University of Groningen

Reconciling methodologically different biodiversity assessments

Van Gemerden, BS; Etienne, RS; Olff, H; Hommel, PWFM; Van Langevelde, F; Gemerden, Barend S. van; Hommel, Patrick W.F.M.

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Link to publication in University of Groningen/UMCG research database

RECONCILING METHODOLOGICALLY DIFFERENT BIODIVERSITY ASSESSMENTS

Barend S. van Gemerden,[1,2,5] Rampal S. Etienne,[3,6] Han Olff,[3] Patrick W. F. M. Hommel,[4] and Frank van Langevelde[1]

[1]*Department of Environmental Studies, Wageningen University and Research Centre, Bornsesteeg 69, 6708 PD Wageningen, The Netherlands*
[2]*Biosystematics Group, Wageningen University and Research Centre, Box 8010, 6700 ED Wageningen, The Netherlands*
[3]*Community and Conservation Ecology Group, University of Groningen, Box 14, 9750 AA Haren, The Netherlands*
[4]*Centre for Ecosystem Studies, Wageningen University and Research Centre, Box 47, 6700 AA Wageningen, The Netherlands*

Abstract. Ongoing large-scale habitat disturbance requires quick identification of conservation priorities such as targeting sites rich in species and/or endemics. Biodiversity assessments are time consuming and expensive, so surveys often rely on partial sampling. Optimal use should be made of all currently available sources of information, but methodological differences between surveys hamper direct comparison. Because diversity depends on spatial scale, diversity characteristics of different sites are best compared on the basis of species–area relationships. As a result of the incompleteness of sampling, the observed species–area relationship deviates from the ''true'' species–area relationship.

In this paper, we identify five key factors affecting the shape of the species–area relationship due to incomplete sampling: (1) the total spatial extent of the observations, (2) the spatial distribution of the observations, (3) the proportion of the total extent sampled, (4) the proportion of the individuals in the sampled area included in the survey, and (5) the proportion of the included individuals successfully identified. We outline how methodologically different surveys can be combined to optimize the use of existing data in the evaluation of conservation needs, particularly for tropical forests.

As an illustration, we analyzed four methodologically different botanical surveys in the same area of old growth lowland forest in South Cameroon with the aim of reconciling these surveys. The four surveys were (1) reconnaissance scale vegetation mapping, (2) detailed botanical assessment (all individuals), (3) incomplete botanical assessment (10% individuals), and (4) herbarium collections.

By correcting for the five key factors we were able to match the results of the four different biodiversity surveys. The five key factors affected the recorded number of species and endemics differently; partial sampling of extent (3) and individuals (4) and partial identification of individuals (5) were the three most important factors.

We conclude that reconciliation of biodiversity assessments is possible if the differences between methods can be accounted for. We advocate reliable documentation of survey methods, especially the five key factors, because it greatly enhances the potential of combining methodologically different surveys for comparative biodiversity analyses.

Key words: biodiversity assessments; Cameroon; conservation priorities; endemics; plants; species–area relations; species richness; tropical forests.

INTRODUCTION

Biodiversity conservation, particularly in the tropics, is an issue of increasing importance as ongoing large-scale habitat disturbance poses a major threat to the survival of many species (Vitousek et al. 1997). Planners are challenged to balance conservation efforts with societal demands for natural resources and conservation funds. Therefore, not all threatened species and

ecosystems can be protected and priorities for biodiversity conservation need to be set (Myers et al. 2000). There is a growing literature on reserve site selection theory (Margules et al. 1988, Csuti et al. 1997, Ando et al. 1998, Margules and Pressey 2000, Olff et al. 2002) and spatial optimization for ecological management (Hof and Bevers 1998, van Langevelde et al. 2000). These studies present approaches to select sites that represent the highest possible number of species. Moreover, so-called gap analyses are conducted to identify gaps in the representation of biodiversity in reserve sites (Scott et al. 1993). Here, it is again critical what sites are selected for inclusion in the network of reserve sites.

Species richness, level of endemism, and exposure to threats are generally accepted as a robust set of criteria for setting conservation priorities (Hawthorne 1996, de Groot 1992, Myers et al. 2000, ter Steege 2000), although there is obviously more to conservation than these criteria (e.g., in secondary forests; van Gemerden et al. 2003a). However, a major obstacle for setting balanced conservation priorities is the incompleteness of such biodiversity information (Faith and Walker 1996, Howard et al. 1998). For example, Lombard et al. (1997) propose a reserve selection in the species-rich Agulhas Plain in South Africa and use, in absence of better data, herbarium records combined with fragmented observations on species distribution. Often, surrogate data are used in conservation planning to identify areas deserving high priority for protection such as key species, indicator species, or umbrella species (Faith and Walker 1996, Simberloff 1998). In contrast to temperate regions, only little ecological knowledge is currently available for the majority of tropical species. Best studied are larger mammals and birds, and species conservation plans for these groups can be based on their distribution patterns and ecological requirements (e.g., Mickleburgh et al. 1992, Oates 1996). For the majority of tropical plant and insect species, neither species ecology nor distribution patterns are sufficiently known for such analyses. At present, only few forest areas have been systematically surveyed for even a limited number of taxonomic groups (e.g., Howard et al. 1998, 2000). Moreover, complete biodiversity assessments of tropical forests are extremely expensive and time-consuming, and arguably beyond the capacity of the global research community (Howard et al. 1998, Lawton et al. 1998). The ongoing rapid habitat disturbance in many tropical regions now implies that conservation priorities need to be identified quickly. Therefore, conservation planners are urgently challenged to set priorities on the presently available, yet incomplete, information on biodiversity patterns in tropical forests (Gaston and Rodrigues 2003). In this paper, we present an approach for efficient use of the available information, using tropical forest data as our prime example.

The main sources of information in setting conservation priorities for tropical plants are species distribution maps based on collection localities of herbarium specimens (Lovett et al. 2000, Linder 2001, Poorter et al. 2004). The advantage of herbarium collections for biodiversity assessments is their state-of-the-art taxonomic identification. Moreover, collections are stored and can be re-examined if necessary. However, it is problematic for biodiversity assessments that collections are generally clustered in areas with high collecting effort, often chosen for reasons (historical, practical) other than their high or characteristic diversity (Nelson et al. 1990, Funk et al. 1999, ter Steege et al. 2000, Poorter et al. 2004). In addition, herbarium collectors tend to focus on (flowering) material that is relevant to their research (e.g., taxonomic revisions).

As a result, collections are a poor representation of the whole plant community composition. Moreover, the number of collections stored in herbaria is generally too small to allow for biodiversity analyses at smaller than regional scales (ter Steege 2000)

Fortunately, herbarium collections by taxonomists are not the only source of information on plant diversity of tropical forests. Most tropical forests have had their share of ecologists, vegetation surveyors, foresters, etc., who, in a more or less systematic way, collected data on forest composition. These surveys are generally less precise in a taxonomic sense but have the advantage that they include a larger proportion of the individuals present in the forest, e.g., through plot-based sampling. Moreover, survey effort is generally more equally distributed over the area. Although not always designed for conservation purposes, these surveys contain valuable information on biodiversity. However, so far, few conservation analyses have been based on such surveys (ter Steege 1998). Pressed for quick conservation priorities, it is necessary to fully exploit the information potential of these alternative sources, possibly in conjunction with each other and the more traditional herbarium collection approach. This requires a more formal comparative approach that captures the essential differences between sampling methods, and their consequences for the measurement of diversity.

In the present study, we identify the critical differences between commonly used sampling methods in general. Then we propose a method for reconciling assessments with a special focus on tropical forests. Finally, we show an example of a reconciliation of the results of four methodologically different botanical diversity assessments in the same forest area in southern Cameroon.

