

## Taxonomic survey compared to ecological sampling: are the results consistent for woodland epiphytes?

C. J. ELLIS and B. J. COPPINS

**Abstract:** Field survey by a taxonomist or specialist biologist ('taxonomic survey') provides a comprehensive inventory of species in a habitat. Common and conspicuous species are rapidly recorded and search effort can be targeted to inconspicuous or rare species. However, the subjective nature of taxonomic survey limits its usefulness in ecological monitoring and analysis. In contrast, 'ecological sampling', focused on the standardized use of repeated sub-units such as quadrats, is designed to quantify the observational error of results, allowing for more robust statistical treatment. Nevertheless, the spatial extent of recording will be lower during ecological sampling, and rarities might be missed. Despite their differences, these two approaches are often assumed to be congruent for decision making. Taxonomic survey is commonly used to identify priority sites for conservation (including species-rich sites, or those with many rare/threatened species) while ecological sampling is used to design conservation strategy by relating species richness or composition to habitat dynamics. If these contrasting approaches are indeed congruent, then trends in species richness and community composition, detected by ecological sampling, will mirror the results of taxonomic survey so that management confidently protects the attributes for which a site was prioritized. This study performed both taxonomic survey and ecological sampling for lichen epiphytes in 13 woodland study sites in Scotland. To understand the procedure of taxonomic survey, fieldwork by a professional taxonomist was structured by effort into 15-minute time intervals. As expected, taxonomic survey discovered more species per site, while ecological sampling (allowing a measure of species frequency) resolved greater variation in community composition. However, the patterns of richness and species composition obtained from the different methods were correlated, suggesting an overall high degree of congruence in identifying and then managing priority sites. Furthermore, when exploring the taxonomic survey in detail, we found that a minimum effort of 45 minutes was required to accurately determine species richness differences among contrasting woodland sites.

**Key words:** accumulation curve, inventory, lichens, quadrat, randomized resampling, sampling effort

*Accepted for publication 26 October 2016*

---

### Introduction

Conservation valuation often uses species data to recognize 'biodiversity hotspots' with concentrated species richness (Reid 1998; Myers *et al.* 2000). The recognition of hotspots also incorporates conservation priorities; for example within the UK, localities important for lichens have been selected for protection as Sites of Special Scientific Interest based on the numbers of nationally rare and scarce species (Hodgetts 1992). For each of these metrics (richness/endemism, or threat status), accuracy depends on the level

of expertise that is employed in generating species occurrence and distributional data (Ahrends *et al.* 2011). A gold standard in generating species data is habitat survey by an expertly-trained professional taxonomist or specialist field biologist. By collating the results of this high quality 'taxonomic survey' over time, and across different sites in the region of interest, biogeographical patterns will emerge and metrics such as species rarity can be assessed. Nevertheless, the behaviour of taxonomic surveyors can be idiosyncratic, causing bias. On a larger landscape-scale this includes the spatial effect of the taxonomists' home or work location (Moerman & Estabrook 2006), ease of access to survey sites (Kadmon *et al.* 2004), combined with a preference for examining localities of known

---

C. J. Ellis and B. J. Coppins: Royal Botanic Garden Edinburgh, 20A Inverleith Row, Edinburgh, EH3 5LR, UK. Email: c.ellis@rbge.org.uk

or perceived richness (Dennis & Thomas 2000). On smaller ecological scales, for example in recording stands or large plots, subjectivity in taxonomic survey work makes ecological inference highly problematic because the variability in results generated by different surveyors (observational error) can swamp the variance related to an environmental signal (Archaux *et al.* 2006, 2009).

For these reasons, 'ecological sampling' tends to employ a standardized repeated sub-sampling regime (Kent & Coker 1992; Kent 2012). By focusing measurements across multiple small-scale units (e.g. points or quadrats), sub-sampling can more effectively quantify and reduce the observational error for species abundance or frequency in a habitat, compared to visual assessments made at a larger scale (Sykes *et al.* 1983; Vittoz *et al.* 2010). However, ecological sampling comes with costs and there are important trade-offs between the survey and sub-sampling approaches (McCune & Lesica 1992). Ecological sampling typically records species over a lower spatial area per unit time. Standard models for species abundance lead to the expectation that species with low abundance/frequency will remain hidden beyond a 'veil line' (Preston 1948; Magurran 2004); rarer but potentially important species in a habitat are likely to go unrecorded when using sub-samples (Gray & Azuma 2005; Vittoz & Guisan 2007). Despite the weak detection of rare species, richness values estimated over multiple small quadrats can be more consistent among observers, with less bias, than for larger plots (Archaux *et al.* 2007).

Considering their attributes, there are clearly important differences in the benefits and costs of taxonomic survey and ecological sampling. If the aim is for as complete a species list as possible in a site, taxonomic survey is suited to provide this as a qualitative inventory. A taxonomic expert will quickly record common and conspicuous species and will be able to spend time purposefully locating inconspicuous or rare species, in order to produce a comprehensive list. Ecological sampling will record fewer species but

can better quantify and reduce observational error in the occurrence, abundance/frequency or richness of species, when relating these to environmental controls. While taxonomic survey aims to identify the sites in a landscape that are important to conserve (Reid 1998; Myers *et al.* 2000), ecological sampling aims to provide information on how habitat conservation might be designed, through an understanding of the relationship between community richness and composition, and habitat dynamics (Jüriado *et al.* 2009; Leppik *et al.* 2011). Given these complementary aims, it is important that information provided by each method is indeed congruent; ecological sampling used to understand habitat management should be detecting and explaining the same trends in species richness, composition and site conservation status revealed by taxonomic survey. Alternatively, one might propose a relationship in which taxonomic survey captures common through to rarer species, with ecological sampling skewed only towards common species, disrupting the congruence between these methods. This paper directly compares the results of taxonomic survey and ecological sampling among 13 contrasting woodland sites, to investigate the degree of congruence between the two approaches. We compared epiphyte richness and species composition among taxonomic survey and ecological sampling methods, using species accumulation curves to reduce bias in survey/sampling effort (survey time, or number of trees sampled per stand). We ask the following questions: (1) to what extent do species richness estimates match (in absolute terms) or correlate between taxonomic survey and ecological sampling, either as raw values or with reduced bias for survey/sampling effort using accumulation curves, (2) do patterns in community composition derived from taxonomic survey (species presence-absence) map onto those derived from ecological sampling (frequency values) and (3) given recent interest in time-to-detection as a proxy for a species' frequency (Bornard *et al.* 2014), is there a relationship between the time-to-detection during taxonomic survey, and species frequency estimated using ecological sampling?

These questions are also caveated by investigating the predictive accuracy of taxonomic survey, when estimating ‘total richness’ for a circumscribed woodland stand. We ask, for example, whether the raw number of species observed is an exact measure of, or direct linear function of, total species richness estimated using species accumulation curves, which would imply consistency in survey outcome across multiple sites with contrasting attributes.

## Methods

### Study sites

Thirteen study sites were selected to include a range of woodland types broadly representative of semi-natural habitats in present-day Scotland, positioned over major gradients of environmental variation (Fig. 1). The sites varied in terms of their climate and tree species composition, ranging structurally from ancient woodlands with old-growth characteristics (e.g. sites 9 & 10, Glen Affric & Glen Loy) through to regenerated secondary woodland (e.g. site 6, Coulmony), and from situations with a continuous or gladed canopy to parkland-type or wayside systems (e.g. site 13, Taynish Compartment 5). Site survey/sampling was limited to a single structurally homogenous stand of *c.* 0.25 ha. For continuous canopy woodland these were interior situations (no edge effects),

and for all selected stands there was freedom of movement without a dense understorey layer, and with field-work conducted on dry days.

