Plants, People, Planet Supporting Information

Article title: Extinction Risk and Threats to Plants and Fungi.
Authors: Eimear Nic Lughadha, Steven P Bachman, Tarciso C C Leão, Felix Forest, John M
Halley, Justin Moat, Carmen Acedo, Karen Bacon, Ryan F ABrewer, Gildas Gâteblé, Susana
C Gonçalves, Rafaël Govaerts, Peter Hollingsworth, Irmgard Krisai-Greilhuber, Elton John
de Lirio, Paloma G P Moore, Raquel Negrão, Jean Michel Onana, Landy Rajaovelona,
Henintsoa Razanajatovo, Peter B Reich, Sophie L Richards, Malin C Rivers, Amanda Cooper,
João Iganci, Gwilym P. Lewis, Eric C Smidt, Alexandre Antonelli, Gregory M Mueller,
Barnaby E Walker

The following Supporting Information is available for this article:

Methods S1 Sections 1.1 Red List Coverage of Plants and Fungi and 1.2 -Progress towards

assessing all plants.

Fig S1 Performance of the fuzzy matching process on a random sample of 100 species

names as the threshold on the similarity between names is increased from 0.7 to 1.0 in

increments of 0.025.

Table S1 Key metrics for the ThreatSearch database at both snapshots, showing the

refinement from raw counts of plant assessments to accepted assessments after matching

to the World Checklist of Vascular Plants (WCVP).

Table S2 Breakdown of assessments in both ThreatSearch snapshots by scope and threat

category, after matching to WCVP.

Methods S2 Section 1.3 - Biases and gaps in plant and fungal coverage on the IUCN Red List.

Table S3 The proportion of accepted species in the World Checklist of Vascular Plants with

data available for each of our chosen attributes.

Methods S3 Section 1.4 - Most and least threatened plant groups.

Fig S2 Prior predictive check of overall proportion of vascular plants that are threatened.

1000 samples were drawn from the specified model before priors were updated with the

data. The resulting histogram shows a distribution that matched our prior expectations –

most of the samples were below 0.5, and previous estimates of the proportion of vascular

plants that are threatened (from the Sampled Red List Index and the IUCN Red List) were in

regions of the distribution with a high number of samples.

Fig S3 Posterior predictive checks to judge how well our model fits the data. Samples from

the posterior predictive distribution of the fitted model are used to compare (A) the

predicted distribution (blue lines) of values for the proportion of threatened species in each

cell defined by combinations of all parameters in our model to the observed distribution (black line), as well as (B) the predicted distribution (blue lines) of the number of threatened species in each cell to the observed distribution (black line). The samples were also used to compare (C) the predicted number of threatened species in each cell to the observed number.

Table S4 Mapping of IUCN Red List categories into two groups: "threatened" and "not threatened".

Table S5 Species counts after matching IUCN Red List assessments to WCVP.

Table S6 Comparison of predictive performance of nested models by PSIS-LOO, showing that our final model incorporating all the variables of interest had the most favourable values of both the PSIS-LOO and ELPD.

Table S7 Estimated values of overall intercept (baseline) and fixed effects in our model.

Table S8 Estimated values for each level of the lifeform random effect in our model.

Table S9 Estimated values for different levels of the continent random effect.

Table S10 Exponentiated parameter values for the 37 families with random effect that increased the odds of being threatened.

Table S11 Predicted proportion of threat for the 10 families with the most under-estimated level of threat.

Table S12 Predicted proportion of threat for the 10 families with the most over-estimated level of threat.

Methods S4 Section 4 - Trends in extinction risk in plants and fungi.

Figure S4 Diagram illustrating how AOH is derived from the intersection of the elevation limits, habitat preferences and species range.

Table S13 Red List Index values – see Figure 4A.

Methods S5 Examples of Reassessments.

Methods S6 Methods used for Box 1.

Dataset S1 Breakdown of proportion of vascular plant species in the World Checklist of Vascular Plants and the IUCN Red List defined by our chosen attributes [External Excel worksheet]

Dataset S2 Modelling parameters and predictions [External Excel worksheet]

References

Methods S1 Sections 1.1 Red List Coverage of Plants and Fungi and 1.2 - Progress towards assessing all plants.

A species is considered extinct when no living individual of that species is known to survive. The term can be qualified in various ways to refer to the regional or local extinction (extirpation) of a species from a particular area though it survives elsewhere in the wild, or to species alive in cultivation or storage but no longer surviving in the wild, which are termed Extinct in the Wild. Terms used in our review broadly follow the global IUCN Red List of Threatened Species (IUCN, 2020; hereafter Red List) and its categories and criteria for extinction risk evaluation which are widely recognised and applied internationally (but see also Box 2). Extinction risk assessments considering only a portion of a species range are not included in the quantitative analyses in our review, but sometimes mentioned in examples.

Application of IUCN Red List Guidelines for Reporting on Proportion Threatened

IUCN have developed guidelines on reporting the proportion of species threatened in a group – see:
https://nc.iucnredlist.org/redlist/content/attachment_files/Guidelines_for_Reporting_Proportion_Threatened_ver1_1.pdf

The method accounts for the uncertainty in estimates of proportion of species threatened due to the unknown threat status of Data Deficient species and provides a midpoint or best estimate along with an upper and lower bound, where all DD species are recognised to be, as threatened as data sufficient species (upper bound), or all threatened or all Least Concern (lower bound). Extinct species are excluded from all calculations of proportions of species threatened. Although the guidelines are intended to be applied to completely assessed groups or those assessed by means of a random sampling approach, we considered it informative to use this method to illustrate the uncertainty of our estimate, where this could be appropriately applied.

Therefore, our analysis of the current Red List (2020-1) coverage of plants shows a midpoint estimate of 43.7% threatened, with lower estimate of 40.7%, and upper 47.5%. The corresponding figure for 2016 is a mid-point estimate of 59.3% threatened, with lower estimate of 54.5% and upper estimate of 62.4%. This is consistent with the longer-term

97　trend for the proportions of assessed plants considered threatened to decrease as coverage

98　increases, reflecting IUCN's strategic focus on assessing whole groups rather than species

99　suspected to be threatened. We also used the mid-point estimate to show the threat-level

100　of assessed species in different countries (Fig. 2B).

101　When considering proportion threatened in our analysis of assessments in ThreatSearch and

102　our predictive modelling (section 1.4), we use generalised categories to indicate threat

103　status, data deficiency and extinction that may cannot be fully equated to global IUCN

104　categories and therefore use of the IUCN methodology for reporting percentage threatened

105　would not be appropriate.

106

107　Over 60% of species with digitally available global assessments lack an assessment on the

108　global Red List. Our best estimate, 61.7% takes into account assessments published on the

109　Red List in 1998 and 2000, because these are included in ThreatSearch (see below).

110　Analyses reported in Sections 1.3 & 1.4 are confined to Red List assessments published since

111　2001, following a revision of the categories and criteria (IUCN 2001).

112　The ThreatSearch database collates all digitally-available evidence-based assessments for

113　plants, and is regularly updated by Botanic Gardens Conservation International (BGCI, 2020).

114　A snapshot of ThreatSearch (in November 2016) was used to quantify progress towards

115　GSPC Target 2 (Bachman et al., 2018) after reconciling scientific names to The Plant List 1.1

116　(TPL) and assessment status to defined standards.

