

#### Review

# Bridging the research-implementation gap in IUCN Red List assessments

```
Victor Cazalis , 1,2,* Moreno Di Marco , 3 Stuart H.M. Butchart , 4,5 H. Reşit Akçakaya , 6,7 Manuela González-Suárez , 8 Carsten Meyer , 1,9,10 Viola Clausnitzer , 11 Monika Böhm , 12 Alexander Zizka , 1,13 Pedro Cardoso , 14 Aafke M. Schipper , 15,16 Steven P. Bachman , 17 Bruce E. Young , 18 Michael Hoffmann , 19 Ana Benítez-López , 20 Pablo M. Lucas , 3 Nathalie Pettorelli , 21 Guillaume Patoine , 1,10 Michela Pacifici , 22 Theresa Jörger-Hickfang , 23 Thomas M. Brooks , 24,25,26 Carlo Rondinini , 22 Samantha L.L. Hill , 27 Piero Visconti , 28 and Luca Santini , 3
```

The International Union for Conservation of Nature (IUCN) Red List of Threatened Species is central in biodiversity conservation, but insufficient resources hamper its long-term growth, updating, and consistency. Models or automated calculations can alleviate those challenges by providing standardised estimates required for assessments, or prioritising species for (re-)assessments. However, while numerous scientific papers have proposed such methods, few have been integrated into assessment practice, highlighting a critical research-implementation gap. We believe this gap can be bridged by fostering communication and collaboration between academic researchers and Red List practitioners, and by developing and maintaining user-friendly platforms to automate application of the methods. We propose that developing methods better encompassing Red List criteria, systems, and drivers is the next priority to support the Red List.

### Major challenges for the IUCN Red List

The IUCN Red List of Threatened Species (hereafter 'Red List') provides assessments of extinction risk for >130 000 species of animals, fungi, and plants (Red List version 2021.2)<sup>i</sup>. These assessments are pivotal to inform conservation action, target resources, and monitor global biodiversity trends and conservation effectiveness [1–6]. The Red List also informs international policies and reports [e.g., the Convention on Biological Diversity (CBD), the Intergovernmental Science–Policy Platform on Biodiversity and Ecosystem Services (IPBES), and the Convention on International Trade in Endangered Species (CITES)] by providing information and underpinning analyses on species' status and trends, distributions, threats, and conservation actions. The Red List uses a set of standard quantitative **Red List criteria** (see Glossary) relating to species' population size, trend, and distribution, that are applied by **assessors** to assign species to a **Red List category** of extinction risk [7,8].

Despite its influence, the Red List operates with a largely insufficient budget and staff [9,10], resulting in four major challenges that jeopardize its breadth and currency in the long term. First, assessments are concentrated on vertebrate species [11–13], with few for invertebrates and plants relative to the number of described species and very few for fungi (Figure 1A). This taxonomic imbalance is being slowly reduced by the ongoing expansion of the Red List in accordance with an agreed strategic plan (Figure 1A; [14]). Second, 14% of assessed species ( $n = 19\,394$ ) are classified as Data Deficient due to insufficient information available to apply Red List criteria (Figure 1B), which introduces uncertainty in estimated proportions of threatened species and may preclude some species from receiving appropriate conservation efforts [15–17]. Third, while species should

#### Highlights

The IUCN Red List of Threatened Species plays a central role in monitoring biodiversity and informing conservation actions.

To best inform conservation, the Red List must be frequently updated, become more taxonomically and geographically representative, and be consistent within and among taxonomic groups, but this is hampered by limited resources.

A variety of models and automated calculations has been proposed in the literature to support Red List assessments, for instance using citizen science or remote-sensing data to predict extinction risk.

We highlight a major researchimplementation gap in the application of these methods, which could be bridged by providing assessors with easy access to the most relevant tools, hands-on training, and strengthening communication.

Further efforts are needed to develop relevant methods to prioritise assessments or better predict extinction risk.

<sup>1</sup>German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Puschstr. 4, D-04103 Leipzig, Germany <sup>2</sup>Leipzig University, Ritterstraße 26, 04109 Leipzig, Germany <sup>3</sup>Department of Biology and

04109 Leipzig, Germany

<sup>3</sup>Department of Biology and
Biotechnologies 'Charles Darwin',
Sapienza Università di Roma,
Rome, Italy





be reassessed at least every 10 years (Red List guidelines version 14)<sup>ii</sup>, 18% of assessments (n = 24 764) are currently outdated (Figure 1C). About 2100 species were last assessed 25 years ago, of which more than half are listed as threatened (Figure 1D). Fourth, Red List assessments are conducted inconsistently across and within taxonomic groups [11,12,18], partly because of heterogeneity in available data among species, but also because of variation in the assessment process and criteria application. The Red List guidelines, which aim at reducing the latter by providing detailed information on how to apply the criteria, have expanded and evolved to further clarify the calculation of **Red List parameters** and the resulting assignment of categories (see examples in Table 1), but substantial discrepancies among taxa or regions remain.

In the last decade, many studies have proposed methods to capitalise on the increasing availability of ecological data and remote-sensing products to address the aforementioned challenges, by enabling faster, more rigorous, and more consistent assessments (e.g., [19,20]). In particular, relevant data, tools, and models have been proposed to standardise the estimation of Red List parameters (e.g., Extent of Occurrence or population trends) or predict species' Red List categories. However, while many methods have been published, very few have been implemented in practice [21].

Here, we systematically reviewed recently published methods that aim either at identifying correlates of extinction risk, or at predicting species' extinction risk categories for groups of species using modelling or automated calculation (considering papers published between 2001 and June 2021; see the supplementary information online). We then evaluated their utility from a practical perspective and discussed the main barriers to their uptake in Red Listing. Finally, we suggested how to bridge this important research-implementation gap, and highlighted potential future research directions.

#### Published methods to predict Red List categories

#### Four main objectives of published studies

Of the 98 studies identified in our review, 46% aimed at predicting Red List categories, and we identified three related objectives depending on the species group targeted (Figure 2). The first objective aimed at prioritising or informing first assessments by assigning plausible Red List categories to unassessed species (e.g., [22]; 13% of studies). The second aimed at resolving Data Deficient species' status (e.g., [17]; 11% of studies), by providing information that may enable assigning data sufficient categories to species with no taxonomy uncertainty [15,16]. The third aimed at prioritising or informing reassessments, by highlighting species likely to be misclassified (e.g., [20]; 22% of studies), sometimes also including Data Deficient species.

