



Amphibians over the edge: silent extinction risk of Data Deficient species

Sam D. Howard* and David P. Bickford

Department of Biological Sciences, National University of Singapore, 14 Science Drive 4, 117543 Singapore, Singapore

ABSTRACT

Aim To apply mathematical models to the task of predicting extinction risk for species currently listed as ‘Data Deficient’ (DD) by the International Union for the Conservation of Nature (IUCN). We demonstrate this approach by applying it globally to amphibians, the vertebrate group recognized as being most extinction threatened and having the largest proportion of DD species. We combine model predictions with current extinction risk knowledge to highlight regions of greatest disparity between known and predicted risk, where potential species extinctions may be overlooked.

Location Global.

Methods Using global amphibian distribution data obtained from the IUCN and species trait data, we apply machine learning *randomForest* models to predict extinction risk of DD species from life history traits, environmental variables and habitat loss. These models are trained using data for species that have been assigned to an extinction risk category (other than DD) by the IUCN. We then combine predictions for DD species with IUCN assessment data in a GIS framework to highlight anomalies between current knowledge of amphibian extinction risk and our model predictions.

Results We show that DD amphibian species are likely to be more threatened with extinction than their fully assessed counterparts. Regions in South America, central Africa and North Asia are particularly at risk due to lack of species knowledge and higher extinction risk than currently recognized.

Main conclusions Application of predictive models ranking regions and species most in need of primary research allows prioritization of limited resources in an informed context, minimizing risk of unnoticed species’ extinction.

Keywords

Anura, Caudata, Gymnophiona, IUCN Redlist, machine learning, predictive model.

*Correspondence: Sam D. Howard, Department of Biological Sciences, National University of Singapore, 14 Science Drive 4, 117543 Singapore, Singapore.
E-mail: samdhoward@gmail.com

INTRODUCTION

Increasing human population size and associated environmental modifications are placing the world’s species at ever increasing risk of extinction (Butchart *et al.*, 2010; Hoffmann *et al.*, 2010). Extinction risk is assessed by the International Union for the Conservation of Nature (IUCN) ‘Red List’ via expert assessment of species’ extent of occurrence (EOO), area of occupancy (AOO) and population trend (IUCN, 2011). Population trend is an amalgamation of a wide range of potential data, often varying between species, which may include information relating to population size, population

fragmentation, demographic analysis and local specific threat processes (IUCN, 2011). Species with insufficient information for experts to be confident in assessing extinction risk status are termed ‘Data Deficient’ (DD) and large proportions of species in all but the most-studied groups fall into this category. In addition to DD species, many taxonomic groups contain a large proportion of unevaluated species (IUCN status NE, ‘Not Evaluated’). Despite these limitations, the IUCN Red List is used in a wide range of policy making, conservation management decisions and research agendas (Possingham *et al.*, 2002; Rodrigues *et al.*, 2006). The lack of extinction risk classification for DD species is likely to result

in these species being overlooked by conservation legislation and management.

High profile vertebrate groups, such as mammals, birds and amphibians, have near complete IUCN coverage (IUCN, 2011), while other taxonomic groups have lower assessment completeness. In species groups where existing knowledge is mostly complete (e.g. mammals), there is a wealth of literature exploring species extinction risk at the global scale, and consideration of DD taxa is less of a significant issue (Cardillo *et al.*, 2008; Fritz *et al.*, 2009; Lee & Jetz, 2011). On the other hand, the IUCN has, for example, assessed less than 10,000 species of fish out of an estimated total of thirty thousand species globally (Froese & Pauly, 2012). The prevalence of DD and NE species in the IUCN Red List is likely to increase as the assessment process is extended to cover a broader range of highly diverse, yet lesser known, taxon groups. For these less studied groups, development of tools and protocols that provide species-specific or regional estimates of extinction risk for DD species (e.g. Morais *et al.*, 2013) serves several important needs.

First, excluding birds and mammals, most faunal taxonomic groups have high proportions of DD species that impede understanding of the group's overall extinction risk. Predictions of extinction risk for DD species in these groups will provide a more complete picture of potential extinction threat globally. Second, conservation funds are often limited and therefore focused towards species known to be at high extinction risk (Stuart *et al.*, 2004; Morais *et al.*, 2012), neglecting poorly known taxa (Brito, 2010). In addition, funding limitations tend to bias research towards areas of recognized high species richness, where return on investment in data acquisition is more guaranteed, and there is little incentive to explore areas of deficient biodiversity knowledge (Brito, 2010). By comparing distributions of DD amphibians to newly described species, Brito (2010) provides compelling justification and a mechanism, for directing research towards detecting undescribed species. While predictions of extinction risk do not replace formal classification, they can provide, in a similar manner to the work of Brito (2010), a means of focusing research towards collecting primary data on species predicted to be at high risk. This acts to accelerate formal classification of DD species in a risk-prioritized manner.

Analysis of causes of extinction risk is commonplace in the literature for better-known vertebrate groups, including birds (e.g. Owens & Bennett, 2000; Collen *et al.*, 2006; Lees & Peres, 2008; Lee & Jetz, 2011), mammals (e.g. Collen *et al.*, 2006; Cardillo *et al.*, 2008; Davidson *et al.*, 2009; Fritz *et al.*, 2009), amphibians (e.g. Bielby *et al.*, 2008; Cooper *et al.*, 2008; Sodhi *et al.*, 2008), reptiles (Collen *et al.*, 2006) and fish (Collen *et al.*, 2014). Conclusions of such studies commonly find that species' range size and habitat loss are major drivers of extinction risk and that species-specific factors, for example, body size (Owens & Bennett, 2000; Fritz *et al.*, 2009) or geographical location (Fritz *et al.*, 2009), interact with these to modify extinction risk. In terms of

predicting the status of DD species, it can be expected that similar relationships will be in operation as have been found for fully assessed species.