117

118　For this review, we repeated the above analysis using ThreatSearch datasets from

119　November 2016 and January 2020. Our methods differed only in reconciling names,

120　including basionyms, to the World Checklist of Vascular Plants (WCVP, 2020). In line with the

121　original analysis by Bachman et al. (2018), we did not remove species listed as hybrids from

122　either the ThreatSearch or the WCVP data. Therefore, the total number of species from

123　WCVP used in this analysis was 347,298.

124

125

4

126 For both ThreatSearch datasets, we matched the plant names from assessments to names

127 on the WCVP. As the WCVP is limited to vascular plants, we removed all bryophytes from

128 the analysis.

129 As with any name matching process to reconcile datasets, there is an underlying assumption

130 that the name is being applied to the same taxon concept in the two different datasets.

131 Our name matching process followed a three-step protocol:

132

133 1.   Use the Kew Names Matching Service (KNMS) to match full taxon names, including

134   author strings, to the WCVP.

135 2.   Pass all unmatched taxon names through the KNMS again, but this time without the

136   author string.

137 3.   Use fuzzy matching to find the most similar name on WCVP for all remaining

138   unmatched ThreatSearch names. Our fuzzy matching process followed the

139   PlantMiner process used in the Bachman et al. (2018) as closely as possible.

140

141 The first two steps of our name matching process returned multiple matches for some taxa.

142 We resolved these matches to a single name for each ThreatSearch assessment as far as

143 possible. If the taxon was an infraspecific, we removed all matches that were for the wrong

144 type of infraspecific (e.g. removed all variety names for matches of a subspecies). If there

145 was a single match that was an accepted name, the matches for the taxon were resolved to

146 that. Otherwise, if there was a single match to a name identified as a homotypic synonym,

147 that was the name the match was resolved to. Any assessments with multiple matches

148 remaining after resolution were removed from the dataset.

149

150 We identified the best threshold to use for accepting fuzzy matches following the same

151 protocol as (Bachman et al., 2018). We took a sample of 100 names that were not matched

152 using KNMS and applied our fuzzy matching process to find the closest matching name on

153 WCVP and the similarity to that name. We varied the fuzzy matching threshold from 0.7 to

154 1.0 and tested the accuracy of the matching.

155

156 The fuzzy matching process can have four outcomes:

5

157　•　　True positive (TP) - a name that is correctly matched with a similarity above the

158　　　　chosen threshold

159　•　　False positive (FP) - a name that is incorrectly matched but with a similarity above

160　　　　the chosen threshold

161　•　　True negative (TN) - a name that is incorrectly matched and has a similarity below

162　　　　the chosen threshold

163　•　　False negative (FN) - a name that is correctly matched but with a similarity below the

164　　　　chosen threshold

165　We used these different outcomes to judge the quality of the matching with four different
166　measures. These were:

167　•　　Accuracy - the proportion of all matches that were correct

168
$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$

169　•　　Precision - the proportion of all positive matches that were correct

170
$$Precision = \frac{TP}{TP + FP}$$

171　•　　Sensitivity - proportion of all matches that were correctly identified

172
$$Sensitivity = \frac{TP}{TP + FN}$$

173　•　　Specificity - the proportion of all non-matches that were correctly identified

174
$$Specificity = \frac{TN}{TN + FN}$$

175　We found that a threshold of 0.92, close to the threshold of 0.925 used by Bachman et al.

176　(2018), was the best for our data, as no false positive matches were made at this threshold

177　(**Fig. S1**).

178

179　After the matching process, we updated all homotypic synonyms to their current accepted

180　name as per (WCVP, 2020). We then removed all assessments that were not matched to any

181　accepted name or homotypic synonym in WCVP, as well as all infraspecifics.

182

183　The ThreatSearch database is a compilation of all electronically available conservation

184　assessments of plant taxa. As such, it contains multiple assessments for some taxa, with

185　some assessed at different geographic scales and/or given different threat ratings. Following

186　Bachman et al., we refined our list of assessments for accepted species so that there was at

187　most a single assessment for each accepted species per threat category. For each threat

category that a species was assessed under, we first took the most recent global
assessment. If there were no global assessments for a species, we then took the most recent
assessment of unknown global scope. Finally, if there were no assessments with unknown
global scope, we took the most recent regional assessment.

It should be noted that this process means that some species will have multiple global
assessments that list them under different threat categories. As such, like in Bachman et al.
(2018), our figure for the number of global assessments with matched names (107,282) will
include multiple assessments for some species. The total number of accepted species with a
global assessment is lower, at 98,306 (28.3 % of all accepted species).

Using this cleaned list of assessments, we tabulated numbers of assessments and names
(**Table S1**) for each ThreatSearch database, comparable to Figure 1 in Bachman et al. (2018).
We have also provided a breakdown of the number of ThreatSearch assessments by
geographic scope and threat status (**Table S2**), comparable to Figure 2 in Bachman et al.
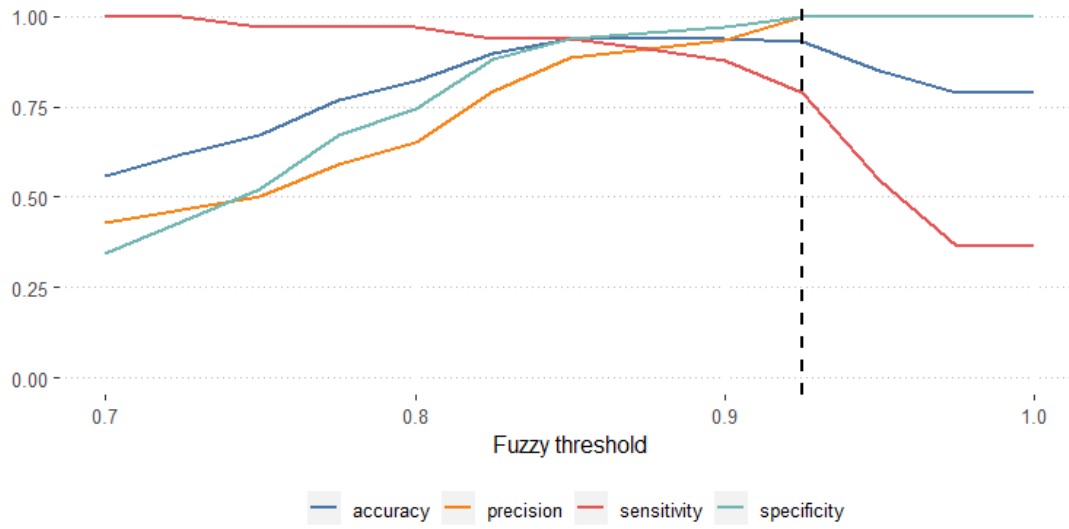(2018).

Fig S1 Performance of the fuzzy matching process on a random sample of 100 species names as the threshold on the similarity between names is increased from 0.7 to 1.0 in increments of 0. 025.

Table S1 Key metrics for the ThreatSearch database at both snapshots, showing the refinement from raw counts of plant assessments to accepted assessments after matching to the World Checklist of Vascular Plants (WCVP).

Quantity	Jan 2020	Nov 2016	Change %
Total assessments	313,565	228,569	37.2
Assessments with matched name	272,412	203,070	34.1
Accepted names with assessments	139,061	120,284	15.6
Global assessments with matched name	107,282	87,116	23.1
Global threatened assessments with matched name	36,503	31,472	16.0

Table S2 Breakdown of assessments in both ThreatSearch snapshots by scope and threat category, after matching to WCVP.

