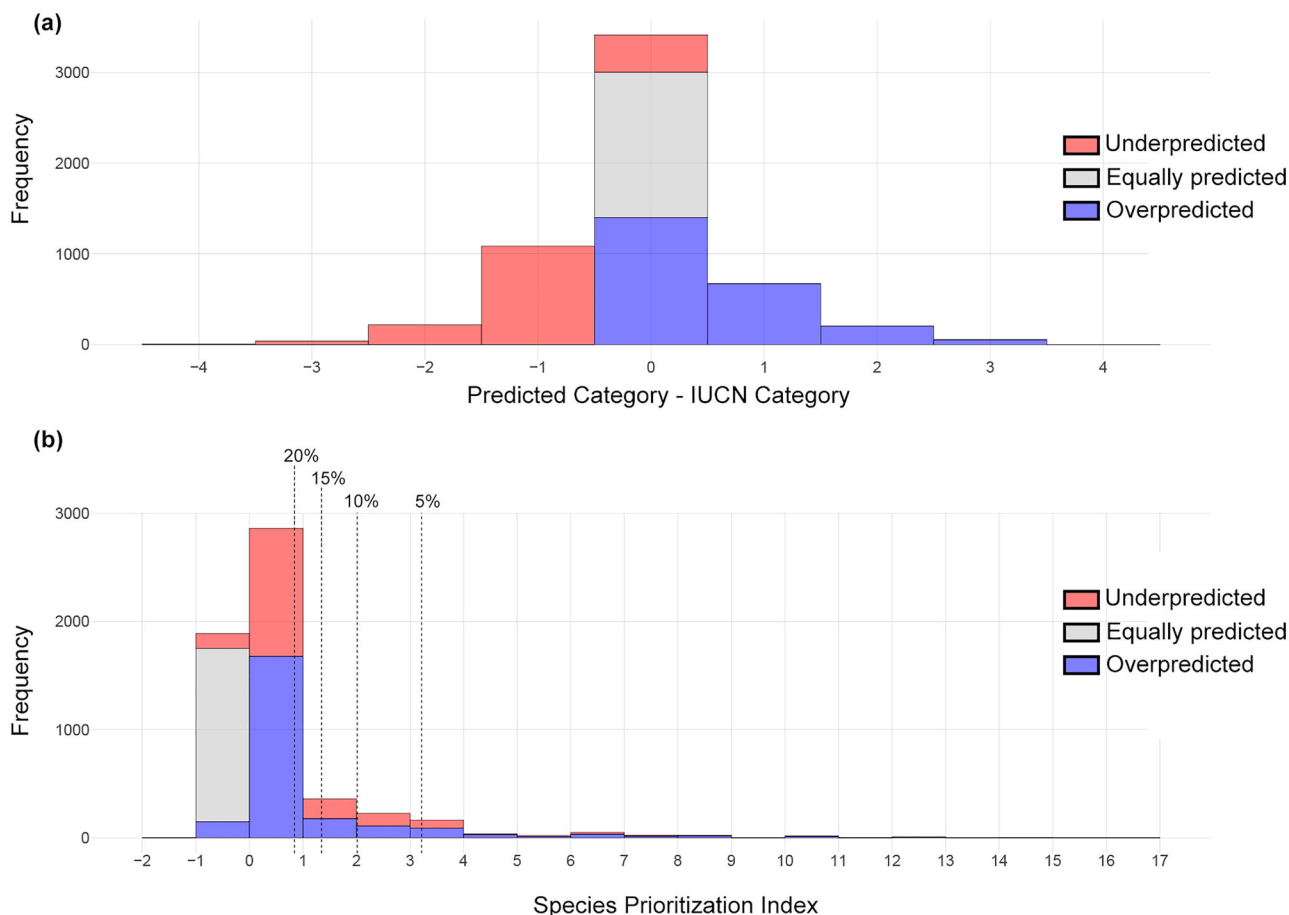


FIGURE 5 (a) Difference between the ensemble prediction and the current International Union for Conservation of Nature Red List category and (b) the species prioritization index (SPI) (percentages in [b] describe the quantiles for the SPI).

Ensemble prediction and species prioritization

Based on the mean error averaged among categories and using the RL categories as a response variable (see above), we excluded 3 models with high errors (≥ 1.00) (Figure 3a; Appendix S26) when we calculated the ensemble prediction: PGLS, RF, and NN models for gymnophionans. The ensemble prediction indicated that 40.22% of species ($n = 2286$) might be threatened with extinction, compared with 38.18% of species ($n = 2170$) assessed as such for the modeled species (Appendices S35). According to our predictions, 40.89% ($n = 2324$) species were overpredicted ($DE > 0$) and 30.95% ($n = 1759$) were underpredicted ($DE < 0$), with 28.17% ($n = 1601$) of species having the same predicted values as the published RL category (Figures 3 & 4a). However, most of DE values were relatively small, only 16.52% ($n = 939$) had a $DE > 1$ or < -1 (1 indicates a difference of one category), and 3.24% ($n = 184$) had a $DE > 2$ or < -2 (Figure 5a).

The distribution of SPI was positively skewed; values ranged from 14.00, indicating a high priority for reassessment, to -0.64 , indicating a low reassessment priority (Figure 4b; Appendix S35). Average SPI_U was concentrated mainly in Central America, the Andes, the Caribbean, the West coast of the United

States, southwestern Europe, Southeast Asia, and southeastern Australia (Figure 6a). Average SPI_O was concentrated mainly in Central America, the Andes, and Southeast Asia (Figure 6b). Rhinodermatidae and Cryptobranchidae had substantially higher SPI_U values than the average SPI, whereas Ichthyophiidae and Siphonopidae had substantially higher SPI_O values than the average SPI (Appendix S36).

Validation of the predicted priorities based on recent reassessments

Species that recently had their RL category changed had higher SPI (mean 1.08 [SD 1.94]) Appendix S37) than species that had not had their RL category changed (0.34 [1.07]). The multinomial model confirmed that species that experienced change in RL category (either in the uplist or downlist group) had a significantly higher SPI than species in the no-change group ($p < 0.001$) (Appendix S38).