METHODS

Main methodological differences between diversity assessments

Botanical diversity assessments aim to identify the kind and number of plant species that inhabit a specific locality. As diversity depends on spatial scale (Rosenzweig 1995), diversity characteristics of different tropical forest sites are best compared on the basis of species–area relations. Species–area relations describe the accumulation of new species with area. Such relations are nested and diversity at smaller spatial scales is embedded in diversity at larger spatial scales. Species–area relations are generally not linear or log-linear; species accumulate most rapidly in small areas (mainly a sampling effect), then more slowly in intermediate areas, and again more rapidly in larger areas, as areas with completely different evolutionary histories are included (Arrhenius 1921, Gleason 1922, Rosenzweig 1995, Hubbell 2001; Fig. 1). The species–area relationship reflects the spatial variation in the density of individuals at small scales, and the spatial distribution
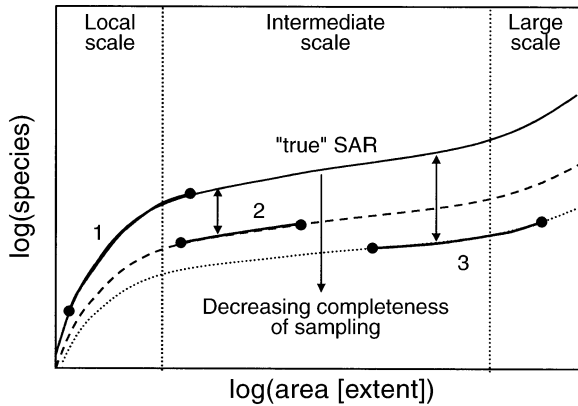
FIG. 1. General shape of species–area relations (SAR). The solid line represents the "true" SAR resulting from complete sampling. Partial sampling results in information shortage (indicated with arrows). Surveys (examples 1–3) generally only cover a limited range of the SAR. The actual magnitude of the scales depends on the taxonomic group considered. For birds, the transition from local to intermediate scale occurs at around 1×10^{-1} km² and from intermediate to global around at 1×10^{6} km², whereas for plant diversity these transitions occur at 1×10^{-4} km² and 1×10^{3} km², respectively (Preston 1960, Shmida and Wilson 1985, Gaston 1994, Hubbell 2001). For tropical forests, we estimate these transitions to occur at 1×10^{-2} km² and 1×10^{5} km².

of individuals of different species. In a given tropical forest environment, mean plant densities (i.e., number of individuals per unit area) are generally very constant (Hubbell 2001) and species–area relations of such environments thus mainly reflect the more interesting spatial arrangement of individuals of different species. At uniform densities, the slope of the species–area relationship is steeper in areas where neighbouring individuals are less likely to be of the same species, i.e., species are less clustered.

The "true" species–area relation should be based on the complete assessment of all individuals. However, such complete biodiversity assessments are impossible to conduct at sufficiently large spatial scales in species-rich and taxonomically poorly studied ecosystems like tropical forests. Survey methods try to overcome these practical limitations by incomplete sampling. The incompleteness of the sampling causes an information shortage on the spatial distribution of species and therefore leads to deviations from the true species–area relation (Fig. 1). Different assessment methods take different decisions on which individuals to include in the survey. We suggest five main factors why different methods produce different information on diversity when applied in the same area. These causes for information shortage are (1) the total extent (E) in which observations are made, (2) the spatial distribution (C) of the observations, (3) the proportion of the total extent sampled (p_e), (4) the proportion of the individuals of the sampled area included in the survey (p_i), and

(5) the proportion of the included individuals that was successfully identified (p_d) (Fig. 2).

Extent (E).—Because it is not possible for practical reasons to taxonomically identify all the individuals in any tropical forest region, most methods assess the diversity in selected sample areas, e.g., plots or transects. These sample areas are distributed in the much larger area that they are meant to represent (Fig. 2). This larger area, hereafter referred to as extent (E), is the appropriate spatial scale to analyze diversity characteristics. Simply constructing species–area relationships by collating the sampled areas underestimates the actual position on the species–area relationship. Therefore, to compare the results of different methods it is necessary to distinguish between sampled area and extent.

Spatial distribution of sample points (C).—Within the extent, the spatial distribution of sample points (C) is likely to influence the observed species richness. Samples closer together are generally more related in species composition. Especially in areas with high beta diversity, clustering of sample points may therefore lead to underestimation of species richness. Spatial distribution of sample points may vary strongly between different methods.

Proportion of total extent sampled (p_e).—Partial sampling of extent implies that observations are only made in selected sites and that no information was collected on the individuals in the area in between these sites (Fig. 2). The proportion of the total extent (p_e) that is actually sampled reflects the size of the sample
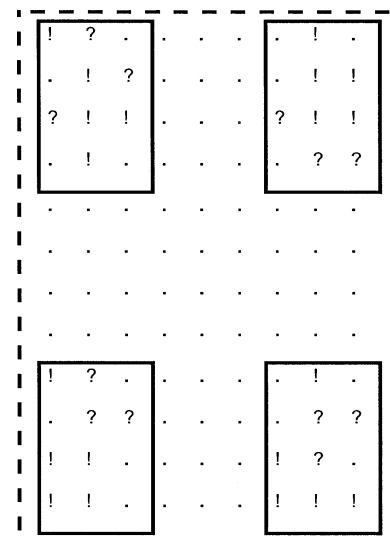

FIG. 2. A schematic showing extent, sample area, and individuals included and identified in a survey based on partial sampling. Extent (represented by the dashed line) envelopes the areas actually sampled (represented by the solid lines). Individuals within sample area are not sampled (·), sampled but not identified (?), or sampled and identified (!).

TABLE 1.   Classification of common botanical assessment methods applied in tropical forests with respect to key factors affecting the observed species richness.

	Key factors			
Assessment method	Total extent of survey (E)	Proportion of total extent sampled (p_e)	Proportion of individuals in sampled area included (p_i)	Proportion of included individuals identified (p_d)
Herbarium collections	large	large	very small	very large
Vegetation mapping	large	small	medium	small
Line intersect methods	medium	small	large	medium
Rapid Botanical Assessment	medium	medium	medium	medium
Para-taxonomic sampling	medium	medium	large	very small
Plot-based sampling	small	small	large	medium

in relation to the area that was not sampled and is likely to have a strong effect on observed species richness.

Proportion of individuals in sampled area included (p_i).—Sample methods also vary with respect to the proportion of individuals (p_i) in the sampled areas that are included in the survey. In sampling methods that are not plot based, especially, often not all individuals are included. In other methods, cut-off levels are also often applied for practical reasons (e.g., not all size classes or taxonomic groups included). Differences between methods may result in different information shortages on the spatial distribution of species and may hamper comparison of the results (Table 1).

Proportion of included individuals successfully identified (p_d).—Finally, methods may vary in the proportion of individuals that is successfully identified (p_d). Species identification in highly diverse but taxonomically incompletely studied tropical forests requires much expertise. For the majority of surveys in tropical regions only limited botanical expertise, reference collections and identification keys are available. Moreover, some surveys rely on nonspecialists or use parataxonomic classification of species (Oliver and Beattie 1996*a*, *b*, Hellier et al. 1999, Danielsen et al. 2000, Kerr et al. 2000). An additional problem is that in tropical forests at any time only a fraction of the individual plants are flowering or fruiting. As classic identification keys are mainly based on reproductive organs, the absence of flowers and fruits complicates identification. Methods differ greatly in the efforts made to identify the individuals surveyed (Table 1).

Reconstruction of species–area relationship from incomplete data

The true diversity (S_{tr}) in the total extent is based on a complete survey of individuals and their complete identification. The true species–area relationship gives the accumulation of species with area (extent) when all individuals are sampled in a nested design (Fig. 1)

$$S_{tr} = f_1(E). \tag{1}$$

Assessment methods based on partial sampling of individuals yield an observed diversity (S_{obs}). Reconciling methods implies reconstructing S_{tr} from S_{obs} values while correcting for the different components of in-

formation shortage; i.e., C, p_e, p_i, and p_d. The observed number of species will be some proportion r of the real number of species:

$$S_{obs} = rS_{tr} \tag{2}$$

and therefore we have

$$S_{tr} = S_{obs}/r. \tag{3}$$

A first approximation of r is the product of four reduction factors, representing the effects of respectively not sampling all the area in the total extent (r_e), spatial clustering of sample points (r_c), not sampling all individuals in the sampled area (r_i), and not successfully identifying all sampled individuals (r_d)

$$r = r_c \times r_e \times r_i \times r_d. \tag{4}$$

Finally, we need to find specific formulas (f_i) to calculate these reduction factors from the components of information shortage

$$r_c = f_2(k_c, C) \tag{5a}$$

$$r_e = f_3(k_e, p_e) \tag{5b}$$

$$r_i = f_4(k_i, p_i) \tag{5c}$$

$$r_d = f_5(k_d, p_d). \tag{5d}$$

In these functions, k_c, k_e, k_i, and k_d are constants that represent the consequences for observed species richness of, respectively, the distribution of observations (C), incomplete sampling of area (p_e) and individuals (p_i), and partial determination (p_d). The values of k_c, k_e, k_i, and k_d may vary between regions and therefore need to be determined case by case. Clustering of sample points has theoretically a negative effect on observed species richness. Therefore, r_c can be approximated as a linear decline to fit maximum species richness on highest observed uniformity of data points. All other reduction factors can theoretically be described as power functions in which all species will be observed at complete sampling. Combining Eqs. 3, 4, and 5 yields the predicted species richness, corrected for the information shortage of the specific method, as

$$S_{pred} = S_{obs}/(r_c \times r_e \times r_i \times r_d) \tag{6}$$

TABLE 2. Characteristics of the four studied botanical surveys of old-growth lowland forest in southern Cameroon.

