



# The effect of sampling effort and methodology on range size estimates of poorly-recorded species for IUCN Red List assessments

Charles J. Marsh<sup>1,2</sup> · Mindy M. Syfert<sup>1</sup> · Elina Aletrari<sup>1,3</sup> · Yoni Gavish<sup>2</sup> · William E. Kunin<sup>2</sup> · Neil Brummitt<sup>1</sup>

Received: 15 February 2022 / Revised: 21 December 2022 / Accepted: 30 December 2022 /  
Published online: 30 January 2023  
© The Author(s) 2023

## Abstract

Geographic range size is the most commonly implemented criterion of species' extinction risk used in IUCN Red List assessments, especially for poorly-recorded species. IUCN applies two contrasting range size measures to capture different facets of a species' distribution: Extent of Occurrence (EOO; Criterion B1) is the area bounding all known occurrences and is a proxy for the spatial autocorrelation of risk, while the Area of Occupancy (AOO; Criterion B2) is the area occupied within this boundary and is related to population size at finer grains. Various methods have been proposed to measure both EOO and AOO. We evaluate the impact of applying four methods for each of Criterion B1 and of B2, as well as key parameter choices, on the Red List status of 227 poorly-recorded neotropical pteridophyte species. Between 2 and 100% of species would be considered threatened depending on methodology. The minimum convex polygon method of estimating EOO was relatively robust to sampling effort for all but the least-recorded species. The IUCN-recommended method for estimating AOO of summing occupied 2 × 2 km grid cells was very strongly correlated with the total number of records. It is likely that only a small fraction of species can be adequately assessed using this method, and we recommend caution applying the method to poorly-recorded species in particular, where models predicting occupancy in unsampled areas (e.g. species distribution models) may provide more accurate assessments. It is vital that methodological information is retained with assessments, and comparisons should only be made between assessments utilising equivalent methods.

**Keywords** Conservation · Area of occupancy · AOO · Extent of occurrence · EOO

---

Communicated by Dirk Sven Schmeller.

---

This article belongs to the Topical Collection: Biodiversity protection and reserves.

---

✉ Charles J. Marsh  
charlie.marsh@mailbox.org

Extended author information available on the last page of the article

## Introduction

The IUCN Red List of Threatened Species (hereafter Red List) is the most widely applied index of conservation status, and a key tool for guiding conservation priorities and policy. It utilises a common currency, extinction risk, approximated by five complementary criteria that describe different aspects of a species' population and distribution (Mace et al. 2008).

The most commonly applied risk criterion, under Criterion B, is geographic range size, on the assumption that range size and extinction risk are correlated (Gaston and Fuller 2009) and is applied in conjunction with subcriteria of other high-risk factors such as population fragmentation. Two aspects of the range can be summarized (Gaston 1991, 1994), the Extent of Occurrence (EOO), which is the geographical spread of the distribution of the species, and the Area of Occupancy (AOO), which is the amount of that range that is occupied. The two measures are useful as they reflect processes occurring at different spatial scales. EOO reflects limits of broad physiological tolerances, and for the context of Red List assessments is a measure of the spatial autocorrelation of risk (rather than geographic range per se; IUCN 2019). AOO, however, is more closely correlated with habitat availability and population size (Hartley and Kunin 2003).

To estimate either range size measure requires only readily available occurrence data, and as a consequence geographic range size is the most widely used measure of extinction risk. For example, some 57% of species overall (Keith et al. 2018) and 59% of plant species assessed for the IUCN Sampled Red List Index (SRLI) for Plants (Brummitt et al. 2015a) use range size to assign their Red List status. This is particularly true for poorly-recorded species, i.e. those with few occurrence records and/or low spatial coverage of those records across their distribution. Such species comprise the majority of species on earth, especially in tropical regions (e.g. Brummitt et al. 2015a) for which data is generally insufficient for their assessment using other criteria (Bland et al. 2017).

In common with all Red List criteria, a set of thresholds are in place to translate EOO and AOO estimates to a predicted extinction risk, represented by the Red List threat categories. Range size thresholds were determined through a combination of trial-and-error and empirical testing (Mace et al. 2008) in order to produce a probability of extinction comparable with that of other criteria. The thresholds of EOO and AOO for risk classification are coupled by a ratio shared across all threat categories, set at 10:1; i.e., a species with an EOO of 100 km<sup>2</sup> is broadly deemed to have the same risk of extinction as one with an AOO of 10 km<sup>2</sup>. The thresholds are fixed regardless of taxon and methods used, which is important as nearly all methods for measuring range size are dependent upon the scale of parameters selected. Although parameter choice can have very large effects of range size estimates (Figs. S3, S4, S5), there are no clear processes for defining their values in the majority of cases. There has been further criticism of the thresholds as they were defined largely considering large vertebrates, and may be inappropriate for other taxa with very different home range sizes, population densities and dispersal abilities (Cardoso et al. 2011). Range shape is also not considered but may affect extinction risk as certain spatial configurations are more robust to edge effects and other threatening processes (Lucas et al. 2019).

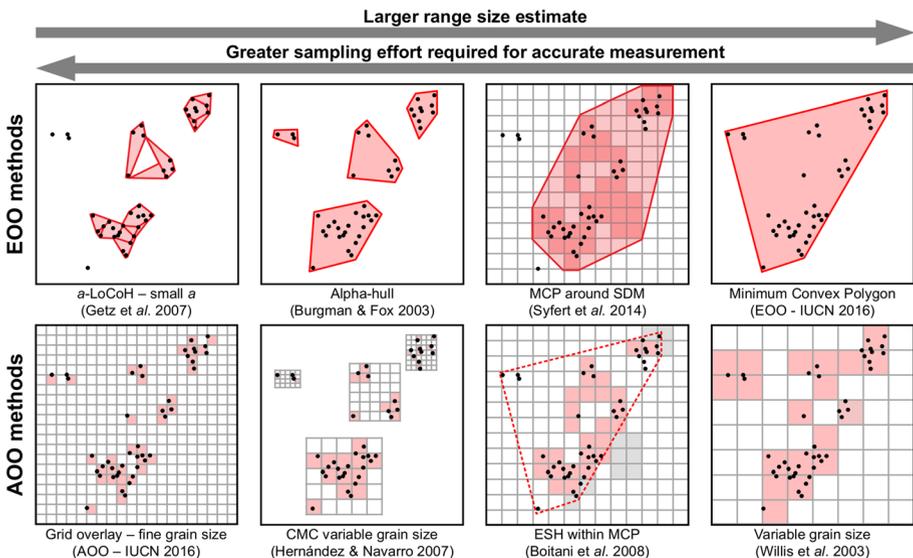
Furthermore, exactly how best to measure these two aspects of range sizes has also proven problematic. This has led to multiple methods being proposed, followed by multiple publications clarifying misconceptions and stressing the discrepancy between practice and IUCN guidelines (e.g. Collen et al. 2016). Here, we explore a range of methods proposed for assessing species under Criterion B, summarised below, and how they differ in how

they delimit the geographic ranges of species. In particular, we focus on their sensitivity to the spatial coverage of sampling effort, which could lead to incorrect estimates of extinction risk, particularly for poorly-recorded species.

**Extent of occurrence (EOO; criterion B1)**

EOO is applied to “measure the degree to which risks from threatening factors are spread spatially across the taxon’s geographical distribution” (IUCN 2019, pp. 48–49). The smaller the EOO, the more likely that all populations will undergo simultaneous extinctions (Collen et al. 2016). IUCN guidelines and Joppa et al. (2016) explicitly recommend that EOO should be measured using the minimum convex polygon (MCP), the smallest polygon that incorporates all known occurrences, regardless of whether populations are considered to be discontinuous. Accurately estimating the MCP therefore requires records that accurately define the extremities of the species’ range. Where data coverage is too poor to guarantee this, then species distribution models (SDMs) or similar have been suggested to predict the distribution edges (de Castro Pena et al. 2014; Syfert et al. 2014; Breiner et al. 2017).