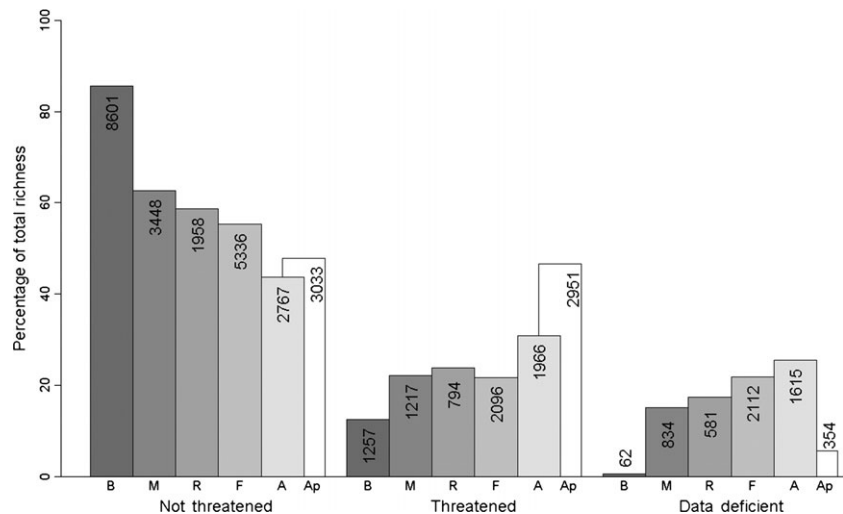
Amphibians are the most threatened group of vertebrates with a third of currently known species endangered with extinction (IUCN, 2008) and the majority experiencing population declines (Stuart *et al.*, 2004). A number of recent extinctions have been attributed to the fungal disease chytridiomycosis and at least one in synergy with global warming (Pounds *et al.*, 2006). In addition, threats from habitat loss (Stuart *et al.*, 2004; Cushman, 2006; Becker *et al.*, 2007), wild harvesting for food (Warkentin *et al.*, 2009) and trade (Stuart *et al.*, 2006), and increases in ultraviolet radiation (Kiesecker *et al.*, 2001) are linked to population declines (Beebe & Griffiths, 2005).

Knowledge of natural history, geographical distribution and population trend is incomplete for many amphibian species, and new species are regularly being described. While a large proportion of fully assessed amphibian species are at risk, amphibians are also the most DD group of vertebrates (see Fig. 1), with a quarter of species (IUCN, 2008) unable to be assessed based on IUCN criteria (IUCN, 2011). Total numbers of amphibian species vary with taxonomic authority and date of the species list, with recent estimates ranging from 6771 species (Frost, 2011) to 6919 species (IUCN, 2011) or 7057 species (AmphibiaWeb, 2012). However, new amphibian species are still being described at a rapid rate, and true richness is therefore likely to be substantially above current figures (Glaw & Köhler, 1998; Dubois, 2005; Kohler *et al.*, 2005; Rowley *et al.*, 2010).

Amphibian declines and drivers of these declines, notably habitat loss and increased disease susceptibility, have been examined by several studies (Kiesecker *et al.*, 2001; Stuart *et al.*, 2004; Pounds *et al.*, 2006; Sodhi *et al.*, 2008; Hof *et al.*, 2011), but DD amphibians have been excluded from all of these analyses. In their recent model-based projection, Hof *et al.* (2011) predicted impacts of three major extinction processes on amphibians by the year 2080: climate change, land-use change and chytrid fungal disease. While their models are based on total amphibian richness, including DD species, they did not specifically differentiate between these categories. Attempts to address the status of DD vertebrates have been rare and largely restricted to birds at the global scale (Butchart & Bird, 2010), with just one recent study addressing local-scale amphibian status (Morais *et al.*, 2013). By modelling EOO against time since discovery, Morais *et al.* (2013) provide a method of predicting extinction risk for DD species of Brazilian amphibians based solely on locality data, providing an important tool to focus conservation efforts. However, while 'area' is often an important predictor of species extinction risk (e.g. Sodhi *et al.*, 2008), there are many other factors that contribute, and inclusion of data relating to other threat processes is likely to improve the accuracy of predictions, especially at broader spatial scales.

Existing extinction threats to amphibians are high, and projections indicating that factors underlying these threats

Figure 1 Broad extinction risk categories for vertebrate groups. Shaded bars show the percentage of International Union for the Conservation of Nature (IUCN) assessed species richness in each broad IUCN threat category for birds (B), mammals (M), reptiles (R) and fish (F), respectively, with extinct species included in threatened. IUCN-assessed amphibian extinction risk (A) is shown overlapping results incorporating our model predictions (Ap). Numbers associated with bars represent species richness for each group/class.



will increase in severity suggest a worsening scenario (Hof *et al.*, 2011). To fully evaluate the significance of this potential extinction crisis, there is a need to quantify and address extinction threat for the large portion of amphibians that are currently overlooked. There is unlikely to be enough time and person-power to collect basic data on all DD amphibians before species declines occur. Along with continued protection of species already recognized as being at risk, there needs to be a prioritization mechanism for investigating the status of un-assessed species before their population declines, potentially to extinction, without being noticed. We here address this challenge for DD amphibians at the global scale. We use data for species that have been assigned to an extinction risk category (other than DD) by the IUCN to fit a model relating extinction risk to life history traits, environmental variables and habitat loss. We then use this model to predict the extinction risk of DD species and employ these results, alongside existing IUCN data, to assess whether DD amphibians can be considered as at greater or lesser risk than fully assessed species. By comparing spatial anomalies between fully assessed species and our model predictions, we provide a tool that should help focus limited conservation resources towards areas with a concentration of DD species predicted to be at high risk of extinction.