Scope	Category	Jan 2020	Nov 2016
global	extinct	831	517
global	threatened	36,503	31,472
global	possibly threatened	21,518	15,881
global	not threatened	40,341	29,771
global	data deficient	8,089	9,475
unknown	extinct	1,474	1,070
unknown	threatened	13,768	13,069
unknown	possibly threatened	8,095	5,830
unknown	not threatened	18,594	18,126
unknown	data deficient	3,550	3,640
not global	extinct	163	113
not global	threatened	3,432	2,018
not global	possibly threatened	1,304	952
not global	not threatened	8,267	7,818
not global	data deficient	1,533	1,102

Methods S2 Section 1.3 - Biases and gaps in plant and fungal coverage on the Red List.

We used comprehensive data from the WCVP (2020) to quantify biases in Red List coverage. We considered attributes likely to have influenced assessment priorities: the lifeform of the species, the taxonomic family to which it is assigned, on what continent the species occurs, whether the species is endemic to a single botanical country, and if the species had a recorded human use. Our definitions of continent and botanical country followed, respectively, levels 1 and 3 of the World Geographical Scheme for Recording Plant Distributions (WGSRPD; Brummitt, 2001).

We excluded both infraspecific and hybrid taxa from our analysis because they are not usually assessed. As such, the total number of accepted species from WCVP in our analysis was 340,024.

We sourced data on species lifeform, taxonomic family, and native range from the World Checklist of Vascular Plants. We used a list of the Useful Plant species (Diazgranados et al., 2020) to identify plants with a recorded human use. All assessment information was downloaded from the Red List of Threatened Species (IUCN, 2020).

Species lifeform is recorded in WCVP following the Raunkiaer system with some modifications. For ease of analysis, and comparison, we standardised these lifeform descriptions to follow those used in (Humphreys et al., 2019). The lifeform description for a species in WCVP can comprise multiple categories when the lifeform of a species is variable. In these cases, we took the first category in the list as the most frequent lifeform for that species and applied the appropriate mapping.

In WCVP, the native range of a species is recorded at level 3 of the WGSRPD. At level 3, these units correspond to what might be called "Botanical Countries", with some large countries like the USA sub-divided into states. We classified species as endemic if they only occurred in a single WGSRPD level 3 region.

We used level 1 of the WGSRPD to assign species as native to a particular continent. The definitions of North and South America are different from common political definitions, and as such are named Northern and Southern America to emphasise this distinction. Species that occurred in more than one continent were assigned to a separate "worldwide" category.
In order to identify which species had a recorded human use, we standardised the database of useful plant names to the WCVP taxonomic backbone using the same name matching procedure as used in our analysis of progress towards to GSPC target 2. We counted recorded uses of infraspecifics as a recorded use for the parent species if that parent species was accepted in WCVP.

While compilation of species names in WCVP has been officially completed (with some minor

works ongoing), compilation of geographic and lifeform data is less complete. This means that of the 340,024 accepted species in our dataset, all had taxonomic and use data available but 97.3 % had geographic data available, and only 69.5% had lifeform data (**Table S3**). As a result of this missing information, we could only carry out our analysis of biases and gaps in coverage of the IUCN Red List for the 234,808 (69.1%) of accepted vascular plant species that we had complete attribute data for.

To be able to analyse the biases and gaps of the Red List, we matched the names of all vascular plant assessments from the Red List of Threatened Plants to our WCVP data. In this case, we did not count assessments of infraspecifics as assessments for the parent species, as the range and threats to an infraspecific can be very different than those for the parent species. We also excluded all assessments carried out under Red List Categories and Criteria earlier than Version 3.1, treating species represented only by a pre-2001 assessment as unassessed. After this process, we were left with 33,029 accepted species in WCVP with global Red List assessments. Of these assessed species, 28,239 (85.5%) had complete attribute information, and so were included in our analysis.

The subdivision of some countries into multiple botanical countries by the WGSRPD system may have reduced the number of species assessed on the Red List that we identified as single-country endemics. To check if this subdivision made a difference to our analysis of biases and gaps in the Red List following the reconstruction of countries used by Gallagher et al. (2020). This demonstrated that using WGSRPD level 3 endemics underestimated the proportion of species on the Red List as endemic (49.1% for WGSRPD level 3 vs. 56.8% for countries reconstructed following Gallagher et al.). However, both analyses showed an under-representation of endemic species on the Red List compared to the proportion of species in WCVP that are endemic (56.0% for WGSRPD level 3 vs. 64.2% for countries reconstructed following Gallagher et al.).

We also found clear impacts of targeted assessment programmes on the taxonomic coverage of the Red List. The most under-represented families include some of the most species-rich, including Asteraceae, Orchidaceae, Poaceae, and Lamiaceae (together almost 25 % of all vascular plants). The most over-represented families include those targeted by assessment programmes, such as Cactaceae (all species covered by a global assessment; Goettsch et al., 2015), Myrtaceae (*Eucalyptus* assessment complete (Fensham et al., 2020) and *Myrcia* assessment underway), and Fabaceae (Leguminosae) (Global Legume Assessment part of Red List strategy since 2013).

A complete comparison of over and under representation of groups defined by combinations of our chosen attributes can be found in **Dataset S1**.

Table S3 The proportion of accepted species in the World Checklist of Vascular Plants with data available for each of our chosen attributes.

Attribute	Species with data / %
endemic	97.3
useful	100.0
lifeform	69.5
continent	97.3
family	100.0

Methods S3 Section 1.4 - Most and least threatened plant groups.

Given the biases illustrated above, the collection of plant species assessed for the Red List is not a representative sample of all plant diversity. This unrepresentativeness has previously been identified as preventing the use of the Red List directly as a "barometer of life" to estimate the current level of threat to plant species. One way to get this estimate is to assess the conservation status of a representative sample of all plant species, as has been done for the Sampled Red List Index. Another is to use modern statistical modelling techniques to use all plant assessments available on the Red List to correct for the unrepresentative sampling.

Multilevel regression and post-stratification (MRP) is a statistical modelling method that has been used successfully in recent years to predict the outcomes of political elections from unrepresentative polls (Wang et al., 2015). MRP works by first identifying variables that affect the inclusion and coverage of the unrepresentative sampling. In political polling this might be variables such as respondents' age and ethnicity. For our purposes, these variables are the attributes identified in our analysis of the biases and gaps of the Red List.

The combination of values of the chosen variables defines a set of individual cells. For instance, one individual cell in our analysis would be woody perennial species in the Fabaceae family from Europe that are endemic to a single region and have a recorded use.

In MRP, a multilevel regression is fit to the number of individuals from the unrepresentative sample to estimate the quantity of interest for each cell - in our case this is the probability that a species in that cell is threatened. Using a multilevel model means that some or all of the chosen variables are incorporated into the regression model as random effects. Using random effects allows partial pooling of information between levels of the variables in the models. This means that estimates for cells represented by a small number of individuals will be pulled towards the overall mean value for all cells.

Once the multilevel regression has been fit to unrepresentative sample, it is then used to get predictions of the value of interest for the overall population. This is done by tabulating all individuals in the population into the cells defined by our chosen variables. This cross-tabulation is called the post-stratification cells. The fit model is applied to these post-stratification cells to get predictions of the value of interest for the individual cells in the population. These can then be aggregated to get predictions of the mean value for the total population or for subpopulations of interest.

We applied MRP to the Red List data using the attributes identified in the previous section as the variables to define our modelling cells. We fit a multilevel regression model to species that

have been assessed for the Red List that had data available for all of our chosen variables (**Table S3**). We then applied our model to make predictions for post-stratification cells defined by all accepted species on the WCVP. We fit our modelling using a Bayesian framework that allowed us to propagate uncertainty estimates through to our predictions. This also allowed us to make predictions for species on WCVP that had values for our chosen variables that were not included in the data used to fit the model, or had values missing for certain variables.