Characteristic	Data set 1, vegetation mapping†	Data set 2, diversity assessment (100%)‡	Data set 3, diversity assessment (10%)‡	Data set 4, herbarium collections§,‖
Total no. plots	83	44	20	55
Plot size (m²)	100	625	625	740
Total sample area (ha)	0.83	2.75	1.25	3.75
Maximum extent (km²)	1235	219	180	896
No. species recorded	320	767	207	378
Morphospecies (%)	14.4	29.9	26.0	0
Endemic species (%)	16.3	16.6	15.6	14.8
Key factors				
Spatial distribution of samples				
Fisher's I	0.287	0.813	0.837	0.689
C standardized	0	0.526	0.550	0.402
Proportion of extent sampled ($p_e \times [1 \times 10^{-5}]$)¶	0.25	6.86	5.78	2.11
Proportion of individuals included (p_i)	0.9	1.0	0.1	0.0175
Proportion of individuals identified (p_d)	0.54	0.76	0.76	0.72

  † *Source:* van Gemerden and Hazeu (1999).
  ‡ *Source:* B. S. van Gemerden, *unpublished data.*
  § *Source:* Extract databases of Wageningen (Herbarium Vadense) and Kribi (Tropenbos/IRAD) in 2000.
  ‖ *Source:* The herbarium data set included 640 fully identified collections. These collections form 72% of the collections actually made in the area. The remaining 28% were not identified yet or were not stored at the Kribi or Wageningen herbaria. The area surveyed was estimated as 800 m² per 17.5 collections.
  ¶ Proportion of the extent sampled at an extent of ~100 km² (see Fig. 2).

CASE STUDY: METHODOLOGICALLY DIFFERENT BOTANICAL DIVERSITY ASSESSMENTS IN THE SAME FOREST AREA IN SOUTHERN CAMEROON

Study site

The study was conducted in the Bibindi–Akom II–Lolodorf region, south Cameroon (3° N, 10° E; 1700 km²). The climate is humid tropical with two distinct wet seasons (March–May and August–November) and two relatively drier periods. The average annual rainfall is 2000 mm (Waterloo et al. 2000). Average monthly temperatures vary between 22.9 and 27.5°C (Olivry 1986). The parent material consists of Precambrian metamorphic rocks and old volcanic intrusions (Franqueville 1973). Topography varies from flat erosional plains to rolling uplands with isolated hills and mountains. Altitude varies from 50 m to 1000 m above sea level. Soils range from moderately acid sandy clay loam to highly clayey and strongly acid and classify as Haplic Acrisols and Plinthic and Xanthic Ferrasols (van Gemerden and Hazeu 1999). Evergreen forests of the Atlantic Biafrian type largely cover the area (Letouzey 1968, 1985). These forests are characteristically rich in Leguminosae–Caesalpinioideae and have a closed canopy at 30–40 m with emergents often surpassing 55 m. The area is rich in plant species. So far, approximately 1600 species have been recorded in the area of which 1264 species have been identified to species level (B. S. van Gemerden, *unpublished data*). The recorded species include 261 species that are endemic to the lower Guinea forest region (Nigeria–Gabon) of which 51 species are restricted to the forests of Cameroon.

Data sets

The vegetation in the area has been surveyed by different methods by a number of projects (Letouzey, 1968, 1985, van Gemerden and Hazeu 1999, Guedje 2002, van Gemerden et al. 2003a, b). The present study examines four surveys that vary in extent (E), spatial distribution of observations (C), proportion of the extent sampled (p_e), proportion of individuals sampled (p_i), and proportion of sampled individuals identified (p_d). The selected methods represent often-used approaches of vegetation survey in tropical forest regions. In order to minimize the effect of spatial heterogeneity, we focused on undisturbed forests between 50 m and 700 m altitude. According to a reconnaissance landscape ecological survey, these forests are quite similar in terms of general floristic composition, soil properties, and landforms (van Gemerden and Hazeu 1999). We first present the characteristics of the four data sets needed to calculate C, p_e, p_i, and p_d (Table 2).

The first data set is an area-wide vegetation survey made for mapping the most important aspects of vegetation at scale 1:100 000 (van Gemerden and Hazeu 1999). Relatively homogenous tracts of land were identified on aerial photographs and joint descriptions of vegetation, soil and landform of the most important units were made in the field. In 83 localities, external foliage cover was estimated of the most important plant species in 10 × 10 m plots (100 m², total sample area 0.83 ha). Plots were more or less evenly distributed over the area (Fig. 3). Mosses, ferns, epiphytes, and seedlings were not included in the survey. Plant identification was done in the field with the help of a field botanist and a local tree spotter. Plant material was

collected from unknown species. Identification at the National Herbarium of Cameroon (Yaoundé) and Limbé Botanic Garden focused on material of the most abundant species. Abundant or characteristic species that could not be identified were categorized as morphospecies. In the total survey of 0.83 ha, 320 plant species were recorded.

The second data set used is a detailed botanical assessment of the area (B. S. van Gemerden, *unpublished data*). Vegetation was sampled in 44 plots of 25×25 m (625 m^2) representing a total sample area of 2.75 ha. Plots were grouped in four sample areas that represented the most important variation in vegetation, soils and landforms (Fig. 3). In the plots, all plants (except individuals of woody species less than 50 cm tall) have been identified. In the field, the most common and readily identifiable species were directly named and plant material was collected of all other species. Voucher material was processed at the Kribi Herbarium (Tropenbos-Cameroon Herbarium) and sent to the National Herbarium of Cameroon (IRAD Yaoundé) and the National Herbarium of the Netherlands—Wageningen University branch for identification by specialists. Collections that could not be identified to species level were systematically categorized as morphospecies. In the total survey of 2.75 ha, 767 species were recorded.

The third data set is a subset of the previous survey and includes the 20 plots (625 m^2, total sample area 1.25 ha) in which all individuals were enumerated. Plots were clustered in four localities (Fig. 3) Ten percent of the individuals per plot were randomly selected and thus a plot-based data set with low sampling intensity was constructed. Field methods and plant identification followed the procedures of set 2. In the total survey of 1.25 ha, 207 species were recorded.

The fourth data set consists of all botanical specimens that were collected in the area and stored in the National Herbarium of the Netherlands–Wageningen University branch, or the field herbarium of the Tropenbos-Cameroon Programme. The herbarium collections were made by a variety of collectors from 1885 to 2001. Only collections that were, according to the collection notes, made in undisturbed forest sites and that were accurately georeferenced were included. In addition, collections that were made as part of a more or less systematic vegetation survey were omitted. This set includes a total of 640 collections made in 55 different localities that were grouped in a few disjoint clusters (Fig. 3). The herbarium database only includes completely and reliably identified specimens, i.e., mostly fertile material that has been checked by specialists. The total number of species in this data set is 378.

Data analysis

Only species that, within a set, were uniquely named (i.e., were comparable between plots) were included in the analyses of species richness. This implied that fully
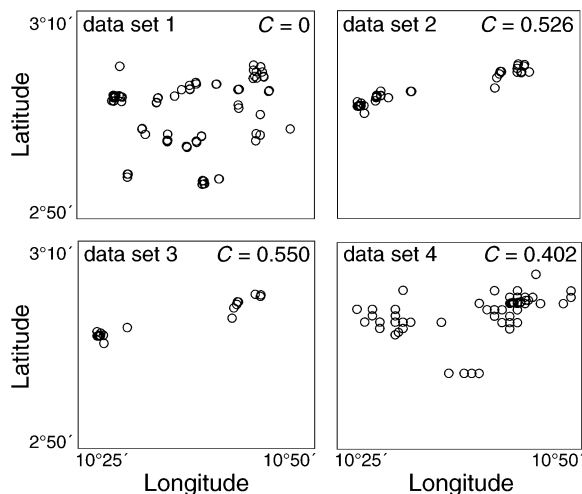

Fig. 3. The spatial distribution of observations in the four data sets included in this case study. C is a measurement of the distribution of observations.

identified species and morphospecies were included and other less precise identifications were omitted, such as field names that were found or were likely to refer to small clusters of morphologically similar species. Species richness was defined as the total number of species (including morphospecies). Of special concern to biodiversity conservation are endemic species, i.e., species with a restricted geographical distribution range. In the present study, we defined endemic species as species restricted to the Lower Guinean forest region (Nigeria, Cameroon, Equatorial Guinea, and Gabon). A further refinement of endemic status is problematic at present, as this part of Africa is generally poorly explored botanically.