METHODS

Data sources

We used the amphibian trait database compiled by Sodhi *et al.* (2008) that provides life history and geographical data on amphibian species described prior to 2005. We have not updated these data to reflect newly described species as such species are highly likely to be DD simply due to their novelty. We used the most recent amphibian extinction risk data and range shape polygons available from the IUCN (IUCN, 2011) and compared the 5718 species present in the 2008 trait dataset to the current IUCN species list and updated

species names accordingly. We excluded 67 species that have been taxonomically split or synonymized since 2008, and could no longer be matched to existing trait data with certainty. We updated EOO, from IUCN range shapefiles, and extinction classification for the remaining 5651 species to reflect changes since 2005. We used the most recent amphibian range shapefiles (IUCN, 2011) to update additional geographical variables (i.e. latitude, see Table S1 in Supporting Information) and trait data remained unchanged. A full list of variables used in our analysis can be found in supplementary material (Table S1) and the original article describing this dataset (Sodhi *et al.*, 2008). The resulting amphibian database we analysed contained 5651 amphibian species described prior to 2005 with the most updated spatial and trait data, representing 82% of current amphibian diversity (Table S2). Of these species, 4402 had been fully assessed for extinction risk by the IUCN, while the remaining 1249 were listed as DD. Categories of extinction risk for the assessed species were as follows: least concern (LC), near threatened (NT), vulnerable (VU), endangered (EN), critically endangered (CR), extinct in the wild (EW) and extinct (EX).

Modelling

Data for the IUCN-assessed species were used to fit a model relating assessed extinction risk to area, geographical and life history variables. All analyses were based on *randomForest* classification models (Breiman, 2001, 2003), implemented in the *RANDOMFOREST* package (Liaw & Wiener, 2002) in R version 2.13.1. (R Development Core Team, 2011). *RandomForests* are a machine learning classification and regression tree (CART) model, using recursive binary partitioning to split predictor variables, replicated across numerous bootstrapped trees grown to maximum size (with no tree pruning). Individual *randomForest* trees are constructed from a random subset of available data (2/3 of the data), and 'out-of-bag' error (OOB) is determined by assessing each tree against the accuracy of predictions for the excluded (1/3)

data subset; thus, each tree in the ‘forest’ gives a measure (OOB error) of self-assessment in classification accuracy. In addition, each node within a tree is assessed against a random sample of available predictor variables (Mtry parameter in *randomForest*). All trees within the forest are then compared, and aggregated by tree averaging, with the final classification of species resulting from an ensemble vote across all trees in the forest. *RandomForests* are nonparametric and relatively free of assumptions regarding the form of relationships between predictor and response variables, allowing modelling of hierarchical and nonlinear relationships. In addition, *randomForests* are not prone to data over-fitting (Breiman, 2001; Robnik-Sikonja, 2004).

RandomForest models are controlled by two adjustable parameters: forest size and the number of variables considered at each node within forest trees. We assessed forest size using the subset of IUCN-assessed amphibian species with no missing data from our database (1621 species). We observed no difference between forests of 500, 1000 or 1500 trees in terms of variable importance and graphs of classification error, suggesting 500-tree forests are sufficient. However, we opted for a more cautious approach and used 1000-tree forests in all of the analyses presented here. The number of variables considered at each splitting node (Mtry) was assessed following suggested methods (Breiman, 2003). We compared by eye, performance graphs of 1000-tree forests on the data subset with Mtry parameters of 2, 5 and 10. Observing no noticeable difference in performance with varying Mtry, we used a default Mtry of five variables.

Data completeness

The *randomForest* modelling process does not accept missing data, and some species–trait pairs in the dataset were missing due to lack of knowledge of species ecology. We assessed the amount of missing data within the database in terms of the number of species with any missing data points, and in terms of the proportion of missing data within the total species–trait matrix (Table S3). We concluded that exclusion of species with any missing data (71% of species) was too extreme when compared to the proportion of actual missing data points (2.3% of the data matrix). We performed exploratory analysis on the dataset of 1621 species with complete trait data, excluding species with missing values (analysis not shown) to assess potential model suitability.

The model containing area (EOO), geographical and trait data performed best (Table S4). We used this model structure to impute missing data values using *randomForest* (Breiman, 2001, 2003). Firstly, missing data were assigned values based on medians (continuous) or modes (factoral) from the dataset. A *randomForest* was then constructed from this ‘first guess’ data and imputed values updated based on proximity-weighted average of non-missing cases or the class with the highest average proximity for categorical variables. This process was iterated five times. *RandomForest* imputation has been shown to be robust in assigning meaningful and

accurate values to missing data points (Breiman, 2003; Cutler *et al.*, 2007; Pantanowitz & Marwala, 2009; Latifi & Koch, 2012). All subsequent analyses were performed on imputed data, restricted to species’ life history characteristics, because spatial data were complete. This imputation method, restricted to species life history traits, assigns values to missing data through iteration based on a proximity matrix and a final imputed value will be an ‘average’ value of the missing variable across species with similar non-missing variable combinations. This imputation method has been employed in other studies of species extinction risk (Musters *et al.*, 2013).

RandomForest models self-assess accuracy through OOB error estimates, and we compared the predictions for IUCN-assessed species to their IUCN classification using a chi-squared test on row/column totals of the OOB error matrix (combining EX and EW categories to account for zero values from model EW class prediction). We determined the direction of model error by averaging the number of misclassified species multiplied by the numeric difference between their assessed and predicted classification.

The *randomForest* model trained on the data for fully assessed species was then used to predict the extinction risk of DD species based on area, geographical and life history variables.