By applying MRP to the Red List data we were able to both estimate the probability that a species in each of the individual cells defined by our variables is threatened and to identify groups of plants that have an over or under estimated level of threat based on the current Red List.

For ease of analysis, and as per Darrah et al. (2017), we grouped the Red List Categories into two categories: threatened and not threatened (**Table S4**). We defined our threatened category following the IUCN definition: all species assessed as either Critically Endangered, Endangered, or Vulnerable. We also included species assessed in the extinct categories, Extinct and Extinct in the Wild. We defined our not threatened category as the IUCN categories Near Threatened and Least Concern. We counted species assessed as Data Deficient as species that still require assessments i.e. equivalent to Not Evaluated.

Of the accepted species in our WCVP dataset, just over 9% had an assessment other than Data Deficient on the Red List, and almost 38% of those were assessed as threatened (**Table S5**).

Our target variable was whether a species is threatened or not. As this is a binary choice with a single trial for each species, much like flipping a coin a single time, we modelled this with a Bernoulli likelihood.

$$threatened_i \sim \text{Bernoulli}(p_i)$$

We modelled the probability that a species was threatened as being linearly related to our chosen variables, by using the logit link function. We included a general intercept term as the baseline odds that a species is threatened. We then included fixed effects for the terms indicating if a species was endemic to a WGSRPD or not and if a species had a recorded use. We included the terms for taxonomic family, continent, and lifeform as random effects.

$$logit(p_i) = \alpha + \beta_E Endemic + \beta_U Useful + \alpha_{family[i]}^{family} + \alpha_{continent[i]}^{continent} + \alpha_{lifeform[i]}^{lifeform}$$

We modelled each random effect as being drawn from common normal distributions.

$$\alpha_{family[i]}^{family} \sim \text{Normal}(0, \sigma_{family})$$

$$\alpha_{family[i]}^{family} \sim \text{Normal}(0, \sigma_{family})$$

$$\alpha_{lifeform[i]}^{lifeform} \sim \text{Normal}(0, \sigma_{lifeform})$$

To ensure that our final model was the best fit to the data, and that each term contributed to an increase in predictive power, we fit nested models by adding in each term in the order defined above. All models were fit with Bayesian sampling using the brms package (Bürkner, 2017), which is a high-level interface to the Stan statistical programming language (Carpenter et al., 2017).

We chose weakly informative priors for all parameters in our models. For fixed effect parameters, these were normal(0, 1) distributions. For the scale parameters on the hierarchical terms in our model, these were half-normal(0, 1) distributions. We also chose a weakly informative prior for the intercept parameter, but also used our knowledge that the baseline probability of being threatened for all species was unlikely to be above 0.5, and so chose the student_t(3, -1, 1) distribution.

We drew samples directly from the priors before fitting to our data (the prior predictive distribution) and checked the resulting histogram of values for the overall proportion of

threatened species (**Fig. S2**). This displayed a low number of samples at zero and one, with most of the samples falling below 0.5. Previous estimates of the proportion of vascular plant species that are threatened, from the Sampled Red List Index (SRLI) and the Red List, fall in regions of the histogram with a high number of samples. This demonstrated to us that our chosen priors resulted in plausible values for the proportion of vascular plants that are threatened.

We sampled 4 chains for 2000 samples each with 1000 warmup samples for every model. All models fit well with no divergent samples.

We compared the models using Pareto-smoothed importance sampling values (PSIS-LOO) and expected log predictive density (ELPD) differences (**Table S6**). These are measures of how well the models predict new data. PSIS-LOO is an information criterion that approximates cross-validation, with smaller values indicating a model that makes better predictions. Higher values of ELPD indicate a model that makes more reliable predictions. Our final model, including the full set of our variables, had the most favourable value of both PSIS-LOO and ELPD.

We further evaluated our MRP model by drawing 1000 samples from the posterior predictive distribution. We compared the distribution of values for the proportion of threatened species in each cell defined by the combination of all parameters in our model to the observed distribution (**Fig S3A**). This in general showed good agreement, although suggested that our model was under-predicting the number of cells with a high proportion of threatened species. We also compared the predicted distribution of the number of threatened species per cell to the observed distribution (**Fig S3B**), which suggested that our model was under-predicting the number of cells without any species. Comparing the posterior predictions of the number of threatened species in each cell to the observed number of threatened species (**Fig S3C**) confirmed that while our model performed well overall, it did not perform as well for cells with few to no threatened species.

The estimated coefficients in our model indicate how much each variable increases the log-

odds that a species is threatened. Exponentiating each coefficient indicates how much each variable multiplicatively increases the odds that a species is threatened.

The fixed effects in our model represent the baseline odds that a species is threatened, the increase in odds if a species is endemic to a single region, and the increase in odds if a species has a recorded use (**Table S7**).

The model fits a single value for each level of the random effects. These values indicate the factor that membership of each group changes the odds of being threatened by. For instance, no lifeform significantly increases or decreases the odds that a species is threatened (**Table S8**).

Similarly, there was no significant change in the probability of being threatened for most continents. Exceptions to this were species from Africa, which had an increased probability of being threatened, and Australasia, which had a decreased probability of being threatened (**Table S9**).

The large number of different levels for families makes it difficult to inspect each one, but most had small or no clear effect on the probability that a species is threatened. There were some families, however, that clearly increased the probability of being threatened (**Table S10**.

We used the predictions from our model to identify which groups of species might have an over or under-estimated level of threat based on those species assessed for the IUCN Red List. To do this, we calculated the difference between the observed proportion of threatened species in each group on the IUCN Red List and our sampled predictions. We judged families as under-estimated if the 95 % credible interval of the difference fell below zero, and over-estimated if the 95 % credible interval fell above zero. We measured the size of over or under-estimation by dividing the difference by the width of the 95 % credible interval.

All parameter estimates for our fitted model have been included in **Dataset S2**, along with the count of species in each post-stratification cell and the predicted proportions of threatened

species in each cell.

Although this method helps to correct for uneven sampling across the defined variables, it does make two assumptions that might limit the conclusions that can be drawn from it.

The first assumption is that we know the composition of our population. We are mainly drawing our information about the composition of our population from WCVP. This has been compiled from a range of data sources and has had extensive quality checks, but the quality of our predictions in part relies on the quality of the WCVP. This is most relevant in predictions for species that are missing data for one or more variables. While our model is able to make predictions for these species, these particular predictions will be very uncertain and that may limit the inferences we can make about the estimates for the variables that are missing, particularly about lifeforms (**Table S3**).