For each survey, we calculated sample area (A), extent (E), species richness, and the spatial distribution of the plots (C). The sample area (A) is the area that was actually included in the survey. For data sets 1, 2, and 3 it is simply the total area covered by plots. For set 4 (herbarium collections) we calculated the sample area on the basis of the number of collections made. On an average collecting day a herbarium botanist collects 15–20 herbarium specimens while "scanning" an area of approximately 7.5 ha (B. S. van Gemerden, *personal observations*; P. Tchouto, *personal communication*; F. J. Breteler, *personal communication*). However, not all individuals in this area will be thoroughly examined, as the herbarium collector's primary interest are flowering and fruiting plants. Based on a small trial, we estimated that 1000 plants are scrutinized during an average day collecting, representing an effective sample area of 800 m^2. However, as not all collectors spend equal time collecting, the area they covered varies. We grouped the collections by locality and estimated the sample area per locality on the basis of the above estimates. A further refinement of these esti-

mates was necessary because some collections made by a specific collector on a specific date were not stored in the herbarium database. This was the result of the fact that herbaria generally do not incorporate unidentified material. In addition, the Wageningen herbarium may not have duplicates of all collections made in the area by (especially) collectors affiliated to other herbaria. As collections are generally uniquely and consecutively numbered, we identified the total number of collections per collector in a specific locality by assuming that missing intermediate numbers were also collected in the same locality. We calculated that some 28% of the collections made were not stored in the herbarium database. The total number of collections per locality was corrected for the proportion of missing material to get a more accurate estimate of the sampled area for data set 4.

We defined the extent (E) as the surface of the polygon that included a given number of plots plus a buffer of 50 m. The size of the buffer reflects common practice in vegetation sampling that plots are not located too close to sudden changes in vegetation or environment. Calculations were performed with the GIS Arcinfo (ESRI, Redlands, California, USA).

For each data set, we calculated species richness and endemic species richness. To construct species-extent curves for each data set, we plotted species richness and endemic species richness against the extent of an increasing number of plots. We constructed series of randomly selected plots to obtain series of randomized plot orders. We used values for extent and species richness averaged over the total number of runs per data set. The number of series of randomized plot orders varied per data set due to computational limitations, i.e., 325 series for data set 1, 795 series for data set 2, 2123 series for data set 3, and 331 series for data set 4. In data set 3, many repetitions were chosen because in this set also a random subset of 10% of the individuals was selected per series. Standard deviation of extent and species richness were very small for all data sets indicating that the numbers of runs used were amply sufficient.

We characterized the spatial distribution of the plots (C) by calculating Fisher's I index; i.e., the ratio of the standard deviation of all point-to-point distances to the average point-to-point distance (Cressie 1993). This index produces high values for clustered point distributions and low values for uniform distributions. We scaled the index values to the observed minimum in the data sets to obtain a range relevant to the present analyses, as $C = I_i - I_{min}$.

To analyze the effects of partial sampling on species richness, we constructed a simulation data set. This simulation data set is based on the 20 enumerated plots of data set 3 and therefore does not contain simulated data. We call it a simulation data set, however, because we manipulated it by varying one parameter and keeping all other parameters constant. In this way, we could

identify the separate effects of partial sampling of respectively C, p_e, p_i, and p_d. We proceeded as follows.

In all 20 enumerated plots of data set 3, we randomly selected 267 fully identified individuals. We isolated the effect of spatial distribution of sample points on species richness by calculating C for all combinations of two plots together with four fixed plots forming the outer perimeter of the extent. As a result the spatial distribution, C varied while p_e, p_i, and p_d remained constant. Likewise, the effect of the proportion of the total extent sampled (p_e) on species richness was analysed by selecting the plots forming the outer perimeter of the extent and randomly adding plots to increase p_e while keeping extent E, p_i, and p_d constant. Average species richness values over 100 randomized runs were used in the analyses to average out variation in C.

For each plot, we analysed the effect of the proportion of individuals included in the sampled area (p_i) on species richness. Per plot, individuals were randomly chosen and average species richness values over 1000 runs at fixed proportions were used in the analyses. We analysed the effect of partial identification of included individuals (p_d) on species richness by randomly classifying different numbers of species per plot as either identified or unidentified. In field situations, individuals are generally quite accurately grouped by botanical species and therefore individuals within a group were assumed to have the same identification status. In the analyses, the average numbers of individuals representing 5, 10, 20, 40, and 55 identified species per plot over five random selections were used.

Results

Method characteristics

Although all data sets represented botanical surveys done in the same area of old growth forest at low altitudes, the observed number of species varied considerably between data sets (Table 2). Most species were recorded in data set 2 (767 species), while data set 3 had only 207 species. Despite these differences, the proportion of endemic species was surprisingly constant between sets; i.e., between 14.8 and 16.6%. Hence, this already provides a partial answer to one of our objectives (estimating endemic species richness) and we can focus our attention on our other objective (estimating total species richness) on which the former depends.

The sampled area (A) per data set varied from 0.83 ha (set 1) to 3.75 ha (set 4), while the maximum extent (E) varied from 180 km^2 (set 3) to 1236 km^2 (set 1) (Table 2). Distribution of sample points was most evenly spread in data set 1 (Fisher's $I = 0.287$, $C = 0$), while data set 3 was highly clustered ($C = 0.402$) (Fig. 3; Table 2). The proportion of the extent covered by sampling (p_e) varied with extent and between data sets (Fig. 4). In general, p_e decreased in all sets with increasing extent until a critical value was reached after
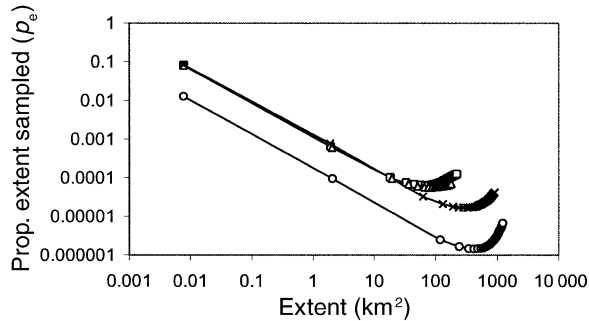
FIG. 4. The relation between extent and proportion of the extent sampled for the different survey methods studied. Symbols are: data set 1, circles; data set 2, squares; data set 3, triangles; data set 4, crosses.

which p_e increased again. Data set 3 did not show an increase in p_e at larger extents but leveled off. The critical value varied between sets and was lowest in data set 2 (≈ 100 km²) and highest in data set 1 (≈ 650 km²).

The proportion of individuals included in the surveyed sites (p_i) varied between data sets. Data set 2 included all individuals in the sampled area, while data set 3 included, by definition, only 10% of the individuals. The prescribed methodology for the vegetation mapping survey (data set 1) was to focus on the most characteristic features of the vegetation in the plot (i.e., dominant species per vegetation stratum). However, in the field an estimated 90% of the individuals were actually screened (B. S. van Gemerden, *personal observations*). In data set 4, material was collected of, on average, 17.5 individuals out of the 1000 individuals that were examined and therefore the estimated proportion of individuals included in the herbarium data set is 0.0175.

The proportion of surveyed individuals that were successfully identified (p_d) was generally high. In data sets 2 and 3, more than 70% of the individuals were successfully identified to species level and additionally 6% were morphocategorized. In data set 4, all individuals were identified to species level. However, some 28% of the collection numbers that were assumed to be collected in the area were not stored in the databases of the Kribi and Wageningen herbaria and therefore p_d was effectively 0.72. In the vegetation survey on which data set 1 was based, no information was recorded on the number of individuals. Therefore, p_d was estimated using general individuals–area relations (individuals = $1.2548 \times \text{area}^{0.9963}$; $r^2 = 1.0$; all individuals data set 3; area measure in m²) and species–individuals relations (species = $4.3927 \times \text{individuals}^{0.5037}$, $r^2 = 0.99$; all identified individuals data set 3). The estimated total number of individuals per 100 m² was 123.4. In data set 1, 90% of the individuals were included in sampling, i.e., 111.0 individuals were sampled per 100-m² plot. The species–individuals relation predicted that 111.0 individuals represent 47.1 species. However, in

data set 1, only 34.5 species were found on average per plot, i.e., only 60.0 individuals were identified in this survey. According to these estimations, the proportion of individuals identified in data set 1 was only 54%.