Additionally, 54% of studies aimed at understanding correlates of extinction risk using Red List categories as a proxy for risk (Figure 2). These studies showed, for instance, that mammals with high weaning age, small geographic range size, and high human population density within their geographic range were particularly likely to be categorised as threatened [23]. We define this objective as fundamental, in the sense that it does not aim to assist Red List assessments directly, but rather contributes to understanding vulnerability to extinction, which in turn may guide the development of predictive approaches.

#### Two main approaches to predict Red List categories

To meet the objectives mentioned previously, studies have relied on two main approaches (Figure 2): (i) the modelling or automated calculation of Red List parameters, then used to apply Red List criteria (criteria-explicit); or (ii) using correlates of extinction risk to predict Red List categories with no explicit use of criteria (category-predictive).

<sup>4</sup>BirdLife International, David Attenborough Building, Pembroke Street, Cambridge CB2 3QZ, UK <sup>5</sup>Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, UK <sup>6</sup>Department of Ecology and Evolution, Stony Brook University, New York, USA <sup>7</sup>IUCN Species Survival Commission (SSC), Gland, Switzerland 8Ecology and Evolutionary Biology, School of Biological Sciences, University of Reading, Whiteknights, Reading, UK <sup>9</sup>Institute of Geosciences and Geography, Martin Luther University Halle-Wittenberg, Halle (Saale), Germany <sup>10</sup>Institute of Biology, Leipzig University, Leipzig, Germany <sup>11</sup>Senckenberg Research Institute, Goerlitz, Germany 12 Global Center for Species Survival, Indianapolis Zoological Society, Indianapolis, USA <sup>13</sup>Department of Biology, Philipps-University Marburg, Karl-von-Frisch Str. 8, D-35043 Marburg, Germany <sup>14</sup>Laboratory for Integrative Biodiversity Research (LIBRe), Finnish Museum of Natural History Luomus, University of Helsinki, Helsinki, Finland <sup>15</sup>Department of Environmental Science, Radboud Institute for Biological and Environmental Sciences (RIBES). Radboud University, Nijmegen, The Netherlands <sup>16</sup>PBL Netherlands Environmental Assessment Agency, The Hague, The Netherlands <sup>17</sup>Conservation Assessment and Analysis. Royal Botanic Gardens. Kew, Richmond, UK <sup>18</sup>NatureServe, Arlington, VA, USA <sup>19</sup>Conservation and Policy, Zoological Society of London, London, NW1 4RY, <sup>20</sup>Integrative Ecology Group, Estación Biológica de Doñana (EBD-CSIC), Sevilla, Spain <sup>21</sup>Zoological Society of London, Institute of Zoology, London, UK <sup>22</sup>Global Mammal Assessment Programme, Department of Biology and Biotechnologies 'Charles Darwin', Sapienza Università di Roma, Rome, Italy <sup>23</sup>German Centre for Integrative Biodiversity Research Halle-Jena-Leipzig (iDiv), Martin Luther University of Halle-Wittenberg, Halle (Saale), Germany <sup>24</sup>IUCN, Gland, Switzerland <sup>25</sup>World Agroforestry Center (ICRAF), University of The Philippines Los Baños, Los Baños, Laguna, Philippines

<sup>26</sup>Institute for Marine and Antarctic Studies, University of Tasmania, Hobart,

Tasmania. Australia



#### Criteria-explicit approach

Criteria-explicit methods mirror the process of assessments by applying Red List criteria based on Red List parameters that have been automatically calculated from data such as species occurrences, species habitat requirements, and remote-sensing products (n = 25; Figure 2). For example, species occurrence data can be used to estimate Extent of Occurrence and Area of Occupancy (e.g., [24]), and several platforms and R packages have been developed to calculate these parameters automatically (e.g., GeoCAT and rCAT [25]; red [26]; ConR [27]; redlistr [28]; rapidLC [29]). These methods are particularly useful if species' geographic distributions have not been mapped although substantial occurrence data exist, and are thus more often used for plant and invertebrate groups. Similarly, abundance data can allow estimating population trends [30], although extensive temporal data are required.

Other studies use habitat and geographic data, often derived from remote-sensing products, to estimate Red List parameters (Figure 2). For example, combining current land cover and digital elevation maps with data on species' habitat preferences and elevational limits allows mapping an estimate of the Area of Habitat of species. This in turn can be used to calculate upper bounds of the Extent of Occurrence and Area of Occupancy [31], and inform application of criteria B and D2 [32,33]. Similarly, land cover time series can be used to estimate past or future trends in suitable habitats within species range, which enables inferring population trends and application of criteria A, B, and C (e.g., [34–36]). Most studies focus on only one or two Red List criteria rather than the full spectrum (Figure 2), although two studies applied each of the criteria, A to D; one focused on past data [20] and the other on future projections [34].

It may perhaps be surprising that criterion E, which is related to quantitative estimates of extinction probability, is rarely considered in these studies. This criterion is also rarely used in assessments (currently only used for four species, always in combination with another criterion). This scarce use of Criterion E results from the large amount of information required (e.g., demographic data or patterns of occupancy used to perform Population Viability Analyses), which is not available for a vast majority of species. This may also explain the lack of relevant multispecies studies targeting Criterion E. We found one single study attempting to apply criterion E on a large set of species [89], with extinction probability estimated from very limited information (generation length and past transition between categories), thus being unreliable at the species level.

#### Category-predictive approach

Category-predictive methods rely on comparative extinction risk analyses using statistical models that link Red List categories with other species-level information (see later; n = 73 studies; Figure 2). These statistical relationships are then used to identify the main drivers of risk (e.g., [37,38]) and/or to predict Red List categories of unassessed species (e.g., [39]), Data Deficient species (e.g., [17]), or species with outdated assessments (e.g., [40]). In addition to species-level predictions, these approaches have estimated and mapped proportions of threatened species for incompletely assessed taxa or regions [24,39].

Many species-level predictors have been used [41], the most common being biological traits (e.g., body mass, weaning age; 86% of studies), and range characteristics (often range size, sometimes insularity, or spatial configuration; 67%; Figure 2). Many studies also included predictors representing levels of human pressure within species' ranges (e.g., human footprint index, river fragmentation; 40%), which are important correlates of extinction risk [38,42]. Other predictors include conservation actions in place (e.g., proportion of species' range overlapping with protected areas; 4%), which may be important covariates of extinction risk [43–45]. Importantly, we found only nine studies using the threats listed in species Red List assessments

<sup>27</sup>United Nations Environment Programme World Conservation Monitoring Centre (UNEP-WCMC), Cambridge, UK

<sup>28</sup>Biodiversity, Ecology and Conservation Group, Biodiversity and Natural Resources Management Programme, International Institute for Applied Systems Analysis, Laxenburg, Austria

\*Correspondence: victor.cazalis@idiv.de (V. Cazalis).



as predictors (e.g., [37,46]), although these can modulate trait-extinction risk relationships (e.g., human consumption more strongly threatens large frogs whereas pet trade threatens small frogs [47]).