The directionality of the changes confirmed the validity of our predictions. Uplisted species had a significantly higher DE (mean 0.13 [SD 1.10]) than species with categories that did not change (0.04 [0.64]) and downlisted species (-0.66 [0.88])

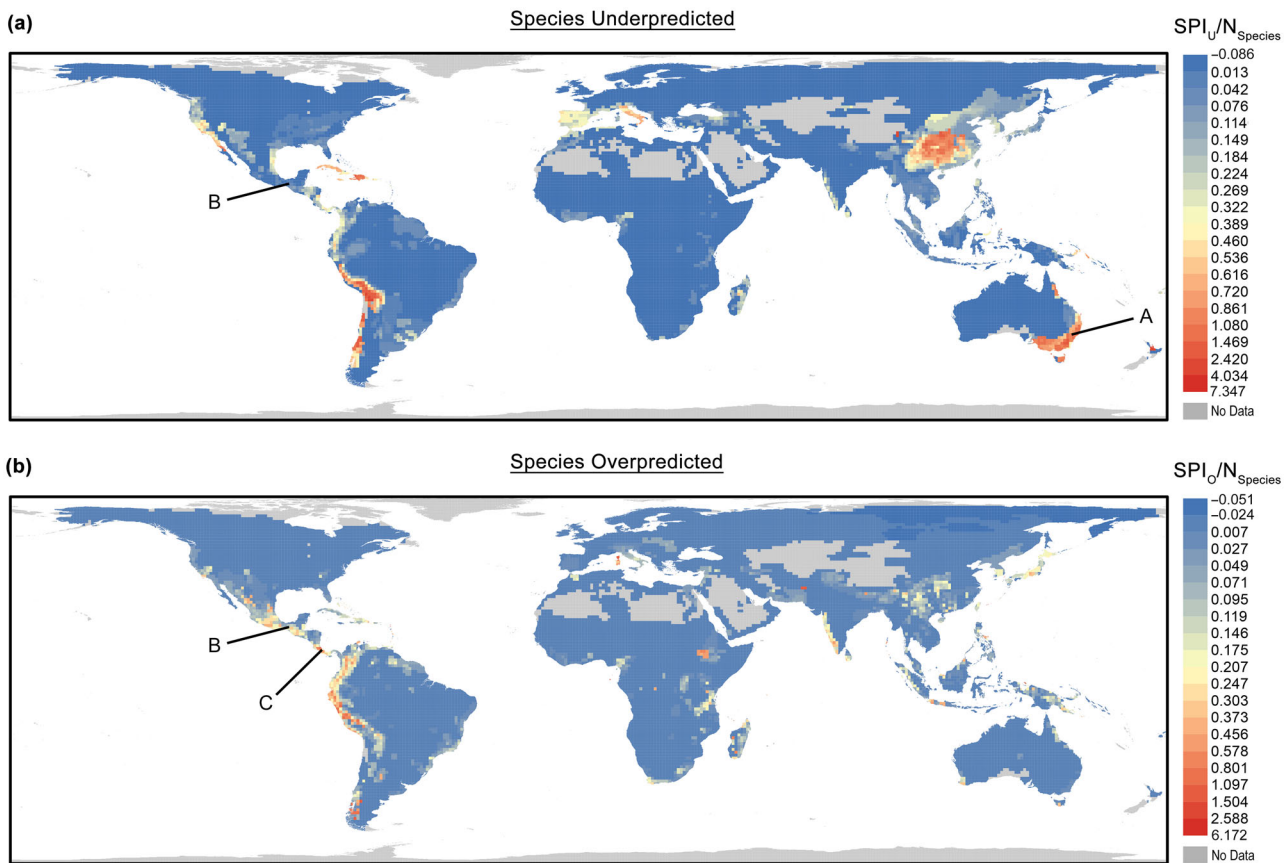


FIGURE 6 Results for the species prioritization index (SPI): (a) underpredicted species (SPI_U) (i.e., species whose red-list category is higher than predicted by the model) and (b) overpredicted species (SPI_O) (i.e., species with an International Union for Conservation of Nature [IUCN] Red List category lower than that predicted by the model). The letters A, B, and C point to the location of the species A, B, and C in Figure 4, which are examples of underpredicted, equally predicted (species with an IUCN Red List category equal to that predicted by the model), and overpredicted species.

(Appendix S39). However, the difference in mean DE between downlisted species and no-change species was not significant ($p > 0.05$).

DISCUSSION

We presented an ensemble of comparative predictive models to assess their potential to prioritize amphibian species for reassessment on the IUCN RL. We proposed an approach to predicting species RL category that combines the predictions of models with sufficient predictive power while accounting for their uncertainty. This approach prioritizes the reassessment of species whose predicted RL category differs the most from its published category on the RL and for which predictions are consistent among models.

Correlates of RL category in amphibians

The most important variable across our models was range size, which is not surprising given the large proportion of species classified as threatened under criterion B (IUCN, 2012). This

variable is commonly used in comparative extinction risk analyses (Chichorro et al., 2019) of, for example, amphibians (Cooper et al., 2008; González-del-Pliego et al., 2019; Sodhi et al., 2008). Other range-related variables were also important, particularly species' realm, which is consistent with studies that show RL category is highly structured in amphibians (González-del-Pliego et al., 2019), as was to some extent the shape of the range (range circularity), as has been reported for terrestrial vertebrates (Lucas et al., 2016, 2019). Among human pressure variables, urbanization and accessibility were the most important factors, probably providing proxies for population declines due to human activities (Cardillo, 2021; González-del-Pliego et al., 2019; Sodhi et al., 2008). A novel and important result was the great importance of climate and climate change covariates. Climate change effects are emerging as a serious threat to this group of species (Cohen et al., 2019; Lertzman-Lepofsky et al., 2020; Loarie et al., 2009; Miller et al., 2018), but the lack of information on species-specific effects of climate change makes it difficult for RL assessors to take climate change into account (Cazalis et al., 2022; Foden & Young, 2016; Mancini et al., 2024). In addition, climate may act as a mediating factor of the effect of direct drivers, such as habitat loss or chytrid fungal infections (Mantyka-pringle et al., 2012; Sonn et al., 2019). Biological

traits were not very important in general. Phylogenetic eigenvectors were not important in our models, and PGLS did not perform better than other models, suggesting that phylogeny may be a proxy of other covariates used in our analyses, such as climate or trait variables not included in previous studies (e.g., González-del-Pliego et al., 2019).