Consequences of incomplete sampling

The simulation data set permitted to assess the individual effects of partial sampling of each of the components C, p_e, p_i, and p_d. The effects of p_i and p_d were assessed per plot. With nonlinear regression, we fitted the data of simulation data set to the model $r_x = p_x^{k_x}$ for each of the variables to obtain the constants k_i, and k_d. The fitted values of k_i and k_d were, respectively, 0.582 ($r^2 = 1$) and 0.949 ($r^2 = 0.98$) (Fig. 5c, d).

The plots in the simulation data set were highly clustered and therefore standardized C values ranged only from 0.33–0.55. This range did not cover the degree of clustering of the different data sets (0–0.55). We estimated maximum species richness at the most uniform distribution (data set 1, $C = 0$) by regressing the species richness data of the simulation set over C. Although regression fit was poor ($y = -66.2C + 244.2$, $r^2 = 0.14$), this result agreed with the maximum cumulative species richness found in six plots in the simulation data set and is supported by the general recognition of low beta diversity in the area. For the present study, we therefore used 244 species as the estimated upper limit of species richness for the range of C values. The value of the site-specific reduction factor k_c in the model $r_c = k_c C + 1$, was -0.271 (linear regression, $r^2 = 0.14$).

In the simulation data set, p_e values ranged only from 2.1×10^{-5} to 6.9×10^{-5} and did not cover the range of p_e values observed in the different surveys. Therefore, the maximum number of species was estimated on the basis of a preliminary checklist of the area and available taxonomic literature (mainly Aubréville and Leroy 1961–1992, 1963–2001, Keay and Hepper 1954–1972, Cable and Cheek 1998). Based on these sources, we estimated the total number of plant species occurring in old growth lowland forest in the Bipindi–Akom II–Lolodorf region to be 3000. However, plots in the simulation data set contained only 267 individuals instead of the average number of 758 individuals in data set 3. Therefore the total number of species expected in the simulations at $p_e = 1$ was estimated using the newly established relation between p_i and proportion species as: maximum species richness = $3000 \times (p_i \text{ simulation set})^{k_i} = 1634$ species. The model $r_e = p_e^{k_e}$ fit the simulation data relatively well (nonlinear regression: $k_e = 0.185$, $r^2 = 0.67$) but showed relatively large deviations in the range of p_e relevant for the present study ($0.2 \times 10^{-5} - 0.08$). A much better fit was obtained by the model $r_e = p_e^a/(p_e^a + e^b)$, and therefore the site-specific reduction factor k_e defined in this case as the logit of r_e ($k_e = \ln[r_e/(1 - r_e)] = a \times \ln[p_e] +$
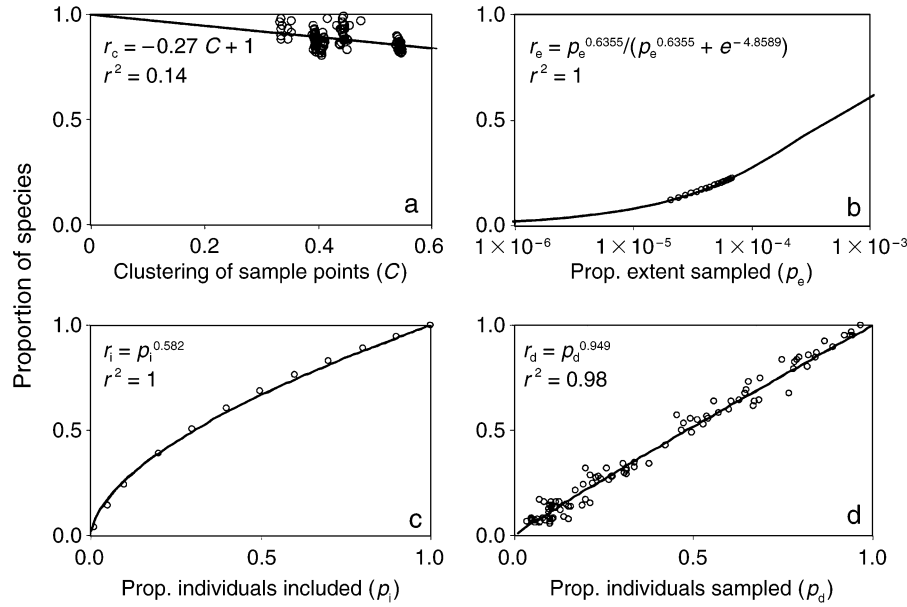
FIG. 5. The effects of partial sampling on observed species richness in simulations with a fully standardized data set. Individual effects of (a) spatial distribution of observations (C), (b) partial sampling of extent (p_e), (c) partial sampling of individuals (p_i), and (d) partial identification of individuals (p_d).

b), was estimated as $0.636 \times \ln(p_e) - 4.859$ ($r^2 = 1$; Fig. 5b).

The fitted values of the site-specific reduction factors indicated that partial sampling of extent had the largest effect on observed species richness. Partial identification and partial sampling of individuals also affected survey results substantially, while the effect of spatial distribution of sample points (C) had a much smaller effect. Our empirical model for Eq. 6 becomes

$$S_{pred} = S_{obs}/[(-0.271C + 1)p_i^{0.582} \times p_d^{0.949} \times p_e^{0.636}]$$

$$\times (p_e^{0.636} + e^{-4.859}). \qquad (7)$$

Reconciling methods

The rate of species accumulation with sampled area was variable between the surveys (Fig. 6a). Scaling to extent made the slopes of the different data sets more similar but the absolute number of species recorded
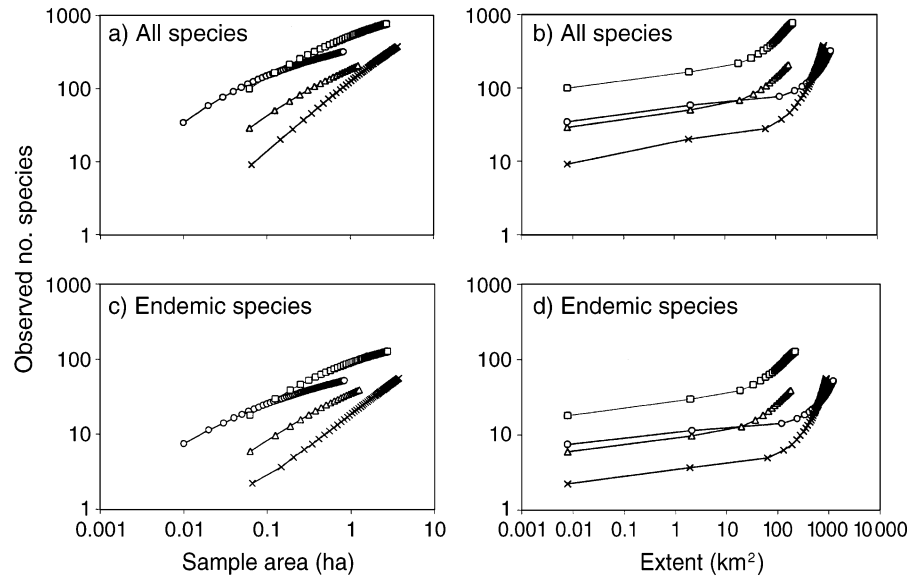