Spatial mapping of results

Mapping of results was conducted in ArcGIS 9.3.1 (ESRI, 2009) with Spatial Analyst and Hawth’s Tools extensions (Beyer, 2004). We attached the most recent IUCN classification of extinction risk for IUCN-assessed species, and the predicted extinction risk from our model for both assessed and DD species, to each IUCN species polygon. For presentation purposes, we assigned ordinal values to the extinction risk classes based on our impression of the relative significance of these categories from a conservation perspective (LC = 1, NT = 3, VU = 5, EN = 7, CR = 9, EW = 15, EX = 20). We removed species not in the database from the IUCN amphibian range polygons and split the resulting polygons into assessed and DD species subsets, and a combination of these. We constructed a 2×2 degree latitude-longitude grid bounded by the maximum/minimum extent of global amphibian class distribution, resulting in 6996 grid cells. The polygon grid was cropped to land areas (excluding oceans) before analysis. Data were spatially projected in Robinson projection to conserve area. We calculated per-cell species richness for all DD amphibians (see Fig. S1). We used the ‘polygon in polygon’ tool (Beyer, 2004) to extract the sum and area-weighted mean (AWM) of assessed extinction risk for IUCN-assessed species and predicted extinction risk for DD species, to each grid cell. We performed the same mapping with the combined data for assessed and DD extinction risk. To identify regions where the extinction risk to DD species is higher than that currently recognized for assessed species, we subtracted assessed extinction risk from

predicted DD risk on a cell by cell basis for both sum and AWM scores. Where differences (dsum and dAWM) are negative (assessed > DD), current knowledge of localized amphibian extinction risk is likely sufficient. Where differences are positive (DD > assessed), overall localized extinction risk is potentially underestimated. It is far more likely for dAWM to be positive than for dsum due to weighting based on species richness ($SR_{\text{assessed}} = 4402 \gg SR_{\text{DD}} = 1249$). These three potential outcomes ($d < 0$, $dAWM > 0$, $dsum > 0$) were merged into a single colour-coded map.

RESULTS

Models incorporating life history and geographical variables and higher taxonomy performed better than those based solely on the IUCN criteria of Area and Population Trend (Table S4). The model with highest accuracy correctly predicts extinction risk of 74% of fully assessed amphibian species. Comparing IUCN extinction risk to that predicted by our model for the same assessed species shows a significant difference in extinction risk distribution ($\chi^2 = 226.3$, d.f. = 6, $P < 0.01$) with the model generally underestimating extinction risk for the 26% mis-classified species (Table S5). Underestimates of extinction risk for fully assessed species are highly likely to be mirrored in predictions for DD species. The correlates of extinction risk identified in our model are similar, albeit based on a different method and updated data, to those identified by Sodhi *et al.* (2008), where species range (EOO) was the single largest contributor to extinction risk in amphibians (see Sodhi *et al.*, 2008 for other insights). Our major interest here, however, is not to re-examine these correlates of extinction risk, but to use the model trained on data for assessed species to predict the conservation status of DD species.

Our model predicts that DD amphibians are, on average, likely to be more threatened by extinction than fully assessed species. The model predicts that 63% of DD amphibians are probably threatened (IUCN classes VU, EN, CR and

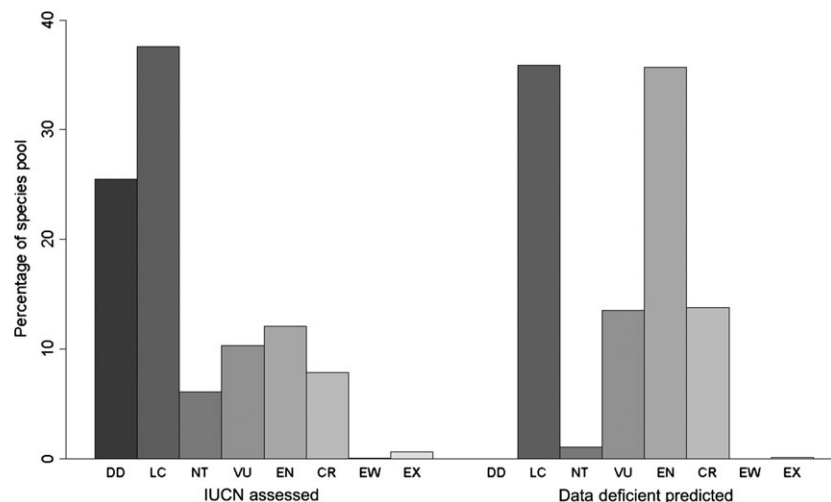
including EW, EX), while 32% of formally assessed species (IUCN, 2008) are classified in these same categories (Fig. 2). Amphibians are already acknowledged as the most threatened vertebrate group, but our model results suggest that overall extinction risk for this group has been considerably underestimated to date (Fig. 1).

Comparing the spatial distribution of extinction risk for fully assessed species to our model-based predictions of extinction risk for DD species highlights regions where there are large gaps in current knowledge (Fig. 3). For cells where fully assessed risk exceeds predicted risk for DD species, we consider these cells to have sufficient current knowledge such that area-based conservation strategies (such as protected areas and habitats) would be likely to adequately protect DD species in that cell. Where predicted risk for DD species exceeds that for fully assessed species, conservation priorities based on current knowledge are likely to be deficient and to overlook significantly threatened species. Where the sum of fully assessed extinction risk is less than model predictions, global recognition of amphibian extinction risk is critically underestimated and urgent measures are needed to formally assess these DD species. Distinct clustering of extinction risk anomalies is evident (Fig. 3), and geographical regions highlighted by this mapping roughly match recognized biodiversity hotspots (Myers *et al.*, 2001). The spatial clustering of extinction risk anomalies is most pronounced in the Andes of Western South America, the Central Africa region, Madagascar and the Indo-Burma/Chinese mountain regions.

DISCUSSION

Prediction of extinction risk for DD species is not intended as a shortcut to bypass formal IUCN assessment, but rather as a tool to focus ongoing data collection and research. Specific extinction risk predictions are provided for individual species (see Data S1), but should not be used to replace formal IUCN assessment. The use of predictive models to fill gaps in knowledge of extinction risk will prove vital to

Figure 2 Distribution of assessed and predicted risk. Extinction risk distribution for International Union for the Conservation of Nature (IUCN) assessed (left) and model predictions for DD (right) amphibians as percentage of total dataset richness by IUCN extinction risk class. DD, Data Deficient; LC, Least Concern; NT, Near Threatened; VU, Vulnerable; EN, Endangered; CR, Critically Endangered; EW, Extinct in the Wild; and EX, Extinct.