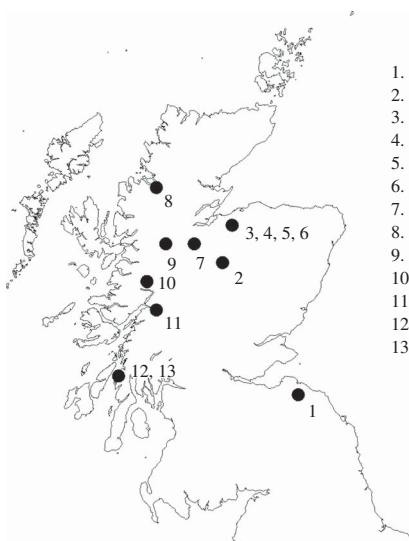
### Taxonomic survey

Having circumscribed the boundaries of a discrete woodland stand at each site, a lichen taxonomist (BJC) aimed to record all epiphytic lichen species, defined in a broad sense; corticolous species on the visible range of tree boles, shrubs and sampled from branches/twigs including debris. The duration of the survey was divided into 15-minute intervals, allowing the accumulation of species to be plotted against survey effort measured as consecutive time-intervals. Lichen recording stopped when the taxonomist considered the species survey to be complete, but with a minimum duration of 60 minutes.

Two values of species richness were calculated per stand. First, the observed number of species recorded by the taxonomist when the survey was considered completed. Second, the value of species richness estimated at the asymptote for each stand, which was derived by fitting an accumulation curve to the increase in species richness over the duration of the survey, using a three-parameter rational function (Tjørve 2003; Dengler 2009):

$$y = (1 + ax) / (b + cx) \quad (\text{Eq. 1})$$

Where *y* is the dependent variable (species richness), *x* is the independent variable (survey effort in 15-minute intervals), and *a*, *b* and *c* are constants. The rational



1. Bell Wood (55.862°N; 2.524°W): birchwood, 901 mm, 1.9 °C
2. Kinara (57.166°N; 3.848°W): salix-alder carr, 831 mm, 1.7 °C
3. Findhorn, Dulsie Bridge 1 (57.446°N; 3.783°W): hazelwood, 855 mm, 1.8 °C
4. Findhorn, Dulsie Bridge 2 (57.445°N; 3.784°W): riparian alder, 855 mm, 1.8 °C
5. Findhorn, Glen Ferness (57.469°N; 3.779°W): aspen stand, 855 mm, 1.8 °C
6. Findhorn, Coulmony (57.514°N; 3.726°W): birchwood, 777 mm, 2.6 °C
7. Inverfarigaig (57.279°N; 4.446°W): riparian ashwood, 950 mm, 2.7 °C
8. Dundonnell (57.836°N; 5.185°W): wayside oak, 1945 mm, 1.9 °C
9. Glen Affric (57.302°N; 4.891°W): birchwood, 1365 mm, 1.7 °C
10. Glen Loy (56.911°N; 5.128°W): pinewood, 2310 mm, 0.5 °C
11. Lost Valley (56.657°N; 4.99°W): birchwood, 2890 mm, 0.4 °C
12. Taynish 1, Bar Mor (56.003°N; 5.638°W): oakwood, 1560 mm, 4.8 °C
13. Taynish 2, Compartment 5 (55.989°N; 5.648°W): wayside sycamore, 1560 mm, 4.8 °C

FIG. 1. Thirteen study sites selected for the comparison between taxonomic survey and ecological sampling of lichen epiphytes. Latitude and longitude are given in parentheses, followed by the dominant tree species, mean annual precipitation and minimum mean temperature for the coldest month using baseline climate data at a 5 km scale (1961–2006).

function describes a hyperbolic (asymptotic) curve assuming therefore that accumulation of species richness during a survey will, on average, be rapid initially and decrease over consecutive time-intervals. Species richness is thus expected to reach the asymptote, at which no further species are detected despite additional survey effort (all species are recorded for the stand). This 'total richness' might be directly quantified as a series of consecutive intervals during which no new species are recorded (in which case the asymptote will be fitted to observed data), or else inferred by extrapolating the fitted curve to the asymptote, reflecting an estimate based on the trend in species accumulation for the observed data. Rational functions were fitted using SigmaPlot v. 11 (Systat 2010).

The procedure of taxonomic survey was investigated by correlating the observed species richness values per stand with the estimated total richness at the asymptote, to explore deviance in potentially unrecorded species. Assuming species accumulation is robustly described using the asymptotic curve-fitting then the following example trends may be detected. First, the survey may stop consistently within the asymptote, based on multiple intervals that yield no new species, and with observed richness equivalent to and providing a direct measure of estimated total richness. Second, if a consistent stopping rule is exerted once the discovery of new species falls below an absolute threshold (i.e. fewer than  $x$  species per time-interval), then observed species richness may correlate with, yet be uniformly lower in absolute terms than, estimated total richness (corresponding to a linear correction factor). Third, a stopping rule may be consistently applied but in relative terms (i.e. at a given percent increase in recorded species, relative to the rate of addition over prior time-intervals) so that observed species richness correlates with, but is non-uniformly related to, absolute values of estimated total richness (corresponding to a non-linear correction factor). Fourth, the observed species richness may be uncorrelated with estimated total richness, corresponding to a stochastic and unpredictable stopping rule for the surveyor.

We also correlated the values of species richness for a given survey effort with values at the asymptote for each stand, to identify the minimum survey effort required to accurately differentiate between the stands in terms of their estimated total richness.

### Ecological sampling

Ecological sampling was carried out by an ecologist (CJE). Sampling proceeded from a point in the centre of the stand, and following a randomized compass bearing the first tree encountered was sampled. The lichen epiphyte community was sampled from the bark surface of the tree bole as presence-absence in a  $7 \times 7$  cm quadrat; this quadrat size made it possible to sample smaller trees, and enabled the number of quadrats to scale more easily with the girth of the tree, so that sampling effort could be correlated with habitat area (product-moment  $r = 0.851$ ;  $P < 0.0001$ , with 108 df.). Quadrats were

positioned randomly on the tree bole with respect to aspect and height between 30 and 175 cm, using a series of constrained random numbers (corresponding to potential aspect and height positions). Having sampled a tree, further random compass bearings were followed, sampling from tree to tree. The ecological sampling was subject to a time constraint matching approximately that of taxonomic survey, by allowing the completion of sampling for a tree when the taxonomist declared the survey to be complete. In practice this lasted for a maximum of c. 1 h beyond the duration of the taxonomic survey.

Two species richness values were calculated for ecological sampling. First, the raw number of species recorded across all the quadrats per stand. Second, richness values were corrected for sampling effort (number of trees) by resampling an accumulation curve (Ellis & Ellis 2013). To do this, a list of lichen epiphyte species per tree was derived from the sampled quadrats. The richness of species per stand was then accumulated among consecutive trees, with randomized resampling 100 times. Species richness at the asymptote was derived by fitting a three-parameter rational function to reconstruct a hyperbolic asymptotic curve (Eq. 1, above). To build confidence in this method, the results from the species accumulation curves were compared to six alternative estimators for total species richness that seek to account for unsampled species, calculated in EstimateS v. 9.1 (Colwell 2013): extrapolation of Colwell *et al.*'s (2012) species accumulation curve, a Michaelis-Menten richness estimator (Raaijmakers 1987), Chao 2 (Chao 1987), the incidence-based coverage estimator, ICE (Chazdon *et al.* 1998) and Jackknife 1 & 2 (Heltshe & Forrester 1983; Smith & van Belle 1984), each with 100 randomizations to estimate 95% confidence intervals.

### Observer bias and herbarium examination

It was our aim to explore how methodological differences between taxonomic survey and ecological sampling affected information on species richness and composition and, in order to do so, to minimize the potential for observer bias that was unrelated to the method used.