FIG. 6. (a, b) Observed total species richness and (c, d) endemic species richness scaled to (a, c) sample area and (b, d) extent of the four surveys studied. Symbols are: data set 1, circles; data set 2, squares; data set 3, triangles; data set 4, crosses.
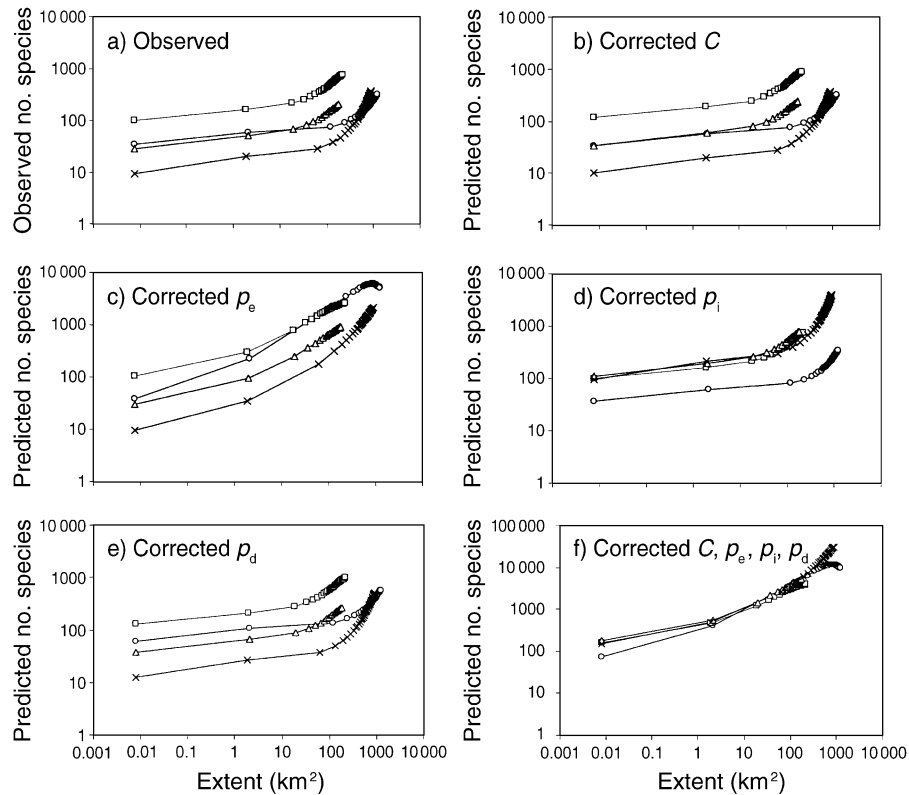
FIG. 7. (a) Observed species richness (all species), (b) predicted species richness corrected for spatial distribution of observations, (c) predicted species richness corrected for partial sampling of extent (p_e), (d) predicted species richness corrected for partial sampling of individuals (p_i), (e) predicted species richness corrected for partial identification of individuals (p_d), and (f) predicted species richness corrected for all key factors. Symbols are: data set 1, circles; data set 2, squares; data set 3, triangles; data set 4, crosses. Note the different y-axis scale for (f).

hindered comparison of the results (Fig. 6b). Accumulation of endemic species with sample area and extent were very similar to the pattern found in all species (Fig. 4c, d). Therefore, we focused in the remainder of the analyses on total species richness.

The results of the different surveys could be made more comparable by calculating the expected number of species while taking into account the methodological differences in sampling (Eq. 7). Spatial distribution of sample points varied little between data sets and because of the low beta diversity in the study region, correcting for C had little effect on the general shape and relative position of the species–extent curves (Fig. 7b). Correcting for partial sampling of extent (p_e) resulted in an overall increase in the rate of species accumulation (Fig. 7c). Despite the correction for p_e the curve of data set 1, and to lesser extent that of data set 2, decreased slightly at very large spatial scales. Correcting for partial sampling of individuals (p_i) merged the curves of data sets 2, 3, and 4, while the curve of data set 1 was considerably lower over all scales of extent (Fig. 7d). Correction for p_i in data set 4 resulted in a slight overestimation in absolute species richness at very large spatial scales. Correcting for partial identification of individuals (p_d) slightly increased species

richness in all data sets but had no effect on the relative position of the different data sets (Fig. 7e).

Correcting for all four factors of information shortage, i.e., C, p_e, p_i, and p_d, greatly improved the comparability of the results of the four surveys over all scales of extent. Together the corrections largely merged the curves of the different data sets (Fig. 7f). The curve of data set 1 was slightly lower than those of the other data sets, while at very large spatial scales a decrease in species numbers was predicted. Predicted species richness in extents between 120 and 130 km² varied between data sets (Table 3). Highest values were predicted in data set 4 while lowest values were found for data set 2. However, the predicted species richness of data set 2 was still 72% of that of data set 1. Uncorrected species richness figures showed a much larger variation with data set 4 having only 7% of the species richness observed in data set 2. Nonetheless, predicted species richness tended to overestimate the actual species richness estimated on the basis of literature and a preliminary checklist of the area. Some 3000 species were expected to occur within the old growth lowland forests in the region at large spatial scales (>100 km²). Predicted species richness at extents from 120 to 130

TABLE 3.   Observed and predicted number of species at extents from 120 to 130 km² for different survey methods in the same area of old-growth lowland forest in South Cameroon.

Data set	Observed		Predicted		
	No. species	Percentage of maximum	No. species	Percentage of maximum	Overestimation (%)†
1	77	15	4257	95	142
2	504	100	3207	72	107
3	160	32	4396	98	147
4	37	7	4464	100	149

† Based on taxonomic literature and a preliminary checklist, actual species richness in the area was estimated at 3000 species.

km² were 107% (data set 2) to 149 % (data set 4) of this number (Table 3).

DISCUSSION

Our analysis of four botanical surveys in the same area of old growth lowland forest in southern Cameroon showed that information relevant to conservation planning such as species richness and endemic species richness were strongly influenced by survey method (Table 2; Fig. 6). The main difference between sampling methods was the selection of the individuals that were included in the survey. We found that important components affecting the general shape of the species–extent relation were spatial distribution of samples (C), proportion of extent sampled (p_e), proportion of individuals included in sampled area (p_i), and proportion of the sampled individuals identified (p_d). In simulations, these factors affected the number of species recorded differently (Fig. 5). Recorded number of species was most sensitive to partial sampling of extent and individuals, and partial identification. Fitting of the site-specific factors to data obtained by simulation enabled the correction of survey findings while taking into account the survey-specific methodological shortcomings. Despite the large variety in sampling methods and their generally poorly described methodology, predicted species numbers were surprisingly similar for all surveys over all scales of extent (Fig. 7). However, correction for the effect of partial sampling resulted in an overestimation of species richness compared to estimations based on current biogeographical insights (Table 3). Nonetheless, the results suggest that the method we propose adequately copes with the essential differences between sampling methods and their consequences for the measurement of diversity. With the proposed framework we advocate a more efficient use of existing sources of botanical information for comparing conservation values of tropical forests.

The present analyses clearly showed the importance of scaling survey results to extent (E) instead of the sum of actually sampled areas (A) (Fig. 6). As stated earlier simply collating samples and suggesting that

they formed a continuous area decouples the species from the actual area and therefore the interpretation of the species–area relation becomes troublesome at best. Especially for comparing different methods it is of paramount importance that the spatial scale reflects the true area in which the species were recorded. Of course, species-sample area curves can also be corrected for partial sampling. However, as extent and spatial distribution of sample points are not explicit in such curves, they can only be corrected for proportion of individuals included (p_i) and proportion of individuals identified (p_d). Correcting for p_i and p_d did not improve comparability of results of the four surveys studied (data not shown) which stresses the importance of scaling the survey results to extent values.

We reconstructed the species–area relation from surveys with different types of information shortage (Figs. 1, 2). Some of these surveys indeed focused on biodiversity assessment while data collection in others was aimed at vegetation mapping or as baseline data for taxonomic revisions. As a result, not all key characteristics important for reconciling assessments were readily available and a series of assumptions were required. It is likely that some of the observed variation is due to inaccurate estimations of sample areas, p_e, p_i, and p_d. Likewise, only one survey included fully enumerated plots (all individuals surveyed) and permitted to construct a standardized simulation data set. However, this set of plots was a-typical with respect to extent and clustering (Table 2; Fig. 3). Despite these shortcomings in the data sets, the results indicated that the majority of the assumptions reflect general trends. For example, correcting for p_i merged the curves of data sets 2, 3, and 4 and suggests that the assumptions underlying the estimate of p_i for data set 4 were accurate.

The identification of reduction factors was straightforward for p_i and p_d as these were based on plots and therefore the effect of partial sampling could be simulated over its full range. The effect of spatial distribution of sample points (C) on observed species richness required the extrapolation on the basis of a limited range of data points while variation in the data was high (Fig. 5). By excluding samples from submontane forest, swamp forest and secondary vegetation in the present analyses, beta diversity in the area was kept low. As a result clustering of sample points will only have had a small effect on observed species richness. A more accurate assessment of the effect of spatial distribution of observation is especially important in areas with a high species turnover. The proportion of the extent sampled had the largest impact on observed species richness (Fig. 6). However, its reduction factor was difficult to estimate, as the range of p_e values in the simulation data set was very small and did not cover the range observed in the data sets. While the model $r_e = p_e^{0.6355}/(p_e^{0.6355} + e^{-4.8589})$ fitted the simulation data perfectly (Fig. 5b), the overestimation of predicted spe-

cies richness in all surveys most likely originates from inaccuracies in the estimation of r_e especially at lower values of p_e. However, we expect that the described effects of partial sampling on observed species richness reflect the general trends for the forests in the study site, although the specific formulation of the reduction factors could likely be improved with additional data.