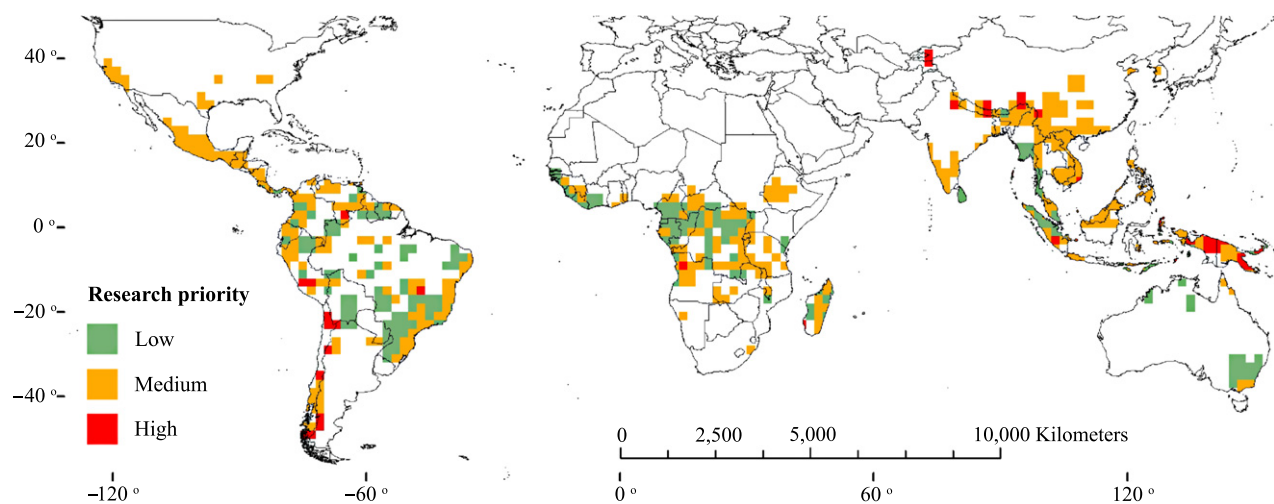


Figure 3 Amphibian extinction risk anomalies. Coloured cells show the distribution of Data Deficient (DD) amphibians in our analysis. Cells are colour-coded to indicate priority areas for formal species assessments, based on anomaly between observed and predicted extinction risk. Cells in green have lower predicted extinction risk of DD species compared with other International Union for the Conservation of Nature (IUCN) categories (by sum and mean). Cells in orange have predicted mean DD risk in excess of IUCN mean risk. Cells in red have the sum of predicted DD risk in excess of the sum of IUCN risk.

keeping pace with escalating threat processes on a global scale, allowing informed focus of limited conservation resources towards formal species assessments.

Our analysis shows that amphibian species are likely to be more threatened by extinction than previously recognized, with DD species predicted to be more threatened, on average, than fully assessed species. The only other analysis predicting extinction risk status of DD amphibians (Morais *et al.*, 2013) reported marginally less risk to re-classified Brazilian species (57% threatened) than shown for our global analysis (63%). Given differences in methods and scope of these two studies, it is difficult to assess this divergence, but both studies predict a far higher extinction risk for DD amphibians than is currently recognized by the IUCN for fully assessed species globally (32% threatened).

Our analysis further highlights areas containing DD species predicted to be at high risk of extinction, which would be otherwise overlooked if conservation planning were based solely on current knowledge of amphibian extinction risk (Fig. 3). The most successful conservation strategies involve habitat preservation and conservation of functional ecosystems (Foley *et al.*, 2005; Hooper *et al.*, 2005; Gardner *et al.*, 2009; Koh & Sodhi, 2010), and this analysis provides a tool to highlight specific areas where such conservation funds might be most effectively targeted, and where formal species assessments are most urgently needed. These areas should be prioritized for future research, to gather basic information necessary to formally assess, monitor and protect the species they contain. Given the numbers of DD amphibians and their predicted threat status, the risk of 'silent extinctions' is too large to ignore.

The specific regions highlighted in our results show considerable overlap with regions identified by a previous spatial

assessment of major amphibian threat processes (Hof *et al.*, 2011). The high-risk regions in South America and Madagascar identified by our study are already recognized for their high species richness and levels of species endemism, respectively. These factors are likely to continue to entice researchers to study these patterns, and additional information on DD species will be collected incidentally, if not by design. Areas in Central Africa and Asia are less recognized and have been largely inaccessible due to human conflict and political isolation. While robust links between DD species richness and the discovery of new species are not well established, it is logical to postulate that low levels of knowledge of species extinction risk are likely to correlate with high levels of undocumented diversity.

The plight of Asian biodiversity has been well discussed previously (Sodhi *et al.*, 2004; Sodhi & Brook, 2008; Bradshaw *et al.*, 2009; Koh & Sodhi, 2010; Rowley *et al.*, 2010; Sodhi *et al.*, 2010), and specifically, amphibian diversity in the Indo-Burma region is highlighted in our analysis. It contains high richness of DD species (Fig. S1), high disparity between recognized and unrecognized extinction risk (Fig. 3), overlaps with all major threat processes (Hof *et al.*, 2011) and has high rates of species discovery (Rowley *et al.*, 2010), but is less well recognized as a priority region for conservation.

The island of New Guinea shows general deficiency in recognized amphibian extinction risk but may be considered as the last true wilderness area and so gaps in ecological knowledge are more recognized. Elsewhere, there are noticeable gaps in the distribution of DD amphibians. Regions of the Amazon basin in South America, Central Africa and South-East Asia all contain grid cells with no DD species (Fig. S1). This pattern can be partially explained by our exclusion of

species described after 2005 (compare Fig. 3 to Fig. S1), where recently described species may occur in the empty cells. Alternatively, empty cells can represent regions where basic species distribution information is completely lacking. For example, the island of Borneo shows noticeable DD/extinction anomaly towards the Northern Malaysian states of Sabah and Sarawak (Fig. 3), but no DD species are recorded to the South (Kalimantan, Indonesia). This is probably an artefact of incomplete sampling rather than a reflection of good species knowledge. Additional research in such areas is likely to extend ranges of assessed and DD species and likely to reveal undocumented diversity.