Thus, the two fieldworkers (BJC & CJE) might be expected to have contrasting abilities to find and identify obscure or cryptic species, and/or may apply a divergent and unequal search effort towards species with certain attributes. We note, however, that the less constrained survey method, which is potentially the most open to observer bias, was undertaken by the more experienced fieldworker (BJC) who has >40 years' experience recording lichens in Scotland, and is a UK taxonomic authority (Coppins *et al.* 2002; Smith *et al.* 2009). Nevertheless, the survey approach did include the routine collection of specimens for herbarium examination which, based on experience, would be injudicious to identify in the field. Furthermore, an experienced researcher (c. 10 years' field experience recording lichens in Scotland) was responsible for the quadrat

sampling (CJE) which focused the search effort into a far smaller area, so that when thoroughly examined with a hand-lens (combined with speculative sampling of bark crevices etc), minimized the chance of species being overlooked. In practice, field identification during quadrat sampling was limited to species that can be easily named in the field using current British concepts (e.g. *Evernia prunastri*, *Hypogymnia physodes*, *Ramalina farinacea* etc.), with a majority of the morphological variability encountered in a given quadrat (e.g. sterile crusts) routinely sampled for herbarium examination.

Species for herbarium examination were identified using light microscopy, chemical spot tests and thin-layer chromatography (Orange *et al.* 2001). Unless otherwise stated, nomenclature follows Smith *et al.* (2009) and concepts were standardized through the discussion of all taxonomic entities (field identified, or sampled) between BJC and CJE.

### Statistical analysis

To compare trends in species richness we used the root-mean-square error (RMSE) to measure deviation in estimated richness values between the taxonomic survey and the ecological sampling, for their raw data and asymptotic values. To estimate nestedness between the taxonomic survey and ecological sampling, we also calculated the number of unique species that were not recorded by the alternative method.

We tested whether the time-to-detection during taxonomic survey is a proxy for species frequency measured using ecological sampling. Frequency values were transformed to normality using an arcsine-square-root-transformation and compared to the time-interval during which a species was recorded, using an unbalanced ANOVA with site identity as a blocking effect (calculated using the base statistical package in R (2013)). Data were presented as box plots and averaged as the mean frequency values per time-interval among stands.

Trends in species composition were based on species presence-absence for the taxonomic survey, and to provide an appropriate contrast were based on frequency of occurrence for the ecological sampling. In each case, ordination by nonmetric multidimensional scaling (NMDS) used the Sorenson's (Bray-Curtis) dissimilarity metric to construct a pairwise distance matrix among the stands (McCune & Grace 2002). An optimum NMDS solution was identified in PC-ORD v.6 (McCune & Mefford 2011), based on 250 separate runs for one to six axes with up to 500 iterations, and selecting the solution which minimized instability assessed over 100 runs (with an instability criterion of 0.00001), minimized stress, and maximized significance against a permutation test with 250 randomizations (McCune & Grace 2002; McCune & Mefford 2011). The congruence in species composition, when compared between the survey and sampling methods, was examined using a pairwise correlation of their respective ordination scores (Pearson's product-moment  $r$ , implemented using the base statistical package in R (2013)).

### Results

The taxonomic survey recorded a total of 281 species across the 13 woodland stands (Supplementary material, available online). The mean richness value was 67, varying between 45 (Bell Wood) and 94 (Compartment 5, Taynish). The rate at which species were added during taxonomic survey declined with increased survey effort (Fig. 2A). Accumulation curves plotted to describe the rate of species addition were all highly significant, in all cases adjusted- $r^2 > 0.99$  and  $P < 0.0001$  (Fig. 2B), suggesting that curve-fitting using the hyperbolic rational function provided a reasonable estimate for total species richness at the asymptote. The estimated richness at the asymptote was higher than, though strongly correlated with, the observed value of species richness (Fig. 2C). However, the number of potential unrecorded species increased with the estimated total richness, suggesting that a stopping rule applied during taxonomic survey will meet the requirement of a non-linear correction factor. With exhaustive survey effort, between 8 and 52 additional species may have been recorded from individual stands during taxonomic survey (for Bell Wood and Glen Loy, respectively), representing an increase of 17.8% and 55.3% over the observed richness values. Comparing the number of species recorded for a given survey effort with the estimated total richness at an asymptote, we can conclude that a survey effort of 45 minutes (three consecutive 15-minute intervals) appeared to be the minimum required to accurately determine richness differences among the woodland stands (Fig. 2D). A survey effort that is shorter than 45 minutes remained in a phase of rapid species accumulation, when data recording was saturated. Differences in species richness will be poorly resolved until the survey leaves the point at which additional species are continuously being recorded and an effort to discern rarer niche-specialist species starts to reveal trends among the different stands.

Ecological sampling recorded 175 species across the woodland stands (Supplementary material, available online) and, comparing the

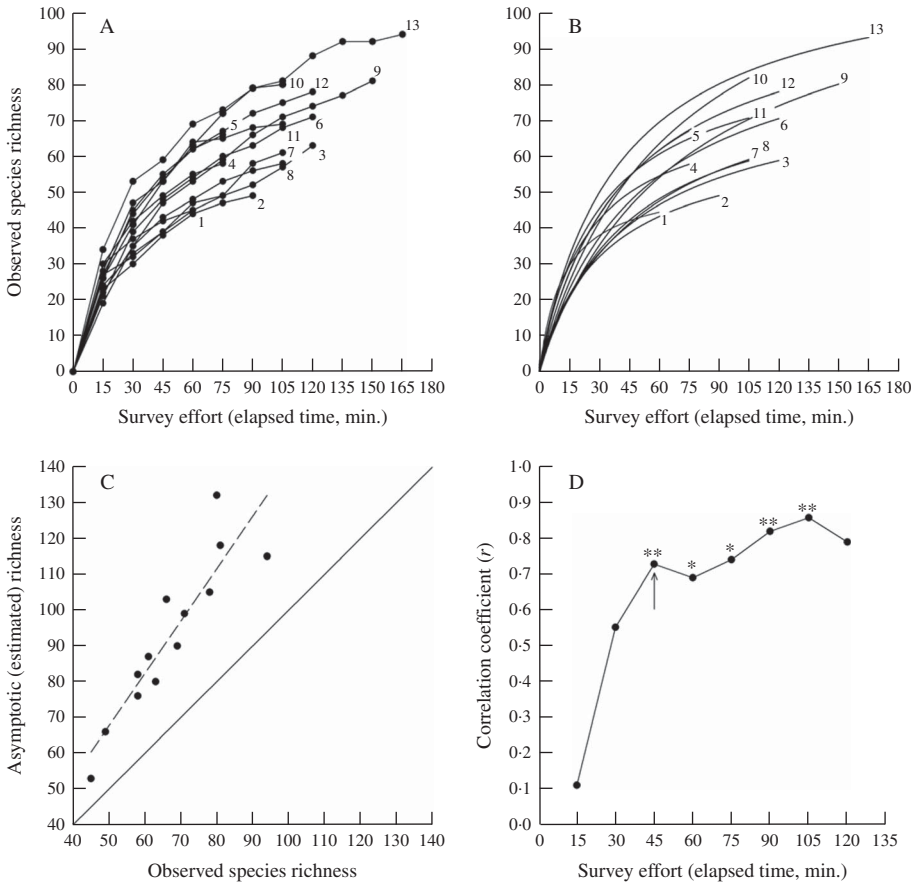


FIG. 2. A, observed accumulation of lichen species with taxonomic survey effort; B, rational functions fitted to the observed data (cf. A); C, relationship between observed species richness at completion of the inventory (cf. A) and estimated total species richness as the asymptote derived from modelled accumulation curves (cf. B), together with line of equality (solid) and line of best fit (dashed;  $r = 0.906, P < 0.001$  with 11 df.); D, correlation coefficients for the observed richness at different survey efforts (elapsed time) and the asymptotic richness ( $* = P < 0.01$ ,  $** = P < 0.005$ ). The arrow indicates the minimum viable survey effort to best distinguish differences in richness among woodland stands. Numbers labelling line graphs in A and B refer to site numbers quoted in Fig. 1.