Approaches to estimate true species richness from incomplete samples are amply available, including Jackknife estimators and species accumulation curves (Colwell and Coddington 1994, Gotelli and Colwell 2001, Brose et al. 2003, Melo et al. 2003). However, these approaches do not explicitly acknowledge the five key factors that we identified and they are designed for biodiversity assessments, whereas our approach is also designed for data that are not originally collected for this purpose. Furthermore, in contrast to most previous approaches we specifically focus on species–area (or rather extent) relations. Our Eq. 6 can be viewed as a new formula to estimate species richness which can be calibrated using real data (as we did with our simulation data set constructed from real data) and led to Eq. 7. Therefore no assumptions about relative abundance patterns are needed. The only caveat is that the data set used for calibration must be representative for all spatial scales.

Ongoing large-scale habitat disturbance in tropical forest regions, in combination with limited conservation funds and generally weak law enforcement, urges conservation planners to identify areas of high biological value (Frumhoff and Losos 1998, Myers et al. 2000). A wide range of criteria has been proposed to identify the value of forests for biodiversity conservation, including species diversity and level of endemism (de Groot 1992, Hawthorne 1996, Frumhoff and Losos 1998, Myers et al. 2000). However, assessment of these criteria in forest environments is expensive and time-consuming and so far few regions are sufficiently explored to make accurate estimations for even a limited number of species groups (Howard et al. 1998, 2000, Lawton et al. 1998). In this study, we found that for terrestrial vascular plants total species richness and endemic species richness were highly correlated. Despite the large methodological differences between surveys, the proportion of endemic plant species detected was surprisingly constant (i.e., 14.8–16.6%) in the old growth forests of the study area. Our findings support the analyses of Myers et al. (2000) that at least for plants endemic richness is an accurate indication of total species richness. In addition, recovery of endemic species was much slower in moderately disturbed logging gaps compared to highly disturbed shifting cultivation fields (van Gemerden et al. 2003a). Therefore, focus on endemic species in conservation assessments of tropical forests may be a cost-effective alternative to full species surveys (cf. Lawton et al. 1998).

The survey methods that we studied differed in the selection of individuals included in the assessment. As a result, they also varied considerably in sampling effort. We estimated the average required sampling effort to conduct the described surveys on the basis of field experience. As involvement of scientists is generally the most expensive element in this kind of surveys, we focused on the required scientist input to survey, collect and identify individuals (Table A1 in Appendix). On a per area basis, sampling effort was highest in detailed botanical assessment (data set 2) and the collection of herbarium specimens (data set 4). The lowest effort was required for incomplete botanical assessment (method 3, 10% of the individuals surveyed). Most labor-intensive were field sampling and plant identification. Proportion of specimens that was successfully identified (e.g., to species level) is likely to increase with identification effort. Data sets 1, 2, and 3 showed a similar response to increased sampling effort, while the recorded number of species in data set 4 was significantly lower (Fig. A1 in Appendix). However, the rate of increase was high in data set 4 compared to the other sets. Moreover, the curves of the plot based surveys tended to level off at higher values of sampling effort, while rate of increase remained constant in the herbarium set. This effect was especially strong in data set 1, and even more so for observed endemic species richness. Apparently, the capacity to continue to detect new species declines much more rapidly for surveys that relied partially on the classification of species by non-specialists, i.e., field identifications and morpho-categorization based to some extent on parataxonomy. The framework we propose for reconciliation of botanical assessments implicitly assumes that detection capacity is similar between data sets. Especially surveys based entirely on classifications by non-specialists, such as those promoted for surveys of vertebrates by Oliver and Beatie (1996a, bM), and Kerr et al. (2000), are likely to show different detection trends compared to classifications by specialists. To cope with such methodological differences the present framework for reconciliation requires further development.

CONCLUSIONS

The main focus of the present study is the effect of partial sampling on observed species richness, and how methodologically different surveys of plant diversity in especially tropical forests can be compared. The reconciliation of methodologically different assessments of plant diversity can contribute to our understanding of patterns of diversity and endemism in tropical forest regions that are currently poorly explored. As biodiversity assessments in floristically diverse forests are generally time-consuming and cost-ineffective, making optimal use of all currently available sources of biodiversity information could contribute to both quick and accurate assessment of conservation needs. The suggested method largely contributes to reserve site selection studies where the distribution and richness of species and endemics are crucial prerequisites (e.g.,

Freitag and van Jaarsveld 1997, 1998, Margules et al. 1988). For such studies, incomplete and fragmented data collected by different methods, are often available. This first attempt to reconcile four methodologically different surveys in old growth lowland forest in southern Cameroon suggests that species–area relations can be reconstructed from incomplete sample data if the key characteristics of the methods and the site can be statistically described. With these species–area relations local biodiversity can be estimated and the expected effect of increasing size and amount of reserve sites on the protected biodiversity can be calculated. Reliable documentation of these components of surveys would greatly enhance their use for comparative biodiversity analyses.

Literature Cited

Ando, A., J. Camm, S. Polasky, and A. Solow. 1998. Species distributions, land values, and efficient conservation. Science **279**:2126–2128.

Arrhenius, O. 1921. Species and area. Journal of Ecology **9**: 95–99.

Aubréville, A., and J.-F. Leroy, editors. 1961–1992. Flore du Gabon. Muséum National d'Histoire Naturelle, Paris, France.

Aubréville, A., and J.-F. Leroy, editors. 1963–2001. Flore du Cameroun. Muséum National d'Histoire Naturelle, Paris, France.

Brose, U., N. D. Martinez, and R. J. Williams. 2003. Estimating species richness: sensitivity to sample coverage and insensitivity to spatial patterns. Ecology **84**:2364–2377.

Cable, S., and M. Cheek. 1998. The plants of Mount Cameroon: a conservation checklist. Royal Botanic Gardens, Kew, UK.

Colwell, R. K., and J. A. Coddington. 1994. Estimating terrestrial biodiversity through extrapolation. Philosophical Transactions of the Royal Society of London Series B-Biological Sciences **345**:101–118.

Cressie, N. A. C. 1993. Statistics for spatial data. Wiley, New York, New York, USA.

Csuti, B., S. Polasky, P. H. Williams, R. L. Pressey, J. D. Camm, M. Kershaw, A. R. Kiester, B. Downs, R. Hamilton, M. Huso, and K. Sahr. 1997. A comparison of reserve selection algorithms using data on terrestrial vertebrates in Oregon. Biological Conservation **80**(1):83–97.

Danielsen, F., D. S. Balete, M. K. Poulsen, M. Enghoff, C. M. Nozawa, and A. E. Jensen. 2000. A simple system for monitoring biodiversity in protected areas of a developing country. Biodiversity and Conservation **9**(12):1671–1705.

de Groot, R. S. 1992. Functions of nature. Evaluation of nature in environmental planning, management and decision making. Wolters-Noordhof, Groningen, The Netherlands.

Faith, D. P., and P. A. Walker. 1996. Environmental diversity: on the best-possible use of surrogate data for assessing the relative biodiversity of sets of areas. Biodiversity and Conservation **5**:399–415.

Franqueville, A. 1973. Atlas regional Sud-Ouest 1. République du Cameroun. ORSTOM, Yaoundé, Cameroon.

Freitag, S., and A. S. van Jaarsveld. 1997. Relative occupancy, endemism, taxonomic distinctiveness and vulnerability: prioritising regional conservation actions. Biodiversity and Conservation **6**:211–232.

Freitag, S., and A. S. van Jaarsveld. 1998. Sensitivity of selection procedures for priority conservation area to survey extent, survey intensity, and taxonomic knowledge. Proceedings of the Royal Society of London Series B **265**: 1475–1482.

Frumhoff, P. C., and E. C. Losos. 1998. Setting priorities for conserving biological diversity in tropical timber production forests. A policy report. The Union of Concerned Scientists, The Center for Tropical Forest Science, Smithsonian Institution, Washington, D.C., USA.

Funk, V. A., Z. M. F. Zermoglio, and N. Nasir. 1999. Testing the use of specimen collection data and GIS in biodiversity exploration and conservation decision making in Guyana. Biodiversity and Conservation **8**:727–751.

Gaston, K. J. 1994. Rarity. Chapman and Hall, London, UK.

Gaston, K. J., and A. S. L. Rodrigues. 2003. Reserve selection in regions with poor biological data. Conservation Biology **17**(1):188–195.