The second assumption is that the species that have been assessed in each cell are a random subsample of the population for that cell. Another way of stating this is that we have accounted for all of the major variables that could affect inclusion in our unrepresentative sample. One variable which has definitely influenced inclusion in our sample is motivation for assessment, and specifically, the imperative to assess species considered highly likely to be threatened so that they can benefit from the degree of protection afforded by being categorised as Endangered or Critically Endangered. We could not account for this variable as did not have sufficient data to determine the motive for assessment for most of the species assessed on the Red List. Another obvious example of a variable that we have not accounted for is the genus a species belongs to. Assessment efforts are often focused on particular genera and may bias which species are assessed within a family e.g. all species of the genus Coffea (source of coffee) have been assessed from the large family Rubiaceae. This may affect the predictions that we make using our model.
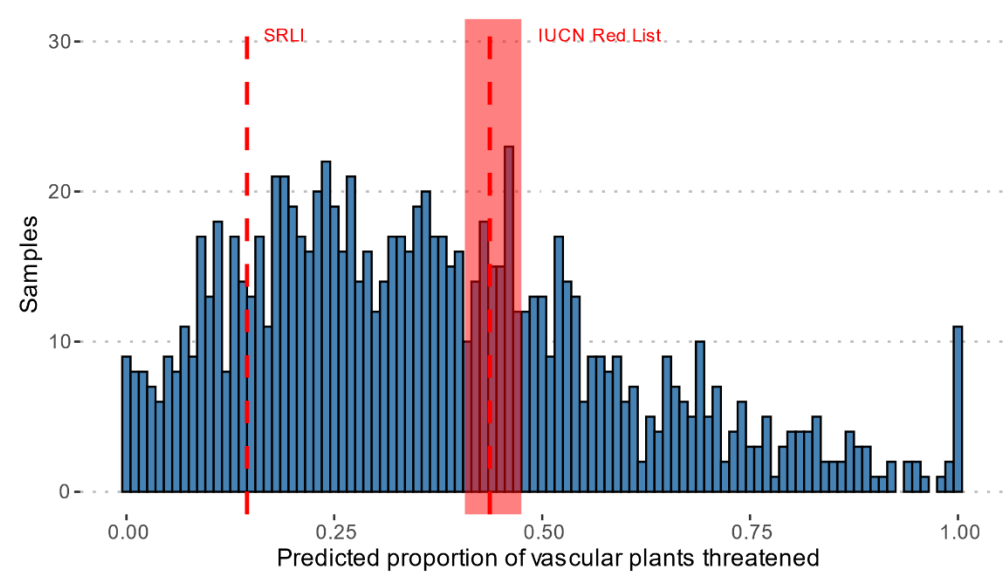
Fig S2 Prior predictive check of overall proportion of vascular plants that are threatened. 1000 samples were drawn from the specified model before priors were updated with the data. The resulting histogram shows a distribution that matched our prior expectations – most of the samples were below 0.5, and previous estimates of the proportion of vascular plants that are threatened (from the Sampled Red List Index and the Red List) were in regions of the distribution with a high number of samples.
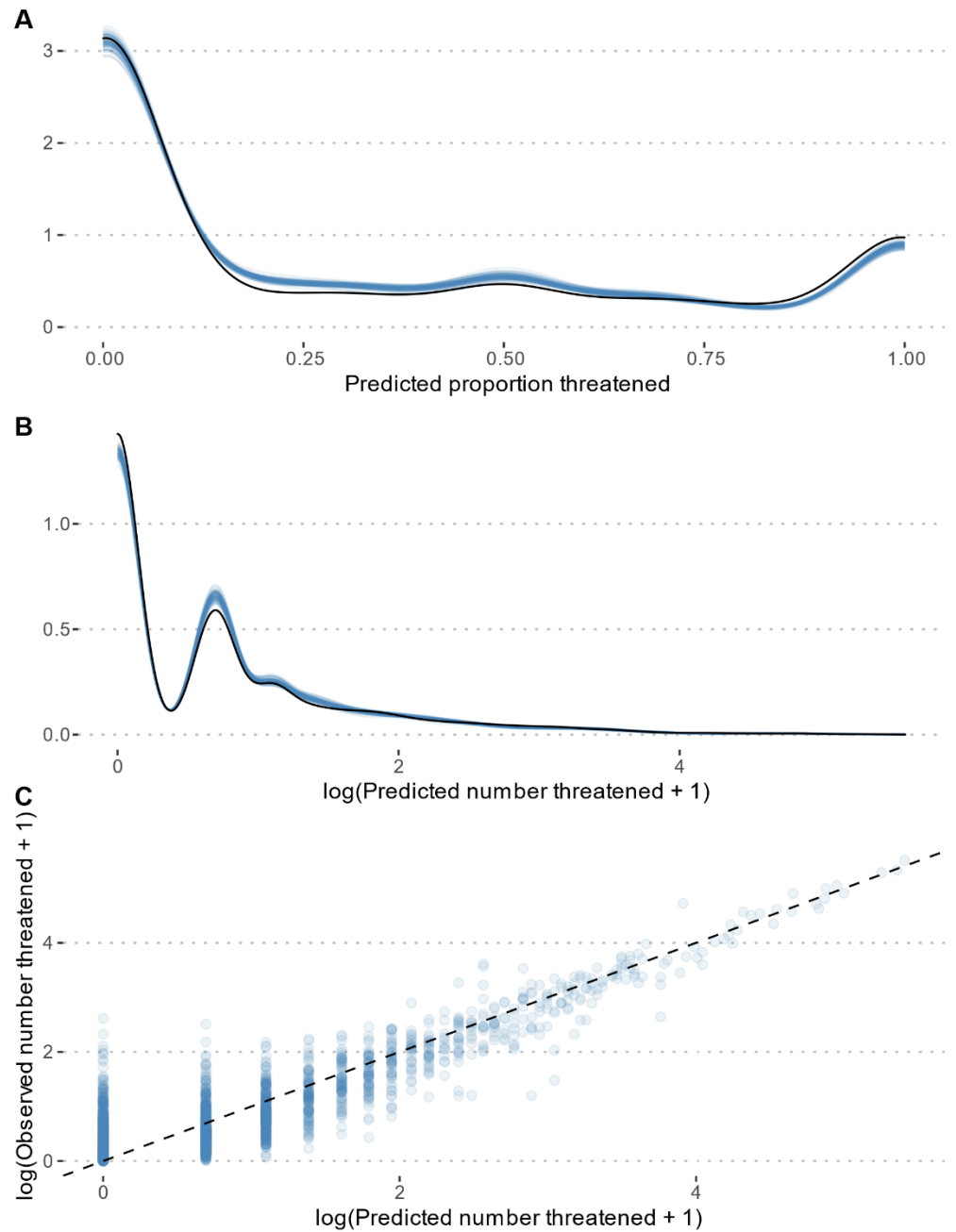
Fig S3 Posterior predictive checks to judge how well our model fits the data. Samples from the posterior predictive distribution of the fitted model are used to compare (A) the predicted distribution (blue lines) of values for the proportion of threatened species in each cell defined by combinations of all parameters in our model to the observed distribution (black line), as well as (B) the prediction distribution (blue lines) of the number of threatened species in each cell to

the observed distribution (black line). The samples were also used to compare (C) the predicted number of threatened species in each cell to the observed number.

Table S4 Mapping of Red List categories into two groups: "threatened" and "not threatened".

Red List Category	Our Category
Extinct	threatened
Extinct in the Wild	threatened
Critically Endangered	threatened
Endangered	threatened
Vulnerable	threatened
Near Threatened	not threatened
Least Concern	not threatened
Data Deficient	-

Table S5 Species counts after matching Red List assessments to WCVP.

Quantity	Number of species
Total accepted species in WCVP	340,024
Species with assessment other than DD	30,835
Threatened species	11,662

Table S6 Comparison of predictive performance of nested models by PSIS-LOO, showing that our final model incorporating all the variables of interest had the most favourable values of both the PSIS-LOO and ELPD.

Model	PSIS-LOO	ΔELPD	$SE_{\Delta ELPD}$
final	24,645.5 ± 172.7	0.00	0.00
continent	24,713.8 ± 172.5	-34.16	8.68
family	25,437.4 ± 170.6	-395.94	27.53
useful	26,957.1 ± 165.4	-1,155.79	47.41
endemic	27,559.2 ± 165.2	-1,456.87	53.24

Table S7 Estimated values of overall intercept (baseline) and fixed effects in our model.

