Three in four undescribed plant species are threatened with extinction

Matilda Brown (✉ m.brown2@kew.org)
Royal Botanic Gardens, Kew  https://orcid.org/0000-0003-2536-8365

Steven Bachman
Royal Botanic Gardens, Kew  https://orcid.org/0000-0003-1085-6075

Eimear Nic Lughadha
Royal Botanic Gardens Kew  https://orcid.org/0000-0002-8806-4345

Brief Communication

Keywords:

Posted Date: June 16th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2953333/v2

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Version of Record: A version of this preprint was published at New Phytologist on August 15th, 2023. See the published version at https://doi.org/10.1111/nph.19214.
Abstract

Many thousands of plant species remain unknown to science, most being range-restricted and/or rare – both factors that increase extinction risk. Here, we show that more than 75% of species described after 2020 would qualify as threatened if assessed for the IUCN Red List. We recommend that newly described species are assumed to be threatened and encourage taxonomists to partner with assessors at the time of description to formally evaluate the extinction risk of new species using the IUCN Red List Categories and Criteria.

Main

Plant species new to science are described at a rate of approximately 2000 per year (Cheek et al., 2020). This rate has remained constant for decades, suggesting that there are still tens of thousands of species currently unknown to science – some estimates suggest that over 20% of plant species (up to 100,000) are yet to be discovered (Pimm & Joppa, 2015; Corlett, 2016; Heywood, 2017). Recently described plants have been shown to have almost twice the number of extinctions over time as those described pre-1900 (Humphreys et al., 2019). Similarly, the relative probability of extinction, as categorised using the International Union for Conservation of Nature's Red List of Threatened Species (hereafter Red List; IUCN, 2012), for newly described species is assumed to be high because these species have restricted ranges or are known from a small number of individuals (Joppa et al., 2011; Scheffers et al., 2012; Pimm et al., 2014). Indeed, most new plant species have extremely narrow ranges (Cheek et al., 2020). These same factors (range and number of individuals) underpin recent findings that vertebrate species described today are almost three times more likely to be threatened (Red List categories Vulnerable, Endangered, Critically Endangered) than those described in the 18th century (Liu et al., 2022). However, the relationship between extinction risk and year of description in plants has, until now, never been empirically characterised.

We analysed data from the World Checklist of Vascular Plants (Govaerts et al., 2021) and the IUCN Red List (IUCN, 2022) to characterise the relationship between year of description and species extinction risk across lifeform and climatic partitions using Bayesian logistic regression. We provide the first quantitative estimates of the extinction risk of plants described after 2020 (including those yet to be described).

We found a clear relationship between year of description and probability of being threatened (Fig 1), with marginal statistical support for the effects of lifeform and climate (Supporting Information Table S1, Fig. S1). Here we focus on the simplest model (threatened ~ year) because the coefficients for this model fall within the credibility intervals of the more complex models (Fig. S1) making our recommendations consistent across all models. Overall, we estimate that 77.3% (CI=76.6-78%) of species described in 2020 meet the criteria to be assessed as threatened, with an increase of 0.0165 logits per year (CI=0.0164-0.1066; Table S1), equivalent to a maximum rate of increase in threat probability of 4% per decade. We also found that species described more recently are more likely to meet the criteria for a higher threat
category (Fig. 1), with 59.8% (CI= 58.9%-60.7%) of species described in 2020 meeting the criteria for Endangered, and 24.2% (CI= 23.3%-25.1%) meeting the criteria to be assessed as Critically Endangered.

Our results are consistent with the hypothesis that species with small ranges and/or population sizes are less likely to be encountered in the wild (and thus less likely to be described) and more likely to be threatened (Gaston, 2003). Further support is provided by the increasing proportion of species described each year that are endemic to a single botanical country (Level 3 of the WGSRPD; Brummitt et al. 2001; a crude proxy for range size); species described before 1900 are generally <50% endemics, but this has increased to 92.5% of species described in 2020 (Fig. S2). Our findings are also consistent with comparable studies on vertebrates, though the relationship is steeper for plants than for animals (Liu et al., 2022). This congruence across kingdoms has implications for other, even less well-known groups. It is possible that, despite disparate modes of species discovery, a similar pattern holds for fungi, where it is thought that just 10% of species (Hawksworth & Lücking, 2017) have been described.

The observed proportion of threatened species in recent years is greater than predicted by our model (2010-2020, Fig. 1). If this underprediction is due to model misspecification, our estimates may be conservative, with true probabilities increasing more rapidly than predicted. However, we believe it is more likely that underprediction of extinction risk for recently described species is due to a bias relating to the lags between a species’ first collection (Bebber et al., 2010), description, and assessment - supported by the smaller proportion of recently described species that have been assessed (Fig. S3). Most recently described species have not yet been assessed and published on the Red List, and there may be additional challenges around obtaining reliable data needed for formal assessment (e.g. distribution or population information), particularly for species that are described based on a single herbarium specimen.

Exceptions to this are where there is obvious and immediate risk to a newly described species, e.g. *Pseudohydrosme ebo*, for which the Critically Endangered status was included in the description (Cheek et al., 2021). Together with 13 other recently described and threatened species, this assessment contributed to the revocation of logging concessions for the area.

