

A peer-reviewed version of this preprint was published in PeerJ on 16 December 2014.

[View the peer-reviewed version](https://doi.org/10.7717/peerj.703) (peerj.com/articles/703), which is the preferred citable publication unless you specifically need to cite this preprint.

Quintero E, Thessen AE, Arias-Caballero P, Ayala-Orozco B. 2014. A statistical assessment of population trends for data deficient Mexican amphibians. PeerJ 2:e703 <https://doi.org/10.7717/peerj.703>

A statistical assessment of population trends for data deficient Mexican amphibians

Esther Quintero, Anne E. Thessen, Paulina Arias-Caballero, Bárbara Ayala-Orozco

Background: Mexico is the fourth richest country in amphibians and the second country with the highest quantity of threatened amphibian species, and this number could be higher as many species are too poorly known to be accurately assigned to a risk category. The absence of a risk status or an unknown population trend can slow or halt conservation action, so it is vital to develop tools that in the absence of specific demographic data can assess a species' risk of extinction, population trend, and to better understand which variables increase their vulnerability. Recent studies have demonstrated that the risk of species decline depends on extrinsic and intrinsic trait, thus including both of them for assessing extinction might render more accurate assessment of threat. **Methods:** In this study harvested data from the Encyclopedia of Life (EOL) and the published literature for Mexican amphibians and used these data to assess the population trend of some of the Mexican species that have been assigned to the Data Deficient category of the IUCN using Random Forests, a Machine Learning method that gives a prediction of complex processes and identifies the most important variables that account for the predictions. **Results:** Our results show that most data deficient Mexican amphibians have decreasing population trends. We found that Random Forests is a solid and accurate way to identify species with decreasing population trends when no demographic data is available. Moreover, we point the most important variables that make species more vulnerable for extinction. This exercise is a very valuable first step in assigning conservation priorities for poorly known species.

A Statistical Assessment of Risk Status for Data Deficient Mexican Amphibians

Esther Quintero, Subcoordinación de Especies Prioritarias, Dirección General de Análisis y Prioridades, Comisión Nacional para el Conocimiento y Uso de la Biodiversidad, Mexico D.F., Mexico

Anne E. Thessen, The Data Detektiv, Waltham, MA, USA

Paulina Arias-Caballero, Subcoordinación de Especies Prioritarias, Dirección General de Análisis y Prioridades, Comisión Nacional para el Conocimiento y Uso de la Biodiversidad, México D.F., México

Bárbara Ayala, Subcoordinación de Especies Prioritarias, Dirección General de Análisis y Prioridades, Comisión Nacional para el Conocimiento y Uso de la Biodiversidad, México D.F., México

Corresponding author: Esther Quintero, Comisión Nacional para el Conocimiento y Uso de la Biodiversidad, Liga Periférico - Insurgentes Sur, Núm. 4903, Col. Parques del Pedregal, Tlalpan, 14010, México, D.F. Tel. + 52 55 5004 5000 email esther.quintero@conabio.gob.mx

Introduction

Amphibians are keystone to conservation and excellent bioindicators. Their extinction will trigger cascading effects on the ecosystem (Gardner 2001; Wake 1991; Wyman 1990). Among all terrestrial vertebrates, amphibians are the most threatened group with more “rapidly declining species” (Stuart et al. 2004). Mexico is the fourth richest country in amphibian species (Ochoa-Ochoa & Flores-Villela 2006) with around 375 documented species, although this number could be greatly underestimated (Flores-Villela & Canseco-Márquez 2004). At the same time, Mexico is the second country with the highest quantity of threatened amphibian species, 211 according to the IUCN (IUCN 2014), and this number could be even higher, as many species are too poorly known to be accurately assigned to a risk category. In Mexico, 38 (10%) amphibian species are currently listed as DD (Data Deficient) by the IUCN (2014) because specific data about a species are missing (i.e. geographic distribution, threats, population status, etc.). The absence of an IUCN status can slow or halt conservation action, which for some species could have irreversible consequences. Therefore, it is of vital importance to develop tools that allow assessing species’ risk of extinction in the absence of specific demographic data, as well as to better understand which variables increase vulnerability to extinction.