Lifeform	Estimate	Lower 95% CI	Upper 95% CI
Intercept	0.199	0.095	0.392
endemic	6.186	0.260	6.649
useful	0.362	0.332	0.396

Table S8 Estimated values for each level of the lifeform random effect in our model.

Lifeform	Estimate	Lower 95% CI	Upper 95% CI
annual	0.836	0.396	1.384
epiphyte	0.565	0.257	0.907
herbaceous perennial	1.182	0.552	1.906
woody perennial	1.142	0.543	1.864

Table S9 Estimated values for different levels of the continent random effect.

Continent	Estimate	Lower 95% CI	Upper 95% CI
Africa	1.997	1.222	2.962
Antarctica	0.774	0.244	1.773
Asia-Temperate	1.227	0.748	1.848
Asia-Tropical	0.942	0.576	1.409
Australasia	0.239	0.139	0.365
Europe	1.470	0.892	2.216
Northern America	1.260	0.765	1.913
Pacific	1.203	0.730	1.820
Southern America	0.956	0.580	1.427
Worldwide	0.557	0.335	0.840

Table S10 Exponentiated parameter values for the 37 families with random effect that increased the odds of being threatened.

Family	Estimate	Lower 95% CI	Upper 95% CI
Aristolochiaceae	11.340	3.931	27.375
Taxaceae	9.072	3.876	18.239
Melastomataceae	7.097	2.145	17.522
Magnoliaceae	6.462	4.429	9.231
Loranthaceae	6.085	2.919	11.597
Heliconiaceae	5.558	1.767	14.220
Campanulaceae	5.538	3.707	7.973
Zamiaceae	4.862	3.370	6.797
Araucariaceae	4.663	2.145	9.172
Berberidaceae	4.569	1.863	9.619
Dipterocarpaceae	4.544	3.281	6.128
Begoniaceae	4.150	1.963	8.038
Fouquieriaceae	3.883	1.148	9.938
Cycadaceae	3.862	2.361	6.044
Gesneriaceae	3.671	2.195	5.903
Bromeliaceae	3.529	2.535	4.810
Marantaceae	3.526	1.410	7.517
Burmanniaceae	3.387	1.245	7.636
Cupressaceae	3.162	2.024	4.681
Geraniaceae	3.065	1.096	7.238
Thymelaeaceae	2.955	1.669	4.916
Cyclanthaceae	2.807	1.194	5.795
Orchidaceae	2.759	2.196	3.418
Calceolariaceae	2.536	1.140	5.028
Brassicaceae	2.263	1.401	3.461
Caryophyllaceae	2.096	1.206	3.457
Rosaceae	2.032	1.490	2.669
Solanaceae	1.983	1.234	3.036
Araceae	1.966	1.416	2.638
Acanthaceae	1.954	1.488	2.510
Pinaceae	1.841	1.264	2.569

Burseraceae	1.653	1.142	2.293
Myrtaceae	1.507	1.197	1.857
Arecaceae	1.461	1.112	1.878
Malvaceae	1.366	1.047	1.751
Zingiberaceae	1.351	1.036	1.717
Lauraceae	1.329	1.035	1.671

Table S11 Predicted proportion of threat for the 10 families with the most under-estimated level of threat.

Family	Accepted species	Assessed species	Threatened species	Observed proportion threatened	Predicted proportion threatened	Scaled residual
Myrtaceae	6079	1249	340	0.272	0.413	-4.665
Lauraceae	3331	647	192	0.297	0.409	-2.536
Fagaceae	959	283	38	0.134	0.254	-1.881
Poaceae	11556	674	155	0.230	0.286	-1.665
Barbeuiaceae	1	1	0	0.000	0.619	-1.625
Solanaceae	2762	292	67	0.229	0.380	-1.621
Cyperaceae	5603	666	115	0.173	0.229	-1.617
Asparagaceae	3188	290	120	0.414	0.492	-1.432
Melastomataceae	5719	542	268	0.494	0.718	-1.356
Stilbaceae	39	6	0	0.000	0.383	-1.350

Table S12 Predicted proportion of threat for the 10 families with the most over-estimated level of threat.

Family	Accepted species	Assessed species	Threatened species	Observed proportion threatened	Predicted proportion threatened	Scaled residual
Frankeniaceae	78	1	1	1.000	0.255	2.628
Balsaminaceae	1047	101	76	0.752	0.406	2.327
Polypodiaceae	4150	104	70	0.673	0.200	2.253
Asteliaceae	37	2	2	1.000	0.362	2.220
Arecaceae	2534	425	235	0.553	0.448	2.189
Calyceraceae	52	1	1	1.000	0.391	2.022
Dennstaedtiaceae	240	9	8	0.889	0.275	2.016
Triuridaceae	60	2	2	1.000	0.436	2.008
Ebenaceae	760	143	62	0.434	0.301	1.999
Dilleniaceae	527	44	18	0.409	0.192	1.949

Methods S4 Section 4 - Trends in extinction risk in plants and fungi.

As a first-step towards developing trend data for the global sampled Red List Index for plants we analysed two of the plant groups (Monocots and Legumes) from the original sample (Brummitt et al., 2015). From a total of 3,000 species, we selected all species with an occurrence in either Brazil or Madagascar (N = 400), two plant mega-diversity countries that featured in earlier versions of the SOTWP reports (RBG Kew, 2016, 2017).

Estimating genuine changes in Red List status

Habitat loss is the dominant threat to plants (Bachman et al., 2016). The increase in availability of remotely sensed products such as forest cover change (Hansen et al., 2013) has facilitated the analysis of habitat loss for extinction risk assessments (Buchanan et al., 2008; Tracewski et al., 2016). We adapted the approach of (Santini et al., 2019) to generate habitat models for each species of monocot and legume in our sample and explore how these changed over time. We calculated the area of habitat (AOH) following (Brooks et al., 2019) where AOH is the

intersection of the species' geographic range, the known elevation limits and habitat preferences (**Fig. S4**). The AOH is not a metric directly used in Red List assessments as it does not equate to either of the most commonly used spatial metrics: extent of occurrence (EOO) and area of occupancy (AOO) (IUCN Petitions and Standards Committee, 2019). However, the AOH can be used to estimate the upper bounds of the area of occupancy (AOO) if measured at the reference scale of 2 × 2 km (Brooks et al., 2019).

Calculating area of habitat (AOH)

To determine the geographic range of each species, occurrence data that formed the basis of the 2015 sampled RLI assessment were extracted from a BRAHMS database. Multiple resources were consulted, including Floras and monographs, as well as the Global Biodiversity Information Facility (GBIF.org 2020) to determine if additional occurrences were available, either due to new collection of the sampled species, or due to the increase in digitisation of existing collections. Georeferencing was carried out using gazetteers and Google Earth, and a simplified point radius method was used to assign coordinate precision (Wieczorek et al., 2004). Geographic ranges were generated by drawing a circular buffer around each occurrence point with a radius equal to the coordinate precision. Buffers were then dissolved to form a single polygon. In cases where an occurrence was reported from a known shape such as a forest reserve, the geometry of the boundary was recorded in well-known text (WKT) and was subsequently merged with the species range polygon.