Notable among our predictions for individual DD species is the predicted extinction of the rhacophorid frog *Pseudophilautus* (*Philautus* in our database) *semiruber*, known only from the type locality in Sri Lanka, although new information (Meegaskumbura *et al.*, 2012) shows this species to be extant as of 2005, where it is still likely to be highly threatened.

The analysis we have conducted here was made possible only through the construction of a life history database of amphibian species (Sodhi *et al.*, 2008). While we have been able to address some sources of extrinsic threat, information on others such as climate change, disease and recently emerging habitat threats is scarce, but likely to improve in the future. Current initiatives (see IUCN, 2011) directed towards locating and cataloguing spatial datasets of extrinsic threat processes will prove invaluable in regional, global and species-specific spatial modelling. For example, Amphibian chytrid fungal disease, a major potential source of extinction risk, has been recently added to the notifiable diseases list, and quantitative spatial data on occurrence may become available in the near future (Schloegel *et al.*, 2010).

While the focus on extrinsic threat processes is important, these extrinsic processes are likely to operate at a species-specific level, interacting with species' life history traits. For example, hunting pressure (for food or trade harvest) is unlikely to affect all species in an area equally, but logically is likely to be influenced by species' body size (Milner *et al.*, 2007; Laurance *et al.*, 2008) and other trait interactions at the species level. Importance of integrating species traits with extrinsic processes has been shown in Australian amphibians (Murray *et al.*, 2011), where interactions between recognized high risk traits (small range size) and extrinsic threat processes (such as disease) create specific mechanisms for species decline at the guild level. Many more potential interactions are likely when assessing extrinsic threats across a wide range of species at broad geographical scales. In addition to collecting extrinsic threat data, we therefore feel there is equal, or potentially greater, benefit to be gained by compiling databases of relevant life history traits. Life history data are unlikely to change significantly over time, and such databases would require minimal maintenance.

While our analysis has considered only amphibians, many of the processes responsible for DD status are not restricted to amphibians and the areas we highlight are therefore likely to be important for many other taxonomic groups. We

highlight and reiterate that our approach and results are intended as a tool to focus conservation effort and limited research funds towards areas where they are most likely to be most beneficial, following the conclusions of Brito (2010). Promoting research into the conservation status of poorly known areas and DD species is vital to improving our overall understanding of extinction risk as a foundation for effective conservation management.

Caveat

Missing data points were not evenly distributed between IUCN-assessed and DD species (Table S3). While imputation through randomForest has been shown to be robust (Breiman, 2001), a greater number of points were estimated for DD than for assessed species. Models predicting extinction risk rely heavily on EOO and by definition; the certainty of EOO will be lower for DD species than for known taxa. One potential effect of focused research on DD species would be to assess and expand EOO data and doing so would significantly improve species' extinction risk predictions. While we have been conservative in the inclusion of DD species for analysis, only including those described prior to 2005 to allow sufficient time for species' knowledge to be acquired (especially for EOO), and the model is conservative (underestimating extinction risk for IUCN-assessed species), it is still possible that our predicted extinction risk is an overestimate (due to effects of measured EOO being inaccurate). The models were based on the most complete and up to date information available and represent the best possible estimates of extinction risk for DD amphibians, with the major assumption that current knowledge is complete.

ACKNOWLEDGEMENTS

We thank N. S. Sodhi, A. C. Diesmos, T.M. Lee, L.P. Koh, B. W. Brook, C. H. Sekercioglu and C. J. A. Bradshaw for access to their amphibian trait dataset. Comments from J. Phelps, S. Goutte, T. Webb and S. Poo have significantly improved on earlier versions of this manuscript. Funding was provided by the National University of Singapore through a research scholarship to SDH and Grant R-154-000-383-133 to DPB.

REFERENCES

- AmphibiaWeb (2012) *Information on amphibian biology and conservation*. [web application]. Available at: <http://amphibia-web.org/> (accessed 8 November 2012).
- Becker, C.G., Fonseca, C.R., Haddad, C.F.B., Batista, R.F. & Prado, P.I. (2007) Habitat split and the global decline of amphibians. *Science*, **318**, 1775–1777.
- Beebe, T.J.C. & Griffiths, R.A. (2005) The amphibian decline crisis: a watershed for conservation biology? *Biological Conservation*, **125**, 271–285.