The first Global Amphibian Assessment (Stuart et al. 2004), found that amphibian declines are not random, but associated to ecological traits (i.e. stream associated species), geographic distribution (i.e. montane areas in the Neotropics, Australia and New Zealand), and specific taxonomic groups (i.e. Leptodactylidae, Bufonidae, Ambystomatidae, Hylidae, and Ranidae). Moreover, they divided the causes of decline in three groups: over-exploitation, defined as those declining due to heavy extraction (concentrated in species in East and Southeastern Asia);

reduced habitat, defined as those that were suffering from extreme habitat loss (concentrated in Southeast Asia, West Africa, and the Caribbean); and enigmatic declines, those that are declining even though suitable area remains (restricted mostly to South America, Mesoamerica, Puerto Rico and Australia). Enigmatic declines were found to be positively associated with streams at high elevation in the tropics, and chytridiomycosis emerged as the most likely culprit.

Chytridiomycosis is a fungal disease caused by *Batrachochytrium dendrobatidis*, and has been related to the decline of at least 43 species of amphibians in Latin America (Lips et al. 2006). In México, there is an association between higher elevations (from 939 to 3200 m) and the prevalence of the infection. It has been found in pristine and disturbed areas alike, but does not seem very common throughout tropical rain forests or lowland deserts (Frías-Alvarez et al. 2008). The reason for this marked preference for high areas with temperate climates may be that the optimal range of growth for this fungus is between 17-25 C (J.S. et al. 2004; Longcore et al. 1999). A geographical survey for the presence of chytridiomycosis in Mexico found the presence of the fungus in sites that have reported “enigmatic declines” in amphibian populations (Frías-Alvarez et al. 2008). The finding by these authors suggests that chytridiomycosis is a likely cause behind many of these enigmatic declines.

Recent studies (Murray et al. 2011; Tingley et al. 2013) have demonstrated that the risk of species decline depends on the specific threats they face, such as habitat loss, presence of invasive species, and pathogens (extrinsic traits), and the species’ own biological ability to cope with these threats, such as clutch size and body size (intrinsic traits). Thus, including intrinsic traits along with extrinsic threats for assessing extinction might render more accurate assessment

of threat (Tingley et al. 2013), and thus improve allocation of resources (Cardillo & Meijaard 2012).

One of the most recognized efforts to assign risk categories to species is that of the International Union for Conservation of Nature (IUCN), which recognizes seven different extinction risk categories for evaluated species: two of them are for species that are already extinct (Extinct and Extinct in the wild), three are those considered as threatened categories (Critically Endangered, Endangered, and Vulnerable), two are for those species that are not yet threatened (Near Threatened and Least Concern), whereas the last one is for those with not enough information to be evaluated (Data Deficient). The IUCN also lists species that have not yet been evaluated (Not Evaluated).

IUCN's criteria for assigning a threat category to species are "quantitative in nature", but the data quality and the uncertainty attached to any evaluation vary. Estimations, inferences, projections and suspected facts based on related data are acceptable, as long as they can be supported and specified in the documentation. The Data Deficient category (DD) is assigned to those species in which the available data is not enough to determine a threat category, not even indirectly, for example through the status of their habitat or other causal factors (IUCN 2012). Only approximately 75,000 out of the 2 million described species are evaluated by the IUCN and one sixth of them are Data Deficient (<http://www.iucnredlist.org/about/summary-statistics>), with 25% of all amphibians classified as such (Stuart et al. 2004). By lacking a threat status, Data Deficient species are not taken into account for conservation programs, potentially placing them

at a higher risk of extinction. Thus, it is clear that a more automated method of evaluating risk that can use a wider variety of available data and still give accurate results is needed.