Elevation ranges for each species were estimated based on interpretation of reported elevation from existing Red List assessments, elevation values reported in occurrence data (i.e. noted at time of collection), and elevation values summarised from a spatial query of a 1km resolution digital elevation model (DEM) at each occurrence location. Similarly, habitat preferences following the IUCN habitat classification scheme (IUCN Petitions and Standards Committee, 2019) were derived from existing Red List assessments and, where recorded, from occurrence data or other literature sources e.g. floras and monographs.

To calculate the area of habitat (AOH), the minimum and maximum elevation limits for each species were used to query suitable area of elevation from a 1km Digital Elevation Model raster (Jarvis et al., 2008). Area of habitat was determined using a 300m global land cover dataset (ESA, 2017). The ESA CCI dataset contains global land cover maps in raster format for the years 1992 – 2018. Habitat codes for species, based on the IUCN habitat classification scheme, were cross-walked to the ESA CCI land cover classification scheme following (Santini et al., 2019) and each annual land cover raster was reclassified to match the species' habitat preferences. To facilitate processing, the habitat layers were resampled to match the 1km DEM using a nearest neighbour method. The elevation and habitat rasters were combined and then masked to the species range polygon to return the area of habitat (AOH). The AOH raster maps were then used as the basis for the calculation of maximum AOO. The centroid of each 1km AOH cell was converted to a point and the AOOsimp function in the rCAT package (Moat & Bachman, 2017) was used to calculate AOO at the reference scale of 2km (IUCN Petitions and Standards Committee, 2019).

To estimate changes in a species' area of habitat (AOH) over time we calculated AOH for each year from 1992 – 2018 inclusive. To enable Red List Index values to be calculated for the years 2000, 2010 and 2020 we forecasted AOH values for the years 2019 and 2020 using the Holt Winters algorithm (R Core Team, 2019).

A modified approach to calculate AOH was employed for all species classified as Least Concern because the extremely large distributions and high number of occurrences could not be efficiently processed using the buffer method. For Least Concern species only, rather than buffered points, the species ranges were calculated as a convex hull of all occurrence data in the native range. All other steps in the approach were the same.

Examples of digital resources utilised to gather distribution, elevation and habitat data for species:

Plants of the World Online: http://plantsoftheworldonline.org/;

Flora do Brasil: http://floradobrasil.jbrj.gov.br/;

African Plant Database: https://www.ville-ge.ch/musinfo/bd/cjb/africa/recherche.php;

Catalogue of the Plants of Madagascar: http://legacy.tropicos.org/Project/Madagascar;

IUCN Red List https://www.iucnredlist.org/;

ColPlanta: http://colplanta.org/;

JSTOR Plants: https://plants.jstor.org/;

Muséum national d'Histoire naturelle:

https://science.mnhn.fr/institution/mnhn/collection/p/item/search/form?lang=en_US;

Useful Tropical Plants: http://tropical.theferns.info/;

SpeciesLink: http://www.splink.org.br/;

Herbário Virtual A. de Saint-Hilaire: http://hvsh.cria.org.br/

Calculating Red List Index values

To calculate the Red List Index values we filtered AOH values of all species for the years 2000, 2010 and 2020. We used the red package (Cardoso, 2017) to calculate all RLI values. Red List Index values for other groups are available on request from the IUCN Global Species Programme Red List Unit, Cambridge, UK. redlist@iucn.org.

Percentage change in AOH between 2000 and 2020 was calculated for each species (**Figure 4D**). Species were then ranked in order of the % change from those that gained AOH to those that lost AOH over this time period.

The approach outlined here is not equivalent to a full Red Listing methodology because only one aspect of threat and one criterion are being considered. The RLI values are also not appropriate to serve as country level indicators for Brazil and Madagascar as the samples were designed as part of a global assessment. However, as a subset of the global sample, they do give a preliminary view of changes in RLI value over time. The analysis of land cover data is challenging due to the need to crosswalk IUCN habitat codes to a global classification scheme derived from remotely sensed data. Further work is needed to examine the strengths and weaknesses of this approach.

All analysis and processing was carried out using the R statistical language version 3.6.3 (R Core Team, 2019). Scripts can be found at the following repository:

https://github.com/stevenpbachman/SOTWPF_extinction_SRLI

The ESA CCI land cover maps for the years 1992 – 2015 were downloaded in August 2019 and the 2016 – 2018 maps were downloaded in October 2019, from:
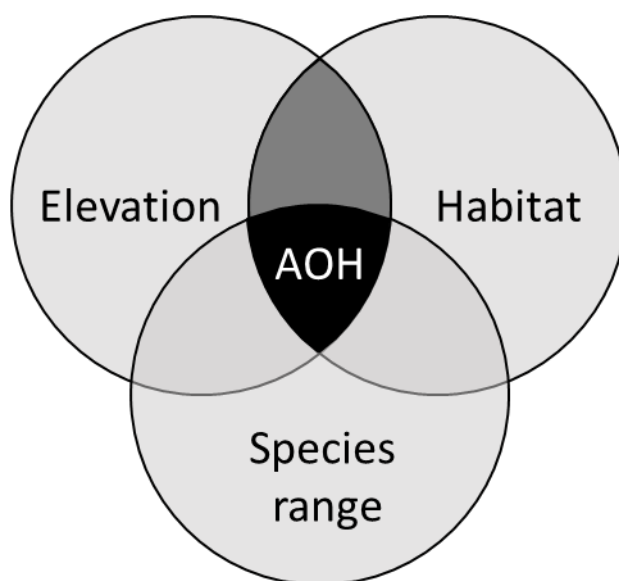
http://maps.elie.ucl.ac.be/CCI/viewer/download.php

Figure S4 Diagram illustrating how AOH is derived from the intersection of the elevation limits, habitat preferences and species range.

Table S13 Red List Index values – see Figure 3A.

Group	2000	2010	2020	# species
Madagascar	0.826506	0.8216867	0.8192771	83
Brazil	0.9280757	0.9280757	0.9280757	317
Total				*400*
Monocots	0.8696133	0.8685083	0.8674033	181
Legumes	0.9378995	0.9369863	0.9369863	219
Total				*400*

Methods S5. Examples of Reassessments.

Co-authors of this paper based in areas of relatively high plant diversity reported vastly differing attitudes to reassessing plant species already on the Red List but rarely mentioned them as a priority.

For example, reassessments of many species endemic to New Caledonia will not be prioritised simply because their original assessment is ten years old or more. Instead, they will be re-assessed when the larger group (e.g. family) of which they are a member is assessed as a whole (Gâteblé pers comm.).

In Cameroon, the first tropical African country to complete an assessment of all its plant species (Onana & Cheek, 2011), there is reported to be little or no interest in reassessments.

In Brazil, where the majority of plant species lack a formal assessment, re-assessments of some species have been undertaken in response to specific initiatives e.g. Global Tree Assessment (Rivers, 2017).

A strategic review of resources allocated to reassessments could help ensure that the value of investment in plant extinction risk assessments is not eroded through lack of update.

Methods S6 Methods used for Box 1.

We analysed the geographical distribution of the modern extinctions database (Humphreys et al. 2019) to reveal areas (botanical countries) which have experienced the greatest proportional excess of actual plant extinctions over expectations. We used the binomial test (Exact Binomial Test in R, $p < 0.001$) and assumed background extinction ~ 0.17%, following methods in Forest et al. 2018. Total numbers of species per botanical country were extracted from Plants of the World Online.