Although “strongly discouraged” in the IUCN guidelines (IUCN 2019), multiple alternative methods have been developed that differ regarding how discontinuous populations are defined (Gaston and Fuller 2009; Joppa et al. 2016), such as alpha-hulls ( $\alpha$ -hull; Burgman and Fox 2003) and the local nearest neighbour convex hull (LoCoH; Getz et al. 2007). When the range is split into more separate populations then EOO estimates becomes smaller (Fig. S3a), but a greater sampling effort is required to provide an accurate assessment (Fig. 1) and therefore the greater the likelihood of underestimating EOO for poorly-recorded species. As the distinctions become very fine then it may result in the measure



**Fig. 1** Estimates of EOO (top row) and AOO (bottom row) differ depending on the method of measurement. Methods to the right produce larger range size estimates, whereas those to the left produce higher resolution estimates, but require greater sampling effort for an accurate assessment

more closely matching the definition of AOO rather than the original purpose of EOO, blurring the distinction between the two measures.

### Area of occupancy (AOO; criterion B2)

As AOO is closely related to population size when measured at sufficiently fine resolution (Gaston 2003), it is considered a general measure of robustness to threatening processes (Gaston and Fuller 2009), as well as an indicator of habitat specialisation, since species specialised for restricted habitats have a higher risk of extinction (IUCN 2019). Critically, AOO scales with the grain size with which it is measured (Kunin 1998), the occupancy-area relationship (OAR, Figs. S3, S4). The smaller the grain size, the smaller the AOO estimate, and the greater the sampling effort required to accurately estimate its value (Fig. 1), but the more closely this will equate to true population size (Hartley and Kunin 2003).

Due to this scale dependency of AOO, to allow for comparability between assessments, IUCN chose to standardise AOO measurements by requiring a grain size of  $2 \times 2$  km (IUCN 2019). This grain size was chosen not because it is a scale that best reflects extinction risks and threatening processes, but because it would be impossible to classify species as Critically Endangered if cells are larger than  $10 \text{ km}^2$ —a single occupied grid cell with cell width  $> 3.16$  km would exceed the  $10 \text{ km}^2$  threshold for AOO between the Critically Endangered and Endangered categories under criterion B2. In fact, the grain size at which AOO most closely correlates with true extinction risk depends primarily on the magnitude and frequency of the threat, with a grain size of around 1/10th the area affected by threat processes optimal in simulations (Keith et al. 2018). As different threats have different spatial patterns, Murray et al. (2017) argue that AOO should be measured at different grain sizes to reflect this.

IUCN recommends estimating AOO as the summed area of occupied grid cells overlaying known occurrences (hereafter ‘grid overlay’). Accurately measuring AOO therefore requires the species to be mapped across its entire range at the grain size selected—if there are sampling gaps where the species is present but not recorded then AOO will be underestimated. Therefore, AOO is likely to be particularly sensitive to incomplete data, especially at finer grains (Marsh et al. 2019). A minimum of 500 spatially independent records at  $2 \times 2$  km grain size would be required to exceed the threshold for the category Vulnerable (Rivers et al. 2011), and so poorly-recorded species will automatically appear to be threatened using this grain size if conditions for subcriteria are also met. Where a species is believed to be poorly-recorded, rather than genuinely range-restricted, a range of estimates are recommended, including the grid intersections with all suitable habitat, provided it is well-known, which largely equates to EOO (IUCN 2019).

Increasing the grain size at which AOO is measured may help alleviate issues with incomplete sampling (Marsh et al. 2019) but would then exceed the required grain size and so any estimate would need to be rescaled to  $2 \times 2$  km. The IUCN guidelines suggest a power law method for downscaling coarse-grain AOO estimates to finer grains (IUCN 2019), although more sophisticated methods are now available (Groom et al. 2018). Alternatively, multiple variations on utilising SDMs have been proposed to predict which unsampled cells are likely to be occupied (Harris and Pimm 2008; Jetz et al. 2008; Boitani et al. 2008; Marcer et al. 2013; Ocampo-Peñuela et al. 2016; Breiner et al. 2017).

Modifications of the grid overlay method include utilising circular buffers (Breiner and Bergamini 2018) and hexagons (Keith et al. 2018; Moat et al. 2018), and optimising the grid origin and orientation (Moat et al. 2018). Other alternative methods, such as

the variable grain method (Willis et al. 2003) and the cartographic method by conglomerates (CMC; Hernández and Navarro 2007), allow grain sizes to exceed the IUCN threshold (Fig. 1), applying a species-specific grain size related to the species' distribution characteristics. The impacts on Red List assignments of varying the grain size have been investigated (Hernández and Navarro 2007; Gaston and Fuller 2009; Roberts et al. 2016 and see Figs. S3, S4 in the Supplementary Material), but under current guidelines (IUCN 2019) any such method should subsequently rescale their estimates back to a  $2 \times 2$  km grain size.

## Study aim

Clearly there can be large discrepancies in estimates for both EOO and AOO between methods, or when multiple parameter values are applied for the same method. Inevitably there will be consequent problems comparing estimates of threat between species or across time (Collen et al. 2016). Furthermore, some methods will be more susceptible than others to incomplete data. In this study we compare estimates of EOO and AOO between methods, and for different parameter values within methods, and their consequence for Red List category assignments of a typical poorly-recorded taxon. We particularly focus on how data availability and incomplete sampling may influence the predictions of threatened status. Whereas most previous studies have focussed exclusively on either EOO or AOO (e.g. Joppa et al. 2016), here we compare both together, as would typically occur during the Red List assessment process.

## Methods

We investigated the application of IUCN Criterion B to 227 species of neotropical-endemic pteridophyte species with a minimum of three records, a subset of the species selected randomly as part of the Sampled Red List Index (SRLI) for Plants (Brummitt et al. 2015a). 15,177 records were harvested from museum data and online databases, reviewed and edited to ensure geo-referencing accuracy using range and bearing calculators, Google Earth and standard online gazetteers (see details within Brummitt et al. 2016), and re-projected onto a cylindrical equal-area projection. It is expected that, as most pteridophyte species are poorly known and sparsely recorded, for the majority there will be considerable sampling gaps in the data. As many also disperse widely we might expect large but sparse distributions. We generated EOO or AOO estimates for a range of methods for all 227 species using R 3.4.3 (R Core Team 2018). If a species could not be assessed by a given method due to computational limitations or insufficient data it was classified as Data Deficient (DD).

## Calculating EOO

We explored four methods of estimating EOO: (1) the MCP (IUCN 2019); (2) the MCP generated from the predicted presences of a species distribution model (SDM-MCP; Syfert et al. 2014); (3) the alpha-hull ( $\alpha$ -hull; Burgman and Fox 2003); and (4) the adaptive local nearest neighbour convex-hull ( $a$ -LoCoH; Getz and Wilmsers 2004; Getz et al. 2007).

The MCP, the smallest convex polygon that encompasses all occurrences, is recommended by IUCN as the best method of evaluating the spread of risk (IUCN 2019). As outlying occurrences may have a disproportionate effect on the extent of the MCP we also

explored the effect of sequentially removing these points in the supplementary material (Fig. S1). For the SDM-MCP, we also generated the MCP from the predicted presences of a species distribution model (Syfert et al. 2014) using MaxEnt (Version 3.3.3; Phillips et al. 2006). We controlled for geographical sampling bias using all geo-referenced plant occurrences available on GBIF. Ecological variables were selected based on a combination of correlation, principal components and cluster analyses to reduce multi-collinearity (see Syfert et al. 2014; Brummitt et al. 2016 for full details). Spatial extent was fixed as the rectangle encompassing all occurrences plus a 200 km buffer following VanDerWal et al. (2009). We applied a presence–absence threshold that maximised the sum of sensitivity and specificity, and then an MCP was drawn around the centres of cells with predicted presences. In this study we have retained all species' models for comparability, but in a real evaluation process it would be important to evaluate model performance, statistically and/or through expert assessments, and retain only models deemed sufficiently accurate.