In this study we aim to harvest data from the Encyclopedia of Life (EOL) and the published literature for Mexican amphibians and use these data to assess the population trend of some of the Mexican species that have been assigned to the DD category of the IUCN using Random Forests, a Machine Learning method algorithm that gives a prediction of complex processes and identifies the most important variables that account for the predictions (Breiman 2001; Cutler et al. 2007; Murray et al. 2011). A recent assessment of DD mammals using and comparing multiple Machine Learning tools found that Random forests perform very well for this type of predictions (Bland et al. 2014). Focusing on such a vulnerable and ecologically important group as the amphibians not only potentiates our conservation efforts, but also has the potential to improve assessment of other ecologically important groups for which we might lack demographic data.

Methods

Selecting traits for the analysis

In order to assess the population traits of those species listed as Data Deficient, we selected previously identified intrinsic traits that can predispose species to a greater degree of vulnerability, as well as a series of extrinsic traits that have been associated to amphibian decline (Stuart et al. 2004).

PeerJ PrePrints

The extrinsic traits in our analysis were habitat use, habitat loss/degradation (one of the biggest concerns for biodiversity (Assessment 2005; Brooks et al. 2002; Frías-Alvarez et al. 2008; Groombridge 1992; Parra-Olea et al. 1999; Wyman 1990), presence of introduced species, presence of pollution, climatic fluctuations, harvest for pet trade, desiccation of bodies of water, presence of chytridiomycosis, and presence of other diseases that may decimate populations. The intrinsic traits selected for our analysis were snout-vent length, ova size, clutch size, and development type as a way to understand their life history and ecological preferences (Murray et al. 2011).

Automated Data Harvesting from Encyclopedia of Life

Starting with a list of scientific names of amphibians in Mexico (Table S1), relevant data and text were harvested from EOL using TraitBank and the EOL API respectively (Parr et al. 2014). The code written for this project can be found at GitHub (<https://github.com/diatomsRcool/MexicanAmphibians>). Data from EOL TraitBank was retrieved by searching for taxon and measurement and downloaded as a .csv file. The API was used to find the EOL identifier corresponding to each amphibian species. This identifier, in combination with EOL chapters and keywords was used to filter and harvest all relevant text data objects (Table 1). This process identified a subset of text data objects for manual data extraction. Data from TraitBank and the text data objects were added to a master spreadsheet for analysis (Table S1). Data gathered for this study that was not already in TraitBank, was placed in a Darwin Core Archive and uploaded into EOL TraitBank.

Other sources of literature

Data that were not available from EOL were obtained from the literature, and cited in Table S2. Data for threats (habitat loss/degradation, introduced species, pollution, chytridiomycosis, climatic fluctuations, pet trade/harvest, desiccation of habitat, and other diseases), as well as for population trend (decreasing, increasing, stable, unknown) were obtained from the IUCN Red List (IUCN 2014).

Data preparation

A table was prepared with 302 rows and 16 columns (Table S1). Each row represented a species and each column represented a trait of that species. Examination of this master table revealed two traits (ova size and clutch size; Table 2) and four species (*Bolitoglossa chinanteca*, *Dermophis oaxacae*, *Eleutherodactylus marnockii*, and *Eleutherodactylus verruculatus*) to be particularly data deficient (defined as 10 or more missing traits). An additional species was identified as being introduced (*Eleutherodactylus planirostris*). These traits and species were removed from the data set.

All traits were coded into numeric categories (Table 3). Snout to Vent length classifications followed (García & Ceballos 1994). In habitat use, we distinguished permanent water associated from stream associated. Threats were treated as present (1) or absent (0). Chytridiomycosis was recorded as present in cases where it was reported as suspected. Missing data were represented by a blank cell. From this table, we prepared a csv file for missing data imputation in R. The scientific name, IUCN status, and population trend were removed before imputation.

Missing data imputation

We used the mice package (van Buuren & Groothuis-Oudshoorn 2011) in R to impute missing values (<http://cran.r-project.org/web/packages/mice/mice.pdf>). This was necessary because the randomForest function did not tolerate missing values. All data were imported into R as factors. The Snout-Vent Length, Habitat Use, and Development Type were imputed as polytomous logistic regression (polyreg). The other traits were imputed using logistic regression (logreg). Missing Population Trend data were not imputed. The data before imputation can be found in Table S1. A summary of missing data can be found in Table 2 and Table S1. Ten imputations were performed for each missing value. The final imputed value was the mode of the 10 imputations. The data after imputation can be found in Table S3. The data set that includes the imputed data was used for predicting the population trend for those species that were Data Deficient and Not Evaluated according to the IUCN evaluation.