raw species richness values for the alternative methods, the taxonomic survey consistently recorded more species per stand than ecological sampling over an equivalent time period (Fig. 3A). However, the ecological sampling results were not simply a nested subset of the species found by taxonomic survey. The use of quadrats detected unique species across each of the stands, though this was highly variable and ranged from 2 to 17 species not detected by the taxonomic survey (Fig. 3B). On average, the number of species uniquely recorded by ecological sampling was 18.2% of the number

of unique species recorded by taxonomic survey (varying from 5.1% to 43.5%). Assuming the asymptote of the taxonomic survey is the expected total number of species in a stand (Fig. 2B), then the estimation of species richness by ecological sampling was improved by *c.* 30% when using multiple randomizations to accumulate samples between trees, and extrapolating to the asymptote, compared to species richness sampled from the quadrat data alone (Fig. 3C) (i.e. generating a lower root-mean-square error (RMSE)). These estimates, which were made using a rational function (Fig. 4),

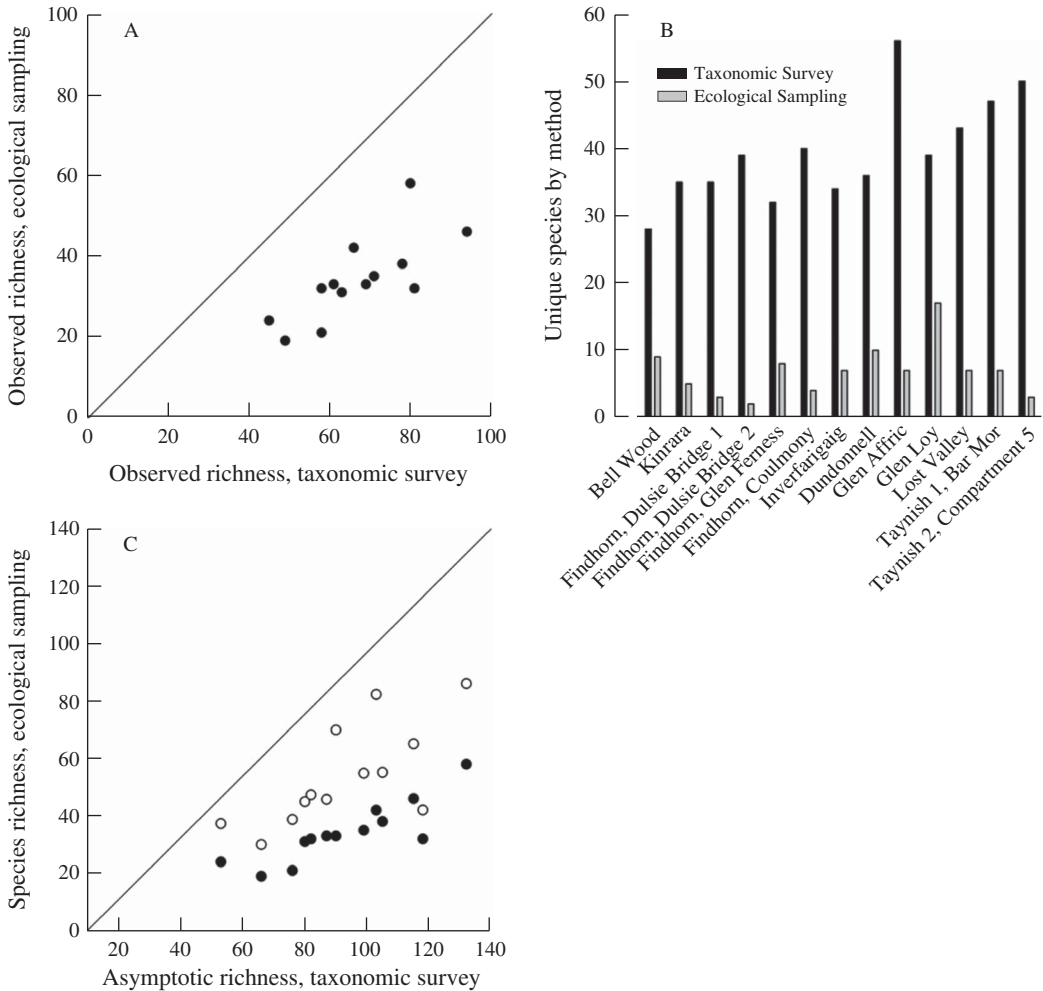
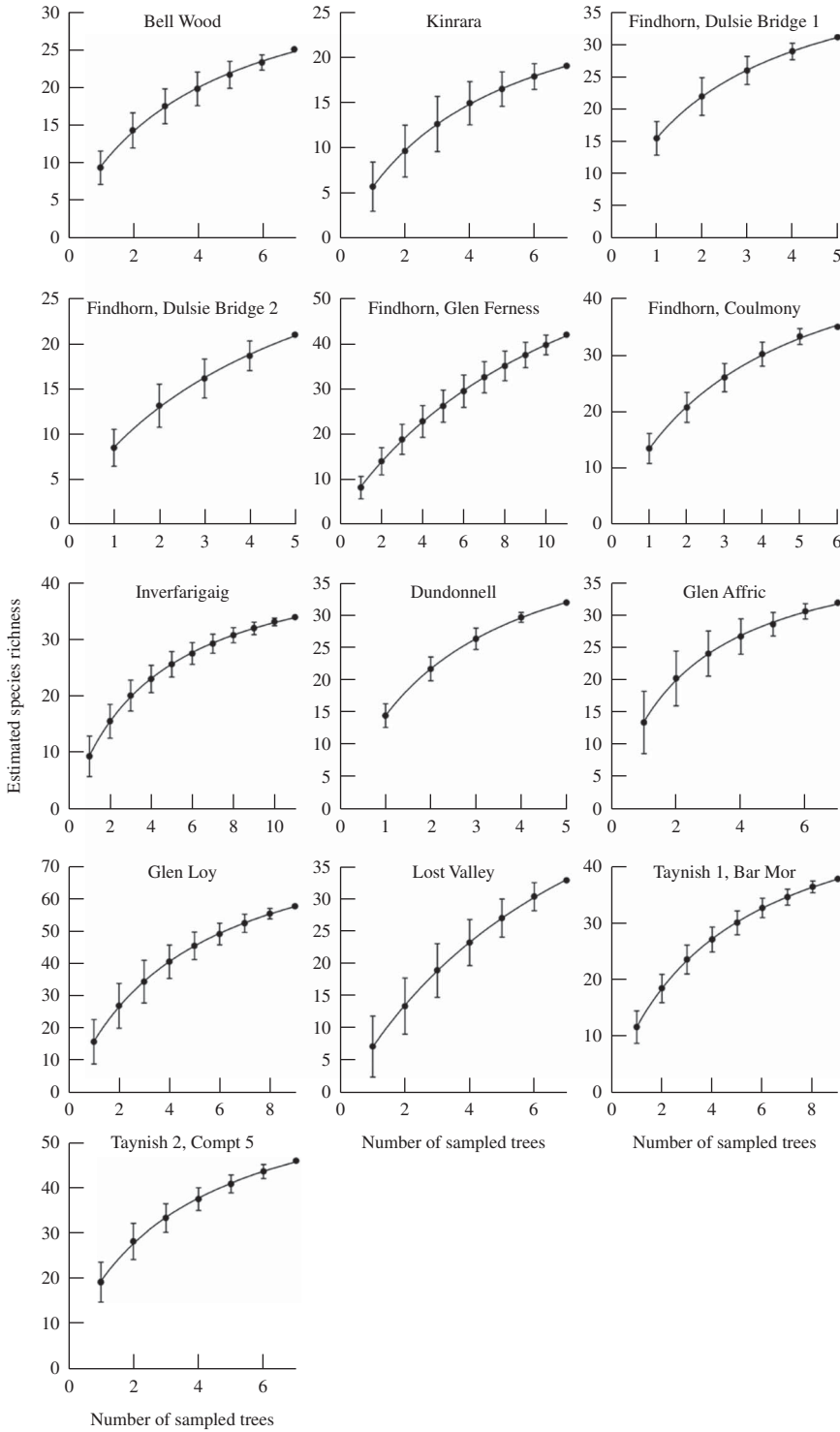