Two main types of models are used in this category-predictive approach: machine learning (e.g., Random Forest [39] or Neural Networks [48]) and statistical linear models (e.g., Generalised Linear Models [49]). Studies comparing their performance in predicting extinction risk are yet too scarce to provide clear guidance on which modelling method is best [50]. An important consideration when building these models is in how to define the extinction risk response variable. Risk can be binary (threatened versus nonthreatened; 43% of studies; e.g., [51]), include individual Red List categories (15%, e.g., [52]), or transforming them in a discrete quantitative variable (39%, e.g., [53] where LC = 1, NT = 2, etc.), or be described as the change in categories between two assessments (3%, e.g., [42]). The preferred option depends on the envisioned applications of the predicted Red List categories. For instance, binary threat predictions are often more accurate [54] and can be sufficiently detailed for a first sorting of species likely to be threatened [29], whereas category-specific models may be needed to inform and prioritise reassessments. When category-specific predictions are needed, using a discrete quantitative variable requires making assumptions about the distance between categories that are generally untested. This could be resolved by using Cumulative Link Mixed Models, which deal with multinomial ordered variables [52,55].

Many studies investigating range size as a correlate of extinction risk have excluded assessments made under criterion B as they could introduce circularity (e.g., because range size is highly correlated with Extent of Occurrence used in criterion B1; see [41,55]). This exclusion is necessary when the objective is fundamental (i.e., to understand if range size correlates with extinction risk), but not necessarily required when the objective is predicting species Red List category.

#### System and taxonomic biases

Our review revealed biases in extinction risk research across taxa and systems, with 73% of studies focusing only on terrestrial species, versus 11% on marine, and 3% on strictly freshwater species (rare examples include [53,56]); 13% cover several systems. Additionally, only one criteria-explicit study focused specifically on marine species and none on freshwater species (Figure 2), possibly because it is less straightforward to derive binary maps of suitable habitat from remote-sensing products for these systems compared to the terrestrial system. Marine and freshwater species, however, are facing particular threats and thus need specific data and methods (e.g., to estimate impacts of dam-induced fragmentation on Area of Habitat; [53]). Studies were also strongly biased towards tetrapod species (74% of studies), while they would be particularly valuable for groups that are less known, such as fishes, invertebrates, plants, and fungi.

#### From research to implementation

The limited uptake of methods developed to support Red List assessments is striking. Perhaps the most widely used tools are platforms and packages that facilitate the use of criterion B from occurrence data, such as GeoCAT [25], which have been cited in 8921 assessments as of early June 2021, or red [26]. Additionally, some studies have been conducted in collaboration with groups undertaking Red List assessments, or have been communicated directly to assessors [20,33,35,36], and have thus informed actual assessments. So far, however, most studies remain research exercises.

#### Overcoming barriers

The important research-implementation gap can be broadly attributed to a lack of communication between extinction risk researchers and Red List practitioners [21]. From the research

#### Glossarv

Assessor: an appointed expert, often a volunteer, who applies the IUCN Red List categories and criteria following associated guidelines, using all relevant data to assess the taxon appropriately. and ensures that the assessment has the required supporting information. Red List categories: ordinal set of extinction risk classes used by the IUCN Red List, including two non-threatened categories [Least Concern (LC) and Near Threatened (NT)], three threatened categories (Vu). Endangered (EN), and Critically Endangered (CR)], and two extinct categories [Extinct in the Wild (EW) and Extinct (EX)]. When data are insufficient to assign a species to one of these categories, it is classified as Data Deficient (DD). Species that have not been assessed yet are classified as Not Evaluated (NE), A subset of Critically Endangered (CR) species are tagged as Possibly Extinct [CR(PE)] or Possibly Extinct in the Wild [CR(PEW)]. Red List criteria: set of five criteria, and

nested subcriteria, associated with quantitative thresholds used to assign Red List categories. These criteria relate to A: population size reduction in the past (A1 and A2), future (A3), or both (A4); B: small geographic range, either in the form of Extent of Occurrence (B1) or Area of Occupancy (B2), combined with severe fragmentation, and/or continuing decline in population, distribution, or habitat quality, and/or extreme fluctuations; C: small population size and decline; D: very small or restricted population; E: quantitative analysis.

Red List quidelines: public document produced by the IUCN Red List Standards and Petitions Committee detailing how to apply the IUCN Red List criteria to assign categories.

Red List parameters: estimates which are compared with the quantitative thresholds listed in the Red List criteria to classify species into Red List categories. For instance, an ongoing reduction in population size of ≥ 30% over the last 10 years (or three generations, whichever is the longer) qualifies a species as Vulnerable (VU) under criterion A2. In this example, the reduction in species' population is the parameter compared with the 30% threshold to apply the criterion. Red List Unit: technical unit working for the IUCN Global Species Program.



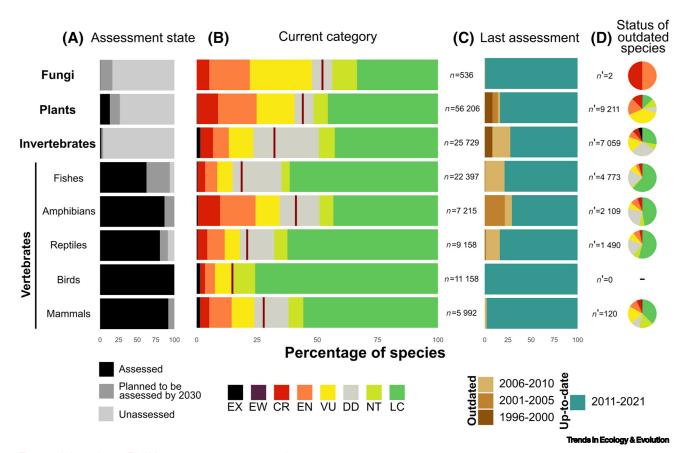


Figure 1. Information on Red List assessments per taxonomic group. (A) Proportion of described species that are currently assessed, planned to be assessed by 2030 (calculated from [14]), or unassessed. (B) Proportion of assessed species in each Red List category (coloured bars) and proportion of these that are threatened (red line), assuming Data Deficient species are threatened in the same proportion as data-sufficient species. (C) Proportion of assessed species with an outdated assessment (year of last assessment coloured in 5-year classes up to 2010; more detail in Figure S1 in the supplemental information online). (D) Distribution of current Red List categories of outdated assessments (colours as in B), n refers to the number of assessed species per group, n' refers to the number of species with outdated assessments. Data were extracted from the version 2021.2 from the Red List, using the rredlist package.