Performance of the different model algorithms

No modeling approach consistently outperformed the others, but several approaches produced good results for the 2 most speciose taxonomic orders (Anura and Caudata). The overall accuracy of our models at the category level was high with predictions diverging by <1 RL category from the published assessments on average. It should be considered, however, that when predictions are made for species in families not included in the training data set, our family-block validation attempts a validation with independent data sets providing a more realistic measure of performance. As expected, random cross-validation led to substantially better estimates, a result consistent with those from previous comparative extinction risk analyses of mammals (Bland et al., 2015), reptiles (Caetano et al., 2022), and plants (Zizka et al., 2021). Compared with the latest analysis on amphibians (González-del-Pliego et al., 2019), which accurately predicted the RL category for 20% of species, our 12 models showed a substantial improvement in the accuracy of the RL category: 54% of species were accurately predicted. The TSS values at the category level for our models were not high, but they were slightly better than previous predictions of RL category in birds and mammals (Santini et al., 2019).

The performance of the different algorithms showed important variations among the 3 amphibian orders. RF and NN models performed better for anurans and caudates, suggesting that sample size may be a limiting factor in these complex models (Tange et al., 2017; Vabalas et al., 2019). This may explain why simpler models, such as the CLM, outperformed the other model algorithms for gymnophionans. The CLM algorithm was the only one in which ordinal factor variables were used as the published IUCN RL categories. The other 3 algorithms required modifications of the response variable. In addition, we modified the predicted values for model validation and comparison between the different model algorithms; thus, those changes may have had an effect on the estimation of errors. We used 4 commonly used model algorithms in comparative analyses of extinction risk (Cazalis et al., 2022). Other less tested but promising model algorithms, such as the recently proposed XGBoost (Chen & Guestrin, 2016), could be added in the future.

Applications, limitations, and future steps

Spatial and taxonomic patterns in SPI can be informative for reassessment prioritization, as well as interpretation of models' limitations. A mismatch between observed and predicted categories can arise from 3 mechanisms: misclassification of the

species due to a lack of or incorrect data; change in RL category of the species; or imperfect model fit. Our models included a number of proxies of intrinsic vulnerability and extrinsic factors that may lead to changes in the parameters used by the RL. Hence, if the model captures indirect relationships between such variables and species RL category, a mismatch can indeed highlight a possible change in the conservation status. The selection of the best-performing models and the prioritization of coherent predictions across algorithms minimize the effects of imperfect model fit mechanism.

An overprediction of RL category can help in the identification of species that may have a higher RL category than the current assessment, even if not currently threatened. This may be the case for many LC or NT species, which may be prioritized for reassessment. This can result from the use of ancillary information in the model (e.g., climate change, land-use change, etc.) that was not directly used in the assessment process because it was unavailable to the assessors or because they were unaware of the information. Overpredicted species that are currently LC or NT could be monitored for emerging threats that would qualify them for a threatened category. Overprediction can also be a modeling artifact; it ignores the complex and composite nature of RL assessments that typically require multiple subcriteria to be met for a species to be considered threatened (Di Marco, 2022) (e.g., restricted range size and continuing population decline).

In contrast, underpredictions are likely a derivative, in most cases, of model simplification. Among the most pressing threats to amphibians are pathogens and invasive species (Scheele et al., 2019; Stuart et al., 2004, 2008), which could not be included explicitly in our models and might explain those mismatches and likely explain the majority of underpredictions. Although less probable, it is conceivable that underpredictions could signal errors in species RL classification, warranting potential revision. Hence, with due caution, underpredicted species could be candidates for downlisting, which may be of interest for zoos and other institutions working in conservation.

Finally, mismatches may also indicate possible inconsistencies in the assessments of different families or genera or in different regions of the world assessed by different groups of experts. Such inconsistencies can, for example, arise from the predominant application of certain criteria or data types used (or not used) in the assessments. All in all, the SPI values for species should be assessed on a case-by-case basis by experts to help define future reassessment priorities. Assessors may undoubtedly have additional factors to consider in the prioritization, which may span from available funding for certain regions or taxonomic groups to groups that have recently undergone taxonomic revision. In these cases, SPI can be used to prioritize preselected groups of species based on other criteria to optimize the efforts required to maintain up-to-date RL assessments under limited available resources (Cazalis et al., 2022; Rondinini et al., 2014).

Our validation with recently reassessed species supports the validity of our approach for guiding future reassessment efforts. Our SPI accurately predicted species that recently had the RL category changed (Appendix S38) and proved partic-

ularly good at differentiating downlisted species from species with categories that did not change (Appendix S39). Amphibians were first comprehensively assessed on 2004 (Stuart et al., 2004), and despite the IUCN commitment to reassess species within 10 years, the global reassessment was published in 2023 (Luedtke et al., 2023), a delay due partially to the high rate of new species descriptions, the taxon's rapid response to environmental stressors, and the limited resources available for reassessment. Considering the challenge of maintaining up-to-date assessments for such a large group of species, our approach can help accelerate and enhance the effectiveness of future reassessment efforts.

Our approach is broadly applicable to all taxa included in the IUCN RL and to regional or national assessments. Using SPI may improve allocation of limited human and economic resources available to maintain an up-to-date IUCN RL and highlight future changes in the conservation status of species, overall allowing for quicker and more effective conservation decisions to minimize biodiversity loss. Overall, our approach would help reduce the taxonomic and geographic bias associated with the reassessed species by reducing the time and costs invested on common species. Furthermore, reducing reassessment efforts might free up resources for new assessments, hence reducing existing geographic and taxonomic biases. Our prioritization index should not be used to exclude species for reassessment, but it can help assessors determine which species should be prioritized for more frequent reassessments. In addition, the prioritization index should be updated for each new reassessment or when new relevant information about predictors is available. Finally, we recommend that developing such reassessment priority studies with specialist groups (in the case of amphibians with the Global Amphibian Assessment team) is useful to ensure they use the results with due consideration of the strengths and weaknesses of the reassessment study. Future steps to improve these models and their usability include the reevaluation of predictions based on new assessments (Di Marco, 2022) and the identification of mechanisms that led to a mismatch between predicted and observed RL categories to identify inherent biases in either the modeling process (e.g., omission of a relevant variable or inclusion of a misleading one) or the assessment process (e.g., omission of a relevant information, such as climate change). Incorporating diseases (e.g., Akçakaya et al., 2023), future climate change scenarios, and real-time threat data (e.g., deforestation alerts) will also be critical next steps in refining the prioritization process for amphibian reassessments.