Predicting Population Trends

We used the randomForest package in R (Liaw & Wiener 2002) to make predictions about the population trends for each species of amphibian (<http://cran.rproject.org/web/packages/randomForest/randomForest.pdf>). This was a three-step process: the first step was using training data, i.e., traits for those species with known population trends, to generate a random forest object. The second step was testing the random forest object on a separate data set of species with known population trends. The third step was using the random forest object to make predictions about population trends for those species listed as Data Deficient and Not Evaluated by the IUCN (IUCN 2014).

The data set of species with known population trends was divided in two to make a training set and a test set. The training data (including the imputed data) was read into R and given to the randomForest (Liaw & Wiener 2002) function, which provided a random forest object as a result. To test the efficacy of the random forest for prediction, we removed the Population Trend data from the test set and made a prediction of population trend for comparison to the observed population trend (Table 4).

We added additional species, as needed, to the data set including all of the species with an unknown population trend (unknown data set) to balance the presence of categories for each trait, a requirement for making predictions. The unknown data set was read into R and given to the randomForest and predict functions in the randomForest package (Liaw & Wiener 2002). To ensure unbiased variable selection (Strobl et al. 2007), we used the cforest (Hothorn et al. 2006a; Strobl et al. 2008; Strobl et al. 2007) and predict functions in the party package (<http://cran.r-project.org/web/packages/party/party.pdf>). To better visualize the actual tree structure, we used the ctree (Hothorn et al. 2006b) function in the party package to visualize the interactions among the most important variables that determine the population trend in the species included in our analysis.

Results and Discussion

Out of the 24 species classified as “Data Deficient” by the IUCN included in our analysis, 18 were predicted to be decreasing, and only five were classified as stable (Table 5). In predicting Population Trend, the most important variables were Habitat Loss/Degradation, Presence of Chytridiomycosis, Development Type, and Habitat Use (Fig. 1). Random Forests and cforest

(results not shown) show the same variables in the top four most important, which means that Random Forests does not have variable selection bias in this analysis. Therefore, we are confident in our results. However, according to cforest, the importance of habitat loss dwarfs the importance of the other three variables. Habitat loss/degradation was the most critical variable according to both types of analyses (Fig. 1), which concurs with the vast amount of information on the cause of species declines (Stuart et al. 2004).

Our randomForest analysis accurately identifies species with decreasing population trend (Table 4). In the test data there were 9 false positives and 8 false negatives (Precision = 0.897. Recall = 0.908; F1 score of 0.903 where 1 is a perfect score) for the “decreasing” category, which means that the method is likely to correctly flag a species as decreasing or stable. The errors are equally likely to be a false positive or negative.

Table 5 also shows the risk status according to the 2010 official Mexican National Red List for the Data Deficient species included in the analysis. Of them, only one, *Bolitoglossa stuarti*, is categorized as Endangered (A), while 5 of them (*Chiropterotriton mosaueri*, *Craugastor taylori*, *Eleutherodactylus maurus*, *Eleutherodactylus pallidus*, and *Eleutherodactylus teretistes*) are considered “Under Special Protection”, which is the lowest risk status of the List. The other 17 species have not been assessed at the national level and thus are not listed. Of all the species on the Mexican Red List, only *Bolitoglossa stuarti* and *Craugastor taylori* are predicted to have a stable population trend, so it would be advisable to assess the other 17 species for the new version of the official Mexican National Red List, which is the only national policy instrument which foresees law enforcement in order to protect Mexican threatened species. It thus becomes

clear why efforts like this are important in order to pinpoint priorities to fill the gaps needed to inform public policy and advance in the conservation of the species.