side, implementation is hindered by misunderstandings or misapplications of Red List criteria in the proposed methods, mismatches between researchers' interests and assessors' needs, or because developed methods do not provide the outputs needed by assessors [57] (Box 1). This may be partly due to researchers being unclear about the most appropriate entry points in the Red List system to discuss and propose change. On the Red List side, assessors may not be able to use potentially relevant tools if these require detailed input data, substantial time, or advanced technical skills and capacity to apply (Box 1). Additionally, some tools have been implemented and used by assessors, but because of a lack of funding, are not being maintained (e.g., the Freshwater Mapping Application used in many assessments had no funding to support development and maintenance at the time of writing).

These barriers could be mitigated in various ways. First, the best means of resolving poor communication between researchers and practitioners is by involving Red List stakeholders early in the development of new approaches and methods, to ensure effective orientation of research efforts and avoid misunderstanding or misapplication of the Red List categories and criteria, or of assessors' needs and constraints [58]. This could include members of Red List Authorities, the



Table 1. Examples of changes made in the Red List guidelines over the last two decades to strengthen consistency and riggur of Red List assessments, with year of inclusion in the Red List guidelines

Red List criterion	Change made in guidelines	Year	Refs
A-E	Using fuzzy arithmetic to propagate data uncertainties and identify the range of plausible Red List categories	2001	[73]
A, C1	Extracting species generation length from databases of calculated and predicted generation lengths for entire taxonomic groups (mammals and birds)	2003, 2011	[74,75]
B2	Measuring Area of Occupancy (AOO) at the reference scale of 2 $\times$ 2 km $$	2003	[76]
B1	Measuring the Extent of Occurrence (EOO) as the area of the minimum convex polygon	2006	[77]
A3	Using ecological niche models and climate projections outputs to infer future reductions resulting from climate change	2010	[78,79]
A2	Calculating 3-generation reduction of species with large fluctuations using statistical models fitted to longer time series	2011	[80,81]
В	Calculating upper bounds of AOO and EOO based on habitat maps and Area of Habitat	2014	[31]
Red List category			
DD	Differentiating (and flagging) three types of Data Deficient	2008	[15]
EX, CR(PE)	Defining (and flagging) species likely but not yet confirmed to be extinct as 'Critically Endangered (Possibly Extinct)' CR(PE).	2008	[82]
EX, CR(PE)	Inferring that a species is extinct based on threats and time series of records and surveys	2019	[83,84]

<sup>&</sup>lt;sup>a</sup>Multiple years indicate stepwise implementation and references related to the issue [the reference may precede change in guidelines (e.g., if it suggested and provided rationale for such change), or follow it (e.g., if it tested or explained such change)]. Red List criteria and categories are detailed in the Glossary.

Red List Committee and its working groups, IUCN Red List Unit or IUCN Standards and Petitions Committee (noting that part of these members are also recognised experts in extinction risk research), or sending a request to the generic IUCN Red List email address when researchers cannot identify the correct entry point. Particular attention must be given to the ultimate outputs to ensure they are useful in practice. On this point, criteria-explicit methods which, by definition, estimate Red List parameters that can be directly used by assessors to apply Red List criteria, seem more useful than category-predictive methods. However, the latter could prove useful to designate priorities for species (re-)assessment (see section on future research directions).

Second, because of the heterogeneity in assessors' backgrounds, uptake of any new method requires easy use. This can be achieved by releasing methods through user-friendly online platforms, such as Shiny Apps (e.g., [29]), and ensuring their long-term maintenance and update with new data and methods. At the same time, any information provided should come with high transparency (so that assessors can understand basic assumptions and limitations of underlying methods), with explicit uncertainty bounds, and be open source. In addition, platforms could benefit from allowing assessors to adjust some methodological choices (e.g., selecting variables to include in a given model) based on their expertise. However, this may come at the expense of consistency and may increase the risk of cherry-picking (e.g., assessors may be tempted to adjust methods to meet the output they expected).

Finally, these platforms should be promoted to assessors, provided with adequate guidance and training (e.g., through webinars, workshops, documentation, or video tutorials), and connected with IUCN database [the Species Information Service (SIS)]. From a longer-term perspective, it



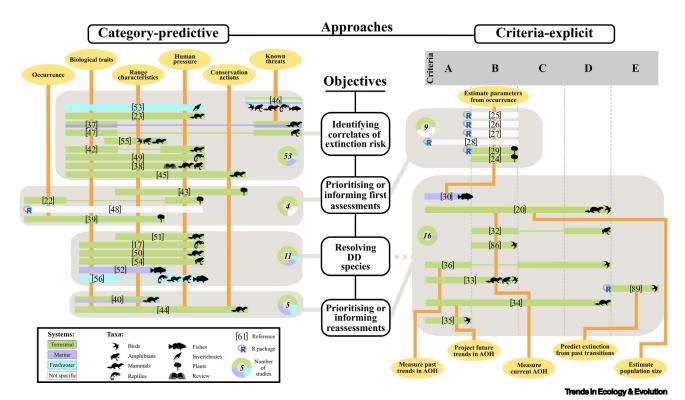


Figure 2. Graphical summary of studies reviewed. Presented are the two approaches and the four objectives of studies developing modelling or automated calculation methods to predict Red List categories. All studies cited in the main text are reported in the figure in brackets (full references in Figure S2 in the supplemental information online); the total number of studies found in the systematic review per approach and objective is given in the doughnut plots. Colours denote the system investigated, with freshwater designating only fully aquatic freshwater species and 'not specific' for R packages that can be applied to any system. Yellow ellipses present the main types of variables used in the category-predictive approach and the main methods used in the criteria-explicit approach. Thin horizontal lines are used to illustrate studies belonging to several adjacent columns (e.g., including criteria B and D, but not C for [32]). Red List criteria are detailed in the Glossary. Grey boxes encompass studies that share the same objective and approach. The broken grey line indicates that only some studies in the grey box share the objective. See references [17,20,22–30,32–40,42–56,86,89]. Abbreviations: AOH, Area of Habitat; DD, Data Deficient.

is also important to enable assessors to provide feedback on these platforms to inform future development, and to track their use (e.g., through citations in assessments).