For the  $\alpha$ -hull method, we created an alpha-hull by calculating the mean side length of a Delaunay triangulation. Connections of distance greater than the mean multiplied by the alpha value are then removed and the area of the subsequent polygons summed. We used the IUCN-recommended alpha value of 2 (IUCN 2019), but the effect of varying alpha is also explored in Fig. S3.

Finally, the  $\alpha$ -LoCoH method is one of the three methods of constructing the LoCoH (Getz et al. 2007). Each local convex hull around a root point is constructed so that the sum of point to root point distances is less than or equal to  $\alpha$ . Therefore convex hulls are smaller where there is a higher density of points and larger in areas with few records. We used the maximum inter-point distance for  $\alpha$  as recommended by Getz et al. (2007). Two alternative methods of generating the LoCoH,  $k$ -LoCoH and  $r$ -LoCoH, are explored in the supplementary material (Fig. S1) along with the scaling of their respective parameters (Figs. S1, S3).

## Calculating AOO

We explored four methods of estimating AOO: (1) summing the area of occupied cells of a grid with a fixed grain size overlaying all occurrence records (grid overlay; IUCN 2019); (2) a grid overlay method that uses a grain size related to the species' distribution characteristics inferred from the occurrence data (variable grain; Willis et al. 2003); (3) the area of predicted occupancies using ecologically suitable habitat models (ESH; Boitani et al. 2008); and (4) the area predicted using occupancy downscaling of the occupancy-area relationship (OAR; Kunin 1998).

The grid overlay method of summing the area of occupied cells of a grid overlaying the occurrence points is the IUCN-recommended method of calculating AOO, which mandates a fixed grain size of  $2 \times 2$  km (IUCN 2019). By contrast, the variable grain method uses a species-specific grain size, set as 1/10th of the largest pairwise inter-point distance (variable grain; Willis et al. 2003). Although cells can therefore exceed the IUCN-mandated size, we explored this method as it has been applied frequently during past assessments, and therefore we want to know if any comparisons between assessments using different grain sizes are valid.

Both these methods require species to be sampled across their entire distributions. However, we expect the majority of species not to have been mapped across their ranges at fine grains which may lead to underestimating AOO. Two further methods attempt to fill these sampling gaps. First, we predict occupancy in unsampled areas using ecologically suitable habitat models (ESH; Boitani et al. 2008; compare also with the Area of Habitat

(AOH) measure in Brooks et al. 2019). The potential range of the species was first defined by the MCP, and  $2 \times 2$  km cells within the MCP removed if they fell outside of the range of values for altitude, land cover (natural or anthropogenic) and available moisture (water deficit) outlined by the spatial locations of the species' records. The AOO estimate was the summed area of remaining cells (see Brummitt et al. 2016 for full details).

Finally, for the occupancy downscaling method, the first step is to create atlas data of the species distribution at a coarse scale that is large enough to reduce, or preferably eliminate, any sampling gaps, and calculate the proportion of cells occupied. The atlas data are aggregated further in order to create three or more estimates of occupancy ('upgraining'; Marsh et al. 2018). Models are then fitted to the OAR generated, and the fitted functions extrapolated down to fine grains. We created atlases by drawing a rectangle that incorporated all records of the species, using the recommendation by Marsh et al. (2019) to use the largest atlas scale that provides three grain sizes for modelling without exceeding the scale of saturation (the grain size at which all cells are occupied) or endemism (the grain size at which only a single cell is occupied). During upgraining, extents were standardised to that of the largest grain size using the method recommended by Groom et al. (2018), which retains all sampled cells from the original atlas ("All sampled" threshold). Various downscaling models are available (Azaele et al. 2012; Barwell et al. 2014); we used the 'simple ensemble' method that averages across the five simplest models that are the least computationally intensive but also most robust (Groom et al. 2018), as implemented in the 'downscale' R package (Marsh et al. 2018).

We also present results for the circular buffer (Breiner and Bergamini 2018) and the cartographic method by conglomerates (Hernández and Navarro 2007) in the supplementary material (Table S1, Fig. S2), as well as the effect of grain size on the grid overlay and variable grain method (Fig. S4).

## Analyses

We examined three aspects of the EOO and AOO estimates. First we looked for potential correlations between EOO and AOO estimates and the number of occurrence records, defined as the number of unique spatial records, excluding multiple individuals from the same location or repeat samples over time. A high correlation indicates that a measure is likely to be susceptible to under-sampling and the result generated will simply be an estimate of sampling effort. The danger is therefore in estimating a species as having a small range when in fact it is simply rarely recorded, whereas a lack of correlation could be considered as one estimate of robustness of the method.

To assess whether any correlation was really an occupancy effect, we repeated each method after subsampling records of the six species with  $> 250$  records. Subsamples were selected randomly using 5–95% of records in ten increments in log-space. We did not repeat this analysis for the SDM-MCP and ESH due to computational limitations, but the impacts of sample size on building niche models are well explored (e.g., van Proosdij et al. 2016).

Second, we examined differences between EOO estimates, and likewise between AOO estimates. Different methods may be employed during the evaluation of Criterion B1 or B2, even within taxa at different time periods, so we wish to examine how comparable the estimates from each method are given equivalent data. For each pairwise comparison we estimated the proportion of species that would be assigned to each Red List status if both

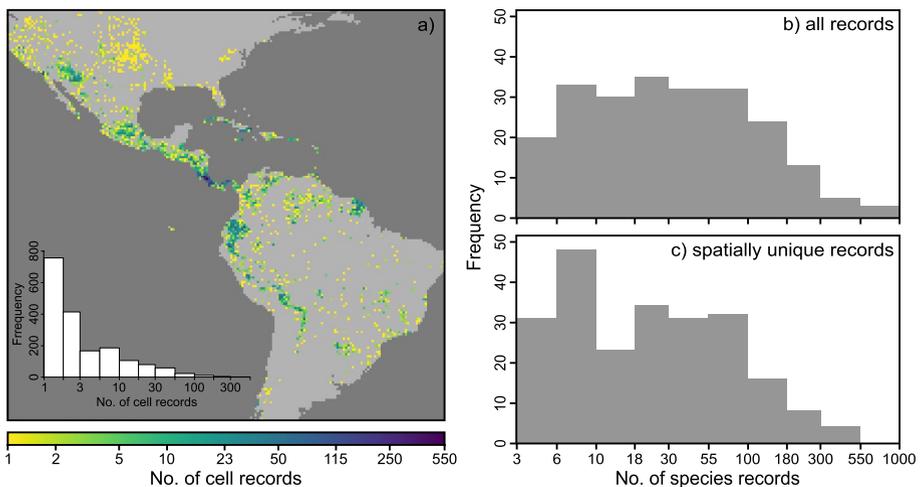
methods were used. Using the precautionary principle we kept the more conservative estimate, as recommended by IUCN (2019).

Finally, we compared all pairwise combinations of EOO and AOO measures. As the ratio between EOO and AOO is set at 10:1, successful combinations of measures should allow either EOO or AOO to produce the most threatened status where applicable. For example, a species distributed across an archipelago should be more threatened for AOO than EOO, whereas a narrow-ranging generalist should be classified as more threatened under EOO. Similarly, specific threats may be more apparent in one measure compared to the other. For example, climate change may result in reductions to EOO as range edges are altered, whereas certain patterns of habitat destruction would result in reductions in AOO with little to no change to EOO. We again evaluated the impact on the proportions of Red List status assignments for each comparison.