In a similar study of Australian amphibians, (Murray et al. 2011) found that Habitat Use (ecological group) was the most important variable to determine population trend, followed by the presence of chytridiomycosis and *Gambusia*, a predatory fish (defined by spatial models of suitability). Contrary to what we did, these authors included range size (EOO), abundance, and testes mass, and the presence of *Gambusia*. Although our study and that of Murray et al. (2011) found different variables as the most important to determine population trend, both studies agree on the fact that habitat use and the presence of chytridiomycosis are some of the most important variables. Moreover, our study concurs with that of Murray et al. (2011) in that by integrating intrinsic and extrinsic factors that are pertinent for the target region, one can get an accurate account of the population trend of a given amphibian species, as well the risk factors that are most pressing for the different ecological groups. In this study we have addressed the variables that are most pressing for Mexican species, and our results show, just as in the Australian case, that this kind of analysis can identify areas to focus limited conservation resources. Another important point is the geographic importance of the analysis. Because Mexican and Australian amphibians are subject to different extrinsic factors, an analysis appropriate for one does not necessarily apply to the other.

Conclusions

The use of Random forests seems to be a very solid and accurate way to identify species with decreasing population trends in the absence of demographic data. The kind of exercise that we

show here is an important first step when planning conservation priorities, as some of the most endangered species might also be those for which most information is lacking, thus falling through the cracks of conservation planning. Moreover, this method has the advantage of not having to depend on aggregated museum locality data that may not have been properly curated by experts, as is the case for some assessment efforts (Hjarding et al. 2014).

Intrinsic factors, such as the ova and clutch size that can give important information about how life history can affect the population trend of a species. Unfortunately, the amount of data we had for those traits was so limited, we felt that including data that had mostly been statistically generated could introduce an extra bias to our analysis. The fact that so little information on the natural history of these endangered species is available is a major challenge that needs to be addressed to successfully prevent their extinction. In addition, our aggregated data set can be used to set data collection priorities to fill in gaps. Fortunately, as we show here, this lack of information should not deter our efforts to assess risk status and assign priorities to their conservation.

Acknowledgements

The authors would like to thank the National Evolutionary Synthesis Center, the Biodiversity Heritage Library and the Encyclopedia of Life for supporting this project.

Funding

This work was supported by a grant "Using Digital Libraries to Discover Biodiversity and Evolution" from the Richard Lounsbery Foundation to C.R. McClain

Reference list

- Assessment ME. 2005. Ecosystems and Human Well-Being: Biodiversity Synthesis. World Resources Institute.
- Bland LM, Collen BEN, Orme CDL, and Bielby JON. 2014. Predicting the Conservation Status of Data-Deficient Species. *Conservation Biology*:n/a-n/a.
- Breiman L. 2001. Random forests. *Machine Learning* 45:15–32.
- Brooks TM, Mittermeier RA, Mittermeier CG, Da Fonseca GAB, Rylands AB, Konstant WR, Flick P, Pilgrim J, Oldfield S, Magin G, and Hilton-Taylor C. 2002. Habitat Loss and Extinction in the Hotspots of Biodiversity
- Pérdida de Hábitat y Extinciones en Áreas Críticas para la Biodiversidad. *Conservation Biology* 16:909-923.
- Cardillo M, and Meijaard E. 2012. Are comparative studies of extinction risk useful for conservation? *Trends in Ecology & Evolution* 27:167-171.
- Cutler DR, Edwards TC, Beard KH, Cutler A, Hess KT, Gibson J, and Lawler JJ. 2007. RANDOM FORESTS FOR CLASSIFICATION IN ECOLOGY. *Ecology* 88:2783-2792.
- Flores-Villela O, and Canseco-Márquez L. 2004. Nuevas especies y cambios taxonómicos para la herpetofauna de México. *Acta Zoologica Mexicana* 20:115-144.
- Frías-Alvarez P, Vredenburg V, Familiar-López M, Longcore J, González-Bernal E, Santos-Barrera G, Zambrano L, and Parra-Olea G. 2008. Chytridiomycosis Survey in Wild and Captive Mexican Amphibians. *EcoHealth* 5:18-26.
- García A, and Ceballos G. 1994. *Guía de campo de los reptiles y anfibios de la costa de Jalisco*. México: Fundación Ecológica de Cuixmala, A.C. .