#### Future research directions

In addition to making developed methods accessible to assessors, further research is needed to create methods that (i) better support the assignment of Red List categories; and (ii) help prioritise assessments and data collection. Before implementation, all methods have to be rigorously validated to measure their performance (Box 2).

#### Supporting assignment of Red List categories

Considering the diversity of threats: with most published methods targeting terrestrial habitat loss (especially in the criteria-explicit approach; Figure 2), it is important to develop methods that focus on the impact of other threats on species extinction risk (e.g., harvesting, pollution, diseases, or invasive species), including those specific to freshwater and marine species (e.g., dams, water pollution, or overfishing). In particular, while climate change is threatening >10 000 species<sup>i</sup> and can significantly increase extinction risk [59], estimating its impact consistently across species is complex<sup>ii</sup> [90]. We need tools providing assessors with species' exposure to past and future climate change (e.g., change in climatic envelope, sea-level rise,



#### Box 1. Main barriers to the implementation of recent methods to predict Red List categories

Misunderstanding of Red List criteria: in many publications, the Red List guidelines are ignored or misinterpreted [79,85], rendering outputs unhelpful for Red List assessments. For instance, considerable confusion has arisen over the interpretation of the slightly ambiguous language around the Extent of Occurrence metric (e.g., [86]), despite attempts to clarify how this should be calculated [31.77].

Divergent interests: there may be differences between what is needed by Red List assessors and what is appealing to researchers. While assessors need tools that give them easy access to basic information (e.g., deforestation rates within species ranges) or readily applicable estimates of Red List parameters, researchers may be more interested in developing sophisticated modelling methods, to increase the novelty of potential publications.

Misaligned output: methods may sometimes output parameters in formats that are not directly usable in Red List assessments. For instance, a model predicting species' Red List categories cannot be used by assessors if it fails to output the specific parameters that assessors must provide to justify categories (e.g., typical of the category-predictive approach).

Lack of data: methods that require extensive species-specific data (e.g., occurrences across range [25] or life-history traits across taxa [50]) cannot be applied to all taxa.

Insufficient skills, capacity, or time: Red List assessors vary in their ability to use technological tools (e.g., GIS or R scripts) and may lack the necessary background, skills, and time to learn how to use newly developed methods if they are not easy to apply (e.g., [20]). For example, the success of GeoCAT [25] is likely due to its user-friendly interface. Specific training on how to use newly developed tools (e.g., courses, tutorials, and fora), is very rarely offered.

Disconnect with the Red List database: all Red List assessments are conducted in the IUCN's online database [the Species Information Service (SIS)]. Uptake of new methods and approaches would be greatly increased if outputs, such as Red List parameters, could readily be integrated into SIS (e.g., through the existing SIS Connect tool).

frequency of extreme climatic events, or ocean acidification), and the ability to integrate this knowledge with information on species' sensitivity to climate change [60-62] in accordance with Red List guidelines [90].

#### Box 2. Best practices to validate methods predicting Red List categories

Model validation is necessary to assess the ability of models to correctly predict species' Red List categories.

In the criteria-explicit approach, validation simply requires comparison of predicted categories with the actual categories from published assessments.

In the category-predictive approach, three main validation methods can be undertaken: (i) temporal block validation is the most recommended method, if applicable (i.e., species have been assessed at least twice), where models are trained on Red List categories from past assessments and validated against current assessments. This is relevant only if changes in categories are 'genuine' (i.e., not due to improved knowledge or other non-genuine reasons, this is specified in Red List data); (ii) phylogenetical or spatial block validation, is the most recommended method when temporal block validation is not applicable, where each independent taxon or region is separately set aside (i.e., not used in model training) and used for validation (e.g., [49]); and (iii) other split sample validation methods randomly split the dataset into training and testing sets (e.g., [50]). This is the least recommended, as accuracy can be overestimated due to the autocorrelation in training and testing samples [87].

For both approaches, we advise systematically reporting confusion matrices and measures of accuracy (i.e., proportion of species correctly categorised), sensitivity (proportion of threatened species correctly categorised), and specificity (proportion of non-threatened species correctly categorised), as they provide key and complementary information [88]. Models with high sensitivity are particularly useful to identify species likely to be threatened, while models with high specificity can rule out species unlikely to be threatened. A model with intermediate specificity and sensitivity is less informative. Additionally, exploring how geographically/taxonomically consistent model performance is may provide important insights on model limitations.

For both approaches, we advise subsetting the species used for validation, keeping only the most accurate assessments, to avoid underestimating the accuracy of the developed methods. We suggest selecting species: (i) with up-to-date assessments; (ii) threatened by processes accounted for in the modelling (e.g., species threatened by habitat loss when validating methods based on Area of Habitat); and (iii) with high certainty in Red List category, although in practice it may be difficult to identify such assessments.



Facilitating the application of criterion E: a wider use of criterion E would have two main advantages: (i) direct incorporation of quantitative analyses in Red List assessments; and (ii) explicit consideration of longer time frames than all other criteria (up to 100 years in the future, regardless of generation length). Methods may build on allometry-driven parameters (e.g., [63]) and population density estimates [64] to inform extinction risk simulations on entire groups of species. Extinction probability could also be estimated by modelling the probability that a species' Area of Habitat disappears in the future, according to climate and land-use change projections<sup>ii</sup>.

Predicting the probability of meeting thresholds: in analogy with the category-predictive approach (i.e., linking extinction risk of multiple species to species-specific data such as biological traits or human pressure in the range), models could be developed to predict the probability of meeting the threshold for a given criterion (e.g., the probability that past population decline is  $\geq$  30% over 10 years), instead of the categories themselves. Such models would thus benefit from the power of multi-species comparisons inherent in category-predictive methods, but provide an output more likely to be useful to assessors.

Accounting for biotic dependencies: informing assessors on biotic dependencies between species (e.g., parasite—host, plant—pollinator, or plant—phytophagous relationships) can lead to better integration of associated coextinction risk in assessments [11], which could affect several thousands of species [65–67]. For instance, the population trend of Barrett's Plant-louse *Trioza barrettae* (an endemic bug from Australia) was estimated based on the population trend of its Critically Endangered and sole known host plant Brown's Banksia *Banksia brownii*, and the louse was consequently categorised as Critically Endangered.

Predicting down-listing: while previously mentioned methods can also identify species warranting down-listing to lower categories of threat, specific research efforts should focus on predicting positive population trends (considering for instance conservation actions undertaken) or range expansions. Such methods may later support assessments of the IUCN Green Status of Species [68,69].