To be formally protected, species must be noticed, described, and assessed. In light of our results, we recommend that all species new to science should be assumed to be threatened until formally assessed. While such an approach might be considered radical, it would in fact be complementary to the International Finance Corporation’s Performance Standard 6 (Biodiversity Conservation and Sustainable Management of Living Natural Resources; International Finance Corporation, 2012) which mandates recognition of Critical Habitat through its significance to Critically Endangered or Endangered species and/or to endemic and/or restricted-range species, a combination of criteria which encompasses the vast majority of newly described plant species (Figure 1, Fig. S2). Our recommended approach is also consistent with the precautionary principle espoused by IUCN (2012). In addition, our findings highlight the continued importance of both taxonomic research, to identify and describe new species, and conservation assessments under internationally recognised schemes (e.g. the Red List) to evaluate the extinction risk of these new species. Ideally, partnerships between taxonomists and experienced IUCN Red
List assessors should describe and assess species simultaneously, thus maximising opportunities for successful conservation interventions.

If, as has been suggested, there are ~100,000 species of plant yet to be described, at least 75,000 of them are likely to be threatened with extinction. Even if this is an overestimate, our recommendation could save many of the tens of thousands of undescribed, threatened species that, without continued taxonomic efforts, assessment, and intervention, may be lost before they are ever known to science.

References


Methods

All data preparation, analysis and visualisation were undertaken using R 4.2.2 (R Core Team, 2022); all materials required to reproduce our results are available at https://github.com/matildabrown/YoD.

[Note to reviewers: this repository will be made public upon acceptance of the manuscript for publication; we have included these materials with our submission for review as ThreeinFour_GitHubSI.zip]

We reconciled the Red List (v2022-2; IUCN, 2022) with the World Checklist of Vascular Plants (v10; Govaerts et al., 2021) using the R package ‘rWCVP’ (Brown et al., 2023). We labelled species assessed as Vulnerable (VU), Endangered (EN), Critically Endangered (CR), Extinct in the Wild (EW) or Extinct (EX) as ‘threatened’, those assessed as Least Concern or Near Threatened as ‘non-threatened’ and those that were Data Deficient or Not Evaluated as missing; this follows the IUCN guidance (https://www.iucnredlist.org/resources/redlistguidelines), though we have included EW and EX as ‘threatened’ as has been done previously (Nic Lughadha et al., 2020), their inclusion being justified by the high proportion of species declared extinct which are subsequently rediscovered (Humphreys et al., 2019). We obtained the year of description for each species from the ‘year of publication’ field in WCVP for its basionym where applicable, or from the same field for the accepted name. We defined species’ climatic zones as tropical, subtropical, temperate or polar using the midpoint of their latitudinal range (at the resolution of Level 3 of the World Geographic Scheme for Recording Plant Distributions (Brummitt et al., 2001), included in the World Checklist of Vascular Plants). Although there are known biases in both the Red List (Nic Lughadha et al., 2020) and species descriptions (Cheek et al., 2020) we are satisfied that these biases are either relatively constant through time or are otherwise unlikely to negate our results (Figs S3-5).

We modelled extinction risk (threatened or non-threatened) as a function of year of description in a Bayesian framework using the ‘brms’ package (Bürkner, 2017). We fitted additional models that included interactions between year and lifeform (woody, annual, epiphyte or herbaceous perennial; following a modified version of the mapping in Humphreys et al., 2019; available at https://github.com/matildabrown/YoD), and year and climate to identify heterogeneity in the relationship between extinction risk and year of description. We performed model selection using Pareto-smoothed importance sampling values (PSIS-LOO; Vehtari et al., 2017) to identify the best-fitting models from this set. To investigate the hypothesis that newly described plants are not only more likely to be threatened, but more likely to be more threatened (i.e. meet the criteria for a higher threat category on the Red List) we used the best model formula to model the probability of threat at two additional thresholds (Endangered or higher; Critically Endangered or higher).

Probability is non-linear in a logistic regression model, but we can interpret the effect size in terms of probability by considering a species with a predicted probability of being threatened of 0.5 (=0 logits). If our linear coefficient is 0.01 logits; the predicted likelihood of the same species being threatened if described one year later is 0+0.01 logits, equivalent to a probability of 0.5025. If described a decade later, the likelihood is 0+0.01*10=0.1 logits, equivalent to a probability of 0.525. At very low or high
probabilities (e.g. if the above example species was described a century later), the rate is 'flattened' by the bounds of probability at 0 and 1. Thus, our maximum rate of increase in probability is 0.0025, or 0.25% per year.

We did not correct for evolutionary history in our main analyses because the aim of this study was to generate predictive estimates of the probability that a newly described species is threatened. Although angiosperm-wide megaphylogenies have recently been reconstructed (Smith & Brown, 2018), we do not expect that species’ positions in these supertrees will necessarily have been inferred at the time of description. We did analyse our data in a phylogenetic context using a recent supertree reconstructed by F. Forest (doi: 10.5281/zenodo.7600341); although models fitted with 'phyloglm'(Ho & Ané, 2014) suggested some phylogenetic signal in the relationship between year of description and extinction risk (alpha = 1 when tree rescaled to total height of 1) we were unable to achieve model convergence and thus reliable results. Nevertheless, the beta coefficient for year in the non-convergent model was 0.0179 (slightly higher than our reported coefficients), so our effect size is likely to be conservative with regard to evolutionary history but not meaningfully different.

**Declarations**

**Author contributions**

MB, SB and ENL conceived and designed the study. MB led analysis and writing of the manuscript with significant contributions from SB and ENL.

**Figures**
Figure 1

Observed proportions (background bars) and predicted probabilities (lines) of threatened species by year of description. Shaded ribbons give the 95% credibility interval for the model coefficients at each level of threatened: VU or higher (yellow, upper line), EN or higher (orange, middle line) and CR or higher (red, lower line). The vertical dashed line at the year 2020 denotes our reference year for predictions, and the last year of data that we used to fit our models.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- ExtendedData.docx