- 289 Gardner T. 2001. Declining amphibian populations: a global phenomenon in conservation
290 biology. *Animal Biodiversity and Conservation* 24:25-44.
- 291 Groombridge B. 1992. *Global biodiversity: status of the Earth's living resources*. World
292 Conservation Monitoring Centre. New York, NY, USA. : Chapman & Hall.
- 293 Hjarding A, Tolley KA, and Burgess ND. 2014. Red List assessments of East African
294 chameleons: a case study of why we need experts. *Oryx*.
- 295 Hothorn T, Buehlmann P, Dudoit S, Molinaro A, and Van Der Laan A. 2006a. Survival
296 Ensembles. *Biostatistics* 7:355-373.
- 297 Hothorn T, Homik K, and Zeileis A. 2006b. Unbiased recursive partitioning: a conditional
298 inference framework. *Journal of Computational and Graphic Statistics* 15:651-674.
- 299 IUCN. 2012. *IUCN Red List Categories and Criteria: Version 3.1*. : Gland, Switzerland and
300 Cambridge, UK: IUCN.
- 301 IUCN. 2014. The IUCN Red List of Threatened Species. Version 2014.1. . Available at
302 <http://www.iucnredlist.org> (accessed May 12 2014).
- 303 J.S. P, S.L. A, and J.E. L. 2004. Physiology of *Batrachochytrium dendrobatidis*, a chytrid
304 pathogen of amphibians. *Mycologia*:9-15.
- 305 Liaw A, and Wiener M. 2002. The randomforest package. *R News* 2:18-22.
- 306 Lips KR, Brem F, Brenes R, Reeve JD, Alford RA, Voyles J, Carey C, Livo L, Pessier AP, and
307 Collins JP. 2006. Emerging infectious disease and the loss of biodiversity in a
308 Neotropical amphibian community. *Proceedings of the National Academy of Sciences of*
309 *the United States of America* 103:3165-3170.
- 310 Longcore JE, Pessier AP, and Nichols DK. 1999. *Batrachochytrium dendrobatidis* gen. et sp.
311 nov., a chytrid pathogenic to amphibians. *Mycologia*:219-227.

312 Murray KA, Rosauer D, McCallum H, and Skerratt LF. 2011. Integrating species traits with
 313 extrinsic threats: closing the gap between predicting and preventing species declines.
 314 *Proceedings of the Royal Society B: Biological Sciences* 278:1515-1523.

315 Ochoa-Ochoa LM, and Flores-Villela O. 2006. *Areas de Diversidad y Endemismo de la*
 316 *Herpetofauna Mexicana*. México, D.F. : Universidad Nacional Autonoma de Me´xico—
 317 Comision Nacional para el Conocimiento y Uso de la Biodiversidad.

318 Parr CS, Wilson N, Leary P, Schulz KS, Lans K, Walley L, Hammock JA, Goddard A, Rice J,
 319 Studer M, Holmes JTG, and Corrigan RJ Jr. 2014. The Encyclopedia of Life v2:
 320 Providing global access to knowledge about life on Earth. *Biodiversity Data Journal*
 321 2:e1079.

322 Parra-Olea G, García-París M, and Wake DB. 1999. Status of some populations of Mexican
 323 salamanders (Amphibia:Plathodontidae). *Revista de Biología Tropical*:217-223.

324 Strobl C, Boulesteix AL, Kneib T, Augustin T, and Zeileis A. 2008. Conditional Variable
 325 Importance for Random Forests. *BMC Bioinformatics* 9:307.

326 Strobl C, Boulesteix AL, Zeileis A, and Hothorn T. 2007. Bias in random forest variable
 327 importance measures: illustrations, sources and a solution. *BMC Bioinformatics* 8.

328 Stuart SN, Chanson JS, Cox NA, Young BE, Rodrigues ASL, Fischman DL, and Waller RW.
 329 2004. Status and Trends of Amphibian Declines and Extinctions Worldwide. *Science*
 330 306:1783-1786.