Assigning a Red List category in this way is only for illustrative purposes to evaluate respective methods. A full assessment would also consider other criteria under which species may be assessed, as well as the requirement that at least two of the three Criterion B subcriteria are fulfilled. Furthermore, although automated procedures can generate range estimates, in reality such outputs are reviewed by experts before assigning Red List categories.

## Results

Records were concentrated in a few well-recorded regions such as Costa Rica (Fig. 2a), whereas most of the region had no records even where pteridophytes are expected to be abundant, such as much of the Amazon basin. The majority of species were rarely recorded (Fig. 2b). 53 species out of the 227 had ten records or fewer, and 50% of species had fewer than 30 records. Of the better-recorded species, 45 had > 100 records and only 4 species had > 500 records. Moreover, 26.5% of the total records were



**Fig. 2** a Map of the 15,177 records of 227 pteridophyte species used in the study. Records were aggregated at a 50 × 50 km cell size using an equal area cylindrical projection. Histograms of b number of records and c number of spatially unique records per species

spatial replicates which provide no additional information for calculating EOO or AOO (Fig. 2c). We were unable to assess 7 (3.08%), 2 (0.88%) and 60 (26.43%) species for the  $\alpha$ -LoCoH, downscaling and ESH methods respectively, due to computational and data limitations (Table 1). These species were assigned as Data-Deficient (DD).

There were large discrepancies in EOO and AOO estimates between methods (Table 1) and within-method if different parameter values were applied (Fig. S3). EOO measures produced less conservative (i.e. less threatened) estimates than AOO, and all four EOO measures estimated most species to have non-threatened ratings. The  $\alpha$ -LoCoH and  $\alpha$ -hull methods estimated the highest proportion of threatened species (24.67% and 27.75% respectively), followed by MCP (13.66%) and finally SDM-MCP (2.20%). Of the AOO measures, the IUCN-recommended grid overlay method and downscaling assigned almost all species as threatened (100% and 98.24% respectively), while the variable grain (4.41%) and ESH (7.05%) methods produced similar results to the EOO methods. For the variable grain method, nearly all species used grain sizes far larger than  $2 \times 2$  km (min. cell width = 0.4 km, max. = 784.7 km, mean = 279.7 km, Fig. S6).

Relationships between the number of records and EOO measures varied mainly in their ability to estimate EOO for species with the fewest records (Fig. 3). The  $\alpha$ -hull method was most robust when randomly removing records within species (Fig. 4). More seriously, the grid overlay method for estimating AOO was particularly linearly correlated with the number of records (Pearson's  $r=0.98$ ). This persisted even when subsampling the most well-recorded species (Fig. 4), where correlations were even higher (Pearson's  $r=0.99876$ – $0.99998$ ), suggesting that the AOO estimates were essentially a measure of sample size rather than true AOO for these species. Removing records had little effect on the occupancy downscaling estimates unless sample sizes were low, but with much higher variability.

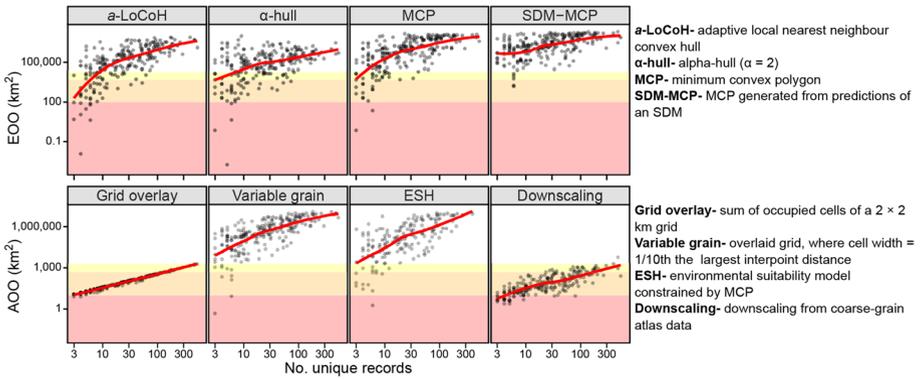
There were broad agreements in EOO estimates and the assignment of Red List categories between EOO methods except for the species with smaller ranges (Fig. 5). Estimates using the MCP approach were always equivalent to or larger than those using the  $\alpha$ -LoCoH or  $\alpha$ -hull methods.

There were much larger differences in AOO estimates between methods (Fig. 6). The grid overlay and occupancy downscaling estimates were similar. If any of these methods were employed 83.7–90.7% of species would be considered Endangered. By contrast, the majority of species would be considered non-threatened using only the ESH method.

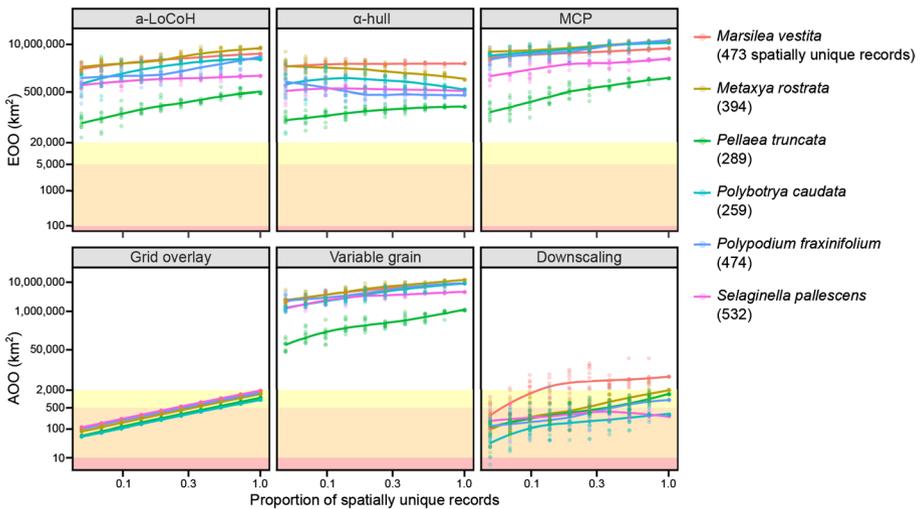
**Table 1** The number of species assigned to IUCN Red List status categories under four measures of AOO and EOO

	EOO				AOO			
	$\alpha$ -LoCoH	$\alpha$ -hull	MCP	SDM-MCP	Grid overlay	Variable grain	Downscaling	ESH
DD	7	0	0	0	0	0	2	60
CR	9	4	2	0	2	2	24	1
EN	25	32	13	2	207	3	171	10
VU	22	27	16	3	18	5	28	5
Non	164	164	196	222	0	217	2	151

DD Data Deficient, CR Critically Endangered, EN Endangered, VU Vulnerable, Non non-threatened