- Beyer, H.L. (2004) *Hawth's analysis tools for ArcGIS*. Available at: <http://www.spatial ecology.com/htools/tooldesc.php> (accessed 5 April 2009).
- Bielby, J., Cooper, N., Cunningham, A.A., Garner, T.W.J. & Purvis, A. (2008) Predicting susceptibility to future declines in the world's frogs. *Conservation Letters*, **1**, 82–90.
- Bradshaw, C.J.A., Sodhi, N.S. & Brook, B.W. (2009) Tropical turmoil: a biodiversity tragedy in progress. *Frontiers in Ecology and the Environment*, **7**, 79–87.
- Breiman, L. (2001) Random forests. *Machine Learning*, **45**, 5–32.
- Breiman, L. (2003) Manual for setting up, using, and understanding random forest V4.0. Available at: http://www.stat.berkeley.edu/~breiman/Using_random_forests_v4.0.pdf (accessed 21 May 2009).
- Brito, D. (2010) Overcoming the Linnean shortfall: data deficiency and biological survey priorities. *Basic and Applied Ecology*, **11**, 709–713.
- Butchart, S.H.M. & Bird, J.P. (2010) Data Deficient birds on the IUCN Red List: what don't we know and why does it matter? *Biological Conservation*, **143**, 239–247.
- Butchart, S.H.M., Walpole, M., Collen, B. *et al.* (2010) Global biodiversity: indicators of recent declines. *Science*, **328**, 1164–1168.
- 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**, 1441–1448.
- Collen, B., Bykova, E., Ling, S., Milner-Gulland, E.J. & Purvis, A. (2006) Extinction risk: a comparative analysis of central Asian vertebrates. *Biodiversity & Conservation*, **15**, 1859–1871.
- Collen, B., Whitton, F., Dyer, E.E., Baillie, J.E.M., Cumberlidge, N., Darwall, W.R.T., Pollock, C., Richman, N.I., Soulsby, A.-M. & Böhm, M. (2014) Global patterns of freshwater species diversity, threat and endemism. *Global Ecology and Biogeography*, **23**, 40–51.
- Cooper, N., Bielby, J., Thomas, G.H. & Purvis, A. (2008) Macroecology and extinction risk correlates of frogs. *Global Ecology and Biogeography*, **17**, 211–221.
- Cushman, S.A. (2006) Effects of habitat loss and fragmentation on amphibians: a review and prospectus. *Biological Conservation*, **128**, 231–240.
- Cutler, D.R., Edwards, T.C., Beard, K.H., Cutler, A., Hess, K.T., Gibson, J. & Lawler, J.J. (2007) Random forests for classification in ecology. *Ecology*, **88**, 2783–2792.
- Davidson, A.D., Hamilton, M.J., Boyer, A.G., Brown, J.H. & Ceballos, G. (2009) Multiple ecological pathways to extinction in mammals. *Proceedings of the National Academy of Sciences USA*, **106**, 10702–10705.
- Dubois, A. (2005) Developmental pathway, speciation and supraspecific taxonomy in amphibians, 1: why are there so many frog species in Sri Lanka? *Alytes*, **22**, 19–37.
- ESRI (2009) *ArcGIS 9.3.1*. Environmental Systems Research Institute, Redlands, CA.
- Foley, J.A., DeFries, R., Asner, G.P., Barford, C., Bonan, G., Carpenter, S.R., Chapin, F.S., Coe, M.T., Daily, G.C., Gibbs, H.K., Helkowski, J.H., Holloway, T., Howard, E.A., Kucharik, C.J., Monfreda, C., Patz, J.A., Prentice, I.C., Ramankutty, N. & Snyder, P.K. (2005) Global consequences of land use. *Science*, **309**, 570–574.
- Fritz, S.A., Bininda-Emonds, O.R.P. & Purvis, A. (2009) Geographical variation in predictors of mammalian extinction risk: big is bad, but only in the tropics. *Ecology Letters*, **12**, 538–549.
- Froese, R. & Pauly, D. (2012) *Fishbase*. Available at: www.fishbase.org (accessed 12 June 2012).
- Frost, D.R. (2011) *Amphibian species of the world: an online reference. Version 5.5 (31 January, 2011)*. Available at: <http://research.amnh.org/vz/herpetology/amphibia/> (accessed 8 November 2012).
- Gardner, T.A., Barlow, J., Chazdon, R., Ewers, R.M., Harvey, C.A., Peres, C.A. & Sodhi, N.S. (2009) Prospects for tropical forest biodiversity in a human-modified world. *Ecology Letters*, **12**, 561–582.
- Glaw, F. & Köhler, J. (1998) Amphibian species diversity exceeds that of mammals. *Herpetological Review*, **29**, 11–12.
- 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**, 516–519.
- Hoffmann, M., Hilton-Taylor, C., Angulo, A. *et al.* (2010) The impact of conservation on the status of the World's vertebrates. *Science*, **330**, 1503–1509.
- Hooper, D.U., Chapin Iii, F.S., Ewel, J.J., Hector, A., Inchausti, P., Lavorel, S., Lawton, J.H., Lodge, D.M., Loreau, M., Naeem, S., Schmid, B., Setälä, H., Symstad, A.J., Vandermeer, J. & Wardle, D.A. (2005) Effects of biodiversity on ecosystem functioning: a consensus of current knowledge. *Ecological Monographs*, **75**, 3–35.
- IUCN (2008) *An analysis of amphibians on the 2008 IUCN Red List*. Available at: <http://www.iucnredlist.org/initiatives/amphibians> (accessed 7 January 2012).
- IUCN (2011) *IUCN Red List of threatened species. Version 2011.2*. Available at: <http://www.iucnredlist.org> (accessed 10 January 2012).
- Kiesecker, J.M., Blaustein, A.R. & Belden, L.K. (2001) Complex causes of amphibian population declines. *Nature*, **410**, 681–684.
- Koh, L.P. & Sodhi, N.S. (2010) Conserving Southeast Asia's imperiled biodiversity: scientific, management, and policy challenges. *Biodiversity and Conservation*, **19**, 913–917.
- Köhler, J., Vietes, D.R., Bonnet, R.M., Garcia, F.H., Glaw, F., Steinke, D. & Vences, M. (2005) New amphibians and global conservation: a boost in species discoveries in a highly endangered vertebrate group. *BioScience*, **55**, 693–696.
- Latifi, H. & Koch, B. (2012) Evaluation of most similar neighbour and random forest methods for imputing forest inventory variables using data from target and auxiliary