#### Prioritising assessments or data collection

Prioritising first assessments: both category-predictive and criteria-explicit approaches can help prioritise assessments to optimise allocation of limited resources [10]. Specifically, for assessors or teams undertaking first-time assessments for large groups of species, these approaches can be used to help provide an initial indication of whether species are likely to be threatened (e.g., [39,43]) or Least Concern (and hence could be fast-tracked [29]).

Prioritising reassessments: given that reassessments rates are currently insufficient to provide updates every 10 years for most groups (Figure 1C), the identification of species most likely to have changed their category is also relevant [20,44]. Additionally, a period of 10 years between assessments may be too long to detect rapid changes in some species' status (e.g., the Mount Gorongosa Pygmy Chameleon, *Rhampholeon gorongosae*, Least Concern in 2014 was Endangered five years later following rapid habitat loss). Identifying which species are most likely to have changed in status since the previous assessment could inform targeted reassessments and thus help to keep the Red List up-to-date. Similarly, it would be useful to develop tools that flag Data Deficient species for which recent increases in data availability may allow application of Red List criteria (e.g., through accumulation of new information on citizen science platforms).

Prioritising data collection: methods that predict species or areas for which data collection would make the biggest difference for Red List assessments can deliver useful information to guide data collection.



For instance, Data Deficient species that are predicted as threatened by category-predictive methods may be prioritised for data collection [50]. Further, predicting where data collection may be the most valuable for conservation (e.g., species that could become data sufficient with few additional data, or regions where collecting contextual information would benefit many species) can also be useful to guide fieldwork efforts [15,70,71]. Synergies with the IUCN Species Monitoring Specialist Group, which aims to produce prioritized lists of existing species data gaps, would be beneficial.

#### Concluding remarks

The multiple approaches reviewed in this paper include some with significant potential to assist Red List assessments. Improved communication between researchers and the Red List community is required to develop the tools and outputs most relevant for assessors. Uptake also requires additional research to tackle key remaining methodological challenges (see Outstanding questions) and deliver practical tools. We believe that further development of such tools, and ensuring their long-term availability to assessors, could constitute an important milestone for the future of the Red List.

Importantly, the proposed methods will neither substitute nor reduce the role of assessors, but rather support them with appropriate and readily usable outputs and techniques. In doing so, these methods may help fast-track or prioritise assessments. However, it is important to note that they will not address the urgent need to increase Red List resources for targeted fieldwork, workshops, tool development, fora, and remunerated assessors.

Increasing resources and embracing new data and methods will enable the Red List to become more taxonomically and geographically representative, data sufficient, up-to-date and consistent, and thus remain the standard and authoritative source of information on species' extinction risk [10]. This is crucial to ensure that the Red List can best guide future conservation actions [1,2], and support accurate monitoring of the effectiveness of global conservation efforts under the post-2020 global biodiversity framework [5,72].

#### **Acknowledgments**

This paper is a joint effort of the working group sRedList kindly supported 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 positions for A.Z. and V.C. M.D.M. and L.S. acknowledge support from the MUR Rita Levi Montalcini program. C.M. acknowledges funding by the Volkswagen Foundation through a Freigeist Fellowship (A118199), and additional support by iDiv, funded by the German Research Foundation (DFG-FZT 118, 202548816). A.B.L. was supported by a Juan de la Cierva-Incorporación grant (IJCI-2017-31419) from the Spanish Ministry of Science, Innovation and Universities. P.C. is supported by Kone Foundation. M.B. and M.H. were supported by a generous grant from the Rufford Foundation. B.E.Y. acknowledges Chevron, Inc. The views expressed in this publication do not necessarily reflect those of IUCN.

#### **Declaration of interests**

No interests to declare.

#### Supplemental information

Supplemental information associated with this article can be found online at https://doi.org/10.1016/j.tree.2021.12.002.

#### Resources

www.iucnredlist.org.

iiwww.iucnredlist.org/documents/RedListGuidelines.pdf.

iihttps://CRAN.R-Project.org/package=rredlist.

whttps://cmsdocs.s3.amazonaws.com/documents/Response%20to%20Ocampo\_Penuela%20et%20al%202016\_ 15Dec2016.pdf

<sup>v</sup>www.iucnredlist.org/resources/summary-statistics

#### Outstanding questions

How can assessments of marine and freshwater species be better supported? Most existing databases and methods are strongly biased towards the terrestrial system. In particular, most criteria-explicit methods are available for the terrestrial system only, as they mostly build from land cover maps. Developing criteria-explicit methods for aquatic systems is needed, making use of specific data (e.g., maps of dams, fishing, or ocean acidification).

How can we better include estimates of the impacts of threats other than habitat loss? Currently, many methods are based on the analysis of land-cover data, and do not consider threats such as invasive species, over-exploitation, pollution, or emerging infectious diseases. Gathering relevant data on these threats and integrating them in both category-predictive and criteriaexplicit methods is needed to better inform future assessments.

How can we provide consistent estimates of climate change impact directly applicable to Red List criteria? Vulnerability to climate change (combining exposure to past and future climate change, sensitivity, and adaptive capacity) is difficult to evaluate by assessors, and urgently needs to be better integrated into assessments.

How can we consistently apply criterion E? Criterion E (i.e., quantitative analyses of extinction probability) is very rarely used, but could provide an effective way of integrating the impact of threats affecting species on a longer-term, as it can accommodate a time window of up to 100 years, regardless of generation time.

How can we better predict species down-listing? Clearly identifying when species extinction risk decreases is also important to keep the Red List updated and to highlight conservation successes. Specific methods, for instance integrating conservation actions, will need to be developed to enable predicting positive population trends or range expansions.