FIG. 3. A, comparison of observed values in species richness for the taxonomic survey and the ecological sampling; B, number of unique species contributed by each method per stand (i.e. the number of species not recorded at the same site by the alternative method); C, comparison of estimated total species richness for a stand (the asymptote for the taxonomic survey, cf. Fig. 2B) with the respective observed (closed circles, RMSE = 60.18) and resampled (asymptotic) (open circles, RMSE = 41.80) values for the ecological sampling. Solid lines in A & C are lines of equality.

fell within a broad range for alternative species richness metrics (Fig. 5). Values tended to be slightly lower, at relatively higher richness values, than estimates made using Colwell *et al.*'s (2012) accumulation curve, Chao 2 or Jackknife 1, yet tended to be higher at these same values than for the Michaelis-Menten estimator, while matching closely across a range of richness values with ICE and Jackknife 2.

Ordination by nonmetric multidimensional scaling (NMDS) was used to compare patterns

of community composition derived from taxonomic survey (species presence-absence in a stand) and ecological sampling (species frequency values). Taxonomic survey generated a statistically significant NMDS solution (instability < 0.00001, stress = 9.7,  $P < 0.005$ ) with two axes (variation for axis one = 65.7%, axis two = 21.4%, cumulative = 87.2%). Ecological sampling generated a statistically significant NMDS solution (instability < 0.00001, stress = 6.04,  $P < 0.01$ ) that had three





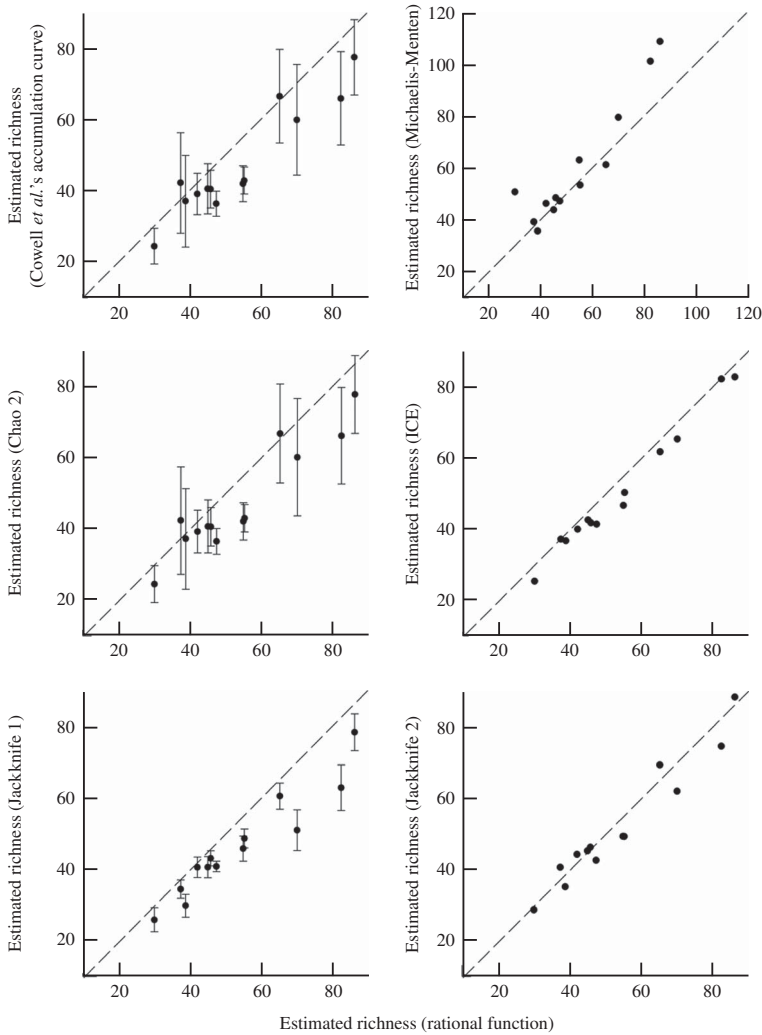


FIG. 5. Comparison of estimated species richness at an asymptote for species accumulation curves employed in this study (cf. Fig. 4) with six alternative estimators popularly applied in the diversity literature. Lines of equality shown as dashed lines.

orthogonal axes (variation for axis one = 56.1%, axis two = 14.5%, axis three = 15.6%, cumulative = 87.2%). Site scores along NMDS axis one for the taxonomic survey were significantly correlated with their equivalent site scores for the ecological sampling (Fig. 6), and this was also the case for NMDS axis two

scores. However, a third NMDS axis recovered for the ecological sampling appeared unique, and was not correlated with site scores for either axis one or two of the taxonomic survey.

Finally, there appeared to be a relationship between the time-interval in which a species

FIG. 4. Resampled species accumulation based on quadrat sub-sampling for the 13 woodland stands (cf. Fig. 1), fitted with a three-parameter rational function to describe an asymptotic curve. In all cases adjusted- $r^2 = 0.99$  with  $P < 0.00001$ .