331 Tingley R, Hitchmough RA, and Chapple DG. 2013. Life-history traits and extrinsic threats
 332 determine extinction risk in New Zealand lizards. *Biological Conservation* 165:62-68.

333 van Buuren S, and Groothuis-Oudshoorn K. 2011. mice: Multivariate Imputation by Chained
 334 Equations. *R Journal of Statistical Software* 45:1-67.

335 Wake DB. 1991. Declining amphibian populations. *Science*:860.
336 Wyman RL. 1990. What's Happening to the Amphibians? *Conservation Biology*:350-352.

337

338

339

340

341

342

343

344

345

346

347

348

349

350

351

352

353

354

355

356

357

358 Figure Titles and Captions

359

360 Figure 1

361 Title: Relative Importance of Variables for Predicting Population Trend

362 Caption: Bar graph showing the relative importance of all variables for predicting population
363 trend. The individual variables are listed on the vertical axis. The horizontal axis shows the
364 decrease in accuracy of the final result if the variable is removed. Important variables have a
365 higher mean decrease in accuracy.

366

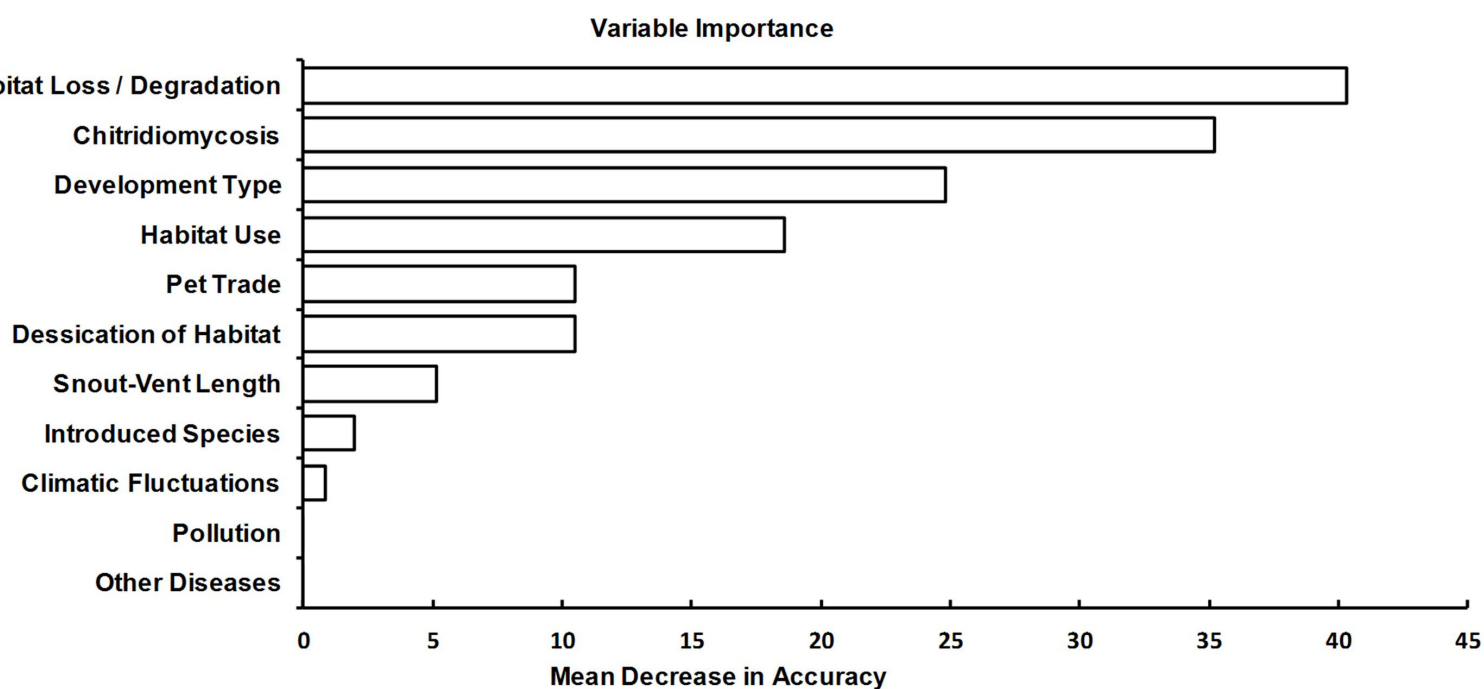


Table 1: EOL chapters and keywords used to filter and harvest relevant text data object for the study.