#### References

- Betts, J. et al. (2020) A framework for evaluating the impact of the IUCN Red List of threatened species. Conserv. Biol. 34, 632–643
- Rodrigues, A. et al. (2006) The value of the IUCN Red List for conservation. Trends Ecol. Evol. 21, 71–76
- 3. Stuart, S.N. et al. (2010) The barometer of life. Science 328, 177
- Hoffmann, M. et al. (2010) The impact of conservation on the status of the world's vertebrates. Science 330, 1503–1509
- Williams, B.A. et al. (2021) A robust goal is needed for species in the post-2020 global biodiversity framework. Conserv. Lett. 14, e12778
- Brummitt, N.A. et al. (2015) Green plants in the red: a baseline global assessment for the IUCN sampled Red List Index for plants. PLoS ONE 10, e0135152
- IUCN (2012) IUCN Red List Categories and Criteria (Version 3.1, 2nd edn), IUCN
- Mace, G.M. et al. (2008) Quantification of extinction risk: IUCN's system for classifying threatened species. Conserv. Biol. 22, 1424–1442
- Juffe-Bignoli, D. et al. (2016) Assessing the cost of global biodiversity and conservation knowledge. PLoS ONE 11, e0160640
- Rondinini, C. et al. (2014) Update or outdate: long-term viability of the IUCN Red List. Conserv. Lett. 7, 126–130
- Cardoso, P. et al. (2011) Adapting the IUCN Red List criteria for invertebrates. Biol. Conserv. 144, 2432–2440
- 12. Nic Lughadha, E. *et al.* (2020) Extinction risk and threats to plants and fundi. *Plants People Planet* 2, 389–408
- 13. Eisenhauer, N. et al. (2019) Recognizing the quiet extinction of invertebrates. *Nat. Commun.* 10, 50
- IUCN (2021) The IUCN Red List 2021-2030 Strategic Plan. Preventing extinction and advancing the recovery of species using The IUCN Red List of Threatened Species, Gland, Switzerland.
- Butchart, S.H.M. and Bird, J.P. (2010) Data Deficient birds on the IUCN Red List: what don't we know and why does it matter? *Biol. Conserv.* 143, 239–247
- Bland, L.M. et al. (2017) Toward reassessing Data-Deficient species. Conserv. Biol. 31, 531–539
- Bland, L.M. and Böhm, M. (2016) Overcoming data deficiency in reptiles. *Biol. Conserv.* 204, 16–22
- Verde Arregoitia, L.D. (2016) Biases, gaps, and opportunities in mammalian extinction risk research: investigating extinction risk in mammals. Mammal Rev. 46, 17–29
- Bachman, S.P. et al. (2019) Progress, challenges and opportunities for Red Listing. Biol. Conserv. 234, 45–55
- Santini, L. et al. (2019) Applying habitat and population-density models to land-cover time series to inform IUCN Red List assessments. Conserv. Biol. 33, 1084–1093
- Cardillo, M. and Meijaard, E. (2012) Are comparative studies of extinction risk useful for conservation? *Trends Ecol. Evol.* 27, 167–171
- Zizka, A. et al. (2021) Automated conservation assessment of the orchid family with deep learning. Conserv. Biol. 35, 897–908
- Cardillo, M. et al. (2008) The predictability of extinction: biological and external correlates of decline in mammals. Proc. B 275, 1441–1448
- Zizka, A. et al. (2020) Biogeography and conservation status of the pineapple family (Bromeliaceae). Divers. Distrib. 26, 183–195
- Bachman, S. et al. (2011) Supporting Red List threat assessments with GeoCAT: geospatial conservation assessment tool. ZK 150, 117–126
- Cardoso, P. (2017) red an R package to facilitate species red list assessments according to the IUCN criteria. BDJ 5, e20530
- Dauby, G. et al. (2017) ConR: an R package to assist largescale multispecies preliminary conservation assessments using distribution data. Ecol. Evol. 7, 11292–11303
- Lee, C.K.F. et al. (2019) Redlistr: tools for the IUCN Red Lists of ecosystems and threatened species in R. Ecography 42, 1050–1055
- Bachman, S. et al. (2020) Rapid Least Concern: towards automating Red List assessments. BDJ 8, e47018
- Sherley, R.B. et al. (2020) Estimating IUCN Red List population reduction: JARA—A decision-support tool applied to pelagic sharks. Conserv. Lett. 13, e12688

- Brooks, T.M. et al. (2019) Measuring terrestrial area of habitat (AOH) and its utility for the IUCN Red List. Trends Ecol. Evol. 34, 977–986
- Ficetola, G.F. et al. (2015) Habitat availability for amphibians and extinction threat: a global analysis. *Divers. Distrib.* 21, 302–311
- Tracewski, Ł. et al. (2016) Toward quantification of the impact of 21st-century deforestation on the extinction risk of terrestrial vertebrates: effects of deforestation on vertebrates. Conserv. Biol. 30, 1070-1079
- Visconti, P. et al. (2016) Projecting global biodiversity indicators under future development scenarios: projecting biodiversity indicators. Conserv. Lett. 9, 5–13
- Bird, J.P. et al. (2012) Integrating spatially explicit habitat projections into extinction risk assessments: a reassessment of Amazonian avifauna incorporating projected deforestation. Divers. Distrib. 18, 273–281
- Buchanan, G.M. et al. (2008) Using remote sensing to inform conservation status assessment: estimates of recent deforestation rates on New Britain and the impacts upon endemic birds. Biol. Conserv. 141, 56–66
- González-Suárez, M. et al. (2013) Which intrinsic traits predict vulnerability to extinction depends on the actual threatening processes. Ecosphere 4, art76
- Murray, K.A. et al. (2014) Threat to the point: improving the value of comparative extinction risk analysis for conservation action. Glob. Change Biol. 20, 483–494
- Pelletier, T.A. et al. (2018) Predicting plant conservation priorities on a global scale. PNAS 115, 13027–13032
- Davidson, A.D. et al. (2009) Multiple ecological pathways to extinction in mammals. PNAS 106, 10702–10705
- 41. Chichorro, F. et al. (2019) A review of the relation between species traits and extinction risk. *Biol. Conserv.* 237, 220–229
- 42. Di Marco, M. et al. (2018) Changes in human footprint drive changes in species extinction risk. Nat. Commun. 9, 4621
- Darrah, S.E. et al. (2017) Using coarse-scale species distribution data to predict extinction risk in plants. Divers. Distrib. 23, 435–447
- Di Marco, M. et al. (2014) Drivers of extinction risk in African mammals: the interplay of distribution state, human pressure, conservation response, and species biology. *Philos. Trans. R.* Soc. B 369, 20130198
- Di Marco, M. et al. (2015) Historical drivers of extinction risk: using past evidence to direct future monitoring. Proc. B 282, 20150928
- 46. Greenville, A.C. *et al.* (2021) Simultaneously operating threats cannot predict extinction risk. *Conserv. Lett.* 14, e12758
- Ruland, F. and Jeschke, J.M. (2017) Threat-dependent traits of endangered frogs. *Biol. Conserv.* 206, 310–313
- Zizka, A. et al. (2021) IUCNN Deep learning approaches to approximate species' extinction risk. Divers. Distrib.
- Böhm, M. et al. (2016) Correlates of extinction risk in squamate reptiles: the relative importance of biology, geography, threat, and range size: extinction risk correlates in squamate reptiles. *Glob. Ecol. Biogeogr.* 25, 391–405
- Bland, L.M. et al. (2015) Predicting the conservation status of data-deficient species: predicting extinction risk. Conserv. Biol. 29, 250–259
- Safi, K. and Pettorelli, N. (2010) Phylogenetic, spatial and environmental components of extinction risk in carnivores: decomposing extinction risk. Glob. Ecol. Biogeogr. 19, 352-362
- Luiz, O.J. et al. (2016) Predicting IUCN extinction risk categories for the world's Data Deficient groupers (Teleostei: Epinephelidae). Conserv. Lett. 9, 342–350
- Bland, L.M. (2017) Global correlates of extinction risk in freshwater crayfish. *Anim. Conserv.* 20, 532–542
- Howard, S.D. and Bickford, D.P. (2014) Amphibians over the edge: silent extinction risk of Data Deficient species. *Divers. Distrib.* 20, 837–846
- Lucas, P.M. et al. (2019) Range area matters, and so does spatial configuration: predicting conservation status in vertebrates. Ecography 42, 1103–1114