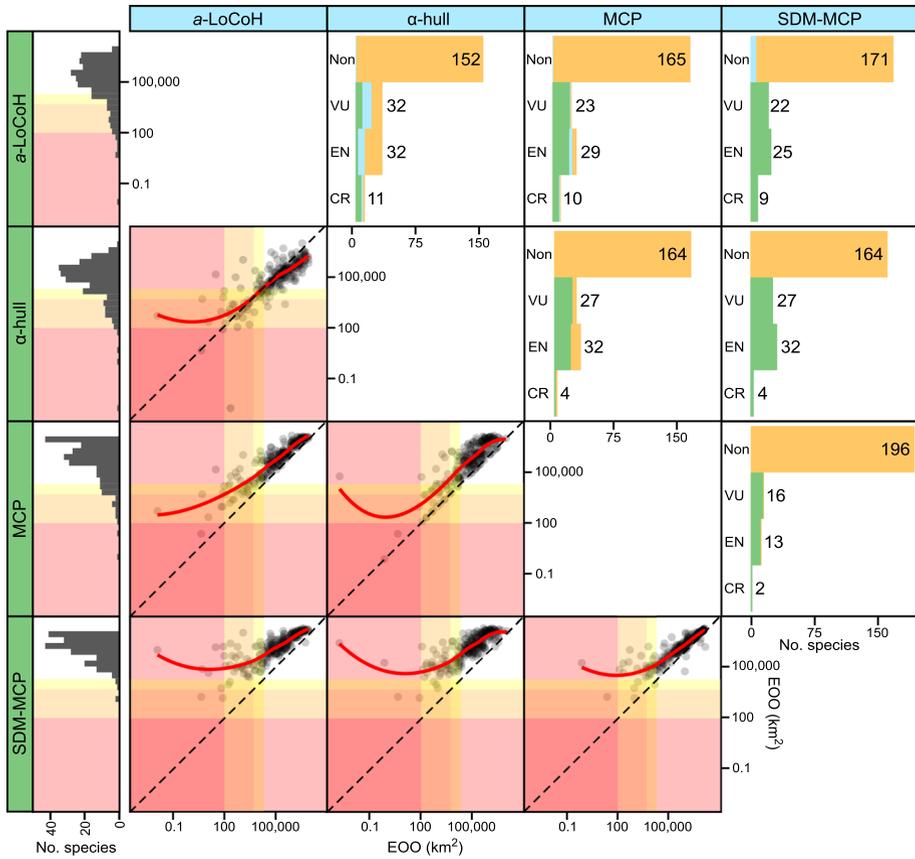


**Fig. 3** Correlations of the number of unique spatial records against estimates of EOO (top row) and AOO (bottom row) evaluated through differing methods described in the top corner. Red lines are loess smoothers. Shaded bands define the thresholds for a species to be assigned a Red List status of Critically Endangered (red), Endangered (orange), and Vulnerable (yellow) and non-threatened (non; white)



**Fig. 4** Correlations of the proportion of randomly retained records against estimates of EOO (top row) and AOO (bottom row) for six well recorded pteridophyte species. Coloured lines are loess smoothers. Shaded bands define the thresholds for a species to be assigned a Red List status of Critically Endangered (CR; red), Endangered (EN; orange), Vulnerable (VU; yellow) and non-threatened (non; white)

There were certain EOO-AOO combinations where a single method always estimated a more threatened category than the other. For example, for the recommended methods of a grid overlay and MCP, nearly every single species would be classified as threatened under AOO only (Figs. 7, S7). This results in large differences in the proportion of threatened species that would be assigned. Broadly, when AOO was measured using the grid overlay or downscaling methods, nearly all species would be considered threatened and EOO would rarely be utilised except for assigning species as Critically Endangered. If AOO was

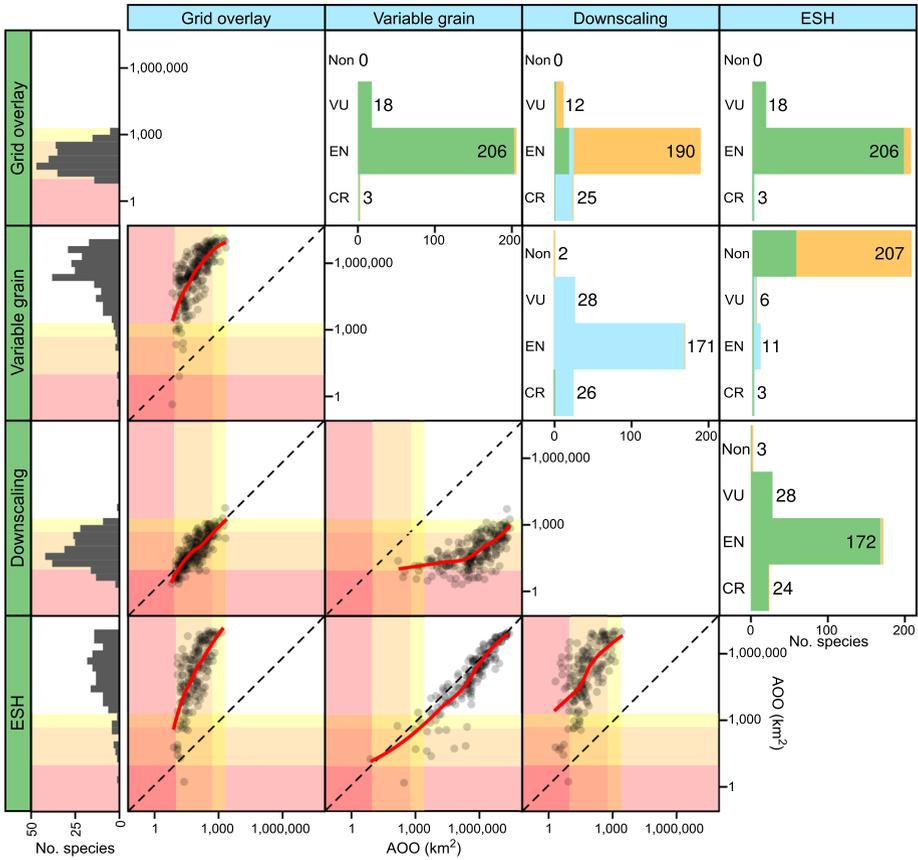


**Fig. 5** Estimates of EOO from pairwise combinations of four possible methods. The lower triangle plots are one method assessed against the other; red lines are loess smoothers and the dashed line is the 1:1 relationship. Histograms for each method are presented far left. Shaded bands define the thresholds for a species to be assigned a Red List status of Critically Endangered (CR; red), Endangered (EN; orange), Vulnerable (VU; yellow) and non-threatened (Non; white). Upper triangles are the number of species assigned to Red List categories if both EOO methods were used, where the most threatened (most conservative) estimate is retained (blue=assigned using the x-axis method only, green=assigned using the y-axis method only, orange=both methods assigned the same status)

measured using the variable grain or ESH methods then the majority of species would be considered non-threatened, and in these cases species considered under a threatened category were generally assigned on the basis of EOO.

## Discussion

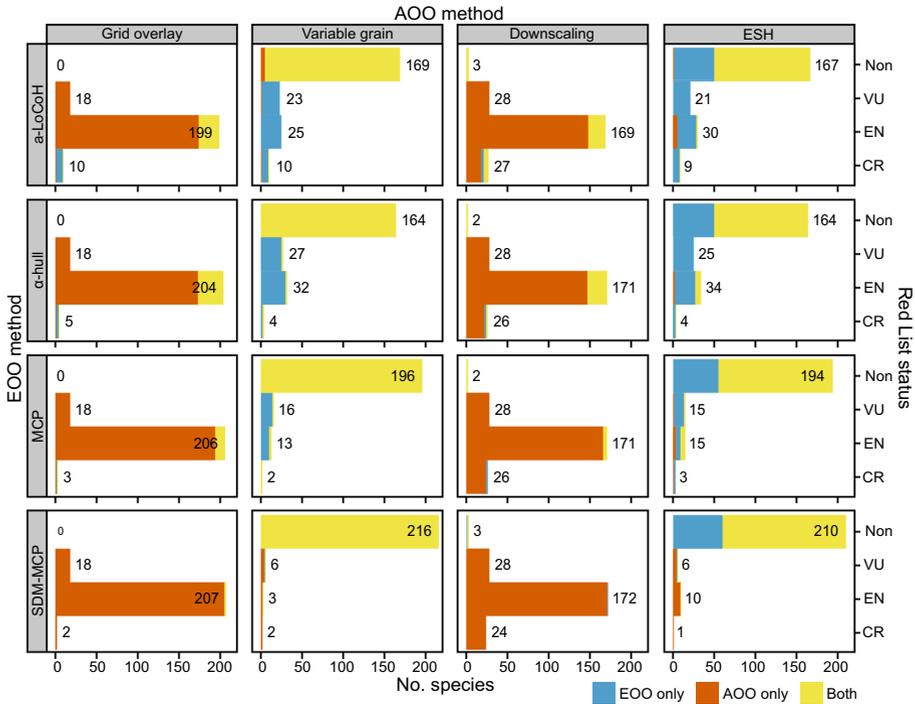
The two measures of range size for assessing extinction risk, Extent of Occurrence (EOO) and Area of Occupancy (AOO), require only spatial records. As a consequence, the majority of species are assigned a Red List status using one or both of these range measures (Gaston and Fuller 2009; Brummitt et al. 2015a; Keith et al. 2018). IUCN strongly