ACKNOWLEDGMENTS

We are sincerely grateful to the working group sRedList for their helpful comments on this manuscript. We also thank the funding of sRedList by sDiv, the Synthesis Centre of the German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig (funded by the German Research Foundation [FZT 118, 202548816], which also funded the position for V.Ca.). L.S. acknowledges support from the MUR Rita Levi Montalcini program. We thank Re:wild, Microsoft Corporation (grant 5313.006-AI4E), the Emergia contract funded by the Junta

de Andalucía (EMERGIA20_00135), and the grant CNS2022-136018 funded by MICIU/AEI/10.13039/501100011033 and by European Union NextGenerationEU/PRTR, which provided funding for the position of P.M.L.G.M. acknowledges support from the Italian Ministry of Education, University and Research, Ptrogramma Operativo Nazionale (CUP B85F21005360001). We also thank Sapienza University of Rome for funding this work (AR22117A859F7D58). We acknowledge M. Clancy, J. G. Martínez-Fonseca, and E. Boza-Oviedo for their photos used in Figure 4.

ORCID

Pablo Miguel Lucas  <https://orcid.org/0000-0003-4517-9748>

Moreno Di Marco  <https://orcid.org/0000-0002-8902-4193>

Victor Cazalis  <https://orcid.org/0000-0003-0850-883X>

Jennifer Luedtke  <https://orcid.org/0000-0001-5162-978X>

Kelsey Neam  <https://orcid.org/0000-0003-1994-4346>

Mary H. Brown  <https://orcid.org/0000-0003-3974-0172>

Penny F. Langhammer  <https://orcid.org/0000-0002-0124-343X>

Giordano Mancini  <https://orcid.org/0000-0001-9208-4253>

Luca Santini  <https://orcid.org/0000-0002-5418-3688>

REFERENCES

- Akçakaya, H. R., Neam, K., Hobin, L., Lötters, S., Martel, A., & Pasmans, F. (2023). Assessing the extinction risks of amphibians impacted by infectious diseases. *Biological Conservation*, 284, 110205.
- Amado, T. F., Martínez, P. A., Pincheira-Donoso, D., & Olalla-Tárraga, M. Á. (2021). Body size distributions of anurans are explained by diversification rates and the environment. *Global Ecology and Biogeography*, 30(1), 154–164.
- Araujo, M. B., & New, M. (2007). Ensemble forecasting of species distributions. *Trends in Ecology & Evolution*, 22(1), 42–47.
- Barton, K. (2009). *Mu-MIn: Multi-model inference*. R package version 0.12.2/r18. <http://R-Forge.R-project.org/projects/mumin/>
- Benítez-López, A., Santini, L., Schipper, A. M., Busana, M., & Huijbregts, M. A. J. (2019). Intact but empty forests? Patterns of hunting-induced mammal defaunation in the tropics. *PLoS Biology*, 17(5), e3000247.
- Bergstra, J., & Bengio, Y. (2012). Random search for hyper-parameter optimization. *Journal of Machine Learning Research*, 13, 281–305.
- Bland, L. M., Bielby, J., Kearney, S., Orme, C. D. L., Watson, J. E. M., & Collen, B. (2017). Toward reassessing data-deficient species. *Conservation Biology*, 31(3), 531–539.
- Bland, L. M., Collen, B., Orme, C. D. L., & Bielby, J. (2015). Predicting the conservation status of data-deficient species. *Conservation Biology*, 29(1), 250–259.
- Borgelt, J., Dorber, M., Høiberg, M. A., & Veronesi, F. (2022). More than half of data deficient species predicted to be threatened by extinction. *Communications Biology*, 5(1), 679.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5–32.
- Brooks, G. C., & Kindsvater, H. K. (2022). Early development drives variation in amphibian vulnerability to global change. *Frontiers in Ecology and Evolution*, 10, 813414.
- Budic, L., Didenko, G., & Dormann, C. F. (2016). Squares of different sizes: Effect of geographical projection on model parameter estimates in species distribution modeling. *Ecology and Evolution*, 6(1), 202–211.
- Burnham, K. P., & Anderson, D. R. (Eds.). (2002). *Model selection and multi-model inference. A practical information-theoretic approach* (2nd ed.). Springer.
- Caetano, G. H. D. O., Chapple, D. G., Grenyer, R., Raz, T., Rosenblatt, J., Tingley, R., Böhm, M., Meiri, S., & Roll, U. (2022). Automated assessment reveals that the extinction risk of reptiles is widely underestimated across space and phylogeny. *PLoS Biology*, 20(5), e3001544.
- Cardillo, M. (2003). Biological determinants of extinction risk: Why are smaller species less vulnerable? *Animal Conservation*, 6, 63–69.