Trait Type	EOL Chapter	Keyword
Intrinsic	Size, Reproduction, Life Cycle	length, clutch, egg, breeding, development, reproduction, hibernation
Extrinsic	Distribution, Habitat	occur, range, inhabit, found, precipitation, wet, arid, dry, moist, temperature, temperate, tropic

Table 2: Number of missing data points for each variable. The 30 “missing” data points for the IUCN status actually refer to the number of Data Deficient and Not Evaluated species.

Trait	Missing Data
Snout-Vent Length	11
Habitat Use	1
Ova Size	276
Development	5
Clutch Size	252
Habitat Loss/Degradation	4
Introduced Species	4
Pollution	4
Chytridiomycosis	4
Climatic Fluctuations	4
Pet Harvest	4
Desiccation of Habitat	4
Other Diseases	4
IUCN status	30
Population Trend	53

Table 3: Numeric categories codes for the traits used in the study.

Trait	Category	Definition
Snout-Vent Length	1	up to 69 mm
	2	70-120 mm
	3	121-171 mm
	4	more than 172 mm
Habitat Use	1	ephemeral pond associated
	2	permanent water associated
	3	stream associated
	4	terrestrial
Development	1	direct development
	2	larval development
	3	paedomorphic
Habitat Loss/Degradation	0	absent
	1	present
Introduced Species	0	absent
	1	present
Pollution	0	absent
	1	present
Chytridiomycosis	0	absent
	1	present
Climatic Fluctuations	0	absent
	1	present
Pet Trade/harvest	0	absent
	1	present
Desiccation of Habitat	0	absent
	1	present
Other Diseases	0	absent
	1	present
Population Trend	0	Decreasing
	1	Stable

Table 4: Confusion matrix obtained using the randomForest and predict functions in the randomForest package (Liaw & Wiener 2002) on the test data to predict population trend.

	PREDICTED	
OBSERVED	Decreasing	Stable
Decreasing	79	9
Stable	8	30

Table 5: Predicted population trend for the 24 species classified as Data Deficient by the IUCN. The categories on the 2010 official Mexican National Red List (NOM-Semarnat-059-2010) are as follows: E, extinct; P, endangered; A, threatened; Pr, under special protection.

Species	Population Trend predicted	Mexican Red List
<i>Bolitoglossa oaxacensis</i>	decreasing	-
<i>Bolitoglossa stuarti</i>	stable	A
<i>Bolitoglossa zapoteca</i>	decreasing	-
<i>Chiropterotriton mosaueri</i>	decreasing	Pr
<i>Craugastor amniscola</i>	decreasing	-
<i>Craugastor occidentalis</i>	stable	-
<i>Craugastor pelorus</i>	decreasing	-
<i>Craugastor taylori</i>	stable	Pr
<i>Eleutherodactylus maurus</i>	decreasing	Pr
<i>Eleutherodactylus pallidus</i>	decreasing	Pr
<i>Eleutherodactylus teretistes</i>	decreasing	Pr
<i>Exerodonta abdivita</i>	decreasing	-
<i>Exerodonta bivocata</i>	stable	-
<i>Lithobates lemosespinali</i>	decreasing	-
<i>Pseudoeurycea amuzga</i>	decreasing	-
<i>Pseudoeurycea maxima</i>	decreasing	-
<i>Pseudoeurycea mixcoatl</i>	decreasing	-
<i>Pseudoeurycea obesa</i>	decreasing	-
<i>Pseudoeurycea quetzalanensis</i>	decreasing	-
<i>Pseudoeurycea tlilicxitl</i>	stable	-
<i>Ptychohyla acrochorda</i>	decreasing	-
<i>Ptychohyla zophodes</i>	decreasing	-
<i>Thorius insperatus</i>	decreasing	-