**Fig. 6** Estimates of AOO from pairwise combinations of four possible methods. The lower triangle plots are one method assessed against the other; red lines are loess smoothers and the dashed line is the 1:1 relationship. Histograms for each method are presented far left. Shaded bands define the thresholds for a species to be assigned a Red List status of Critically Endangered (CR; red), Endangered (EN; orange), Vulnerable (VU; yellow) and non-threatened (Non; white). The upper triangle of plots show the number of species assigned to Red List categories if both AOO methods were used, where the most threatened (most conservative) estimate is retained (blue = assigned using the x-axis method only, green = assigned using the y-axis method only, orange = both methods assigned the same status)

recommends using the MCP for EOO and mandates a grid of 2 × 2 km for measuring AOO (IUCN 2019). Despite IUCN recommendations, drawbacks with each metric have led to the proposal and implementation of many alternative methods of measurement. The majority of these alternative methods can be attributed to two intentions: accounting for discontinuities between populations, particularly with regards to EOO, and accounting for incomplete sampling, where AOO is most affected.

The treatment of discontinuities when measuring EOO relates to what constitutes the ‘range’ of a species (*a*-LoCoH and  $\alpha$ -hull in this study). However, it should be remembered that for Red List assessments, EOO is not a measure of range size per se, but a measure of the spread of potential threats (Mace et al. 2008; Collen et al. 2016; IUCN 2019). Confusion over this intention has led to proposed measures for EOO that can be more accurately described as AOO measurements (e.g. Harris and Pimm 2008; Ocampo-Peñuela et al.



**Fig. 7** Proportion of species in each Red List category assigned through EOO only (blue), AOO only (orange) or both AOO and EOO (yellow) using pairs of combination of measurement. CR Critically Endangered, EN Endangered, VU Vulnerable, Non non-threatened

2016). Similarly, the separation of populations is carried out by selecting scaling parameters (Fig. S3) that discriminate finer and finer divisions, and therefore move the measurement away from the original purpose of EOO and instead approach the definition of AOO.

EOO was relatively robust to data quantity, but AOO was particularly sensitive to incomplete sampling coverage (Figs. 3, 4). If using the recommended grid overlay method, very few species will be mapped accurately across their entire range at a 2x2 km grain size, even for the best-known taxa. Furthermore, a minimum of 500 spatially-unique records are required to assign a non-threatened status. For poorly-recorded taxa such as neotropical pteridophytes, representative of the limited amount of information known for the majority of species, estimates of AOO generated from the grid overlay (and circular buffer methods) simply reflect the number of records (Figs. 3, 4, S2) with an extremely high correlation (0.99893–0.99998). This resulted in every species investigated being considered as threatened using the IUCN-recommended grid overlay method.

Many proposed methodologies are attempts to overcome these sampling gaps. For example, we can delimit areas falling within niche tolerances of a species (ESH in this study), although Brooks et al. (2019) have emphasised that such techniques generate a measure of the extent of habitat (EOH), and therefore are not estimates of either AOO or EOO, but instead lie somewhere in-between. Occupancy downscaling attempts to eliminate sampling gaps by aggregating records at coarser scales, especially if certain properties of sampling bias and species’ distribution characteristics can be incorporated (Marsh et al.

2019), although sampling gaps appeared too severe to be overcome in the species examined here.

Alternatively, methods that predict occupancy in unsampled areas, such as SDMs, are perhaps better able to overcome sampling gaps successfully whilst maintaining the grain sizes required by IUCN. However, generating accurate SDMs is complicated by methodological issues of their own (Araújo and Guisan 2006). Some authors also stress that AOO cannot be inferred from presence-background models of the type most frequently applied, as without absence data the models will not produce true probabilities of occupancy (Guilera-Arroita et al. 2015). Furthermore, SDMs are prone to over-predicting occupancy as they fail to include biological interactions, historical factors and other spatial processes such as dispersal limitation that can prevent species occupancy despite favourable environmental conditions (Araújo and Guisan 2006).

It should also be noted that for real assessments it is likely that much larger quantities of data than were available here would be required to build accurate models (Valavi et al. 2022). Recent methodological advancements make modelling poorly-recorded species more achievable (Jeliazkov et al. 2022), but it is unclear what consequences they have for Red List assessments. Generating models for measuring EOO is likely to be relatively robust to data availability, and Syfert et al. (2014) found that around 20 points were required to generate an EOO estimate similar to when much larger quantities of data were available. Building models for estimating AOO, however, are likely to require greater quantities of presence–absence data, as well as requiring minimising or controlling for spatial and environmental biases in the samples.

Furthermore, the data requirements for generating accurate estimates of model performance are probably even greater than for model building. This may especially be the case if testing data is environmentally or geographically biased, in which case expert assessments may be the only feasible approach to evaluate SDM predictions. Regardless of method, it is important that any range size estimate using SDM predictions should always assess model performance, and where sufficiently reliable models can not be generated then the species should be classed as Data Deficient for the criterion.

It is also likely that data limitations for calculating EOO and AOO are equally or more limiting for the Criterion B subcriteria, for which a species must fulfil at least two to be assigned a threatened status. In the vast majority of cases, we generally won't have a time-series of occurrences for poorly-recorded species. Consequently we will be unable to measure fluctuations (subcriterion c) or observe or project continuing declines in EOO (subcriterion b(i)), AOO (b(ii)), or the number of subpopulation locations (b(iv)) or mature individuals (b(v)). Instead most assessments will most likely rely on estimates or inferred declines in the area, extent or quality of habitat (b(iii)). Similarly, it may be difficult to ascertain if the species has a severely fragmented or a limited number of locations (subcriterion a) from a handful of occurrence points and no data on dispersal ability or metapopulation dynamics, and so likewise any assessment of subcriterion a will be estimated or inferred.

## Ensuring comparability between range estimates

An important function of Red Listing is comparing estimates among species and taxa, assessing changes in status over time, and evaluating whether conservation measures are succeeding (IUCN's Green List of species). However, we have shown that simply applying different methodologies or parameters may result in drastic changes in perceived status

without any true differences in range size or input data. To identify genuine changes it is therefore vital not only to compare separate assessments of the same species under the same criteria (Brummitt et al. 2015b), but furthermore only to compare estimates of range size utilising the same method of measurement (Collen et al. 2016). As long as consistent methodology and parameterisation is maintained, this would also assist the assessment of subcriterion b as there will be an observed time-series of estimates of EOO (for subcriterion b(i)) or AOO (b(ii)). This requires detailed information on the methods and parameters applied to be recorded during any assessment, which must remain associated with that assessment when it is retrieved.