- stands. *International Journal of Remote Sensing*, **33**, 6668–6694.
- Laurance, W.F., Croes, B.M., Guissouegou, N., Buij, R., Dethier, M. & Alonso, A. (2008) Impacts of roads, hunting, and habitat alteration on nocturnal mammals in African rainforests. *Conservation Biology*, **22**, 721–732.
- Lee, T.M. & Jetz, W. (2011) Unravelling the structure of species extinction risk for predictive conservation science. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **278**, 1329–1338.
- Lees, A.C. & Peres, C.A. (2008) Avian life-history determinants of local extinction risk in a hyper-fragmented neotropical forest landscape. *Animal Conservation*, **11**, 128–137.
- Liaw, A. & Wiener, M. (2002) Classification and regression by randomForest. *R News*, **2**, 18–22.
- Meegaskumbura, M., Manamendra-Arachchi, K., Bowatte, G. & Meegaskumbura, S. (2012) Rediscovery of *Pseudophilautus semiruber*, a diminutive shrub frog (Rhacophoridae: Pseudophilautus) from Sri Lanka. *Zootaxa*, **3229**, 58–68.
- Milner, J.M., Nilsen, E.B. & Andreassen, H.P. (2007) Demographic side effects of selective hunting in ungulates and carnivores: review. *Conservation Biology*, **21**, 36–47.
- Morais, A.R., Braga, R.T., Bastos, R.P. & Brito, D. (2012) A comparative analysis of global, national, and state red lists for threatened amphibians in Brazil. *Biodiversity and Conservation*, **21**, 2633–2640.
- Morais, A.R., Siqueira, M.N., Lemes, P., Maciel, N.M., De Marco, P., Jr & Brito, D. (2013) Unraveling the conservation status of Data Deficient species. *Biological Conservation*, **166**, 98–102.
- Murray, K.A., Rosauer, D., McCallum, H. & Skerratt, L.F. (2011) Integrating species traits with extrinsic threats: closing the gap between predicting and preventing species declines. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **278**, 1515–1523.
- Musters, C.J.M., Kalkman, V. & van Strien, A. (2013) Predicting rarity and decline in animals, plants, and mushrooms based on species attributes and indicator groups. *Ecology and Evolution*, **3**, 3401–3414.
- Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A.B. & Kent, J. (2001) Biodiversity hotspots for conservation priorities. *Nature*, **403**, 853–858.
- Owens, I.P.F. & Bennett, P.M. (2000) Ecological basis of extinction risk in birds: habitat loss versus human persecution and introduced predators. *Proceedings of the National Academy of Sciences USA*, **97**, 12144–12148.
- Pantanowitz, A. & Marwala, T. (2009) Evaluating the impact of missing data imputation. *Proceedings of the 5th International Conference on Advanced Data Mining and Applications* (ed. by R. Huang, Q. Yang, J. Pei, J. Gama, X. Meng and X. Li), pp. 577–586. Springer-Verlag, Beijing, China.
- Possingham, H.P., Andelman, S.J., Burgman, M.A., Medellín, R.A., Master, L.L. & Keith, D.A. (2002) Limits to the use of threatened species lists. *Trends in Ecology & Evolution*, **17**, 503–507.
- Pounds, J.A., Bustamante, M.R., Coloma, L.A., Consuegra, J.A., Fogden, M.P.L., Foster, P.N., La Marca, E., Masters, K.L., Merino-Viteri, A., Puschendorf, R., Ron, S.R., Sánchez-Azofeifa, G.A., Still, C.J. & Young, B.E. (2006) Widespread amphibian extinctions from epidemic disease driven by global warming. *Nature*, **439**, 161–167.
- R Development Core Team (2011) *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Robnik-Sikonja, M. (2004) Improving random forests. *Machine learning, ECML 2004 Proceedings* (ed. by J.F. Boulicaut), pp. 359–370. Springer, Berlin.
- Rodrigues, A.S.L., Pilgrim, J.D., Lamoreux, J.F., Hoffmann, M. & Brooks, T.M. (2006) The value of the IUCN Red List for conservation. *Trends in Ecology and Evolution*, **21**, 71–76.
- Rowley, J., Brown, R., Bain, R., Kusri, M., Inger, R., Stuart, B., Wogan, G., Thy, N., Chan-Ard, T., Trung, C.T., Diesmos, A., Iskandar, D.T., Lau, M., Ming, L.T., Makchai, S., Truong, N.Q. & Phimmachak, S. (2010) Impending conservation crisis for Southeast Asian amphibians. *Biology Letters*, **6**, 336–338.
- Schloegel, L.M., Daszak, P., Cunningham, A.A., Speare, R. & Hill, B. (2010) Two amphibian diseases, chytridiomycosis and ranaviral disease, are now globally notifiable to the World Organization for Animal Health (OIE): an assessment. *Diseases of Aquatic Organisms*, **92**, 101–108.
- Sodhi, N.S. & Brook, B.W. (2008) Fragile Southeast Asian biotas. *Biological Conservation*, **141**, 883–884.
- Sodhi, N.S., Koh, L.P., Brook, B.W. & Ng, P.K.L. (2004) Southeast Asian biodiversity: an impending disaster. *Trends in Ecology & Evolution*, **19**, 654–660.
- 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**, e1636.
- Sodhi, N.S., Posa, M.R.C., Lee, T.M., Bickford, D., Koh, L.P. & Brook, B.W. (2010) The state and conservation of Southeast Asian biodiversity. *Biodiversity and Conservation*, **19**, 317–328.
- 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**, 1783–1786.
- Stuart, B.L., Rhodin, A.G.J., Grismer, L.L. & Hansel, T. (2006) Scientific description can imperil species. *Science*, **312**, 1137.
- Warkentin, I.G., Bickford, D., Sodhi, N.S. & Bradshaw, C.J.A. (2009) Eating frogs to extinction. *Conservation Biology*, **23**, 1056–1059.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Text S1 Variable summary.

Table S1 Summary of model variables.

Table S2 Database completeness.

Table S3 Summary of missing data.

Table S4 Model accuracy.

Table S5 Model confusion matrix.

Figure S1 Distribution of Data Deficient amphibians.

Data S1 Predicted status of Data Deficient amphibians.

BIOSKETCHES

Sam D. Howard is a PhD student at NUS, who is broadly interested in the spatial ecology and community composition of amphibians and reptiles in South-East Asia.

David P. Bickford is an evolutionary ecologist, a conservation biologist and a tropical herpetologist. He leads a laboratory with focus on applications for conservation, approaching questions in an organismal and evolutionary perspective.

Author contributions: S.D.H. designed the study and performed statistical and GIS analyses. S.D.H. wrote the manuscript. D.P.B. provided data and edited the manuscript.

Editor: Simon Ferrier