- Cardillo, M. (2021). Clarifying the relationship between body size and extinction risk in amphibians by complete mapping of model space. *Proceedings of the Royal Society B: Biological Sciences*, 288(1944), 20203011.
- Cardillo, M., Mace, G. M., Gittleman, J. L., Jones, K. E., Bielby, J., & Purvis, A. (2008). The predictability of extinction: Biological and external correlates of decline in mammals. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, 275(1641), 1441–1448.
- Cardillo, M., Purvis, A., Sechrest, W., Gittleman, J. L., Bielby, J., & Mace, G. M. (2004). Human population density and extinction risk in the world's carnivores. *PLoS Biology*, 2(7), 0909–0914.
- Carilo Filho, L. M., de Carvalho, B. T., Azevedo, B. K. A., Gutiérrez-Pesquera, L. M., Mira-Mendes, C. V., Solé, M., & Orrico, V. G. D. (2021). Natural history predicts patterns of thermal vulnerability in amphibians from the Atlantic Rainforest of Brazil. *Ecology and Evolution*, 11(23), 16462–16472.
- Cazalis, V., Di Marco, M., Butchart, S. H. M., Akçakaya, H. R., González-Suárez, M., Meyer, C., Clausnitzer, V., Böhm, M., Zizka, A., Cardoso, P., Schipper, A. M., Bachman, S. P., Young, B. E., Hoffmann, M., Benítez-López, A., Lucas, P. M., Pettorelli, N., Patoine, G., Pacifici, M., ... Santini, L. (2022). Bridging the research-implementation gap in IUCN Red List assessments. *Trends in Ecology & Evolution*, 37(4), 359–370.
- Cazalis, V., Santini, L., Lucas, P. M., González-Suárez, M., Hoffmann, M., Benítez-López, A., Pacifici, M., Schipper, A. M., Böhm, M., Zizka, A., Clausnitzer, V., Meyer, C., Jung, M., Butchart, S. H. M., Cardoso, P., Mancini, G., Akçakaya, H. R., Young, B. E., Patoine, G., & Di Marco, M. (2023). Prioritizing the reassessment of data-deficient species on the IUCN Red List. *Conservation Biology*, 37(6), e14139.
- Ceballos, G., Ehrlich, P. R., & Dirzo, R. (2017). Biological annihilation via the ongoing sixth mass extinction signaled by vertebrate population losses and declines. *Proceedings of the National Academy of Sciences of the United States of America*, 114(30), E6089–E6096.
- Chen, C., Chen, C., & Wang, Y. (2019). Ecological correlates of extinction risk in Chinese amphibians. *Diversity and Distributions*, 25(10), 1586–1598.
- Chen, T., & Guestrin, C. (2016). *XGBoost: A scalable tree boosting system*. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, 13–17 August 2016; pp. 785–794. <https://doi.org/10.1145/2939672.2939785>
- Chichorro, F., Juslén, A., & Cardoso, P. (2019). A review of the relation between species traits and extinction risk. *Biological Conservation*, 237, 220–229.
- Christensen, R. H. B. (2020). *Ordinal package version 2020.8-22*. <http://www.cran.r-project.org/package=ordinal/>
- Cohen, J. M., Civitello, D. J., Venesky, M. D., McMahon, T. A., & Rohr, J. R. (2019). An interaction between climate change and infectious disease drove widespread amphibian declines. *Global Change Biology*, 25(3), 927–937.
- Cooper, N., Bielby, J., Thomas, G. H., & Purvis, A. (2008). Macroecology and extinction risk correlates of frogs. *Global Ecology and Biogeography*, 17(2), 211–221.
- Copernicus Climate Change Service (C3S) Climate Data Store (CDS). (2019). *Land cover classification gridded maps from 1992 to present derived from satellite observations version 2.0 and 2.1*. <https://doi.org/10.24381/cds.006f2c9a>
- Di Marco, M. (2022). Reptile research shows new avenues and old challenges for extinction risk modelling. *PLoS Biology*, 20(7), e3001719.
- Di Marco, M., Buchanan, G. M., Szantoi, Z., Holmgren, M., Grottolo Marasini, G., Gross, D., Tranquilli, S., Boitani, L., & Rondinini, C. (2014). Drivers of extinction risk in African mammals: The interplay of distribution state, human pressure, conservation response and species biology. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1643), 20130198.
- Diaz, G. I., Fokoue-Nkoutche, A., Nannicini, G., & Samulowitz, H. (2017). An effective algorithm for hyperparameter optimization of neural networks. *IBM Journal of Research and Development*, 61(4/5), 9:1–9:11.
- Diniz-Filho, J. A. F., de Sant'Ana, C. E. R., & Bini, L. M. (1998). An eigenvector method for estimating phylogenetic data. *Evolution*, 52(5), 1247–1262.
- Dormann, C. F. (2013). *Parametric statistics. Distributions, maximum likelihood and GLM in R*. Springer.
- Duellman, W. E., & Trueb, L. (1994). *Biology of amphibians*. Johns Hopkins University Press.
- Environmental Systems Research Institute (ESRI). (2008). *ArcMap 9.3*. Author.
- Evans, J., & Murphy, M. (2018). *rUtilities*. R package version 2.1-3. <https://cran.r-project.org/package=rUtilities>