References

Bachman, S. P., Fernandez, E., Hargreaves, S., Nic Lughadha, E., Rivers, M., & Williams, E. (2016). *Extinction risk and threats to plants*. Retrieved from https://stateoftheworldsplants.com/2016/report/sotwp_2016.pdf

Bachman, S. P., Nic Lughadha, E. M., & Rivers, M. C. (2018). Quantifying progress toward a conservation assessment for all plants. *Conservation Biology, 32*(3), 516-524. https://doi.org/10.1111/cobi.13071

BGCI. (2020). ThreatSearch online database. from Botanic Gardens Conservation International Available at www.bgci.org/threat_search.php

Brooks, T. M., Pimm, S. L., Akçakaya, H. R., Buchanan, G. M., Butchart, S. H. M., Foden, W., . . . Rondinini, C. (2019). Measuring Terrestrial Area of Habitat (AOH) and Its Utility for the IUCN Red List. *Trends in Ecology & Evolution, 34*(11), 977-986. https://doi.org/10.1016/j.tree.2019.06.009

Brummitt, N., Bachman, S. P., Griffiths-Lee, J., Lutz, M., Moat, J. F., Farjon, A., . . . Nic Lughadha, E. M. (2015). Green Plants in the Red: A Baseline Global Assessment for the IUCN Sampled Red List Index for Plants. *PLoS ONE, 10*(8), e0135152. https://doi.org/10.1371/journal.pone.0135152

Brummitt, R. K. (2001). World Geographic Scheme for Recording Plant Distributions, Edition 2. Retrieved from http://www.tdwg.org/standards/109

Buchanan, G. M., Butchart, S. H. M., Dutson, G., Pilgrim, J. D., Steininger, M. K., Bishop, K. D., & Mayaux, P. (2008). Using remote sensing to inform conservation status assessment: Estimates of recent deforestation rates on New Britain and the impacts upon endemic birds. *Biological Conservation, 141*(1), 56-66. https://doi.org/https://doi.org/10.1016/j.biocon.2007.08.023

Bürkner, P.-C. (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software, 80*(1), 1-28. https://doi.org/10.18637/jss.v080.i01

Cardoso, P. (2017). red - an R package to facilitate species red list assessments according to the IUCN criteria. *Biodiversity Data Journal, 5*(e20530). https://doi.org/10.3897/BDJ.5.e20530

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., . . . Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software, 76*(1), 1-32. https://doi.org/10.18637/jss.v076.i01

Darrah, S. E., Bland, L. M., Bachman, S. P., Clubbe, C. P., & Trias-Blasi, A. (2017). Using coarse-scale species distribution data to predict extinction risk in plants. *Diversity and Distributions, 23*(4), 435-447. https://doi.org/10.1111/ddi.12532

Diazgranados, M., Allkin, B., Black, N., Cámara-Leret, R., Canteiro, C., Hargreaves, S., . . . Ulian, T. (2020). List of the World's Useful Plant species. *Knowledge Network for Biocomplexity, Unpublished*.

ESA. (2017). *Land Cover CCI Product User Guide Version 2. Tech. Rep.* Retrieved from maps.elie.ucl.ac.be/CCI/viewer/download/ESACCI-LC-Ph2-PUGv2_2.0.pdf

Fensham, R. J., Laffineur, B., Collingwood, T. D., Beech, E., Bell, S., Hopper, S. D., . . . White, M. (2020). Rarity or decline: Key concepts for the Red List of Australian eucalypts. *Biological Conservation, 243*, 108455. https://doi.org/https://doi.org/10.1016/j.biocon.2020.108455

Gallagher, R. V., Allen, S., Rivers, M. C., Allen, A. P., Butt, N., Keith, D., . . . Adams, V. M. (2020). Global shortfalls in extinction risk assessments for endemic flora. *bioRxiv*, 2020.2003.2012.984559. https://doi.org/10.1101/2020.03.12.984559

GBIF.org (2020). *GBIF Home Page*. Available from: https://www.gbif.org [13 January 2020].

Goettsch, B., Hilton-Taylor, C., Cruz-Piñón, G., Duffy, J. P., Frances, A., Hernández, H. M., . . . Gaston, K. J. (2015). High proportion of cactus species threatened with extinction. *Nature Plants, 1*(10), 15142. https://doi.org/10.1038/nplants.2015.142

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., . . . Townshend, J. R. G. (2013). High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science, 342*(6160), 850-853. https://doi.org/10.1126/science.1244693

Humphreys, A. M., Govaerts, R., Ficinski, S. Z., Nic Lughadha, E., & Vorontsova, M. S. (2019). Global dataset shows geography and life form predict modern plant extinction and rediscovery. *Nature Ecology & Evolution, 3*(7), 1043-1047. https://doi.org/10.1038/s41559-019-0906-2

IUCN. 2001. IUCN Red List Categories and Criteria: Version 3.1. IUCN Species Survival Commission. IUCN, Gland, Switzerland and Cambridge, U.K.

IUCN. (2020). The IUCN Red List of Threatened Species. Version 2020-1.   Retrieved from https://www.iucnredlist.org

IUCN Petitions and Standards Committee. (2019). *IUCN Standards and Petitions Committee. 2019. Guidelines for Using the IUCN Red List Categories and Criteria. Version 14*. Retrieved from http://www.iucnredlist.org/documents/RedListGuidelines.pdf

Jarvis, A., Reuter, H. I., Nelson, A., & Guevara, E. (2008). Hole-filled SRTM for the globe, Version 4. Retrieved from https://cgiarcsi.community/data/srtm-90m-digital-elevation-database-v4-1/

Moat, J., & Bachman, S. P. (2017). rCAT: Conservation Assessment Tools.

Onana, J. M., & Cheek, M. (2011). *Red Data Book of the Flowering Plants of Cameroon: IUCN Global Assessments*: Royal Botanic Gardens, Kew.

R Core Team. (2019). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

RBG Kew. (2016). *The State of the World's Plants Report, 2016*. London: Royal Botanic Gardens, Kew.

RBG Kew. (2017). *State of the World's Plants, 2017*. London: Royal Botanic Gardens, Kew.

Rivers, M. (2017). The Global Tree Assessment – Red Listing the World's Trees. *BGjournal, 14*(1), 16-19.

Santini, L., Butchart, S. H. M., Rondinini, C., Benítez-López, A., Hilbers, J. P., Schipper, A. M., . . . Huijbregts, M. A. J. (2019). Applying habitat and population-density models to land-cover time series to inform IUCN Red List assessments. *Conservation Biology, 33*(5), 1084-1093. https://doi.org/10.1111/cobi.13279

Tracewski, Ł., Butchart, S. H. M., Di Marco, M., Ficetola, G. F., Rondinini, C., Symes, A., . . . Buchanan, G. M. (2016). Toward quantification of the impact of 21st-century deforestation on the extinction risk of terrestrial vertebrates. *Conservation Biology, 30*(5), 1070-1079. https://doi.org/10.1111/cobi.12715

Wang, W., Rothschild, D., Goel, S., & Gelman, A. (2015). Forecasting elections with non-representative polls. *International Journal of Forecasting, 31*(3), 980-991. https://doi.org/10.1016/j.ijforecast.2014.06.001

WCVP. (2020). World Checklist of Vascular Plants, version 2.0. Available from Royal Botanic Gardens, Kew. Retrieved 01 January 2019 http://wcvp.science.kew.org/

Wieczorek, J., Guo, Q., & Hijmans, R. (2004). The point-radius method for georeferencing locality descriptions and calculating associated uncertainty. *International Journal of Geographical Information Science, 18*(8), 745-767. https://doi.org/10.1080/13658810412331280211