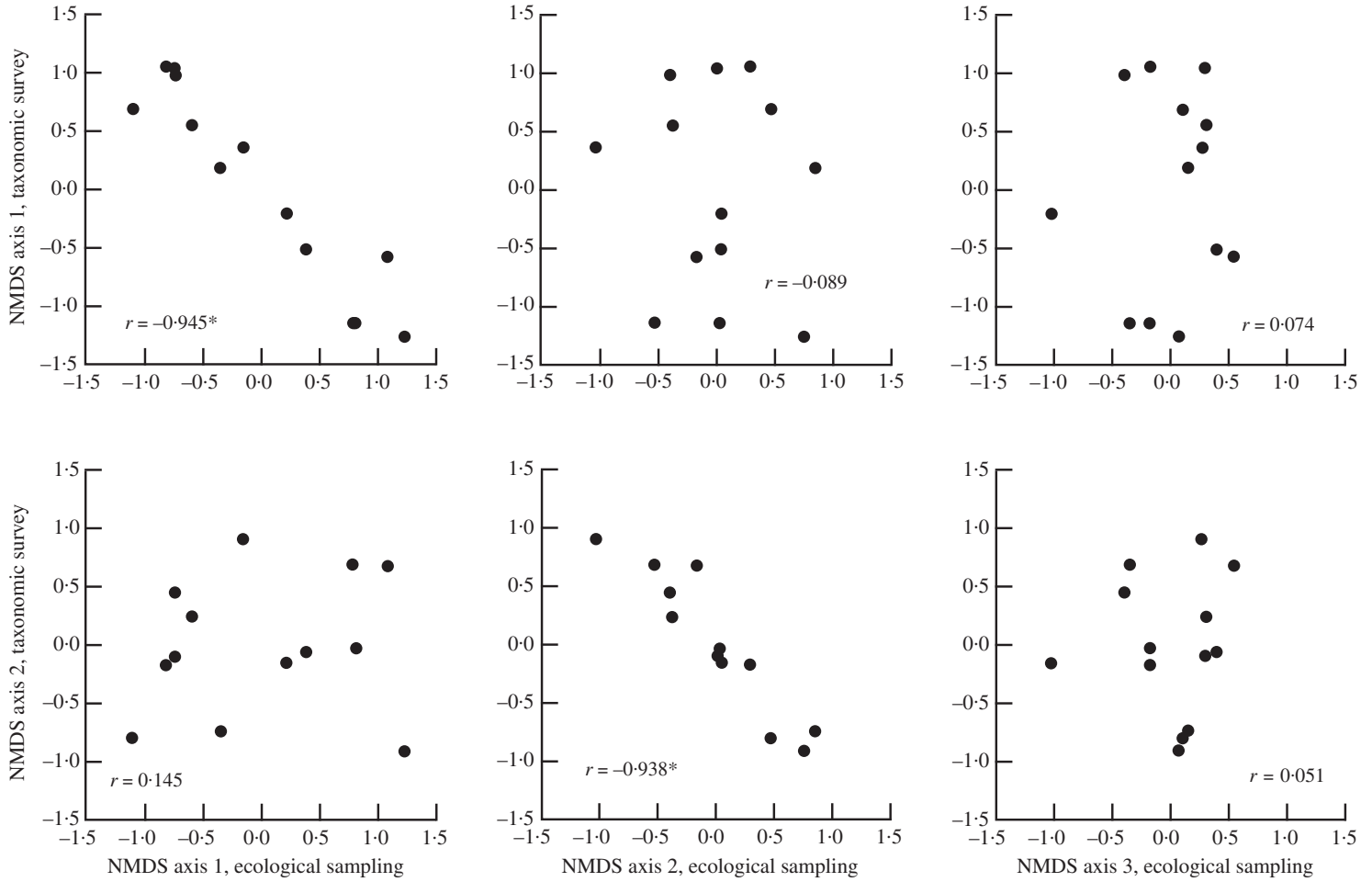


FIG. 6. Correlations for orthogonal ordination axis scores that describe trends in epiphyte community composition, recovered using NMDS. Community structure for ecological sampling used frequency data and that for taxonomic survey used presence-absence data. *r* values with \* are significant at  $P < 0.001$ . All correlations at 12 df.

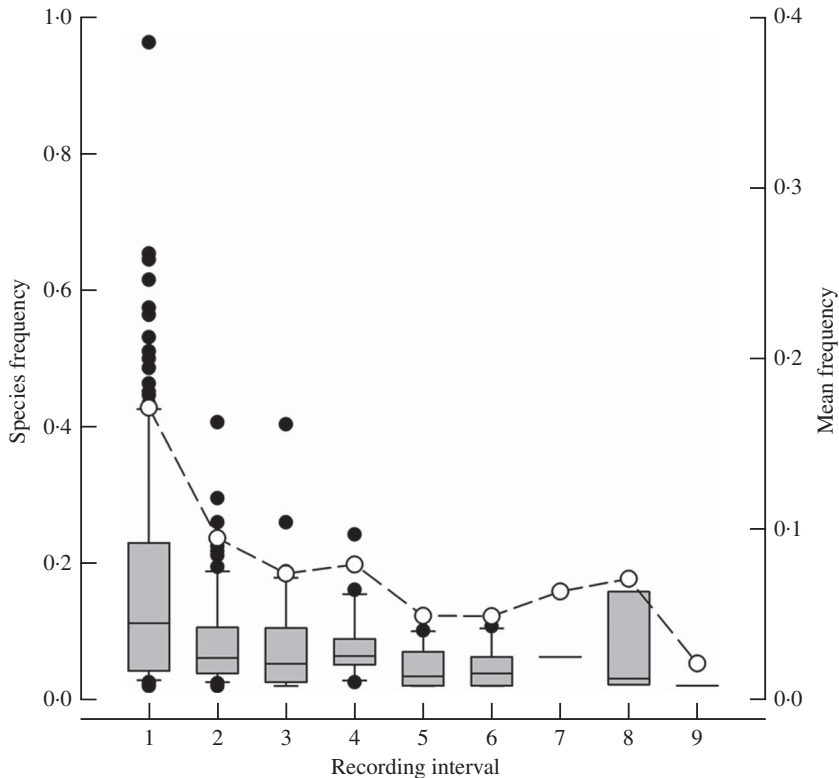


FIG. 7. The frequency of species recorded for consecutive 15-minute intervals (box plots with median and 25th and 75th percentiles, whiskers represent the 10th and 90th percentiles and closed circles indicate outliers. Open circles with dashed line indicate the mean frequency of species per interval. Note the difference in scales.

was recorded during taxonomic survey and its frequency in the stand estimated by ecological sampling (Fig. 7: tested by ANOVA,  $F = 6.94$ ,  $P < 0.0001$ , with 338 df). This was not a straightforward relationship however. Although the mean frequency of a species recorded during the first 15-minute interval was higher on average than at later intervals, there was a very large spread of frequency values for species recorded during the first interval. This suggests that uncommon species were recorded alongside the most frequent species early in the survey. Nevertheless, the species recorded during later intervals tended to have lower maximum frequencies, suggesting that the most common species (and some rarities) were recorded earlier during a survey, and that later time intervals were mostly detecting species with low frequencies.

## Discussion

The study asked whether results generated for the different purposes of taxonomic survey and ecological sampling are congruent for decision making in conservation. This is explored below through the comparison of trends in community richness and composition, within the context of potential subjectivity (human behaviour) during the taxonomic survey.

There was a positive relationship between the numbers of species recorded by taxonomic survey and ecological sampling. Estimates from ecological sampling were always lower than, but more closely approached, the values for taxonomic survey when processed to control for sampling effort (e.g. by using resampling with asymptotic curves). Thus, while ecological sampling detects fewer species

than taxonomic survey (McCune & Lesica 1992), when it is used to explain trends in species richness observed among the woodland stands in Scotland it can be considered congruent with the more complete values of epiphytic richness described using taxonomic survey. This is not an inevitable result given that quadrat sub-sampling was limited to the lower tree bole (up to a summed quadrat area of just 0.23 m<sup>2</sup> for the most heavily sampled stand (Glen Loy)) and represented a small amount of habitat space compared to taxonomic survey applied across the visible range of all tree boles, as well as accessible and fallen branches/twigs. The result implies a strong dependency between ecological scales so that local richness (sub-sampled on the tree bole) is somehow linked to a species pool of potential colonists operating at the stand-scale (Pärtel *et al.* 1996).

Despite the severely limited spatial extent of recording, the focused use of quadrats successfully located a small number of species that had gone unnoticed by the taxonomist surveyor, though these species typically had two properties. Either they were inconspicuous, such as *Jamesiella anastomosans*, *Halecania viridescens*, or *Japezia subaurifera*, or else, if they were more obvious, such as *Parmelia sulcata* or *Ramalina farinacea*, they occurred at very low frequencies in the woodland stand. This is consistent with a pattern of non-detection observed for vascular plants based on a combination of species characteristics and abundance (Vittoz & Guisan 2007; Milberg *et al.* 2008; Chen *et al.* 2013). There were also a small number of instances when some of the most common macrolichen species went unrecorded during taxonomic survey, possibly reflecting a tendency to overlook the commonly recorded and/or a human instinct to filter out early in a search the most obvious 'background' species, in order to better detect nested patterns that can locate diminutive and infrequent rarities. Many of the difficult to locate species unique to the ecological sampling were recorded because bark material from within each quadrat was routinely collected and screened under a dissecting microscope in the herbarium.

While habitat conservation should be a priority, this result nevertheless highlights the potential importance of a limited though speculative sub-sampling of different bark microhabitats, purely for microscopic examination, in order to locate diminutive species that might be easily overlooked under field conditions. The degree to which this is valuable will be context-specific. Familiarity with local diversity can be a critical factor in explaining the number of species detected and identified by surveyors (McCune *et al.* 1997; Archaux *et al.* 2009), and in Scotland the use of quadrat sampling led to marginal gains. In tropical habitats, however, where lichen species are less well known, the use of quadrat sampling can improve species detection over general survey methods (Cáceres *et al.* 2008).