- Ficetola, G. F., Rondinini, C., Bonardi, A., Baisero, D., & Padoa-Schioppa, E. (2015). Habitat availability for amphibians and extinction threat: A global analysis. *Diversity and Distributions*, 21(3), 302–311.
- Foden, W. B., & Young, B. E. (2016). *IUCN SSC guidelines for assessing species' vulnerability to climate change. Version 1.0*. Occasional Paper of the IUCN Species Survival Commission No. 59. IUCN Species Survival Commission.
- Fontana, R. B., Furtado, R., Zanella, N., Debastiani, V. J., & Hartz, S. M. (2021). Linking ecological traits to extinction risk: Analysis of a Neotropical anuran database. *Biological Conservation*, 264, 109390.
- Frost, D. R. (2021). *Amphibian species of the world: An online reference*. <https://amphibiansoftheworld.amnh.org/index.php>
- Gedeon, T. D. (1997). Data mining of inputs: Analysing magnitude and functional measures. *International Journal of Neural Systems*, 8(2), 209–218.
- Genuer, R., Poggi, J.-M., & Tuleau-Malot, C. (2010). Variable selection using random forests. *Pattern Recognition Letters*, 31(14), 2225–2236.
- González-del-Pliego, P., Freckleton, R. P., Edwards, D. P., Koo, M. S., Scheffers, B. R., Pyron, R. A., & Jetz, W. (2019). Phylogenetic and trait-based prediction of extinction risk for data-deficient amphibians. *Current Biology*, 29(9), 1557.e3–1563.e3.
- González-Suárez, M., Lucas, P. M., & Revilla, E. (2012). Biases in comparative analyses of extinction risk: Mind the gap. *The Journal of Animal Ecology*, 81(6), 1211–1222.
- Green, D. M. (2017). Amphibian breeding phenology trends under climate change: Predicting the past to forecast the future. *Global Change Biology*, 23(2), 646–656.
- Harfoot, M. B. J., Johnston, A., Balmford, A., Burgess, N. D., Butchart, S. H. M., Dias, M. P., Hazin, C., Hilton-Taylor, C., Hoffmann, M., Isaac, N. J. B., Iversen, L. L., Outhwaite, C. L., Visconti, P., & Geldmann, J. (2021). Using the IUCN Red List to map threats to terrestrial vertebrates at global scale. *Nature Ecology & Evolution*, 5(11), 1510–1519.
- Henry, E. G., Santini, L., Butchart, S. H. M., González-Suárez, M., Lucas, P. M., Benítez-López, A., Mancini, G., Jung, M., Cardoso, P., Zizka, A., Meyer, C., Akçakaya, H. R., Berryman, A. J., Cazalis, V., & Di Marco, M. (2024). Modelling the probability of meeting IUCN Red List criteria to support reassessments. *Global Change Biology*, 30(1), e17119.
- Herve, M. (2023). *RV AideMemoire: Testing and plotting procedures for biostatistics*. R package version 0.9-83-7. <https://CRAN.R-project.org/package=RV AideMemoire>
- Hof, C., Araújo, M. B., Jetz, W., & Rahbek, C. (2011). Additive threats from pathogens, climate and land-use change for global amphibian diversity. *Nature*, 480(7378), 516–519.
- Hong, H., Xiaoling, G., & Hua, Y. (2016). *Variable selection using Mean Decrease Accuracy and Mean Decrease Gini based on random forest*. Proceedings of the 2016 7th IEEE International Conference on Software Engineering and Service Science (ICSESS), Beijing, China, 26–28 August 2016; pp. 219–224.
- International Union for Conservation of Nature (IUCN). (2012). *IUCN Red List Categories and Criteria. Version 3.1* (2nd ed.). Author.
- International Union for Conservation of Nature (IUCN). (2018). *Mapping standards and data quality for the IUCN Red List Categories and Criteria. Version 1.16*. Author.
- International Union for Conservation of Nature (IUCN). (2021). *IUCN Red List of Threatened Species. Version 2021–2*. www.iucnredlist.org
- International Union for Conservation of Nature (IUCN). (2022). *The IUCN Red List of Threatened Species. Version 2021–3*. <https://www.iucnredlist.org>
- International Union for Conservation of Nature (IUCN). (2023). *IUCN Red List of Threatened Species. Version 2023-1*. Author.
- Jetz, W., & Pyron, R. A. (2018). The interplay of past diversification and evolutionary isolation with present imperilment across the amphibian tree of life. *Nature Ecology & Evolution*, 2(5), 850–858.
- Karger, D. N., Conrad, O., Böhrner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., Zimmermann, N. E., Linder, H. P., & Kessler, M. (2017). Climatologies at high resolution for the earth's land surface areas. *Scientific Data*, 4(1), 170122.
- Karger, D. N., Conrad, O., Böhrner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., Zimmermann, N. E., Linder, H. P., & Kessler, M. (2018). *Data from: Climatologies at high resolution for the earth's land surface areas*. Dryad. <https://doi.org/10.5061/dryad.kd1d4>
- LeDell, E., Gill, N., Aiello, S., Fu, A., Candel, A., Click, C., Kraljevic, T., Nykodym, T., Aboyoun, P., Kurka, M., & Malohlava, M. (2022). *h2o: R inter-*