A further problem is that as records accumulate over time so range size can only remain the same or increase. To identify decreases in range size, existing records have to be removed, for example from sites deemed no longer occupied through re-surveying, inference from remote sensing data, or by excluding records that exceed an (undefined) age threshold. Methods that use data at larger scales, such as downscaling, should be applied with caution as occupancy changes will be manifested over longer time periods within large cells, because extinctions at coarse scales will occur more slowly than localised extinctions at finer grains (Hartley and Kunin 2003), although there is evidence that declining species can show characteristically steep occupancy-area relationships (Wilson et al. 2004). Although, SDM-based methods tend to over-predict prevalence, and thus shift IUCN classification towards less threatened categories (Table 1), considerable potential lies in more realistic presence-absence SDMs that account for additional processes, such as biotic interactions (Gavish et al. 2017) and dispersal limitation, and thus better map the realized distribution rather than the potential one.

## Measuring range size for Red List assessments

The IUCN Red List is the most widely used index of conservation status, with tremendous influence on funding, policy and the international conservation agenda. Criterion B (Geographic Range) is the most widely applied criterion, largely because it is easy to generate *an* estimate, only requiring easily-harvestable occurrence data. However, we stress that it requires a particularly large quantity of data to generate *an accurate* measurement.

We show that for 227 neotropical pteridophyte species, poorly-recorded taxa with incomplete sampling across their distributions but typical of the majority of world's biodiversity, the estimate of EOO and AOO is so strongly affected by method of measurement that it can obscure actual differences in extinction risk. In particular, whereas the measurement of EOO was relatively robust to sampling effort, AOO was especially sensitive to incomplete sampling coverage. In our case, the IUCN-required method of a 2×2 km grid overlay primarily reflected only a measure of the number of records (Figs. 3, 4). We argue that for the vast majority of species, which have only a small quantity of records which exhibit some degree of spatial bias, AOO should not be measured by the IUCN grid overlay method without attempting to correct for incomplete sampling, unless additional detailed ecological information confirms that this species has a naturally restricted range. If sampling gaps cannot be adequately corrected then no AOO estimate should be generated.

Furthermore, empirical or theoretical studies linking measures of EOO or AOO with true probability of extinction are still lacking, and the derivation of thresholds used to assign threatened categories needs greater grounding in theory and observation. Further studies explicitly aiming to link such methods with true probability of extinction must also set thresholds appropriate to that method. The great challenge is to maintain assessments

that are comparable across time whilst also staying up-to-date in an ever-advancing scientific world. Methodological decisions set early on may not reflect best-practice guidelines identified when new scientific methods emerge. The incremental refinement of recommended guidelines plays a crucial role in this context. However, if evidence of fundamental limitations of thresholds or methods mount, while great care should be taken, we should not avoid more extensive amendments for the sake of historical consistency.

## Supporting Information

Additional results figures and analyses are available in the supplementary material, including additional EOO (k-LoCoH, r-LoCoH) and AOO (circular buffer and CMC) methods and explorations of scale-dependencies for the MCP,  $\alpha$ -hull, k-LoCoH, r-LoCoH, a-LoCoH, grid overlay, variable grain and CMC methods.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s10531-023-02543-9>.

**Acknowledgements** We thank Stefan Pinkert for useful comments on earlier drafts.

**Author contributions** All authors conceived the idea and designed the methodology. NB, MMS and EA harvested and georeferenced the pteridophyte records and carried out the SDM and ESH modelling. CJM carried out the analysis and led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

**Funding** This work was financed by the EU BON project ([www.eubon.eu](http://www.eubon.eu)) that is a 7th Framework Programme project funded by the European Union under Contract No. 308454.

**Data availability** R code for calculating EOO and AOO for each method, along with a shapefile of the subset of occurrence points used in the study is freely available at <https://github.com/charliem2003/AOOvEOO>.

## Declarations

**Conflict of interest** The authors have no relevant financial or non-financial interests to disclose.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Araújo MB, Guisan A (2006) Five (or so) challenges for species distribution modelling. *J Biogeogr* 33:1677–1688. <https://doi.org/10.1111/j.1365-2699.2006.01584.x>
- Azaele S, Cornell SJ, Kunin WE (2012) Downscaling species occupancy from coarse spatial scales. *Ecol Appl* 22:1004–1014. <https://doi.org/10.1890/11-0536.1>
- Barwell LJ, Azaele S, Kunin WE, Isaac NJB (2014) Can coarse-grain patterns in insect atlas data predict local occupancy? *Divers Distrib* 20:895–907. <https://doi.org/10.1111/ddi.12203>

- Bland LM, Bielby J, Kearney S et al (2017) Toward reassessing data-deficient species. *Conserv Biol* 31:531–539. <https://doi.org/10.1111/cobi.12850>
- Boitani L, Sinibaldi I, Corsi F et al (2008) Distribution of medium- to large-sized African mammals based on habitat suitability models. *Biodivers Conserv* 17:605–621. <https://doi.org/10.1007/s10531-007-9285-0>
- Breiner FT, Bergamini A (2018) Improving the estimation of area of occupancy for IUCN Red List assessments by using a circular buffer approach. *Biodivers Conserv* 27:2443–2448. <https://doi.org/10.1007/s10531-018-1555-5>
- Breiner FT, Guisan A, Nobis MP, Bergamini A (2017) Including environmental niche information to improve IUCN Red List assessments. *Divers Distrib* 23:484–495. <https://doi.org/10.1111/ddi.12545>
- Brooks TM, Pimm SL, Akçakaya HR et al (2019) Measuring terrestrial area of habitat (AOH) and its utility for the IUCN Red List. *Trends Ecol Evol* 34:977–986. <https://doi.org/10.1016/j.tree.2019.06.009>
- Brummitt N, Bachman SP, Aletrari E et al (2015a) The Sampled Red List Index for Plants, phase II: ground-truthing specimen-based conservation assessments. *Philos Trans R Soc B* 370:20140015. <https://doi.org/10.1098/rstb.2014.0015>
- Brummitt NA, Bachman SP, Griffiths-Lee J et al (2015b) Green plants in the red: a baseline global assessment for the IUCN Sampled Red List Index for Plants. *PLoS ONE* 10:e0135152. <https://doi.org/10.1371/journal.pone.0135152>
- Brummitt N, Aletrari E, Syfert MM, Mulligan M (2016) Where are threatened ferns found? Global conservation priorities for pteridophytes. *J Syst Evol* 54:604–616. <https://doi.org/10.1111/jse.12224>
- Burgman MA, Fox JC (2003) Bias in species range estimates from minimum convex polygons: implications for conservation and options for improved planning. *Anim Conserv* 6:19–28. <https://doi.org/10.1017/S1367943003003044>
- Cardoso P, Borges PAV, Triantis KA et al (2011) Adapting the IUCN Red List criteria for invertebrates. *Biol Conserv* 144:2432–2440. <https://doi.org/10.1016/j.biocon.2011.06.020>
- Collen B, Dulvy NK, Gaston KJ et al (2016) Clarifying misconceptions of extinction risk assessment with the IUCN Red List. *Biol Lett* 12:20150843. <https://doi.org/10.1098/rsbl.2015.0843>
- de Castro Pena JC, Kamino LHY, Rodrigues M et al (2014) Assessing the conservation status of species with limited available data and disjunct distribution. *Biol Conserv* 170:130–136. <https://doi.org/10.1016/j.biocon.2013.12.015>
- Gaston KJ (1991) How large is a species' geographic range? *Oikos* 61:434–438. <https://doi.org/10.2307/3545251>
- Gaston KJ (1994) Measuring geographic range sizes. *Ecography* 17:198–205
- Gaston KJ (2003) The structure and dynamics of geographic ranges. Oxford University Press, Oxford
- Gaston KJ, Fuller RA (2009) The sizes of species' geographic ranges. *J Appl Ecol* 46:1–9. <https://doi.org/10.1111/j.1365-2664.2008.01596.x>
- Gavish Y, Marsh CJ, Kueemleren M et al (2017) Accounting for biotic interactions through alpha-diversity constraints in stacked species distribution models. *Methods Ecol Evol* 8:1092–1102. <https://doi.org/10.1111/2041-210X.12731>
- Getz WM, Wilmers CC (2004) A local nearest-neighbor convex-hull construction of home ranges and utilization distributions. *Ecography* 27:489–505. <https://doi.org/10.1111/j.0906-7590.2004.03835.x>
- Getz WM, Fortmann-Roe S, Cross PC et al (2007) LoCoH: nonparametric kernel methods for constructing home ranges and utilization distributions. *PLoS ONE* 2:e207. <https://doi.org/10.1371/journal.pone.0000207>
- Groom QJ, Marsh CJ, Gavish Y, Kunin WE (2018) How to predict fine resolution occupancy from coarse occupancy data. *Methods Ecol Evol* 9:2273–2284. <https://doi.org/10.1111/2041-210X.13078>
- Guillera-Arroita G, Lahoz-Monfort JJ, Elith J et al (2015) Is my species distribution model fit for purpose? Matching data and models to applications. *Glob Ecol Biogeogr* 24:276–292. <https://doi.org/10.1111/geb.12268>
- Harris G, Pimm SL (2008) Range size and extinction risk in forest birds. *Conserv Biol* 22:163–171. <https://doi.org/10.1111/j.1523-1739.2007.00798.x>
- Hartley S, Kunin WE (2003) Scale dependency of rarity, extinction risk, and conservation priority. *Conserv Biol* 17:1559–1570. <https://doi.org/10.1111/j.1523-1739.2003.00015.x>
- Hernández HM, Navarro M (2007) A new method to estimate areas of occupancy using herbarium data. *Biodivers Conserv* 16:2457–2470. <https://doi.org/10.1007/s10531-006-9134-6>
- IUCN (2019) Guidelines for using the IUCN Red List categories and criteria. Prepared by the Standards and Petitions Subcommittee
- Jeliazkov A, Gavish Y, Marsh CJ et al (2022) Sampling and modelling rare species: conceptual guidelines for the neglected majority. *Glob Change Biol* 28:3754–3777