The results also provided a tentative insight into the strategy employed by a single taxonomist during survey work, while acknowledging that there will be variability among different taxonomist surveyors (Vondrák *et al.* 2016). Although species recording was not exhaustive during a survey, because none of the species accumulation curves plateau within the allotted time-intervals (Fig. 2B), the decision to stop was always sufficiently delayed until a point at which each survey had begun to yield a poor return in new recorded species. Typically, this occurred when an average of 2.8 new species had been recorded during the final 15-minute interval, following 4.5 new species in the penultimate interval, consistent with a law of diminishing returns. Different surveyors follow subtly different patterns in this respect (Vondrák *et al.* 2016). Thus, surveys standardized by a maximum time per unit area may not represent a useful control method if individual surveyors are idiosyncratic in accumulating lists at different rates (Archaux *et al.* 2006) and applying different stopping rules (Klimeš *et al.* 2001). For the single surveyor in this study, the correlation between observed and expected total species richness was not linearly related to a line of equity, and surveys appear to have ceased relatively early for stands where time constraints to achieve an exhaustive species list were the most severe.

For example, this is the case for Glen Loy (site 10, Fig. 1), when a single species recorded during the final 15-minute interval (*Micarea xanthomica*) followed a relatively long and mentally-tiring survey regime in a complex stand (105 minutes). Previous studies have identified mental tiredness as a limiting factor (Archaux *et al.* 2009) and this may have prompted an end to the survey, despite the relatively high accumulation of species during the preceding two 15-minute intervals (nine and seven species, respectively). However, in such cases an informed decision might also have been taken to cease recording, for example because all contrasting microhabitats were known to have been examined, with the strong expectation for none or very few records during additional recording intervals. False discrepancies may therefore arise between observed and expected total richness because aspects of logical forward-thinking were not incorporated into the curve-fitting applied to our sampled data which were used to predict the asymptote. Similarly, taxonomic survey often aims to be realistically comprehensive, rather than exhaustive, and once a woodland stand is felt to have been contextualized in terms of species richness and composition, additional work represents an incremental gain that may be maximized by a future visit when physically and mentally refreshed.

In terms of species composition, the greater information content provided by ecological sampling of abundance or frequency data revealed community patterns undetected by taxonomic survey (presence-absence data) when estimated as the higher number of statistically significant orthogonal ordination axes. However, individual sites received similar ordination scores in relation to one another suggesting that summary differences in terms of species composition are highly likely to be congruent ( $r$  values for the comparison of site scores along NMDS axes one and two  $>0.9$ ) when compared between the taxonomic survey and ecological sampling. This is consistent with the ability of structured survey data to reveal gradients in epiphyte community composition among sites (McCune *et al.* 1997), and the

distribution of sites included here (Fig. 1) will capture the already well-established effect on lichen composition of major biogeographical gradients such as for climate (western oceanic to eastern continental sites) as well as woodland structure (Ellis *et al.* 2015). We also found that time-to-detection in taxonomic survey was statistically related to, but did not provide an entirely reliable proxy for, a species' frequency. This is because the behaviour of a taxonomist surveyor was not random. Conversations on this topic when preparing this paper suggested that a typical mode of operation during fieldwork was as follows: 1) start with one discrete habitat (in this case a tree species) and record until exhausted, 2) move to another habitat (e.g. structures such as branches, or an alternative tree species), and so on through the survey, 3) towards completion, re-walk the woodland stand looking for 'rare niches' that may have been missed, or make a special effort to find inconspicuous species which it is felt ought to be present. This selective surveying can explain the apparent spike in species frequencies at later time-intervals (Fig. 7) because a final re-evaluation of the stand can lead to overlooked assemblages. This behaviour would tend to favour comprehensive inventory, rather than reconstruction of species frequency patterns.

Finally, the methods employed by our study required that taxonomic survey work take on a quantitative character, in order to compare species richness for a timed effort. This represents a greater control over lichen taxonomic survey than is normal and can lead to improved spatial or temporal comparisons relevant to conservation planning (Hunter & Webb 2001). However, there are limits to which these constraints may be beneficial. In terms of species richness, there is evidence from vascular plant survey to suggest that error rates among observers might increase as methods become constrained to smaller sub-sampled units (Klimeš *et al.* 2001; Archaux *et al.* 2007; Vittoz & Guisan 2007), possibly because infrequent or ambiguous specimens that are young, damaged or otherwise atypical cannot be confirmed

through wider search effort and so are misidentified or discounted. The potential for misidentifications to drive differences between taxonomic survey and ecological sampling (Scott & Hallam 2002) was not addressed by our study though we expect this effect to be minimal. Specimens were extensively discussed between the taxonomist and ecologist to purposefully standardize for species concepts. As a more general point, however, it has been suggested that DNA barcoding offers the potential to standardize and improve the accuracy of species detection in vegetation survey (Thompson & Newmaster 2014). However, while DNA barcoding for lichens might improve standards in identification of sampled specimens (Kelly *et al.* 2011), a limiting step remains the expertise to recognize and sample different species under field conditions (McCune *et al.* 1997; Giordani *et al.* 2009). This process of detection has to happen prior to the application of a species name, either based on morphology or confirmed using DNA barcoding, and will remain dependent on taxonomic expertise and field skills.

The study was carried out in contribution to the Scottish Government's Strategic Research Programme and we thank Scottish Natural Heritage and private landowners for permission to access the woodland sites. We thank three anonymous reviewers for helpful comments to improve an earlier version of the manuscript.

#### SUPPLEMENTARY MATERIAL

For supplementary material accompanying this paper visit <https://doi.org/10.1017/S0024282917000056>

#### REFERENCES

- Ahrends, A., Rahbek, C., Bulling, M. T., Burgess, N. D., Platts, P. J., Lovett, J. C., Kindemba, V. W., Owen, N., Sallu, A. N., Marshall, A. R. *et al.* (2011) Conservation and the botanist effect. *Biological Conservation* **144**: 131–140.
- Archaux, F., Gosselin, F., Bergès, L. & Chevalier, R. (2006) Effects of sample time, species richness and observer on the exhaustiveness of plant censuses. *Journal of Vegetation Science* **17**: 299–306.
- Archaux, F., Bergès, L. & Chevalier, R. (2007) Are plant censuses carried out on small quadrats more reliable than larger ones? *Plant Ecology* **188**: 179–190.
- Archaux, F., Camaret, S., Dupouey, J.-L., Ulrich, E., Corcket, E., Bourjot, L., Brêthes, A., Chevalier, R., Dobremez, J. F., Dumas, Y., *et al.* (2009) Can we reliably estimate species richness with large plots? An assessment through calibration training. *Plant Ecology* **203**: 303–315.
- Bornard, C. N., Kéry, M., Bueche, L. & Fischer, M. (2014) Hide-and-seek in vegetation: time-to-detection is an efficient design for estimating detectability and occurrence. *Methods in Ecology and Evolution* **5**: 433–442.
- Cáceres, M. E. S., Lücking, R. & Rambold, G. (2008) Efficiency of sampling methods for accurate estimation of species richness of corticolous micro-lichens in the Atlantic rainforest of northeastern Brazil. *Biodiversity and Conservation* **17**: 1285–1301.
- Chao, A. (1987) Estimating the population size for mark-recapture data with unequal catchability. *Biometrics* **43**: 783–791.
- Chazdon, R. L., Colwell, R. K., Denslow, J. S. & Guariguata, M. R. (1998) Statistical methods for estimating species richness of woody regeneration in primary and secondary rain forests of NE Costa Rica. In *Forest Biodiversity Research, Monitoring and Modeling: Conceptual Background and Old World Case Studies* (F. Dallmeier & J. A. Comiskey, eds): 285–309. Paris: Parthenon Publishing.
- Chen, G., Kéry, M., Plattner, M., Ma, K. & Gardner, B. (2013) Imperfect detection is the rule rather than the exception in plant distribution studies. *Journal of Ecology* **101**: 183–191.
- Colwell, R. K. (2013) *EstimateS: statistical estimation of species richness and shared species from samples. Version 9*. Available at: <http://purl.oclc.org/estimates>
- Colwell, R. K., Chao, A., Gotelli, N. J., Lin, S.-Y., Mao, C. X., Chazdon, R. L. & Longino, J. T. (2012) Models and estimators linking individual-based and sample-based rarefaction, extrapolation, and comparison of assemblages. *Journal of Plant Ecology* **5**: 3–21.
- Coppins, B. J. (2002) *Checklist of Lichens of Great Britain and Ireland*. London: British Lichen Society.
- Dengler, J. (2009) Which function describes the species-area relationship best? A review and empirical evaluation. *Journal of Biogeography* **36**: 728–744.
- Dennis, R. L. H. & Thomas, C. D. (2000) Bias in butterfly distribution maps: the influence of hot spots and recorder's home range. *Journal of Insect Conservation* **4**: 73–77.
- Ellis, C. J. & Ellis, S. C. (2013) Signatures of autogenic succession for an aspen chronosequence. *Journal of Vegetation Science* **24**: 688–701.
- Ellis, C. J., Eaton, S., Theodoropoulos, M. & Elliott, K. (2015) *Epiphyte Communities and Indicator Species. An Ecological Guide for Scotland's Woodlands*. Edinburgh: Royal Botanic Garden Edinburgh.
- Giordani, P., Brunialti, G., Benesperi, R., Rizzi, G., Frati, L. & Modenesi, P. (2009) Rapid biodiversity assessment in lichen diversity surveys: implications for quality assurance. *Journal of Environmental Monitoring* **11**: 730–735.
- Gray, A. N. & Azuma, D. L. (2005) Repeatability and implementation of a forest vegetation indicator. *Ecological Indicators* **5**: 57–71.