Gleason, H. A. 1922. On the relation between species and area. Ecology **3**:158–162.

Gotelli, N. J., and R. K. Colwell. 2001. Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. Ecology Letters **4**:379–391.

Guedje, N. M. 2002. La gestion des populations d'arbres comme outil pour une exploitation durable des produits forestiers non-ligneux: l'exemple de *Garcinia lucida* (sud Cameroun). Tropenbos Cameroon Series 5. The Tropenbos-Cameroon Programme, Kribi, and Université Libre de Bruxelles, Brussels, Belgium.

Hawthorne, W. D. 1996. Holes and the sums of parts in the Ghanaian forest: regeneration, scale and sustainable use. Proceedings of the Royal Society of Edinburgh **104B**:75–176.

Hellier, A., A. C. Newton, and S. O. Gaona. 1999. Use of indigenous knowledge for rapidly assessing trends in biodiversity: a case study from Chiapas, Mexico. Biodiversity and Conservation **8**:869–889.

Hof, J., and M. Bevers. 1998. Spatial optimization for managed ecosystems. Columbia University Press, New York, New York, USA.

Howard, P. C., T. R. B. Davenport, F. W. Kigenyi, P. Viskanic, M. C. Baltzer, C. J. Dickinson, J. Lwanga, R. A. Matthews, and E. Mupada. 1998. Protected area planning in the tropics: Uganda's national systems of forest nature reserves. Conservation Biology **14**:858–875.

Howard, P. C., P. Viskanic, T. R. B. Davenport, F. W. Kigenyi, M. Baltzer, C. J. Dickinson, J. S. Lwanga, R. A. Matthews, and A. Balmford. 2000. Complementarity and the use of indicator groups for reserve selection in Uganda. Nature **394**:472–475.

Hubbell, S. P. 2001. The unified neutral theory of biodiversity and biogeography. Monographs in Population Biology 32. Princeton University Press, Princeton, New Jersey, USA.

Keay, R. J. W., and F. N. Hepper, editors. 1954–1972. Flora of west tropical Africa. Second edition. Crown Agents, London, UK.

Kerr, J. T., A. Sugar, and L. Packer. 2000. Indicator taxa, rapid biodiversity assessment and the nestedness in an endangered ecosystem. Conservation Biology **14**(6):1726–1734.

Lawton, J. H., et al. 1998. Biodiversity inventories, indicator taxa and the effects of habitat modification in tropical forest. Nature **391**:72–76.

Letouzey, R. 1968. Etude phytogégraphique du Cameroun. Ed. P. Lechavelier, Paris, France.

Letouzey, R. 1985. Notice de la carte phytogéographique du Cameroun au 1:500 000. Institut de la Carte Internationale de la Végétation, Toulouse, France.

Linder, H. P. 2001. Plant diversity and endemism in sub–Saharan Tropical Africa. Journal of Biogeography **28**(2): 169–182.

Lombard, A. T., R. M. Cowling, R. L. Pressey, and P. J. Mustart. 1997. Reserve selection in a species-rich and fragmented landscape on the Agulhas Plain, South Africa. Conservation Biology **11**(5):1101–1116.

Lovett, J. C., S. Rudd, J. Taplin, and C. Frimodt-Moller. 2000. Patterns of plant diversity in Africa south of the Sahara and their implications for conservation management. Biodiversity and Conservation **9**(1):37–46.

Margules, C. R., A. O. Nicholls, and R. L. Pressey. 1988. Selecting networks of reserves to maximize biological diversity. Biological Conservation **43**(1):63–76.

Margules, C. R., and R. L. Pressey. 2000. Systematic conservation planning. Nature **405**:243–253.

Melo, A. S., R. S. Pereira, A. J. Santos, G. J. Shepherd, G. Machado, H. F. Medeiros, and R. J. Sawaya. 2003. Comparing species richness among assemblages using sample units: why not use extrapolation methods to standardize different sample sizes? Oikos **101**(2):398–410.

Mickleburgh, S., A. M. Hutson, P. A. Racey, andIUCN/SSC Chiroptera Specialist Group. 1992. Old World fruit bats. An action plan for their conservation. IUCN, Cambridge, UK.

Myers, N., R. A. Mittermeier, C. G. Mittermeier, G. A. B. da Fonseca, and J. Kent. 2000. Biodiversity hotspots for conservation priorities. Nature **403**:853–858.

Nelson, B. W., C. A. C. Ferreira, M. F. Dasilva, and M. L. Kawasaki. 1990. Endemism centers, refugia and botanical collection density in Brazilian Amazonia. Nature **345**:714–716.

Oates, J. F. 1996. African primates: status survey and conservation action plan (revised edition). IUCN, Cambridge, UK.

Olff, H., M. E. Ritchie, and H. H. T. Prins. 2002. Global environmental controls of diversity in large herbivores. Nature **415**:901–904.

Oliver, I., and A. J. Beattie. 1996a. Invertebrate morphospecies as surrogates for species: a case study. Conservation Biology **10**(1):99–109.

Oliver, I., and A. J. Beattie. 1996b. Designing a cost-effective invertebrate survey: a test of methods for rapid assessment of biodiversity. Ecological Applications **6**:594–607.

Olivry, J. C. 1986. Fleuves et rivières du Cameroun. Collection Monographies Hydrologiques d'Orstom 9. Mesres-Orstom, Paris, France.

Poorter, L., F. Bongers, F. N. Kouamé, and W. D. Hawthorne, editors. 2004. Biodiversity of West African forests: an ecological atlas of woody plant species. CAB International, Oxfordshire, UK.

Preston, F. W. 1960. Time and space variation of species. Ecology **41**:611–627.

Rosenzweig, M. L. 1995. Species diversity in space and time. Cambridge University Press, Cambridge, UK.

Scott, J. M., F. Davis, B. Csuti, R. Noss, B. Butterfield, C. Groves, H. Anderson, S. Caicco, F. Derchia, T. C. Edwards, J. Ulliman, and R. G. Wright. 1993. Gap analysis—a geographic approach to protection of biological diversity. Wildlife Monographs **123**:1–41.

Shmida, A., and M. V. Wilson. 1985. Biological determinants of species diversity. Journal of Biogeography **12**:1–20.

Simberloff, D. 1998. Flagships, umbrellas and keystones: is single-species management passé in the landscape era? Biological Conservation **83**:247–257.

ter Steege, H. 1998. The use of forest inventory data for a National Protected Area Strategy in Guyana. Biodiversity and Conservation **7**:1457–1483.

ter Steege, H., editor. 2000. Plant diversity in Guyana—with recommendations for a national protected area strategy. Tropenbos Series 18. The Tropenbos Foundation, Wageningen, The Netherlands.

ter Steege, H., M. J. Jansen-Jacobs, and V. K. Datadin. 2000. Can botanical collections assist in a National Protected Area Strategy in Guyana? Biodiversity and Conservation **9**:215–240.

van Gemerden, B. S., and G. W. Hazeu. 1999. Landscape ecological survey (1:100 000) of the Bipindi–Akom II–Lolodorf region, southwest Cameroon. Tropenbos-Cameroon Documents 1. Tropenbos Cameroon Programme, Wageningen, The Netherlands.

van Gemerden, B. S., H. Olff, M. P. E. Parren, and F. Bongers. 2003b. The pristine forest? Remnants of historical human impacts on current tree species composition and diversity. Journal of Biogeography **30**:1381–1390.

van Gemerden, B. S., G. N. Shu, and H. Olff. 2003a. Recovery of conservation values in Central African forest after logging and shifting cultivation. Biodiversity and Conservation **12**:1553–1570.

van Langevelde, F., A. Schotman, F. Claassen, and G. Sparenburg. 2000. Competing land use in the reserve site selection problem. Landscape Ecology **15**(3):243–256.

Vitousek, P. M., H. A. Mooney, J. Lubchenco, and J. M. Melillo. 1997. Human domination of Earth's ecosystems. Science **277**:494–499.

Waterloo, M. J., J. C. Ntonga, A. J. Dolman, and A. B. Ayangma. 2000. Impact of shifting cultivation and selective logging on the hydrology and erosion of forest land in south Cameroon. Tropenbos-Cameroon Documents 3. Tropenbos-Cameroon Programme, Wageningen, The Netherlands.

APPENDIX

A table and figure showing the scientific effort made in surveying, processing, and identifying plants in four different surveys and the relation of effort to diversity estimates is presented in ESA's Electronic Data Archive: *Ecological Archives* A015-051-A1.