- Jetz W, Sekercioglu CH, Watson JEM (2008) Ecological correlates and conservation implications of over-estimating species geographic ranges. *Conserv Biol* 22:110–119. <https://doi.org/10.1111/j.1523-1739.2007.00847.x>
- Joppa LN, Butchart SHM, Hoffmann M et al (2016) Impact of alternative metrics on estimates of extent of occurrence for extinction risk assessment. *Conserv Biol* 30:362–370. <https://doi.org/10.1111/cobi.12591>
- Keith DA, Akçakaya HR, Murray NJ (2018) Scaling range sizes to threats for robust predictions of risks to biodiversity. *Conserv Biol* 32:322–332. <https://doi.org/10.1111/cobi.12988>
- Kunin WE (1998) Extrapolating species abundance across spatial scales. *Science* 281:1513–1515. <https://doi.org/10.1126/science.281.5382.1513>
- Lucas PM, González-Suárez M, Revilla E (2019) Range area matters, and so does spatial configuration: predicting conservation status in vertebrates. *Ecography* 42:1103–1114. <https://doi.org/10.1111/ecog.03865>
- Mace GM, Collar NJ, Gaston KJ et al (2008) Quantification of extinction risk: IUCN's system for classifying threatened species. *Conserv Biol* 22:1424–1442. <https://doi.org/10.1111/j.1523-1739.2008.01044.x>
- Marcer A, Sáez L, Molowny-Horas R et al (2013) Using species distribution modelling to disentangle realised versus potential distributions for rare species conservation. *Biol Conserv* 166:221–230. <https://doi.org/10.1016/j.biocon.2013.07.001>
- Marsh CJ, Barwell LJ, Gavish Y, Kunin WE (2018) downscale: an R package for downscaling species occupancy from coarse-grain data to predict occupancy at fine-grain sizes. *J Stat Softw* 86:1–20. <https://doi.org/10.18637/jss.v086.c03>
- Marsh CJ, Gavish Y, Kunin WE, Brummitt NA (2019) Mind the gap: can downscaling area of occupancy overcome sampling gaps when assessing IUCN Red List status? *Divers Distrib* 025:1832–1845. <https://doi.org/10.1111/ddi.12983>
- Moat J, Bachman SP, Field R, Boyd DS (2018) Refining area of occupancy to address the modifiable areal unit problem in ecology and conservation. *Conserv Biol* 32:1278–1289. <https://doi.org/10.1111/cobi.13139>
- Murray NJ, Keith DA, Bland LM et al (2017) The use of range size to assess risks to biodiversity from stochastic threats. *Divers Distrib* 23:474–483. <https://doi.org/10.1111/ddi.12533>
- Ocampo-Peñuela N, Jenkins CN, Vijay V et al (2016) Incorporating explicit geospatial data shows more species at risk of extinction than the current Red List. *Sci Adv* 2:e1601367. <https://doi.org/10.1126/sciadv.1601367>
- Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. *Ecol Model* 190:231–259. <https://doi.org/10.1016/j.ecolmodel.2005.03.026>
- R Core Team (2018) R: A language and environment for statistical computing
- Rivers MC, Taylor L, Brummitt NA et al (2011) How many herbarium specimens are needed to detect threatened species? *Biol Conserv* 144:2541–2547. <https://doi.org/10.1016/j.biocon.2011.07.014>
- Roberts DL, Taylor L, Joppa LN (2016) Threatened or data deficient: assessing the conservation status of poorly known species. *Divers Distrib* 22:558–565. <https://doi.org/10.1111/ddi.12418>
- Syfert MM, Joppa L, Smith MJ et al (2014) Using species distribution models to inform IUCN Red List assessments. *Biol Conserv* 177:174–184. <https://doi.org/10.1016/j.biocon.2014.06.012>
- Valavi R, Guillera-Arroita G, Lahoz-Monfort JJ, Elith J (2022) Predictive performance of presence-only species distribution models: a benchmark study with reproducible code. *Ecol Monogr* 92:e01486. <https://doi.org/10.1002/ecm.1486>
- van Proosdij ASJ, Sosef MSM, Wieringa JJ, Raes N (2016) Minimum required number of specimen records to develop accurate species distribution models. *Ecography* 39:542–552. <https://doi.org/10.1111/ecog.01509>
- VanDerWal J, Shoo LP, Graham C, Williams SE (2009) Selecting pseudo-absence data for presence-only distribution modeling: how far should you stray from what you know? *Ecol Model* 220:589–594. <https://doi.org/10.1016/j.ecolmodel.2008.11.010>
- Willis F, Moat J, Paton A (2003) Defining a role for herbarium data in Red List assessments: a case study of Plectranthus from eastern and southern tropical Africa. *Biodivers Conserv* 12:1537–1552
- Wilson RJ, Thomas CD, Fox R et al (2004) Spatial patterns in species distributions reveal biodiversity change. *Nature* 432:393–396

## Authors and Affiliations

Charles J. Marsh<sup>1,2</sup>  · Mindy M. Syfert<sup>1</sup> · Elina Aletrari<sup>1,3</sup> · Yoni Gavish<sup>2</sup>  · William E. Kunin<sup>2</sup> · Neil Brummitt<sup>1</sup>

Mindy M. Syfert  
mmsyfert@gmail.com

Elina Aletrari  
ealetrari@gmail.com

Yoni Gavish  
gavishyoni@gmail.com

William E. Kunin  
w.e.kunin@leeds.ac.uk

Neil Brummitt  
n.brummitt@nhm.ac.uk

<sup>1</sup> Department of Life Sciences, Natural History Museum, London, UK

<sup>2</sup> School of Biology, University of Leeds, Leeds, UK

<sup>3</sup> Department of Geography, King's College London, Strand Campus, London, UK