- Heltshe, J. & Forrester, N. E. (1983) Estimating species richness using the jackknife procedure. *Biometrics* **39**: 1–11.
- Hodgetts, N. G. (1992) *Guidelines for the Selection of Biological SSSIs: Non-Vascular Plants*. Peterborough: Joint Nature Conservation Committee.
- Hunter, M. L. & Webb, S. L. (2001) Enlisting taxonomists to survey poorly known taxa for biodiversity conservation: a lichen case study. *Conservation Biology* **16**: 660–665.
- Jüriado, I., Liira, J., Paal, J. & Suija, A. (2009) Tree and stand level variables influencing diversity of lichens on temperate broad-leaved trees in boreo-nemoral floodplain forests. *Biodiversity and Conservation* **18**: 105–125.
- Kadmon, R., Farber, O. & Danin, A. (2004) Effect of roadside bias on the accuracy of predictive maps produced by bioclimatic models. *Ecology* **14**: 401–413.
- Kelly, L. J., Hollingsworth, P. M., Coppins, B. J., Ellis, C. J., Harrold, P., Tosh, J. & Yahr, R. (2011) DNA barcoding of lichenized fungi demonstrates high identification success in a floristic context. *New Phytologist* **191**: 288–300.
- Kent, M. (2012) *Vegetation Description and Data Analysis*. Chichester: Wiley-Blackwell.
- Kent, M. & Coker, P. (1992) *Vegetation Analysis and Description: A Practical Approach*. Chichester: John Wiley & Sons.
- Klimeš, L., Dančák, M., Hájek, M., Jongepierová, I. & Kučera, T. (2001) Scale-dependent biases in species counts in grassland. *Journal of Vegetation Science* **12**: 699–704.
- Leppik, E., Jüriado, I. & Liira, J. (2011) Changes in stand structure due to the cessation of traditional land use in wooded meadows impoverish epiphytic lichen communities. *Lichenologist* **43**: 257–274.
- Magurran, A. E. (2004) *Measuring Biological Diversity*. Oxford: Blackwell Publishing.
- McCune, B. & Grace, J. B. (2002) *Analysis of Ecological Communities*. Glendon Beach, Oregon: MjM Software Design.
- McCune, B. & Lesica, P. (1992) The trade-off between species capture and quantitative accuracy in ecological inventory of lichens and bryophytes in forests in Montana. *Bryologist* **95**: 296–304.
- McCune, B. & Mefford, M. J. (2011) *PC-ORD v.6: Multivariate Analysis of Ecological Data*. Glendon Beach, Oregon: MjM Software Design.
- McCune, B., Dey, J. P., Peck, J. E., Cassell, D., Heiman, K., Will-Wolf, S. & Neitlich, P. N. (1997) Repeatability of community data: species richness versus gradient scores in large-scale lichen studies. *Bryologist* **100**: 40–46.
- Milberg, P., Bergstedt, J., Fridman, J., Odell, G. & Westerberg, L. (2008) Observer bias and random variation in vegetation monitoring data. *Journal of Vegetation Science* **19**: 633–644.
- Moerman, D. E. & Estabrook, G. F. (2006) The botanist effect: counties with maximal species richness tend to be home to universities and botanists. *Journal of Biogeography* **33**: 1969–1974.
- Myers, N., Mittermeier, R. A., Mittermeier, C. G., da Fonseca, G. A. B. & Kent, J. (2000) Biodiversity hotspots for conservation priorities. *Nature* **403**: 853–858.
- Orange, A., James, P. W. & White, F. J. (2001) *Microchemical Methods for the Identification of Lichens*. London: British Lichen Society.
- Pärtel, M., Zobel, M., Zobel, K. & van der Maarel, E. (1996) The species pool and its relation to species richness: evidence from Estonian plant communities. *Oikos* **75**: 111–117.
- Preston, F. W. (1948) The commonness, and rarity, of species. *Ecology* **29**: 254–283.
- R (2013) *The R Foundation for Statistical Computing*. Vienna, Austria.
- Raaijmakers, J. G. W. (1987) Statistical analysis of the Michaelis-Menten equation. *Biometrics* **43**: 793–803.
- Reid, W. V. (1998) Biodiversity hotspots. *Trends in Ecology and Evolution* **13**: 275–280.
- Scott, W. A. & Hallam, C. J. (2002) Assessing species misidentification rates through quality assurance of vegetation monitoring. *Plant Ecology* **165**: 101–115.
- Smith, C. W., Aptroot, A., Coppins, B. J., Fletcher, A., Gilbert, O. L., James, P. W. & Wolseley, P. A. (eds) (2009) *The Lichens of Great Britain and Ireland*. London: British Lichen Society.
- Smith, E. P. & van Belle, G. (1984) Nonparametric estimation of species richness. *Biometrics* **40**: 119–129.
- Sykes, J. M., Horrill, A. D. & Mountford, M. D. (1983) Use of visual cover estimates as quantitative estimators of some British woodland taxa. *Journal of Ecology* **71**: 437–450.
- Systat (2010) *SigmaPlot v.11.2*. San Jose: Sysstat Software, Inc.
- Thompson, K. A. & Newmaster, S. G. (2014) Molecular taxonomic tools provide more accurate estimates of species richness at less cost than traditional morphology-based taxonomic practices in a vegetation survey. *Biodiversity and Conservation* **23**: 1411–1424.
- Tjørve, E. (2003) Shapes and functions of species-area curves: a review of possible models. *Journal of Biogeography* **30**: 827–835.
- Vittoz, P. & Guisan, A. (2007) How reliable is the monitoring of permanent vegetation plots? A test with multiple observers. *Journal of Vegetation Science* **18**: 413–422.
- Vittoz, P., Bayfield, N., Brooker, R., Elston, D. A., Duff, E. I., Theurillat, J.-P. & Guisan, A. (2010) Reproducibility of species lists, visual cover estimates and frequency methods for recording high-mountain vegetation. *Journal of Vegetation Science* **21**: 1035–1047.
- Vondrák, J., Malíček, J., Palice, Z., Coppins, B., Kukwa, M., Czarnota, P., Sanderson, N. & Acton, A. (2016) Methods for obtaining more complete species lists in surveys of lichen biodiversity. *Nordic Journal of Botany* **34**: 619–626.