- 56. He, F. et al. (2021) Combined effects of life-history traits and human impact on extinction risk of freshwater megafauna. Conserv. Biol. 35, 643-653
- 57. Santini, L. et al. (2021) The interface between macroecology and conservation; existing links and untapped opportunities. Front. Biogeography 13, e53025
- 58. Akçakaya, H.R. (2017) Do you have an idea or concern about the Red List? IUCN SSC Quarterly Report, September 2017. pp. 19-20
- 59. Moat. J. et al. (2019) Least concern to endangered: applying climate change projections profoundly influences the extinction risk assessment for wild Arabica coffee. Glob. Change Biol. 25, 390-403
- 60. Foden, W.B. et al. (2013) Identifying the world's most climate change vulnerable species: a systematic trait-based assessment of all birds, amphibians, and corals. PLoS ONE 8, e65427
- 61. Pacifici, M. et al. (2015) Assessing species vulnerability to climate change. Nat. Clim. Chang. 5, 215-224
- 62. Pacifici, M. et al. (2017) Species' traits influenced their response to recent climate change. Nat. Clim. Chang. 7, 205-208
- 63. Hilbers, J.P. et al. (2017) Setting population targets for mammals using body mass as a predictor of population persistence. Conserv. Biol. 31, 385-393
- 64. Santini, L. et al. (2018) TetraDENSITY: a database of population density estimates in terrestrial vertebrates, Glob, Ecol, Biogeogr. 27. 787-791
- 65. Koh. L.P. et al. (2004) Species coextinctions and the biodiversity crisis, Science 305, 1632-1634
- 66 Kwak M L et al. (2020) Methods for the assessment and conservation of threatened animal parasites. Biol. Conserv. 248, 108696
- 67. Moir, M.L. and Brennan, K.E.C. (2020) Incorporating coextinction in threat assessments and policy will rapidly improve the accuracy of threatened species lists. Biol. Conserv. 249, 108715
- 68. Akçakaya, H.R. et al. (2018) Quantifying species recovery and conservation success to develop an IUCN Green List of Species. Conserv. Biol. 32, 1128-1138
- 69. IUCN (2021) IUCN Green Status of Species (1st edn), IUCN, International Union for Conservation of Nature
- 70. Kennerley, R.J. et al. (2021) Global patterns of extinction risk and conservation needs for Rodentia and Eulipotyphla. Divers. Distrib. 27, 1792-1806
- 71. Stewart, C.L. et al. (2021) Determining ranges of poorly known mammals as a tool for global conservation assessment. Biol. Conserv. 260, 109188
- 72. SCBD (2021) First draft of the post-2020 global biodiversity framework. CBD/WG2020/3/3 5 July 2021, Secretariat of the Convention on Biological Diversity
- 73. Akçakaya, H.R. et al. (2000) Making consistent IUCN classifications under uncertainty. Conserv. Biol. 14, 1001-1013

- 74. Pacifici, M. et al. (2013) Generation length for mammals. Nat. Conserv. 5, 89-94
- 75. Bird, J.P. et al. (2020) Generation lengths of the world's birds and their implications for extinction risk. Conserv. Biol. 34, 1252-1261
- 76. Keith, D.A. et al. (2018) Scaling range sizes to threats for robust predictions of risks to biodiversity. Conserv. Biol. 32, 322-332
- 77. Joppa, L.N. et al. (2015) Impact of alternative metrics on estimates of extent of occurrence for extinction risk assessment; extent of occurrence and extinction risk, Conserv. Biol. 30. 362-370
- 78. Araújo, M.B. et al. (2005) Validation of species-climate impact models under climate change. Glob. Change Biol. 11, 1504-1513
- Akçakaya, H.R. et al. (2006) Use and misuse of the IUCN Red List Criteria in projecting climate change impacts on biodiversity. Glob. Change Biol. 12, 2037–2043
- 80. Porszt, E.J. et al. (2012) Reliability of indicators of decline in abundance. Conserv. Biol. 26, 894-904
- 81. Akçakaya, H.R. et al. (2021) Calculating population reductions of invertebrate species for IUCN Red List assessments. J. Insect Conserv. 25, 377-382
- Butchart, S.H.M. et al. (2006) Going or gone: defining 'Possibly Extinct' species to give a truer picture of recent extinctions. Bull. Br. Omithol. Club 126A, 19
- 83. Keith, D.A. et al. (2017) Inferring extinctions I: a structured method using information on threats. Biol. Conserv. 214,
- Thompson, C.J. et al. (2017) Inferring extinctions II: a practical, iterative model based on records and surveys. Biol. Conserv. 214 328-335
- 85. Collen, B. et al. (2016) Clarifying misconceptions of extinction risk ssessment with the IUCN Red List. Biol. Lett. 12, 20150843
- 86. Ocampo-Peñuela, N. et al. (2016) Incorporating explicit geospatial data shows more species at risk of extinction than the current Red List. Sci. Adv. 2, e1601367
- 87. Roberts, D.R. et al. (2017) Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. Ecography 40, 913-929
- 88. Walker, B.E. et al. (2020) Caution needed when predicting species threat status for conservation prioritization on a global scale. Front. Plant Sci. 11, 520
- 89. Andermann, T. et al. (2021) juch sim: a new program to simulate future extinctions based on IUCN threat status. Fcography 44, 162-176
- 90. Foden, W.B. and Young, B.E. (2016) IUCN SSC Guidelines for Assessing Species' Vulnerability to Climate Change (Version 1.0) Occasional Paper of the ILICN Species Survival Commission No. 59, IUCN Species Survival Commission