- face for the 'H2O' scalable machine learning platform. R package version 3.38.0.1. <https://github.com/h2oai/h2o-3>
- Lertzman-Lepofsky, G. F., Kissel, A. M., Sinervo, B., & Palen, W. J. (2020). Water loss and temperature interact to compound amphibian vulnerability to climate change. *Global Change Biology*, *26*(9), 4868–4879.
- Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R News*, *2*(3), 18–22.
- Liedtke, H. C. (2019). AmphiNom: An amphibian systematics tool. *Systematics and Biodiversity*, *17*(1), 1–6.
- Loarie, S. R., Duffy, P. B., Hamilton, H., Asner, G. P., Field, C. B., & Ackerly, D. D. (2009). The velocity of climate change. *Nature*, *462*(7276), 1052–1055.
- Lu, M., & Ishwaran, H. (2021). Discussion on “Nonparametric variable importance assessment using machine learning techniques” by Brian D. Williamson, Peter B. Gilbert, Marco Carone, and Noah Simon. *Biometrics*, *77*(1), 23–27.
- Lucas, P. M., González-Suárez, M., & Revilla, E. (2019). Range area matters, and so does spatial configuration: Predicting conservation status in vertebrates. *Ecography*, *42*(6), 1103–1114. <https://doi.org/10.1111/ecog.03865>
- Lucas, P. M., González-Suárez, M., & Revilla, E. (2016). Toward multifactorial null models of range contraction in terrestrial vertebrates. *Ecography*, *39*(11), 1100–1108.
- Luedtke, J. A., Chanson, J., Neam, K., Hobin, L., Maciel, A. O., Catenazzi, A., Borzée, A., Hamidy, A., Aowphol, A., Jean, A., Sosa-Bartuano, Á., Fong G, A., de Silva, A., Fouquet, A., Angulo, A., Kidov, A. A., Muñoz Saravia, A., Diesmos, A. C., Tominaga, A., ... Stuart, S. N. (2023). Ongoing declines for the world's amphibians in the face of emerging threats. *Nature*, *622*(7982), 308–314.
- Lukacs, P. M., Burnham, K. P., & Anderson, D. R. (2010). Model selection bias and Freedman's paradox. *Annals of the Institute of Statistical Mathematics*, *62*(1), 117–125.
- Mancini, G., Benítez-López, A., Di Marco, M., Pacifici, M., Rondinini, C., & Santini, L. (2023). Synergistic effects of habitat fragmentation and hunting on the extinction risk of neotropical primates. *Biodiversity and Conservation*, *32*(8), 2655–2669.
- Mancini, G., Santini, L., Cazalis, V., Akçakaya, H. R., Lucas, P. M., Brooks, T. M., Foden, W., & Di Marco, M. (2024). A standard approach for including climate change responses in IUCN Red List assessments. *Conservation Biology*, *38*(3), e14227. <https://doi.org/10.1111/cobi.14227>
- Mantyka-pringle, C. S., Martin, T. G., & Rhodes, J. R. (2012). Interactions between climate and habitat loss effects on biodiversity: A systematic review and meta-analysis. *Global Change Biology*, *18*(4), 1239–1252.
- McKinney, M. L. (2002). Urbanization, biodiversity, and conservation. *Bioscience*, *52*(10), 883–890.
- Miller, D. A. W., Grant, E. H. C., Muths, E., Amburgey, S. M., Adams, M. J., Joseph, M. B., Waddle, J. H., Johnson, P. T. J., Ryan, M. E., Schmidt, B. R., Calhoun, D. L., Davis, C. L., Fisher, R. N., Green, D. M., Hossack, B. R., Rittenhouse, T. A. G., Walls, S. C., Bailey, L. L., Cruickshank, S. S., ... Sigafus, B. H. (2018). Quantifying climate sensitivity and climate-driven change in North American amphibian communities. *Nature Communications*, *9*(1), 3926.
- Murphy, M. A., Evans, J. S., & Storfer, A. (2010). Quantifying Bufo boreas connectivity in Yellowstone National Park with landscape genetics. *Ecology*, *91*(1), 252–261.
- Nakagawa, S., & Freckleton, R. P. (2008). Missing inaction: The dangers of ignoring missing data. *Trends in Ecology & Evolution*, *23*(11), 592–596.
- Nakagawa, S., & Freckleton, R. P. (2011). Model averaging, missing data and multiple imputation: A case study for behavioural ecology. *Behavioral Ecology and Sociobiology*, *65*(1), 103–116.
- National Aeronautics and Space Administration (NASA). (2018). *Gridded Population of the World, Version 4 (GPWv4): Population Density, Revision 11*. <https://sedac.ciesin.columbia.edu/data/set/gpw-v4-population-count-rev11/data-download>
- Newbold, T., Hudson, L. N., Hill, S. L. L., Contu, S., Lysenko, I., Senior, R. A., Borger, L., Bennett, D. J., Choimes, A., Collen, B., Day, J., De Palma, A., Diaz, S., Echeverria-Londono, S., Edgar, M. J., Feldman, A., Garon, M., Harrison, M. L. K., Alhusseini, T., & Purvis, A. (2015). Global effects of land use on local terrestrial biodiversity. *Nature*, *520*, 45–50.
- Oliveira, B. F., São-Pedro, V. A., Santos-Barrera, G., Penone, C., & Costa, G. C. (2017). AmphiBIO, a global database for amphibian ecological traits. *Scientific Data*, *4*(1), 170123.
- Olson, D. M., Dinerstein, E., Wikramanayake, E. D., Burgess, N. D., Powell, G. V. N., Underwood, E. C., D'Amico, J. A., Itoua, I., Strand, H. E., Morrison, J. C., Loucks, C. J., Allnutt, T. F., Ricketts, T. H., Kura, Y., Lamoreux, J. F., Wettengel, W. W., Hedao, P., & Kassem, K. R. (2001). Terrestrial ecoregions of the world: A new map of life on earth. *Bioscience*, *51*(11), 933–938.
- Paradis, E., & Schliep, K. (2019). ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, *35*, 526–528.
- Pelletier, T. A., Carstens, B. C., Tank, D. C., Sullivan, J., & Espindola, A. (2018). Predicting plant conservation priorities on a global scale. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(51), 13027–13032.
- Penone, C., Davidson, A. D., Shoemaker, K. T., Di Marco, M., Rondinini, C., Brooks, T. M., Young, B. E., Graham, C. H., & Costa, G. C. (2014). Imputation of missing data in life-history trait datasets: Which approach performs the best? *Methods in Ecology and Evolution*, *5*(9), 961–970.
- Pereira, H. M., Leadley, P. W., Proença, V., Alkemade, R., Scharlemann, J. P. W., Fernandez-Manjarrés, J. F., Araújo, M. B., Balvanera, P., Biggs, R., Cheung, W. W. L., Chini, L., Cooper, H. D., Gilman, E. L., Guénette, S., Hurtt, G. C., Huntington, H. P., Mace, G. M., Oberdorff, T., Revenga, C., ... Walpole, M. (2010). Scenarios for global biodiversity in the 21st century. *Science*, *330*, 1496–1501.
- Pincheira-Donoso, D., Harvey, L. P., Cotter, S. C., Stark, G., Meiri, S., & Hodgson, D. J. (2021). The global macroecology of brood size in amphibians reveals a predisposition of low-fecundity species to extinction. *Global Ecology and Biogeography*, *30*(6), 1299–1310.
- Pincheira-Donoso, D., Harvey, L. P., Grattarola, F., Jara, M., Cotter, S. C., Tregenza, T., & Hodgson, D. J. (2021). The multiple origins of sexual size dimorphism in global amphibians. *Global Ecology and Biogeography*, *30*(2), 443–458.
- Pincheira-Donoso, D., Meiri, S., Jara, M., Olalla-Tárraga, M. Á., & Hodgson, D. J. (2019). Global patterns of body size evolution are driven by precipitation in legless amphibians. *Ecography*, *42*(10), 1682–1690.
- Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & Team, R. C. (2021). *nlme: Linear and nonlinear mixed effects models*. R package version 3.1-153. <https://CRAN.R-project.org/package=nlme>
- Revell, L. J. (2010). Phylogenetic signal and linear regression on species data. *Methods in Ecology and Evolution*, *1*(4), 319–329.
- Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillerá-Arroita, G., Hauenstein, S., Lahoz-Monfort, J. J., Schröder, B., Thuiller, W., Warton, D. I., Wintle, B. A., Hartig, F., & Dormann, C. F. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography*, *40*(8), 913–929.
- Rohlf, F. J. (2001). Comparative methods for the analysis of continuous variables: Geometric interpretations. *Evolution*, *55*(11), 2143–2160.
- Rondinini, C., Di Marco, M., Visconti, P., Butchart, S. H. M., & Boitani, L. (2014). Update or outdated: Long-term viability of the IUCN Red List. *Conservation Letters*, *7*(2), 126–130.
- Santini, L., Benítez-López, A., Ficetola, G. F., & Huijbregts, M. A. J. (2018). Length–mass allometries in amphibians. *Integrative Zoology*, *13*(1), 36–45.
- Santini, L., Butchart, S. H. M., Rondinini, C., Benítez-López, A., Hilbers, J. P., Schipper, A. M., Cengic, M., Tobias, J. A., & Huijbregts, M. A. J. (2019). Applying habitat and population-density models to land-cover time series to inform IUCN Red List assessments. *Conservation Biology*, *33*(5), 1084–1093.
- Santos, T. (2018). *PVR: Phylogenetic eigenvectors regression and phylogenetic signal-representation curve*. R package version 0.3. <https://cran.r-project.org/web/packages/PVR/index.html>
- Scheele, B. C., Pasmans, F., Skerratt, L. F., Berger, L., Martel, A., Beukema, W., Acevedo, A. A., Burrows, P. A., Carvalho, T., Catenazzi, A., De la Riva, I., Fisher, M. C., Flechas, S. V., Foster, C. N., Frías-Álvarez, P., Garner, T. W. J., Gratwicke, B., Guayasamin, J. M., Hirschfeld, M., ... Canessa, S. (2019). Amphibian fungal panzootic causes catastrophic and ongoing loss of biodiversity. *Science*, *363*(6434), 1459–1463.

- Silva, S. V., Andermann, T., Zizka, A., Kozłowski, G., & Silvestro, D. (2022). Global estimation and mapping of the conservation status of tree species using artificial intelligence. *Frontiers in Plant Science*, *13*, 839792.
- Silvano, D. L., & Segalla, M. V. (2005). Conservation of Brazilian amphibians. *Conservation Biology*, *19*(3), 653–658.
- Sodhi, N. S., Bickford, D., Diesmos, A. C., Lee, T. M., Koh, L. P., Brook, B. W., Sekercioglu, C. H., & Bradshaw, C. J. A. (2008). Measuring the meltdown: Drivers of global amphibian extinction and decline. *PLoS ONE*, *3*(2), e1636.
- Sonn, J. M., Utz, R. M., & Richards-Zawacki, C. L. (2019). Effects of latitudinal, seasonal, and daily temperature variations on chytrid fungal infections in a North American frog. *Ecosphere*, *10*(11), e02892.
- Soto-Saravia, R. A., Garrido-Cayul, C. M., Avaria-Llatureo, J., Benítez-Mora, A., Hernández, C. E., & González-Suárez, M. (2021). Threatened neotropical birds are big, ecologically specialized, and found in less humanized refuge areas. *Avian Conservation and Ecology*, *16*(2), 18. <https://doi.org/10.5751/ACE-01948-160218>
- Stewart, K., Carmona, C. P., Clements, C., Venditti, C., Tobias, J. A., & González-Suárez, M. (2023). Functional diversity metrics can perform well with highly incomplete data sets. *Methods in Ecology and Evolution*, *14*(11), 2856–2872.
- Stuart, S. N., Chanson, J. S., Cox, N. A., Young, B. E., Rodrigues, A. S. L., Fischman, D. L., & Waller, R. W. (2004). Status and trends of amphibian declines and extinctions worldwide. *Science*, *306*(5702), 1783–1786.
- Stuart, S. N., Hoffmann, M., Chanson, J. S., Cox, N. A., Berridge, R. J., Ramani, P., & Young, B. E. (2008). *Threatened amphibians of the world*. Lynx Edicions; IUCN; and Conservation International.
- Tange, R. I., Rasmussen, M. A., Taira, E., & Bro, R. (2017). Benchmarking support vector regression against partial least squares regression and artificial neural network: Effect of sample size on model performance. *Journal of Near Infrared Spectroscopy*, *25*(6), 381–390.
- Tapley, B., Michaels, C. J., Gumbs, R., Böhm, M., Luedtke, J., Pearce-Kelly, P., & Rowley, J. J. L. (2018). The disparity between species description and conservation assessment: A case study in taxa with high rates of species discovery. *Biological Conservation*, *220*, 209–214.
- Thuiller, W., Guéguen, M., Renaud, J., Karger, D. N., & Zimmermann, N. E. (2019). Uncertainty in ensembles of global biodiversity scenarios. *Nature Communications*, *10*(1), 1446.
- Torres, J. F., Gutiérrez-Avilés, D., Troncoso, A., & Martínez-Álvarez, F. (2019). Random hyper-parameter search-based deep neural network for power consumption forecasting. In I. Rojas, H. Joya, & A. Catala (Eds.), *Advances in computational intelligence* (Vol. 11506, pp. 259–269). Springer.
- Urban, M. C. (2015). Accelerating extinction risk from climate change. *Science*, *384*, 571–573.
- Urban, M. C., Bocedi, G., Hendry, A. P., Mihoub, J. B., Peer, G., Singer, A., Bridle, J. R., Crozier, L. G., Meester, D. L., Godsoe, W., Gonzalez, A., Hellmann, J. J., Holt, R. D., Huth, A., Johst, K., Krug, C. B., Leadley, P. W., Palmer, S. C. F., Pantel, J. H., ... Travis, J. M. J. (2016). Improving the forecast for biodiversity under climate change. *Science*, *353*(6304), 00.
- Vabalas, A., Gowen, E., Poliakov, E., & Casson, A. J. (2019). Machine learning algorithm validation with a limited sample size. *PLoS ONE*, *14*(11), e0224365.
- Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with S* (4th ed.). Springer.
- Wake, D. B., & Vredenburg, V. T. (2008). Are we in the midst of the sixth mass extinction? A view from the world of amphibians. *Proceedings of the National Academy of Sciences of the United States of America*, *105*, (supplement_1), 11466–11473.
- Weiss, D. J., Nelson, A., Gibson, H. S., Temperley, W., Peedell, S., Lieber, A., Hancher, M., Poyart, E., Belchior, S., Fullman, N., Mappin, B., Dalrymple, U., Rozier, J., Lucas, T. C. D., Howes, R. E., Tusting, L. S., Kang, S. Y., Cameron, E., Bisanzio, D., ... Gething, P. W. (2018). A global map of travel time to cities to assess inequalities in accessibility in 2015. *Nature*, *553*(7688), 333–336.
- Wieringa, J. G. (2022). Comparing predictions of IUCN Red List categories from machine learning and other methods for bats. *Journal of Mammalogy*, *103*(3), 528–539.
- Williamson, B. D., Gilbert, P. B., Carone, M., & Simon, N. (2021). Nonparametric variable importance assessment using machine learning techniques. *Biometrics*, *77*(1), 9–22.
- Yackulic, C. B., Sanderson, E. W., & Uriarte, M. (2011). Anthropogenic and environmental drivers of modern range loss in large mammals. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(10), 4024–4029.
- Zizka, A., Andermann, T., & Silvestro, D. (2022). IUCNN—Deep learning approaches to approximate species' extinction risk. *Diversity and Distributions*, *28*(2), 227–241.
- Zizka, A., Silvestro, D., Vitt, P., & Knight, T. M. (2021). Automated conservation assessment of the orchid family with deep learning. *Conservation Biology*, *35*(3), 897–908.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Lucas, P. M., Di Marco, M., Cazalis, V., Luedtke, J., Neam, K., Brown, M. H., Langhammer, P. F., Mancini, G., & Santini, L. (2024). Using comparative extinction risk analysis to prioritize the IUCN Red List reassessments of amphibians. *Conservation Biology*, e14316. <https://doi.org/10.1111/cobi.14316>